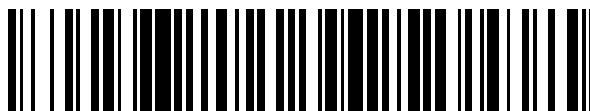


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 632 504**

51 Int. Cl.:

**A61K 38/28** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.11.2008 PCT/US2008/084252**

87 Fecha y número de publicación internacional: **28.05.2009 WO09067636**

96 Fecha de presentación y número de la solicitud europea: **20.11.2008 E 08851822 (0)**

97 Fecha y número de publicación de la concesión europea: **10.05.2017 EP 2217265**

54 Título: **Polipéptidos de insulina modificados y sus usos**

30 Prioridad:

**20.11.2007 US 989388 P**

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

**13.09.2017**

73 Titular/es:

**AMBRX, INC. (100.0%)  
10975 North Torrey Pines Road, Suite 100  
La Jolla, CA 92037, US**

72 Inventor/es:

**MIAO, ZHENWEI;  
KRAWITZ, DENISE;  
KRAYNOV, VADIM;  
TIAN, FENG;  
SIM, BEE-CHENG;  
HO, LILLIAN;  
PUTNAM, ANNA-MARIA, A. HAYS;  
KNUDSEN, NICK y  
DEGUZMAN, MICHAEL**

74 Agente/Representante:

**IZQUIERDO BLANCO, María Alicia**

**ES 2 632 504 T3**

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

**DESCRIPCIÓN****Polipéptidos de insulina modificados y sus usos****5 CAMPO DE LA INVENCION**

Esta invención se refiere a polipéptidos de insulina opcionalmente modificados con al menos un aminoácido no naturalmente codificado.

**10 ANTECEDENTES DE LA INVENCION**

La diabetes afecta actualmente a 246 millones de personas en todo el mundo y se espera que afecte a 380 millones en 2025 (estadística de la Federación Internacional de Diabetes el 15 de noviembre, 2007). La diabetes de tipo 2 constituye alrededor del 85% al 95% de todos los casos de diabetes en los países desarrollados y representa un porcentaje aún mayor en los países en desarrollo. La naturaleza epidémica de la diabetes continúa afectando a un número siempre creciente de personas en todo el mundo, mientras que la conciencia pública sigue siendo baja.

La diabetes es cualquier trastorno caracterizado por una excesiva excreción de orina. La forma más común de diabetes es la diabetes mellitus, un trastorno metabólico en el que hay una incapacidad para oxidar hidratos de carbono debido a perturbaciones en la función de la insulina. La diabetes mellitus se caracteriza por niveles elevados de glucosa en el plasma y la cetoacidosis episódica. Otros síntomas de la diabetes mellitus incluyen sed excesiva, glucosuria, poliuria, lipemia y el hambre. Si no se trata la enfermedad puede conducir a la cetoacidosis fatal. Otras formas de diabetes son la diabetes insípida y la diabetes inestable. La diabetes insípida es el resultado de una deficiencia de la hormona antidiurética. El principal síntoma de diabetes insípida (producción de orina excesiva) resulta de una incapacidad de los riñones para reabsorber agua. Diabetes inestable es una forma que es muy difícil de controlar. Se caracteriza por oscilaciones inexplicables entre la hipoglucemia y acidosis.

Criterios, que establecen un individuo clínicamente por sufrir de diabetes mellitus, incluyen: (1) que tiene un nivel de glucosa en plasma en ayunas en exceso de 126 mg/dl (7 mmol/L). Los niveles normales deben ser menos de 100mg/dl (5,6 mmol/L) o (2) que tiene niveles de glucosa en plasma en exceso de 200 mg/dl (11mmol/L) en dos momentos puntos durante una prueba de tolerancia a la glucosa oral, OGTT, una de las cuales debe estar dentro de 2 horas de la ingestión de glucosa.

Cuanto antes una persona es diagnosticada con diabetes cuanto mejor será la posibilidad de que la persona evite las consecuencias negativas primarias que son insuficiencia renal, ceguera y amputaciones de las extremidades debido a problemas circulatorios. La Asociación Americana de la Diabetes ha examinado la recomendación de que los médicos consideran a los pacientes a ser prediabéticos si su nivel de glucosa en sangre en ayunas está por encima de 100mg/dl, pero inferior a 125 mg/dl y cuyos niveles de glucosa son al menos 140mg/dl, pero inferior a 200 mg/dl después de una prueba oral de tolerancia a la glucosa (OGTT).

La diabetes mellitus es un trastorno clínico heterogéneo con numerosas causas. Existen dos clasificaciones principales de diabetes mellitus, idiopática y secundaria. La diabetes idiopática se divide en dos tipos principales; dependiente de la insulina y insulina no dependiente. La diabetes mellitus dependiente de insulina, IDDM (más comúnmente referido como diabetes de tipo 1 se define por el desarrollo de cetoacidosis en ausencia de terapia con insulina. La diabetes de tipo 1 se manifiesta con mayor frecuencia en la infancia (también llamada diabetes juvenil) y es el resultado de una destrucción autoinmune de las células B del páncreas. La diabetes mellitus no dependiente de insulina, NIDDM (más comúnmente denominada diabetes de tipo 2 se caracteriza por la hiperglucemia persistente, pero pocas veces conduce a la cetoacidosis. La diabetes de tipo 2 se manifiesta generalmente después de los 40 años y por lo tanto tiene el nombre obsoleto de adultos de inicio de la diabetes. La diabetes de tipo 2 puede ser el resultado de defectos genéticos que causan tanto la resistencia a la insulina como la deficiencia de insulina. Hay dos formas principales de diabetes de tipo 2: (1) aparición tardía asociada con la obesidad; y (2) inicio tardío no asociado a la obesidad. Muchos, si no todos, los efectos negativos a largo plazo de vivir con diabetes de tipo 2 se deben a la hiperglucemia persistente. Por esta razón un objetivo principal de la intervención terapéutica en la diabetes de tipo 2 es el de reducir los niveles de glucosa circulante.

La función principal de la insulina es contrarrestar la acción concertada de una serie de hormonas que generan hiperglicemia y para mantener los niveles bajos de glucosa en sangre. Debido a que existen numerosas hormonas de hiperglucemia, trastornos asociados no tratados con insulina generalmente conducen a la hiperglucemia severa y acortamiento de la vida.

Además de su papel en la regulación de metabolismo de la glucosa, la insulina estimula la lipogénesis, disminuye la lipólisis, y aumenta el transporte de aminoácidos en las células. La insulina también modula la transcripción, alterando el contenido celular de numerosos ARNm. Estimula el crecimiento, la síntesis de ADN, y la replicación celular, los efectos que tiene en común con los factores de crecimiento similar a la insulina (IGF) y relaxina.

**RESUMEN DE LA INVENCION**

La invención proporciona polipéptidos de insulina que comprenden un polímero soluble en agua unido por un enlace de oxima covalente con la cadena A de polipéptido de insulina en la posición 14 de la secuencia de aminoácidos de cadena A seleccionada del grupo SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, y SEQ ID NO: H, en el que el aminoácido en la posición 14 de la cadena A es un aminoácido no codificado de forma natural.

En los individuos no diabéticos, el páncreas suministra continuamente los niveles bajos (o basales) de insulina. Los suministros de hígado glucosa al resto del cuerpo cuando el cuerpo está en un estado de ayuno. Esto se denomina típicamente la producción de glucosa hepática. En respuesta a una comida, altos niveles de insulina son liberados por el páncreas. La insulina primero interactúa con el hígado, lo que indica el hígado para detener la producción de glucosa hepática y para comenzar la glucosa absorbente ingerida como parte de la comida. Cierta bolo de liberación de insulina por el páncreas pasa a través del hígado e interactúa con otras células del cuerpo, especialmente del músculo, señalando que absorben y utilizan la glucosa. En un diabético, sin embargo, este trabajo tándem hecho por el páncreas y el hígado puede ser interrumpido y es deseable proporcionar una terapia para los pacientes diabéticos que les ayuden a mantener la glucosa dentro de límites saludables sin experimentar hiper o hipoglucemia y los efectos cortos asociados y a largo plazo de estas fluctuaciones. Por lo tanto, ha sido durante mucho tiempo un objetivo de la terapia de insulina imitar el patrón de secreción endógena de insulina en individuos normales. La demanda fisiológica diaria para la insulina fluctúa y se puede separar en dos fases: (a) la fase de absorción que requiere un pulso de insulina para disponer de la oleada de glucosa en sangre relacionada con la comida, y (b) la fase postabsorbente que requiere una cantidad sostenida de insulina para regular la producción de glucosa hepática para mantener óptima la glucosa en sangre en ayunas. En consecuencia, la terapia eficaz implica generalmente el uso combinado de dos insulinas exógenas: una insulina de hora de la comida de efecto rápido proporcionada por inyecciones de bolo y una insulina basal de efecto largo administrada por inyección una vez o dos veces al día.

En una solicitud de patente asignada a Eli Lilly, la Publicación de Estados Unidos nº 20040242460, una clase de insulinas aciladas se describen para su uso como una terapia de insulina basal de efecto largo. Las insulinas aciladas se preparan mediante acilación, selectivamente con un derivado de ácido graso activado, el grupo amino libre de una insulina monomérica, incluyendo la insulina normal y ciertos análogos de insulina. Derivados de ácidos grasos útiles incluyen compuestos de tipo ácido graso reactivo que tienen al menos un seis (6) longitud de cadena de átomos de carbono y en particular los derivados de ácidos grasos que tienen de 8 a 21 átomos de carbono en su cadena. Insulina monoacilada humana normal, acilada con un derivado de ácido palmítico, es un candidato particularmente prometedor. Insulinas que caen dentro de esta categoría se describen en la solicitud de patente japonesa 1-254.699. Los péptidos de insulina de la presente invención también pueden estar acilados.

Los métodos de acuerdo con los casos de la presente divulgación pueden restaurar la homeostasis de la glucosa a un individuo diabético que tiene un hígado sano (es decir, un hígado capaz de captación de glucosa normal y la producción, pero por la ausencia de la hormona insulina de regulación). Al activar el hígado para regular el nivel de glucosa en la sangre, los métodos de la presente descripción pueden reducir o eliminar la hiperglucemia y/o la hipoglucemia asociada con los métodos convencionales de tratamiento de la diabetes mellitus. Métodos de acuerdo con los casos de la presente descripción también pueden reducir o eliminar algunas, si no todas, las complicaciones microvasculares (por ejemplo, nefropatía, retinopatía y/o neuropatía) y/o complicaciones macrovasculares (por ejemplo, infarto de miocardio y/o derrame) típicamente asociadas con la diabetes mellitus. Además, los métodos de acuerdo a los casos de la presente divulgación pueden reducir o eliminar la hiperinsulinemia asociada con la administración periférica (por ejemplo, subcutánea, intrapulmonar, intranasal, mucosal bucal) de la insulina. Además, los métodos de acuerdo a los casos de la presente divulgación pueden reducir o eliminar la hiperlipidemia asociada con la diabetes mediante la activación del hígado para mejorar su metabolismo de ácidos grasos. La activación adecuada del hígado también puede restaurar otra célula de hígado, las vías metabólicas reguladas genéticamente relacionadas con las complicaciones asociadas con la diabetes mellitus.

En algunas realizaciones, el polipéptido de insulina comprende una o más modificaciones postraduccionales. En algunas realizaciones, la insulina de polipéptido está ligada a un enlazador, polímero, o molécula biológicamente activa. En algunas realizaciones, el polipéptido de insulina está vinculado a un polímero bifuncional, enlazador bifuncional, o al menos un polipéptido de insulina adicional.

En algunas formas de realización, el aminoácido no naturalmente codificado está unido a un polímero soluble en agua. En algunas realizaciones, el polímero soluble en agua comprende un resto de poli(etilenglicol). En algunas formas de realización, el aminoácido no naturalmente codificado está ligado al polímero soluble en agua con un enlazador o está unido al polímero soluble en agua. En algunas realizaciones, la molécula de poli(etilenglicol) es un polímero bifuncional. En algunas realizaciones, el polímero bifuncional está enlazado a un segundo polipéptido. En algunas realizaciones, el segundo polipéptido es un polipéptido de insulina.

En algunas realizaciones, el polipéptido de insulina comprende al menos dos aminoácidos unidos a un polímero soluble en agua que comprende un resto de poli (etilenglicol). En algunas realizaciones, al menos un aminoácido es un aminoácido no naturalmente codificado.

En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina de insulina: en la cadena A antes de la posición 1 (es decir, en el término N), 1,2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 1 o los aminoácidos correspondientes en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, o SEQ ID NO: 11) y/o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 2 o los aminoácidos correspondientes en SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12). En algunas realizaciones, uno o más aminoácidos no naturales codificados están también incorporados en una o más de las siguientes posiciones en la insulina: antes de la posición 1 (es decir, en el término N), 1,2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de la SEQ ID NO: 13 o los aminoácidos correspondientes en SEQ ID NO: 14). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina de insulina: en la cadena A 1, 9, 15 (SEQ ID NO: 1 o las correspondientes posiciones de aminoácidos en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina de insulina: en la cadena B 1, 2, 28 (SEQ ID NO: 2 o las posiciones correspondientes de aminoácidos en SEQ ID NOs: 2, 4, 6, 8, 10, 12). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina de insulina: en la cadena A 1, 4, 5, 8, 9, 12, 15, 18 (SEQ ID NO: 1 o las correspondientes posiciones de aminoácidos en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina de insulina: en la cadena B 1, 3, 4, 21, 22, 28, 29 (SEQ ID NO: 2 o las posiciones correspondientes de aminoácidos en SEQ ID NOs: 2, 4, 6, 8, 10, 12).

En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones de insulina o un análogo de insulina: Una cadena de las posiciones 8, 9, 10, (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11); posiciones de cadena 1, 17, 21, 25, 28 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12).

En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está unido a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: en la cadena A antes de la posición 1 (es decir, en el término N), 1,2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 1 o los aminoácidos correspondientes en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, o SEQ ID NO: 11) y/o en la cadena B antes de la posición 1 (es decir, en el término N), 1,2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 2 o los aminoácidos correspondientes en SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a: antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6,7,8,9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o los aminoácidos correspondientes en SEQ ID NO: 14). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a: en la cadena A 1, 4, 5, 8, 9, 12, 14, 15, 18 (SEQ ID NO: 1 o las correspondientes posiciones de aminoácidos en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a: en la cadena B 1, 3, 4, 21, 22, 28, 29 (SEQ ID NO: 2 o las posiciones de aminoácidos correspondientes en SEQ ID NOs: 2, 4, 6, 8, 10, 12).

En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está unido a un polímero soluble en agua, incluyendo, pero no limitado a, posiciones: posiciones de la cadena A 8, 9, 10, 14 (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11); posiciones de cadena 1, 17, 21, 25, 28 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO 12).

En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir

insulina, o la insulina eptides glargina pólipo: posiciones de la cadena B 1,2, 9, 10, 28, 29, 30 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: posiciones de la cadena B 1,2, 9, 10, 28, 29, 30 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12).

También se describe un polipéptido de insulina que tiene uno o más aminoácidos codificados no naturales incorporados en uno o más de la siguiente posición en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o con polipéptidos de glargina de insulina: cadena A 21, o posiciones de la cadena B 1, 10, 13, 16, 28, 29, 31, o 32 (es decir, en el extremo carboxilo de la proteína) (SEQ ID NO: 1 a SEQ ID NO: 12). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a, posiciones: cadena A 21, o posiciones de la cadena B 1, 10, 13, 16, 28, 29, 31, o 32 (es decir, en el extremo carboxilo de la proteína) (SEQ ID NO: 1 a SEQ ID NO: 12).

En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina insulina: posiciones de la cadena A 8, 9, 10, (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: posiciones de la cadena A 8, 9, 10 (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11).

También se describe un polipéptido de insulina que tiene uno o más aminoácidos no naturales codificados incorporados en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina de insulina: posiciones de la cadena B 17, 21, 25, 28 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a las posiciones: posiciones de cadena B 17, 21, 25, 28 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12).

En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina de insulina: posiciones de la cadena A 4, 8, 9, 10, 18, 21 (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: posiciones de la cadena A 4, 8, 9, 10, 14, 18, 21 (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11).

En algunos casos de la descripción, uno o más aminoácidos no naturales codificados son incorporados en una o más de las siguientes posiciones en cualquiera de la insulina, análogos de insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina insulina: posiciones de cadena B 1, 5, 17, 21, 25, 29, 30 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12). En algunos casos, el aminoácido no naturalmente codificado en una o más de estas posiciones está ligada a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: posiciones de la cadena B 1, 5, 17, 21, 25, 29, 30 (SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12).

En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en la insulina: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 1) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, (SEQ ID NO: 2). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en la insulina lispro: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 3) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, (SEQ ID NO: 4). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en la insulina aspart: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 5) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, (SEQ ID NO: 6). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en la insulina glulisina: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 7) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, (SEQ ID NO: 8).

En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en la insulina detemir: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 9) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, (SEQ ID NO: 10). En algunas realizaciones, uno o más aminoácidos no naturales codificados también se incorporan en una o más de las siguientes posiciones en la insulina glargina: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 11) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 12).

En algunas realizaciones, el polipéptido de insulina de la invención puede tener un aminoácido no naturalmente codificado en una o más de estas posiciones en cualquiera de la insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos insulina glargina es ligado a un agua soluble de polímero, incluyendo, pero no limitado a, las posiciones: antes de la posición 1 (es decir, en el término N), en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 1 o los correspondiente aminoácidos en la SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, o SEQ ID NO: 11) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 2 o los aminoácidos correspondientes en SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12).

En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones en el polipéptido de insulina está unido a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: antes de la posición 1 (es decir, en el término N), en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 1) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, (SEQ ID NO: 2). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones en el polipéptido de insulina lispro está ligado a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 3) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 4). En algunas realizaciones, los aminoácidos no naturales codificados de en una o más de estas posiciones en el polipéptido de insulina aspart está ligado a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 5) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 6). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones en el polipéptido de insulina glulisina está ligado a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, (es decir, en el carboxilo terminal de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 7) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 8). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones en el polipéptido de insulina detemir está ligado a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el término de carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 9) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 10). En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones en el polipéptido de insulina glargina está ligado a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, (es decir, en el carboxilo terminal de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 11) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, (SEQ ID NO: 12).

También se describen casos en los que uno o más aminoácidos no naturales codificados se incorporan en una o más de las siguientes posiciones de cualquiera de los de la insulina, insulina lispro, insulina glulisina, detemir insulina, o polipéptidos de glargina Insulina: 28, 36, 76, 80, 107, 108, 111, 8, 15, 19, 108, 111, 113, 155, y cualquier combinación de los mismos (SEQ ID NO: 1 y SEQ ID NO: 2; SEQ ID NO: 3 y SEQ ID NO: 4; SEQ ID NO: 5 y SEQ ID NO: 6; SEQ ID NO: 7 y SEQ ID NO: 8; SEQ ID NO: 9 y SEQ ID NO: 10; o SEQ ID NO: 11 y SEQ ID NO: 12).

Los métodos de la presente divulgación se podrían utilizar para producir cristales de análogos de insulina que

contienen uno o más aminoácidos no natural codificados. Patente de Estados Unidos N° 7.193.035 asignada a Sanofi-Aventis Deutschland GmbH, describe la formación y el uso de cristales de un análogo de insulina.

### Formulaciones

Una formulación puede contener una mezcla de dos o más de una insulina, un análogo de insulina, una insulina acilada, o análogo de insulina acilado con al menos uno de los componentes de la mezcla que contiene un aminoácido no naturalmente codificado. En otro caso, las formulaciones que contienen una mezcla de dos o más de insulina, un análogo de insulina, una insulina acilada, o análogo de insulina acilado con al menos uno de los componentes de la mezcla que contiene un aminoácido no naturalmente codificado incluye también al menos uno de agua polímero soluble unido a al menos uno de los aminoácidos no natural codificados.

La presente descripción también incluye mezclas heterogéneas en las que los polipéptidos de insulina y análogos de insulina se preparan por los métodos descritos en este documento y luego se mezclan de modo que una formulación puede administrarse a un paciente en necesidad del mismo que contiene, por ejemplo, 25% polipéptido de insulina que contienen un aminoácido no naturalmente codificado en la posición 28 de la cadena B que se ha pegilado, 25% polipéptido de insulina que contiene un aminoácido no naturalmente codificado en la posición 10 de la cadena B, dicho aminoácido no naturalmente codificado acoplado a un polímero soluble en agua, y 50% de polipéptido de insulina en el que un aminoácido no naturalmente codificado se produce en la posición 31 de la cadena B de la insulina (SEQ ID NO: 2; alternativamente SEQ ID NOs: 4, 6, 8, 10, o 12). Todas las diferentes mezclas de diferentes cantidades porcentuales de variantes de polipéptido de insulina en el que los polipéptidos de insulina incluyen una variedad (1) con PEGs de diferente tamaño, o (2) PEGs se incluyen en diferentes posiciones en la secuencia. Esto pretende ser un ejemplo y no debería en modo alguno interpretarse como una limitación a las formulaciones hechas posibles por la presente invención y serán evidentes para los expertos en la técnica. En una realización adicional, las variantes de polipéptido de insulina que se incluyen en la mezcla de formulación serán elegidas por sus tiempos de disociación que varían de modo que la formulación puede proporcionar una liberación sostenida de insulina para un paciente en necesidad del mismo.

formulaciones de la presente invención pueden incluir un glucagón.

### Otras realizaciones de la invención presente incluyen formulación para inhalación

En una realización adicional de la presente invención, es posible utilizar la tecnología descrita en el presente documento para la producción de análogos de insulina con un aumento de las propiedades farmacocinéticas y farmacodinámicas para uso del paciente por medio de la administración al pulmón, lo que resulta en niveles elevados en sangre de la insulina que se sustentan durante al menos 6 horas, y más típicamente por lo menos 8, 10, 12, 14, 18, 24 horas o mayor de posadministración. Otra realización de la presente invención permite mezclas ventajosas de los análogos de insulina adecuados para las formulaciones terapéuticas diseñadas para ser administradas a los pacientes como un inhalante.

En algunas realizaciones de la presente descripción, los siguientes sitios en la molécula de insulina nativa pueden estar sustituidos con aminoácidos no naturales codificados y, opcionalmente modificados adicionalmente por unión covalente de un polímero soluble en agua, tal como PEG: el término 2 C de las cadenas A y B, Arg22B, His10B, His5A, Glu4A, Glu17A, Glu13B y Glu21B.

Además de la insulina nativa, la presente invención proporciona polipéptidos de insulina no nativos y análogos de insulina que tienen uno o más aminoácidos no naturales codificados sustituidos o insertados en la secuencia de señal, que también puede proporcionar un sitio para la incorporación de uno o más polímeros solubles en agua, tal como PEG. Esta realización de la invención es particularmente útil para la introducción de sitios de pegilación adicionales, dentro de la molécula de insulina, por ejemplo, para la formación de una insulina de PEG, teniendo resistencia mejorada a la degradación enzimática. Este enfoque proporciona una mayor flexibilidad en el diseño de un conjugado de insulina optimizado que tiene el equilibrio deseado de actividad, estabilidad, solubilidad y propiedades farmacológicas. Las mutaciones pueden llevarse a cabo, es decir, por mutagénesis específica, en cualquier número de posiciones dentro de la molécula de insulina. PEG para uso en la presente invención pueden poseer una variedad de estructuras: lineal, bifurcada, ramificada, y similares. Típicamente, el PEG es activado con un grupo activador adecuado apropiado para el acoplamiento de un sitio o sitios deseados en la molécula de insulina. Un PEG activado poseerá un grupo reactivo en un extremo para la reacción con la insulina. Derivados representativos activos y métodos de PEG para la conjugación de estos agentes a un fármaco tal como insulina son conocidos en la técnica y además se describe en Zalipsky, s., *et al.*, "Use of Functionalized Poly(Ethylene Glycols) for Modification of Polypéptidos " in Polyethylene Glycol Chemistry: Biotechnical and Biomedical Applications, J.M. Harris, Plenus Press, Nueva York (1992), y en Advanced Drug Reviews, 16: 157-182 (1995).

En una realización particular de la invención, la porción de PEG del conjugado carece de uno o más grupos lipófilos eficaces para modificar significativamente la naturaleza soluble en agua del polímero o del conjugado de polímero-insulina. Es decir, la porción de polímero o no de insulina de un conjugado de la invención puede contener un grupo de átomos que se considera más lipófilo que hidrofílico (por ejemplo, una cadena de carbono que tiene de

aproximadamente 2 a 812 átomos de carbono), sin embargo, si la presencia de un grupo o grupos no es eficaz para alterar significativamente la naturaleza hidrófila del polímero o del conjugado, a continuación, tal resto puede estar contenido en los conjugados de la invención. Es decir, a través de mutaciones de sitio específico de la insulina, polipéptidos de insulina y análogos de insulina, un conjugado de insulina de la invención en sí mismo puede exhibir hidrófilo, en lugar de lipófila o anfífila. En ciertas realizaciones de la invención en las que un resto lipófilo puede estar presente, el resto preferiblemente no se posiciona en un extremo de una cadena de PEG.

PEG ramificados para su uso en los conjugados de la invención incluyen los descritos en la Publicación de Patente Internacional WO 96/21469. Generalmente, PEGs ramificados se pueden representar por la fórmula R (PEG--OH).sub.n, donde R representa la molécula central "núcleo" y .sub.n representa el número de brazos. Los PEG ramificados tienen un núcleo central desde el que se extienden 2 o más brazos "PEG". En una configuración ramificada, el núcleo de polímero ramificado posee un sitio reactivo único para la unión a la insulina. Los PEG ramificados para su uso en la presente invención comprenderán típicamente menos de 4 brazos de PEG, y más preferiblemente, comprenderán menos de 3 brazos de PEG. Los PEG ramificados ofrecen la ventaja de tener un único sitio reactivo, junto con una nube de polímero más grande, más densa que sus contrapartes de PEG lineales. Un tipo particular de PEG ramificado se puede representar como (MeO-PEG).sub.p R--X, donde p es igual a 2 o 3, R es una estructura de núcleo central tal como lisina o glicerol que tiene 2 ó 3 brazos de PEG unidos a la misma, y X representa cualquier grupo funcional adecuado que es o que puede ser activado para el acoplamiento a la insulina. Un PEG particularmente ramificado preferido es mPEG2-NHS (Shearwater Corporation, Alabama) que tiene la estructura de mPEG2-lisina-succinimida.

] En aún otra arquitectura ramificada, "PEG colgante" tiene grupos reactivos para el acoplamiento de proteínas posicionadas a lo largo del esqueleto de PEG en lugar de al final de las cadenas de PEG. Los grupos reactivos que se extienden desde el esqueleto de PEG para el acoplamiento a la insulina pueden ser los mismos o diferentes. Estructuras de PEG colgante pueden ser útiles, pero son generalmente menos preferidas, particularmente para composiciones para inhalación.

Alternativamente, la porción PEG de un conjugado de insulina PEG puede poseer una estructura en forma de horquilla que tiene un resto ramificado en un extremo de la cadena polimérica y dos grupos reactivos libres (o cualquier múltiplo de 2) unido al resto ramificado para la unión a la insulina. PEGs bifurcados ejemplares se describen en la Publicación de Patente Internacional nº WO 99/45964. El polietilenglicol en forma de horquilla puede incluir opcionalmente un grupo "R" o alquilo en el extremo opuesto de la cadena polimérica. Más específicamente, un conjugado PEG-insulina bifurcada de acuerdo con la invención tiene la fórmula: R-PEG-L(Y-insulina)<sub>n</sub>, donde R es alquilo, L es un punto de ramificación hidrolíticamente estable e Y es un grupo de enlace que proporciona la unión química del polímero en forma de horquilla a la insulina, y n es un múltiplo de 2. L puede representar un único grupo de "núcleo", como "CH", o puede comprender una cadena más larga de átomos. Los grupos L ejemplares incluyen lisina, glicerol, pentaeritritol, o sorbitol. Típicamente, el átomo de rama particular dentro del resto de ramificación es de carbono.

En una realización particular de la invención, la unión del PEG en horquilla a la molécula de insulina, (Y), es hidrolíticamente estable. En una realización preferida, n es 2. Restos adecuados Y, antes de la conjugación con un sitio reactivo en la insulina, incluyen pero no se limitan a ésteres activos, carbonatos activos, aldehídos, isocianatos, isotiocianatos, epóxidos, alcoholes, maleimidias, vinilsulfonas, hidrazidas ditiopiridinas, y iodacetamidas. La selección de un grupo de activación adecuado dependerá del sitio previsto de unión en la molécula de insulina y puede ser determinada fácilmente por un experto en la técnica. El grupo Y correspondiente en el conjugado PEG-insulina resultante es la que resulta de la reacción del polímero en forma de horquilla activada con un sitio reactivo adecuado en la insulina. La identidad específica de la articulación final será evidente para un experto en la técnica. Por ejemplo, si el reactivo PEG bifurcado contiene un éster activado, tal como una succinimida o éster de maleimida, la conjugación a través de un sitio de amina en la insulina dará lugar a la formación del enlace amida correspondiente. Estos polímeros bifurcados particulares son particularmente atractivos ya que proporcionan conjugados que tienen una relación molar de insulina a PEG de 2:1 o mayor. Tales conjugados pueden ser menos propensos a bloquear el sitio receptor de la insulina, mientras que todavía proporciona la flexibilidad en el diseño para proteger la insulina frente a la degradación enzimática, por ejemplo, por la enzima degradante de insulina.

En una realización relacionada, un conjugado PEG-insulina de horquilla se pueden utilizar en la presente invención, representado por la fórmula: R-[PEG-L(Y-insulina)<sub>2</sub>]<sub>n</sub>. En este caso R representa un aminoácido no naturalmente codificado que tiene al menos un conjugado PEG-diinsulina unido al mismo. Específicamente, los polímeros bifurcados preferidas de acuerdo con este aspecto de la invención son aquellos en los que n se selecciona del grupo que consiste en 1, 2, 3, 4, 5, y 6. En una realización alternativa, el enlace químico entre el aminoácido no natural dentro de la insulina, polipéptido de insulina o análogo de insulina y el punto de ramificación del polímero puede ser degradable (es decir, hidrolíticamente inestable). Alternativamente, uno o más enlaces degradables pueden estar contenidos en la cadena principal del polímero para permitir la generación *in vivo* de un conjugado PEG-insulina que tiene una cadena de PEG más pequeño que en el conjugado se administra inicialmente. Por ejemplo, un conjugado grande y relativamente inerte (es decir, que tiene una o más cadenas de PEG de alto peso molecular unido a la misma, por ejemplo, una o más cadenas de PEG que tienen un peso molecular mayor que aproximadamente 10,000, en donde el conjugado no posee esencialmente ninguna bioactividad) puede estar administrado, el cual



entonces o bien en el pulmón o en el torrente sanguíneo, es hidrolizado para generar un conjugado bioactivo que posee una porción de la cadena de PEG originalmente presente. Tras la escisión *in vivo* de la unión hidrolíticamente degradable, ya sea insulina libre (dependiendo de la posición del enlace degradable) o insulina que tiene una etiqueta de polietileno pequeño unido a la misma, se libera y se absorbe más fácilmente a través del pulmón y/o se distribuye en la sangre.

En una característica de esta realización de la invención, el polímero-conjugado intacto, antes de la hidrólisis, se degrada mínimamente después de la administración, de manera que la hidrólisis del enlace escindible es eficaz para gobernar la velocidad lenta de liberación de insulina activa en el torrente sanguíneo, en contraposición a la degradación enzimática de la insulina antes de su liberación en la circulación sistémica.

Enlaces fisiológicamente escindibles apropiados incluyen, pero no se limitan a éster, éster de carbonato, carbamato, sulfato, fosfato, éter de aciloxialquilo, acetal, y cetal. Tales conjugados deben poseer un enlace fisiológicamente escindible que es estable durante el almacenamiento y después de la administración. Por ejemplo, un conjugado de enlace-insulina de escisión de PEG debe mantener su integridad a la fabricación de la composición farmacéutica final, tras la disolución en un vehículo de administración apropiado, si se emplea, y tras la administración independientemente de la vía.

### **Insulina de cadena única**

Cuando se centra en la resistencia periférica a la insulina, el fármaco de elección es una tiazolidindiona, que es un tipo de agente de sensibilización a la insulina. La troglitazona (TRG), por ejemplo, es un agente antidiabético oralmente activo de la serie química de tiazolidinadiona. Este fármaco ha demostrado invertir la resistencia a la insulina en pacientes con NIDDM y tolerancia alterada a la glucosa, y puede mejorar la acción de la insulina en numerosos modelos de roedores genéticos y adquiridos de resistencia a la insulina. Los efectos antihiperoglucémicos de TRG resultado de su capacidad para aumentar la eliminación de glucosa dependiente de insulina y reducir la producción de glucosa hepática. Se cree que, mediante la mejora la acción de la insulina, el tratamiento de TRG resulta en la euglucemia a un nivel de insulina circulante inferior. En este sentido, los estudios en roedores normales y diabéticos y ensayos clínicos humanos no han revelado la hipoglucemia como una complicación del tratamiento con tiazolidinedionas. Por otro lado, la administración de estos fármacos a animales diabéticos normales o deficientes en insulina no alteran la glucosa en plasma o la insulina o tolerancia a la glucosa, aunque la sensibilidad a la insulina, no obstante, se incrementó.

La estructura de dos cadenas de la insulina permite a la insulina llevar a cabo conformaciones múltiples, y varios hallazgos han indicado que la insulina tiene la propensión al cambio conformacional considerable y que las restricciones en el potencial de dicho cambio disminuyen considerablemente la afinidad del receptor de la insulina para los ligandos. La proinsulina tiene una afinidad 100 veces menor para el receptor de insulina de la insulina nativa. El bloqueo del residuo aminoácido A1 en la insulina también se traduce en mala unión al receptor, de acuerdo con el dogma de que un N-terminal libre de la cadena A y C-terminal libre de la cadena B de la insulina son importantes para la unión al receptor de insulina.

La estabilidad física y química heredada de la molécula de insulina es una condición básica para la terapia con insulina de la diabetes mellitus. Estas propiedades básicas son fundamentales para la formulación de insulina y para los métodos de administración de insulina aplicables, así como para condiciones de almacenamiento de preparaciones farmacéuticas. El uso de soluciones en la administración de insulina expone la molécula a una combinación de factores, por ejemplo temperatura elevada, interfases variables aire-líquido-sólido, así como fuerzas de cizallamiento, que pueden resultar en conformación irreversible cambia por ejemplo la fibrilación. Esto es particularmente relevante para las soluciones de insulina en las bombas de infusión, ya sea usado externamente o implantado, que expone la molécula a una combinación de estos factores, así como las fuerzas de cizallamiento desde el movimiento de la bomba por períodos prolongados de tiempo. En consecuencia, la fibrilación es especialmente una preocupación cuando se utilizan bombas de infusión como sistema de administración de insulina. Además, la solubilidad de la insulina está influenciada por múltiples factores y muestra clara reducción en el intervalo de pH de 4,2 y 6,6. La zona de precipitación pH generalmente impone limitaciones para la formulación, pero también se ha utilizado deliberadamente en el desarrollo y formulación de ciertos análogos.

Por lo tanto, en otra realización de la presente invención, la modificación tal como se define en las reivindicaciones se incorpora en una única insulina de cadena o análogo de insulina de cadena única. Insulinas monocatenarias con actividad de la insulina se describen en EP 1,193,272. Estas insulinas monocatenarias tienen un C-Péptido modificado de 5-18 aminoácidos y tienen una actividad de insulina de hasta un 42%. EP 1,193,272 da a conocer los siguientes C-péptidos modificados que conectan B30 con A21: Gly-Gly-Gly-Pro-Gly-Lys-Arg, Arg-Arg-Gly-Pro-Gly-Gly-Gly, Gly-Gly-Gly-Gly-Gly-Lys-Arg, Arg-Arg-Gly-Gly-Gly-Gly-Gly, Gly-Gly-Ala-Pro-Gly-Asp-Val-Lys-Arg, Arg-Arg-Ala-Pro-Gly-Asp-Val-Gly-Gly, Gly-Gly-Tyr-Pro-Gly-Asp-Val-Lys-Arg, Arg-Arg-Tyr-Pro-Gly-Asp-Val-Gly-Gly, Gly-Gly-His-Pro-Gly-Asp-Val-Lys-Arg, y Arg-Arg-His-Pro-Gly-Asp-Val-Gly-Gly. EP 741,188 da a conocer insulinas cadena sencilla con un C-Péptido modificado que tiene de 10-14 residuos de aminoácidos y que tiene actividad de la insulina de 14 a 34% y que tiene los siguientes péptidos de conexión Gln-Pro-Leu-Ala-Leu-Glu-Gly-Ser-Leu-Gln-Lys-Arg y Gly-Tyr-Gly-Ser-Ser-Ser-Arg-Ala-Pro-Gln-Thr. WO 95/16708 da a conocer insulinas de cadena sencilla con un

péptido de conexión de 1-15 residuos de aminoácidos y sin Lys o Arg como el residuo de aminoácido C-terminal en el péptido de conexión. WO 95/16708 da a conocer las siguientes secuencias C-péptido Gly-Tyr-Gly-Ser-Ser-Ser-Arg-Arg-Ala-Pro-Gln-Thr y Gly-Tyr-Gly-Ser-Ser-Ser-Ala-Ala-Ala-Pro-Gln-Thr. Se ha informado de que estas insulinas monocatenarias tienen una actividad de insulina, y también una afinidad al receptor de IGF-1 bastante alta.

5 Publicación de Patente de Estados Unidos Novo Nordisk, Inc. n° 20070129284 1491 describe análogos de insulina de cadena sencilla por introducción de un C-Péptido entre las cadenas B y A para disminuir la flexibilidad molecular y de forma concomitante reducir la propensión de fibrilación y limitar o modificar la zona de precipitación de pH. Un ejemplo de la insulina monocatenaria descrita por Novo Nordisk, Inc., comprende las cadenas B y A de insulina o análogos o derivados de los mismos conectados por un péptido de conexión humana, en el que el péptido de conexión tiene 5-11 residuos de aminoácidos, siempre que si el péptido de conexión contiene dos restos adyacentes básicos de aminoácidos entonces al menos uno de los residuos de aminoácidos naturales en la cadena B y/o A es sustituido con otro residuo de aminoácido codificable o al menos un residuo de lisina en la cadena A, en la Cadena B o en el péptido de conexión se ha modificado químicamente por acilación o el péptido de conexión no es una de las siguientes secuencias Gly-Gly-Gly-Pro-Gly-Lys-Arg, Arg-Arg-Gly-Pro-Gly-Gly-Gly, Gly-Gly-Gly-Gly-Gly-Lys-Arg, Arg-Arg-Gly-Gly-Gly-Gly-Gly, Gly-Gly-Ala-Pro-Gly-Asp-Val-Lys-Arg, Arg-Arg-Ala-Pro-Gly-Asp-Val-Gly-Gly, Gly-Gly-Tyr-Pro-Gly-Asp-Val-Lys-Arg, Arg-Arg-Tyr-Pro-Gly-Asp-Val-Gly-Gly, Gly-Gly-His-Pro-Gly-Asp-Val-Lys-Arg, o Arg-Arg-His-Pro-Gly-Asp-Val-Gly-Gly.

20 En algunas realizaciones, el polipéptido de la invención comprende una o más sustitución, adición o delección de aminoácidos no codificados naturalmente en la secuencia señal. En algunas realizaciones, el polipéptido de la invención comprende una o más sustitución, adición o delección de aminoácidos no naturalmente codificados, en la secuencia de señal de la insulina o cualquiera de los análogos de insulina o polipéptidos descritos en esta memoria. En algunas realizaciones, el polipéptido de la invención comprende una o más sustitución, adición o delección de aminoácidos no naturalmente codificados, en la secuencia de señal, así como una o más sustitución, adición o delección de aminoácidos no naturalmente codificados, en la secuencia de señal de insulina o cualquiera de los análogos de insulina o polipéptidos descritos en esta memoria. En algunas realizaciones, uno o más aminoácidos no naturales se incorporan en la secuencia líder o señal para insulina o cualquiera de los análogos de insulina o polipéptidos descritos en esta memoria.

30 En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que modula la afinidad del polipéptido de insulina para un receptor de polipéptido de insulina o pareja de unión, incluyendo, pero no limitado a, una proteína, polipéptido, molécula pequeña o ácido nucleico. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que aumenta la estabilidad del polipéptido de insulina en comparación con la estabilidad de la insulina correspondiente en la sustitución, adición o eliminación. La estabilidad y/o solubilidad puede medirse usando un número de diferentes ensayos conocidos por los expertos normales en la técnica. Tales ensayos incluyen, pero no se limitan a SE-HPLC y RP-HPLC. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que modula la inmunogenicidad del polipéptido de insulina en comparación con la inmunogenicidad de la insulina correspondiente en la sustitución, adición o eliminación. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que modula la vida media en suero o circulación del polipéptido de insulina en comparación con la vida media en suero o el tiempo de circulación de la insulina correspondiente en la sustitución, adición o eliminación.

45 En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que aumenta la solubilidad acuosa del polipéptido de insulina en comparación con la solubilidad acuosa de la insulina correspondiente en la sustitución, adición o eliminación. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que aumenta la solubilidad del polipéptido de insulina producida en una célula huésped en comparación con la solubilidad de la insulina correspondiente en la sustitución, adición o eliminación. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que aumenta la expresión del polipéptido de insulina en una célula huésped o aumenta la síntesis *in vitro* cuando se compara con la expresión o síntesis de la insulina correspondiente en la sustitución, adición, o supresión. El polipéptido de insulina que comprende esta sustitución retiene la actividad agonista y conserva o mejora los niveles de expresión en una célula huésped. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que aumenta la resistencia de la proteasa del polipéptido de insulina en comparación con la resistencia a la proteasa de la insulina correspondiente en la sustitución, adición o eliminación. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que modula la actividad de transducción de señales del receptor de insulina en comparación con la actividad del receptor tras la interacción con el polipéptido de insulina correspondiente en la sustitución, adición o eliminación. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que modula su unión a otra molécula tal como un receptor cuando se compara con la unión del polipéptido de insulina correspondiente en la sustitución, adición o eliminación. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que modula su actividad antiviral en comparación con la actividad antiviral del polipéptido de insulina correspondiente en la sustitución, adición o eliminación. En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que aumenta su actividad de metabolización de la glucosa en comparación con la actividad de metabolización de la glucosa del polipéptido de insulina correspondiente en la sustitución, adición o eliminación.

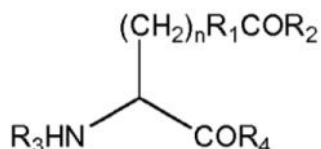
En algunas realizaciones, el polipéptido de insulina comprende una sustitución, adición o delección que aumenta la compatibilidad del polipéptido de insulina con conservantes farmacéuticas (por ejemplo, m-cresol, fenol, alcohol bencílico) en comparación con la compatibilidad de la insulina correspondiente en la sustitución, adición, o eliminación. Esta mayor compatibilidad permitiría la preparación de una formulación farmacéutica conservada que mantiene las propiedades fisicoquímicas y la actividad biológica de la proteína durante el almacenamiento.

En algunas realizaciones, uno o más enlaces de ingeniería se crean con uno o más aminoácidos no naturales. El enlace intramolecular puede ser creado de muchas maneras, incluyendo pero no limitado a, una reacción entre dos aminoácidos en la proteína en condiciones adecuadas (uno o ambos aminoácidos pueden ser un aminoácido no natural); una reacción con dos aminoácidos, cada uno de los cuales pueden ser codificados de forma natural o no natural, con un engarce, polímero, u otra molécula en condiciones adecuadas; etc.

En algunas realizaciones, una o más sustituciones de aminoácidos en el polipéptido de insulina puede ser con uno o más de origen natural o aminoácidos no naturalmente codificados. En algunas realizaciones las sustituciones de aminoácidos en el polipéptido de insulina pueden ser aminoácidos naturalmente o no naturalmente codificados, siempre que el polipéptido de insulina comprende un polímero soluble en agua unido por un enlace covalente de oxima con la cadena A de polipéptido de insulina en la posición 14 de la amino cadena A secuencia de ácido seleccionado del grupo SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, y SEQ ID NO: H, en el que el aminoácido en la posición 14 de la cadena A se un aminoácido no codificado de forma natural. En algunas realizaciones, uno o más aminoácidos sustituciones en el polipéptido de insulina puede ser con uno o más de origen natural aminoácidos, y adicionalmente al menos una sustitución es con un aminoácido no naturalmente codificado siempre que el polipéptido de insulina comprende un polímero soluble en agua unido por un enlace covalente de oxima con el polipéptido de insulina cadena A en la posición 14 de la secuencia de aminoácidos de cadena A seleccionada del grupo SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, y SEQ ID NO: 11, en el que el aminoácido en la posición 14 de la cadena A es un aminoácido no codificado de forma natural.

En algunas formas de realización, el aminoácido no naturalmente codificado comprende un grupo carbonilo, un grupo acetilo, un grupo aminooxi, un grupo hidrazina, un grupo hidrazida, un grupo semicarbazida, un grupo azida o un grupo alquino.

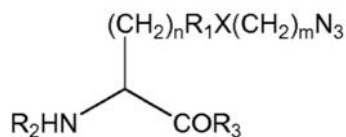
En algunas formas de realización, el aminoácido no naturalmente codificado comprende un grupo carbonilo. En algunas formas de realización, el aminoácido no naturalmente codificado tiene la estructura:



en la que n es 0-10; R1 es un alquilo, arilo, alquilo sustituido, o arilo sustituido; R2 es H, un alquilo, arilo, alquilo sustituido, y arilo sustituido; y R3 es H, un aminoácido, un polipéptido, o un grupo de modificación amino terminal, y R4 es H, un aminoácido, un polipéptido, o un grupo de modificación terminal carboxi.

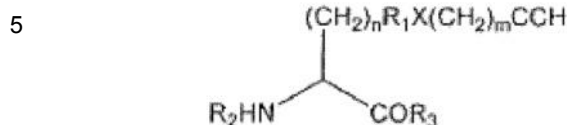
En algunas formas de realización, el aminoácido no naturalmente codificado comprende un grupo aminooxi. En algunas formas de realización, el aminoácido no naturalmente codificado comprende un grupo hidrazida. En algunas formas de realización, el aminoácido no naturalmente codificado comprende un grupo hidrazina. En algunas realizaciones, el residuo de aminoácido no naturalmente codificado comprende un grupo semicarbazida.

En algunas realizaciones, el residuo de aminoácido no naturalmente codificado comprende un grupo azida. En algunas formas de realización, el aminoácido no naturalmente codificado tiene la estructura:



en donde n es 0-10; R1 es un alquilo, arilo, alquilo sustituido, arilo sustituido o no está presente; X es O, N, S o no está presente; m es 0-10; R2 es H, un aminoácido, un polipéptido, o un grupo de modificación amino terminal, y R3 es H, un aminoácido, un polipéptido, o un grupo de modificación terminal carboxi.

En algunas formas de realización, el aminoácido no naturalmente codificado comprende un grupo alquino. En algunas formas de realización, el aminoácido no naturalmente codificado tiene la estructura:



15 en donde n es 0-10; R1 es un alquilo, arilo, alquilo sustituido, o arilo sustituido; X es O, N, S o no está presente; m es 0-10, R2 es H, un aminoácido, un polipéptido, o un grupo de modificación amino terminal, y R3 es H, un aminoácido, un polipéptido, o un grupo de modificación terminal carboxi.

20 En algunas realizaciones, el polipéptido es un agonista de polipéptido de insulina, agonista parcial, antagonista, antagonista parcial, o agonista inverso. En algunas realizaciones, el agonista de polipéptido de insulina, agonista parcial, antagonista, antagonista parcial, o agonista inverso comprende un aminoácido no naturalmente codificado unido a un polímero soluble en agua. En algunas realizaciones, el polímero soluble en agua comprende un resto de poli(etilenglicol). En algunas realizaciones, el agonista de polipéptido de insulina, agonista parcial, antagonista, antagonista parcial, o agonista inverso comprende un amino ácido no naturalmente codificado y uno o más modificación postraduccional, enlazador, polímero, o molécula biológicamente activa.

25 La presente divulgación también proporciona ácidos nucleicos aislados que comprenden un polinucleótido que se hibrida en condiciones rigurosas ácidos nucleicos que codifican polipéptidos de insulina de SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, y 12. La presente descripción también proporciona ácidos nucleicos aislados que comprenden un polinucleótido que se hibrida en condiciones rigurosas de ácidos nucleicos que codifican polipéptidos de insulina de SEQ ID NOs: 1 y 2. La presente descripción también proporciona ácidos nucleicos aislados que comprenden un polinucleótido o polinucleótidos que hibridan en condiciones rigurosas a los polinucleótidos que codifican polipéptidos mostrados como SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 y 12 donde el polinucleótido comprende al menos un codón selector. La presente divulgación también proporciona ácidos nucleicos aislados que comprenden un polinucleótido o polinucleótidos que se hibridan en condiciones rigurosas a los polinucleótidos que codifican polipéptidos mostrados como SEQ ID NOs: 1 y 2 en el que el polinucleótido comprende al menos un codón selector. La presente descripción proporciona ácidos nucleicos aislados que comprende un polinucleótido que codifica los polipéptidos que se muestran como SEQ ID NOs: 1 y 2. La presente descripción también proporciona ácidos nucleicos aislados que comprende un polinucleótido que codifica los polipéptidos que se muestran como SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, y 12. La presente descripción proporciona ácidos nucleicos aislados que comprenden un polinucleótido que codifica los polipéptidos que se muestran como SEQ ID NOs: 1 y 2 con uno o más aminoácidos no naturales codificados. La presente divulgación también proporciona ácidos nucleicos aislados que comprende un polinucleótido que codifica los polipéptidos que se muestran como SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, y 12 con uno o más aminoácidos no naturales codificados. Es fácilmente evidente para los expertos ordinarios en la técnica que un número de diferentes polinucleótidos puede codificar cualquier polipéptido de la presente invención.

45 En algunas realizaciones, el codón selector se selecciona del grupo que consiste de un codón ámbar, el codón ocre, codón ópalo, un codón único, un codón raro, un codón de cinco bases, y un codón de cuatro bases.

50 La presente descripción también proporciona métodos de fabricación de un polipéptido de insulina unido a un polímero soluble en agua. En algunos casos de la descripción, el método comprende la puesta en contacto de un polipéptido de insulina aislada que comprende un aminoácido no naturalmente codificado con un polímero soluble en agua que comprende un resto que reacciona con el aminoácido no naturalmente codificado. En algunas formas de realización, el aminoácido no naturalmente codificado incorporado en el polipéptido de insulina es reactivo frente a un polímero soluble en agua que es de otra manera no reactivo hacia cualquiera de los 20 aminoácidos comunes. En algunas formas de realización, el aminoácido no naturalmente codificado incorporado en el polipéptido de insulina es reactivo frente a un enlazador, polímero, o molécula biológicamente activa que es de otra manera no reactiva hacia cualquiera de los 20 aminoácidos comunes.

60 En todas las realizaciones de la invención, el polímero soluble en agua ligada al cadena A de polipéptido de insulina en la posición 14 de la secuencia de aminoácidos de cadena A seleccionada del grupo SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, y SEQ ID NO: 11, en el que el aminoácido en la posición 14 de la cadena A es un aminoácido no naturalmente codificado está unido por un enlace covalente de oxima. En algunas realizaciones, el polipéptido de insulina conectado al polímero soluble en agua se prepara haciendo reaccionar un polipéptido de insulina que comprende un aminoácido que contiene carbonilo con una molécula de poli(etilenglicol) que comprende un aminooxi, hidrazina, hidrazida o un grupo semicarbazida. En algunas realizaciones, el grupo

aminooxi, hidrazina, hidrazida o semicarbazida se une a la molécula poli(etilenglicol) a través de un enlace amida. En algunas realizaciones, el grupo aminooxi, hidrazina, hidrazida o semicarbazida se une a la molécula poli(etilenglicol) a través de un enlace carbamato.

5 En algunas realizaciones, el polipéptido de insulina conectado al polímero soluble en agua se prepara haciendo reaccionar una molécula de poli(etilenglicol) que comprende un grupo carbonilo con un polipéptido que comprende un aminoácido no naturalmente codificado que comprende un grupo aminooxi, hidrazina, hidrazida o semicarbazida.

10 El polipéptido de insulina puede estar enlazado a un polímero soluble en agua por reacción de un polipéptido de insulina que comprende un aminoácido que contiene alquilona con una molécula de poli(etilenglicol) que comprende un resto azida. En algunos casos, el grupo azida o alquino está ligado a la molécula poli(etilenglicol) a través de un enlace amida.

15 El polipéptido de insulina puede estar enlazado a un polímero soluble en agua por reacción de un polipéptido de insulina que comprende un aminoácido que contiene azida con una molécula de poli(etilenglicol) que comprende un resto de alquino. En algunos casos, el grupo azida o alquino está ligado a la molécula poli(etilenglicol) a través de un enlace amida.

20 En algunas realizaciones, la molécula de poli(etilenglicol) tiene un peso molecular de entre aproximadamente 0,1 kDa y aproximadamente 100 kDa. En algunas realizaciones, la molécula de poli(etilenglicol) tiene un peso molecular de entre 0,1 kDa y 50 kDa.

25 En algunas realizaciones, la molécula de poli(etilenglicol) es un polímero ramificado. En algunas realizaciones, cada rama del polímero ramificado de poli(etilenglicol) tiene un peso molecular de entre 1 kDa y 100 kDa, o entre 1 kDa y 50 kDa.

30 En algunas realizaciones, el polímero soluble en agua ligado al polipéptido de insulina comprende un resto de polialquilenglicol. En algunas realizaciones, el residuo de aminoácido no naturalmente codificado incorporado en el polipéptido de insulina comprende un grupo carbonilo, un grupo aminooxi, un grupo hidrazida, una hidrazina, un grupo semicarbazida, un grupo azida o un grupo alquino. En algunas realizaciones, el residuo de aminoácido no naturalmente codificado incorporado en el polipéptido de insulina comprende un resto carbonilo y el polímero soluble en agua comprende un aminooxi, hidrazida, hidrazina, o resto de semicarbazida. En algunas realizaciones, el residuo de aminoácido no naturalmente codificado incorporado en el polipéptido de insulina comprende un resto alquino y el polímero soluble en agua comprende un resto azida. En algunas realizaciones, el residuo de aminoácido no naturalmente codificado incorporado en el polipéptido de insulina comprende un resto azida y el polímero soluble en agua comprende un resto alquino.

35 La presente divulgación también proporciona composiciones que comprenden un polipéptido de insulina que comprende un aminoácido no naturalmente codificado y un vehículo farmacéuticamente aceptable. En algunos casos de la descripción, el aminoácido no naturalmente codificado está ligado a un polímero soluble en agua.

40 La presente divulgación también proporciona células que comprenden un polinucleótido que codifica el polipéptido de insulina que comprende un codón selector. En algunos casos de la descripción, las células comprenden una sintetasa ortogonal ARN y/o un ARNt ortogonal para la sustitución de un aminoácido no naturalmente codificado en el polipéptido de insulina.

45 La presente descripción también proporciona métodos de fabricación de un polipéptido de insulina que comprende un aminoácido no naturalmente codificado. En algunos casos de la descripción, los métodos comprenden células que comprenden un polinucleótido o polinucleótidos que codifican un polipéptido de insulina, una sintetasa ortogonal ARN y/o un cultivo de ARNt ortogonal en condiciones que permiten la expresión del polipéptido de insulina; y purifican el polipéptido de insulina de las células y/o medio de cultivo.

50 La presente descripción también proporciona métodos para aumentar la vida media terapéutica, vida media en suero o el tiempo de circulación de polipéptidos de insulina. La presente divulgación también proporciona procedimientos de modulación de la inmunogenicidad de los polipéptidos de insulina. En algunos casos de la descripción, los métodos comprenden la sustitución de un aminoácido no naturalmente codificado por uno cualquiera o más aminoácidos en polipéptidos naturales de insulina y/o la unión del polipéptido de insulina a un enlazador, un polímero, un polímero soluble en agua, o una molécula biológicamente activa.

55 La presente divulgación también proporciona procedimientos de tratamiento de un paciente en necesidad de tal tratamiento con una cantidad eficaz de una molécula de insulina de la presente descripción. En algunos casos de la descripción, los métodos comprenden administrar al paciente una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende un polipéptido de insulina que comprende un aminoácido no naturalmente codificado y un vehículo farmacéuticamente aceptable. En algunos casos de la descripción, el aminoácido no naturalmente codificado está ligado a un polímero soluble en agua. En algunos casos de la descripción, el polipéptido de insulina es glicosilado. En algunas realizaciones, el polipéptido de insulina no está glicosilado.

60 La presente divulgación también proporciona procedimientos de tratamiento de un paciente en necesidad de tal tratamiento con una cantidad eficaz de una molécula de insulina de la presente descripción. En algunos casos de la descripción, los métodos comprenden administrar al paciente una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende un polipéptido de insulina que comprende un aminoácido no naturalmente codificado y un vehículo farmacéuticamente aceptable. En algunos casos de la descripción, el aminoácido no naturalmente codificado está ligado a un polímero soluble en agua. En algunos casos de la descripción, el polipéptido de insulina es glicosilado. En algunas realizaciones, el polipéptido de insulina no está glicosilado.

65 La presente divulgación también proporciona procedimientos de tratamiento de un paciente en necesidad de tal tratamiento con una cantidad eficaz de una molécula de insulina de la presente descripción. En algunos casos de la descripción, los métodos comprenden administrar al paciente una cantidad terapéuticamente eficaz de una composición farmacéutica que comprende un polipéptido de insulina que comprende un aminoácido no naturalmente codificado y un vehículo farmacéuticamente aceptable. En algunos casos de la descripción, el aminoácido no naturalmente codificado está ligado a un polímero soluble en agua. En algunos casos de la descripción, el polipéptido de insulina es glicosilado. En algunas realizaciones, el polipéptido de insulina no está glicosilado.

5 Los polipéptidos de insulina que comprenden una secuencia mostrada en SEQ ID NO: 1 y 2, o cualquier otra secuencia de polipéptido de insulina (un ejemplo no limitante de estos sería SEQ ID NOs: 3 a 12) excepto que al menos un aminoácido está sustituido por un aminoácido no naturalmente codificado se describen. En algunos casos, el aminoácido no naturalmente codificado está ligado a un polímero soluble en agua. En algunos casos, el polímero soluble en agua comprende un resto de poli(etilenglicol). En algunos casos, el aminoácido no naturalmente codificado comprende un grupo carbonilo, un grupo aminoóxi, un grupo hidrazida, un grupo hidrazina, un grupo semicarbazida, un grupo azida o un grupo alquino.

10 La presente divulgación también proporciona composiciones farmacéuticas que comprenden un vehículo farmacéuticamente aceptable y un polipéptido de insulina que comprende la secuencia mostrada en SEQ ID NOs: 1 a 12, o cualquier otra secuencia de polipéptido de insulina, en el que al menos un aminoácido está sustituido por un aminoácido no naturalmente codificado. La presente divulgación también proporciona composiciones farmacéuticas que comprenden un vehículo farmacéuticamente aceptable y un polipéptido de insulina que comprende una cadena A y B (por ejemplo, SEQ ID NOs: 1 y 2 harían la insulina, SEQ ID NOs 3 y 4 harían insulina lispro, etc.) la secuencia mostrada en SEQ ID NOs: 1 a 12, o cualquier otra secuencia de polipéptido de insulina, en la que al menos un aminoácido está sustituido por un aminoácido no naturalmente codificado. La presente divulgación también proporciona composiciones farmacéuticas que comprenden un vehículo farmacéuticamente aceptable y un polipéptido de insulina que comprende la secuencia mostrada en SEQ ID NO: 1 y 2. La presente descripción también proporciona composiciones farmacéuticas que comprenden un vehículo farmacéuticamente aceptable y un polipéptido de insulina que comprende la secuencia mostrada en SEQ ID NO: 1 a 12. En algunas realizaciones, el aminoácido no naturalmente codificado comprende un resto sacárido. En algunas realizaciones, el polímero soluble en agua está relacionado con el polipéptido mediante un resto de sacárido. En algunas realizaciones, un enlazador, polímero, o molécula biológicamente activa está relacionada con el polipéptido de insulina a través de un resto de sacárido.

La presente divulgación también proporciona un polipéptido de insulina que comprende un polímero soluble en agua unido por un enlace covalente con el polipéptido de insulina en un solo aminoácido. En algunos casos, el polímero soluble en agua comprende un resto de poli(etilenglicol). En algunos casos, el aminoácido unido covalentemente al polímero soluble en agua es un aminoácido no naturalmente codificado presente en el polipéptido.

La presente divulgación proporciona un polipéptido de insulina que comprende al menos un enlazador, polímero, o molécula biológicamente activa, en donde que dicho enlazador, polímero, o molécula biológicamente activa se une al polipéptido a través de un grupo funcional de un aminoácido no naturalmente codificado ribosomalmente incorporado en el polipéptido. En algunas realizaciones, el polipéptido se monopegiló. La presente divulgación también proporciona un polipéptido de insulina que comprende un enlazador, polímero, o molécula biológicamente activa que se une a uno o más aminoácidos no naturalmente codificados que dicho aminoácido no naturalmente codificado se incorpora en el ribosoma en el polipéptido en sitios preseleccionados. En todos los casos, los polipéptidos de insulina de la invención comprenden un polímero soluble en agua unida por un enlace covalente de oxima con el cadena A de polipéptido de insulina en la posición 14 de la secuencia de aminoácidos de cadena A seleccionada del grupo SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, y SEQ ID NO: 11, en donde el aminoácido en la posición 14 de la cadena A es un aminoácido no naturalmente codificado.

Se incluyen dentro del alcance de esta invención el líder de la insulina o la secuencia de señal, un ejemplo de la cual se puede considerar como proinsulina. La secuencia líder o señal heteróloga seleccionada debe ser una que sea reconocida y procesada, por ejemplo, sistema de secreción de la célula huésped para secretar y, posiblemente, escindirse por una peptidasa de señal, por la célula huésped. Un método para tratar una afección o trastorno con insulina o un polipéptido de insulina o análogo de la presente invención está destinado a implicar el tratamiento con insulina con o sin un péptido de señal o líder.

La presente descripción también proporciona métodos para inducir un aumento en el metabolismo de la glucosa, comprendiendo dicho método la administración de insulina a dichas células en una cantidad eficaz para inducir un aumento en la actividad metabólica de la glucosa.

55 En otra realización, la conjugación del polipéptido de insulina que comprende uno o más aminoácidos no naturales codificados a otra molécula, incluyendo, pero no limitado a, PEG, proporciona insulina sustancialmente purificada debido a la reacción química única utilizada para la conjugación con el aminoácido no natural. La conjugación de la insulina que comprende uno o más aminoácidos no naturales codificados a otra molécula, tales como PEG, puede realizarse con otras técnicas de purificación llevadas a cabo antes o después de la etapa de conjugación para proporcionar insulina sustancialmente pura.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

65 La Figura 1 muestra dos modelos de la estructura cristalina de la insulina se muestran junto con la secuencia de aminoácidos de la insulina.

La Figura 2 es un modelo de la estructura cristalina de la insulina. Se muestran los sitios seleccionados para la sustitución con un aminoácido no naturalmente codificado.

La Figura 3 es un dibujo de la proinsulina primaria.

La Figura 4 muestra dibujos de insulina, jumalog, novoLog, glargina y detemir.

La Figura 5 muestra la expresión de proinsulina de insulina de plásmido pVK6-LisPro, así como la secuencia del inserto lispro.

La Figura 6 muestra un análisis en gel de SDS-PAGE con polipéptidos de insulina de la presente invención incluyen aquellos sustituidos en cadena A de Q15 (SEQ ID NO: 3), la cadena A de Y14 A (SEQ ID NO: 3), la cadena B de R22 (SEQ ID NO: 4), la cadena B de FI (SEQ ID NO: 4), la cadena A de G1 (SEQ ID NO: 3), la cadena A de S9 (SEQ ID NO: 3), y la cadena B de K28 (SEQ ID NO: 4), cada sustitución hecha con un aminoácido no natural (pAF).

La Figura 7 muestra polipéptidos de insulina SEC-HPLC de pegilados y no pegilados.

La Figura 8 muestra un diagrama de un plásmido utilizado para la clonación en Pichia.

La Figura 9 muestra un área bajo el gráfico de barras de la curva por dos polipéptidos de insulina de la invención presente, A14pAF-PEG-A21N y A14pAF-PEG-A21G.

La Figura 10 muestra un gel de los resultados de la reacción de replegamiento para el polipéptido de insulina A21G, como se discute adicionalmente en el Ejemplo 36.

La Figura 11 muestra un HPLC de proinsulina de QHP purificada y replegada a partir del Ejemplo 36.

La Figura 12 muestra un HPLC de la reacción de PEGilación descrita en el Ejemplo 36.

## DEFINICIONES

Tal como se utiliza aquí y en las reivindicaciones adjuntas, las formas singulares "un", "una", "el" y "la" incluyen la referencia en plural a menos que el contexto indique claramente lo contrario. Así, por ejemplo, la referencia a "insulina" o "polipéptido de insulina" y varias formas con guiones y sin guiones es una referencia a una o más de tales proteínas e incluye equivalentes de los mismos conocidos por los expertos normales en la técnica, y así sucesivamente.

El término "sustancialmente purificado" se refiere a un polipéptido de insulina que puede ser sustancial o esencialmente libre de componentes que normalmente acompañan o interaccionan con la proteína como se encuentra en su entorno natural, es decir, una célula nativa, o célula huésped en el caso de polipéptidos de forma recombinante de insulina producida. El polipéptido de insulina que puede ser sustancialmente libre de material celular incluye preparaciones de proteína tienen menos de aproximadamente 30%, menos de aproximadamente 25%, menos de aproximadamente 20%, menos de aproximadamente 15%, menos de aproximadamente 10%, menos de aproximadamente 5%, menos de aproximadamente 4%, menos de aproximadamente 3%, menos de aproximadamente 2%, o menos de aproximadamente 1% (en peso seco) de proteína contaminante. Cuando el polipéptido de insulina o variante de la misma se produce recombinantemente por las células huésped, la proteína 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. Cuando el polipéptido de insulina o variante del mismo se produce recombinantemente por las células huésped, la proteína 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 500mg/l, aproximadamente 200mg/l, sobre 100mg/l, aproximadamente 50mg/l, aproximadamente 10mg/l, o aproximadamente 1mg/l o menos del peso seco de las células. Por lo tanto, polipéptido de insulina "sustancialmente purificado" como el producido por los métodos de la presente descripción pueden tener un nivel de pureza de al menos aproximadamente 30%, al menos aproximadamente 35%, al menos aproximadamente 40%, al menos aproximadamente 45%, al menos aproximadamente 50%, al menos aproximadamente 55%, al menos aproximadamente 60%, al menos aproximadamente 65%, al menos aproximadamente 70%, específicamente, un nivel de pureza de al menos aproximadamente 75%, 80%, 85%, y más específicamente, un nivel de pureza de al menos aproximadamente 90%, un nivel de pureza de al menos aproximadamente 95%, un nivel de pureza de al menos aproximadamente 99% o mayor tal como se determina mediante métodos apropiados tales como el análisis SDS/PAGE, RP-HPLC, SEC, y la electroforesis capilar.

Una "célula huésped recombinante" o "célula huésped" se refiere a una célula que incluye un polinucleótido exógeno, independientemente del método empleado para la inserción, por ejemplo, captación directa, transducción, f-acoplamiento, u otros métodos conocidos en la técnica para crear células de huésped recombinantes. El

polinucleótido exógeno puede mantenerse como un vector no integrado, por ejemplo, un plásmido, o alternativamente, puede integrarse en el genoma del huésped.

Tal como se utiliza aquí, el término "medio" o "medios de comunicación" incluye cualquier medio de cultivo, solución, sólido, semisólido, o soporte rígido que pueden apoyar o contener cualquier célula huésped, incluyendo células huésped bacterianas, células huésped de levadura, células huésped de insectos, células de la planta huésped, las células huésped eucariotas, células huésped de mamífero, células CHO, células huésped procariotas, E. coli, o células huésped de pseudomonas, y contenidos celulares. Así, el término puede abarcar medio en el que la célula huésped ha sido cultivado, por ejemplo, medio en el que se ha secretado el polipéptido de insulina, incluido el medio ya sea antes o después de una etapa de proliferación. El término también puede abarcar tampones o reactivos que contienen lisados de células huésped, tales como en la facilidad en la que el polipéptido de insulina se produce intracelularmente y las células huésped se lisan o se interrumpen para liberar el polipéptido de insulina.

"Agente reductor," Tal como se usa en el presente documento con respecto al replegamiento de proteínas, se define como cualquier compuesto o material que mantiene los grupos de sulfhidrilo en el estado reducido y reduce los enlaces de disulfuro intra o intermoleculares. Los agentes reductores adecuados incluyen, pero no se limitan a, ditioneitol (DTT), 2-mercaptoetanol, ditioneitol, cisteína, cisteamina (2-aminoetanol), y glutatión reducido. Es fácilmente evidente para los expertos ordinarios en la técnica que una amplia variedad de agentes reductores son apropiados para uso en los métodos y composiciones de la presente descripción.

"Agente oxidante", tal como se usa aquí respecto al replegamiento de proteínas, se define como cualquier compuesto o material que es capaz de eliminar un electrón de un compuesto que está siendo oxidado. Los agentes oxidantes adecuados incluyen, pero no se limitan a, glutatión oxidado, cistina, cistamina, ditioneitol oxidado, eritritol oxidado y oxígeno. Es fácilmente evidente para los expertos ordinarios en la técnica que una amplia variedad de agentes oxidantes son adecuados para uso en los métodos de la presente descripción.

"Agente desnaturalizante" o "desnaturalizante", tal como se usa aquí, se define como cualquier compuesto o material que provocará un despliegue reversible de una proteína. La fuerza de un agente desnaturalizante o desnaturalizante será determinado tanto por las propiedades como la concentración del agente desnaturalizante en particular o desnaturalizante. Agentes desnaturalizantes adecuados o desnaturalizantes pueden ser caotropes, detergentes, disolventes orgánicos, disolventes miscibles en agua, lípidos de fosfo, o una combinación de dos o más de tales agentes. caotropes adecuados incluyen, pero no se limitan a, urea, guanidina, y tiocianato de sodio. Detergentes útiles pueden incluir, pero no se limitan a, detergentes fuertes tales como sulfato de dodecilo sódico, o éteres de polioxietileno (por ejemplo, detergentes Tween o Triton), Sarkosyl, detergentes no iónicos leves (por ejemplo, digitonina), detergentes suaves catiónicos tales como N->2,3-(Dioleoyl)-propilo-N,N,N-trimetilamoniaco, detergentes iónicos suaves (por ejemplo colato de sodio o desoxicolato de sodio) o detergentes de ion híbrido, incluyendo, pero no limitado a, sulfobetaínas (Zwittergent), 3-(3-cloramidopropilo)dimetilamoni-1-propano sulfato (CHAPS), y 3-(3-cloramidopropilo)dimetilamoni-2-hidroxi-1-sulfonato de propano (CHAPSO). Disolventes orgánicos, miscibles con agua tales como acetonitrilo, alcanoles inferiores (especialmente alcanoles C2 - C4 tales como dimetilol o isopropanol), o alcanodiolos inferiores (especialmente alcanodiolos C2 - C4, tales como etilenglicol) pueden ser utilizados como desnaturalizantes. Los fosfolípidos útiles en la presente invención pueden ser fosfolípidos de origen natural tales como fosfatidilodimetilolamina, fosfatidilcolina, fosfatidilserina, y fosfatidilinositol o derivados de fosfolípido sintéticos o variantes como dihexanoilfosfatidilcolina o diheptanoilfosfatidilcolina.

"Replegamiento," tal como se usa en el presente documento describe cualquier proceso, reacción o método que transforma el enlace de disulfuro que contiene polipéptidos de un estado incorrectamente plegado o desplegado a una conformación nativa o correctamente plegada con respecto a enlaces de disulfuro.

"Coplegamiento," tal como se usa aquí, se refiere específicamente a procesos de replegamiento, reacciones, o métodos que emplean al menos dos polipéptidos que interactúan entre sí y dan como resultado la transformación de polipéptidos no plegados o incorrectamente plegados a polipéptidos nativos, correctamente plegados.

El término "proinsulina", tal como se usa en el presente documento es una proteína correctamente en cruz de la fórmula: B-C-A

en donde:

A es la cadena A de la insulina o un derivado funcional de la misma;

B es la cadena B de la insulina o un derivado funcional de la misma que tiene un grupo de amino epsilon; y

C es el péptido de conexión de proinsulina. Preferentemente, la proinsulina es la cadena A de la insulina humana, la cadena B de la insulina humana, y C es el péptido de conexión natural. Cuando la proinsulina es la secuencia natural, la proinsulina posee tres grupos amino libres: fenilo-alanina(1) (.alfa.-grupo amino), lisina(29) (.epsilon.-grupo amino) y lisina(64) (.epsilon.-grupo amino).

El término "análogo de insulina" tal como se usa en el presente documento es una proteína correctamente en cruz que exhibe actividad de la insulina de la fórmula

A-B

en donde:



A es la cadena A de la insulina o un derivado funcional de la cadena A de insulina; y  
B es la cadena B de la insulina o un derivado funcional de la cadena B de insulina que tiene un.ε-grupo amino y al menos uno de A o B contiene una modificación de aminoácidos de la secuencia natural.

5 En la presente memoria, siempre que el término insulina se usa en un plural o un sentido genérico se pretende abarcar ambas insulinas de origen natural y análogos de insulina y derivados de los mismos. Por "polipéptido de insulina" tal como se usa en el presente documento, se quiere decir un compuesto que tiene una estructura molecular similar a la de la insulina humana, incluyendo los puentes de disulfuro entre Cys.sup.A7 y Cys.sup.B7 y entre Cys.sup.A20 y Cys.sup.B19 y un puente de disulfuro interno entre Cys.sup.A6 y Cys.sup.A11, y que tienen actividad de la insulina.

10 El término "insulina" tal como se usa en el presente documento, se refiere a insulina humana, cuya secuencia y estructura espacial de aminoácidos son bien conocidas. La insulina humana está compuesta de una cadena A de veintinueve aminoácidos y una Cadena B de treinta aminoácidos que están reticulados por enlaces de disulfuro. Una insulina adecuadamente reticulada contiene tres puentes de disulfuro: uno entre la posición 7 de la cadena A y la posición 7 de la Cadena B, un segundo entre la posición 20 de la cadena A y la posición 19 de la Cadena B, y un tercero entre las posiciones 6 y 11 de la cadena A [Nicol, D.S.H.W. y Smith, L.F., Nature, 187,483-485 (1960)].

15 péptidos de insulina incluyendo, pero no limitado a, insulina, humano; insulina, porcino; IGF-I, humano; factor de crecimiento similar a la insulina 11 (69-84); factor de crecimiento proinsulinLike 11 (68-102), humano; factor de crecimiento similar a insulina 11 (105-128), humano; [AspB28]insulina, humana; [LysB28] insulina, humana; [LeuB28] insulina, humana; [ValB28] insulina, humana; [AlaB28] insulina, humana; [AspB28, ProB29] insulina, humana; [LysB28, ProB29] insulina, humana; [LeuB28, ProB29] insulina, humana; [ValB28, ProB29] insulina, humana; [AlaB28, ProB29] insulina, humana; [GlyA21] insulina, humana; [GlyA21 GlnB3] insulina, humana; [AlaA21] insulina, humana; [AlaA21 Gln.sup.B3] insulina, humana; [GlnB3] insulina, humana; [GlnB30] insulina, humana; [GlyA21 GluB30] insulina, humana; [GlyA21 GlnB3 GluB30] insulina, humana; [G1nB3 GluB30] insulina, humana; B22B30 insulina, humana; B23B30 insulina, humana; B25B30 insulina, humana; B26B30 insulina, humana; B27B30 insulina, humana; B29B30 insulina, humana; la cadena A de la insulina humana, y la cadena B de la insulina humana.

20 El término "análogo de insulina" significa una proteína que tiene una cadena A y una Cadena B que tienen sustancialmente las mismas secuencias de aminoácidos como la Cadena A y/o Cadena B de la insulina humana, respectivamente, pero difiere de la cadena A y Cadena B de la insulina humana por que tiene una o más deleciones de aminoácidos, uno o más reemplazos de aminoácidos, y/o una o más adiciones de aminoácidos que no destruyen la actividad de la insulina del análogo de insulina. Un análogo de insulina que tiene un punto isoeléctrico que es "mayor que" el punto isoeléctrico de la insulina es un tipo de análogo de insulina. Otro tipo de análogo de insulina es un "análogo de insulina monomérico."

25 Un "análogo de insulina monomérico" es un análogo que actúa rápidamente de la insulina humana, incluyendo, por ejemplo, insulina humana en la que Pro en la posición B28 se sustituye con Asp, Lys, Leu, Val, o Ala, y en el que Lys en la posición B29 es Lys o está sustituido con Pro. Otro análogo de insulina monomérico, también conocido como des(B27) insulina humana, es insulina humana en la que Thr en la posición 27 de la Cadena B se elimina. Análogos de insulina monoméricos se describen en Chance, R.E., *et al.*, Patente de Estados Unidos. n° 5.514.646, concedida el 7 de mayo de 1996; Brems, D.N., *et al.* Protein Engineering, 5, 527-533 (1992); Brange, J.J.V., *et al.*, publicación EPO n° 214.826 (publicada el 18 de marzo de 1987).; y Brange, J.J.V., *et al.*, Current Opinion in Structural Biology, 1, 934-940 (1991). Los análogos de insulina monoméricos empleados en las presentes formulaciones se reticulan adecuadamente en las mismas posiciones que en la insulina humana.

30 Péptidos de insulina incluyendo, pero no limitado a, insulina, humano; insulina, porcino; TFIG, humano; factor de crecimiento similar a la insulina II (69-84); factor de crecimiento similar a insulina II (68-102), humano; factor de crecimiento similar a insulina 11 (105-128), humano; [AspB28] insulina, humana; [LysB28] insulina, humana; [LeuB28] insulina, humana; [ValB28] insulina, humana; [AlaB28] insulina, humana; [AspB28, ProB29] insulina, humana; [LysB28, ProB29] insulina, humano; [LeuB28, ProB29] insulina, humana; [ValB28, ProB29] insulina, humana; [AlaB28, ProB29] insulina, humana; [GlyA21] insulina, humana; [GlyA21 GlnB3] insulina, humana; [AlaA21] insulina, humana; [AlaA21 Gln. sup.B3] insulina, humana; [GlnB3] insulina, humana; [GlnB30] insulina, humana; [GlyA21 GluB30] insulina, humana; [G1yA21 G1nB3 GluB30] insulina, humana; [GlnB3 GluB30] insulina, humano; B22B30 insulina, humana; B23B30 insulina, humana; B25B30 insulina, humana; B26B30 insulina, humana; B27B30 insulina, humana; B29B30 insulina, humana; la cadena A de la insulina humana, y la cadena B de la insulina humana.

35 En un aspecto adicional, la descripción proporciona ácidos nucleicos recombinantes que codifican las proteínas variantes, los vectores de expresión que contienen los ácidos nucleicos variantes, las células huésped que comprenden los ácidos nucleicos variantes y/o vectores de expresión, y métodos para producir las proteínas variantes. En un aspecto adicional, la descripción proporciona el tratamiento de un trastorno sensible a la insulina mediante la administración a un paciente de una proteína variante, por lo general con un vehículo farmacéutico, en una cantidad terapéuticamente eficaz. En un aspecto adicional, la descripción proporciona métodos para modular la

inmunogenicidad (en particular reducir la inmunogenicidad) de polipéptidos de insulina mediante la alteración de epítomos de MHC de clase II.

5 El término "polipéptido de insulina" también incluye las sales farmacéuticamente aceptables y profármacos, y profármacos de las sales, polimorfos, hidratos, solvatos, fragmentos biológicamente activos, variantes biológicamente activas y estereoisómeros de la insulina en estado natural, así como agonistas, variantes miméticas, y antagonistas de las fusiones en estado natural de insulina y polipéptido de la mismas fusiones que comprenden aminoácidos adicionales en el extremo amino, extremo carboxilo terminal, o ambos, están abarcados por el término "polipéptido de insulina." fusiones ejemplares incluyen, pero no se limitan a, por ejemplo, insulina de metionilo en la  
10 que una metionina está enlazada al término N de la insulina que resulta de la expresión recombinante de la forma madura de la insulina que carece del péptido líder o señal o parte del mismo (una metionina está enlazada al término N de la insulina que resulta de la expresión recombinante), las fusiones con el propósito de purificación (incluyendo, pero no limitado a polihistidina o epítomos de afinidad), fusiones con albúmina de suero, péptidos y fusiones de unión con las proteínas del suero tales como albúmina de suero. Patente de Estados Unidos nº 5.750.373, describe un método para seleccionar nuevas proteínas tales como la hormona del crecimiento y variantes de fragmento de anticuerpo que tienen propiedades de unión alterada por sus respectivas moléculas de receptor. El método comprende la fusión de un gen que codifica una proteína de interés al dominio carboxi terminal de la proteína de recubrimiento del gen III del fago filamentoso M13. Moléculas quiméricas que comprenden insulina y una o más  
15 otras moléculas. La molécula quimérica puede contener regiones o fragmentos de una o ambas de la insulina y la otra molécula específica. Cualquiera de estos fragmentos se pueden preparar a partir de las proteínas por métodos bioquímicos estándar, o mediante la expresión de un polinucleótido que codifica el fragmento. La insulina, o un fragmento de la misma, pueden producirse como una proteína de fusión que comprende albúmina de suero humano (HSA), Fe, o una porción del mismo. Tales construcciones de fusión son apropiadas para potenciar la expresión de la insulina, o fragmento de la misma, en una célula huésped eucariota. Porciones HSA ejemplares incluyen el polipéptido N-terminal (aminoácidos 1-369, 1-419, y las longitudes intermedias que empiezan con el aminoácido 1), como se describe en la patente de los Estados Unidos. nº 5.766.883, y la publicación WO 97/24445. Otros polipéptidos quiméricos pueden incluir una proteína HSA con insulina, o fragmentos de los mismos, que se adjuntan a cada uno de los extremos de terminal C y terminal N de la HSA. Tales construcciones de HSA se describen en la patente de los Estados Unidos nº 5.876.969. Otras fusiones pueden ser creadas por la fusión de la insulina con a) la parte Fc de una inmunoglobulina; b) un análogo de la parte Fc de una inmunoglobulina; y c) fragmentos de la porción Fc de una inmunoglobulina.

Varias referencias describen la modificación de polipéptidos mediante la conjugación del polímero o glicosilación. El término "polipéptido de insulina" incluye polipéptidos conjugados con un polímero tal como PEG y puede estar  
35 compuesto de uno o más derivitizaciones de cisteína, lisina, u otros residuos. Además, el polipéptido de insulina puede comprender un enlazador o polímero, en el que el aminoácido al que el engarce o polímero está conjugado puede ser un aminoácido no natural según la presente invención, o puede conjugarse con un aminoácido naturalmente codificado mediante el uso de técnicas conocidas en la técnica tales como el acoplamiento a la utilización de lisina o cisteína.

40 El término "polipéptido de insulina" también incluye la insulina glicosilada, tales como, pero no limitado a, los polipéptidos glicosilados en cualquier posición de aminoácido, formas glicosiladas N-enlazadas o O-enlazadas del polipéptido. Las variantes que contienen cambios de nucleótido único también se consideran como variantes biológicamente activas de polipéptido de insulina. Además, también se incluyen variantes de empalme. El término "polipéptido de insulina" también incluye heterodímeros de polipéptido de insulina, homodímeros, heteromultímeros, o homomultímeros de uno cualquiera o más de polipéptidos de insulina o cualquier otro polipéptido, proteína, hidratos de carbono, polímeros, moléculas pequeñas, enlazadores, ligando, u otra molécula biológicamente activa de cualquier tipo, unidos por medios químicos o expresados como una proteína de fusión, así como los análogos de polipéptidos que contienen, por ejemplo, deleciones específicas u otras modificaciones aún mantienen la actividad biológica.

55 El término "polipéptido de insulina" o "insulina" abarca polipéptidos de insulina que comprenden una o más sustituciones de aminoácidos, adiciones o supresiones. polipéptidos de insulina de la presente invención pueden estar compuestos de modificaciones con uno o más aminoácidos naturales en conjunción con una o más modificaciones de aminoácido no natural. Se han descrito las sustituciones ejemplares en una amplia variedad de posiciones de aminoácidos en polipéptidos de insulina natural, incluyendo pero no limitado a sustituciones que modulan la estabilidad farmacéutica, que modulan una o más de las actividades biológicas del polipéptido de insulina, tales como, pero no limitado a, aumentar la actividad agonista, aumentar la solubilidad del polipéptido, disminuir la susceptibilidad de la proteasa, convertir el polipéptido en un antagonista, etc., y están abarcados por el término "polipéptido de insulina." En algunas realizaciones, el antagonista de la insulina comprende un aminoácido no naturalmente codificado unido a un polímero soluble en agua que está presente en una región de unión al receptor de la molécula de insulina.

65 En algunas realizaciones, los polipéptidos de insulina comprenden además una adición, sustitución o deleción que modula la actividad biológica del polipéptido de insulina. En algunas realizaciones, los polipéptidos de insulina comprenden además una adición, sustitución o deleción que modula la actividad antiviral del polipéptido de insulina.

En algunas realizaciones, los polipéptidos de insulina comprenden además una adición, sustitución o delección que aumenta la actividad antiviral del polipéptido de insulina. Por ejemplo, las adiciones, sustituciones o delecciones pueden modular una o más propiedades o actividades de la insulina. Por ejemplo, las adiciones, sustituciones o delecciones pueden modular la afinidad por el receptor de insulina, modular la vida media circulante, modular la semivida terapéutica, modular la estabilidad del polipéptido, modular la escisión por proteasas, modular la dosis, modular la liberación o la biodisponibilidad, facilitan la purificación, o mejoran o alteran una ruta particular de administración. Del mismo modo, los polipéptidos de insulina pueden comprender secuencias de escisión de la proteasa, grupos reactivos, dominios de unión de anticuerpos (incluyendo pero no limitado a, FLAG o poli-His) o de otras secuencias basadas en la afinidad (incluyendo pero no limitado a, FLAG, poli-His, GST, etc.) o moléculas unidas (incluyendo pero no limitado a, biotina) que mejoran la detección (incluyendo pero no limitado a, GFP), la purificación u otros rasgos del polipéptido.

El término "polipéptido de insulina" también abarca homodímeros, heterodímeros, homomúltimeros, y heteromúltimeros que están enlazados, incluyendo pero no limitado a los enlazados directamente a través de cadenas laterales de aminoácidos no naturalmente codificados, ya sea para las mismas o diferentes cadenas laterales de aminoácidos no naturalmente codificados, a las cadenas laterales de aminoácidos naturalmente codificados, o indirectamente a través de un enlazador. Ejemplos de adaptadores que incluyen pero no se limitan a compuestos orgánicos pequeños, polímeros solubles en agua de una variedad de longitudes, tales como poli(etilenglicol) o polidextrano, o polipéptidos de varias longitudes.

Un "aminoácido no naturalmente codificado" se refiere a un aminoácido que no es uno de los 20 aminoácidos comunes o pirrolisina o selenocisteína. Otros términos que pueden ser utilizados como sinónimos con el término "aminoácido no naturalmente codificado" son "aminoácido no natural", "aminoácido no natural", "aminoácido no natural", y versiones de diversas maneras con guiones y sin guiones de los mismos. El término "aminoácido no naturalmente codificado" también incluye, pero no se limita a los ácidos, aminoácidos que se producen por la modificación (por ejemplo modificaciones postraduccionales) de un aminoácido codificado de forma natural (incluyendo pero no limitado a, los 20 aminoácidos comunes o pirrolisina y selenocisteína), pero no a sí mismos se incorporan naturalmente en una cadena polipeptídica en crecimiento por el complejo de la traducción. Ejemplos de tales aminoácidos no naturales incluyen, pero no se limitan a, N-acetilglucosaminilo-L-serina, N-acetilglucosaminilo-L-treonina, y O-fosfotirosina.

Un "grupo de modificación terminal amino" se refiere a cualquier molécula que puede unirse al extremo amino terminal de un polipéptido. De manera similar, un "grupo de modificación terminal carboxi" se refiere a cualquier molécula que se puede conectar a la terminal carboxi de un polipéptido. Grupos de modificación terminal incluyen, pero no se limitan a, diversos polímeros solubles en agua, péptidos o proteínas tales como albúmina de suero, u otros restos que aumentan la vida media en suero de los péptidos.

Los términos "grupo funcional", "resto activo", "grupo de activación", "grupo saliente", "sitio reactivo", "grupo químicamente reactivo" y "resto químicamente reactivo" son utilizados en la técnica y aquí se refieren a porciones o unidades definibles de una molécula distintas. Los términos son un tanto sinónimos en las técnicas químicas y se utilizan en el presente documento para indicar las porciones de moléculas que realizan alguna función o actividad y son reactivas con otras moléculas.

El término "enlace" o "enlazador" se utiliza aquí para referirse a grupos o enlaces que se forman normalmente como resultado de una reacción química y típicamente son enlaces covalentes. Enlaces hidrolíticamente estables significan que los enlaces son sustancialmente estables en agua y no reaccionan con agua a valores de pH útiles, incluyendo pero no limitado a, en condiciones fisiológicas durante un período prolongado de tiempo, quizás incluso indefinidamente. Enlaces hidrolíticamente inestables o degradables significan que los enlaces son degradables en agua o en soluciones acuosas, incluyendo por ejemplo, sangre. Enlaces enzimáticamente inestables o degradables significan que el enlace puede ser degradado por una o más enzimas. Tal como se entiende en la técnica, el PEG y polímeros relacionados pueden incluir enlaces degradables en la cadena principal del polímero o en el grupo enlazador entre la cadena principal del polímero y uno o más de los grupos funcionales terminales de la molécula de polímero. Por ejemplo, los enlaces de éster formados por la reacción de ácidos carboxílicos PEG o ácidos carboxílicos activados por PEG con grupos de alcohol en un agente biológicamente activo generalmente se hidrolizan en condiciones fisiológicas para liberar el agente. Otros enlaces hidrolíticamente degradables incluyen, pero no se limitan a, enlaces carbonato; enlaces imina como resultado de la reacción de una amina y un aldehído; enlaces éster fosfato formados mediante la reacción de un alcohol con un grupo fosfato; enlaces de hidrazona que son 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, incluyendo pero no limitado a un extremo de un polímero tal como PEG, y un grupo carboxilo de un péptido; y enlaces oligonucleotídicos formados por un grupo de fosforamida, incluyendo pero no limitados al extremo de un polímero, y un grupo de hidroxilo 5' de un oligonucleótido.

El término "molécula biológicamente activa", "resto biológicamente activo" o "agente biológicamente activo" cuando se usa en este documento significa cualquier sustancia que puede afectar a cualquier propiedad física o bioquímica de un sistema biológico, vía, molécula, o la interacción relativa a un organismo, incluyendo pero no limitado a, virus,

bacterias, bacteriófagos, transposones, prión, insectos, hongos, plantas, animales, y seres humanos. En particular, tal como se usa en el presente documento, las moléculas biológicamente activas incluyen, pero no se limitan a, cualquier sustancia destinada al diagnóstico, cura, mitigación, tratamiento o prevención de enfermedades en seres humanos u otros animales, o para mejorar de otra manera el bienestar físico o mental de los seres humanos o animales. Ejemplos de moléculas biológicamente activas incluyen, pero no se limitan a, péptidos, proteínas, enzimas, fármacos de moléculas pequeñas, las vacunas, los inmunógenos, drogas duras, drogas blandas, hidratos de carbono, átomos inorgánicos o moléculas, colorantes, lípidos, nucleósidos, radionúclidos, oligonucleótidos, toxoides, toxinas, células procarióticas y eucarióticas, virus, polisacáridos, ácidos nucleicos y partes de los mismos obtenidas o derivadas de virus, bacterias, insectos, animales o cualquier otro tipo de célula o células, liposomas, micropartículas y micelas. Los polipéptidos de insulina pueden ser añadidos en una formulación micelar; véase la patente de los Estados Unidos. nº 5.833.948. Las clases de agentes biológicamente activos que son adecuados para uso con la invención incluyen, pero no están limitados a, fármacos, profármacos, radionucleidos, agentes de formación de imágenes, polímeros, antibióticos, fungicidas, agentes antivirales, agentes antiinflamatorios, agentes antitumorales, agentes cardiovasculares, agentes anti ansiedad, hormonas, factores de crecimiento, agentes esteroideos, toxinas microbianamente derivadas, y similares.

Un "polímero bifuncional" se refiere a un polímero que comprende dos grupos funcionales discretos que son capaces de reaccionar específicamente con otros restos (incluidos pero no limitados a, grupos de amino laterales de ácido) para formar enlaces covalentes o no covalentes. Un enlazador bifuncional que tiene un grupo funcional reactivo con un grupo en un componente biológicamente activo particular, y otro grupo reactivo con un grupo en un segundo componente biológico, se puede usar para formar un conjugado que incluye el primer componente biológicamente activo, el enlazador bifuncional y el segundo componente biológicamente activo. Se conocen muchos procedimientos y moléculas enlazadoras para la unión de diversos compuestos a péptidos. Véase, por ejemplo, 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 multifuncional" se refiere a un polímero que comprende dos o más grupos funcionales discretos que son capaces de reaccionar específicamente con otros restos (incluidos, pero no limitados a, amino grupos laterales de ácido) para formar enlaces covalentes o no covalentes. Un polímero bifuncional o polímero multifuncional pueden ser de cualquier longitud deseada o el peso molecular, y puede ser seleccionado para proporcionar una separación deseada particular o de conformación entre una o más moléculas unidas a la insulina y su receptor o insulina.

Cuando los grupos sustituyentes se especifican mediante sus fórmulas químicas convencionales, escritas de izquierda a derecha, que igualmente abarcan los sustituyentes químicamente idénticos que resultarían de escribir la estructura de derecha a izquierda, por ejemplo, la estructura CH<sub>2</sub>O es equivalente a la estructura OCH<sub>2</sub>.

El término "sustituyentes" incluye pero no se limita a "sustituyentes no interferentes". "Sustituyentes no interferentes" son aquellos grupos que dan compuestos estables. Sustituyentes que no interfieren adecuados o radicales incluyen, pero no se limitan a, halo, C1-C10 alquilo, C2-C10 alqueno, C2-C10 alquino, C1-C10 alcoxi, C1-C12 aralquilo, C1-C12 alcarilo, C3-C12 cicloalquilo, C3-C12 cicloalqueno, fenilo, fenilo sustituido, toluoilo, xileno, bifenilo, C2-C12 alcoxialquilo, C2-C12 alcoxiarilo, C7-C12 ariloxialquilo, C7-C12 oxiarilo, C1-C6 alquilsulfino, C1-C10 alquilsulfonilo, --(CH<sub>2</sub>)<sub>m</sub> --O--(C1-C10 alquilo) en la que m es de 1 a 8, arilo, arilo sustituido, alcoxi sustituido, fluoroalquilo, radical heterocíclico, sustituido radical heterocíclico, nitroalquilo, --NO<sub>2</sub>, --CN, --NRC(O)--(C1-C10 alquilo), --C(O)(C1-C10 alquilo), C2-C10 alquilo tioalquilo, -C(O)O--(C1-C10 alquilo), --OH, --SO<sub>2</sub>, =S, --COOH, --NR<sub>2</sub>, carbonilo, --C(O)(C1-C10 alquilo) CF<sub>3</sub>, --C(O)CF<sub>3</sub>, --C(O)NR<sub>2</sub>, --(C1-C10 arilo)-S--(C6-C10 arilo), --C(O)--(C1-C10 arilo), --(CH<sub>2</sub>)<sub>m</sub> --O--(-CH<sub>2</sub>)<sub>m</sub> --O--(C1-C10 alquilo) en donde cada m es de 1 a 8, --C(O)NR<sub>2</sub>, --C(S)NR<sub>2</sub>, --SO<sub>2</sub>NR<sub>2</sub>, --NRC(O) NR<sub>2</sub>, --NRC(S) NR<sub>2</sub>, sales de los mismos, y similares. Cada R, tal como se usa en el presente documento, es H, alquilo o alquilo sustituido, arilo o arilo sustituido, aralquilo o alcarilo.

El término "halógeno" incluye flúor, cloro, yodo y bromo.

El término "alquilo", por sí mismo o como parte de otro sustituyente, significa, a menos se indique lo contrario, una cadena lineal o ramificada, o radical de hidrocarburo cíclico, o combinación de los mismos, que puede estar completamente saturado, mono o poliinsaturado y puede incluir radicales divalentes y multivalentes, que tiene el número de átomos de carbono designados (es decir C1-C10 significa de 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, (ciclohexilo)metilo, ciclopropilo-metilo, 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. Los ejemplos de grupos de alquilo insaturados 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 sus homólogos superiores y isómeros. El término "alquilo", a menos que se indique lo contrario, también pretende incluir aquellos derivados de alquilo definidos con más detalle a continuación, como "heteroalquilo". Los grupos de alquilo que se limitan a grupos hidrocarburo se denominan "homoalquilo".

El término "alquilenos" por sí mismo o como parte de otro sustituyente significa un radical divalente derivado de un alcano, como se ejemplifica, pero no se limita, por las estructuras -CH<sub>2</sub>CH<sub>2</sub>- y -CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>-, e incluye además aquellos grupos descritos más adelante como "heteroalquilenos". Típicamente, un grupo alquilo (o alquilenos) tendrá

de 1 a 24 átomos de carbono, con aquellos grupos que tienen 10 o menos átomos de carbono que son una realización particular de los métodos y composiciones descritos en el presente documento. Un "alquilo inferior" o "alquilen inferior" es un grupo alquilo de cadena más corta o alquilen, que tiene generalmente ocho o menos átomos de carbono.

5 Los términos "alcoxi", "alquilamino" y "alquiltio" (o tioalcoxi) se usan en su sentido convencional, y se refieren a aquellos grupos de alquilo unidos al resto de la molécula mediante un átomo de oxígeno, un grupo amino, o un átomo de azufre, respectivamente.

10 El término "heteroalquilo", por sí mismo o en combinación con otro término, significa, a menos que se indique otra cosa, una cadena lineal o ramificada estable, o radical de hidrocarburo cíclico, o combinaciones de los mismos, que consiste en el número indicado de átomos de carbono y al menos una heteroátomo seleccionado del grupo que  
 15 consiste en O, N, Si y S, y en donde los átomos de nitrógeno y azufre pueden estar opcionalmente oxidados y el heteroátomo de nitrógeno puede estar opcionalmente cuaternizado. El heteroátomo O, N y S y Si puede 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. Los ejemplos incluyen, pero no se limitan a, -CH<sub>2</sub>-CH<sub>2</sub>-O-CH<sub>3</sub>, -CH<sub>2</sub>-CH<sub>2</sub>-NH-CH<sub>3</sub>, -CH<sub>2</sub>-CH<sub>2</sub>-N(CH<sub>3</sub>)-CH<sub>3</sub>, -CH<sub>2</sub>-S-CH<sub>2</sub>CH<sub>3</sub>, CH<sub>2</sub>-CH<sub>2</sub>-S(O)CH<sub>3</sub>, -CH<sub>2</sub>-CH<sub>2</sub>-S(O)<sub>2</sub>-CH<sub>3</sub>, -CH=CH-O-CH<sub>3</sub>, -Si(CH<sub>3</sub>)<sub>3</sub>, -CH<sub>2</sub>-CH=N-OCH<sub>3</sub>, y CH=CH-N(CH<sub>3</sub>)-CH<sub>3</sub>. Hasta dos heteroátomos pueden ser consecutivos, tales como, por ejemplo, -CH<sub>2</sub>-NH-OCH<sub>3</sub> y -CH<sub>2</sub>-O-Si(CH<sub>3</sub>)<sub>3</sub>. Del mismo modo, el término "heteroalquilen" por sí mismo o como parte de  
 20 otro sustituyente significa un radical divalente derivado de heteroalquilo, como se ejemplifica, pero no limitado por, -CH<sub>2</sub>-CH<sub>2</sub>-S-CH<sub>2</sub> CH<sub>2</sub> y -CH<sub>2</sub>-S-CH<sub>2</sub>-CH<sub>2</sub>-NH-CH<sub>2</sub>-. Para los grupos heteroalquilen, los heteroátomos iguales o diferentes también pueden ocupar cualquiera o ambos extremos de la cadena (incluyendo pero no limitado a, alquilenoxi, alquilendioxo, alquilenamino, alquilendiamino, aminoalquiloeno, y similares). Aún más, para los grupos alquilen y heteroalquilen de unión, ninguna orientación del grupo de enlace está implicada por la dirección en la que está escrita la fórmula del grupo de unión. Por ejemplo, la fórmula -C(O)2R' representa tanto -C(O)2R' como -R'C(O)2.

Los términos "cicloalquilo" y "heterocicloalquilo", por sí mismos o en combinación con otros términos, representan, a menos que se indique lo contrario, versiones cíclicas de "alquilo" y "heteroalquilo", respectivamente. Por lo tanto, un cicloalquilo o heterocicloalquilo incluyen, uniones de anillos parcialmente insaturados y completamente insaturados saturados. Adicionalmente, para heterocicloalquilo, un heteroátomo puede ocupar la posición en la que el heterociclo está unido al resto de la molécula. Los ejemplos de cicloalquilo incluyen, pero no se limitan a, ciclopentilo, ciclohexilo, 1-ciclohexenilo, 3-ciclohexenilo, cicloheptilo, y similares. Los ejemplos de heterocicloalquilo incluyen, pero no se limitan a, 1-(1,2,5,6-tetrahidropiridilo), 1-piperidinilo, 2-piperidinilo, 3-piperidinilo, 4-morfolinilo, 3-morfolinilo, tetrahidrofurano-2-ilo, tetrahidrofurano-3-ilo, tetrahidrotieno-2-ilo, tetrahidrotieno-3-ilo, 1-piperazinilo, 2-piperazinilo, y similares. Además, el término abarca estructuras bicíclicas y de anillo tricíclico. Del mismo modo, el término "heterocicloalquilen" por sí mismo o como parte de otro sustituyente significa un radical divalente derivado de heterocicloalquilo, y el término "cicloalquilen" por sí mismo o como parte de otro sustituyente significa un radical divalente derivado de cicloalquilo.

40 Tal como se utiliza aquí, el término "polímero soluble en agua" se refiere a cualquier polímero que es soluble en disolventes acuosos. La unión de polímeros solubles en agua a polipéptidos de insulina puede resultar en cambios incluyendo, pero no limitado al aumento o la vida media en suero modulado, o vida media terapéutica aumentada o modulada con respecto a la forma no modificada, la inmunogenicidad modulada, características moduladas de asociación física, tales como la agregación y la formación de multímero, de unión al receptor alterado, unión alterada a una o más parejas de unión, y una alteración de la dimerización del receptor o multimerización. El polímero soluble en agua puede o no puede tener su propia actividad biológica, y puede ser utilizado como un enlazador para unir la insulina a otras sustancias, incluyendo, pero no limitado a uno o más polipéptidos de insulina, o una o más moléculas biológicamente activas. Los polímeros adecuados incluyen, pero no se limitan a, polietilenglicol, propionaldehído de polietilenglicol, mono C1-C10 alcoxi o derivados ariloxi del mismo (descritos en la Patente de Estados Unidos n° 5.252.714), monometoxipolietilenglicol, polivinilpirrolidona, alcohol polivinílico, poliaminoácidos, diviniléter anhídrido maleico, N-(2-Hidroxipropilo) metacrilamida, dextrano, derivados de dextrano, incluyendo sulfato de dextrano, glicol de polipropileno, óxido de polipropileno/copolímero de óxido de etileno, polioli polioxiethylado, heparina, fragmentos de heparina, polisacáridos, oligosacáridos, glucanos, celulosa y derivados de celulosa, incluyendo pero no limitado a metilcelulosa y derivados de la celulosa de carboximetilo, almidón y derivados de almidón, polipéptidos, polialquilen glicol y derivados de los mismos, copolímeros de polialquilen glicoles y derivados de los mismos, éteres de polivinilo etilo, y alfa-beta-polilo (2-hidroxietilo) DL-aspartamida, y similares, o mezclas de los mismos. Ejemplos de polímeros solubles en agua incluyen, pero no se limitan a, polietilenglicol y albúmina de suero.

60 Tal como se utiliza aquí, el término "polialquilen glicol" o "poli(alqueno)" se refiere a polietilenglicol (poli(etilenglicol)), glicol de polipropileno, glicol de butileno, y sus derivados. El término "polialquilen glicol" abarca tanto los polímeros lineales y ramificados y los pesos moleculares medios de entre 0,1 kDa y 100 kDa. Otros ejemplos de realización se enumeran, por ejemplo, en catálogos de proveedores comerciales, tales como catálogo de Shearwater Corporation "Polyethylene Glycol and Derivatives for Biomedical Applications" de (2001).

El término significa "arilo", a menos que se indique lo contrario, un sustituyente poliinsaturado, aromático, hidrocarburo que puede ser un solo anillo o múltiples anillos (incluyendo pero no limitado a, de 1 a 3 anillos) que están fusionados juntos o unidos covalentemente. El término "heteroarilo" se refiere a grupos arilo (o anillos) que contienen de uno a cuatro heteroátomos seleccionados de N, O y S, donde los átomos de nitrógeno y azufre están opcionalmente oxidados, y el átomo de nitrógeno está opcionalmente cuaternizado. Un grupo heteroarilo puede estar unido al resto de la molécula a través de un heteroátomo. Ejemplos no limitantes de grupos de arilo y heteroarilo incluyen fenilo, 1-naftilo, 2-naftilo, 4-bifenilo, 1-pirrolilo, 2pirrolilo, 3-pirrolilo, 3-pirazolilo, 2-imidazolilo, 4-imidazolilo, pirazinilo, 2-oxazolilo, 4-oxazolilo, 2-fenilo-4-oxazolilo, 5-oxazolilo, 3-isoxazolilo, 4-isoxazolilo, 5-isoxazolilo, 2-tiazolilo, 4-tiazolilo, 5-tiazolilo, 2-furilo, 3-furilo, 2-tienoilo, 3-tienilo, 2-piridilo, 3-piridilo, 4-piridilo, 2-pirimidilo, 4-pirimidilo, 5-benzotiazolilo, purinilo, 2-benzimidazolilo, 5-indolilo, 1-isoquinolilo, 5-isoquinolilo, 2-quinoxalinilo, 5-quinoxalinilo, 3-quinolilo, y 6-quinolilo. Los sustituyentes para cada uno de los sistemas de anillos arilo y heteroarilo indicados anteriormente se seleccionan del grupo de sustituyentes aceptables descrito a continuación.

Por razones de brevedad, el término "arilo" cuando se usa en combinación con otros términos (incluyendo pero no limitado a, ariloxi, ariltioxi, arilalquilo) incluye tanto arilo y anillos de heteroarilo como se define anteriormente. Así, el término "arilalquilo" se entiende que incluye aquellos radicales en los que un grupo arilo está unido a un grupo alquilo (incluyendo pero no limitado a, bencilo, fenetilo, piridilmetilo y similares) incluyendo aquellos grupos de alquilo en los que un átomo de carbono (incluyendo pero no limitado a, un grupo metileno) ha sido reemplazado por, por ejemplo, un átomo de oxígeno (incluyendo pero no limitado a, fenoximetilo, 2-piridilooximetilo, 3-(1-naftiloxi)propilo, y similares).

Cada uno de los términos anteriores (incluyendo pero no limitado a, "alquilo", "heteroalquilo", "arilo" y "heteroarilo") se entiende que incluyen ambas formas del radical indicado sustituido y no sustituido. Los ejemplos de sustituyentes para cada tipo de radical se proporcionan a continuación.

Los sustituyentes para los radicales alquilo y heteroalquilo (incluyendo aquellos grupos que se denominan a menudo alquilenilo, alqueniilo, heteroalquilenilo, heteroalqueniilo, alquinilo, cicloalquilo, heterocicloalquilo, cicloalqueniilo, y heterocicloalqueniilo) pueden ser uno o más de una variedad de grupos seleccionados de, pero no limitado a:  $-OR'$ ,  $=O$ ,  $=NR'$ ,  $=NOR'$ ,  $-NR'R''$ ,  $-SR'$ ,  $-halógeno$ ,  $-SiR'R''R'''$ ,  $-O-C(O)R'$ ,  $-C(O)R'$ ,  $-CO_2R'$ ,  $-CONR'R''$ ,  $-OC(O)NR'R''$ ,  $-NR''C(O)R'$ ,  $NR''C(O)NR''R'''$ ,  $-NR''C(O)_2R'$ ,  $-NR-C(NR'R''R''')=NR''''$ ,  $NR C(NR'R'')=NR''''$ ,  $-S(O)R'$ ,  $-S(O)_2R'$ ,  $-S(O)_2NR'R''$ ,  $NRSO_2R'$ ,  $-CN$  y  $-NO_2$  en un número que va de cero a  $(2m'+1)$ , donde  $m'$  es el número total de átomos de carbono en tal radical.  $R'$ ,  $R''$ ,  $R'''$  y  $R''''$  se refieren cada uno independientemente a hidrógeno, heteroalquilo sustituido o no sustituido, arilo sustituido o no sustituido, incluyendo pero no limitado a, arilo sustituido con 1-3 halógenos, alquilo sustituido o no sustituido, alcoxi o tioalcoxi, o grupos arilalquilo. Cuando un compuesto de la invención incluye más de un grupo R, por ejemplo, cada uno de los grupos R se selecciona independientemente como son cada uno grupos  $R'$ ,  $R''$ ,  $R'''$  y  $R''''$  cuando más de uno de estos grupos está presente. Cuando  $R'$  y  $R''$  están unidos al mismo átomo de nitrógeno, se pueden combinar con el átomo de nitrógeno para formar un anillo de 5, 6, o 7 miembros. Por ejemplo,  $NR'R''$  está destinado a incluir, pero no limitarse a, 1-pirrolidinilo y 4-morfolinilo. De la discusión anterior de sustituyentes, un experto en la técnica entenderá que el término 'alquilo' se entiende que incluye grupos que incluyen átomos de carbono unidos a grupos distintos de grupos hidrógeno, tales como haloalquilo (incluyendo pero no limitado a,  $-CF_3$  y  $-CH_2CF_3$ ) y acilo (incluyendo, pero no limitado a,  $-C(O)CH_3$ ,  $-C(O)CF_3$ ,  $-C(O)CH_2-O-CH_3$ , y similares).

De manera similar a los sustituyentes descritos para el radical alquilo, los sustituyentes para los grupos arilo y heteroarilo son variados y se seleccionan de, pero no se limitan a: halógeno,  $OR'$ ,  $=O$ ,  $=NR'$ ,  $=NOR'$ ,  $-NR'R''$ ,  $-SR'$  - halógeno,  $-SiR'R''R'''$ ,  $OC(O)R'$ ,  $-C(O)R'$ ,  $CO_2R'$ ,  $-CONR'R''$ ,  $-OC(O)NR'R''$ ,  $-NR''C(O)R'$ ,  $NR''C(O)NR''R'''$ ,  $-NR''C(O)_2R'$ ,  $NR-C(NR'R''R''')=NR''''$ ,  $NR C(=NR'R'')=NR''''$ ,  $-S(O)R'$ ,  $-S(O)_2R'$ ,  $-S(O)_2NR'R''$ ,  $NRSO_2R'$ ,  $-CN$  y  $-NO_2$ ,  $-R'$ ,  $-N_3$ ,  $-CH(Ph)_2$ , fluoro(C1-C4)alcoxi, y fluoro(C1-C4)alquilo, en un número que oscila de cero al número total de valencias abiertas sobre el sistema de anillo aromático; y donde  $R'$ ,  $R''$ ,  $R'''$  y  $R''''$  se seleccionan independientemente de hidrógeno, alquilo, heteroalquilo, arilo y heteroarilo. Cuando un compuesto de la invención incluye más de un grupo R, por ejemplo, se selecciona cada uno de los grupos R independientemente como son cada uno de los grupos  $R'$ ,  $R''$ ,  $R'''$  y  $R''''$  cuando más de uno de estos grupos está presente.

Tal como se utiliza aquí, el término "vida media modulada en suero" significa el cambio positivo o negativo en la circulación de vida media de una insulina relativa modificada a su forma no modificada. La vida media en suero se mide tomando muestras de sangre en diversos puntos temporales después de la administración de la insulina, 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 vida media en suero. Vida media aumenta en suero tiene deseablemente al menos aproximadamente dos veces, pero un aumento más pequeño puede ser útil, por ejemplo donde se permite un régimen de dosificación satisfactoria o evita un efecto tóxico. En algunas realizaciones, el aumento es al menos aproximadamente tres veces, al menos aproximadamente cinco veces, o al menos aproximadamente diez veces.

El término "vida media terapéutico modulada" Tal como se usa en este documento significa el cambio positivo o negativo en la vida media de la cantidad terapéuticamente eficaz de la insulina, con respecto a su forma no modificada. La vida media terapéutica se mide midiendo propiedades farmacocinéticas y/o farmacodinámicas de la

molécula en diversos puntos temporales después de la administración. El aumento de la vida media terapéutica permite de manera deseable un régimen de dosificación beneficioso particular, una dosis total beneficiosa particular, o evita un efecto no deseado. En algunas realizaciones, el aumento de la vida media terapéutica resultante de mayor potencia, aumento o disminución de la unión de la molécula modificada a su diana, aumento o disminución de ruptura de la molécula mediante enzimas tales como proteasas, o un aumento o disminución en otro parámetro o mecanismo de acción de la molécula no modificada o un aumento o disminución en el aclaramiento mediado por receptor de la molécula.

El término "aislado", cuando se aplica a un ácido nucleico o proteína, denota que el ácido nucleico o proteína es libre de al menos algunos de los componentes celulares con los que está asociado en el estado natural, o que el ácido nucleico o proteína se ha concentrado a un nivel mayor que la concentración de su producción *in vivo* o *in vitro*. Puede estar en un estado homogéneo. Sustancias aisladas pueden estar en un estado seco o semiseco, o en solución, incluyendo, pero no limitado a, una solución acuosa. Puede ser un componente de una composición farmacéutica que comprende vehículos y/o excipientes farmacéuticamente aceptables adicionales. La pureza y la homogeneidad se determinan típicamente utilizando técnicas de química analítica tales como electroforesis en gel de poliacrilamida o cromatografía líquida de alto rendimiento. Una proteína que es la especie predominante presente en una preparación se purifica sustancialmente. En particular, un gen aislado se separa de los marcos de lectura abiertos que flanquean el gen y codifican una proteína distinta del gen de interés. El término "purificado" se refiere a que un ácido nucleico, o proteína da lugar a esencialmente una banda en un gel electroforético. En particular, esto puede significar que el ácido nucleico o proteína es al menos 85% puro, al menos 90% puro, al menos 95% puro, al menos el 99% o más puro.

El término "ácido nucleico" se refiere a deoxiribonucleótidos, deoxiribonucleótidos, ribonucleótidos, o ribonucleótidos y polímeros de los mismos, ya sea en forma individual o de doble hebra. A menos que se limite específicamente, el término abarca ácidos nucleicos que contienen análogos conocidos de nucleótidos naturales que tienen propiedades de unión similares al ácido nucleico de referencia y se metabolizan de una manera similar a nucleótidos de origen natural. A menos que se limite específicamente de otra manera, el término también se refiere a análogos de oligonucleótido incluyendo APN (ácido peptidonucleico), análogos de ADN utilizados en la tecnología antisentido (fosfortioatos, fosforamidatos, y similares). A menos que se indique lo contrario, una secuencia de ácido nucleico particular también abarca implícitamente variantes modificadas de manera conservadora del mismo (incluyendo pero no limitado a, sustituciones de codones degenerados) y secuencias complementarias así como la secuencia indicada explícitamente. Específicamente, las sustituciones de codon degeneradas se pueden conseguir mediante la generación de secuencias en las que la tercera posición de uno o más codones seleccionados (o todos) está sustituida con residuos de base mezclada y/o desoxiinosina (Batzer et al., *Nucleic Acid Res.* 19: 5081 (1991); Ohtsuka et al., *J. Biol Chem* 260: 2605-2608 (1985); Rossolini et al., *Mol. Cell. Probes* 8: 91-98 (1994)).

Los términos "polipéptido", "péptido" y "proteína" se utilizan indistintamente en este documento para referirse a un polímero de residuos de aminoácidos. Es decir, una descripción dirigida a un polipéptido que se aplica igualmente a la descripción de un péptido y una descripción de una proteína, y viceversa. Los términos se aplican a polímeros de origen natural de aminoácidos, así como polímeros de aminoácidos en los que uno o más residuos de aminoácidos es un aminoácido no naturalmente codificado. Tal como se usa en este documento, los términos abarcan cadenas de aminoácidos de cualquier longitud, incluyendo proteínas de longitud completa, en las que los residuos de aminoácidos están unidos por enlaces peptídicos covalentes.

El término "aminoácido" se refiere a aminoácidos naturales y no naturales, así como análogos de aminoácidos y miméticos de aminoácidos que funcionan de una manera similar a los aminoácidos de origen natural. Aminoácidos codificados de forma natural son los 20 aminoácidos comunes (alanina, arginina, asparagina, ácido aspártico, cisteína, glutamina, ácido glutámico, glicina, histidina, isoleucina, leucina, lisina, metionina, fenilo-alanina, prolina, serina, treonina, triptófano, tirosina y valina) y pirrolisina y selenocisteína. Análogos de aminoácidos se refieren a compuestos que tienen la misma estructura química básica que un aminoácido de origen natural, es decir, un carbono que está unido a un hidrógeno, un grupo carboxilo, un grupo amino, y un grupo R, tal como, homoserina, norleucina, sulfóxido de metionina, sulfonio de metilo de metionina. Tales análogos tienen grupos R modificados (por ejemplo, norleucina) o cadenas principales peptídicas modificadas, pero conservan la misma estructura química básica que un aminoácido de origen natural. La referencia a un aminoácido incluye, por ejemplo, L-aminoácidos proteogénicos de origen natural; D-aminoácidos, aminoácidos químicamente modificados tales como variantes y derivados de aminoácidos; aminoácidos no proteogénicos de origen natural tales como  $\beta$ -alanina, ornitina, etc.; y compuestos sintetizados químicamente que tienen propiedades conocidas en la técnica por ser características de los aminoácidos. Ejemplos de aminoácidos no naturales incluyen, pero no se limitan a, ametilo aminoácidos (por ejemplo,  $\alpha$ -alanina-metilo), D-aminoácidos, aminoácidos de tipo histidina (por ejemplo, 2-amino-histidina,  $\beta$ -hidroxihistidina, homohistidina,  $\alpha$ -fluorometilo-histidina y  $\alpha$ -metilo-histidina), aminoácidos que tienen un metiloeno adicional en la cadena lateral (aminoácidos "homo"), y aminoácidos en los que un grupo funcional de ácido carboxílico en la cadena lateral se sustituye con un grupo de ácido sulfónico (por ejemplo, ácido cisteico). La incorporación de ácidos no naturales aminoácidos, incluyendo aminoácidos no nativos sintéticos, aminoácidos sustituidos, o uno o más D-aminoácidos en las proteínas de la presente invención puede ser ventajosa en un número de diferentes maneras. Péptidos que contienen D-aminoácidos, etc., presentan una mayor estabilidad *in vitro* o *in vivo* en comparación con homólogos que contienen L-aminoácido. Por lo tanto, la construcción de péptidos, etc., incorporando D-aminoácidos

puede ser particularmente útil cuando se desea o se requiere una mayor estabilidad intracelular. Más específicamente, D-péptidos, etc., son resistentes a las peptidasas endógenas y proteasas, proporcionando así una mejor biodisponibilidad de la molécula, y tiempos de vida prolongados *in vivo* cuando tales propiedades son deseables. Además, D-péptidos, etc., no pueden ser procesados de manera eficiente para complejo de mayor histocompatibilidad de clase II-presentación restringida a células auxiliares T, y son por lo tanto, menos probables de inducir respuestas inmunes humorales en todo el organismo.

Los aminoácidos pueden ser conocidos en la presente memoria por cualquiera de sus símbolos de tres letras comúnmente conocidos o por los símbolos de una letra recomendados por la Comisión de Nomenclatura Bioquímica IUPAC-IUB. Igualmente, los nucleótidos se pueden conocer por sus códigos de una letra comúnmente aceptados.

"Variantes modificadas de manera conservadora" se aplican tanto a secuencias de aminoácidos como a ácidos nucleicos. Con respecto a secuencias particulares de ácidos nucleicos, "variantes modificadas de forma conservadora" se refieren a aquellos ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o cuando el ácido nucleico no codifica una secuencia de aminoácidos, a secuencias esencialmente idénticas. 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 los aminoácidos de alanina. Por lo tanto, en cada posición donde una alanina es especificada por un codón, el codón puede alterarse a cualquiera de los codones correspondientes descritos sin alterar el polipéptido codificado. Tales variaciones de ácido nucleico son "variaciones silenciosas", que son una especie de variaciones modificadas de manera conservadora. Cada secuencia de ácido nucleico en el presente documento que codifica un polipéptido también describe cada posible variación silenciosa del ácido nucleico. Un experto en la técnica reconocerá que cada codón en un ácido nucleico (excepto AUG, que es ordinariamente el único codón para metionina, y TGG, que es ordinariamente el único codón para triptófano) puede modificarse para producir una molécula funcionalmente idéntica. Por consiguiente, cada variación silenciosa de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

En relación a secuencias de aminoácidos, alguien de experiencia ordinaria en la técnica reconocerá que las sustituciones individuales, deleciones o adiciones a un ácido nucleico, péptido, polipéptido o secuencia de proteína que altera, añade o elimina un único aminoácido o un pequeño porcentaje de aminoácidos en la secuencia codificada es una "variante modificada de forma conservadora", donde la alteración da como resultado la deleción de un aminoácido, la adición de un aminoácido, o la sustitución de un aminoácido con un aminoácido químicamente similar. Las tablas de sustitución conservadora que proporcionan aminoácidos funcionalmente similares son conocidas por personas de experiencia ordinaria en la técnica. Tales variantes modificadas de manera conservadora se suman a y no excluyen variantes polimórficas, homólogos entre especies, y alelos de la invención.

Las tablas de sustitución conservadoras que proporcionan aminoácidos funcionalmente similares son conocidas por personas de experiencia ordinaria en la técnica. Los ocho grupos siguientes contienen cada uno aminoácidos que son sustituciones conservadoras entre sí:

- 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);

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

Los términos "identidad", "idéntico" o porcentaje en el contexto de dos o más ácidos nucleicos o secuencias polipeptídicas, se refieren a dos o más secuencias o subsecuencias que son iguales. Las secuencias son "sustancialmente idénticas" si tienen un porcentaje de residuos de aminoácidos o nucleótidos que son iguales (es decir, aproximadamente el 60% de identidad, aproximadamente el 65%, aproximadamente el 70%, aproximadamente el 75%, aproximadamente el 80%, aproximadamente el 85%, aproximadamente el 90%, o aproximadamente el 95% de identidad sobre una región especificada), cuando se comparan y alinean para correspondencia máxima sobre una ventana de comparación, o región designada medida utilizando uno de los siguientes algoritmos de comparación de secuencias (u otros algoritmos disponibles para las personas de experiencia ordinaria en la técnica) o mediante alineamiento manual e inspección visual. Esta definición también se refiere al complemento de una secuencia de prueba. La identidad puede existir más de una región que es al menos aproximadamente 50 aminoácidos o nucleótidos de longitud, o más de una región que es 75-100 aminoácidos o nucleótidos de longitud, o, si no se especifica, a través de toda la secuencia de un polinucleótido o polipéptido. Un polinucleótido que codifica un polipéptido de la presente invención, incluyendo homólogos de especies distintas a la



humana, puede obtenerse mediante un procedimiento que comprende las etapas de selección de una biblioteca en condiciones de hibridación rigurosas con una sonda marcada que tiene una secuencia de polinucleótidos de la divulgación o un fragmento del mismo y aislar clones de ADNc de longitud completa y genómica que contienen dicha secuencia de polinucleótidos. Tales técnicas de hibridación son bien conocidas para el experto en la materia.

Para la comparación de secuencias, típicamente una secuencia actúa como secuencia de referencia, a la que las secuencias de prueba se comparan. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de ensayo y de referencia se introducen en un ordenador, las coordenadas de subsecuencia se designan, si es necesario, y se designan los parámetros del programa de algoritmo de secuencia. Los parámetros del programa por defecto se pueden utilizar, o parámetros alternativos pueden ser designados. El algoritmo de comparación de secuencias calcula entonces la identidad de secuencia en porcentaje para las secuencias de ensayo con relación a la secuencia de referencia, basándose en los parámetros del programa.

Una "ventana de comparación", tal como se usa en el presente documento, incluye la referencia a un segmento de una cualquiera del número de posiciones contiguas seleccionadas del grupo que consiste en de 20 a 600, por lo general aproximadamente 50 a aproximadamente 200, más habitualmente de aproximadamente 100 a aproximadamente 150 en que una secuencia puede compararse a una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias se alinean de manera óptima. Los métodos de alineación de secuencias para comparación son conocidos por personas de experiencia ordinaria en la técnica. La alineación óptima de secuencias para la comparación puede llevarse a cabo, incluyendo pero no limitado a, por el algoritmo de homología local de Smith y Waterman (1970) *Adv. Appl. Math.* 2:482c, mediante el algoritmo de alineación de homología de Needleman y Wunsch (1970) *J. Mol. Biol.* 48: 443, por la búsqueda del método de similitud de Pearson y Lipman (1988) *Proc. Nacional. Acad. Sci. EE.UU.* 85: 2444, mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el Paquete de Software de Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante alineamiento manual e inspección visual (véase, por ejemplo, Ausubel *et al.*, *Current Protocols in Molecular Biology* (1995 suplemento)).

Un ejemplo de un algoritmo que es adecuado para determinar el porcentaje de identidad de secuencia y similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (1997) *Nuc. Acids Res.* 25: 3389-3402, y Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-410, respectivamente. El software para realizar análisis BLAST está disponible públicamente a través del Centro Nacional de Información Biotecnológica disponible en el World Wide Web en [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov). Parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de H, una expectativa (E) de 10, M=5, N=-4 y una comparación de ambas cadenas. 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. EE.UU.* 89: 10915) alineamientos (B) de 50, expectativa (E) de 10, M=5, N=-4, y una comparación de ambas cadenas. El algoritmo BLAST se realiza normalmente con el filtro de "baja complejidad" desactivada.

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. EE.UU.* 90: 5873-5787). Una medida de similitud proporcionada por el algoritmo BLAST es la suma de probabilidad más pequeña (P(N)), que proporciona una indicación de la probabilidad por la cual una coincidencia entre dos secuencias de nucleótidos o aminoácidos se produciría por casualidad. Por ejemplo, un ácido nucleico se considera similarmente a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es menor que aproximadamente 0,2, o menor que aproximadamente 0,01, o menor que aproximadamente 0,001.

La frase "selectivamente (o específicamente) se hibrida a" se refiere a la unión, formación de dúplex o hibridación de una molécula sólo con una secuencia de nucleótidos particular en condiciones de hibridación rigurosas cuando esa secuencia está presente en una mezcla compleja (incluyendo pero no limitado a, ADN celular o biblioteca, total o ARN).

La frase "condiciones de hibridación rigurosas" se refiere a la hibridación de secuencias de ADN, ARN, APN, u otros miméticos de ácidos nucleicos, o combinaciones de los mismos en condiciones de baja fuerza iónica y alta temperatura como se conoce en la técnica. Típicamente, en condiciones rigurosas una sonda se hibridará con su subsecuencia diana en una mezcla compleja de ácido nucleico (incluyendo pero no limitado a, ADN celular o biblioteca, total o ARN), pero no se hibrida con otras secuencias en la mezcla compleja. Las condiciones rigurosas son dependientes de secuencia y serán diferentes en diferentes circunstancias. Las secuencias más largas se hibridan específicamente a temperaturas más altas. Una extensa guía para la hibridación de ácidos nucleicos se encuentra en Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principales of hybridization and the strategy of nucleic acid assays" (1993). Generalmente, las condiciones rigurosas se seleccionan para ser aproximadamente 5-10o C más bajas que el punto de fusión térmico (Tm) para la secuencia específica a una fuerza iónica y pH definidos. La Tm es la temperatura (bajo una fuerza iónica, pH, y concentración nucleica) a la que 50% de las sondas complementarias a la diana se hibridan con la secuencia diana en equilibrio (como las secuencias diana están presentes en exceso, a Tm, el 50% de las sondas están ocupadas en equilibrio). Las condiciones rigurosas pueden ser aquellas en los que la concentración de sal es

menor de aproximadamente 1,0 M de ión sodio, típicamente aproximadamente 0,01 a 1,0 M de concentración de ión sodio (u otras sales) a pH de 7,0 a 8,3 y la temperatura es al menos aproximadamente 30°C para sondas cortas (incluyendo pero no limitado a, 10 a 50 nucleótidos) y al menos aproximadamente 60°C para sondas largas (incluyendo pero no limitado a, mayor que 50 nucleótidos). Las condiciones restrictivas también pueden ser logradas con la adición de agentes desestabilizantes tales como formamida. Para la hibridación selectiva o específica, una señal positiva puede ser al menos dos veces el fondo, opcionalmente 10 veces la hibridación de fondo. Condiciones de hibridación rigurosas ejemplares pueden ser las siguientes: 50% de formamida, 5X SSC, y SDS al 1%, incubándose a 42°C, o 5X SSC, SDS al 1%, incubándose a 65°C, con lavado en 0,2X SSC, y 0,1% de SDS a 65°C. Tales lavados pueden llevarse a cabo para 5, 15, 30, 60, 120, o más minutos.

Tal como se utiliza aquí, el término "eucariota" se refiere a organismos pertenecientes al dominio filogenético Eucarya, tales como animales (incluyendo pero no limitado a, mamíferos, insectos, reptiles, aves, etc.), ciliados, plantas (incluyendo pero no limitado a, monocotiledóneas, dicotiledóneas, algas, etc.), hongos, levaduras, flagelados, microsporidios, protistas, etc.

Tal como se utiliza aquí, el término "no eucarionte" se refiere a organismos no eucariontes. Por ejemplo, un organismo no eucarionte puede pertenecer a la Eubacteria (incluyendo pero no limitado a, *Escherichia coli*, *Thermus thermophilus*, *Bacillus stearothermophilus*, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, etc.) dominio filogenético, o la Archaea (incluyendo pero no limitado a, *Methanococcus jannaschii*, metanobacteria termoautotrófica, halobacteria tal como *Haloferax volcanii* y especies de halobacteria NRC-1, *Archaeoglobus fulgidus*, *Pirococcus furiosus*, *Pirococcus horikoshii*, *AeuroPirum Pemix*, etc.) dominio filogenético.

El término "sujeto" tal como se usa aquí, se refiere a un animal, en algunas realizaciones a un mamífero, y en otras realizaciones un humano, que es el objeto de tratamiento, observación o experimento. Un animal puede ser un animal de compañía (por ejemplo, perros, gatos, y similares), animales de granja (por ejemplo, vacas, ovejas, cerdos, caballos, y similares) o un animal de laboratorio (por ejemplo, ratas, ratones, cobayas, y similares).

El término "cantidad eficaz" Tal como se usa aquí se refiere a aquella cantidad del polipéptido de aminoácido no natural modificado que se administra que aliviará en cierto grado uno o más de los síntomas de la enfermedad, afección o trastorno a tratar. Las composiciones que contienen el polipéptido de aminoácido no natural modificado descrito en la presente memoria se pueden administrar para el tratamiento profiláctico, y/o tratamientos de mejora y terapéuticos.

Los términos "mejora" o "mejorar" significa aumentar o prolongar, ya sea en potencia o duración un efecto deseado. Así, en lo que respecta a mejorar el efecto de los agentes terapéuticos, el término "mejora" se refiere a la capacidad de aumentar o prolongar, ya sea en potencia o duración, el efecto de otros agentes terapéuticos en un sistema. Una "cantidad de mejora efectiva", tal como se usa aquí, se refiere a una cantidad adecuada para mejorar el efecto de otro agente terapéutico en un sistema deseado. Cuando se usa en un paciente, las cantidades eficaces para este uso dependerán de la gravedad y curso de la enfermedad, trastorno o afección, la terapia previa, el estado de salud del paciente y la respuesta a los fármacos, y el juicio del médico tratante.

El término "modificado", tal como se usa aquí se refiere a cualquier cambio realizado en un polipéptido dado, tales como cambios en la longitud del polipéptido, la secuencia de aminoácidos, la estructura química, modificación cotraduccional o modificación postraduccional de un polipéptido. El término forma "(modificada)" significa que los polipéptidos que se discuten están opcionalmente modificados, es decir, los polipéptidos en discusión pueden ser modificados o sin modificar.

El término "postraduccionalmente modificado" se refiere a cualquier modificación de un aminoácido natural o no natural que se produce en tal aminoácido después de que se ha incorporado en una cadena de polipéptido. El término abarca, a modo de ejemplo solamente, modificaciones cotraduccionales *in vivo*, modificaciones cotraduccionales *in vitro* (tales como en un sistema de traducción libre de células), después de la traducción en las modificaciones *in vivo*, y postraduccionales en las modificaciones *in vitro*.

En las aplicaciones profilácticas, las composiciones que contienen el polipéptido de insulina se administran a un paciente susceptible a o de otro modo en riesgo de una enfermedad, trastorno o condición. Tal cantidad se define por ser una "cantidad profilácticamente eficaz". En este uso, las cantidades precisas también dependen del estado de salud del paciente, peso, y similares. Se considera dentro de la técnica la determinación de tales cantidades profilácticamente eficaces por experimentación rutinaria (por ejemplo, ensayo clínico de un aumento de la dosis).

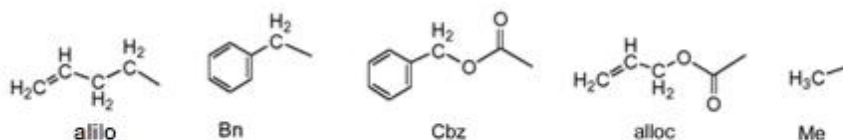
El término "protegido" se refiere a la presencia de un "grupo protector" o resto que previene la reacción del grupo funcional químicamente reactivo en determinadas condiciones de reacción. El grupo protector variará dependiendo del tipo de grupo químicamente reactivo que se protege. Por ejemplo, si el grupo químicamente reactivo es una amina o una hidrazida, el grupo protector se puede seleccionar del grupo de *tert*-butiloxicarbonilo (tBoc) y 9-fluorenilometoxicarbonilo (Fmoc). Si el grupo químicamente reactivo es un tiol, el grupo protector puede ser ortopiridilodisulfuro. 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 *tert*-

butilo. Otros grupos protectores conocidos en la técnica también se pueden usar en o con los métodos y composiciones descritas en este documento, incluidos los grupos fotolábiles como Nvoc y MeNvoc. Otros grupos protectores conocidos en la técnica también se pueden usar en o con los métodos y composiciones descritos en el presente documento.

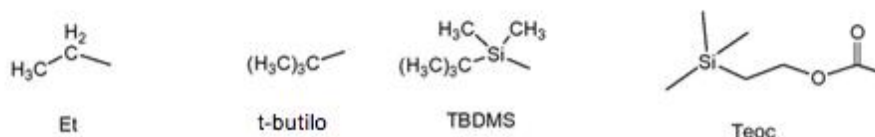
5

A modo de ejemplo solamente, grupos de bloqueo/protectores pueden seleccionarse de:

10

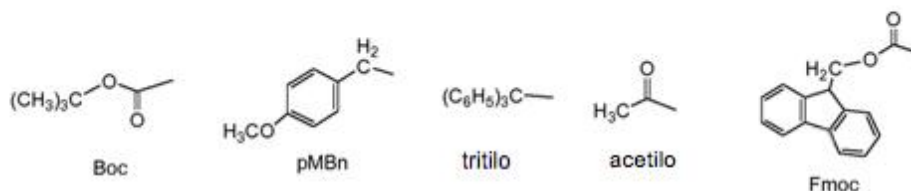


15



20

25



30

Otros grupos protectores se describen en Greene y Wuts, Protective Groups in Organic Synthesis, 3ª Ed., John Wiley & Sons, Nueva York, NY, 1999.

35

En las aplicaciones terapéuticas, las composiciones que contienen el polipéptido de aminoácido no natural modificado se administran a un paciente que ya padece una enfermedad, afección o trastorno, en una cantidad suficiente para curar o al menos detener parcialmente los síntomas de la enfermedad, trastorno o condición. Tal cantidad se define para ser una "cantidad terapéuticamente eficaz", y dependerá de la gravedad y curso de la enfermedad, trastorno o afección, la terapia previa, el estado de salud del paciente y la respuesta a los fármacos, y el juicio del médico tratante. Se considera dentro de la experiencia de la técnica de una persona determinar dichas cantidades terapéuticamente eficaces por experimentación rutinaria (por ejemplo, ensayo clínico de un aumento de la dosis).

40

El término "tratar" se usa para hacer referencia a tratamiento profiláctico y/o tratamientos terapéuticos.

45

Polipéptidos de aminoácidos no naturales codificados presentados en el presente documento pueden incluir compuestos isotópicamente etiquetados con uno o más átomos reemplazados por un átomo que tiene una masa atómica o número másico diferente de la masa atómica o número másico encontrado habitualmente en la naturaleza. Ejemplos de isótopos que se pueden incorporar en los presentes compuestos incluyen isótopos de hidrógeno, carbono, nitrógeno, oxígeno, flúor y cloro, tales como 2H, 3H, 13C, 14C, 15N, 18O, 17O, 35S, 18F, 36Cl, respectivamente. Ciertos compuestos isotópicamente etiquetados descritos en el presente documento, por ejemplo aquellos en los que los isótopos radiactivos tales como 3H y 14C están incorporados, pueden ser útiles en ensayos de distribución tisular de sustrato y/o drogas. Además, la sustitución con isótopos tales como deuterio, es decir, 2H, puede proporcionar ciertas ventajas terapéuticas resultantes de una mayor estabilidad metabólica, por ejemplo aumento de la vida media in vivo o menores requisitos de dosificación.

50

55

Todos los isómeros incluyendo pero no limitado a diastereómeros, enantiómeros y mezclas de los mismos se consideran como parte de las composiciones descritas en este documento. En realizaciones adicionales, los polipéptidos de aminoácidos no naturales codificados se metabolizan tras la administración a un organismo en necesidad de producir un metabolito que se utiliza entonces para producir un efecto deseado, incluyendo un efecto terapéutico deseado. En realizaciones adicionales son metabolitos activos de polipéptidos de aminoácidos no naturales codificados.

60

En algunas situaciones, los polipéptidos de aminoácidos no naturales codificados pueden existir como tautómeros. Además, los polipéptidos de aminoácidos no naturales codificados descritos en este documento pueden existir en

65

formas no solvatadas así como solvatadas con disolventes farmacéuticamente aceptables tales como agua, dimetilol, y similares. También se consideran las formas solvatadas que se describe en este documento. Personas de experiencia ordinaria en la técnica reconocerán que algunos de los compuestos en el presente documento pueden existir en varias formas tautoméricas. Todas estas formas tautoméricas se consideran como parte de las composiciones descritas en este documento.

A menos que se indique lo contrario, se emplean métodos convencionales de espectroscopia de masas, RMN, HPLC, química de proteínas, bioquímica, técnicas de ADN recombinante y farmacología, dentro de la experiencia de la técnica.

## DESCRIPCIÓN DETALLADA

### I. Introducción

Los polipéptidos de insulina que comprenden al menos un aminoácido no natural se proporcionan en la invención. En ciertas realizaciones de la invención, el polipéptido de insulina con al menos un aminoácido no natural incluye al menos una modificación postraducciona. En una realización, al menos una modificación postraducciona comprende la unión de una molécula incluyendo, pero no limitado a, una etiqueta, un colorante, un polímero, un polímero soluble en agua, un derivado de polietilenglicol, un fotoentrecruzador, un radionúclido, un compuesto citotóxico, un fármaco, un marcador de afinidad, un marcador 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 metales, un cofactor, un ácido graso, un carbohidrato, un polinucleótido, un ADN, un ARN, un polinucleótido antisentido, un sacárido, un dendrímero soluble en agua, una ciclodextrina, un ácido ribonucleico inhibitorio, un biomaterial, una nanopartícula, un marcador de espín, un fluoróforo, un resto que contiene metal, un resto radiactivo, un grupo novedoso funcional, un grupo que forma covalente o no covalente interactúa con otras moléculas, un resto excitable de radiación actínica, un resto fotoisomerizable, biotina, un derivado de la 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 enlazado a carbono, un agente redoxactivo, un tioaminoácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo quimioluminiscente, un grupo electrón denso, un grupo magnético, un grupo de intercalación, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, un marcador detectable, una molécula pequeña, un punto cuántico, un nanotransmisor, un nucleótido radio, un radiotransmisor, una agente de captura de neutrones, o cualquier combinación de los anteriores o cualquier otro compuesto deseable o sustancia, que comprende un segundo grupo reactivo a al menos un aminoácido no natural que comprende una primer grupo utilizando metodología química reactiva que se conoce por una persona de experiencia ordinaria en la técnica que son adecuados para los grupos reactivos particulares. Por ejemplo, el primer grupo reactivo es un resto de alquínilo (incluyendo pero no limitado a, en el aminoácido no natural de p-propargiloxifenilalanina, donde el grupo propargilo también a veces se denomina un resto de acetileno) y el segundo grupo reactivo es un resto azido, y se utilizan metodologías de química de cicloadición [3 + 2]. En otro ejemplo, el primer grupo reactivo es el resto azido (incluyendo pero no limitado a, en aminoácido no natural de p-azido-L-fenilalanina) y el segundo grupo reactivo es el resto alquínilo. En ciertas realizaciones del polipéptido de insulina modificada de la presente invención, al menos un aminoácido no natural (incluyendo pero no limitados a, aminoácido no natural que contiene un grupo funcional cEto) que comprende al menos una modificación postraducciona, se utiliza donde al menos una modificación post-traducciona comprende un resto sacárido. En ciertas realizaciones, la modificación postraducciona se hace *in vivo* en una célula eucariota o en una célula no eucariótica. Un enlazador, polímero, polímero soluble en agua, u otra molécula pueden unir la molécula al polipéptido. La molécula puede estar enlazada directamente al polipéptido.

En ciertas realizaciones, la proteína incluye al menos una modificación postraducciona que se hace *in vivo* por una célula huésped, donde la modificación postraducciona no se hace normalmente por otro tipo de célula huésped. En ciertas realizaciones, la proteína incluye al menos una modificación postraducciona que se hace *in vivo* por una célula eucariótica, donde la modificación postraducciona no se hace normalmente por una célula no eucariótica. Ejemplos de modificaciones posteriores a la traducción incluyen, pero no se limitan a, glicosilación, acetilación, acilación, modificación de lípidos, palmitoilación, además palmitato, la fosforilación, la modificación de enlace a glicolípidos, y similares.

En algunas realizaciones, el polipéptido de insulina comprende uno o más aminoácidos no naturales codificados para la glicosilación, acetilación, acilación, modificación de lípidos, palmitoilación, además palmitato, fosforilación, o modificación de enlace glicolípido del polipéptido. En algunas realizaciones, el polipéptido de insulina comprende uno o más aminoácidos no naturales codificados para la glicosilación del polipéptido. En algunas realizaciones, el polipéptido de insulina comprende uno o más aminoácidos naturalmente codificados para la glicosilación, acetilación, acilación, modificación de lípidos, palmitoilación, además palmitato, fosforilación, o modificación de enlace glicolípido del polipéptido. En algunas realizaciones, el polipéptido de insulina comprende uno o más aminoácidos naturalmente codificados para la glicosilación del polipéptido.

En algunas realizaciones, el polipéptido de insulina comprende una o más adiciones y/o sustituciones que mejoran la

glicosilación del polipéptido de aminoácidos no naturales codificados. En algunas realizaciones, el polipéptido de insulina comprende una o más deleciones que mejoran la glicosilación del polipéptido. En algunas realizaciones, el polipéptido de insulina comprende una o más adiciones y/o sustituciones que mejoran la glicosilación en un aminoácido diferente en el polipéptido de aminoácidos no naturales codificados. En algunas realizaciones, el polipéptido de insulina comprende una o más deleciones que mejoran la glicosilación en un aminoácido diferente en el polipéptido. En algunas realizaciones, el polipéptido de insulina comprende una o más adiciones y/o sustituciones que mejoran la glicosilación en un aminoácido no naturalmente codificado en el polipéptido de aminoácidos no naturalmente codificados. En algunas realizaciones, el polipéptido de insulina comprende una o más adiciones y/o sustituciones que mejoran la glicosilación en un aminoácido codificado de forma natural en el polipéptido de aminoácidos no naturalmente codificados. En algunas realizaciones, el polipéptido de insulina comprende una o más adiciones codificadas naturalmente de aminoácidos y/o sustituciones que mejoran la glicosilación en un aminoácido diferente en el polipéptido. En algunas realizaciones, el polipéptido de insulina comprende una o más adiciones y/o sustituciones que mejoran la glicosilación en un aminoácido codificado de forma natural en el polipéptido de aminoácidos no naturalmente codificados. En algunas realizaciones, el polipéptido de insulina comprende una o más adiciones y/o sustituciones que mejoran la glicosilación en un aminoácido no naturalmente codificado en el polipéptido de aminoácidos no naturalmente codificados.

En una realización, la modificación postraducciona comprende la unión de un oligosacárido a una asparagina por un enlace GlcNAc-asparagina (incluyendo pero no limitado a, donde el oligosacárido comprende (GlcNAc-Man)<sub>2</sub>-Man-GlcNAc-GlcNAc, y similares). En otra realización, la modificación postraducciona comprende adjunto de un oligosacárido (incluyendo pero no limitado a, Gal-GalNAc, Gal-GlcNAc, etc.) a una serina o treonina por una GalNAc-serina, un GalNAc-treonina, un GlcNAc-serina, o un enlace GlcNAc-treonina. En ciertas realizaciones, una proteína o polipéptido de la invención pueden comprender una secuencia de secreción o localización, una etiqueta de epítipo, una etiqueta FLAG, una etiqueta de poli-Histidina, una fusión GST, y/o similares. Los ejemplos de señales de secuencias de secreción incluyen, pero no se limitan a, una secuencia de señal de secreción procariota, una señal de secuencia de secreción eucariota, una señal de secuencia de secreción eucariota 5'-optimizada para la expresión bacteriana, una nueva secuencia de señal de secreción, la señal de secuencia de secreción de liasa de pectato, secuencia de señal de secreción Omp A, y una secuencia de señal de secreción de fagos. Los ejemplos de secuencias de señal de secreción, incluyen, pero no se limitan a, STII (procariota), Fd Gill y M13 (fago), Bgl2 (levadura), y la secuencia de señal bla derivada de un transposón. Cualquiera de tales secuencias puede modificarse para proporcionar un resultado deseado con el polipéptido, incluyendo pero no limitado a, la sustitución de una secuencia de señal con una secuencia de señal diferente, sustituyendo una secuencia líder con una secuencia líder diferente, etc.

La proteína o polipéptido de interés 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 anti naturales. Los aminoácidos no naturales pueden ser iguales o diferentes, por ejemplo, puede haber 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más sitios diferentes en la proteína que comprenden 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más diferentes aminoácidos no naturales. En ciertas realizaciones, al menos uno, pero menos que todos, de un aminoácido particular presente en una versión de origen natural de la proteína está sustituido con un aminoácido no natural.

La presente descripción proporciona métodos y composiciones a base de insulina que comprende al menos un aminoácido no naturalmente codificado. Introducción de al menos un aminoácido no naturalmente codificado en la insulina puede permitir la aplicación de químicas de conjugación que implican reacciones químicas específicas, incluyendo, pero no limitado a, con uno o más aminoácidos no naturales codificados mientras que no reacciona con los 20 aminoácidos que ocurren comúnmente. En algunas realizaciones, la insulina que comprende el aminoácido no naturalmente codificado está ligado a un polímero soluble en agua, tal como polietilenglicol (PEG), a través de la cadena lateral del aminoácido no naturalmente codificado. Esta descripción proporciona un método altamente eficiente para la modificación selectiva de proteínas con derivados de PEG, que implica la incorporación selectiva de aminoácidos no genéticamente codificados, incluyendo, pero no limitado a, aquellos aminoácidos que contienen grupos funcionales o sustituyentes no encontrados en los 20 aminoácidos naturalmente incorporados, incluyendo pero no limitado a una cetona, una azida o resto acetileno, en proteínas en respuesta a un codon selector y la posterior modificación de los aminoácidos con un derivado de PEG adecuadamente reactivo. Una vez incorporado, las cadenas laterales de los aminoácidos se pueden modificar mediante la utilización de metodologías de química conocidas por personas de experiencia ordinaria en la técnica por ser adecuadas para los grupos funcionales particulares o sustituyentes presentes en el aminoácido no naturalmente codificado. Metodologías de química conocidas de una amplia variedad son adecuadas para su uso en la presente descripción para incorporar un polímero soluble en agua en la proteína. Dichas metodologías incluyen, pero no se limitan a una reacción de cicloadición de Huisgen [3 + 2] (véase, por ejemplo, Padwa, A. en *Comprehensive Organic Synthesis*, Vol. 4, (1991) Ed. Trost, B.M., Pergamon, Oxford, p. 1069-1109; y, Huisgen, R. en *1,3-Dipolar Cycloaddition Chemistry*, (1984) Ed Padwa, A., Wiley, Nueva York, p 1-176) con, incluyendo pero no limitado a, los derivados de acetileno o azida, respectivamente.

Debido a que el método de cicloadición de Huisgen [3 + 2] implica una cicloadición en lugar de una reacción de sustitución nucleófila, las proteínas se pueden modificar con extremadamente alta selectividad. La reacción puede llevarse a cabo a temperatura ambiente en condiciones acuosas con excelente regioselectividad (1,4> 1,5) mediante

la adición de cantidades catalíticas de sales a la mezcla de reacción Cu(I). Véase, por ejemplo, Tornado, *et al.*, (2002) J. Org. Chem. 67: 3057-3064; y, Rostovtsev, *et al.*, (2002) Angew. Chem. Int. Ed. 41: 2596-2599; y WO 03/101972. Una molécula que se puede añadir a una proteína de la invención a través de una cicloadición [3 + 2] incluye prácticamente cualquier molécula con un grupo adecuado funcional o sustituyente incluyendo, pero no limitado a un azido o derivado de acetileno. Estas moléculas se pueden añadir a un aminoácido no natural con un grupo acetileno, incluyendo pero no limitados a, p-propargiloxifeniloalanina, o grupo azido, incluyendo pero no limitados a pazidofeniloalanina, respectivamente.

El anillo de cinco miembros que resulta de la cicloadición de Huisgen [3 + 2] generalmente no es reversible en entornos reductores y es estable frente a la hidrólisis para períodos prolongados en ambientes acuosos. En consecuencia, las características físicas y químicas de una amplia variedad de sustancias pueden ser modificadas bajo exigentes condiciones acuosas con los derivados de PEG activos de la presente invención. Aún más importante, porque restos de azida y acetileno son específicos para uno de otro (y no reaccionan, por ejemplo, con cualquiera de los 20 aminoácidos comunes, genéticamente codificados), las proteínas pueden ser modificados en uno o más sitios específicos con extremadamente alta selectividad.

La invención también proporciona derivados solubles e hidrolíticamente estables en agua de derivados de PEG y polímeros hidrófilos relacionados que tienen uno o más restos de acetileno o azida. Los derivados de polímero de PEG que contienen restos de acetileno son altamente selectivos para acoplar con restos de azida que se han introducido selectivamente en proteínas en respuesta a un codón selector. Del mismo modo, los derivados de polímero de PEG que contienen restos de azida son altamente selectivos para acoplar con restos de acetileno que se han introducido selectivamente en proteínas en respuesta a un codón selector.

Más específicamente, los restos de azida comprenden, pero no se limitan a, azidas de alquilo, azidas de arilo y derivados de estas azidas. Los derivados de las azidas de alquilo y arilo pueden incluir otros sustituyentes siempre que se mantenga la reactividad de acetileno específico. Los restos de acetileno comprenden acetilenos de alquilo y arilo y derivados de cada uno. Los derivados de los acetilenos de alquilo y arilo pueden incluir otros sustituyentes siempre que se mantenga la reactividad de azida específica.

La presente invención proporciona conjugados de sustancias que tienen una amplia variedad de grupos funcionales, sustituyentes o restos, con otras sustancias, incluyendo pero no limitado a una etiqueta; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotoentrecruzador; un radionucleido; un compuesto citotóxico; una droga; una etiqueta de afinidad; un marcador 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 metales; un cofactor; un ácido graso; un carbohidrato; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido; un dendrímero soluble en agua; una ciclodextrina; un ácido ribonucleico inhibitorio; un biomaterial; una nanopartícula; un marcador de espín; un fluoróforo, un resto que contiene metal; un resto radioactivo; un grupo funcional novelo; un grupo que covalentemente o no covalentemente interactúa con otras moléculas; un resto de radiación actínica excitable; un resto fotoisomerizable; biotina; un derivado de biotina; un análogo de biotina; un resto de la incorporación de un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargado; un azúcar enlazado a carbono; un agente redoxactivo; un amino tioácido; un resto tóxico; un resto marcado isotópicamente; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un de electrón denso; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; un marcador detectable; una molécula pequeña; un punto cuántico; un nanotransmisor; un radionucleótido; un radiotransmisor; un agente de captura de neutrones; o cualquier combinación del otro compuesto o sustancia deseable. La presente invención también incluye conjugados de sustancias que tienen restos de azida o acetileno con derivados de polímeros de PEG que tienen los correspondientes restos de acetileno o azida. Por ejemplo, un polímero de PEG que contiene un resto azida se puede acoplar a una molécula biológicamente activa en una posición en la proteína que contiene un aminoácido no genéticamente codificado teniendo una funcionalidad de acetileno. El enlace por el que el PEG y la molécula biológicamente activa se acoplan incluye pero no se limita a productos de cicloadición de Huisgen [3 + 2].

Está bien establecido en la técnica que el PEG se puede utilizar para modificar las superficies de los biomateriales (véase, por ejemplo, Patente de Estados Unidos 6.610.281; Mehvar, R., J. Pharm Pharm Sci, 3(1): 125-136 (2000)). La invención también incluye biomateriales que comprenden una superficie que tiene una o más sitios reactivos de azida o de acetileno y uno o más de los polímeros que contienen azida o acetileno de la invención acoplados a la superficie a través del enlace de cicloadición de Huisgen [3 + 2]. Biomateriales y otras sustancias también se pueden acoplar a derivados de polímeros activados por azida o acetileno a través de una unión distinta de la azida o un enlace de acetileno, tal como a través de un enlace que comprende un ácido carboxílico, amina, alcohol o un resto tiol, para salir de la azida o resto acetileno disponible para reacciones posteriores.

La descripción incluye un método de síntesis de polímeros que contiene azida y acetileno de la divulgación. En el caso del derivado de PEG que contiene azida, la azida puede ser unida directamente a un átomo de carbono del polímero. Alternativamente, el derivado de PEG que contiene azida se puede preparar uniendo un agente de unión que tiene el resto azida en un extremo a un polímero activado convencional, de modo que el polímero resultante tiene el resto azida en su término. En el caso del derivado de PEG que contiene acetileno, el acetileno puede estar

unido directamente a un átomo de carbono del polímero. Alternativamente, el derivado de PEG que contiene acetileno se puede preparar uniendo un agente de unión que tiene el resto de acetileno en un extremo a un polímero activado convencional, de modo que el polímero resultante tiene el resto acetileno en su término.

5 Más específicamente, en el caso del derivado de PEG que contiene azida, un polímero soluble en agua que tiene al menos un resto hidroxilo activo se somete a una reacción para producir un polímero sustituido que tiene un resto más reactivo, tal como un mesilato, tresilato, tosilato o grupo saliente halógeno. La preparación y uso de derivados de PEG que contienen haluros de ácido de sulfonilo, átomos de halógeno y otros grupos salientes son conocidos por personas de experiencia ordinaria en la técnica. El polímero sustituido resultante se somete entonces a una reacción para sustituir el resto más reactivo por un resto azida en el extremo del polímero. Alternativamente, un polímero soluble en agua que tiene al menos un resto nucleófilo o electrófilo activo se somete a una reacción con un agente de unión que tiene una azida en un extremo de manera que se forma un enlace covalente entre el polímero PEG y el agente de enlace y el resto azida se coloca en el extremo del polímero. Restos nucleófilos y electrófilos, incluyendo aminas, tioles, hidrazidas, hidrazinas, alcoholes, carboxilatos, aldehídos, cetonas, tioésteres y similares, son conocidos por personas de experiencia ordinaria en la técnica.

Más específicamente, en el caso del derivado de PEG que contiene acetileno, un polímero soluble en agua que tiene al menos un resto hidroxilo activo se somete a una reacción para desplazar un halógeno u otro grupo saliente activado a partir de un precursor que contiene un resto acetileno. Alternativamente, un polímero soluble en agua que tiene al menos un resto nucleófilo o electrófilo activo se somete a una reacción con un agente de unión que tiene un acetileno en un extremo de manera que se forma un enlace covalente entre el polímero PEG y el agente de unión y el resto acetileno está posicionado en el extremo del polímero. El uso de restos de halógeno, grupo saliente activado, restos nucleófilos y electrófilos en el contexto de la síntesis orgánica y la preparación y uso de derivados de PEG está bien establecido para los profesionales en la técnica.

La descripción también proporciona un método para la modificación selectiva de proteínas para añadir otras sustancias a la proteína modificada, incluyendo pero no limitado a polímeros solubles en agua tales como derivados de PEG y PEG que contienen un resto de azida o acetileno. Los derivados de PEG de azida y que contiene acetileno se pueden utilizar para modificar las propiedades de superficies y moléculas en las que la biocompatibilidad, estabilidad, solubilidad y la falta de inmunogenicidad son importantes, mientras que al mismo tiempo proporciona un medio más selectivo de la fijación de los derivados de PEG a las proteínas que fue previamente conocido en la técnica.

#### **Métodos de ácido nucleico recombinante general para uso con la invención**

En numerosos casos de la presente descripción, los ácidos nucleicos que codifican un polipéptido de insulina de interés se aislaron, se clonaron y a menudo se alteraron usando métodos recombinantes. Se usan tales casos, incluyendo pero no limitados a, para la expresión de proteínas o durante la generación de variantes, derivados, casetes de expresión, u otras secuencias derivadas de un polipéptido de insulina. En algunos casos, las secuencias que codifican los polipéptidos de la invención están unidas operativamente a un promotor heterólogo.

Una secuencia de nucleótidos que codifica un polipéptido de insulina que comprende un aminoácido no naturalmente codificado puede ser sintetizada sobre la base de la secuencia de aminoácidos del polipéptido padre, incluyendo pero no limitado a, que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 1 y SEQ ID NO: 2 y luego cambiando la secuencia de nucleótidos a fin de efectuar la introducción (es decir, la incorporación o sustitución) o eliminación (es decir, deleción o sustitución) del residuo de aminoácido pertinente. La secuencia de nucleótidos puede ser convenientemente modificada por mutagénesis de sitio dirigido de acuerdo con métodos convencionales. Alternativamente, la secuencia de nucleótidos se puede preparar por síntesis química, incluyendo, pero no limitado a, mediante el uso de un sintetizador de oligonucleótidos, en el que los oligonucleótidos se diseñan basándose en la secuencia de aminoácidos del polipéptido deseado, y preferiblemente seleccionando aquellos codones que están favorecidos en la célula huésped en la que se producirá el polipéptido recombinante.

Esta divulgación utiliza técnicas de rutina en el campo de la genética recombinante. Textos básicos que describen los métodos generales de uso en esta invención incluyen Sambrook et al., *Molecular Cloning, A Laboratory Manual*. (3ª ed 2001); Krieglger, *Gene Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (Ausubel et al., eds., 1994).

Los textos generales que describen técnicas de biología molecular incluyen Berger y Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology*, volumen 152, Academic Press, Inc., San Diego, CA (Bergcr); Sambrook et al., *Molecular Cloning A Laboratory Manual* (2ª Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, 1989 ("Sambrook") and *Current Protocols in Molecular Biology*, F.M. Ausubel et al., Eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (complementado por 1999) ("Ausubel"). Estos textos describen mutagénesis, el uso de vectores, promotores y muchos otros temas relevantes relacionados con, incluyendo pero no limitado a, la generación de genes o polinucleótidos que incluyen codones selectores para la producción de proteínas que incluyen aminoácidos no naturales, ARNt ortogonales, sintetisas ortogonales y pares de los mismos.

Varios tipos de mutagénesis se utilizan en la invención para una variedad de propósitos, incluyendo pero no limitados a, para producir nuevas sintetetasas o ARNt, para mutar moléculas de trina, para mutar polinucleótidas que codifican sintetetasas, para producir bibliotecas de ARNt, para producir bibliotecas de sintetetasas, para producir codones selectores, para insertar codones selectores que codifican aminoácidos no naturales en una proteína o polipéptido de interés. Incluyen, pero no se limitan a mutagénesis de sitio dirigido de punto aleatorio, recombinación homóloga, barajado de ADN u otros métodos de mutagénesis recursivos, construcción quimérica, mutagénesis usando uracilo que contiene plantillas, mutagénesis de oligonucleótido dirigido, mutagénesis de ADN fosforotioatemodificado, la mutagénesis utilizando ADN dúplex con huecos o similares, mutagénesis mediada por PCT, o cualquier combinación de los mismos. Métodos adecuados adicionales incluyen el punto de reparación de genes, mutagénesis usando cepas huésped deficientes en reparo, selección restringida y restricción-purificación, mutagénesis por delección, mutagénesis por síntesis génica total, reparación de doble cadena de la rotura, y similares. La mutagénesis, incluyendo, pero no limitado a, la participación de construcciones quiméricas, también se incluyen en la presente descripción. En una realización, la mutagénesis puede ser guiada por la información conocida de la molécula de origen natural o alterado o molécula mutada de origen natural, incluyendo, pero no limitado a secuencia, comparaciones de secuencias, las propiedades físicas, estructura secundaria, terciaria, o cuaternaria, estructura cristalina o similares.

Los textos y ejemplos que se encuentran en el presente documento describen estos procedimientos. Información adicional se encuentra en las siguientes publicaciones y referencias citadas aquí: Ling et al., Approaches to DNA mutagenesis: an overview, *Anal Biochem.* 254-(2): 157-178 (1997); Dale et al., 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. y Lilley, D.M.J. eds, Springer Verlag, Berlin) (1987); Kunkel, Rapid end efficient site-specific mutagenesis without phenotypic selection, *Proc. Natl. Acad. Sci. EE.UU.* 82: 488-492 (1985); Kunkel et al., Rapid and efficient site-specific mutagenesis without phenotypic selection, *Methods in Enzymol.* 154, 367-382 (1987); Bass et al., Mutant TRP repressors with new DNA-binding specificities, *Science* 242: 240-245 (1988); Zoller y Smith, Oligonucleotide-directed mutagenesis using M13-derived vectores: an efficient and general procedure for the production of point mutations in any DNA fragment, *Nucleic Acids Res.* 10: 6487-6500 (1982); Zoller y Smith, Oligonucleotide-directed mutagenesis of DNA fragments clones into M13 vectors, *Methods in Enzymol.* 100: 468-500 (1983); Zoller y 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 et al., The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, *Nucl. Acids Res.* 13: 8749-8764 (1985); Taylor et al., The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA, *Nucl. Acids Res.* 13: 8765-8785 (1985); Nakamaye y Eckstein, Inhibition of restriction endonuclease Nci-1 cleavage by phosphorothioate groups and its application to oligonucleotide directed mutagenesis, *Nucl. Acids Res.* 14: 9679-9698 (1986); Sayers et al., 5'-3' Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, *Nucl. Acids Res.* 16: 791-802 (1988); Sayers et al., 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 et al., The gapped duplex DNA approach to oligonucleotide-directed mutation construction of mutations via gapped duplex DNA, *Methods in Enzymol.* 12: 9441-9456 (1984); Kramer y Fritz Oligonucleotide-directed construction os mutations via gapped duplex DNA, *Methods in Enzymol.* 154: 350-367 (1987); Kramer et al., Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, *Nucl. Acids Res.* 16: 7207 (1988); Fritz et al., Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro, *Nucl. Acids Res.* 16: 6987-6999 (1988); Kramer et al., Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*, *Cell* 38: 879-887 (1984); Carter et al., 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 y Henikoff, Use of oligonucleotides to generate large deletions, *Nucl. Acids Res.* 14: 5115 (1986); Wells et al., Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, *Phil. Trans. R. Soc. Lond. A* 317: 415 423-(1986); Nambiar et al., Total synthesis and cloning of a gene coding for the ribonuclease S protein, *Science* 223: 1299-1301 (1984); Sakmar y 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 et al., Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, *Gene* 34: 315-323 (1985); Grundström et al., 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. EE.UU.*, 83: 7177-7181 (1986); Arnold, Protein engineering for unusual environments, *Current Opinion in Biotechnology* 4:450-455 (1993); Sieber, et al., *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 métodos anteriores se pueden encontrar en *Methods in Enzymology*, Volumen 154, que también describe controles útiles para solucionar problemas con diferentes métodos de mutagénesis.



Los oligonucleótidos, por ejemplo, para su uso en la mutagénesis de la presente descripción, por ejemplo, mutando bibliotecas de sintetasas, o alterando ARNt, son típicamente sintetizados químicamente de acuerdo con el método del triéster de fosoramidita en fase sólida descrito por Beaucage y Caruthers, *Tetrahedron Letts.* 22 (20): 1859-1862, (1981) por ejemplo, usando un sintetizador automatizado, como se describe en Needham-VanDevanter et al., *Nucleic Acids Res.* 12: 6159-6168 (1984).

La descripción también se refiere a células huésped eucariotas, células huésped no eucarióticas, y organismos para la incorporación *in vivo* de un aminoácido no natural a través de pares ortogonales de ARNt/RS. Las células huésped se modifican por ingeniería genética (incluyendo, pero no limitado a, transformada, transducida o transfectada) con los polinucleótidos de la divulgación o construcciones que incluyen un polinucleótido de la divulgación, incluyendo pero no limitado a, un vector de la invención, que puede ser, por ejemplo, un vector de clonación o un vector de expresión. Por ejemplo, las regiones codificantes para la ARNt ortogonal, la sintetasa ortogonal de ARNt, y la proteína a ser derivatizada se unen operativamente a elementos de control de expresión génica que son funcionales en la célula huésped deseada. El vector puede ser, 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 por métodos convencionales incluyendo electroporación (Fromm *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 82, 5824 (1985)), infección por vectores virales, de alta velocidad de penetración balística por pequeñas partículas con el ácido nucleico dentro de la matriz de pequeñas perlas o partículas, o sobre la superficie (Klein *et al.*, *Nature* 327, 7073 (1987)), y/o similares. Las técnicas adecuadas para la transferencia de ácido nucleico en células *in vitro* incluyen el uso de liposomas, microinyección, fusión celular, DEAE-dextran, el método de precipitación de fosfato de calcio, etc. Técnicas de transferencia génica *in vivo* incluyen, pero no se limitan a, transfección con vectores virales (típicamente) los vectores retrovirales y transfección mediada por liposoma de revestimiento de cubierta viral [Dzau et al., *Trends in Biotechnology* 11: 205 210 (1993)]. En algunas situaciones, puede ser deseable proporcionar la fuente de ácido nucleico con un agente que se dirige a las células diana, tales como un anticuerpo específico para una proteína de membrana de superficie celular o la célula diana, un ligando para un receptor en la célula diana, etc. Cuando se emplean liposomas, las proteínas que se unen a una proteína de membrana de superficie celular asociada con la endocitosis pueden usarse para dirigir y/o facilitar la captación, por ejemplo proteínas de la cápside o fragmentos de las mismas trópicas para un tipo de célula particular, anticuerpos para proteínas que experimentan internalización en el ciclo, proteínas que dirigen la localización intracelular y potencian la vida media intracelular.

Las células huésped manipuladas pueden cultivarse en medios nutrientes convencionales modificados según sea apropiado para actividades tales como, por ejemplo, etapas de detección, la activación de promotores o selección de transformantes. Estas células opcionalmente pueden cultivarse en organismos transgénicos. Otras referencias útiles, incluyendo pero no limitado al aislamiento de células y el cultivo (por ejemplo, para el aislamiento posterior de ácidos nucleicos) incluyen Freshney (1994) *Culture of Animal Cells*, a *Manual of Basic Technique*, tercera edición, Wiley-Liss, Nueva York y las referencias citadas en la misma; Payne *et al.* (1992) *Plant Cell, Tissue and Organ Culture in Liquid Systems* John Wiley & Sons, Inc., Nueva York, NY; Gamborg y Phillips (eds.) (1995) *Plant Cell, Tissue and Organ Culture*; *Fundamental Methods Springer Lab Manual*, Springer-Verlag (Berlin Heidelberg Nueva York) y Atlas and Parks (eds.) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL.

Varios métodos bien conocidos de introducción de ácidos nucleicos diana en las células están disponibles, cualquiera de los cuales puede ser utilizado en la divulgación. Estos incluyen: fusión de las células receptoras con protoplastos bacterianos que contienen el ADN, electroporación, bombardeo de proyectil, y la infección con vectores virales (descrito más adelante), etc. Las células bacterianas pueden ser usadas para amplificar el número de plásmidos que contienen construcciones de ADN de esta revelación. Las bacterias se cultivan hasta la fase logarítmica y los plásmidos dentro de las bacterias se pueden aislar por una variedad de métodos conocidos en la técnica (véase, por ejemplo, Sambrook). Además, los kits están disponibles comercialmente para la purificación de plásmidos a partir de bacterias, (véase, por ejemplo, EasyPrep™, FlexiPrep™, ambos de Pharmacia Biotech; StrataClean™ de Stratagene; y, QIAprep™ de Qiagen). Los plásmidos aislados y purificados entonces se manipulan adicionalmente para producir otros plásmidos, que se utilizan para transfectar células o se incorporan en vectores relacionados para infectar organismos. Los vectores típicos contienen terminadores de transcripción y traducción, transcripción y secuencias de iniciación de la traducción, y promotores útiles para la regulación de la expresión del ácido nucleico diana particular. Los vectores comprenden opcionalmente casetes genéricos de expresión que contienen al menos una secuencia terminadora independiente, secuencias que permiten la replicación del casete en eucariotas o procariontes, o ambas, (incluyendo pero no limitado a, vectores lanzaderos) y marcadores de selección tanto para sistemas procariontes y eucariotas. Los vectores son adecuados para la replicación e integración en procariontes, eucariotas, o ambos. Véase, Gillam y Smith, *Gene* 8:81 (1979); Roberts, et al., *Nature*, 328: 731 (1987); Schneider, E., *et al.*, *Protein Expr. Purif.* 6 (1): 10-14 (1995); Ausubel, Sambrook, Berger (todos supra). Se proporciona un catálogo de bacterias y bacteriófagos útiles para la clonación, por ejemplo, por la ATCC, por ejemplo, el catálogo ATCC de bacterias y bacteriófagos (1992) Gherna *et al.* (eds) publicado por la ATCC. Procedimientos básicos adicionales para la secuenciación, la clonación y otros aspectos de biología molecular y consideraciones teóricas subyacentes también se encuentran en Watson *et al.* (1992) *Recombinant DNA Second Edition Scientific American Books*, NY. Además, esencialmente cualquier ácido nucleico (y prácticamente cualquier ácido nucleico marcado, ya sea estándar o no estándar) puede ser personalizado o estándar a partir de cualquiera de una variedad de fuentes comerciales, tales como la Midland Certified Reagent Company (Midland, TX disponible en el internet en

mrcr.com), la Great American Gene Company (Ramona, CA disponible en el internet en genco.com, ExpressGen Inc. (Chicago, IL disponible en el internet en expressgen.com), Operon Technologies Inc. (Alameda, CA) y muchos otros.

## 5 CODONES SELECTORES

Los codones selectores de la divulgación expanden el marco de codón genético de la maquinaria biosintética de proteínas. Por ejemplo, un codón selector incluye, pero no se limite a un único codón de tres bases, un codón sin sentido, tales como un codón de parada, incluyendo, pero no limitado a, un codón ámbar (UAG), un codón ocre, o un codón de ópalo (UGA), un codón no natural, un codón de cuatro o más de bases, un codón raro, o similares. Es fácilmente evidente para expertos ordinarios en la técnica que existe una amplia gama en el número de codones selectores que puede ser introducido en un gen o polinucleótido deseado, incluyendo, pero no limitado a, uno o más, dos o más, tres o más, 4, 5, 6, 7, 8, 9, 10 o más en un solo polinucleótido que codifica al menos una porción del polipéptido de insulina. También es fácilmente evidente para expertos ordinarios en la técnica que existe una amplia gama en el número de codones selectores que puede ser introducido en un gen o polinucleótido deseado, incluyendo, pero no limitado a, uno o más, dos o más, tres o más, 4, 5, 6, 7, 8, 9, 10 o más total encontrados en las secuencias de polinucleótidos de cadena A y cadena B y que codifican al menos una porción del polipéptido de insulina.

En una realización, los métodos implican el uso de un codón selector que es un codón de terminación para la incorporación de uno o más aminoácidos no naturales *in vivo*. Por ejemplo, un O-ARNt se produce que reconoce el codón de parada, incluyendo pero no limitado a, UAG, y es aminoacilado por una O-RS con un aminoácido no natural deseado. Esta O-ARNt no es reconocida por las sintetasas de aminoacilo-ARNt naturales del huésped. La mutagénesis de sitio dirigido convencional puede ser utilizada para introducir el codón de parada, incluyendo pero no limitado a, TAG, en el sitio de interés en un polipéptido de interés. Véase, por ejemplo, Sayers, JR, *et al.* (1988), '5'3' Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis. *Nucleic Acids Res.*, 16: 791-802. Cuando 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.

La incorporación de aminoácidos no naturales *in vivo* se puede realizar sin perturbación significativa de la célula huésped eucariota. Por ejemplo, debido a que la eficacia de supresión para el codón UAG depende de la competencia entre el O-ARNt, incluyendo, pero no limitado al supresor ámbar ARNt, y un factor de liberación eucariota (incluyendo pero no limitado a, eRF) (que se une a un codón de parada e inicia la liberación del péptido creciente del ribosoma), la eficacia de supresión puede modularse por, incluyendo pero no limitados a, al aumento del nivel de expresión de O-ARNt, y/o el supresor ARNt.

Los aminoácidos no naturales también pueden ser codificados con codones raros. Por ejemplo, cuando la concentración de arginina en una reacción de síntesis de proteínas *in vitro* se reduce, el codón raro de arginina, AGG, ha demostrado ser eficaz para la inserción de Ala mediante un ARNt sintético acilado con alanina. Véase, por ejemplo, Ma et al., *BioChemistry*, 32:7939 (1993). En este caso, el ARNt sintético compite con el ARNt Arg de origen natural, que existe como una especie menor en *Escherichia coli*. Algunos organismos no usan todos los codones de triplete. Un codón AGA no asignado en *Micrococcus luteus* se ha utilizado 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). Componentes de la presente divulgación se pueden generar para el uso de estos codones raros *in vivo*.

Los codones selectores también comprenden codones extendidos, incluyendo pero no limitados a, cuatro o más codones base, tal como, cuatro, cinco, seis o más codones base. Los ejemplos de codones de cuatro bases incluyen, pero no se limitan a, AGGA, CUAG, UAGA, CCGU y similares. Ejemplos de cinco codones base incluyen, pero no se limitan a, AGGAC, CCCCCU, CCCUC, CUACU, UAGGC y similares. Una característica de la divulgación incluye el uso de codones extendidos basados en la supresión del marco de lectura. Cuatro o más codones base pueden insertarse, incluyendo, pero no limitado a, uno o múltiples aminoácidos no naturales en la misma proteína. Por ejemplo, en presencia de O-ARNt mutados, incluyendo pero no limitado a ARNt supresores de cambio de marco especiales, con bucles de anticodón, por ejemplo, con al menos 8-10 nt bucles anticodón, el codón de base de cuatro o más se lee como un solo aminoácido. En otros casos, los bucles de anticodón pueden decodificar, incluyendo, pero no limitado a, al menos un codón de cuatro bases, al menos un codón de cinco bases, o al menos un codón de seis bases o más. Puesto que hay 256 posibles codones de cuatro bases, múltiples aminoácidos no naturales pueden ser codificados en la misma célula usando un codón de cuatro o más bases. Véase, Anderson et al., (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, codones de cuatro bases se han utilizado para incorporar aminoácidos no naturales en proteínas usando métodos biosintéticos *in vitro*. Véase, por ejemplo, Ma et al., (1993) *BioChemistry*, 32: 7939; y Hohsaka *et al.*, (1999) *J. Am. Chem. Soc.*, 121: 34. CGGG y AGGU fueron utilizados para incorporar simultáneamente 2-naftiloalanina y un derivado NBD de lisina en estreptavidina *in vitro* con dos ARNt supresores de desplazamiento del

marco químicamente acilados. Véase, por ejemplo, Hohsaka *et al.*, (1999) J. Am. Chem. Soc., 121: 12194. En un estudio *in vivo*, Moore *et al.* examinaron la capacidad de los derivados ARN<sup>t</sup>Leu con anticodones NCUA para suprimir codones UAGN (N puede ser U, A, G, o C), y se encontró que la UAGA cuadruplete puede ser decodificada por un ARN<sup>t</sup>Leu con un anticodón UCUA con una eficacia de 13 a 26 % con poca decodificación en la trama 0 o -1. Véase, Moore *et al.*, (2000) J. Mol. Biol. 298: 195. En una realización, los codones extendidos basados en codones raros o codones sin sentido se pueden utilizar en la presente descripción, que pueden reducir la lectura de sentido erróneo y la supresión del marco de lectura en otros sitios no deseados.

Para un sistema dado, un codón selector también puede incluir uno de los tres codones de bases naturales, donde el sistema endógeno no usa (o raramente utiliza) el codón de base natural. Por ejemplo, esto incluye un sistema que carece de un ARN<sup>t</sup> que reconoce el codón de tres bases naturales, y/o un sistema en el que el codón de tres bases es un codón raro.

Los codones selectores incluyen opcionalmente pares de bases no naturales. Estos pares de bases no naturales expanden adicionalmente el alfa-beto genético existente. Un par de bases adicional aumenta el número de codones de triplete de 64 a 125. Las propiedades de pares de bases terceros incluyen el apareamiento de bases estable y selectivo, la incorporación enzimática eficaz en ADN con alta fidelidad por una polimerasa, y la continua extensión del cebador eficiente después de la síntesis del par de bases naciente no natural. Descripciones de pares de bases no naturales que se pueden adaptar para métodos y composiciones incluyen, por ejemplo, Hirao, *et al.*, (2002) An unnatural base pair for incorporating amino acid analogues into protein, Nature Biotechnology, 20: 177.182. Véase, también, Wu, Y., *et al.*, (2002) J. Am. Chem. Soc. 124: 14626-14630. Otras publicaciones relevantes se enumeran a continuación.

Para el uso *in vivo*, el nucleósido no natural es permeable a las membranas y es fosforilado para formar el trifosfato correspondiente. Además, el aumento de la información genética es estable y no se destruye por enzimas celulares. Los esfuerzos anteriores por Benner y otros aprovecharon patrones de unión de hidrógeno que son diferentes de aquellos en los pares canónicos WatsonCrick, el ejemplo más notable de las cuales es el isoC: par isoG. Véase, por ejemplo, Switzer *et al.*, (1989) J. Am. Chem. Soc., 111: 8322; y Piccirilli *et al.*, (1990) Nature, 343: 33; Kool, (2000) Curr. Opin. Chem. Biol., 4: 602. Estas bases en general se desajustan a un cierto grado con bases naturales y no pueden ser enzimáticamente replicadas. Kool y sus colaboradores demostraron que las interacciones de embalaje hidrófobas entre bases pueden sustituir a los enlaces de hidrógeno para conducir a 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 esfuerzo para desarrollar un par de bases no naturales que satisface todos los requisitos anteriores, Schultz, Romesberg y colaboradores han sintetizado y estudiado sistemáticamente una serie de bases hidrófobas no naturales. Se ha encontrado que un par PICS:PICS es más estable que pares de bases naturales y se puede incorporar de manera eficiente en el ADN por el fragmento Klenow de polimerasa de ADN de *Escherichia coli* 1 (KF). Véase, por ejemplo, McMinn *et al.*, (1999) J. Am. Chem. Soc., 121: 11585-6; y Ogawa *et al.*, (2000) J. Am. Chem. Soc., 122: 3274. Un par 3MN:3MN puede ser sintetizado por KF con eficiencia y selectividad suficiente para la función biológica. Véase, por ejemplo, Ogawa *et al.*, (2000) J. Am. Chem. Soc., 122: 8803. Sin embargo, ambas bases actúan como un terminador de cadena para la replicación adicional. Una polimerasa de ADN mutante ha evolucionado recientemente que se puede utilizar para replicar el mismo par PICS. Además, un par 7AI puede ser replicado. Véase, por ejemplo, Tae *et al.*, (2001) J. Am. Chem. Soc., 123: 7439. Un par de metalobase novelo, Dipic:Py, también se ha desarrollado, que forma un par estable después de la unión Cu(II). Véase, Meggers *et al.*, (2000) J. Am. Chem. Soc., 122:10714. Debido a que los codones extendidos y los codones no naturales son intrínsecamente ortogonales a los codones naturales, los métodos de la descripción pueden aprovechar esta propiedad para generar ARN<sup>t</sup> ortogonales para los mismos.

Un sistema de derivación de la traducción también se puede utilizar para incorporar un aminoácido no natural en un polipéptido deseado. En un sistema de derivación de la traducción, una secuencia grande se incorpora en un gen pero no se traduce en proteína. La secuencia contiene una estructura que sirve como una señal para inducir el ribosoma para saltar sobre la secuencia y reanudar la traducción aguas abajo de la inserción.

En ciertos casos, la proteína o polipéptido de interés (o parte del mismo) en los métodos y/o composiciones de la divulgación está codificada por un ácido nucleico. Típicamente, el ácido nucleico comprende al menos un codón selector, al menos dos codones selectores, al menos tres codones selectores, al menos cuatro codones selectores, al menos cinco codones selectores, al menos seis codones selectores, al menos siete codones selectores, al menos ocho codones selectores, al menos nueve codones selectores, diez o más codones selectores.

Genes que codifican para proteínas o polipéptidos de interés se pueden mutagenizar usando métodos conocidos para un experto normal en la técnica y descritos en este documento para incluir, por ejemplo, uno o más codones selectores para la incorporación de un aminoácido no natural. Por ejemplo, un ácido nucleico para una proteína de interés se mutagenizó para incluir uno o más codones selectores, que prevé la incorporación de uno o más aminoácidos no naturales. La invención incluye cualquiera de tales variantes, incluyendo, pero no limitado a, mutante, versiones de cualquier proteína, por ejemplo, incluyendo al menos un aminoácido no natural. Del mismo modo, la descripción también incluye ácidos nucleicos correspondientes, es decir, cualquier ácido nucleico con uno o más codones selectores que codifica uno o más aminoácidos no naturales.

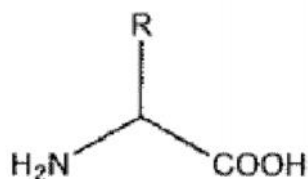
Moléculas de ácido nucleico que codifica una proteína de interés, tales como un polipéptido de insulina pueden mutarse fácilmente para introducir una cisteína en cualquier posición deseada del polipéptido. La cisteína se utiliza ampliamente para introducir moléculas reactivas, polímeros solubles en agua, proteínas, o una amplia variedad de otras moléculas, en una proteína de interés. Los métodos adecuados para la incorporación de cisteína en una posición deseada de un polipéptido son conocidos para los expertos ordinarios en la técnica, tales como los descritos en la Patente de EE.UU. nº 6.608.183, y técnicas de mutagénesis estándar.

#### **Aminoácidos no naturalmente codificados**

Una muy amplia variedad de aminoácidos no naturales codificados son adecuados para uso en la presente invención. Cualquier número de aminoácidos no naturales codificados se puede introducir en un polipéptido de insulina. En general, los aminoácidos no naturales codificados introducidos son sustancialmente químicamente inertes frente a los 20 aminoácidos comunes, genéticamente codificados (es decir, alanina, arginina, asparagina, ácido aspártico, cisteína, glutamina, ácido glutámico, glicina, histidina, isoleucina, leucina, lisina, metionina, fenilalanina, prolina, serina, treonina, triptófano, tirosina y valina). En algunas realizaciones, los aminoácidos no naturales codificados incluyen grupos funcionales de cadena lateral que reaccionan de manera eficiente y selectiva con grupos funcionales que no se encuentran en los 20 aminoácidos comunes (incluyendo, pero no limitado a, azido, cetona, aldehído y grupos aminooxi) para formar conjugados estables. Por ejemplo, un polipéptido de insulina que incluye un aminoácido no naturalmente codificado que contiene un grupo funcional azido se puede hacer reaccionar con un polímero (incluyendo pero no limitado a, poli(etilenglicol) o, alternativamente, un segundo polipéptido que contiene un resto alquino para formar un conjugado estable resultante para la reacción selectiva de la azida y los grupos funcionales de alquino para formar un producto de cicloadición de Huisgen [3 + 2].

La estructura genérica de un ácido de alfa-amino se ilustra del siguiente modo (Fórmula I):

I



Un aminoácido no naturalmente codificado es típicamente cualquier estructura que tiene la fórmula anteriormente enumerada en donde el grupo R es cualquier sustituyente distinto de uno utilizado en los veinte aminoácidos naturales, y puede ser adecuado para uso en la presente invención. Debido a que los aminoácidos no naturales codificados de la invención típicamente difieren de los aminoácidos naturales sólo en la estructura de la cadena lateral, los aminoácidos no naturales codificados forman enlaces de amida con otros aminoácidos, incluyendo pero no limitado a, naturalmente o no naturalmente codificados, en donde se forman en polipéptidos de origen natural.

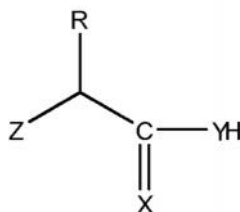
Sin embargo, los aminoácidos no naturales codificados tienen grupos de cadena lateral que los distinguen de los aminoácidos naturales. Por ejemplo, R comprende opcionalmente un alquilo, arilo, acilo, ceto, azido, hidroxilo, hidracina, ciano, halo, hidrazida, alqueniilo, alquiloni, éter, tiol, seleno, sulfonilo, borato, boronato, fosfo, fosfono, fosfina, heterociclo, enona, imina, aldehído, éster, tioácido, hidroxilamina, grupo amino, o similares, o cualquier combinación de los mismos. Otros aminoácidos no naturales de interés que pueden ser adecuados para su uso en la presente invención incluyen, pero no se limitan a, aminoácidos que comprenden un agente de reticulación fotoactivable, aminoácidos marcados por espín, aminoácidos fluorescentes, aminoácidos de unión a metal, aminoácidos que contienen metal, aminoácidos radiactivos, aminoácidos con nuevos grupos funcionales, aminoácidos que covalentemente o no covalentemente interactúan con otras moléculas, y/o aminoácidos fotoisomerizables, aminoácidos que comprenden biotina o un análogo de biotina, aminoácidos glicosilados tales como serina sustituida por azúcar, otros aminoácidos modificados por hidratos de carbono, aminoácidos que contienen ceto, aminoácidos que comprenden glicol de polietileno o poliéter, aminoácidos sustituidos por el átomo pesado, aminoácidos químicamente escindibles y/o fotoescindibles, aminoácidos con cadenas laterales alargadas en comparación con aminoácidos naturales, incluyendo, pero no limitado a, poliéteres o hidrocarburos de cadena larga, incluyendo pero no limitados a, mayor de aproximadamente 5 o mayor de aproximadamente 10 carbonos, aminoácidos de enlace de carbono que contienen azúcar, aminoácidos redoxactivos, tioácido de amino que contiene aminoácidos, y aminoácidos que comprenden uno o más restos tóxicos.

Aminoácidos no naturalmente codificados ejemplares que pueden ser adecuados para su uso en la presente invención y que son útiles para reacciones con polímeros solubles en agua incluyen, pero no se limitan a, aquellos con grupos reactivos de carbonilo, aminooxi, hidrazina, hidrazida, semicarbazida, azida y alquino. En algunas

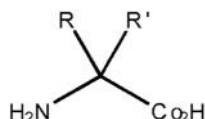
realizaciones, los aminoácidos no naturales codificados comprenden un resto de sacárido. Ejemplos de tales aminoácidos incluyen *N*-acetilo-L-glucosaminilo-L-serina, *N*-acetilo-L-galactosaminilo-L-serina, *N*-acetilo-L-glucosaminilo-L-treonina, *N*-acetilo-L-glucosaminilo-L-asparagina y *O*-mannosaminilo-L-serina. Ejemplos de tales aminoácidos incluyen también ejemplos en los que enlaces N o O naturales entre el aminoácido y el sacárido se sustituyen por un enlace covalente no se encuentran comúnmente en la naturaleza incluyendo, pero no limitado a, un alqueno, una oxima, un tioéter, una amida y similares. Ejemplos de tales aminoácidos también incluyen sacáridos que no se encuentran comúnmente en las proteínas naturales tales como 2-deoxiglucosa, 2-deoxigalactosa y similares.

Muchos de los aminoácidos no naturales codificados proporcionados en este documento están disponibles comercialmente, por ejemplo, de Sigma-Aldrich (St. Louis, MO, EE.UU.), Novabiochem (una división de EMD Biosciences, Darmstadt, Alemania), o Peptech (Burlington, MA, EE.UU.). Los que no están disponibles comercialmente se sintetizan opcionalmente como se proporciona aquí o se utilizan métodos estándar conocidos por personas de experiencia ordinaria en la técnica. Para técnicas de síntesis orgánica, véase, por ejemplo, Organic Chemistry de Fessenden y Fessenden, (1982, segunda edición, Willard Grant Press, Boston Mass.); Advanced Organic Chemistry por March (Tercera Edición, 1985, Wiley and Sons, Nueva York); y Advanced Organic Chemistry por Carey y Sundberg (Tercera edición, partes A y B, 1990, Plenum Press, Nueva York). Véase, también, la Patente de Estados Unidos Nos. 7.045.337 y 7.083.970. Además de aminoácidos no naturales que contienen nuevas cadenas laterales, los aminoácidos no naturales que pueden ser adecuados para su uso en la presente invención también comprenden opcionalmente estructuras de cadena principal modificadas, incluyendo, pero no limitado a, como se ilustra por las estructuras de la Fórmula II y III:

II



III



en la que Z comprende típicamente OH, NH<sub>2</sub>, SH, NH-R' o S-R'; X e Y, que pueden ser iguales o diferentes, comprenden típicamente S o O, y R y R', que son opcionalmente iguales o diferentes, se seleccionan típicamente de la misma lista de constituyentes para el grupo R descrito anteriormente para aminoácidos antinaturales que tienen la Fórmula I, así como hidrógeno. Por ejemplo, aminoácidos no naturales de la invención comprenden opcionalmente sustituciones en el grupo amino o carboxilo como se ilustra por las Fórmulas II y III. Los aminoácidos no naturales de este tipo incluyen, pero no se limitan a, ácidos α-hidroxi, α-tioácidos, α-aminotiocarboxilatos, incluyendo pero no limitados a, con cadenas laterales correspondientes a los veinte aminoácidos naturales comunes o cadenas laterales no naturales. Además, las sustituciones en el α-carbono opcionalmente incluyen, pero no se limitan a, aminoácidos L, D, o α-α-disustituidos tales como D-glutamato, D-alanina, D-metil tirosina, ácido aminobutírico, y similares. Otras alternativas estructurales incluyen ácidos cíclicos aminoácidos, tales como análogos de prolina así como análogos de prolina de anillo de 3, 4, 6, 7, 8, y 9 miembros, aminoácidos β e γ tales como β-alanina sustituida y ácido γ-aminobutírico.

Muchos aminoácidos no naturales se basan en aminoácidos naturales, tales como tirosina, glutamina, fenilo-alanina, y similares, y son adecuados para uso en la presente invención. Análogos de tirosina incluyen, pero no se limitan a, tirosinas parasustituidas, tirosinas ortosustituidas, y tirosinas meta sustituidas, donde la tirosina sustituida comprende, incluyendo, pero no limitado a, un grupo ceto (incluyendo pero no limitado a, un grupo acetilo), un grupo de benzoílo, un grupo amino, una hidracina, una hidroxiamina, un grupo tiol, un grupo carboxi, un grupo isopropilo, un grupo metilo, una cadena lineal C 6 - C20 o hidrocarburo ramificado, un hidrocarburo saturado o insaturado, un grupo O-metilo, un grupo poliéter, un grupo nitro, un grupo alquililo o similares. Además, también se contemplan anillos de arilo polisustituidos. Análogos de glutamina que pueden ser adecuados para su uso en la presente

invención incluyen, pero no se limitan a, derivados  $\alpha$ -hidroxi, derivados  $\gamma$ -sustituidos, derivados cíclicos y derivados de glutamina amida sustituidos. Análogos de fenilo-alanina ejemplares que pueden ser adecuados para su uso en la presente invención incluyen, pero no se limitan a, fenilo-alaninas parasustituidas, feniloalaninas ortosustituidas, y fenilo-alaninas metasustituidas, donde el sustituyente comprende, incluyendo, pero no limitado a, un grupo hidroxilo, un grupo metoxi, un grupo metilo, un grupo alilo, un aldehído, un azido, un yodo, un bromo, un grupo ceto (incluyendo pero no limitado a, un grupo acetilo), un benzoílo, un grupo alquinilo, o similares. Los ejemplos específicos de aminoácidos no naturales que pueden ser adecuados para su uso en la presente invención incluyen, pero no se limitan a, una *p*-acetilo-L-fenilo-alanina, un *O*-metilo-L-tirosina, una L3-(2-naftilo)alanina, una 3-metilo-fenilalanina, un *O*-4-alilo-L-tirosina, 4-propilo- L-tirosina, un tri-*O*-acetilo-GlcNAc  $\beta$ -serina, un L-Dopa, una fenilo-alanina fluorada, una isopropilo- L-feniloalanina, una *p*-azido-L- fenilo-alanina, un *p*-acilo-L-feniloalanina, una *p*-benzoílo-L- fenilo-alanina, una-L-fosfoserina, una fosfonoserina, una fosfonotirosina, un *p*-yodo-feniloalanina, una *p*-bromofeniloalanina, una *p*-amino-L-feniloalanina, una isopropilo- L-feniloalanina y un *p*-propargiloxi-feniloalanina, y similares. Ejemplos de estructuras de una variedad de aminoácidos no naturales que pueden ser adecuados para uso en la presente invención se proporcionan en, por ejemplo, el documento WO 2002/085923 titulado "*In vivo* incorporation of unnatural amino acids." Véase también Kiick et al., (2002) Incorporation of azidas into recombinant proteins for chemoselective modification by the Staudinger ligation, PNAS 99: 19-24, para análogos de metionina adicional. Solicitud Internacional n° PCT/U506/47822 titulada "Compositions Containing, Methods Involving and Uses of Non-natural Amino Acids and Polypeptides", describe la alquilación reductora de restos de amina aromáticos, incluyendo, pero no limitado a, *p*-amino-feniloalanina y animación reductora.

En una realización, se proporcionan composiciones de polipéptido de insulina que incluyen un aminoácido no natural (tales como *p*-(propargiloxi)-fenilo-alanina). Varias composiciones que comprenden *p*-(propargiloxi)-fenilo-alanina) y, incluyendo, pero no limitado a, proteínas y/o células, también se proporcionan. En un aspecto, una composición que incluye el aminoácidos antinaturales de *p*-(propargiloxi)-fenilalanina, incluye además un ARNt ortogonal. El aminoácido no natural puede ser unido (incluyendo pero no limitado a, covalentemente) a la ARNt ortogonal, incluyendo, pero no limitado a, unido covalentemente a la ARNt ortogonal aunque un enlace de amino-acilo, unido covalentemente a 3'OH o 2'OH de un azúcar de ribosa terminal de ARNt ortogonal, etc.

Los restos químicos a través de aminoácidos no naturales que se pueden incorporar en las proteínas ofrecen una variedad de ventajas y manipulaciones de la proteína. Por ejemplo, la reactividad única de un grupo funcional ceto permite la modificación selectiva de proteínas con cualquiera de un número de reactivos que contienen hidrazina o hidroxilamina *in vitro* e *in vivo*. Un aminoácido no natural de átomo pesado, por ejemplo, puede ser útil para la eliminación de datos de la estructura de rayos X. La introducción de sitio específico de átomos pesados utilizando aminoácidos no naturales también proporciona selectividad y la flexibilidad en la elección de las posiciones de los átomos pesados. Aminoácidos fotorreactivos no naturales (incluyendo pero no limitado a, aminoácidos con benzofenona y arilazidas (incluyendo pero no limitado a, fenilazida) cadenas laterales), por ejemplo, permitir fotorreticulación de proteína eficiente *in vivo* e *in vitro*. Ejemplos de aminoácidos no naturales fotorreactivos incluyen, pero no se limitan a, pazidofeniloalanina y *p*-benzoílo-feniloalanina. La proteína con los aminoácidos no naturales fotorreactivos a continuación se pueden reticular a voluntad por excitación del control temporal que proporciona el grupo fotorreactivo. En un ejemplo, el grupo metilo de un aminoácido no natural puede estar sustituido con un grupo metilo marcado isotópicamente, como una sonda de la estructura local y la dinámica, incluyendo pero no limitado a, con el uso de resonancia magnética nuclear y la espectroscopia vibracional. Grupos funcionales de alquinilo o azido, por ejemplo, permiten la modificación selectiva de proteínas con moléculas a través de una reacción de cicloadición [3 + 2].

Un aminoácido no natural incorporado en un polipéptido en el extremo amino puede estar compuesto de un grupo R que es cualquier sustituyente distinto de uno utilizado en los veinte aminoácidos naturales y un segundo grupo reactivo diferente entre el grupo NH<sub>2</sub> normalmente presente en  $\alpha$ -aminoácidos (véase la fórmula I). Un aminoácido no natural similar puede incorporarse en el terminal de carboxilo con un segundo grupo reactivo diferente del grupo COOH normalmente presente en aminoácidos (véase la Fórmula I).

Los aminoácidos no naturales de la invención pueden ser seleccionados o diseñados para proporcionar características adicionales no disponibles en los veinte aminoácidos naturales. Por ejemplo, aminoácido no natural puede estar diseñado o seleccionado para modificar las propiedades biológicas de una proteína, por ejemplo, en donde se incorporan opcionalmente. Por ejemplo, las siguientes propiedades pueden ser opcionalmente modificadas por inclusión de un aminoácido no natural en una proteína: toxicidad, biodistribución, solubilidad, estabilidad, por ejemplo, térmica, hidrolítica, oxidativa, resistencia a la degradación enzimática, y similares, facilidad de purificación y procesamiento, propiedades estructurales, propiedades espectroscópicas, química y/o propiedades fotoquímicas, actividad catalítica, potencial redox, vida media, capacidad para reaccionar con otras moléculas, por ejemplo, de forma covalente o no covalente, y similares.

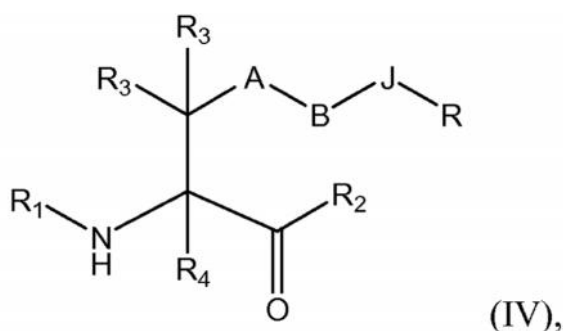
#### **ESTRUCTURA Y SÍNTESIS DE AMINOÁCIDOS NO NATURALES: CARBONILO, DE TIPO CARBONILO, CARBONILO ENMASCARADO, GRUPOS DE CARBONILO PROTEGIDOS, Y GRUPOS DE HIDROXILAMINA**

En la presente invención la insulina está unida a un polímero soluble en agua, por ejemplo, un PEG, por un enlace de oxima.

Muchos tipos de aminoácidos no naturales codificados son adecuados para la formación de enlaces de oxima. Estos incluyen, pero no se limitan a, aminoácidos no naturales codificados que contienen un carbonilo, dicarbonilo, o grupo de hidroxilamina. Tales aminoácidos se describen en la Publicación de Patente de Estados Unidos Nos. 2006/0194256, 2006/0217532, y 2006/0217289 y WO 2006/069246 con el título "Compositions containing, methods involving, and uses of non-natural amino acids and polypéptidos". Aminoácidos no naturalmente codificados también se describen en la Patente de Estados Unidos nº 7.083.970 y la Patente de Estados Unidos nº 7.045.337.

Algunas realizaciones de la invención utilizan polipéptidos de insulina que están sustituidos en una o más posiciones con un aminoácido para-acetilfeniloalanina. La síntesis de p-acetilo-(+/-)-fenilalanina y m-acetilo-(+/-)-fenilo-alanina se describen en Zhang, Z., et al., *Biochemistry* 42: 6735-6746 (2003). Otros aminoácidos que contienen carbonilo o dicarbonilo se pueden preparar de manera similar por una persona de experiencia ordinaria en la técnica. Además, síntesis no limitativa ejemplaria de aminoácido no natural que se incluyen en el presente documento se presentan en las FIGS. 4, 24-34 y 36-39 de la Patente de EE.UU. nº 7.083.970.

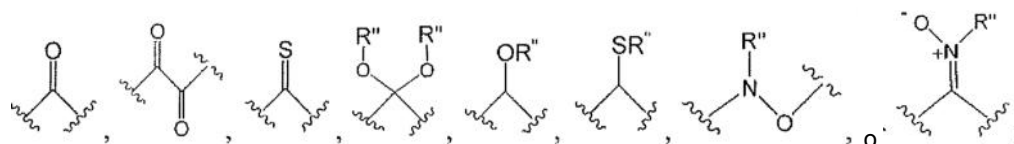
Los aminoácidos con un grupo reactivo electrofílico permiten una variedad de reacciones para unir moléculas a través de reacciones de adición nucleófila entre otros. Tales grupos reactivos electrofílicos incluyen un grupo carbonilo (incluido un grupo ceto y un grupo dicarbonilo), un grupo de tipo carbonilo (que tiene una reactividad similar a un grupo carbonilo (incluido un grupo ceto y un grupo dicarbonilo) y es estructuralmente similar a un grupo carbonilo), un grupo carbonilo enmascarado (que se puede convertir fácilmente en un grupo carbonilo (incluido un grupo ceto y un grupo dicarbonilo)), o un grupo carbonilo protegido (que tiene una reactividad similar a un grupo carbonilo (incluido un grupo ceto y un grupo dicarbonilo) tras la desprotección). Tales aminoácidos incluyen aminoácidos que tienen la estructura de Fórmula (IV):



donde:

A es opcional, y cuando está presente es alquileo inferior, alquileo inferior sustituido, cicloalquileo inferior, cicloalquileo inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, alquinileno, heteroalquileo inferior, heteroalquileo sustituido, heterocicloalquileo inferior, heterocicloalquileo inferior sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno sustituido, aralquileo o aralquileo sustituido;

B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileo inferior, alquileo inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileo, heteroalquileo inferior sustituido, -O-, -O-(alquileo o alquileo sustituido)-, -S-, -S-(alquileo o alquileo sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>(alquileo o alquileo sustituido)-, -C(O)-, -C(O)-(alquileo o alquileo sustituido)-, -C(S)-, -C(S)-(alquileo o alquileo sustituido)-, -N(R')-, -NR'-(alquileo o alquileo sustituido)-, -C(O)N(R')-, -CON(R')-(alquileo o alquileo sustituido)-, -CSN(R')-, -CSN(R')-(alquileo o alquileo sustituido)-, -N(R')CO-(alquileo o alquileo sustituido)-, -N(R')C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N-, y -C(R')<sub>2</sub>-N(R')-N(R')-, donde cada R' es independientemente H, alquilo, o alquilo sustituido;



J es

R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;  
 cada R" es independientemente H, alquilo, alquilo sustituido, o un grupo protector, o cuando más de un grupo R"  
 está presente, dos R" opcionalmente forman un heterocicloalquilo;

5 R<sub>1</sub> es opcional y, cuando está presente, es H, un protector de amino grupo, resina, aminoácido, polipéptido o  
 polinucleótido; y

R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o  
 polinucleótido;

10 cada uno de R<sub>3</sub> y R<sub>4</sub> es independientemente H, halógeno, alquilo inferior, o alquilo inferior sustituido, o dos grupos  
 R<sub>3</sub> R<sub>3</sub> y R<sub>4</sub> o forman opcionalmente un cicloalquilo o un heterocicloalquilo; o los grupos -A-B-J-R juntos forman un  
 cicloalquilo o heterocicloalquilo bicíclico o tricíclico que comprende al menos un grupo carbonilo, incluido un grupo  
 dicarbonilo, grupo carbonilo protegido, incluyendo un grupo protegido de dicarbonilo, o grupo carbonilo  
 enmascarado, incluido un grupo de dicarbonilo enmascarado;

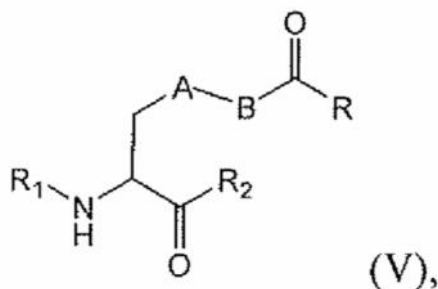
15 o el grupo -J-R juntos forman un anillo cicloalquilo o heterocicloalquilo monocíclico o bicíclico que comprende al  
 menos un grupo carbonilo, incluido un grupo dicarbonilo, grupo carbonilo protegido, incluido un grupo dicarbonilo  
 protegido, o un grupo carbonilo enmascarado, incluido un grupo dicarbonilo enmascarado;

con la condición de que cuando A es fenileno y cada R<sub>3</sub> es H, B está presente; y que cuando A es -(CH<sub>2</sub>)<sub>4</sub> y cada  
 R<sub>3</sub> es H, B no es -NHC(O)(CH<sub>2</sub>CH<sub>2</sub>); y que cuando A y B están ausentes y cada R<sub>3</sub> es H, R no es metilo.

[235] Además, teniendo la estructura de Fórmula (V), se incluyen:

20

25



30

en donde:

35

A es opcional, y cuando está presente es alquileno inferior, alquileno inferior sustituido, cicloalquiloene inferior,  
 cicloalquiloene inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, alquilileno, heteroalquileno  
 inferior, heteroalquileno sustituido, heterocicloalquileno inferior, heterocicloalquileno inferior sustituido, arileno,  
 arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno, aralquileno o aralquileno sustituido;

40

B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileno inferior,  
 alquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileno, heteroalquileno  
 inferior sustituido, -O-, -O-(alquileno o alquileno sustituido), -S-, -S-(alquileno o alquileno sustituido), -S(O)<sub>k</sub>- donde k  
 es 1, 2, o 3, -S(O)<sub>k</sub>(alquileno o alquileno sustituido), -C(O)-, -C(O)-(alquileno o alquileno sustituido)-, -C(S)-, -C(S)-  
 (alquileno o alquileno sustituido)-, -N(R')-, -NR'-(alquileno o alquileno sustituido)-, -C(O)N(R')-, -CON(R')-(alquileno o  
 alquileno sustituido), -CSN(R')-, -CSN(R')-(alquileno o alquileno sustituido)-, -N(R')CO-(alquileno o alquileno  
 sustituido)-, -N(R')C(O)O-, S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R'), -N(R')=N-, -C(R')=N-,  
 -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N-, y -C(R')<sub>2</sub>-N(R')-N(R'), donde cada R' es independientemente H, alquilo,  
 o alquilo sustituido; R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

45

50

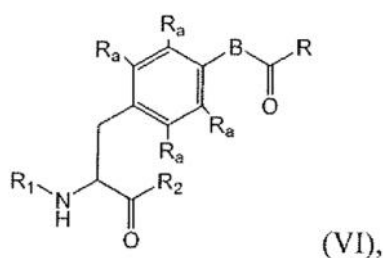
R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo amino, resina, aminoácido, polipéptido o polinucleótido; y  
 R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo éster, resina, aminoácido, polipéptido o polinucleótido  
 protector;

con la condición de que cuando A es fenileno, B está presente; y que cuando A es (CH<sub>2</sub>)<sub>4</sub>, B no es -NHC(O)(CH<sub>2</sub>-  
 CH<sub>2</sub>)-; y de que cuando A y B están ausentes, R no es metilo.

55

Además, los aminoácidos que tienen la estructura de Fórmula (VI) se incluyen:

60



65



en donde:

5 B es un enlazador seleccionado del grupo que consiste en alquileo inferior, alquileo inferior sustituido, alquenileno, alquenileno inferior sustituido, heteroalquileo inferior, -O-, -O-(alquileo o alquileo sustituido)-, -S-, -S-(alquileo o alquileo sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>(alquileo o alquileo sustituido)-, -C(O)-, -C(O)-(alquileo o alquileo sustituido)-, -C(S)-, -C(S)-(alquileo o alquileo sustituido)-, -N(R')-, -NR'-(alquileo o alquileo sustituido)-, -C(O)N(R')-, -CON(R')-(alquileo o alquileo sustituido)-, -CSN(R')-, -CSN(R')-(alquileo o alquileo sustituido)-, -N(R')CO-(alquileo o alquileo sustituido)-, -N(R')C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, -N(R')-N=, -C(R')=N-, -C(R')=NN(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N-, y -C(R')<sub>2</sub>N(R')-N(R')-, donde cada R' es independientemente H, alquilo, o alquilo sustituido.

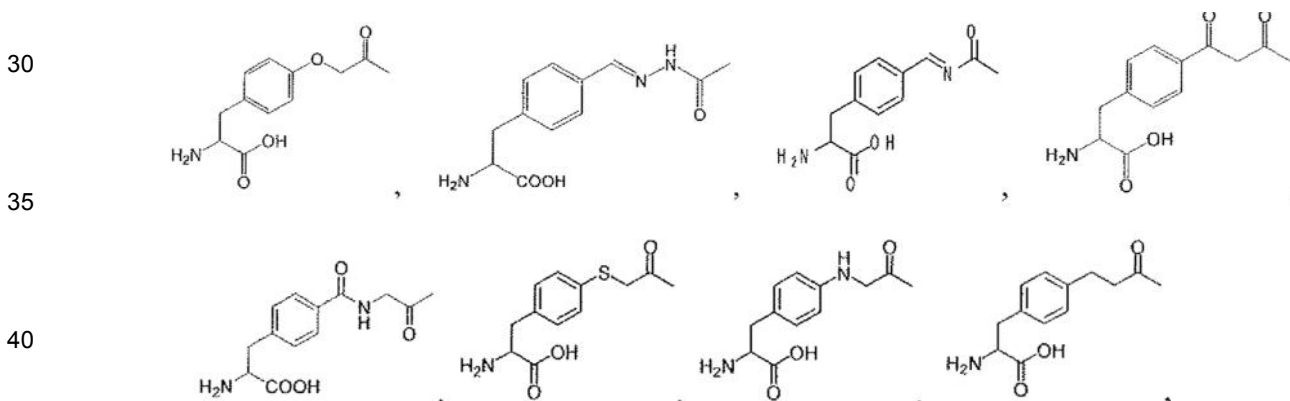
15 R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

20 R<sup>2</sup> es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

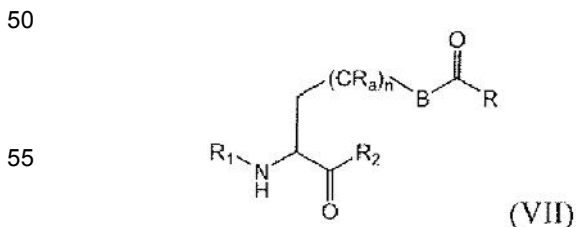
25 cada R<sub>a</sub> se selecciona independientemente del grupo que consiste en H, halógeno, alquilo, alquilo sustituido, -N(R')<sub>2</sub>, -C(O)<sub>k</sub>R', donde k es 1, 2, o 3, -C(O)N(R')<sub>2</sub>, -OR' y -S(O)<sub>k</sub>R', donde cada R' es independientemente H, alquilo o alquilo sustituido.

Además, se incluyen los siguientes aminoácidos:



45 compuestos son un grupo opcionalmente de amino protegido, carboxilo protegido o una sal protegida del mismo. Además, cualquiera de los siguientes aminoácidos no naturales pueden incorporarse en un polipéptido de aminoácidos no naturales.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (VII):



60 donde

65 B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileo inferior, alquileo inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileo inferior, -O-, -O-(alquileo o alquileo sustituido)-, -S-, -S-(alquileo o alquileo sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>(alquileo o alquileo sustituido)-, -C(O)-, -C(O)-(alquileo o alquileo sustituido)-, -C(S)-, -C(S)-(alquileo o alquileo sustituido)-, -N(R')-, -NR'-(alquileo o alquileo sustituido)-, -

C(O)N(R')-, -CON(R')-(alquileo o alquileo sustituido)-, -CSN(R')-, -CSN(R')-(alquileo o alquileo sustituido)-, -N(R')CO-(alquileo o alquileo sustituido)-, -N(R')-C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N-, y -C(R')<sub>2</sub>-N(R')-N(R')-, donde cada R' es independientemente H, alquilo, o alquilo sustituido;

5 R es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

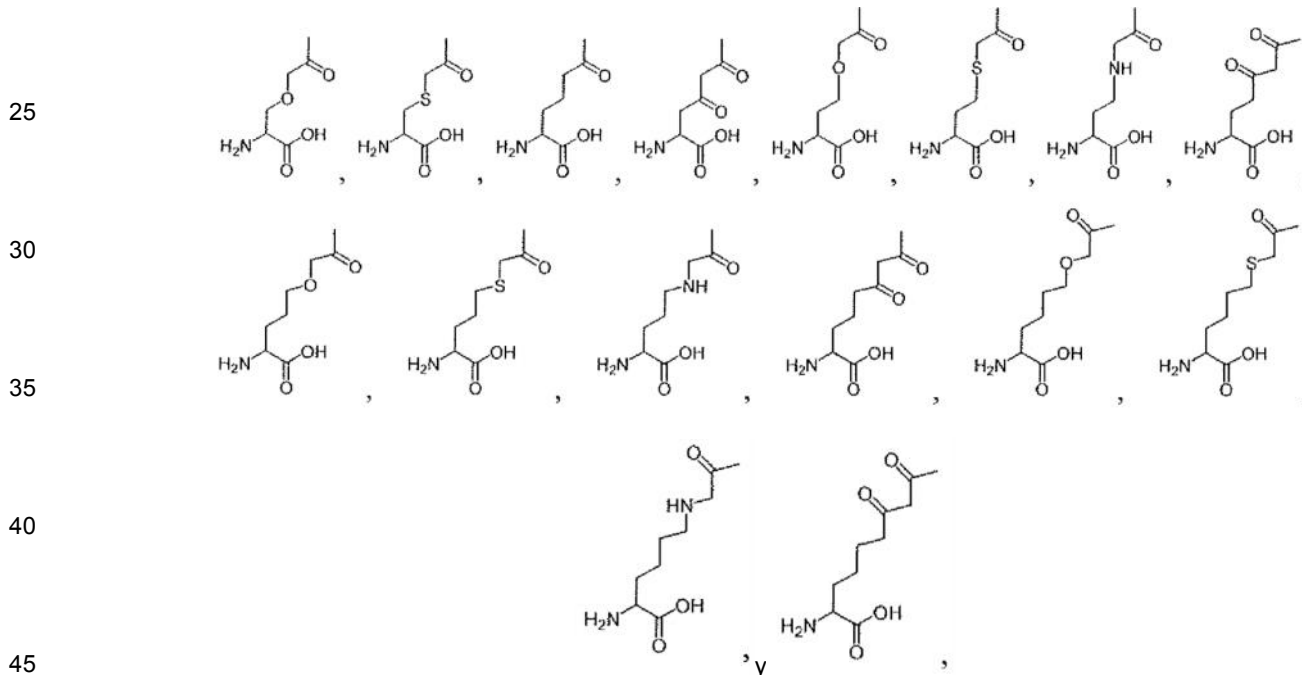
R<sub>1</sub> es opcional, y cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

10 R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo éster, resina, aminoácido, polipéptido o polinucleótido;

15 cada R<sub>a</sub> se selecciona independientemente del grupo que consiste en H, halógeno, alquilo, alquilo sustituido, -N(R')<sub>2</sub>, -C(O)<sub>k</sub>R', donde k es 1, 2, o 3, -C(O)N(R')<sub>2</sub>, -OR' y -S(O)<sub>k</sub>R', donde cada R<sub>1</sub> es independientemente H, alquilo, o alquilo sustituido; y n es de 0 a 8; con la condición de que cuando A es -(CH<sub>2</sub>)<sub>4</sub>-, B no es -NHC(O)(CH<sub>2</sub>-CH<sub>2</sub>)-,

Además, se incluyen los siguientes aminoácidos:

20

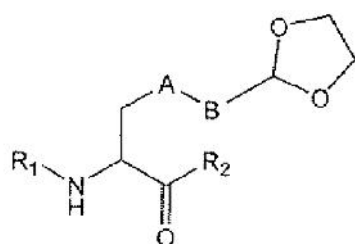


50 en donde dichos compuestos están opcionalmente amino protegidos, carboxilo opcionalmente protegido, amino opcionalmente protegido y carboxilo protegido, o una sal del mismo. Además, estos aminoácidos no naturales y cualquiera de los siguientes aminoácidos no naturales se pueden incorporar en un polipéptido de aminoácido de amino no natural.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (VIII):

55

60



(VIII),

65

en donde A es opcional, y cuando está presente es alquileo inferior, alquileo inferior sustituido, cicloalquileo,

cicloalquileno inferior sustituido, alquenileno, alquenileno inferior sustituido, alquinileno, heteroalquileno inferior, heteroalquileno sustituido, heterocicloalquileno inferior, heterocicloalquilona inferior sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno sustituido, aralquileno o aralquileno sustituido;

5 B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileno inferior, alquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileno, heteroalquileno inferior sustituido, -O-, -O-(alquileno o alquileno sustituido)-, -S-, -S-(alquileno o alquileno sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>- (alquileno o alquileno sustituido)-, -C(O)-, -C(O)-(alquileno o alquileno sustituido)-, -C(S)-, -C(S)-(alquileno o alquileno sustituido)-, -N(R')-, -NR'-(alquileno o alquileno sustituido)-, -C(O)N(R')-, -CON(R')(alquileno o alquileno sustituido)-, -CSN(R')-, -CSN(R')-(alquileno o alquileno sustituido)-, -N(R')CO-(alquileno o alquileno sustituido)-, -N(R')C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, -N(R')=N, -C(R')=N, -C(R')=NN(R')-, -C(R')=N=N-, -C(R')<sub>2</sub>-N=N-, y -C(R')<sub>2</sub>-N(R')-N(R')-, donde cada R' es independientemente H, alquilo, o alquilo sustituido;

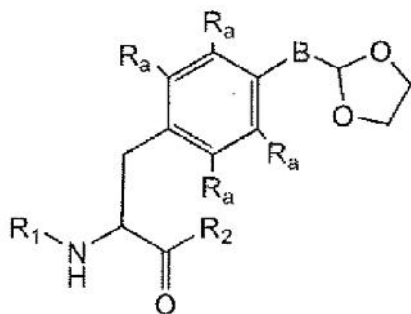
15 R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido.

20

Además, los siguientes aminoácidos que tienen la estructura de Fórmula (IX) están incluidos:

25



30

(IX),

35

B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileno inferior, alquileno inferior sustituido, alquenileno, alquenileno inferior sustituido, heteroalquileno, heteroalquileno inferior sustituido, -O-, -O-(alquileno o alquileno sustituido)-, -S-, -S-(alquileno o alquileno sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>(alquileno o alquileno sustituido)-, -C(O)-, -C(O)-(alquileno o alquileno sustituido)-, -C(S)-, -C(S)-(alquileno o alquileno sustituido)-, -N(R')-, -NR'-(alquileno o alquileno sustituido)-, -C(O)N(R')-, -CON(R')-(alquileno o alquileno sustituido)-, -CSN(R')-, -CSN(R')-(alquileno o alquileno sustituido)-, -N(R')CO-(alquileno o alquileno sustituido)-, -N(R')C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, N(R')S(O)<sub>k</sub>N(R')-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N=N-, -C(R')<sub>2</sub>-N=N-, y -C(R')<sub>2</sub>-N(R')-N(R')-, donde cada R<sub>1</sub> es independientemente H, alquilo, o alquilo sustituido;

R es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

50 R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido, o polinucleótido; y

R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido;

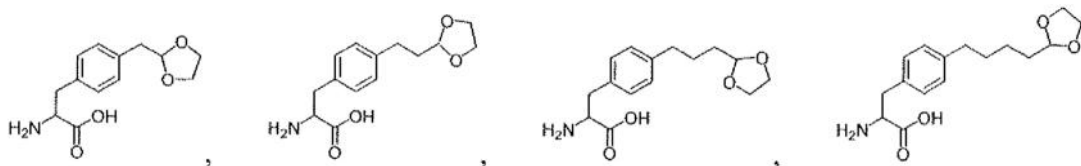
55 en la que cada R<sub>a</sub> se selecciona independientemente del grupo que consiste en H, halógeno, alquilo, alquilo sustituido, -N(R')<sub>2</sub>, -C(O)<sub>k</sub>R', donde k es 1, 2, o 3, -C(O)N(R')<sub>2</sub>, OR', y -S(O)<sub>k</sub>R' donde cada R<sub>1</sub> es independientemente H, alquilo, o alquilo sustituido.

Además, se incluyen los siguientes aminoácidos:

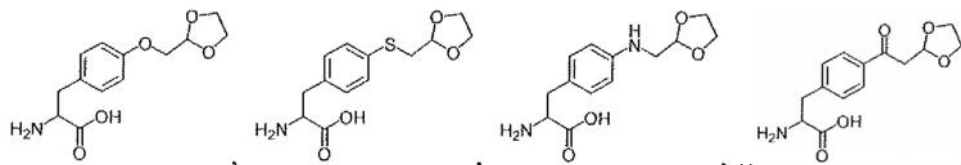
60

65

5



10

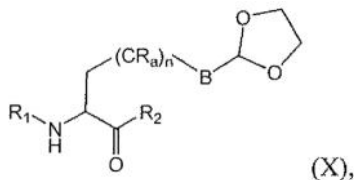


15

tales compuestos están opcionalmente amino protegidos, opcionalmente carboxilo protegidos, opcionalmente amino protegidos y carboxilo protegidos, o una sal de los mismos. Además, estos aminoácidos no naturales y cualquiera de los siguientes aminoácidos no naturales pueden incorporarse en un polipéptido de aminoácidos no naturales.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (X):

25



30

en donde B es opcional, y cuando está presente un enlazador seleccionado del grupo que consiste en alquileo inferior, alquileo inferior sustituido, alquilenilo inferior, alquilenilo inferior sustituido, heteroalquileo inferior, heteroalquileo inferior sustituido, -O-, -O-(alquileo o alquileo sustituido)-, -S-, -S-(alquileo o alquileo sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, S(O)<sub>k</sub>(alquileo o alquileo sustituido)-, -C(O)-, -C(O)-(alquileo o alquileo sustituido)-, -C(S)-, -C(S)-(alquileo o alquileo sustituido)-, -N(R')-, -NR'-(alquileo o alquileo sustituido)-, -C(O)N(R')-, -CON(R')-(alquileo o alquileo sustituido)-, -CSN(R')-, -CSN(R')-(alquileo o alquileo sustituido)-, -N(R')CO-(alquileo o alquileo sustituido)-, -N(R')C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, -N(R')-N=, -C(R')=N, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N, y -C(R')<sub>2</sub>-N(R')-N(R')-, donde cada R' es, independientemente, H, alquilo, o alquilo sustituido;

R es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

R1 es opcional, y cuando está presente, es H, un grupo amino, resina, aminoácido, polipéptido, o polinucleótido protector; y

45

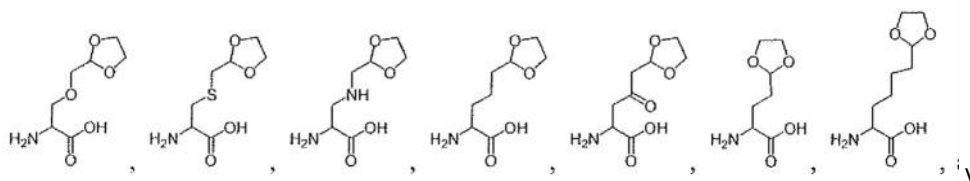
R2 es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido, o polinucleótido;

cada Ra se selecciona independientemente del grupo que consiste de H, halógeno, alquilo, alquilo sustituido, -N(R')<sub>2</sub>, -C(O)kR', donde k es 1, 2, o 3, -C(O)N(R')<sub>2</sub>, -OR' y -S(O)kR', donde cada R' es, independientemente, H, alquilo, o alquilo sustituido; y n es de 0 a 8.

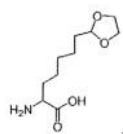
50

Además, se incluyen los siguientes aminoácidos:

55



60

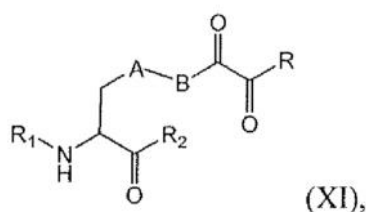


65

en donde dichos compuestos están opcionalmente amino protegidos, opcionalmente carboxilo protegidos, opcionalmente amino protegidos y carboxilo protegidos, o una sal de los mismos. Además, estos aminoácidos no naturales y cualesquiera de los siguientes aminoácidos no naturales pueden incorporarse en un polipéptido de aminoácidos no naturales.

Además de las estructuras de monocarbonilo, los aminoácidos no naturales descritos en el presente documento pueden incluir grupos tales como dicarbonilo, tipo dicarbonilo, dicarbonilo enmascarado y grupos protegidos de dicarbonilo.

Por ejemplo, se incluyen los siguientes aminoácidos que tienen la estructura de fórmula (XI):



en donde A 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;

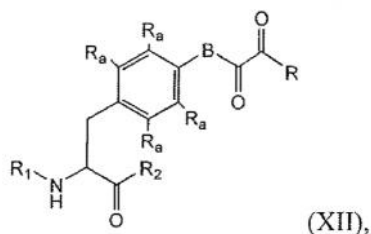
B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileno inferior, alquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileno, heteroalquileno inferior sustituido, -O-, -O-(alquileno o alquileno sustituido)-, -S-, -S-(alquileno o alquileno sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>-(alquileno o alquileno sustituido)-, C(=O), -C(O)-(alquileno o alquileno sustituido)-, -C(S)-, -C(S)-(alquileno o alquileno sustituido)-, -N(R')-, -NR'-(alquileno o alquileno sustituido)-, -C(O)N(R')-, -CON(R')-(alquileno o alquileno sustituido)-, -CSN(R')-, -CSN(R')-(alquileno o alquileno sustituido)-, -N(R')CO-(alquileno o alquileno sustituido)-, N(R')-C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N-, y C(R')<sub>2</sub> N(R') N(R')-, en donde cada R' es independientemente H, alquilo, o alquilo sustituido;

R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

R1 es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido, o polinucleótido; y

R2 es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (XII):



B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileno inferior, alquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileno inferior, heteroalquileno inferior sustituido, -O-, -O-(alquileno o alquileno sustituido)-, -S-, -S-(alquileno o alquileno sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>(alquileno o alquileno sustituido), -C(O)-, -C(O)-(alquileno o alquileno sustituido)-

, -C(S)-, -C(S)-(alquileo o alquileo sustituido)-, -N(R')-, -NR'-(alquileo o alquileo sustituido), -C(O)N(R')-, C-O-N(R')-(alquileo o alquileo sustituido), -CSN(R')-, -CSN(R')-(alquileo o alquileo sustituido), -N(R')CO-(alquileo o alquileo sustituido)-, -N(R')-C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N-, y -C(R')<sub>2</sub>-N(R')-N(R')-, donde cada R' es independientemente H, alquilo, o alquilo sustituido;

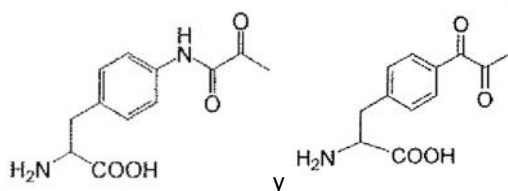
R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

R1 es opcional, y cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido, o polinucleótido; y

R2 es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

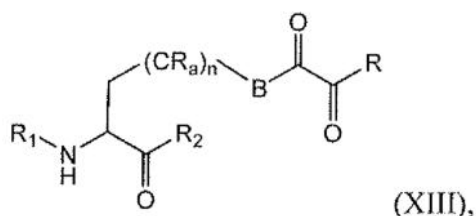
en la que cada R<sub>a</sub> se selecciona independientemente del grupo que consiste en H, halógeno, alquilo, alquilo sustituido, -N(R')<sub>2</sub>, -C(O)kR', donde k es 1, 2, o 3, -C(O)N(R')<sub>2</sub>, OR', y -S(O)kR', donde cada R' es independientemente H, alquilo o alquilo sustituido.

Además, se incluyen los siguientes aminoácidos:



en donde dichos compuestos están opcionalmente amino protegidos, opcionalmente carboxilo protegidos, opcionalmente amino protegidos y carboxilo protegidos, o una sal de los mismos. Además, estos aminoácidos no naturales y cualquiera de los siguientes aminoácidos no naturales pueden incorporarse en un polipéptido de aminoácidos no naturales.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (XIII):



en el que B es opcional, y cuando está presente es un enlazador seleccionado del grupo que consiste en alquileo inferior, alquileo inferior sustituido, alquilenilo inferior, sustituido alquilenilo inferior, heteroalquileo inferior sustituido heteroalquileo inferior, -O-, -O-(alquileo o alquileo sustituido)-, -S-, -S-(alquileo o alquileo sustituido)-, -S(O)<sub>k</sub>- donde k es 1, 2, o 3, -S(O)<sub>k</sub>- (alquileo o alquileo sustituido)-, -C(O)-, -C(O)-(alquileo o alquileo sustituido)-, -C(S)-, -C(S)-(alquileo o alquileo sustituido)-, -N(R')-, -NR'-(alquileo o alquileo sustituido)-, -C(O)N(R')-, -CON(R')-(alquileo o alquileo sustituido)-, -CSN(R')-, -CSN(R')-(alquileo o alquileo sustituido)-, -N(R')CO-(alquileo o alquileo sustituido)-, -N(R')-C(O)O-, -S(O)<sub>k</sub>N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)<sub>k</sub>N(R')-, -N(R')-N, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')<sub>2</sub>-N=N, y -C(R')<sub>2</sub>-N(R')-N(R')-, donde cada R' es independientemente H, alquilo, o alquilo sustituido;

R es H, alquilo, alquilo, cicloalquilo, o cicloalquilo sustituido;

R1 es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

R2 es opcional y, cuando está presente, es OH, un grupo de éster, resina, aminoácido, polipéptido o polinucleótido protector;

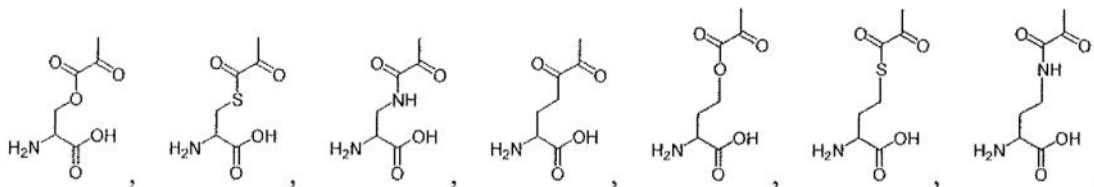
cada R<sub>a</sub> se selecciona independientemente del grupo que consiste de H, halógeno, alquilo, alquilo sustituido, -N(R')<sub>2</sub>, -C(O)kR', donde k es 1, 2, o 3, -C(O)N(R')<sub>2</sub>, OR', y -S(O)kR' donde cada R1 es, independientemente, H, alquilo, o

alquilo sustituido; y n es de 0 a 8.

Además, se incluyen los siguientes aminoácidos:

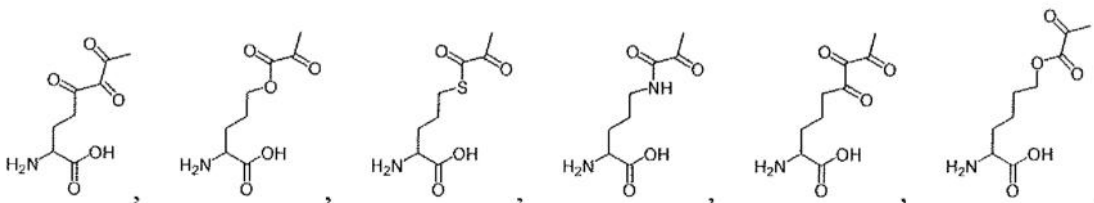
5

10



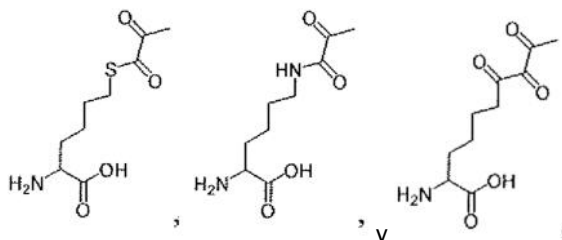
15

20



25

30



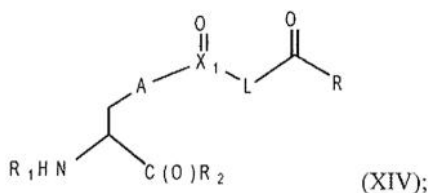
en donde dichos compuestos están opcionalmente amino protegidos, opcionalmente carboxilo protegidos, opcionalmente amino protegidos y carboxilo protegidos, o una sal del mismo. Además, estos aminoácidos no naturales y cualquiera de los siguientes aminoácidos no naturales pueden incorporarse en un polipéptido de aminoácidos no naturales.

35

Además, los siguientes aminoácidos que tienen la estructura de fórmula (XIV) están incluidos:

40

45



donde:

50

A es opcional, y cuando está presente es alquilenio inferior, alquilenio inferior sustituido, cicloalquilenio inferior sustituido, alquilenio inferior, alquilenio inferior sustituido, alquilenio, heteroalquilenio inferior, heteroalquilenio sustituido, heterocicloalquilenio inferior, heterocicloalquilenio inferior sustituido, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno sustituido, aralquilenio, o aralquilenio sustituido;

55

R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

R1 es opcional, y cuando está presente, es H, un grupo amino, resina, aminoácido, polipéptido o polinucleótido; y

60

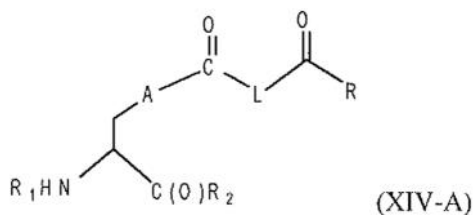
R2 es opcional y, cuando está presente, es de aceite, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

X1 es C, S, o S (O); y L es alquilenio, alquilenio sustituido, -N(R')-(alquilenio) o N(R')(alquilenio sustituido), donde R' es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido.

65

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (XIV-A):

5



10

donde:

15 A es opcional, y cuando está presente es alquileno inferior, alquileno inferior sustituido, cicloalquileno, 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;

R es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;-

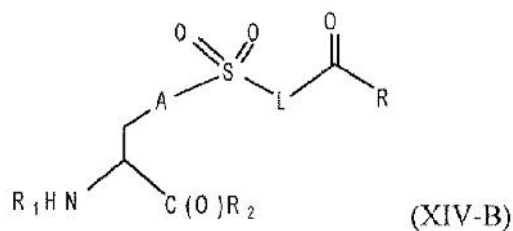
20 R1 es opcional, y cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

25 R2 es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

L es alquileno, alquileno sustituido, -N(R')-(alquileno) o -N(R')-(alquileno sustituido)-, donde R' es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido.

30 Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (XIV-B):

35



40

donde:

45 A es opcional, y cuando está presente es alquileno inferior, alquileno inferior sustituido, cicloalquileno, 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;

R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

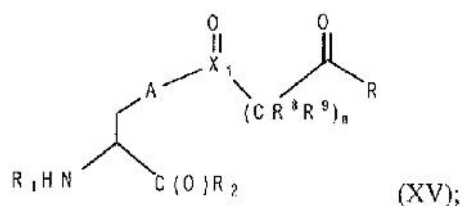
50 R1 es opcional, y cuando está presente, es H, un grupo amino, resina, aminoácido, polipéptido o polinucleótido; y

R2 es opcional y, cuando está presente, es OH, un grupo de éster, resina, aminoácido, polipéptido o polinucleótido;

55 L es alquileno, alquileno sustituido, -N(R')-(alquileno) o -N(R')-(alquileno sustituido)-, en donde R1 es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (XV):

60



65



donde:

A 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, hclcroarileno, heteroarileno sustituido, alcarileno, alcarileno, aralquileno o aralquileno sustituido;

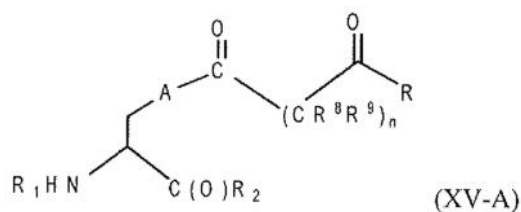
R es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo amino, resina, aminoácido, polipéptido o polinucleótido; y

R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

X<sub>1</sub> es C, S, o S(O); y n es 0, 1, 2, 3, 4, o 5; y cada R<sup>8</sup> y R<sup>9</sup> en cada grupo CR<sup>8</sup>R<sup>9</sup> se selecciona independientemente del grupo que consiste en H, alcoxi, alquilamina, halógeno, alquilo, arilo, o cualquier R<sup>8</sup> y R<sup>9</sup> juntos pueden formar =O o un cicloalquilo, o cualquiera de los grupos de R<sup>8</sup> adyacentes pueden formar juntos un cicloalquilo.

Además, los siguientes aminoácidos que tienen la estructura de Fórmula (XV-A) están incluidos:



donde:

A es opcional, y cuando está presente es alquileno inferior, alquileno inferior sustituido, cicloalquileno, 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, aralquileno, o aralquileno sustituido;

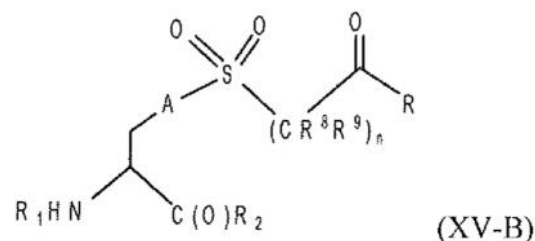
R es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo amino, resina, aminoácido, polipéptido o polinucleótido; y

R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo éster, resina, aminoácido, polipéptido o polinucleótido;

n es 0, 1, 2, 3, 4, o 5; y cada R<sup>8</sup> y R<sup>9</sup> en cada grupo CR<sup>8</sup>R<sup>9</sup> se selecciona independientemente del grupo que consiste en H, alcoxi, alquilamina, halógeno, alquilo, arilo, o cualquier R<sup>8</sup> y R<sup>9</sup> pueden juntos formar =O o un cicloalquilo, o cualquiera a grupos R<sup>8</sup> adyacentes pueden juntos formar un cicloalquilo.

Además, están incluidos los siguientes aminoácidos que tienen la estructura de Fórmula (XV-B):



donde:

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

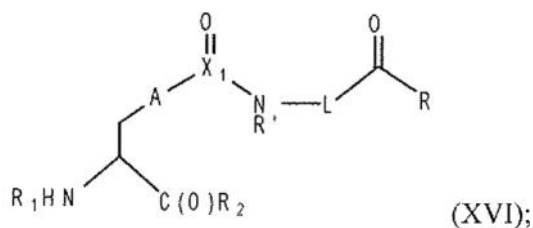
R es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

R es opcional, y cuando está presente, es H, un grupo amino, resina, aminoácido, polipéptido o polinucleótido; y

R<sub>2</sub> es opcional y, cuando está presente, es OH, un éster grupo, resina, aminoácido, polipéptido o polinucleótido;

n es 0, 1, 2, 3, 4, o 5; y cada R<sup>8</sup> y R<sup>9</sup> en cada grupo CR<sup>8</sup>R<sup>9</sup> se selecciona independientemente del grupo que consiste en H, alcoxi, alquilamina, halógeno, alquilo, arilo, o cualquier R<sup>8</sup> y R<sup>9</sup> pueden juntos formar =O o un cicloalquilo, o cualesquiera grupos R<sup>8</sup> adyacentes pueden juntos formar un cicloalquilo.

Además, los siguientes aminoácidos que tienen la estructura de Fórmula (XVI) están incluidos:



donde:

A 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, aralquileno, o aralquileno sustituido;

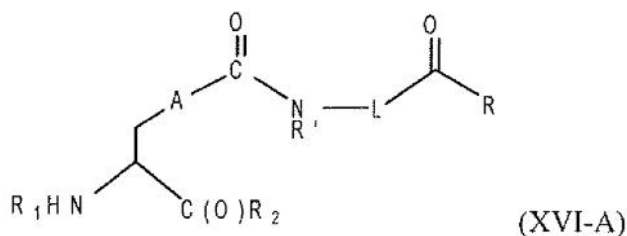
R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

R<sub>2</sub> es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

X<sub>1</sub> es C, S, o S(O); y L es alquileno, alquileno sustituido, -N(R')-(alquileno) o -N(R')-(alquileno sustituido)-, en donde R' es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (XVI-A):



donde:

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

R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

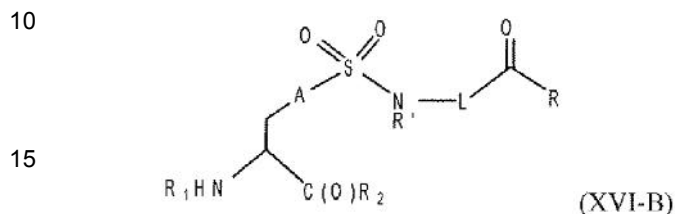
R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo de amino, resina, aminoácido, polipepti de, o polinucleótido

protector; y

R<sup>2</sup> es opcional y, cuando está presente, es OH, un éster grupo, resina, aminoácido, polipéptido o polinucleótido;

5 L es alquileo, alquileo sustituido, -N(R')-(alquileo) o -N(R')-(alquileo sustituido)-, en donde R<sub>1</sub> es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido.

Además, se incluyen los siguientes aminoácidos que tienen la estructura de Fórmula (XVI-B):



donde:

20 A es opcional, y cuando está presente es alquileo inferior, alquileo inferior sustituido, cicloalquileo inferior, cicloalquileo inferior sustituido, alquilenilo inferior, alquilenilo inferior sustituido, alquinileno, heteroalquileo inferior, heteroalquileo sustituido, heterocicloalquileo inferior, heterocicloalquileo inferior sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno sustituido, aralquileo, o aralquileo sustituido;

25

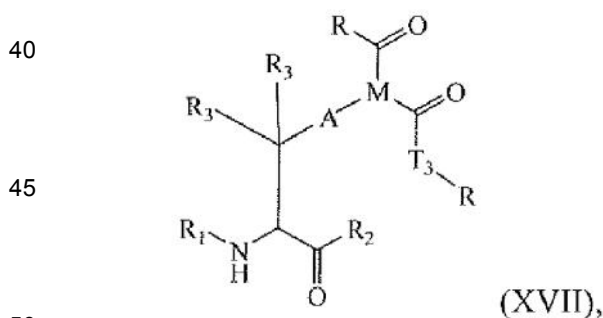
R es H, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

30 R<sub>1</sub> es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y R<sup>2</sup> es opcional y, cuando está presente, es el aceite, un grupo de éster, resina, aminoácido, polipéptido o polinucleótido;

L es alquileo, alquileo sustituido, -N(R')-(alquileo) o -N(R')-(alquileo sustituido)-, en donde R<sub>1</sub> es H, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido.

35

Además, los aminoácidos que tienen la estructura de Fórmula (XVII) están incluidos:



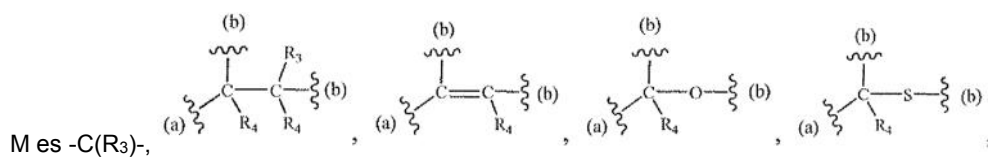
en donde:

55 A es opcional, y cuando está presente es alquileo inferior, alquileo inferior sustituido, cicloalquileo, cicloalquileo inferior sustituido, alquilenilo inferior, alquilenilo inferior sustituido, alquinileno, heteroalquileo inferior, heteroalquileo sustituido, heterocicloalquileo inferior, heterocicloalquileo inferior sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno, aralquileo o aralquileo sustituido;

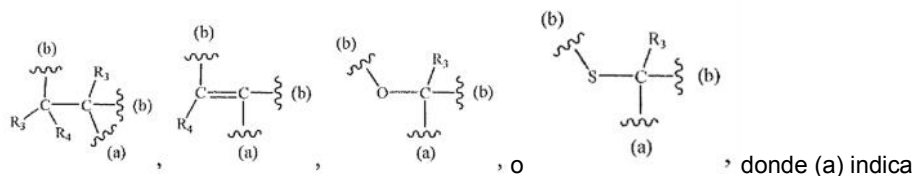
60

65

5



10



15

20 el enlace al grupo A y (b) indica el enlace a los respectivos grupos de carbonilo,  $R_3$  y  $R_4$  se eligen independientemente de H, halógeno, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido, o  $R_3$  y  $R_4$  o dos grupos  $R_3$  o dos grupos  $R_4$  opcionalmente forman un cicloalquilo o un heterocicloalquilo;

R es H, halógeno, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

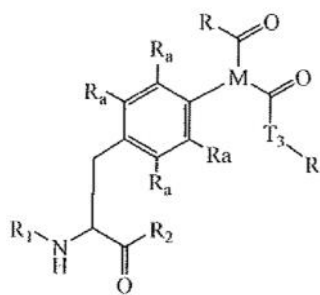
25  $T_3$  es un enlace, C(R)(R), O, o S, y R es H, halógeno, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

$R_1$  es opcional y, cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipeptide, o polinucleótido; y

30  $R^2$  es opcional y, cuando está presente, es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido.

Además, los aminoácidos que tienen la estructura de Fórmula (XVIII) están incluidos:

35



40

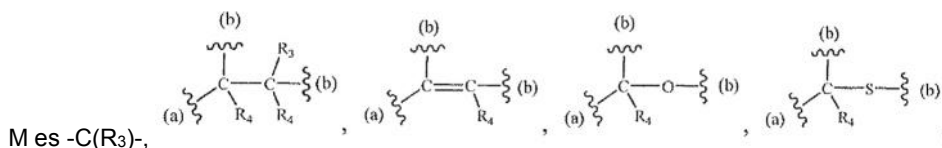
45

(XVIII),

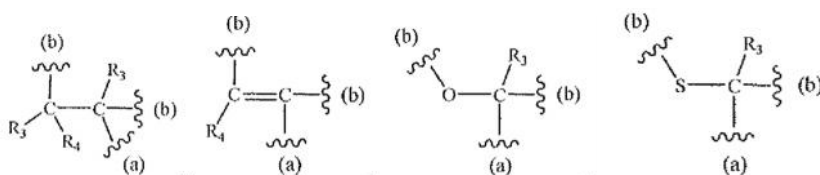
donde:

50

55



60



65

el enlace al grupo A y (b) indica el enlace a los respectivos grupos de carbonilo, R<sub>3</sub> y R<sub>4</sub> se eligen independientemente entre H, halógeno, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido, o R<sub>3</sub> y R<sub>4</sub> o dos grupos R<sub>3</sub> o dos grupos R<sub>4</sub> forman opcionalmente un cicloalquilo o un heterocicloalquilo;

5 R es H, halógeno, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido;

T<sub>3</sub> es un enlace, C(R)(R), O, o S, y P. es H, halógeno, alquilo, alquilo sustituido, cicloalquilo, o cicloalquilo sustituido;

10 R<sub>1</sub> es opcional, y cuando está presente, es H, un grupo protector de amino, resina, aminoácido, polipéptido, o polinucleótido; y

R<sub>2</sub> es opcional, y cuando está presente, es OH, un éster grupo, resina, aminoácido, polipéptido o polinucleótido;

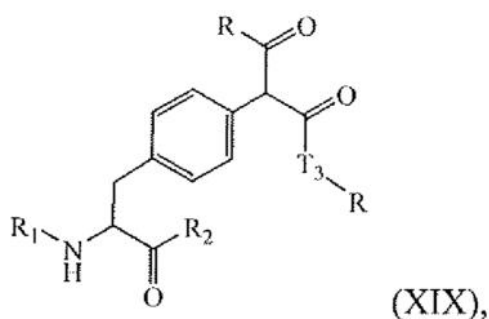
15 cada R<sub>a</sub> se selecciona independientemente del grupo que consiste de H, halógeno, alquilo, alquilo sustituido, -N(R')<sub>2</sub>, -C(O)<sub>k</sub>R', donde k es 1, 2, o 3, -C(O)N(R')<sub>2</sub>, -OR' y -S(O)<sub>k</sub>R', donde cada R' es independientemente H, alquilo o alquilo sustituido.

Además, los aminoácidos que tienen la estructura de Fórmula (XIX) están incluidos:

20

25

30



donde:

35 R es H, halógeno, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido; y

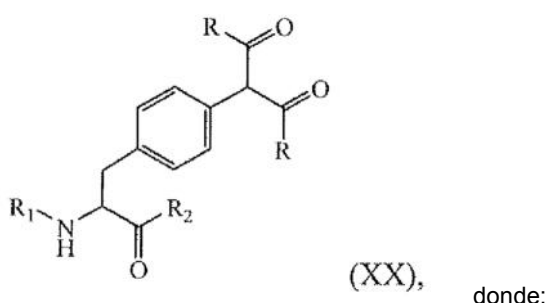
T<sub>3</sub> es O, o S.

Además, los aminoácidos que tienen la estructura de Fórmula (XX) se incluyen:

40

45

50



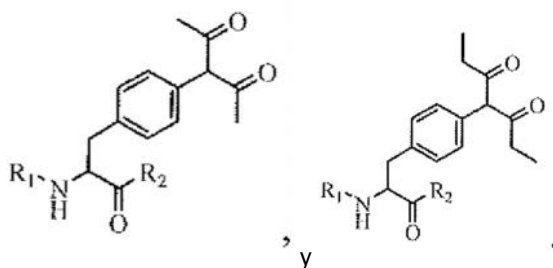
donde:

R es H, halógeno, alquilo, alquilo sustituido, cicloalquilo o cicloalquilo sustituido.

55 Además, los siguientes aminoácidos que tienen estructuras de fórmula (XXI) están incluidos:

60

65



En algunas realizaciones, un polipéptido que comprende un aminoácido no natural se modifica químicamente para generar un carbonilo reactivo o grupo de dicarbonilo funcional. Por ejemplo, una funcionalidad de aldehído útil para reacciones de conjugación se puede generar a partir de una funcionalidad que tiene grupos amino e hidroxilo adyacentes. Cuando la molécula biológicamente activa es un polipéptido, por ejemplo, una serina o treonina de término N (que puede estar normalmente presente o puede estar expuesto a través de digestión química o enzimática) se puede utilizar para generar una funcionalidad de aldehído en condiciones de escisión oxidativa suaves utilizando peryodato. Véase, por ejemplo, Gaertner, et. al., *Bioconjug. Chem.* 3: 262-268 (1992); Geoghegan, K. & Stroh, J., *Bioconjug. Chem.* 3: 138-146 (1992); Gaertner *et al.*, *J. Biol. Chem.* 269: 7224-7230 (1994). Sin embargo, los métodos conocidos en la técnica se limitan al aminoácido en el término N del péptido o proteína.

En la presente invención, un aminoácido no natural que lleva grupos hidroxilo y amino adyacentes se pueden incorporar en el polipéptido como una funcionalidad de aldehído "enmascarado". Por ejemplo, 5-hidroxilisina lleva un grupo de hidroxilo adyacente a la amina de épsilon. Las condiciones de reacción para generar el aldehído típicamente implican la adición de un exceso molar de metaperyodato de sodio en condiciones suaves para evitar la oxidación en otros sitios dentro del polipéptido. El pH de la reacción de oxidación es típicamente de aproximadamente 7,0. Una reacción típica implica la adición de aproximadamente 1,5 molar en exceso de peryodato meta de sodio a una solución tamponada del polipéptido, seguido de incubación durante aproximadamente 10 minutos en la oscuridad. Véase, p.ej. Patente de EE.UU. nº 6.423.685.

La funcionalidad de carbonilo o dicarbonilo se puede hacer reaccionar selectivamente con un reactivo que contiene hidroxilamina en condiciones suaves en disolución acuosa para formar el enlace de oxima correspondiente que es estable en condiciones fisiológicas. Véase, por ejemplo, Jencks, W.P., *J. Am. Chem. Soc.* 81, 475-481 (1959); Shao, J. y Tam, J.P., *J. Am. Chem. Soc.* 117: 3893-3899 (1995). Además, la reactividad única del grupo carbonilo o dicarbonilo permite la modificación selectiva en presencia de las otras cadenas laterales de aminoácidos. Véase, por ejemplo, Cornish, V.W., *et al.*, *J. Am. Chem. Soc.* 118: 8150-8151 (1996); Geoghegan, KF y Stroh, JG, *Bioconjug. Chem.* 3:138-146 (1992); Mahal, L.K., *et al.*, *Science* 276:1125-1128 (1997).

### **Estructura y síntesis de aminoácidos no naturales: Aminoácidos que contienen hidroxilamina**

#### **SÍNTESIS QUÍMICA DE AMINOÁCIDOS NO NATURALES**

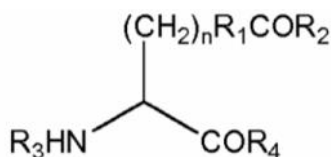
Muchos de los aminoácidos no naturales adecuados para su uso en la presente invención están disponibles comercialmente, por ejemplo, de Sigma (EE.UU.) o Aldrich (Milwaukee, WI, EE.UU.). Los que no están disponibles comercialmente se sintetizan opcionalmente como se describe aquí o como se describe en varias publicaciones o usando procedimientos estándar conocidos por los expertos normales en la técnica. Para las técnicas de síntesis orgánica, véase, por ejemplo, *Organic Chemistry* de Fessenden y Fessenden, (1982, Segunda Edición, Willard Grant Press, Boston Mass.); *Advanced Organic Chemistry* por March (Tercera Edición, 1985, Wiley and Sons, Nueva York); y *Advanced Organic Chemistry* por Carey y Sundberg (Tercera edición, partes A y B, 1990, Plenum Press, Nueva York). Otras publicaciones que describen la síntesis de aminoácidos no naturales incluyen, por ejemplo, el documento WO 2002/085923 titulado "In vivo incorporation of Unnatural Amino Acids;" Matsoukas *et al.*, (1995) *J. Med. Chem.* 38, 4660-4669.; Rey, F.E. y Kidd, D.A.A. (1949) An New Synthesis of Glutamine and of  $\gamma$ -Dipeptides of Glutamic Acid from Phthylated Intermediates. *J. Chem. Soc.* 3315-3319.; Friedman, O.M. y Chatterji, R. (1959) Synthesis of Derivates of Glutamine as Model Substates for Anti-Tumor Agents. *J. Am. Chem. Soc.* 81, 3750-3752; Craig, J.C. *et al.* (1988) Absolute Configuration of the Enantiomers of 7-Cloro-4 [[4-(diethylamino)-1-methylbutyl]amino]quinoline (Chloroquine). *J. Org. Chem.* 53, 1167-1170; Azoulay, M., Vilmont, M. & Frappier, F. (1991) Glutamine analogues as Potential Antimalarials, *Eur. J. Med. Chem.* 26, 201-5; Koskinen, A.M.P. y Rapoport, H. (1989) Synthesis of 4-Substituted Prolines as Conformationally Constrained Amino Acid Analogues. *J. Org. Chem.* 54, 1859-1866; Christie, B.D. y Rapoport, H. (1985) Synthesis of Optically Pure Pipecolates from L-Asparagine. Application to the Total Synthesis of (+)-Apovincamine through Amino Acid Decabonylation and Iminium Ion Cyclization. *J. Org. Chem.* 50: 1239-1246; Barton *et al.*, (1987) Synthesis of Novel  $\alpha$ -Amino-Acids and Derivatives Using Radical Chemistry: Synthesis of L- and D- $\alpha$ -Amino-Adipic Acids, L- $\alpha$ -aminopimelic Acid and Appropriate Unsaturated Derivatives. *Tetrahedron* 43: 4297-4308; y, Subasinghe *et al.*, (1992) Quisqualic acid analogues: synthesis of beta-heterocyclic 2-aminopropanoic acid derivatives and their activity at a novel quisqualate-sensitized site. *J. Med. Chem.* 35: 4602-7. Véase también, la Publicación de Patente de Estados Unidos nº 2004/0198637 titulada "Protein Arrays".

#### **A. Grupos de reactivos de carbonilo**

Los aminoácidos con un grupo reactivo carbonilo permiten una variedad de reacciones para enlazar moléculas (incluyendo pero no limitado a, PEG u otras moléculas solubles en agua) a través de la adición nucleófila o reacciones de condensación aldólica entre otros.

Aminoácidos ejemplares que contienen carbonilo pueden ser representados del siguiente modo:

5



10 en donde n es 0-10; R<sub>1</sub> es un alquilo, arilo, alquilo sustituido, o arilo sustituido; R<sup>2</sup> es H, alquilo, arilo, alquilo sustituido, y arilo sustituido; y R<sub>3</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación amino terminal, y R<sub>4</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación terminal carboxi. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo y R<sup>2</sup> es un alquilo sencillo (es decir, metilo, etilo, o propilo) y el resto de cetona se coloca en la posición *para* con respecto a la cadena lateral de alquilo. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo y R<sup>2</sup> es un alquilo sencillo (es decir, metilo, etilo, o propilo) y el resto de cetona se coloca en la posición *meta* respecto a la cadena lateral de alquilo.

La síntesis de p-acetilo-(+/-)-fenilo-alanina y m-acetilo-(+/-)-fenilo-alanina se describe en Zhang, Z., et al., BioChemistry, 42: 6735-6746 (2003). Otros aminoácidos que contienen carbonilo se pueden preparar de manera similar por una persona de experiencia ordinaria en la técnica.

En algunas realizaciones, un polipéptido que comprende un aminoácido no naturalmente codificado se modifica químicamente para generar un grupo funcional de carbonilo reactivo. Por ejemplo, una funcionalidad de aldehído útil para reacciones de conjugación se puede generar a partir de una funcionalidad que tiene grupos amino e hidroxilo adyacentes. Cuando la molécula biológicamente activa es un polipéptido, por ejemplo, una serina N-terminal o treonina (que puede estar normalmente presente o puede estar expuesta a través de digestión química o enzimática) se pueden utilizar para generar una funcionalidad de aldehído en condiciones de escisión oxidativa suaves utilizando peryodato. Véase, por ejemplo, Gaertner, *et al.*, Bioconjug. Chem. 3: 262-268 (1992); Geoghegan, K. & Stroh, J., Bioconjug. Chem. 3: 138-146 (1992); Gaertner *et al.*, -1-Biol. Chem. 269: 7224-7230 (1994). Sin embargo, los métodos conocidos en la técnica se limitan al aminoácido en el término N del péptido o proteína.

Un aminoácido no naturalmente codificado que lleva grupos hidroxilo y amino adyacentes se puede incorporar en el polipéptido como una funcionalidad de aldehído "enmascarado". Por ejemplo, 5-hidroxisisina lleva un grupo de hidroxilo adyacente a la amina de épsilon. Las condiciones de reacción para generar el aldehído típicamente implican adición de un exceso molar de metaperyodato de sodio en condiciones suaves para evitar la oxidación en otros sitios dentro del polipéptido. El pH de la reacción de oxidación es típicamente de aproximadamente 7,0. Una reacción típica implica la adición de aproximadamente 1,5 de exceso molar de peryodato meta de sodio a una solución tamponada del polipéptido, seguido de incubación durante aproximadamente 10 minutos en la oscuridad. Véase, por ejemplo, la Patente de EE.UU. n° 6.423.685.

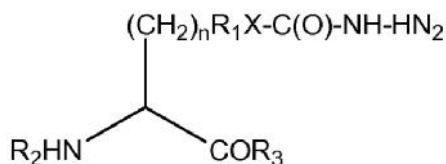
La funcionalidad carbonilo se puede hacer reaccionar selectivamente con una hidrazina, hidrazida, hidroxilamina, o reactivo semicarbazida contiene azida en condiciones suaves en disolución acuosa para formar la correspondiente hidrazona, oxima, o enlaces de semicarbazona, respectivamente, que son estables en condiciones fisiológicas. Véase, por ejemplo, Jencks, WP, J. Am. Chem. Soc. 81, 475-481 (1959); Shao, J. y Tam, JP, 01 Am. Chem. Soc. 117: 3893-3899 (1995). Además, la reactividad única del grupo carbonilo permite la modificación selectiva en presencia de las otras cadenas laterales de aminoácidos. Véase, por ejemplo, Cornish, V.W., et al., J Am. Chem. Soc. 118: 8150-8151 (1996); Geoghegan, K.F. y Stroh, J.G., Bioconjug. Chem. 3: 138-146 (1992); Mahal, L.K., et al., Science 276: 1125-1128 (1997).

## 50 B. Grupos reactivos de hidrazina, hidrazida o semicarbazida

Aminoácidos no naturales codificados que contienen un grupo nucleófilo, tal como una hidrazina, hidrazida o semicarbazida, permiten la reacción con una variedad de grupos electrófilos para formar conjugados (incluyendo pero no limitado a PEG u otro polímeros solubles en agua).

Hidrazina, hidrazida o semicarbazida ejemplares que contienen los aminoácidos pueden ser representados del siguiente modo:

60



65

en donde n es 0-10; R<sub>1</sub> es un alquilo, arilo, alquilo sustituido, o arilo sustituido o no está presente; X, es O, N, o S o no está presente; R<sup>2</sup> es H, un aminoácido, un polipéptido, o un grupo de modificación de terminal de amino, y R<sub>3</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación terminal de carboxi.

- 5 En algunas realizaciones, n es 4, R<sub>1</sub> no está presente, y X es N. En algunas realizaciones, n es 2, R<sub>1</sub> no está presente, y X no está presente. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo, X es O, y el átomo de oxígeno está posicionado *para* con respecto al grupo alifático en el anillo de arilo.

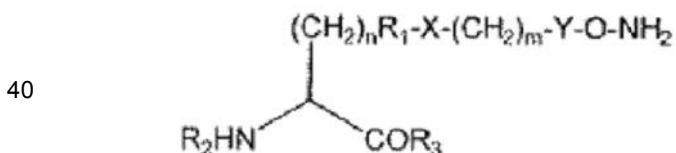
10 Aminoácidos que contienen hidrazida, hidrazina, y semicarbazida están disponibles de fuentes comerciales. Por ejemplo, L-glutamato-γ-hidrazida está disponible de Sigma Chemical (St. Louis, MO). Otros aminoácidos no comercialmente disponibles se pueden preparar por alguien de habilidad normal en la técnica. Véase, por ejemplo, la patente de los Estados Unidos. n° 6.281,211.

15 Polipéptidos que contienen aminoácidos no naturalmente codificados que llevan funcionalidades de hidrazida, hidrazina o semicarbazida se pueden hacer reaccionar de manera eficiente y selectiva con una variedad de moléculas que contienen aldehídos u otros grupos funcionales con reactividad química similar. Véase, por ejemplo, Shao, J. y Tam, J., *J. Am. Chem. Soc.* 117: 3893-3899 (1995). La reactividad única de grupos funcionales de hidrazida, hidrazina y semicarbazida los hace significativamente más reactivos frente a los aldehídos, cetonas y otros grupos electrofílicos en comparación con los grupos nucleófilos presentes en los 20 aminoácidos comunes (incluyendo, pero no limitado al grupo de hidroxilo de grupos de serina o treonina o el amino de la lisina y el término N).

### C. Aminoácidos que contienen amino

25 Aminoácidos no naturalmente codificados que contienen un grupo de aminooxi (también llamados hidroxilamina) permite la reacción con una variedad de grupos electrófilos para formar conjugados (incluyendo pero no limitado a, con PEG u otros polímeros solubles en agua). Como hidrazinas, hidrazidas y semicarbazidas, la nucleofilicidad mejorada del grupo aminooxi permite que reaccione de manera eficiente y selectiva con una variedad de moléculas que contienen aldehídos u otros grupos funcionales con reactividad química similar. Véase, por ejemplo, Shao, J. y Tam, J., *J. Am. Chem. Soc.* 117: 3893-3899 (1995); H. Cuelgue y C. Bertozzi, *Acc. Chem. Res.* 34: 727-736 (2001). Considerando que el resultado de la reacción con un grupo hidrazina es la hidrazona correspondiente, sin embargo, una oxima resulta generalmente de la reacción de un grupo aminooxi con un grupo que contiene carbonilo tal como una cetona.

35 Aminoácidos ejemplares que contienen grupos aminooxi pueden ser representados de la siguiente manera:



45 en donde n es 0-10; R<sub>1</sub> es un alquilo, arilo, alquilo sustituido, o arilo sustituido o no está presente; X es O, N, S o no está presente; m es 0-10; Y = C(O) o no está presente; R<sup>2</sup> es H, un aminoácido, un polipéptido, o un grupo de modificación amino terminal, y R<sub>3</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación terminal carboxi. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo, X es O, m es 1, e Y está presente. En algunas realizaciones, n es 2, R<sub>1</sub> y X no están presentes, m es 0, e Y no está presente.

50 Aminoácidos que contienen amino se pueden preparar a partir de precursores de aminoácidos fácilmente disponibles (homoserina, serina y treonina). Véase, por ejemplo, M. Carrasco y R. Brown, *J. Org. Chem.* 68: 8853-8858 (2003). Ciertos aminoácidos que contienen amino, tal como L-2-amino-4-(aminooxi)ácido butírico, se han aislado de fuentes naturales (Rosenthal, G., *Life Sci.* 60: 1635-1641 (1997). Otros aminoácidos que contienen amino se pueden preparar por un experto ordinario en el arte.

### D. Grupos reactivos de azida y alquino

60 La reactividad única de grupos funcionales de azida y alquino los hace extremadamente útiles para la modificación selectiva de polipéptidos y otras moléculas biológicas. Azidas orgánicas, azidas particularmente alifáticas, y alquinos son generalmente estables hacia las condiciones químicas reactivas comunes. En particular, los grupos funcionales tanto de azida como de alquino son inertes hacia las cadenas laterales (es decir, grupos R) de los 20 aminoácidos comunes encontrados en polipéptidos naturales. Cuando se pone en estrecha proximidad, sin embargo, la naturaleza "por resorte" de los grupos de azida y alquino se revela y reaccionan de forma selectiva y eficiente a través de reacción de cicloadición de Huisgen [3 + 2] para generar el triazol correspondiente. Véase, por ejemplo, Chin J., et al., *Science* 301: 964-7 (2003); Wang, Q., *et al.*, *J. Am. Chem. Soc.* 125, 3192-3193-(2003); Chin, J.W., et



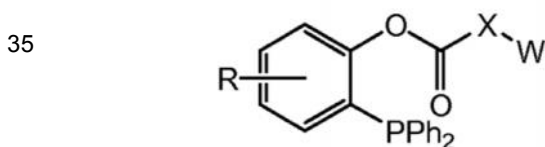
al., J. Am. Chem. Soc. 124:9026-9027 (2002).

Debido a que la reacción de cicloadición de Huisgen implica una reacción selectiva de cicloadición (véase, por ejemplo, Padwa, A., en *COMPREHENSIVE ORGANIC SYNTHESIS*, Vol 4, (ed. Trost, B.M., 1991), p 1069-1109; Huisgen, R. en *1,3-DIPOLAR CYCLOADDITION CHEMISTRY*, (ed. Padwa, A., 1984), p. 1-176) en lugar de una sustitución nucleofílica, la incorporación de aminoácidos no naturalmente codificados que llevan cadenas laterales que contienen azida y alquino permite que se modifiquen selectivamente los polipéptidos resultantes en la posición del aminoácido no naturalmente codificado. La reacción de cicloadición que implica que se realice polipéptido de insulina que contiene azida o alquino a temperatura ambiente en condiciones acuosas mediante la adición de Cu(II)(incluyendo pero no limitado a, en forma de una cantidad catalítica de CuSO<sub>4</sub>) en presencia de un agente reductor para la reducción de Cu(II) a Cu(I), in situ, en cantidad catalítica. Véase, por ejemplo, Wang, Q., et al., J. Am. Chem. Soc. 125, 3192-3193 (2003); Tornøe, C.W., et al., J. Org. Chem. 67: 3057-3064 (2002); Rostovtsev, et al., Angew. Chem. Int. Ed. 41: 2596-2599 (2002). Agentes reductores ejemplares incluyen, incluyendo pero no limitado a, ascorbato, cobre metálico, la quinina, la hidroquinona, vitamina K, glutatión, cisteína, Fe<sup>2+</sup>, Co<sup>2+</sup>, y un potencial eléctrico aplicado.

En algunos casos, cuando se desea una reacción de cicloadición de Huisgen [3 + 2] entre una azida y un alquino, el polipéptido de insulina comprende un aminoácido no naturalmente codificado que comprende un resto alquino y el polímero soluble en agua que se ha de unir al aminoácido comprende un resto de azida. Alternativamente, la reacción inversa (es decir, con el resto de azida en el aminoácido y el resto alquino presente en el polímero soluble en agua) también puede llevarse a cabo.

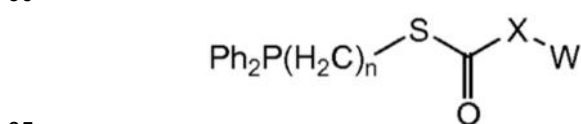
El grupo funcional de azida también puede hacerse reaccionar selectivamente con un polímero soluble en agua que contiene un éster de arilo y apropiadamente funcionalizado con un resto de fosfina de arilo para generar un enlace de amida. El grupo de fosfina de arilo reduce la azida in situ y la amina resultante luego reacciona eficientemente con un enlace éster proximal para generar la correspondiente amida. Véase, por ejemplo, E. Saxon y C. Bertozzi, *Science* 287, 2007-2010 (2000). El aminoácido que contiene azida puede ser o bien una azida de alquilo (incluyendo pero no limitado a, 2-amino-6-azido-1-ácido hexanoico) o una azida de arilo (p-azido-fenilalanina).

Polímeros solubles en agua ejemplar que contienen un éster de arilo y un resto de fosfina se pueden representar del siguiente modo:



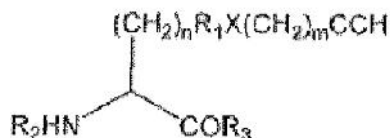
40 en donde X puede ser O, N, S o no está presente, Ph es fenilo, W es un polímero soluble en agua y R puede ser H, alquilo, arilo, grupos sustituidos por alquilo y sustituidos por arilo. Los grupos R ejemplares incluyen pero no se limitan a -CH<sub>2</sub>, -C(CH<sub>3</sub>), -OR', -NR'R'', -SR', -halógeno, -C(O)R', -CONR'R'', -S(O)<sub>2</sub>R', -S(O)<sub>2</sub>NR'R'', -CN y -NO<sub>2</sub>. R', R'', R''' y R'''' cada uno independientemente se refieren a hidrógeno, heteroalquilo sustituido o no sustituido, arilo sustituido o no sustituido, incluyendo pero no limitado a, arilo sustituido con 1-3 halógenos, alquilo sustituido o no sustituido, alcoxi o grupos de tioalcoxi, o grupos de arilalquilo. Cuando un compuesto de la invención incluye más de un grupo R, por ejemplo, cada uno de los grupos R se selecciona independientemente como son cada uno grupos R', R'', R''' y R'''' cuando más de uno de estos grupos está presente. Cuando R' y R'' están unidos al mismo átomo de nitrógeno, se pueden combinar con el átomo de nitrógeno para formar un anillo de 5, 6, o 7 miembros. Por ejemplo, -NR'R'' está destinado a incluir, pero no se limita a, 1-pirrolidinilo y 4-morfolinilo. De la discusión anterior de sustituyentes, un experto en la técnica entenderá que el término 'alquilo' incluye grupos que incluyen átomos de carbono unidos a grupos distintos de grupos de hidrógeno, tales como haloalquilo (incluyendo pero no limitado a, -CF<sub>3</sub> y -CH<sub>2</sub>CF<sub>3</sub>) y acilo (incluyendo pero no limitados a, -C(O)CH<sub>3</sub>, -C(O)CF<sub>3</sub>, -C(O)CH<sub>2</sub>OCH<sub>3</sub>, y similares).

55 El grupo funcional de azida también puede hacerse reaccionar selectivamente con un polímero soluble en agua que contiene un tioéster y apropiadamente funcionalizado con un resto de fosfina de arilo para generar un enlace de amida. El grupo de fosfina de arilo reduce la azida in situ y la amina resultante reacciona entonces de manera eficiente con el enlace de tioéster para generar la correspondiente amida. Polímeros solubles en agua ejemplar que contienen un tioéster y un resto de fosfina pueden ser representados del siguiente modo:



en donde n es 1-10; X puede ser O, N, S o no está presente, Ph es fenilo, y W es un polímero soluble en agua.

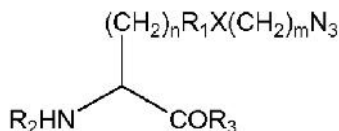
Aminoácidos ejemplares que contienen alquino pueden ser representados del siguiente modo:



en donde n es 0-10; R<sub>1</sub> es un alquilo, arilo, alquilo sustituido, o arilo sustituido o no está presente; X es O, N, S o no está presente; m es 0-10, R<sub>2</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación terminal amino, y R<sub>3</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación terminal carboxi. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo, X no está presente, m es 0 y el resto acetileno se coloca en la posición para con respecto a la cadena lateral de alquilo. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo, X es O, m es 1 y el grupo propargilo se coloca en la posición para con respecto a la cadena lateral de alquilo (es decir, O-propargilo-tirosina). En algunas realizaciones, n es 1, R<sub>1</sub> y X no están presentes y m es 0 (es decir, proparilglicina).

Aminoácidos que contienen alquino están disponibles comercialmente. Por ejemplo, propargilglicina está disponible comercialmente de Peptech (Burlington, MA). Alternativamente, aminoácidos que contienen alquino se pueden preparar de acuerdo con métodos estándar. Por ejemplo, *p*-propargiloxifeniloalanina se puede sintetizar, por ejemplo, como se describe en Deiters, A., et al., J. Am. Chem. Soc. 125: 11782-11783 (2003), y 4-alquino-L-feniloalanina se pueden sintetizar como se describe en Kayser, B., et al., Tetrahedron 53-(7): 2475-2484 (1997). Otros aminoácidos que contienen alquino se pueden preparar por alguien de habilidad normal en la técnica.

Aminoácidos ejemplares que contienen azida pueden ser representados del siguiente modo:



en donde n es 0-10; R<sub>1</sub> es un alquilo, arilo, alquilo sustituido, arilo sustituido o no está presente; X es O, N, S o no está presente; m es 0-10; R<sub>2</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación terminal amino, y R<sub>3</sub> es H, un aminoácido, un polipéptido, o un grupo de modificación terminal carboxi. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo, X no está presente, m es 0 y el resto azida se coloca en posición para con la cadena lateral de alquilo. En algunas realizaciones, n es 0-4 y R<sub>1</sub> y X no están presentes, y m=0. En algunas realizaciones, n es 1, R<sub>1</sub> es fenilo, X es O, m es 2 y el resto β-azidoetoxi se coloca en la posición para con respecto a la cadena lateral de alquilo.

Aminoácidos que contienen azida están disponibles de fuentes comerciales. Por ejemplo, 4-azidofeniloalanina se puede obtener de Chem-Impex International, Inc. (Wood Dale, IL). Para los aminoácidos que contienen azida que no están disponibles comercialmente, el grupo azida se puede preparar relativamente fácilmente usando métodos estándar conocidos por personas de experiencia ordinaria en la técnica, incluyendo pero no limitado a, a través de desplazamiento de un grupo saliente adecuado (incluyendo pero no limitado a, haluro, mesilato, tosilato) o por medio de la apertura de una lactona adecuadamente protegida. Véase, por ejemplo, Advanced Organic Chemistry por March (Tercera Edición, 1985, Wiley and Sons, Nueva York).

### E. Grupos reactivos de aminotiol

La reactividad única de grupos funcionales de aminotiol beta-sustituidos los hace extremadamente útiles para la modificación selectiva de polipéptidos y otras moléculas biológicas que contienen grupos de aldehído mediante la formación de la tiazolidina. Véase, por ejemplo, J. Shao y J. Tam, J. Am. Chem. Soc. 1995, 117 (14) 3893-3899. En algunas realizaciones, aminoácidos de aminotiol beta-sustituidos pueden incorporarse en los polipéptidos de insulina y después se hacen reaccionar con polímeros solubles en agua que comprenden una funcionalidad de aldehído. En algunas realizaciones, un polímero soluble en agua, conjugado de droga u otra carga útil pueden acoplarse a un polipéptido de insulina que comprende un aminoácido aminotiol beta-sustituido a través de la formación de la tiazolidina.

### F. Grupos reactivos adicionales

Grupos reactivos adicionales y aminoácidos no naturalmente codificados, incluyendo pero no limitado a para-amino-fenilalanina, que pueden incorporarse en los polipéptidos de insulina de la invención se describen en las siguientes

solicitudes de patente: Publicación de Patente de Estados Unidos nº 2006/0194256, la Publicación de Patente de Estados Unidos nº 2006/0217532, patente de EE.UU. nº 2006/0217289, patente provisional de EE.UU. nº 60/755.338; Patente de Estados Unidos provisional nº 60/755.711; Patente de Estados Unidos provisional nº 60/755.018; Solicitud de Patente Internacional nº PCT/USO6/49397; WO 2006/069246; Patente de Estados Unidos provisional nº 60/743.041; Patente de Estados Unidos provisional nº 60/743.040; Solicitud de Patente Internacional nº PCT/USO6/47822; Patente de Estados Unidos provisional nº 60/882.819; Patente de Estados Unidos provisional nº 60/882.500; y Patente provisional de Estados Unidos nº 60/870.594. Estas aplicaciones también discuten grupos reactivos que pueden estar presentes en PEG u otros polímeros, incluyendo, pero no limitado a, grupos de hidroxilamina (aminooxi) para la conjugación.

## CAPTACIÓN CELULAR DE AMINOÁCIDOS NO NATURALES

La captación de aminoácidos no naturales por una célula es una cuestión que se considera típicamente en el diseño y la selección de aminoácidos no naturales, incluyendo, pero no limitado a, para la incorporación en una proteína. Por ejemplo, la alta densidad de carga de aminoácidos sugiere que es poco probable que estos compuestos sean permeables a las células. Los aminoácidos naturales son absorbidos en la célula eucariótica a través de una colección de sistemas de transporte a base de proteína. Un examen rápido se puede realizar que evalúa qué aminoácidos no naturales, en su caso, se absorben por las células. Véase, por ejemplo, los ensayos de toxicidad en, por ejemplo, la Publicación de Patente de Estados Unidos nº US 2004/0198637 titulada "Protein Arrays"; y Liu, D.R. & Schultz, P.G. (1999) Progress toward the evolution of an organism with an expanded genetic code. PNAS Estados Unidos 96: 4780-4785. Aunque la captación se analiza fácilmente con diversos ensayos, una alternativa para diseñar aminoácidos no naturales que son susceptibles a rutas de captación celulares consiste en proporcionar rutas biosintéticas para crear aminoácidos *in vivo*.

## BIOSÍNTESIS DE AMINOÁCIDOS NO NATURALES

Muchas rutas biosintéticas ya existen en las células para la producción de aminoácidos y otros compuestos. Mientras que un método biosintético para un aminoácido no natural concreto puede no existir en la naturaleza, incluyendo pero no limitado a, en una célula, la descripción proporciona tales métodos. Por ejemplo, las rutas biosintéticas para aminoácidos no naturales se generan opcionalmente en la célula huésped mediante la adición de nuevas enzimas o modificando las vías de la célula huésped existentes. Nuevas enzimas adicionales son opcionalmente enzimas de origen natural o enzimas desarrolladas artificialmente. Por ejemplo, la biosíntesis de *p*-amino-feniloalanina (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 se pueden introducir en una célula eucariota mediante la transformación de la célula con un plásmido que comprende los genes. Los genes, cuando se expresa en la célula, proporcionan una ruta enzimática para sintetizar el compuesto deseado. Ejemplos de los tipos de enzimas que se añaden opcionalmente se proporcionan en los ejemplos a continuación. Secuencias de enzimas adicionales se encuentran, por ejemplo, en Genbank. Enzimas artificialmente evolucionadas también se añaden opcionalmente en una célula de la misma manera. De esta manera, la maquinaria celular y los recursos de una célula se manipulan para producir aminoácidos no naturales.

Una variedad de métodos están disponibles para la producción de nuevas enzimas para su uso en rutas biosintéticas o para la evolución de rutas existentes. Por ejemplo, la recombinación recursiva, incluyendo pero no limitado a, como el desarrollado por Maxygen, Inc. (disponible en la World Wide Web en maxygen.com), se usa opcionalmente para desarrollar nuevas enzimas y vías. Véase, 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. EE.UU., 91: 10747-10751. Del mismo modo DesignPath™, desarrollado por Genencor (disponible en el internet en genencor.com) se utiliza opcionalmente para la ingeniería por vía metabólica, incluyendo pero no limitado al diseño de una ruta para crear O-metilo-L-tirosina en una célula. Esta tecnología reconstruye vías existentes en organismos huésped utilizando una combinación de nuevos genes, incluyendo, pero no limitado a, aquellos identificados a través de la genómica funcional, y la evolución molecular y diseño. Diversa Corporation (disponible en el internet en diversa.com) también proporciona la tecnología para el examen rápido de bibliotecas de genes y vías de genes, incluyendo, pero no limitado a la creación de nuevas vías.

Típicamente, el aminoácido no natural producido con una ruta biosintética diseñada de la divulgación se produce en una concentración suficiente para la biosíntesis eficaz de proteínas, incluyendo, pero no limitado a, una cantidad celular natural, pero no en un grado que pueda afectar a la concentración de los otros aminoácidos o recursos celulares de escape. Las concentraciones típicas producidas *in vivo* de esta manera son de aproximadamente 10 min a aproximadamente 0,05 mM. Una vez que una célula se transforma con un plásmido que comprende los genes utilizados para producir enzimas deseadas para una ruta específica y un aminoácido no natural, selecciones *in vivo* se utilizan opcionalmente para optimizar aún más la producción del aminoácido no natural tanto para la síntesis de la proteína ribosómica como el crecimiento celular.

## POLIPÉPTIDOS CON AMINOÁCIDOS NO NATURALES

5 La incorporación de un aminoácido no natural se puede hacer para una variedad de propósitos, incluyendo pero no limitado a, la adaptación de los cambios en la estructura de la proteína y/o función, el cambio de tamaño, acidez, nucleofilia, enlaces de hidrógeno, hidrofobia, accesibilidad de sitios diana de proteasa, apuntando a un resto  
10 (incluyendo pero no limitado a, para una matriz de proteína), la adición de una molécula biológicamente activa, uniendo un polímero, adjuntando un radionúclido, la modulación de la vida media sérica, la modulación de la penetración del tejido (por ejemplo, tumores), la modulación de transporte activo, modulando tejido, célula o especificidad de órganos o la distribución, la modulación de la inmunogenicidad, modulación de la resistencia a la proteasa, etc. Las proteínas que incluyen un aminoácido no natural puede tener propiedades mejoradas o incluso  
15 totalmente nuevas catalíticas o biofísicas. Por ejemplo, las siguientes propiedades se modifican opcionalmente por la inclusión de un aminoácido no natural en una proteína: toxicidad, biodistribución, propiedades estructurales, propiedades espectroscópicas, químicas y/o propiedades fotoquímicas, capacidad catalítica, vida media (incluyendo pero no limitado a, vida media en suero), capacidad de reaccionar con otras moléculas, incluyendo, pero no limitado a, de forma covalente o no covalente, y similares. Las composiciones que incluyen proteínas que incluyen al menos un aminoácido no natural son útiles para, incluyendo, pero no limitado a, terapéuticos nuevos, diagnósticos, enzimas catalíticas, enzimas industriales, proteínas de unión (incluyendo pero no limitado a, anticuerpos), y incluyendo pero no limitado al 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.

20 También se describe aquí una composición que incluye al menos una proteína con al menos uno, incluyendo pero no limitado 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. Los aminoácidos no naturales pueden ser iguales o diferentes, incluyendo, pero no limitados a, puede haber 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 o más sitios diferentes en la proteína que comprenden 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 o más diferentes aminoácidos no naturales. En otro  
25 aspecto, una composición incluye una proteína con al menos uno, pero menos que 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 pero no limitados a 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  
30 aminoácidos no naturales pueden ser iguales, diferentes o una combinación de un aminoácido no natural múltiple del mismo tipo con al menos un aminoácido no natural diferente.

Las proteínas o polipéptidos de interés con al menos un aminoácido no natural son una característica de la descripción. La invención también incluye polipéptidos o proteínas con al menos un aminoácido no natural producido  
35 utilizando las composiciones y métodos de la descripción. Un excipiente (incluyendo pero no limitado a un excipiente farmacéuticamente aceptable) también puede estar presente con la proteína.

Mediante la producción de proteínas o polipéptidos de interés con al menos un aminoácido no natural en las células eucariotas, las proteínas o polipéptidos típicamente incluyen modificaciones postraduccionales eucariotas. En ciertas realizaciones, una proteína incluye al menos un aminoácido no natural y al menos una modificación postraducciona  
40 que se hace *in vivo* por una célula eucariota, cuando la modificación postraducciona no se hace por una célula procariota. Por ejemplo, la modificación postraducciona incluye, pero no se limita a acetilación, acilación, modificación de lípidos, palmitoilación, además de palmitato, fosforilación, modificación de enlace glicolípido, glicosilación, y similares. En un aspecto, la modificación postraducciona incluye la unión de un oligosacárido (incluyendo pero no limitado a, (GlcNAc-Man)<sub>2</sub>-Man-GlcNAc-GlcNAc) a una asparagina por un enlace GlcNAc-asparagina. Véase la Tabla 1 que enumera algunos ejemplos de oligosacáridos de enlace N de proteínas eucariotas (residuos adicionales también pueden estar presentes, que no se muestran). En otro aspecto, la modificación postraducciona incluye la unión de un oligosacárido (incluyendo pero no limitado a, Gal-GalNAc, Gal-GlcNAc, etc.) a una serina o treonina por un enlace de GalNAc-serina o GalNAc-treonina, o un enlace de GlcNAc-serina o GlcNAc-treonina.  
50

55

Tabla 1: Ejemplos de oligosacáridos a través de enlaces GLCNAC

Tipo	Estructura de base
MANOSA ALTA	$\begin{array}{l} \text{Man}\alpha 1-6 \\ \text{Man}\alpha 1-3 \end{array} \rightarrow \text{Man}\alpha 1-6 \rightarrow \text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}\beta 1-\text{Asn}$
HÍBRIDO	$\begin{array}{l} \text{Man}\alpha 1-6 \\ \text{GlcNAc}\beta 1-2 \text{ — } \text{Man}\alpha 1-3 \end{array} \rightarrow \text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}\beta 1-\text{Asn}$
COMPLEJO	$\begin{array}{l} \text{GlcNAc}\beta 1-2 \text{ — } \text{Man}\alpha 1-6 \\ \text{GlcNAc}\beta 1-2 \text{ — } \text{Man}\alpha 1-3 \end{array} \rightarrow \text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}\beta 1-\text{Asn}$
XILOSA	$\begin{array}{l} \text{Man}\alpha 1-6 \\ \text{Xyl}\beta 1-2 \end{array} \rightarrow \text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}\beta 1-\text{Asn}$

En otro aspecto más, la modificación postraduccional incluye procesamiento proteolítico de los precursores (incluyendo pero no limitado al precursor de calcitonina, al precursor del péptido relacionado con el gen de la calcitonina, hormona preproparatiroidea, preproinsulina, proinsulina, prepro-opiomelanocortina, proopiomelanocortina y similares), el montaje en una proteína de múltiples subunidades o ensamblaje macromolecular, traducción a otro sitio en la célula (incluyendo pero no limitado a orgánidos, tales como el retículo endoplasmático, el aparato de Golgi, el núcleo, lisosomas, peroxisomas, mitocondrias, cloroplastos, vacuolas, etc., o a través de la vía secretora). En ciertas realizaciones, la proteína comprende una secuencia de secreción o localización, una etiqueta de epítipo, una etiqueta FLAG, una etiqueta de poli-Histidina, una fusión GST, o similar.

Una de las ventajas de un aminoácido no natural es que presenta restos químicos adicionales que se pueden utilizar para añadir moléculas adicionales. Estas modificaciones pueden realizarse *in vivo* en una célula eucariota o no eucariótica, o *in vitro*. Por lo tanto, en ciertas realizaciones, la modificación postraduccional es a través del aminoácido no natural. Por ejemplo, la modificación postraduccional puede ser a través de una reacción nucleofílica electrofílica. La mayoría de las reacciones que actualmente se utilizan para la modificación selectiva de proteínas implican la formación del enlace covalente entre parejas de reacción nucleófilas y electrófilas, incluyendo, pero no limitado a la reacción de  $\alpha$ -halocetonas con cadenas laterales de histidina o cisteína. La selectividad en estos casos se determina por el número y la accesibilidad de los residuos nucleofílicos en la proteína. En las proteínas de la invención, otras reacciones más selectivas se pueden utilizar tal como la reacción de un ácido de cetamino no natural con hidrazidas o compuestos aminooxi, *in vitro* e *in vivo*. Véase, por ejemplo, Cornish, *et al.*, (1996) J. Am. Chem. Soc., 118: 8150-8151; Mahal, *et al.*, (1997) Science, 276: 1125-1128; Wang, *et al.*, (2001) Science 292: 498-500; Chin, *et al.*, (2002) J. Am. Chem. Soc. 124: 9026-9027; Chin, *et al.*, (2002) Proc. Natl. Acad. Sci., 99: 11020-11024; Wang, *et al.*, (2003) Proc. Natl. Acad. Sci., 100: 56-61; Zhang, *et al.*, (2003) BioChemistry, 42: 6735-6746; y, Chin, *et al.*, (2003) Science, 301: 964-7. Esto permite el etiquetado selectivo de prácticamente cualquier proteína con una serie de reactivos incluyendo fluoróforos, agentes de reticulación, derivados de sacáridos y moléculas citotóxicas. Véase también, la Patente de Estados Unidos nº 6.927.042 titulada "Glycoprotein synthesis". Modificaciones postraduccionales, incluyendo pero no limitado a, por un aminoácido de azido, también se pueden hacer a través de la ligadura de Staudinger (incluyendo pero no limitado a reactivos de triarilfosfina). Véase, por ejemplo, Kiick *et al.*, (2002) Incorporation of azidas into recombinant proteins for chemoselective modification by the Staudinger ligation, PNAS 99: 1924.

Esta descripción proporciona otro método altamente eficiente para la modificación selectiva de proteínas, lo que implica la incorporación genética de aminoácidos no naturales, incluyendo, pero no limitado a contener una azida o

alquinilo en proteínas en respuesta a un codón selector. Estas cadenas laterales de los aminoácidos se pueden modificar por, incluyendo pero no limitado a, una reacción de cicloadición de Huisgen [3 + 2] (véase, por ejemplo, Padwa, A. en *Comprehensive Organic Synthesis*, Vol. 4, (1991) Ed. Trost, B.M., Pergamon, Oxford, p 1069-1109; y, Huisgen, R. en *1,3-Dipolar Cycloaddition Chemistry*, (1984) Ed Padwa, A., Wiley, Nueva York, p 1-176) con, incluyendo pero no limitado a, derivados de alquinilo o azida, respectivamente. Debido a que este método implica una cicloadición en lugar de una sustitución nucleófila, las proteínas se pueden modificar con extremadamente alta selectividad. Esta reacción puede llevarse a cabo a temperatura ambiente en condiciones acuosas con excelente regioselectividad (1,4 > 1,5) mediante la adición de cantidades catalíticas de sales a la mezcla de reacción Cu(I). Véase, por ejemplo, Tornøe, *et al.*, (2002) *J. Org. Chem.* 67: 3057-3064; y, Rostovtsev, *et al.*, (2002) *Angew. Chem. Int. Ed.* 41: 2596-2599. Otro método que se puede utilizar es el intercambio de ligandos en un compuesto bisarsénico con un motivo de tetracisteína, véase, por ejemplo, Griffin, *et al.*, (1998) *Science* 281: 269-272.

Una molécula que se puede añadir a una proteína de la invención a través de una cicloadición [3 + 2] incluye prácticamente cualquier molécula con un derivado de azida o alquinilo. Las moléculas incluyen, pero no se limitan a, colorantes, fluoróforos, agentes de reticulación, derivados de sacáridos, polímeros (incluyendo, pero no limitado a, los derivados de polietilenglicol), agentes de reticulación, compuestos citotóxicos, marcadores de afinidad, derivados de biotina, resinas, perlas, una segunda proteína o polipéptido (o más), polinucleótido(s) (incluyendo pero no limitado a, ADN, ARN, etc.), quelantes de metales, cofactores, ácidos grasos, carbohidratos, y similares. Estas moléculas se pueden añadir a un aminoácido no natural con un grupo alquinilo, incluyendo pero no limitados a, p-propargiloxifeniloalanina, o grupo azido, incluyendo pero no limitados a, pazidofeniloalanina, respectivamente.

#### **Generación *in vivo* de los polipéptidos de insulina que comprende los aminoácidos no naturalmente codificados**

Los polipéptidos de insulina de la invención pueden ser generados *in vivo* usando sintetetas ARNt y ARNt modificadas para añadir a para sustituir aminoácidos que no están codificados en los sistemas en estado natural.

Los métodos para generar sintetetas de ARNt y ARNt que utilizan aminoácidos que no son codificados en los sistemas en estado natural se describen en, por ejemplo, Patente de Estados Unidos Nos. 7.045.337 y 7.083.970. Estos métodos implican la generación de una maquinaria de traducción que funciona independientemente de las sintetetas y los ARNt endógenos al sistema de traducción (y por lo tanto a veces se denominan "ortogonales"). Típicamente, el sistema de traducción comprende un ARNt ortogonal (O-ARNt) y una sintetasa de ARNt de aminoacilo ortogonal (O-RS). Típicamente, la O-RS aminoacila de forma preferente el O-ARNt con al menos un aminoácido no natural en el sistema de traducción y el O-ARNt reconoce al menos un codón selector que no es reconocido por otros ARNt en el sistema. Así pues, el sistema de traducción inserta el aminoácido no naturalmente codificado en una proteína producida en el sistema, en respuesta a un codón selector codificado, con lo que se produce la "sustitución" de un aminoácido en una posición en el polipéptido codificado.

Una amplia variedad de ARNt ortogonales y sintetetas de ARNt de aminoacilo se han descrito en la técnica para la inserción de aminoácidos particulares sintéticos en polipéptidos, y son generalmente adecuados para uso en la presente invención. Por ejemplo, sintetetas cetó-específicas O-ARNt/aminoacilo/ARNt se describen en Wang, L., *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 100: 56-61 (2003) y Zhang, Z. *et al.*, *Biochem.* 42-(22): 6735-6746 (2003). O-RS ejemplares, o porciones de los mismos, son codificados por secuencias de polinucleótido e incluyen secuencias de aminoácidos descritas en la Patente de los Estados Unidos Nos. 7.045.337 y 7.083.970. Las correspondientes moléculas de O-ARNt para uso con la OSR también se describen en la Patente de los Estados Unidos Nos. 7.045.337 y 7.083.970. Ejemplos adicionales de pares de sintetasa O-ARNt/aminoacilo se describen en el documento WO 2005/007870, WO 2005/007624; y WO 2005/019415.

Un ejemplo de un sistema de sintetasa O-ARNt/aminoacilo de azida específica se describe en Chin, J.W., *et al.*, *J. Am. Chem. Soc.* 124: 9026-9027 (2002). Secuencias O-RS ejemplares para p-azido-L-Phe incluyen, pero no se limitan a, secuencias de nucleótidos SEQ ID NOs: 14-16 y 29-32 y las secuencias de aminoácidos SEQ ID NOs: 46-48 y 61-64 como se describe en la Patente de Estados Unidos nº 7.083.970. Secuencias ejemplares de O-ARNt adecuadas para uso en la presente invención incluyen, pero no se limitan a, secuencias de nucleótidos SEQ ID NOs: 13, como se describe en la Patente de Estados Unidos nº 7.083.970. Otros ejemplos de pares específicos de sintetasa O-ARNt/aminoacilo-ARNt para aminoácidos particulares no naturalmente codificados se describen en la Patente de Estados Unidos nº 7.045.337. O-RS y O-ARNt que incorporan aminoácidos que contienen tanto cetó como azida en *S. cerevisiae* se describen en Chin, J.W., *et al.*, *Science* 301: 964-967 (2003).

Se ha informado de varios otros pares ortogonales. Glutaminilo (véase, por ejemplo, Liu, D.R., y Schultz, P.G. (1999) *Proc. Natl. Acad. Sci. EE.UU.* 96: 4780-4785), aspartilo (véase, por ejemplo, Pastrnak, M., *et al.*, (2000) *Helv. Chim. Acta* 83: 2277-2286), y tirosilo (véase, por ejemplo, Ohno, S., *et al.*, (1998). *J. Biochem (Tokio, Jpn.)* 124: 1065-1068; y, Kowal, A.K., *et al.*, (2001) *Proc. Natl. Acad. Sci. EE.UU.* 98: 2268-2273) sistemas derivados de *S. cerevisiae* de ARNt y sintetetas se han descrito para la incorporación potencial de aminoácidos no naturales en *E. coli*. Sistemas derivados de la glutaminilo de *E. coli* (véase, por ejemplo, Kowal, A.K., *et al.*, (2001) *Proc. Natl. Acad. Sci. EE.UU.* 98: 2268-2273) y sintetetas de tirosilo (véase, por ejemplo, Edwards, H., y Schimmel, P. (1990) *Mol. Cell. Biol.* 10: 1633-1641) se han descrito para uso en *S. cerevisiae*. El sistema de tirosilo *E. coli* se ha utilizado para la

incorporación de 3-yodo-L-tirosina *in vivo*, en células de mamífero. Véase, Sakamoto, K., *et al.*, (2002) *Nucleic Acids Res.* 30: 4692-4699.

El uso de sintetetasas de O-ARNt/aminoacilo-ARNt implica la selección de un codón específico que codifica el aminoácido no naturalmente codificado. Mientras que cualquier codón se puede utilizar, es generalmente deseable seleccionar un codón que rara vez o nunca se utiliza en la célula en la que se expresa la sintetasa O-ARNt/aminoacilo. Por ejemplo, los codones ejemplares incluyen codón sin sentido tales como codones de parada (ámbar, ocre y ópalo), cuatro o más codones base y otros codones de tres bases naturales que no se utilizan o se utilizan poco.

Codón selector específico se puede introducir en las posiciones apropiadas en la secuencia de codificación de polinucleótido de insulina usando métodos de mutagénesis conocidos en la técnica (incluyendo pero no limitados a, mutagénesis específica de sitio, mutagénesis de casete, la restricción de la selección de mutagénesis, etc.).

Los métodos para generar componentes de la maquinaria biosintética de proteínas, tales como los pares O-SR, O-ARNt, y O-ARNt/O-RS ortogonales que se pueden utilizar para incorporar un aminoácido no naturalmente codificado se describen en Wang, L., *et al.*, *Science* 292: 498-500 (2001); Chin, JW, *et al.*, *J. Am. Chem. Soc.* 124: 9026-9027 (2002); Zhang, Z. y otros, *Biochemistry* 42: 6735-6746 (2003). Métodos y composiciones para la incorporación *in vivo* de aminoácidos no naturalmente codificados se describen en la Patente de Estados Unidos nº 7.045.337. Los métodos para seleccionar un par de sintetasa ARNt-ARNt ortogonal para su uso en sistema de traducción *in vivo* de un organismo se describen también en la patente de EE.UU. Nos. 7.045.337 y 7.083.970. Publicación PCT nº WO 04/035743, titulada "Site Specific Incorporation of Keto Amino Acids into Proteins", describe pares RS y ARNt ortogonales para la incorporación de aminoácidos ceto. Publicación PCT nº WO 04/094593 titulada "Expanding the Eukaryotic Genetic Code," describe pares de RS y ARNt ortogonales para la incorporación de Aminoácidos no naturalmente codificados en células huésped eucariotas.

Los métodos para producir al menos una sintetasa (O-RS) de aminoacilo-ARNt ortogonal recombinante comprenden: (a) generar una biblioteca de RS (opcionalmente mutante) derivado de al menos una sintetasa (RS) de aminoacilo de un primer organismo, incluyendo, pero no limitado a, un organismo procariota, tal como *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, o similares, o un organismo eucariota; (b) seleccionar (y/o detectar) la biblioteca de RS (opcionalmente RS mutante) para los miembros que aminoacilan un ARNt ortogonal (O-ARNt) en presencia de un aminoácido no naturalmente codificado y un aminoácido natural, proporcionando de ese modo un depósito de RS activo (opcionalmente mutante); y/o, (c) seleccionar (opcionalmente a través de selección negativa) del depósito para RS activas (incluyendo pero no limitado a, RS mutante) que preferentemente aminoacilan el O-ARNt en ausencia del aminoácido no naturalmente codificado, proporcionando con ello al menos una O-RS recombinante; en el que al menos una O-RS recombinante aminoacila de forma preferente el O-ARNt con el aminoácido no naturalmente codificado.

En un caso, el RS es una RS inactiva. Las RS inactivas pueden ser generadas mediante la mutación de una RS activa. Por ejemplo, las RS inactivas pueden ser generadas mediante la mutación de 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, incluyendo pero no limitado a, alanina.

Las bibliotecas de RS mutantes pueden generarse usando diversas técnicas conocidas en la técnica, incluyendo pero no limitado a diseño racional basado en estructura de proteína de RS tridimensional, o la mutagénesis de nucleótidos de RS en una técnica de diseño aleatorio o racional. Por ejemplo, RS mutantes pueden generarse por mutaciones de sitio específico, mutaciones aleatorias, mutaciones de recombinación que generan diversidad, construcciones quiméricas, diseño racional y por otros métodos descritos en este documento o conocidos en la técnica.

En un caso, la selección (y/o el examen) de la biblioteca de RS (opcionalmente RS mutantes) para los miembros que están activos, incluyendo, pero no limitado a, que aminoacilan un ARNt ortogonal (O-ARNt) en presencia de un aminoácido no naturalmente codificado y un aminoácido natural, incluye: la introducción de una selección positiva o de detección de marcadores, incluyendo, pero no limitado a, un gen de resistencia a los antibióticos, o similares, y la biblioteca de RS (opcionalmente mutantes) en una pluralidad de células, en donde la selección positiva y/o de detección de marcadores comprende al menos un codón selector, incluyendo pero no limitado a, un color ámbar, ocre, o codón ópalo; el crecimiento de la pluralidad de células en presencia de un agente de selección; identificación de las células que sobreviven (o muestran una respuesta específica) en presencia de la selección y/o el agente de detección mediante la supresión de al menos un codón selector en la selección positiva o de detección de marcadores, proporcionando de este modo un subconjunto de células seleccionadas positivamente que contiene el depósito de RS activas (opcionalmente mutantes). Opcionalmente, la selección y/o la concentración del agente de selección se puede variar.

En un aspecto, el marcador de selección positiva es un gen de la acetiltransferasa de cloranfenicol (CAT) y el codón

selector es un codón de parada ámbar en el gen CAT. Opcionalmente, el marcador de selección positiva es un gen  $\beta$ -lactamasa y el codón selector es un codón de parada ámbar en el gen  $\beta$ -lactamasa. En otro aspecto el marcador de selección positiva comprende un marcador de selección fluorescente o luminiscente o un marcador de detección a base de afinidad (incluyendo pero no limitado a un marcador de superficie celular).

5 En un caso, la selección negativa o detección del depósito para RS activas (opcionalmente mutantes) que preferentemente aminoacilan el O-ARNt en ausencia del aminoácido no naturalmente codificado incluye: la introducción de una selección negativa o de detección de marcadores con el depósito de RS activas (opcionalmente mutantes) de la selección positiva o detección en una pluralidad de células de un segundo organismo, en el que la  
10 selección negativa o marcador de detección comprende al menos un codón selector (incluyendo pero no limitado a, un gen de resistencia a los antibióticos, incluyendo pero no limitado a, un gen de acetiltransferasa de cloranfenicol (CAT)); y, la identificación de las células que sobreviven o muestran una respuesta de detección específica en un primer medio complementado con el aminoácido no naturalmente codificado y un agente de detección o de  
15 selección, pero no sobreviven o no muestran la respuesta específica en un segundo medio no complementado con aminoácido no naturalmente codificado y el agente de selección o de detección, proporcionando de este modo las células supervivientes o células cribadas con al menos un O-RS recombinante. Por ejemplo, un protocolo de identificación CAT opcionalmente actúa como una selección positiva y/o una proyección negativa en la determinación de los O-RS recombinantes apropiados. Por ejemplo, un grupo de clones se replica opcionalmente en  
20 placas de crecimiento que contienen CAT (que comprende al menos un codón selector) ya sea con o sin uno o más aminoácidos no naturalmente codificados. Las colonias que crecen exclusivamente en las placas que contienen aminoácidos no naturalmente codificados se consideran por contener O-RS recombinantes. En un aspecto, la concentración del agente de selección (y/o detección) es variado. En algunos aspectos el primer y segundo organismos son diferentes. Por lo tanto, el primer y/o segundo organismo opcionalmente comprende: un procariota, un eucariota, un mamífero, una Escherichia coli, un hongo, una levadura, una arqueobacteria, una eubacteria, una  
25 planta, un insecto, un protista, etc. En otros casos, el marcador de selección comprende un marcador de selección fluorescente o luminiscente o un marcador de detección a base de afinidad.

En otro caso, la detección o la selección (incluyendo pero no limitado a, la selección negativa) del depósito para RS activo (opcionalmente mutante) incluye: el aislamiento del depósito de RS mutante activa de la etapa de selección  
30 positiva (b); la introducción de una selección negativa o de detección de marcadores, en la que la selección negativa o marcador de detección comprende al menos un codón selector (incluyendo pero no limitado a, un gen marcador tóxico, incluyendo, pero no limitado a, un gen de barnasa de ribonucleasa, que comprende al menos un codón selector), y el depósito de RS activa (opcionalmente mutante) en una pluralidad de células de un segundo organismo; y la identificación de las células que sobreviven o muestran una respuesta de detección específica en un  
35 primer medio no complementado con el aminoácido no naturalmente codificado, pero no sobreviven o muestran una respuesta de detección específica de un segundo medio complementado con el aminoácido no naturalmente codificado, proporcionando de ese modo células que sobreviven o se detectan con al menos un O-RS recombinante, en donde al menos una O-RS recombinante es específica para el aminoácido no naturalmente codificado. En un aspecto, al menos un codón selector comprende aproximadamente dos o más codones selectores. Tales casos  
40 pueden incluir opcionalmente, en el que al menos un codón selector comprende dos o más codones selectores, y en el que el primer y el segundo organismo son diferentes (incluyendo pero no limitado a, cada organismo es opcionalmente, incluyendo, pero no limitado a, una procariota, una eucariota, un mamífero, una Escherichia coli, un hongo, una levadura, arqueobacterias, eubacterias, una planta, un insecto, un protista, etc.). Además, algunos aspectos se incluyen en donde el marcador de selección negativa comprende un gen de la barnasa de ribonucleasa  
45 (que comprende al menos un codón selector). Otros aspectos se incluyen en donde el marcador de detección comprende opcionalmente un marcador de detección fluorescente o luminiscente o un marcador de detección a base de afinidad. En los casos del presente documento, las proyecciones y/o selecciones incluyen opcionalmente la variación de la proyección y/o rigurosidad de selección.

50 En un caso, los métodos para producir al menos una sintetasa (O-RS) de aminoacilo-ARNt ortogonal recombinante puede comprender además: (d) aislar al menos una O-RS recombinante; (e) generar un segundo conjunto de O-RS (opcionalmente mutado) derivado de al menos una O-RS recombinante; y, (f) repetir las etapas (b) y (c) hasta que se obtenga una O-RS mutada que comprende una capacidad para aminoacilar preferentemente el O-ARNt. Opcionalmente, las etapas (d)-(f) se repiten, incluyendo, pero no limitado a, al menos aproximadamente dos veces.  
55 En un aspecto, el segundo conjunto de O-RS mutadas derivadas de al menos una O-RS recombinante pueden generarse por mutagénesis, incluyendo, pero no limitado a, mutagénesis aleatoria, mutagénesis específica de sitio, recombinación o una combinación de los mismos.

60 La rigurosidad de las etapas de selección/detección, incluyendo, pero no limitado a, la selección positiva/detección de la etapa (b), la selección negativa/detección de la etapa (c) o las etapas de selección/detección tanto positivas como negativas (b) y (c), en los métodos anteriormente descritos, opcionalmente incluye variar la rigurosidad de selección/detección. En otro caso, la selección positiva/detección de la etapa (b), la selección negativa/detección de la etapa (c) o las etapas de selección/detección tanto positivas como negativas (b) y (c) comprenden el uso de un reportero, en el que se detecta el reportero mediante separación de células activadas por fluorescencia (FACS) o en  
65 el que el reportero se detecta mediante luminiscencia. Opcionalmente, el reportero se muestra en una superficie celular, en una presentación en fagos o similares y seleccionarse basándose en la afinidad o actividad catalítica que



implica el aminoácido no naturalmente codificado o un análogo. En un caso, la sintetasa mutada se muestra en una superficie celular, en una presentación de fagos o similares.

Los métodos para producir un ARNt ortogonal recombinante (O-ARNt) incluyen: (a) generar una biblioteca de ARNt mutantes derivadas de al menos un ARNt, incluyendo pero no limitado a, un supresor ARNt, de un primer organismo; (b) seleccionar (incluyendo pero no limitado a, selección negativa) o detección de la biblioteca para ARNt (opcionalmente mutantes) que se aminoacilan por una sintetasa de aminoacilo (RS) de un segundo organismo en ausencia de una RS desde el primer organismo, proporcionando de este modo un conjunto de ARNt (opcionalmente mutante); y, (c) selección o detección del depósito de ARNt (opcionalmente mutante) para los miembros que están aminoacilados por una RS ortogonal introducida (O-RS), proporcionando de este modo al menos un O-ARNt recombinante; en el que el al menos un O-ARNt recombinante reconoce un codón selector y no es la eficiencia reconocida por las RS desde el segundo organismo y está preferentemente aminoacilado por la O-RS. En algunos casos al menos un ARNt es un supresor de ARNt y/o comprende un único codón de tres bases naturales y/o bases no naturales, o es un codón sin sentido, 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 parada ópalo. En un caso, el O-ARNt recombinante posee una mejora de la ortogonalidad. Se apreciará que en algunos casos, O-ARNt se importa opcionalmente en un primer organismo a partir de un segundo organismo sin la necesidad de modificación. En diversos casos, el primer y segundo organismo son o bien el mismo o diferentes y se seleccionan opcionalmente de, incluyendo pero no limitados a, procariontes (incluyendo pero no limitado a, *Methanococcus jannaschii*, *Metanobacteria thermoautotrophicum*, *Escherichia coli*, *Halobacterium*, etc.), eucariotas, mamíferos, hongos, levaduras, archaeobacteria, eubacteria, plantas, insectos, protistas, etc. Además, el ARNt recombinante se aminoacila opcionalmente por un aminoácido no naturalmente codificado, en el que el aminoácido no naturalmente codificado se biosintetiza *in vivo* ya sea naturalmente o por medio de manipulación genética. El aminoácido no naturalmente codificado se añade opcionalmente a un medio de crecimiento durante al menos el primer o segundo organismo.

En un aspecto, la selección (incluyendo pero no limitado a, selección negativa) o la detección de la biblioteca para ARNt (opcionalmente mutantes) que se aminoacilan por una sintetasa de aminoacilo-ARNt (etapa (b)) incluye: la introducción de un gen marcador tóxico, en donde el gen marcador tóxico comprende al menos uno de los codones selectores (o un gen que conduce a la producción de un agente tóxico o estático o un gen esencial para el organismo en el que dicho gen marcador comprende al menos un codón selector) y la biblioteca de ARNt (opcionalmente mutantes) en una pluralidad de células del segundo organismo; y, la selección de células supervivientes, en las que las células supervivientes contienen el depósito de ARNt (opcionalmente mutantes) que comprenden al menos un ARNt ortogonal o ARNt no funcional. Por ejemplo, las células supervivientes se pueden seleccionar mediante el uso de un ensayo de densidad celular de relación de comparación.

En otro aspecto, el gen marcador tóxico puede incluir dos o más codones selectores. En otro ejemplo de los métodos, el gen marcador tóxico es un gen de barnasa de ribonucleasa, donde el gen de la barnasa de ribonucleasa comprende al menos un codón ámbar. Opcionalmente, el gen de la barnasa de ribonucleasa puede incluir dos o más codones ámbar.

En un caso, la selección o detección del depósito de ARNt (opcionalmente mutantes) para los miembros que están aminoacilados por una RS ortogonales introducidas (O-RS) pueden incluir: la introducción de una selección positiva o detección de gen marcador, en el que el gen marcador positivo comprende un gen de resistencia a los medicamentos (incluyendo pero no limitado al gen de lactamasa  $\square$ , que comprende al menos uno de los codones selectores, tal como al menos un codón de parada 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 depósito de ARNt (opcionalmente mutantes) en una pluralidad de células del segundo organismo; y, la identificación de las células que sobreviven o se criban cultivadas en presencia de un agente de selección o de detección, incluyendo, pero no limitado a, un antibiótico, proporcionando de este modo un conjunto de células que poseen el al menos un ARNt recombinante, donde se aminoacila al menos un ARNt recombinante por la O-RS y se inserta un aminoácido en un producto de traducción codificada por el gen marcador positivo, en respuesta a al menos un codón selector. En otro caso, la concentración del agente de selección y/o detección es variada.

Se proporcionan los métodos para generar pares O-ARNt/O-RS específicos. Los métodos incluyen: (a) generar una biblioteca de ARNt mutantes derivadas de al menos un ARNt a partir de un primer organismo; (B) seleccionar negativamente o detección de la biblioteca para ARNt (opcionalmente mutantes) que se aminoacilan por una sintetasa de aminoacilo (RS) de un segundo organismo en ausencia de una RS desde el primer organismo, proporcionando de ese modo un depósito de ARNt (opcionalmente mutantes); (c) seleccionar o detectar el depósito de ARNt (opcionalmente mutantes) para los miembros que están aminoacilados por una RS ortogonales introducidas (O-RS), proporcionando de este modo al menos un O-ARNt recombinante. Al menos un O-ARNt recombinante reconoce un codón selector y no es la eficiencia reconocida por RS desde el segundo organismo y está preferentemente aminoacilado por la O-RS. El método también incluye (d) la generación de una biblioteca de RS (opcionalmente mutante) derivada de al menos una sintetasa de aminoacilo (RS) de un tercer organismo; (e) la selección o detección de la biblioteca de RS mutantes para los miembros que preferentemente aminoacilan al menos un O-ARNt recombinante en presencia de un aminoácido no naturalmente codificado y un aminoácido

natural, proporcionando de ese modo un depósito de RS activa (opcionalmente mutante); y, (f) la selección o detección del depósito para RS activa (opcionalmente mutante) que preferentemente aminoacila al menos un O-ARNt recombinante en ausencia del aminoácido no naturalmente codificado, proporcionando de ese modo al menos un par determinado O-ARNt/O-RS negativamente, en donde al menos un par específico O-ARNt/O-RS comprende al menos un O-RS recombinante que es específico para el aminoácido no naturalmente codificado y al menos un O-ARNt recombinante. Pares específicos O-ARNt/O-RS producidos por los métodos están incluidos. Por ejemplo, el par específico O-ARNt/O-RS puede incluir, incluyendo pero no limitado a, un par muARNtTyr-mutTyrRS, tal como un par muARNtTyr-SS12TyrRS, un par muARNtLeu-mutLeuRS, un par muARNtThr-mutThrRS, un par muARNtGlu-mutGluRS, o similares. Además, tales métodos incluyen en donde el primer y tercer organismo son iguales (incluyendo, pero no limitado a, *Methanococcus jannaschii*).

Los métodos para seleccionar un par de sintetasa ARNt-ARNt ortogonal para su uso en un sistema de traducción *in vivo* de un segundo organismo también se incluyen en la presente descripción. Los métodos incluyen: la introducción de un gen marcador, una sintetasa ARNt y una sintetasa de aminoacilo-ARNt (RS) aisladas o derivadas de un primer organismo en un primer conjunto de células del segundo organismo; introducir el gen marcador y el ARNt en un conjunto de células duplicado de un segundo organismo; y, la selección para las células en el primer conjunto que no logran sobrevivir en el conjunto de células duplicado o detección de células que muestran una respuesta de detección específica que no dan tal respuesta en el conjunto de células por duplicado, en el que el primer conjunto y el conjunto de células duplicado cultivado en presencia de un agente de selección o de detección, en donde las células supervivientes o cribadas comprenden el par de sintetasa ARNt-ARNt ortogonal para su uso en el sistema de traducción *in vivo* del segundo organismo. En un caso, la comparación y la selección o detección incluye un ensayo de complementación *in vivo*. La concentración del agente de selección o detección puede ser variada.

Los organismos de la presente divulgación comprenden una variedad de organismo y una variedad de combinaciones. Por ejemplo, el primer y el segundo organismo de los métodos de la presente divulgación pueden ser iguales o diferentes. En un caso, los organismos son, opcionalmente, un organismo procarionta, incluyendo, pero no limitado a, *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, o similar. Alternativamente, los organismos comprenden opcionalmente un organismo eucariótico, incluyendo, pero no limitado a, las plantas (incluyendo pero no limitado a, las plantas complejas tales como monocotiledóneas, o dicotiledóneas), algas, protistas, hongos (incluyendo pero no limitado a, levadura, etc.), animales (incluyendo pero no limitado a, mamíferos, insectos, artrópodos, etc.), o similares. En otro ejemplo, el segundo organismo es un organismo procarionta, incluyendo, pero no limitado a, *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *Halobacterium*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, o similares. Alternativamente, el segundo organismo puede ser un organismo eucariótico, incluyendo, pero no limitado a, una levadura, una célula animal, una célula de la planta, un hongo, una célula de mamífero, o similares. En diversos casos el primer y segundo organismos son diferentes.

#### **Localización de los aminoácidos no naturales en polipéptidos de insulina**

La presente invención contempla la incorporación de uno o más aminoácidos no naturales en polipéptidos de insulina. Uno o más aminoácidos no naturales se pueden incorporar en una posición particular que no interrumpe la actividad del polipéptido. Esto se puede lograr haciendo sustituciones "conservadoras", incluyendo, pero no limitado a, la sustitución de aminoácidos hidrófobos por aminoácidos hidrofóbicos, aminoácidos voluminosos para los aminoácidos voluminosos, aminoácidos hidrófilos para los aminoácidos hidrofílicos y/o insertar el aminoácido no natural en una ubicación que no se requiere para la actividad. Polipéptidos de insulina de acuerdo con la invención comprenden un polímero soluble en agua unido por un enlace de oxima covalente a la cadena A de polipéptido de insulina en la posición 14 de la secuencia de aminoácidos de cadena A seleccionada del grupo SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, y SEQ ID NO: 11, en donde el aminoácido en la posición 14 de la cadena A es un aminoácido no codificado de forma natural.

Una variedad de enfoques bioquímicos y estructurales se puede emplear para seleccionar los sitios deseados para la sustitución con un aminoácido no naturalmente codificado dentro del polipéptido de insulina. Es fácilmente evidente para expertos ordinarios en la técnica que cualquier posición de la cadena de polipéptido es adecuado para la selección para incorporar un aminoácido no naturalmente codificado, y la selección puede estar basada en el diseño racional o por selección aleatoria por cualquiera o ninguna finalidad deseada particular. La selección de sitios deseados puede ser para la producción de una molécula de insulina que tiene cualquier propiedad o actividad deseada, incluyendo, pero no limitado a, agonistas, superagonistas, agonistas inversos, antagonistas, moduladores de unión a receptores, moduladores de la actividad del receptor, dímero o la formación de multímero, ningún cambio en la actividad o propiedad en comparación con la molécula nativa, o manipular cualquier propiedad física o química del polipéptido tales como la solubilidad, la agregación, o la estabilidad. Por ejemplo, las ubicaciones en el polipéptido necesarias para la actividad biológica de polipéptidos de insulina pueden identificarse utilizando el análisis de mutación puntual, barrido de alanina, mutagénesis de saturación y la detección de la actividad biológica, o métodos de exploración de homólogos conocidos en la técnica. Otros métodos pueden ser utilizados para identificar los residuos de modificación de polipéptidos de insulina incluyen, pero no se limitan a, la secuencia de perfiles (Bowie y Eisenberg, *Science* 253(5016): 164-70, (1991)), las selecciones de bibliotecas de rotámeros

(Dahiyat y Mayo, *Protein Sci* 5 (5): 895-903 (1996); Dahiyat y Mayo, *Science* 278 (5335): 82-7 (1997); Desjarlais y Handel, *Protein Science* 4: 2006-2018 (1995); Harbury et al., *PNAS EE.UU.* 92(18): 8408-8412 (1995); Kono et al., *Proteins: Structure, Function and Genetics* 19: 244-255 (1994); Hellinga y Richards, *PNAS EE.UU.* 91: 5803-5807 (1994)); y par de residuos potenciales (Jones, *Protein Science* 3: 567-574, (1994)), y el diseño racional usando la tecnología de Protein Design Automation®. (Véase la patente de Estados Unidos nº 6.188.965; 6.269.312; 6.403.312; WO98/47089). Los residuos que son críticos para la bioactividad de la insulina, los residuos que están involucrados con la estabilidad farmacéutica, epítopos de anticuerpo, o los residuos de unión al receptor pueden ser mutados. Patente de Estados Unidos nº 5.580.723; 5.834.250; 6.013.478; 6.428.954; y 6.451.561, describen métodos para el análisis sistemático de la estructura y función de los polipéptidos tales como la insulina mediante la identificación de dominios activos que influyen en la actividad del polipéptido con una sustancia diana. Residuos distintos de los identificados como críticos para la actividad biológica por alanina o mutagénesis de detección de homólogo pueden ser buenos candidatos para la sustitución con un aminoácido no naturalmente codificado en función de la actividad deseada para el polipéptido. Alternativamente, los sitios identificados como críticos para la actividad biológica también pueden ser buenos candidatos para la sustitución con un aminoácido no naturalmente codificado, dependiendo de nuevo de la actividad deseada buscada para el polipéptido. Otra alternativa consistiría únicamente en hacer sustituciones de serie en cada posición en la cadena de polipéptido con un aminoácido no naturalmente codificado y observar el efecto sobre las actividades del polipéptido. Es fácilmente evidente para expertos ordinarios en la técnica que cualquier medio, técnica o método para seleccionar una posición para la sustitución con un aminoácido no natural en cualquier polipéptido es adecuado para su uso en la presente invención.

La estructura y la actividad de mutantes de polipéptidos de insulina que contienen deleciones también pueden ser examinadas para determinar las regiones de la proteína que son susceptibles de ser tolerantes con la sustitución con un aminoácido no naturalmente codificado. De una manera similar, la digestión de la proteasa y los anticuerpos monoclonales se pueden utilizar para identificar regiones de la insulina que son responsables de la unión del receptor de insulina. Una vez que los residuos que son propensos a ser intolerantes a la sustitución con aminoácidos no naturalmente codificados han sido eliminados, el impacto de las sustituciones propuestas en cada una de las posiciones restantes puede ser examinada. Los modelos pueden ser generados a partir de las estructuras cristalinas tridimensionales de otros miembros de la familia de la insulina y los receptores de insulina. Protein Data Bank (PDB, disponible en el internet en [rcsb.org](http://rcsb.org)) es una base de datos centralizada que contenga datos estructurales tridimensionales de grandes moléculas de proteínas y ácidos nucleicos. Los modelos pueden hacerse mediante la investigación de la estructura secundaria y terciaria de los polipéptidos, si los datos estructurales tridimensionales no están disponibles. Por lo tanto, los expertos en la técnica pueden identificar fácilmente las posiciones de aminoácidos que pueden ser sustituidas con aminoácidos no naturalmente codificados.

En algunas realizaciones, los polipéptidos de insulina de la invención comprenden uno o más aminoácidos no naturalmente codificados situados en una región de la proteína que no interrumpe la estructura del polipéptido.

Residuos ejemplares de incorporación de un aminoácido no naturalmente codificado pueden ser aquellos que están excluidos de potenciales regiones de unión a receptor, pueden ser total o parcialmente expuestos al disolvente, tienen interacciones de enlaces de hidrógeno mínimas o ausentes con los residuos cercanos, pueden estar mínimamente expuestos a los residuos en las inmediaciones reactivas, pueden estar en una o más de las caras expuestas, puede ser un sitio o sitios que se yuxtaponen a una segunda insulina, u otra molécula o fragmento del mismo, puede estar en regiones que son altamente flexibles, o estructuralmente rígidas, tal como se predijo por la estructura tridimensional, secundaria, terciaria, o cuaternaria de la insulina, unida o no unida a su receptor, o acoplado o no acoplado a otra molécula biológicamente activa, o puede modular la conformación de la propia insulina o un dímero o multímero que comprende una o más insulina, por alteración de la flexibilidad o rigidez de la estructura completa como se desee.

Un experto en la técnica reconoce que este tipo de análisis de la insulina permite la determinación de qué residuos de aminoácidos son expuestos a la superficie en comparación con los residuos de aminoácidos que están enterrados dentro de la estructura terciaria de la proteína. Por lo tanto, es una forma de realización de la presente invención para sustituir un aminoácido no naturalmente codificado para un aminoácido que es un residuo expuesto en la superficie.

En algunas realizaciones, uno o más aminoácidos no naturalmente codificados también se incorporan en una o más de las siguientes posiciones en la insulina: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 1) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, (SEQ ID NO: 2).

Un examen de la estructura cristalina de la insulina y su interacción con el receptor de la insulina puede indicar qué residuos de aminoácidos tienen cadenas laterales que son totalmente o parcialmente accesibles al solvente. La cadena lateral de un aminoácido no naturalmente codificado en estas posiciones pueden apuntar hacia afuera de la superficie de la proteína y fuera del disolvente.

En algunas formas de realización, el aminoácido no naturalmente codificado en una o más de estas posiciones está unido a un polímero soluble en agua, incluyendo, pero no limitado a, las posiciones: en la cadena A antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína), y cualquier combinación de los mismos (SEQ ID NO: 1) o en la cadena B antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 (SEQ ID NO: 2).

Una amplia variedad de aminoácidos no naturalmente codificados puede ser sustituida por, o incorporada en, una posición dada en un polipéptido de insulina. En general, se selecciona un aminoácido no naturalmente codificado en particular para su incorporación sobre la base de un examen de la estructura cristalina tridimensional de un polipéptido de insulina u otro miembro de familia de la insulina o análogo de insulina con su receptor, una preferencia por sustituciones conservativas (es decir, aminoácidos no naturalmente codificados a base de arilo, tales como *p*-acetilo-feniloalanina o *O*-propargiltirosina sustituyendo Phe, Tyr o Trp), y la química de conjugación específica que se desea introducir en el polipéptido de insulina (por ejemplo, la introducción de 4-azidofeniloalanina si se quiere efectuar una cicloadición de Huisgen [3 + 2] con un polímero soluble en agua que lleva un resto alquino o una formación de enlace amida con un polímero soluble en agua que lleva un éster arilo que, a su vez, incorpora un resto fosfina).

En un caso, el método además incluye la incorporación el aminoácido no natural en la proteína, en el que el aminoácido no natural comprende un primer grupo reactivo; y poner en contacto la proteína con una molécula (incluyendo pero no limitado a, una etiqueta, un colorante, un polímero, un polímero soluble en agua, un derivado de polietilenglicol, un fotoentrecruzador, un radionucleida, un compuesto citotóxico, un fármaco, un marcador de afinidad, un marcador de fotoafinidad, un compuesto reactivo, una resina, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metal, un cofactor, un ácido graso, un carbohidrato, un polinucleótido, un ADN, un ARN, un polinucleótido antisentido, un sacárido, un dendrímero soluble en agua, una ciclodextrina, un ácido ribonucleico inhibitorio, un biomaterial, una nanopartícula, un marcador de espín, un fluoróforo, un resto que contiene metal, un resto radiactivo, un grupo funcional novela, un grupo que interactúa covalentemente o no covalentemente con otras moléculas, un resto excitable de radiación actínica, un resto fotoisomerizable, biotina, un derivado de la biotina, un análogo de biotina, un resto de incorporación de un átomo pesado, un grupo químicamente escindible, un grupo fotoescindible, una cadena lateral alargada, un azúcar de enlace de carbono, un agente redoxactivo, un tioácido de amino, un resto tóxico, un resto isotópicamente marcado, una sonda biofísica, un grupo fosforescente, un grupo quemiluminescente, un grupo denso de electrones, un grupo magnético, un grupo de intercalación, un cromofosforo, un agente de transferencia de energía, un agente biológicamente activo, un marcador detectable, una molécula pequeña, un punto cuántico, un nanotransmisor, un radionucleótido, un radiotransmisor, un agente de captación de neutrona, o cualquier combinación de lo anterior, o cualquier otro compuesto deseable o sustancia) que comprende un segundo grupo reactivo. El primer grupo reactivo reacciona con el segundo grupo reactivo para unir la molécula al aminoácido no natural a través de una cicloadición [3 + 2]. En una realización, el primer grupo reactivo es un resto alquino o azido y el segundo grupo reactivo es un azido o alquino. Por ejemplo, el primer grupo reactivo es el resto alquino (incluyendo pero no limitado a aminoácido no natural *p*-propargiloxifeniloalanina) y el segundo grupo reactivo es el resto azido. En otro ejemplo, el primer grupo reactivo es el resto azido (incluyendo pero no limitado a, en el aminoácido no natural *p*-azido-L-feniloalanina) y el segundo grupo reactivo es el resto alquino.

En algunos casos, la sustitución de aminoácido no naturalmente codificado se puede combinar con otras adiciones, sustituciones o deleciones en el polipéptido de insulina para afectar a otros rasgos biológicos del polipéptido de insulina. En algunos casos, las otras adiciones, sustituciones o deleciones pueden aumentar la estabilidad (incluyendo pero no limitado a, resistencia a la degradación proteolítica) del polipéptido de insulina o aumentar la afinidad del polipéptido de insulina para su receptor. En algunos casos, las otras adiciones, sustituciones o deleciones pueden aumentar la estabilidad farmacéutica del polipéptido de insulina. En algunos casos, las otras adiciones, sustituciones o deleciones pueden mejorar la actividad antiviral del polipéptido de insulina. En algunos casos, las otras adiciones, sustituciones o deleciones pueden aumentar la solubilidad (incluyendo pero no limitado a, cuando se expresa en *E. coli* u otras células huésped) del polipéptido de insulina. En algunas realizaciones, adiciones, sustituciones o deleciones pueden aumentar la solubilidad de polipéptido de insulina tras la expresión en *E. coli* u otras células huésped recombinantes. En algunas realizaciones, los sitios se seleccionan para la sustitución con un aminoácido codificado de forma natural o no natural, además de otro sitio para la incorporación de un aminoácido no natural que resulta en el aumento de la solubilidad del polipéptido tras la expresión en *E. coli* u otras células huésped recombinantes. En algunas realizaciones, los polipéptidos de insulina comprenden otra adición, sustitución o deleción que modula la afinidad por el receptor polipéptido de insulina, proteínas, o ligando asociado de unión, modula la transducción de señal después de la unión al receptor de insulina, modula la vida media circulante, modula la liberación o biodisponibilidad, facilita la purificación, o mejora o altera una ruta particular de administración. En algunas realizaciones, los polipéptidos de insulina comprenden una adición, sustitución o deleción que aumentan la afinidad de la variante de insulina para su receptor. Del mismo modo, los polipéptidos de insulina pueden comprender secuencias de escisión química o enzima, secuencias de escisión de la proteasa, grupos reactivos, dominios de enlace de anticuerpos (incluyendo pero no limitado a, FLAG o poli-His) u otras secuencias basadas en afinidad (incluyendo, pero no limitado a, FLAG, poli-His, GST, etc.) o moléculas enlazadas (incluyendo, pero no limitado a, biotina) que mejoran la detección (incluyendo, pero no limitado a, GFP), purificación,

transporte a través de los tejidos o membranas celulares, liberación de profármaco o la activación, reducción del tamaño de la insulina, u otros rasgos del polipéptido.

En algunas realizaciones, la sustitución de un aminoácido no naturalmente codificado genera un antagonista de la insulina. En algunas realizaciones, un aminoácido no naturalmente codificado está sustituido o añadido en una región implicada con la unión al receptor. En algunas realizaciones, los antagonistas de insulina comprenden al menos una sustitución que causa la insulina para actuar como un antagonista. En algunas realizaciones, el antagonista de la insulina comprende un aminoácido no naturalmente codificado unido a un polímero soluble en agua que está presente en una región de unión al receptor de la molécula de insulina.

En algunos casos, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, o más aminoácidos están sustituidos con uno o más aminoácidos no naturalmente codificados. En algunos casos, el polipéptido de insulina incluye además 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, o más sustituciones de uno o más aminoácidos no naturalmente codificados para aminoácidos naturales. Por ejemplo, en algunas realizaciones, uno o más residuos en la insulina están sustituidos con uno o más aminoácidos no naturalmente codificados. En algunos casos, uno o más residuos no naturalmente codificados están enlazados a uno o menos PEGs de peso molecular lineal o ramificado, mejorando de este modo la afinidad de unión y vida media comparable en suero en relación con las especies unidas a un solo PEG de peso molecular superior.

En algunas realizaciones, hasta dos de los siguientes residuos de insulina están sustituidos con uno o más aminoácidos no naturalmente codificados.

#### **Expresión en no eucariotas y eucariotas**

Para obtener expresión de alto nivel de un polinucleótido de insulina clonado, típicamente se subclona polinucleótidos que codifican un polipéptido de insulina de la invención en un vector de expresión que contiene un promotor fuerte para dirigir la transcripción, un terminador de transcripción/traducción, y si para un ácido nucleico que codifica una proteína, un sitio de unión al ribosoma para la iniciación de la traducción. Promotores bacterianos adecuados son conocidos por personas de experiencia ordinaria en la técnica y se describen, por ejemplo, en Sambrook *et al.* y Ausubel *et al.*

Sistemas de expresión bacteriana para expresar polipéptidos de insulina de la invención están disponibles, incluyendo pero no limitado a, *E. coli*, *Bacillus sp.*, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, y *Salmonella* (Palva *et al.*, *Gene* 22: 229-235 (1983); Mosbach *et al.*, *Nature* 302: 543-545 (1983)). Kits para tales sistemas de expresión están disponibles comercialmente. Sistemas de expresión eucariotas para células de mamíferos, levaduras, y células de insectos son conocidos por personas de experiencia ordinaria en la técnica y también están disponibles comercialmente. En los casos en los que sintetasas de ARNt ortogonal y ARNt de aminoácido (descritas anteriormente) se utilizan para expresar los polipéptidos de insulina de la invención, las células huésped para la expresión se seleccionan basándose en su capacidad de usar los componentes ortogonales. Células huésped ejemplares incluyen bacterias gram positivas (incluyendo pero no limitado a *B. brevis*, *B. subtilis*, o *Streptomyces*) y las bacterias gram negativas (*E. coli*, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa*, *Pseudomonas putida*), así como la levadura y otras células eucariotas. Las células que comprenden pares de O-ARNt/O-RS se pueden utilizar como se describe aquí.

Una célula huésped eucariota o célula huésped no eucariótica de la presente descripción proporciona la capacidad de sintetizar proteínas que comprenden aminoácidos no naturales en grandes cantidades útiles. En un aspecto, la composición incluye opcionalmente, incluyendo, pero no limitado 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 la proteína que comprende un aminoácido no natural, o una cantidad que se puede conseguir con métodos *in vivo* de producción de proteínas (detalles en producción de proteína recombinante y purificación se proporcionan en el presente documento). En otro aspecto, la proteína está opcionalmente presente en la composición a una concentración de, incluyendo pero no limitado a, al menos 10 microgramos de proteína por litro, al menos 50 microgramos de proteína por litro, al menos 75 microgramos de proteína por litro, al menos 100 microgramos de proteína por litro, al menos 200 microgramos de proteína por litro, al menos 250 microgramos de proteína por litro, al menos 500 microgramos de proteína por litro, al menos 1 miligramo de proteína por litro, o al menos 10 miligramos de proteína por litro o más, incluyendo pero no limitado a, un lisado celular, un tampón, un tampón farmacéutico u otra suspensión líquida (incluyendo pero no limitado a, en un volumen de, incluyendo pero no limitado a, en cualquier lugar de aproximadamente 1 ml a aproximadamente 100 L o más). La producción de grandes cantidades (incluyendo pero no limitado a, mayor que generalmente posible con otros métodos, incluyendo, pero no limitado a, la traducción *in vitro*) de una proteína en una célula eucariota que incluye al menos un aminoácido no natural es una característica de la revelación.

Una célula huésped eucariota o célula huésped no eucariótica de la presente descripción proporciona la capacidad de biosintetizar proteínas que comprenden aminoácidos no naturales en grandes cantidades útiles. Por ejemplo, las proteínas que comprenden un aminoácido no natural se pueden producir a una concentración de, incluyendo pero no limitados a, 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 de células, lisado de células, medio de cultivo, un tampón, y/o similares.

Un número de vectores adecuados para la expresión de insulina están disponibles comercialmente. Vectores de expresión útiles para huéspedes eucarióticos, incluyen, pero no se limitan a, vectores que comprenden secuencias de control de la expresión de SV40, virus del papiloma bovino, adenovirus y citomegalovirus. Tales vectores incluyen pCADN3,1(+)\Hyg (Invitrogen, Carlsbad, Calif., EE.UU.) y pCI-neo (Stratagene, La Jolla, Calif., EE.UU.). Plásmidos bacterianos, tales como plásmidos de *E. coli*, incluyendo pBR322, pET3a y pET12a, plásmidos de huésped más amplios, tales como RP4, ADN de fagos, por ejemplo, los numerosos derivados del fago lambda, por ejemplo, NM989, y otros fagos ADN, tales como M13 y fagos de ADN de cadena sencilla filamentosa pueden utilizarse. El plásmido 2µ y derivados de los mismos, el vector POT1 (Pat. nº 4,931,373), el vector pJSO37 descrito en (Okkels, Ann. Nueva York Acad. Sci. 782, 202-207, 1996) y pPICZ A, B o C (Invitrogen) se puede usar con células huésped de levadura. Para células de insecto, los vectores incluyen, pero no se limitan a, pVL941, pBG311 (Cate *et al.*, "Isolation of Bovine and Human Genes for Mullerian Inhibiting Substance and Expression of the Human Gene in Animal Cells", Cell, 45, pp. 685-98 (1986), pBluebac 4.5 y pMelbac (Invitrogen, Carlsbad, CA).

La secuencia de nucleótidos que codifica un polipéptido de insulina puede o no puede también incluir una secuencia que codifica un péptido de señal. El péptido de señal está presente cuando el polipéptido ha de secretarse de las células en las que se expresa. Tal péptido de señal puede ser cualquier secuencia. El péptido de señal puede ser procariota o eucariota. Coloma, M (1992) J. Imm. Methods 152: 89-104 describen un péptido de señal para su uso en células de mamífero (péptido de señal de cadena ligera Ig kappa murino). Otros péptidos de señal incluyen, pero no se limitan al péptido de señal α-factor de *S. cerevisiae* (patente de EE.UU. nº 4.870.008), el péptido de señal de la amilasa salival del ratón (O. Hagenbuchle *et al.*, Nature 289, 1981, pp. 643-646), un péptido de señal de carboxipeptidasa modificada (L. A. Valls *et al.*, Cell 48, 1987, pp. 887-897), el péptido de señal BAR1 de levadura (WO 87/02670), y el péptido de señal de levadura de la proteasa aspártica 3 (YAP3) (cf. M. Egel-Mitani *et al.*, Yeast 6, 1990, pp. 127-137).

Ejemplos de células huésped de mamífero adecuados son conocidos por personas de experiencia ordinaria en la técnica. Tales células huésped pueden ser células de ovario de hámster chino (CHO), (por ejemplo CHO-K1; ATCC CCL-61), células de mono verde (COS) (por ejemplo, COS 1 (ATCC CRL-1650), COS 7 (ATCC CRL-1651)); células de ratón (por ejemplo, NS/O), líneas celulares de riñón de hámster pequeño (BHK) (por ejemplo ATCC CRL-1632 o ATCC CCL-10), y células humanas (por ejemplo, HEK 293 (ATCC CRL-1573)), así como células vegetales en cultivo de tejidos. Estas líneas celulares y otras están disponibles de depositarios públicos tales como la American Type Culture Collection, Rockville, Md. Con el fin de proporcionar una mejor glicosilación del polipéptido de insulina, una célula huésped de mamífero puede ser modificada para expresar la sialiltransferasa, por ejemplo 1,6-sialiltransferasa, por ejemplo como se describe en la patente de los Estados Unidos. nº 5.047.335.

Los métodos para la introducción de ADN exógeno en células huésped de mamífero incluyen, pero no se limitan a, transfección mediada por calcio fosfato, electroporación, transfección mediada por DEAE-dextrano, transfección mediada por liposoma, vectores virales y métodos de transfección descritos por Life Technologies Ltd, Paisley, Reino Unido usando Lipofectamine 2000 y Roche Diagnostics Corporation, Indianapolis, EE.UU. utilizando FuGENE 6. Estos métodos son bien conocidos en la técnica y se describen por Ausbel *et al.* (eds.), 1996, Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, Estados Unidos. El cultivo de células de mamífero se puede realizar de acuerdo con métodos establecidos, por ejemplo como se describe en (Animal Cell Biotechnology, Methods and Protocols, editado por Nigel Jenkins, 1999, Human Press Inc. Totowa, NJ., EE.UU. y Harrison Mass. y Rae IF, General Techniques of Cell Culture, Cambridge University Press, 1997).

### **Sistemas de expresión, cultivo, y aislamiento**

Polipéptidos de insulina pueden ser expresados en cualquier número de sistemas de expresión adecuados incluyen, por ejemplo, levadura, células de insecto, células de mamífero y bacterias. Una descripción de los sistemas de expresión ejemplares se proporciona a continuación.

**Levadura** Tal como se usa aquí, el término "levadura" incluye cualquiera de las diversas levaduras capaces de expresar un gen que codifica un polipéptido de insulina. Tales levaduras incluyen, pero no se limitan a, levaduras ascosporógenas (Endomicetales), levaduras basidiosporógenas y levaduras pertenecientes al grupo de hongos imperfectos (Blastomicetos). Las levaduras ascosporógenas se dividen en dos familias, Spermophthoraceae y Saccharomycetaceae. Esta última está compuesta por cuatro subfamilias, Schizosaccharomycoidae (por ejemplo, género Schizosaccharomyces), Nadsonioideae, Lipomycoideae y Saccharomycoidae (por ejemplo, géneros Pichia, Kluyveromyces y Saccharomyces). Las levaduras basidiosporógenas incluyen los géneros Leucosporidium, Rhodosporidium, Sporidiobolus, Filobasidium, y Filobasidiella. Las levaduras que pertenecen al grupo de hongos imperfectos (Blastomicetos) se dividen en dos familias, Sporobolomycetaceae (por ejemplo, géneros Sporobolomyces y Bullera) y Cryptococcaceae (por ejemplo, género Candida).

De particular interés para su uso con la presente descripción son especies dentro de los géneros *Pichia*, *Kluyveromyces*, *Saccharomyces*, *Schizosaccharomyces*, *Hansenula*, *Torulopsis* y *Candida*, incluyendo, pero no limitado a, *P. pastoris*, *P. guilliermondii*, *S. cerevisiae*, *S. carlsbergensis*, *S. diastolicus*, *S. douglasii*, *S. kluyveri*, *S. norbensis*, *S. oviformis*, *K. lactis*, *K. fragilis*, *C. albicans*, *C. maltosa*, y *H. polimorfa*.

La selección de la levadura adecuada para la expresión de polipéptidos de insulina está dentro del conocimiento de un experto normal en la técnica. En la selección de huéspedes de levadura para la expresión, hospedantes adecuados pueden incluir los que demuestran tener, por ejemplo, buena capacidad de secreción, baja actividad proteolítica, buena capacidad de secreción, buena producción de proteína soluble, y la robustez general. Levaduras son generalmente disponibles de una variedad de fuentes, incluyendo, pero no limitado a, Yeast Genetic Stock Center, Departamento de Biofísica y Física Médica de la Universidad de California (Berkeley, CA), y de la American Type Culture Collection ("ATCC")(Manassas, VA).

El término "huésped de levadura" o "célula huésped de levadura" incluye levadura que puede ser, o ha sido, usada 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 al progenitor original, debido a mutación accidental o deliberada. La progenie de la célula parental que es suficientemente similar al padre a caracterizarse por la propiedad relevante, tales como la presencia de una secuencia de nucleótidos que codifica un polipéptido de insulina, están incluidos en la progenie prevista por esta definición.

Vectores de expresión y transformación, incluyendo replicones extracromosómicos o vectores de integración, se han desarrollado para la transformación en muchas huéspedes de levadura. Por ejemplo, los vectores de expresión se han desarrollado para *S. cerevisiae* (Sikorski et al., GENETICS (1989) 122: 19; Ito et al., J. BACTERIOL (1983) 153: 163; Hinnen et al, PROC. NATL. ACAD. SCI. EE.UU. (1978) 75: 1929); *C. albicans* (Kurtz et al., MOL. CELL. BIOL. (1986) 6: 142); *C. maltosa* (Kunze et al., J. BASIC MICROBIOL (1985) 25: 141); *H. polimorfa* (Gleeson et al., J. GEN MICROBIOL (1986) 132: 3459; Roggenkamp et al., MOL GENETICS AND GENOMICS (1986) 202: 302); *K. fragilis* (Das et al., J. BACTERIOL (1984) 158: 1165); *K. lactis* (De Louvencourt et al., J. BACTERIOL (1983) 154: 737; Van den Berg et al., BIOTECHNOLOGY (NY) (1990) 8: 135); *P. guilliermondii* (Kunze et al., J. BASIC MICROBIOL (1985) 25: 141); *P. pastoris* (patente de Estados Unidos nº 5.324.639; 4.929.555; y 4.837.148; Cregg et al., MOL. CELL. BIOL. (1985) 5: 3376); *Schizosaccharomyces pombe* (Beach et al., NATURE (1982) 300: 706); y *Y. lipolitica*; *A. nidulans* (Ballance y otros, BIOCHEM BIOPHYS RES COMMUN (1983) 112: 284-89; Tilburn y otros, GENE (1983) 26: 205-221; y Yelton et al., PROC NATL ACAD SCI. EE.UU. (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*, *Toly pocladium* (WO 91/00357).

Secuencias de control para vectores de levaduras son conocidas por personas de experiencia ordinaria en la técnica e incluyen, pero no se limitan a, regiones promotoras de genes como deshidrogenasa de alcohol (ADH) (EP 0 284 044); enolasa; glucoquinasa; glucosa-6-isomerasa de fosfato; gliceraldehído-3-fosfato-dehidrogenasa (GAP o GAPDH); hexoquinasa; fosfofructoquinasa; 3-fosfoglicerato mutasa; y quinasa de piruvato (PyK) (EP 0 329 203). El gen de levadura PHO5, que codifica la fosfatasa ácida, también puede proporcionar secuencias promotoras útiles (Miyano-hara et al., PROC. NATL. ACAD. SCI. EE.UU. (1983) 80: 1). Otras secuencias promotoras adecuadas para su uso con huéspedes de levadura pueden incluir los promotores para la quinasa 3-fosfoglicerato (Hitzeman et al., J. BIOL CHEM (1980) 255: 12073); y otras enzimas glicolíticas, tales como descarboxilasa de piruvato, isomerasa de triosafofosfato, e isomerasa de fosfoglucosa (Holland et al., bioquímica (1978) 17: 4900; Hess et al., J. ADV ENZYME REG (1969) 7: 149). Promotores de levadura inducibles que tienen la ventaja adicional de transcripción controlada por las condiciones de crecimiento pueden incluir las regiones promotoras para deshidrogenasa de alcohol 2; isocitocromo C; fosfatasa ácida; metalotioneína; gliceraldehído-3-fosfato deshidrogenasa; enzimas degradativas asociadas con el metabolismo del nitrógeno; y enzimas responsables de la utilización de maltosa y galactosa. Vectores y promotores adecuados para uso en la expresión en levaduras se describen adicionalmente en EP 0 073 657.

Potenciadores de levadura también se pueden usar con promotores de levadura. En adición, promotores sintéticos también pueden funcionar como promotores de levaduras. Por ejemplo, las secuencias de activación aguas arriba (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 unida a la región de activación de la transcripción de GAP. Véase la patente de los Estados Unidos Nos. 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, combinados con la región de activación transcripcional de un gen de enzima glicolítica tal como GAP o PyK. Véase el documento EP 0 164 556. Además, un promotor de levadura puede incluir promotores de origen natural de origen no de levadura que tienen la capacidad de enlazar polimerasa de ARN de levadura e inician la transcripción.

Otros elementos de control que pueden comprender parte de los vectores de expresión de levadura incluyen

terminadores, por ejemplo, de GAPDH o los genes de enolasa (Holland et al., J. BIOL CHEM (1981) 256: 1385). Además, el origen de replicación del origen del plásmido  $2\mu$  es adecuado para la levadura. Un gen de selección adecuado para uso en levaduras es el gen *trp1* presente en el plásmido de levadura. Véase Tschumper et al., GENE (1980) 10: 157; Kingsman et al., GENE (1979) 7: 141. El gen *trp1* proporciona un marcador de selección para una cepa mutante de levadura que carece de la capacidad de crecer en triptófano. De manera similar, las cepas de levadura deficientes en *Leu2* (ATCC 20.622 o 38.626) se complementan mediante plásmidos conocidos que portan el gen *Leu2*.

Los procedimientos para introducir ADN exógeno en huéspedes de levadura son conocidos por personas de experiencia ordinaria en la técnica, y típicamente incluyen, pero no se limitan a, ya sea la transformación de esferoplastos o de células huésped de levadura intactas tratadas con cationes alcalinos. Por ejemplo, la transformación de la levadura se puede llevar a cabo de acuerdo con el método descrito en Hsiao et al., Proc. NATL. ACAD. SCI. Estados Unidos (1979) 76: 3829 y Van Solingen et al., J. BACT. (1977) 130: 946. Sin embargo, otros procedimientos para introducir ADN en células tales como la inyección nuclear, electroporación o fusión de protoplastos también se pueden usar como se describe generalmente en SAMBROOK ET AL., MOLECULAR CLONING: A LAB. MANUAL (2001). Células huésped de levadura pueden entonces ser cultivadas utilizando técnicas estándar conocidas por los expertos normales en la técnica.

Otros métodos para la expresión de proteínas heterólogas en células huésped de levadura son conocidos por personas de experiencia ordinaria en la técnica. Véase en general la Publicación de Patente de Estados Unidos nº 20020055169, Patente de Estados Unidos 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; Patente de Estados Unidos reexaminada 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. Véase también Gellissen et al., ANTONIE VAN LEEUWENHOEK (1992) 62(1-2):79-93; Romanos et al., YEAST (1992) 8(6): 423-488; Goeddel, METHODS IN ENZYMOLOGY (1990) 185: 3-7.

Las cepas huésped de levadura pueden cultivarse en fermentadores durante la etapa de amplificación utilizando métodos de fermentación por lotes de alimentación estándar conocidos por personas de experiencia ordinaria en la técnica. Los métodos de fermentación pueden adaptarse para tener en cuenta las diferencias en la ruta de utilización de carbono de un huésped de levadura en particular o modo de control de la expresión. Por ejemplo, la fermentación de un huésped de levadura *Saccharomyces* puede requerir una única alimentación de glucosa, la fuente de nitrógeno compleja (por ejemplo, hidrolizados de caseína), y los suplementos de vitaminas múltiples. Por el contrario, la levadura metilotrófica *P. pastoris* puede requerir glicerol, metanol, y piensos minerales, pero sólo sales de amoníaco sencillo (nitrógeno) para el crecimiento y la expresión óptima. Véase, por ejemplo, la Patente de Estados Unidos nº 5.324.639; Elliott et al., J. PROTEIN CHEM. (1990) 9:95; y Fieschko et al., BIOTECH. BIOENG. (1987) 29: 1113.

Tales métodos de fermentación, sin embargo, pueden tener ciertas características comunes independientes de la cepa huésped de levadura empleada. Por ejemplo, un crecimiento nutriente limitante, típicamente de carbono, se puede añadir al fermentador durante la fase de amplificación para permitir el crecimiento máximo. Además, los métodos de fermentación generalmente emplean un medio de fermentación diseñado para contener cantidades adecuadas de carbono, nitrógeno, sales basales, fósforo, y otros nutrientes menores (vitaminas, minerales y sales, etc.). Ejemplos de medios de fermentación adecuados para uso con *Pichia* se describen en la Patente de los Estados Unidos Nos. 5.324.639 y 5.231,178.

**Células de insectos infectados con Baculovirus** El término "huésped de insecto" o "célula huésped de insecto" se refiere a un insecto que puede ser, o ha sido, usado como receptor para vectores recombinantes u otro ADN de transferencia. El término incluye la progenie de la célula huésped de insecto original que ha sido transfectada. 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 al progenitor original, debido a mutación accidental o deliberada. La progenie de la célula parental que es suficientemente similar al padre a caracterizarse por la propiedad relevante, tales como la presencia de una secuencia de nucleótidos que codifica un polipéptido de insulina, están incluidos en la progenie prevista por esta definición. La expresión de Baculovirus de polipéptidos de insulina es útil en la presente invención y el uso de la tecnología de ADNr, polipéptidos o precursores de los mismos porque la insulina puede ser biosintetizada en cualquier número de células huésped incluyen bacterias, células de mamífero, células de insectos, levaduras u hongos. Una instancia de la presente descripción incluye la biosíntesis de la insulina, la insulina modificada, los polipéptidos de insulina, o análogos de insulina en bacterias, levaduras o células de mamífero. Otro ejemplo de la presente descripción implica la biosíntesis realizada en *E. coli* o una levadura. Ejemplos de biosíntesis en células de mamífero y animales transgénicos se describen en Hakola, K. [Molecular and Cellular Endocrinology, 127: 59-69, (1997)].

La selección de células de insecto adecuadas para la expresión de polipéptidos de insulina es conocida por los expertos ordinarios en la técnica. Varias especies de insectos están descritas en la técnica y están comercialmente disponibles, incluyendo *Aedes aegypti*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, y *Trichoplusia ni*. En la selección de insectos hospedadores para la expresión, hospedantes adecuados pueden incluir



los que se muestran a tener, entre otras cosas, buena capacidad de secreción, baja actividad proteolítica, y la robustez general. Insectos están generalmente disponibles de una variedad de fuentes, incluyendo, pero no limitado a Insect Genetic Stock Center, Departamento de Biofísica y Física Médica de la Universidad de California (Berkeley, CA); y la American Type Culture Collection ("ATCC") (Manassas, VA).

En general, los componentes de un sistema de expresión de insecto infectado por 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 se expresa; un baculovirus de tipo silvestre con secuencias homólogas al fragmento específico al baculovirus en el vector de transferencia (esto permite la recombinación homóloga del gen heterólogo en el genoma del baculovirus); y células huésped de insecto apropiadas y medio de crecimiento. Los materiales, métodos y técnicas utilizadas en la construcción de vectores, transfección de células, recogida de placas, las células que crecen en cultivo, y similares son conocidos en la técnica y están disponibles manuales que describen estas técnicas.

Después de insertar el gen heterólogo en el vector de transferencia, el vector y el genoma viral de tipo silvestre se transfectan en una célula huésped de insecto donde se recombinan el vector y el genoma viral. El virus recombinante empaquetado se expresa y se identifican y purifican placas recombinantes. Materiales y métodos para sistemas de expresión de baculovirus/células de insecto están disponibles comercialmente en forma de kit de, por ejemplo, Invitrogen Corp. (Carlsbad, CA). Estas técnicas son generalmente conocidas por los expertos ordinarios en la técnica y se describen completamente en SUMMERS AND SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN NO. 1555 (1987). Véase también, RICHARDSON, 39 METHODS IN MOLECULAR BIOLOGY: BACULOVIRUS EXPRESSION PROTOCOLS (1995); AUSUBEL ET AL., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY 16.916.11 (1994).; KING AND POSSEE, THE BACULOVIRUS SYSTEM: A LABORATORY GUIDE (1992); y O'REILLY ET AL., BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992).

En efecto, la producción de diversas proteínas heterólogas usando sistemas de expresión de baculovirus/células de insectos se conoce por los expertos ordinarios en la técnica. Véase, por ejemplo, la Patente de Estados Unidos 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; 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.

Los vectores que son útiles en sistemas de expresión de baculovirus/células de insectos son conocidos en la técnica e incluyen, por ejemplo, la expresión de insectos y vectores de transferencia derivados del baculovirus virus de polihedrosis nuclear de autographacalifomica (AcNPV), que es un vector de expresión viral independiente de auxiliares. Vectores de expresión virales derivados de este sistema normalmente utilizan el potente promotor del gen de polihedrina viral para conducir la expresión de genes heterólogos. Véase, en general, O'Reilly ET AL., BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992).

Antes de insertar el gen extraño en el genoma de baculovirus, los componentes anteriormente descritos, que comprenden un promotor, líder (si se desea), secuencia de interés, y secuencia de terminación de la transcripción, son típicamente montados en una construcción de transposición intermedia (vector de transferencia). Las construcciones de sustitución intermedias se mantienen a menudo en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos) capaz de mantenimiento estable en un huésped, tal como las bacterias. El replicón tendrá un sistema de replicación, permitiendo de este modo que se mantenga en un huésped adecuado para clonación y amplificación. Más específicamente, el plásmido puede contener la señal de poliadenilación de polihedrina (Miller, ANN REV MICROBIOL (1988) 42: 177) y un gen procariota de resistencia a ampicilina (amp) y origen de replicación para selección y propagación en *E. coli*.

Un vector de transferencia comúnmente utilizado para introducir genes extraños en AcNPV es pAc373. Muchos otros vectores, conocidos por los expertos en la técnica, también se han diseñado incluyendo, por ejemplo, pVL985, que alteran el codón de inicio de polihedrina de ATG a ATT, y que introduce un sitio de clonación de BamHI 32 pares de bases aguas abajo del ATT. Véase Luckow y Summers, VIROLOGY 170: 31 (1989). Otros vectores comercialmente disponibles incluyen, por ejemplo, pBlueBaC4.5/V5His; pBlueBacHis2; pMelBac; pBlueBaC4.5 (Invitrogen Corp., Carlsbad, CA).

Después de la inserción del gen heterólogo, el vector de transferencia y el genoma de baculovirus de tipo silvestre se cotransfectan en una célula huésped de insecto. Los métodos para introducir ADN heterólogo en el sitio deseado en el virus baculovirus son conocidos en la técnica. Véase SUMMERS Y SMITH, TEXAS AGRICULTURAL EXPERIMENTAL STATION BULLETIN NO. 1555 (1987); Smith *et al.*, MOL. CELL. BIOL. (1983) 3:2156; Luckow y Summers, VIROLOGY (1989) 170: 31. Por ejemplo, la inserción puede ser en un gen tal como el gen de polihedrina, mediante recombinación homóloga de doble entrecruzamiento; la inserción también puede ser en un sitio de enzima de restricción por ingeniería genética en el gen de baculovirus deseado. Véase Miller *et al.*, BIOESSAYS (1989)

11(4):91.

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

Vectores de expresión de baculovirus contienen habitualmente un promotor de baculovirus. Un promotor de baculovirus es cualquier secuencia de ADN capaz de unir una polimerasa de ARN 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 normalmente se coloca proximal al extremo 5' de la secuencia de codificación. Esta región de iniciación de la transcripción incluye típicamente un sitio de unión de polimerasa de ARN 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, es usualmente distal al gen estructural. Por otra parte, la expresión puede ser regulada o constitutiva.

Genes estructurales, abundantemente transcritos en momentos tardíos en el ciclo de infección, proporcionan secuencias promotoras particularmente útiles. Los ejemplos incluyen secuencias derivadas del gen que codifica la proteína poliédrica vírica (FRIESEN ET AL., *The Regulation of Baculovirus Gene Expression* en THE MOLECULAR BIOLOGY OF BACULOVIRUSES (1986); EP 0 127 839 y 0 155 476) y el gen que codifica la proteína p10 (Vlak et al., J. GEN. VIROL (1988) 69:765).

El vector de expresión de baculovirus recién formado se empaqueta en un baculovirus recombinante infeccioso y placas posteriormente cultivadas se pueden purificar por técnicas conocidas por personas de experiencia ordinaria en la técnica. Véase Miller et al., BIOESSAYS (1989) 11 (4): 91; SUMMERS AND SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN NO. 1555 (1987).

Vectores de expresión de baculovirus recombinantes se han desarrollado para la infección en varias células de insecto. Por ejemplo, los baculovirus recombinantes se han desarrollado para, *inter alia*, *Aedes aegypti* (ATCC n° CCL-125), *Bombyx mori* (ATCC n° CRL-8910), *Drosophila melanogaster* (ATCC n° 1963), *Spodoptera frugiperda*, y *Trichoplusia ni*. Véase Wright, NATURE (1986) 321:718; Carbonell *et al.*, J. VIROL. (1985) 56: 153; Smith *et al.*, MOL. CELL. BIOL. (1983) 3:2156. Véase, en general, Frasar *et al.*, IN VITRO CELL. DEV. BIOL. (1989) 25: 225. Más específicamente, las líneas celulares utilizadas para sistemas de vectores de expresión de baculovirus comúnmente incluyen, pero no se limitan a, Sf9 (*Spodoptera frugiperda*) (ATCC n° CRL-1711), Sf21 (*Spodoptera frugiperda*) (Invitrogen Corp., Cat. n° 11497-013 (Carlsbad, CA)), Tri-368 (*Trichoplusia ni*), y HighFive™ BTI-TN-5B1-4 (*Trichoplusia ni*).

Las células y los medios de cultivo están comercialmente disponibles para la expresión tanto directa como de fusión de polipéptidos heterólogos en un baculovirus/expresión, y la tecnología de cultivo celular es generalmente conocida por los expertos normales en la técnica.

**E. Coli, especies Pseudomonas y otras Procariotas** Técnicas de expresión bacteriana son conocidas por personas de experiencia ordinaria en la técnica. Una amplia variedad de vectores están disponibles para su uso en huéspedes bacterianas. Los vectores pueden ser de copia única o vectores de multicopia bajas o altas. Los vectores pueden servir para la clonación y/o expresión. En vista de la amplia literatura sobre vectores, la disponibilidad comercial de muchos vectores, e incluso los manuales que describen vectores y sus mapas de restricción y características, no se requiere una extensa discusión aquí. Como es bien conocido, los vectores normalmente implican marcadores que permiten la selección, cuyos marcadores pueden proporcionar la resistencia de agente citotóxico, prototrofia o inmunidad. Con frecuencia, una pluralidad de marcadores está presente, que prevén diferentes características.

Un promotor bacteriano es cualquier secuencia de ADN capaz de unión de polimerasa de ARN 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 normalmente se coloca proximal al extremo 5' de la secuencia codificante. Esta región de iniciación de la transcripción incluye típicamente un sitio de unión de polimerasa de ARN 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 la polimerasa de ARN adyacente en el que comienza la síntesis de ARN. El operador permite la transcripción regulada negativa (inducible), como una proteína represora del gen puede unirse al operador y de este modo inhibir la transcripción de un gen específico. La expresión constitutiva puede ocurrir en ausencia de elementos reguladores negativos, tales como el operador. Además, la regulación positiva se puede lograr mediante una secuencia de activador gen de la proteína de unión, que, si está presente es usualmente proximal (5') a la secuencia de unión a polimerasa de ARN. Un ejemplo de una proteína activadora del gen es la proteína activadora de catabolito (CAP), que ayuda a iniciar la transcripción del operón lac en *Escherichia coli* (*E. coli*) [Raibaud *et al.*, ANNU. REV. GENET. (1984) 18:173]. Por lo tanto, la expresión regulada puede ser positiva o negativa, potenciando o reduciendo la transcripción.

Las secuencias que codifican enzimas de la ruta metabólica proporcionan secuencias promotoras particularmente útiles. Los ejemplos incluyen secuencias promotoras derivadas de enzimas que metabolizan azúcares, tales como galactosa, lactosa (lac) [Chang *et al.*, Nature (1977) 198: 1056], y maltosa. Ejemplos adicionales incluyen secuencias promotoras derivadas de enzimas biosintéticas tales como triptófano (trp) [Goeddel *et al.*, Nuc. Acids Res. (1980) 8: 4057; Yelverton *et al.*, Nucl. Acids Res. (1981) 9: 731; Patente de Estados Unidos. n° 4.738.921; EP Pub. Nos. 036 776 y 121 775]. El sistema promotor  $\beta$ -galactosidasa (bla) [Weissmann (1981) "The cloning of interferon and other mistakes". En Interferon 3 (Ed I. Gresser.), nacteriófago lambda PL [Shimatake *et al.*, NATURE (1981) 292: 128] y sistemas promotores T5 [Patente de Estados Unidos. n° 4.689.406] también proporcionan secuencias promotoras útiles. Los métodos preferidos de la presente divulgación utilizan promotores fuertes, tales como el promotor de T7 para inducir polipéptidos de insulina en niveles altos. Ejemplos de tales vectores son conocidos por personas de experiencia ordinaria en la técnica e incluyen la serie pET29 de Novagen, y los vectores pPop descritos en WO99/05297. Tales sistemas de expresión producen altos niveles de polipéptidos de insulina en el huésped sin comprometer parámetros de viabilidad de la célula huésped o de crecimiento. pET19 (Novagen) es otro vector conocido en la técnica.

Además, los promotores sintéticos que no se producen en la naturaleza también funcionan como promotores bacterianos. Por ejemplo, las secuencias de activación de la transcripción de un promotor bacteriano o de bacteriófago pueden unirse con las secuencias operón de otro promotor bacteriano o de bacteriófago, creando un promotor híbrido sintético [Patente de Estados Unidos. n° 4.551.433]. Por ejemplo, el promotor lac es un promotor híbrido trp-lac compuesto de ambas secuencias de operón promotor trp y lac que está regulado por el represor lac [Amann y otros, GENE (1983) 25: 167; de Boer *et al.*, PROC. NATL. ACAD. SCI. (1983) 80:21]. Además, un promotor bacteriano puede incluir promotores de origen natural de origen no bacteriano que tienen la capacidad de unirse a la polimerasa de ARN bacteriana e iniciar la transcripción. Un promotor de origen natural de origen no bacteriano también se puede acoplar con una polimerasa de ARN compatible para producir altos niveles de expresión de algunos genes en procariontes. El sistema polimerasa/promotor de bacteriófago T7 ARN es un ejemplo de un sistema promotor acoplado [Studier *et al.*, J. MOL. BIOL. (1986) 189:113; Tabor *et al.*, Proc Natl. Acad. Sci. (1985) 82:1074]. Además, un promotor híbrido también puede estar compuesto de un promotor de bacteriófago y una región operadora de *E. coli* (EP Pub. n° 267 851).

Además de una secuencia promotora en funcionamiento, un sitio de unión al ribosoma eficaz también es útil para la expresión de genes extraños en procariontes. En *E. coli*, el sitio de unión al ribosoma se denomina la secuencia Shine-Dalgarno (SD) e incluye un codón de iniciación (ATG) y una secuencia de 3-9 nucleótidos de longitud localizada 3-11 nucleótidos aguas arriba del codón de iniciación [Shine *et al.*, NATURE (1975) 254: 34]. Se cree que la secuencia SD promueve la unión de ARNm al ribosoma por el apareamiento de bases entre la secuencia SD y la 3' y de *E. coli* 16S ARNr [Steitz *et al.* "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 procariontes con un sitio de unión a ribosoma débil [Sambrook *et al.* "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 célula bacteriana que puede ser, o ha sido, usada como receptor para vectores recombinantes u otro ADN de transferencia. El término incluye la progenie de la célula huésped bacteriana original que ha sido transfectada. 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 al progenitor original, debido a mutación accidental o deliberada. La progenie de la célula parental que es suficientemente similar al padre a caracterizarse por la propiedad relevante, tales como la presencia de una secuencia de nucleótidos que codifica un polipéptido de insulina, están incluidos en la progenie prevista por esta definición.

La selección de bacterias huésped adecuadas para la expresión de polipéptidos de insulina es conocida por los expertos ordinarios en la técnica. En la selección de huéspedes bacterianos para la expresión, hospedantes adecuados pueden incluir los que se muestran tener, entre otras cosas, una buena capacidad de formación de cuerpos de inclusión, baja actividad proteolítica, y la robustez general. Los huéspedes bacterianos están generalmente disponibles de una variedad de fuentes, incluyendo, pero no limitado a, la genética bacteriana Stock Center, Departamento de Biofísica y Física Médica de la Universidad de California (Berkeley, CA); y la American Type Culture Collection ("ATCC") (Manassas, VA). Fermentación industrial/farmacéutica generalmente usan bacteria derivada de cepas K (por ejemplo W3110) o a partir de bacterias derivadas de las cepas B (por ejemplo, BL21).

Estas cepas son particularmente útiles debido a que sus parámetros de crecimiento son extremadamente bien conocidos y robusto. Además, estas cepas son no patógenas, que es comercialmente importante para la seguridad y razones ambientales. Otros ejemplos de *E. coli* adecuados anfitriones incluyen, pero no se limitan a las cepas de BL21, DH10B, o derivados de los mismos. En otro ejemplo de los métodos de la presente descripción, el huésped de *E. coli* es una proteasa menos cepa incluyendo, pero no limitado a, OMP y LON. La cepa de célula huésped puede ser una especie de *Pseudomonas*, incluyendo pero no limitado a, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa* y *Pseudomonas putida*. *Pseudomonas fluorescens* biovar 1, cepa designada MB 101, se sabe que es útil para la producción recombinante y está disponible para los procesos de producción de proteínas terapéuticas. Los ejemplos de un sistema de expresión de *Pseudomonas* incluyen el sistema disponible de The Dow Chemical Company como una cepa huésped (Midland, MI disponible en el internet en dow.com).

Una vez que se ha establecido una cepa de célula huésped recombinante (es decir, la construcción de expresión se ha introducido en la célula huésped y células huésped con la construcción de expresión apropiada son aisladas), la cepa de célula huésped recombinante se cultiva en condiciones apropiadas para la producción de polipéptidos de insulina. Como será evidente para un experto en la técnica, el método de cultivo de la cepa de la célula huésped recombinante será dependiente de la naturaleza de la construcción de expresión utilizada y la identidad de la célula huésped. Cepas huésped recombinantes se cultivan normalmente usando métodos que son conocidos por los expertos normales en la técnica. Las células huésped recombinantes son típicamente cultivadas en medio líquido que contiene fuentes asimilables de carbono, nitrógeno y sales inorgánicas y, opcionalmente, que contienen vitaminas, aminoácidos, factores de crecimiento, y otros suplementos de cultivo proteínico conocidos por personas de experiencia ordinaria en la técnica. Los medios líquidos para el cultivo de células huésped pueden contener opcionalmente antibióticos o antifúngicos para evitar el crecimiento de microorganismos indeseables y/o compuestos, incluyendo, pero no limitado a, los antibióticos para seleccionar las células huésped que contienen el vector de expresión.

Las células huésped recombinantes pueden cultivarse en lotes o formatos continuos, ya sea con la recolección de células (en el caso en el que el polipéptido de insulina se acumula intracelularmente) o la recolección de sobrenadante de cultivo en lotes o formatos continuos. Para la producción en células huésped procariontes, se prefieren cultivo discontinuo y recogida de células.

Los polipéptidos de insulina de la presente invención se purifican normalmente después de la expresión en sistemas recombinantes. El polipéptido de insulina se puede purificar a partir de células huésped o medio de cultivo mediante una variedad de métodos conocidos en la técnica. Polipéptidos de insulina producidos en células huésped bacterianas pueden ser poco solubles o insolubles (en la forma de cuerpos de inclusión). En una realización de la presente invención, sustituciones de aminoácidos pueden ser fácilmente hechas en el polipéptido de insulina que se seleccionan con el fin de aumentar la solubilidad de la proteína producida de manera recombinante utilizando los métodos descritos en el presente documento, así como los conocidos en la técnica. En el caso de la proteína insoluble, la proteína puede ser recogida a partir de lisados de células huésped por centrifugación y puede además ser seguido por homogeneización de las células. En el caso de proteínas poco solubles, compuestos, incluyendo, pero no limitado a, polietilimina (PEI) pueden ser añadidos para inducir la precipitación de proteína parcialmente soluble. La proteína precipitada puede entonces ser convenientemente recogida por centrifugación. Las células huésped recombinantes pueden ser interrumpidas o homogeneizadas para liberar los cuerpos de inclusión dentro de las células usando una variedad de métodos conocidos por los expertos normales en la técnica. Disrupción de células huésped o de homogeneización puede llevarse a cabo utilizando técnicas bien conocidas incluyendo, pero no limitado a, la alteración enzimática celular, sonicación, homogeneización, o la interrupción de liberación de alta presión. En un caso del método de la presente descripción, la técnica de liberación de alta presión se utiliza para romper las células huésped de *E. coli* para liberar los cuerpos de inclusión de los polipéptidos de insulina. Al manipular cuerpos de inclusión de polipéptido de insulina, puede ser ventajoso minimizar el tiempo de homogeneización en repeticiones en fin de maximizar el rendimiento de cuerpos de inclusión sin pérdida debido a factores tales como solubilización, cizallamiento mecánico o proteólisis.

El polipéptido de insulina insoluble o precipitado puede entonces ser solubilizado utilizando cualquiera de un número de agentes de solubilización adecuados conocidos en la técnica. El polipéptido de insulina puede ser solubilizado con urea o hidrocloreuro de guanidina. El volumen del polipéptido de insulina solubilizada debe reducirse al mínimo de manera que los lotes grandes se pueden producir usando tamaños de lote convenientemente manejables. Este factor puede ser significativo en un entorno comercial a gran escala donde el huésped recombinante puede cultivarse en lotes que tienen un volumen de miles de litros. Además, al fabricarse el polipéptido de insulina en un entorno comercial a gran escala, en particular para usos farmacéuticos humanos, la evitación de productos químicos agresivos que pueden dañar la maquinaria y el recipiente, o el producto de proteína en sí, se debe evitar, si es posible. Se ha demostrado en el método de la presente descripción que el agente desnaturizante más suave urea se puede utilizar para solubilizar los cuerpos de inclusión del polipéptido de insulina en lugar del agente desnaturizante más duro hidrocloreuro de guanidina. El uso de urea reduce significativamente el riesgo de daños en el equipo de acero inoxidable utilizado en el proceso de fabricación y purificación de polipéptido de insulina, solubilizando con ello de manera eficiente los cuerpos de inclusión del polipéptido de insulina.

En el caso de la proteína de la insulina soluble, la insulina puede ser secretada en el espacio periplásmico o en el

medio de cultivo. Además, la insulina soluble puede estar presente en el citoplasma de las células huésped. Se puede desear concentrar la insulina soluble antes de realizar etapas de purificación. Las técnicas estándar conocidas por personas de experiencia ordinaria en la técnica pueden utilizarse para concentrar la insulina soluble a partir de, por ejemplo, lisados celulares o medio de cultivo. Además, las técnicas estándar conocidas por personas de experiencia ordinaria en la técnica pueden utilizarse para romper las células huésped y liberar insulina soluble del citoplasma o en el espacio periplásmico de las células huésped.

Cuando se produce polipéptido de insulina como una proteína de fusión, la secuencia de fusión puede eliminarse. La eliminación de una secuencia de fusión puede llevarse a cabo por escisión enzimática o química. La eliminación enzimática de secuencias de fusión puede llevarse a cabo utilizando métodos conocidos por los expertos normales en la técnica. La elección de la enzima para la eliminación de la secuencia de fusión será determinada por la identidad de la fusión, y las condiciones de reacción se especificarán por la elección de la enzima, como será evidente para un experto normal en la técnica. escisión química se puede realizar usando reactivos conocidos por personas de experiencia ordinaria en la técnica, incluyendo pero no limitado a, bromuro de cianógeno, la proteasa TEV, y otros reactivos. El polipéptido de insulina escindido puede purificarse a partir de la secuencia de fusión escindida por métodos conocidos por los expertos normales en la técnica. Tales métodos serán determinados por la identidad y propiedades de la secuencia de fusión y el polipéptido de insulina, tal como será evidente para un experto normal en la técnica. Los métodos para la purificación pueden incluir, pero no se limitan a, cromatografía de exclusión de tamaño, cromatografía de interacción hidrófoba, cromatografía de intercambio iónico o diálisis o cualquier combinación de los mismos.

El polipéptido de insulina también puede purificarse para eliminar el ADN de la solución de proteína. ADN puede eliminarse por cualquier método adecuado conocido en la técnica, tales como precipitación o cromatografía de intercambio iónico, pero puede ser removido por precipitación con un agente de precipitación con ácido nucleico, tal como, pero no limitado a, sulfato de protamina. El polipéptido de insulina puede ser separado del ADN precipitado utilizando métodos estándar bien conocidos incluyendo, pero no limitado a, centrifugación o filtración. La eliminación de moléculas de ácido nucleico huésped es un factor importante en un entorno en el que el polipéptido de insulina ha de utilizarse para el tratamiento de los seres humanos y los métodos de la presente descripción reducen ADN de células huésped para niveles farmacéuticamente aceptables.

Métodos para la fermentación a pequeña escala o gran escala también se pueden utilizar en la expresión de proteínas, incluyendo pero no limitados a, fermentadores, matraces de agitación, biorreactores de lecho fluidizado, biorreactores de fibras huecas, sistemas de cultivo de botella con ruedas, y sistemas de biorreactor de tanque agitado. Cada uno de estos métodos se pueden realizar en una proceso por lotes, lote alimentado o de modo continuo.

Polipéptidos de insulina humana de la invención generalmente se pueden recuperar mediante métodos estándar en la técnica. Por ejemplo, medio de cultivo o lisado de células se pueden centrifugar o filtrar para eliminar los desechos celulares. El sobrenadante puede ser concentrado o diluido a un volumen deseado o diafiltrado en un tampón adecuado para acondicionar la preparación para una purificación adicional. La purificación adicional del polipéptido de insulina de la presente invención incluye la separación de formas desamidadas y recortadas de la variante de polipéptido de insulina de la forma intacta.

Cualquiera de los siguientes procedimientos ejemplares puede ser empleado para la purificación de polipéptidos de insulina de la invención: cromatografía de afinidad; cromatografía de intercambio aniónico o catiónico (utilizando, incluyendo, pero no limitado a, DEAE SEPHAROSE); cromatografía sobre sílice; cromatografía líquida de alto rendimiento (HPLC); HPLC de fase inversa; filtración en gel (usando, incluyendo pero no limitado a, SEPHADEX G-75); cromatografía de interacción hidrófoba; cromatografía de exclusión de tamaño; cromatografía de quelato de metal; ultrafiltración/diafiltración; precipitación con etanol; precipitación con sulfato de amoníaco; cromatoenfoque; cromatografía de desplazamiento; procedimientos electroforéticos (incluyendo pero no limitado a enfoque isoeléctrico preparativo), solubilidad diferencial (incluyendo pero no limitado a, precipitación con sulfato amónico), SDS-PAGE, o extracción.

Las proteínas de la presente invención, incluyendo pero no limitado a, proteínas que comprenden aminoácidos no naturales, los péptidos que comprenden aminoácidos no naturales, anticuerpos para proteínas que comprenden aminoácidos no naturales, anticuerpos a proteínas que comprenden aminoácidos no naturales, etc. se pueden purificar, ya sea parcial o sustancialmente a homogeneidad, de acuerdo con procedimientos estándar conocidos y utilizados por los expertos en la técnica. En consecuencia, los polipéptidos de la invención pueden recuperarse y purificarse mediante cualquiera de una serie de métodos conocidos por los expertos normales en la técnica, incluyendo, pero no limitados a, sulfato de amoníaco o precipitación de etanol, extracción con ácido o base, cromatografía en columna, cromatografía de columna de afinidad, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de hidroxilapatita, cromatografía de lectina, electroforesis en gel y similares. Etapas de repliegamiento de proteína se pueden utilizar, según se desee, en la fabricación de proteínas maduras correctamente plegadas. Cromatografía líquida de alto rendimiento (HPLC), cromatografía de afinidad u otros métodos adecuados se pueden emplear en las etapas de purificación finales en donde se desea alta pureza. En un caso, los anticuerpos producidos contra los aminoácidos

no naturales (o proteínas o péptidos que comprenden aminoácidos no naturales) se utilizan como reactivos de purificación, incluyendo, pero no limitado a la purificación a base de afinidad de proteínas o péptidos que comprenden uno o más aminoácidos antinaturales. Una vez purificados, parcialmente o hasta homogeneidad, según se desee, los polipéptidos se utilizan opcionalmente para una amplia variedad de utilidades, incluyendo pero no limitados a, como componentes del ensayo, terapéutica, profilaxis, diagnóstico, reactivos de investigación, y/o como inmunógenos para la producción de anticuerpos. Los anticuerpos generados contra polipéptidos de la presente invención se pueden obtener mediante la administración de los polipéptidos o fragmentos que llevan epitopos, o células a un animal, preferiblemente un animal no humano, usando protocolos de rutina. Un experto en la técnica podría generar anticuerpos usando una variedad de técnicas conocidas. También, ratones transgénicos, u otros organismos, incluyendo otros mamíferos, se pueden usar para expresar anticuerpos humanizados. Los anticuerpos anteriormente descritos pueden emplearse para aislar o identificar clones que expresan el polipéptido o para purificar los polipéptidos. Los anticuerpos contra polipéptidos de la presente invención también se pueden emplear para tratar enfermedades.

Los polipéptidos de la presente invención también pueden usarse como vacunas. Por consiguiente, en un aspecto adicional, la presente descripción se refiere a un método para inducir una respuesta inmunológica en un mamífero que comprende inocular al mamífero con un polipéptido de la presente invención, adecuado para producir anticuerpo y/o respuesta inmune de células T, incluyendo, por ejemplo, células T que producen citoquina o células T citotóxicas, para proteger a dicho animal de la enfermedad, tanto si la enfermedad ya está establecida dentro del individuo o no. Una respuesta inmunológica en un mamífero también puede ser inducida por un método que comprende la administración de un polipéptido de la presente invención a través de un vector que dirige la expresión del polinucleótido y que codifica el polipéptido *in vivo* con el fin de inducir una respuesta tan inmunológica para producir anticuerpo para proteger a dichos animales de enfermedades de la descripción. Una forma de administrar el vector consiste en acelerarlo en las células deseadas como un recubrimiento sobre partículas o de otra manera. Tal vector de ácido nucleico puede comprender ADN, ARN, un ácido nucleico modificado o un híbrido de ADN/ARN. Para su uso como una vacuna, un polipéptido o un vector de ácido nucleico será proporcionado normalmente como una formulación de vacuna (composición). La formulación puede comprender además un vehículo adecuado. Puesto que un polipéptido puede descomponerse en el estómago, puede administrarse por vía parenteral (por ejemplo, subcutánea, intramuscular, intravenosa, o intradérmica). Las formulaciones adecuadas para administración parenteral incluyen soluciones de inyección estériles acuosas y no acuosas que pueden contener antioxidantes, tampones, bacteriostáticos y solutos que hacen la formulación isotónica con la sangre del receptor; y suspensiones acuosas y no acuosas estériles que pueden incluir agentes de suspensión o agentes espesantes. La formulación de vacuna también puede incluir sistemas adyuvantes para potenciar la inmunogenicidad de la formulación que son conocidos por los expertos normales en la técnica. La dosificación dependerá de la actividad específica de la vacuna y puede determinarse fácilmente por experimentación de rutina.

### **Expresión en sistemas alternativos**

Se han empleado varias estrategias para introducir aminoácidos no naturales en proteínas en células hospedadoras no recombinantes, células huésped mutagenizadas, o en sistemas libres de células. Estos sistemas también son adecuados para uso en la fabricación de los polipéptidos de insulina de la presente invención. La derivatización de los aminoácidos con cadenas laterales reactivas, tales como Lys, Cys y Tyr resultaron en la conversión de lisina para N<sub>2</sub>-acetilo-lisina. La síntesis química también proporciona un método sencillo para incorporar aminoácidos no naturales. Con el reciente desarrollo de ligación enzimática y ligadura química nativa de fragmentos de péptidos, es posible hacer las proteínas más grandes. Véase, por ejemplo, P.E. Dawson y S.B.H. Kent, Annu. Rev. Biochem., 69: 923-(2000). Ligación química de péptidos y la ligadura química nativa se describen en la Patente de Estados Unidos n° 6.184.344, la Publicación de Patente de Estados Unidos n° 2004/0138412, la Publicación de Patente de Estados Unidos n° 2003/0208046, WO 02/098902, y WO 03/042235. Se añade un método biosintético *in vitro* general en el que un supresor ARNt acilado químicamente con el aminoácido no natural deseado a un extracto *in vitro* capaz de soportar la biosíntesis de proteínas, se ha usado para incorporar por sitio específico más de 100 aminoácidos no naturales en una variedad de proteínas de prácticamente cualquier tamaño. Véase, 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. Griffith, 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). Una amplia gama de grupos funcionales se ha introducido en las proteínas para estudios de estabilidad de la proteína, plegamiento de proteínas, mecanismo de la enzima, y la transducción de señales.

Además de otras referencias citadas en el presente documento, una variedad de métodos de plegamiento de purificación/proteínas son conocidos por personas de experiencia ordinaria en la técnica, incluyendo, pero no limitado a, los expuestos en 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 *et al.* (1996) Protein Methods, 2ª edición Wiley-Liss, NY; Walker, (1996) The Protein Protocols Handbook Humana Press, NJ, Harris y Angal, (1990) Protein Purification Applications: A Practical Approach IRL Press at Oxford, Oxford, Inglaterra; Harris y Angal, Protein Purification Methods: A Practical Approach IRL Press at Oxford, Oxford, Inglaterra; Scopes, (1993) Protein Purification:

Principles and Practice 3ª edición Springer Verlag, NY; Janson y Ryden, (1998) Protein Purification: High Resolution Methods and Applications, Segunda edición Wiley-VCH, NY; y Walker (1998), Protein Protocols on CD-ROM Humana Press, NJ; y las referencias citadas en el mismo.

5 Una de las ventajas de producir una proteína o polipéptido de interés con un aminoácido no natural en una célula huésped eucariota o célula huésped no eucariótica es que típicamente las proteínas o polipéptidos se pliegan en sus conformaciones nativas. Sin embargo, en ciertas realizaciones de la invención, los expertos en la técnica reconocerán que, después de la síntesis, expresión y/o purificación, las proteínas o péptidos pueden poseer una conformación diferente de las conformaciones deseadas de los polipéptidos relevantes. En un aspecto de la divulgación, la proteína o polipéptido expresado es opcionalmente desnaturalizado y luego renaturalizado. Esto se logra utilizando métodos conocidos en la técnica, incluyendo pero no limitado a, mediante la adición de una chaperonina a la proteína o polipéptido de interés, mediante la solubilización de las proteínas en un agente caotrópico tal como guanidina HCl, utilizando la isomerasa de disulfuro de proteína, etc.

15 En general, es deseable de vez en cuando desnaturalizar y reducir polipéptidos expresados y luego para hacer que los polipéptidos se plieguen en la conformación preferida. Por ejemplo, guanidina, urea, DTT, DTE, y/o una chaperonina se puede añadir a un producto de traducción de interés. Métodos de reducción, desnaturalización y renaturalización de proteínas son conocidos por personas de experiencia ordinaria en la técnica (véase, las referencias anteriores, y Debinski, et al. (1993) J. Biol Chem., 268: 14065-14070; Kreitman y Pastan (1993) Bioconjug Chem., 4: 581-585; y Buchner, et al., (1992) Anal Biochem, 205: 263-270). Debinski, *et al.*, Por ejemplo, describen la desnaturalización y reducción de proteínas de cuerpos de inclusión en guanidina-DTE. Las proteínas pueden replegarse en un tampón redox que contiene, incluyendo pero no limitados a, glutatión oxidado y L-arginina. Reactivos de replegamiento pueden fluir o de otra manera moverse en contacto con uno o más polipéptidos u otro producto de expresión, o viceversa.

25 En el caso de la producción procariota de polipéptido de insulina, el polipéptido de insulina producido de este modo puede ser mal plegado y por lo tanto carece o tiene una actividad biológica reducida. La bioactividad de la proteína puede ser restaurada por "replegado". En general, el polipéptido de insulina mal plegada se repliega mediante la solubilización (donde el polipéptido de insulina también es insoluble), despliegue y la reducción de la cadena polipeptídica usando, por ejemplo, uno o más agentes caotrópicos (por ejemplo, urea y/o guanidina) y un agente reductor capaz de la reducción de los enlaces de disulfuro (por ejemplo, ditiotretitol, DTT o 2-mercaptoetanol, 2-ME). A una concentración moderada de agente caotrópico, se añade entonces un agente oxidante (por ejemplo, oxígeno, cistina o cistamina), que permite la reformación de los enlaces de disulfuro. Polipéptido de insulina puede ser replegado usando métodos estándar conocidos en la técnica, tales como los descritos en la patente de los Estados Unidos. Nos. 4.511.502, 4.511.503, y 4.512.922. El polipéptido de insulina también puede ser coplegado con otras proteínas para formar heterodímeros o heteromultímeros.

30 Después del replegamiento, la insulina puede ser purificada adicionalmente. La purificación de la insulina se puede conseguir utilizando una variedad de técnicas conocidas por personas de experiencia ordinaria en la técnica, incluyendo cromatografía de interacción hidrófoba, cromatografía de exclusión por tamaño, cromatografía de intercambio iónico, cromatografía líquida de alto rendimiento de fase inversa, cromatografía de afinidad, y similares, o cualquier combinación de los mismos. La purificación adicional puede incluir también una etapa de secado o precipitación de la proteína purificada.

45 Después de la purificación, la insulina puede intercambiarse en diferentes tampones y/o concentrarse mediante cualquiera de una variedad de métodos conocidos en la técnica, incluyendo, pero no limitado a, la diafiltración y diálisis. La insulina que se proporciona como una única proteína purificada puede estar sujeta a la agregación y precipitación.

50 La insulina purificada puede ser al menos 90% pura (medida por cromatografía de fase inversa de alto rendimiento líquido, RP-HPLC, o electroforesis en gel poliacrilamida de dodecilsulfato de sodio, SDS-PAGE) o al menos 95% pura, o al menos 98% pura, o por lo menos 99% o más pura. Independientemente del valor numérico exacto de la pureza de la insulina, la insulina es suficientemente pura para su uso como un producto farmacéutico o para un procesamiento adicional, tal como conjugación con un polímero soluble en agua tal como PEG.

55 Ciertas moléculas de insulina pueden ser usadas como agentes terapéuticos en la ausencia de otros ingredientes activos o proteínas (distintas de excipientes, portadores, y estabilizadores, albúmina de suero y similares), o pueden formar complejos con otra proteína o un polímero.

60 **Métodos de purificación generales** Cualquiera de una variedad de etapas de aislamiento se puede realizar en el lisado celular, extracto, medio de cultivo, los cuerpos de inclusión, el espacio periplásmico de las células huésped, el citoplasma de las células huésped, o de otro material, que comprende polipéptido de insulina o en cualquier mezcla de polipéptido de insulina que resultan de cualesquiera etapas de aislamiento incluyendo, pero no limitados a, cromatografía de afinidad, cromatografía de intercambio iónico, cromatografía de interacción hidrófoba, cromatografía de filtración en gel, cromatografía líquida de alto rendimiento ("HPLC"), HPLC de fase inversa ("RP-HPLC"), adsorción de lecho expandida, o cualquier combinación y/o repetición de los mismos y en cualquier orden

apropiado.

Equipos y otros materiales necesarios utilizados en la realización de las técnicas descritas en el presente documento están disponibles comercialmente. Bombas, colectores de fracciones, monitores, grabadoras y sistemas completos están disponibles en, por ejemplo, Applied Biosystems (Foster City, CA), BioRad Laboratories, Inc. (Hercules, CA), y Amersham Biosciences, Inc. (Piscataway, NJ). Materiales cromatográficos incluyendo, pero no limitado a, materiales de matriz de intercambio, los medios y tampones también están disponibles a partir de tales empresas.

El equilibrio, y otros etapas en los procesos de cromatografía de columna descrita en el presente documento como el lavado y la elución, puede ser más rápidamente realizado usando un equipo especializado tal como una bomba. Bombas comercialmente disponibles incluyen, pero no se limitan a, HILOAD® Bomba P-50, Bomba Peristáltica P-1, Bomba P-901, y Bomba P-903 (Amersham Biosciences, Piscataway, NJ).

Ejemplos de colectores de fracciones incluyen RediFrac Fraction Collector, FRAC-100 y FRAC-200 Fraction Collectors, y SUPERFRAC® Fraction Collector (Amersham Biosciences, Piscataway, NJ). Mezcladores también están disponibles para formar gradientes de pH y concentración lineal. mezcladores disponibles comercialmente incluyen el mezclador de gradiente GM-1 y mezcladoras en línea (Amersham Biosciences, Piscataway, NJ).

El procedimiento cromatográfico se puede monitorizar utilizando cualquier monitor disponible comercialmente. Tales monitores pueden usarse para recopilar información como UV, 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 conductividad del monitor (Amersham Biosciences, Piscataway, NJ). De hecho, sistemas completos están disponibles comercialmente, incluidos los diversos sistemas AKTA® de Amersham Biosciences (Piscataway, NJ).

En un caso de la presente descripción, por ejemplo, el polipéptido de insulina puede disminuirse y desnaturalizarse mediante la desnaturalización primero del polipéptido de insulina resultante purificado en urea, seguido de dilución en tampón TRIS que contiene un agente reductor (tal como DTT) a un pH adecuado. En otro ejemplo, el polipéptido de insulina se desnaturaliza en urea en un intervalo de concentración de entre aproximadamente 2 M a 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 este ejemplo puede entonces incubarse. En un caso, se incubaba la mezcla de replegamiento a temperatura ambiente durante cuatro a veinticuatro horas. La mezcla de polipéptido de insulina reducida y desnaturalizada puede entonces estar más aislada o purificada.

Como se ha indicado en el presente documento, el pH de la primera mezcla de polipéptido de insulina se puede ajustar antes de realizar cualesquiera etapas de aislamiento posteriores. Además, la primera mezcla de polipéptido de insulina o cualquier mezcla posterior de los mismos pueden concentrarse usando técnicas conocidas en la técnica. Por otra parte, el tampón de elución que comprende la primera mezcla de polipéptido de insulina o cualquier mezcla de los mismos posterior puede ser intercambiada por un tampón adecuado para la siguiente etapa de aislamiento usando técnicas conocidas por los expertos normales en la técnica.

**Cromatografía de intercambio iónico** En un caso, y como una etapa adicional opcional, cromatografía de intercambio iónico, se pueden realizar en la primera mezcla de polipéptido de insulina. Véase, en general ION EXCHANGE CHROMATOGRAPHY: PRINCIPLES AND METHODS (Cat. nº 1811-1421, Amersham Biosciences (Piscataway, NJ)). Columnas de intercambio iónico comercialmente disponibles incluyen columnas HITRAP®, HIPREP® y HILOAD® (Amersham Biosciences, Piscataway, NJ). Tales columnas utilizan intercambiadores aniónicos fuertes tales como Q SEPHAROSE® Fast Flow, Q SEPHAROSE® High Performance, y Q SEPHAROSE XL; fuertes intercambiadores de cationes tales como SP SEPHAROSE® High Performance, SP SEPHAROSE® Fast Flow, y SP SEPHAROSE® XL; intercambiadores de aniones débiles tales como DEAE SEPHAROSE® Fast Flow; y intercambiadores de cationes débiles tales como CM SEPHAROSE® Fast Flow (Amersham Biosciences, Piscataway, NJ). Cromatografía en columna de intercambio de aniones o de cationes se puede realizar en el polipéptido de insulina en cualquier etapa del proceso de purificación para aislar el polipéptido de insulina sustancialmente purificada. La etapa de cromatografía de intercambio catiónico puede realizarse usando cualquier matriz de intercambio catiónico adecuado. matrices de intercambio catiónico útiles incluyen, pero no se limitan a, materiales de matriz de intercambio de catión fibrosos, porosos, no porosos, microgranulares, moldeados, o reticulados. Tales materiales de matriz de intercambio catiónico incluyen, pero no se limitan a, celulosa, agarosa, dextrano, poliacrilato, polivinilo, poliestireno, sílice, poliéter, o materiales compuestos de cualquiera de los anteriores.

La matriz de intercambio catiónico puede ser cualquier intercambiador de cationes adecuado, incluyendo intercambiadores de cationes fuertes y débiles. intercambiadores de cationes fuertes pueden ionizarse en un amplio intervalo de pH y, por tanto, pueden ser capaces de insulina de unión a lo largo de un amplio intervalo de pH. Intercambiadores de cationes débiles, sin embargo, pueden perder ionización como una función de pH. Por ejemplo, un intercambiador de cationes débil puede perder carga cuando el pH cae por debajo de aproximadamente pH 4 o pH 5. intercambiadores fuertes de cationes adecuados incluyen, pero no se limitan a, grupos cargados funcionales tales como sulfopropilo (SP), sulfonato de metilo (S), o sulfoetilo (SE). La matriz de intercambio catiónico puede ser un intercambiador de cationes fuerte, preferentemente con un intervalo de pH de enlace de insulina aproximadamente 2,5 a aproximadamente 6,0. Alternativamente, el intercambiador de cationes fuerte puede tener



un intervalo de pH de unión de insulina de aproximadamente pH 2,5 a aproximadamente pH 5,5. La matriz de intercambio catiónico puede ser un intercambiador de cationes fuerte que tiene un pH de enlace a la insulina de aproximadamente 3,0. Alternativamente, la matriz de intercambio catiónico puede ser un intercambiador de cationes fuerte, preferentemente con un intervalo de pH de enlace de insulina de aproximadamente 6,0 a aproximadamente 8,0. La matriz de intercambio catiónico puede ser un intercambiador catiónico fuerte teniendo preferiblemente un intervalo de pH de enlace de insulina de aproximadamente 8,0 a aproximadamente 12,5. Alternativamente, el intercambiador de cationes fuerte puede tener un intervalo de pH de enlace de insulina de aproximadamente pH 8,0 a aproximadamente pH 12,0.

Antes de cargarse la insulina, la matriz de intercambio catiónico puede equilibrarse, por ejemplo, el uso de varios volúmenes de columna de una solución ácida diluida débil, por ejemplo, cuatro volúmenes de columna de 20 mM de ácido acético, pH 3. Después del equilibrio, la insulina puede ser añadida y la columna puede lavarse de una a varias veces, antes de la elución de la insulina sustancialmente purificada, utilizando también una solución de ácido débil tal como ácido acético débil o solución de ácido fosfórico. Por ejemplo, aproximadamente 2-4 volúmenes de columna de 20 mM de ácido acético, pH 3, se pueden usar para lavar la columna. Lavados adicionales que utilizan, por ejemplo, 2-4 volúmenes de columna de 0,05 M de acetato sódico, pH 5,5, o 0,05 M de acetato sódico mezclado con 0,1 M de cloruro de sodio, pH 5,5, también se pueden usar. Alternativamente, usando métodos conocidos en la técnica, la matriz de intercambio catiónico puede equilibrarse utilizando varios volúmenes de columna de una solución diluida, base débil.

De forma alternativa, la insulina sustancialmente purificada se puede eluir por contacto de la matriz de intercambiador de cationes con un tampón que tiene un pH suficientemente bajo o la fuerza iónica para desplazar la insulina a partir de la matriz. El pH del tampón de elución puede variar de aproximadamente pH 2,5 a aproximadamente pH 6,0. Más específicamente, el pH del tampón de elución puede variar de aproximadamente pH 2,5 a aproximadamente pH 5,5, aproximadamente pH 2,5 a aproximadamente pH 5,0. El tampón de elución puede tener un pH de aproximadamente 3,0. Además, la cantidad de tampón de elución puede variar ampliamente y generalmente estará en el intervalo de aproximadamente 2 a aproximadamente 10 volúmenes de columna.

Después de la adsorción del polipéptido de insulina a la matriz de intercambiador de cationes, polipéptido de insulina sustancialmente purificada se puede eluir poniendo en contacto la matriz con un tampón que tiene un pH suficientemente alto o la fuerza iónica para desplazar el polipéptido de insulina a partir de la matriz. Los tampones adecuados para uso en alta elución de pH de polipéptido de insulina sustancialmente purificada pueden incluir, pero no se limitan a, citrato, fosfato, formiato, acetato, HEPES, y tampones de MES que varían en concentración desde al menos aproximadamente 5 min hasta al menos aproximadamente 100 mM.

**Cromatografía de fase reversa** RP-HPLC se puede realizar para purificar proteínas siguiendo los protocolos adecuados son conocidos por los expertos normales en la técnica. Véase, por ejemplo, Pearson *et al.*, ANAL BIOCHEM. (1982) 124: 217-230 (1982); Rivier *et al.*, J. Chrom. (1983) 268: 112-119; Kunitani *et al.*, J. CHROM. (1986) 359: 391-402. RP-HPLC se puede realizar en el polipéptido de insulina para aislar polipéptido de insulina sustancialmente purificada. En este sentido, resinas derivatizadas por sílice con funcionalidades de alquilo con una amplia variedad de longitudes, incluyendo, pero no limitado a, al menos aproximadamente C3 a al menos aproximadamente C30, al menos aproximadamente C a al menos aproximadamente C20, o al menos aproximadamente C3 a al menos alrededor de C18, se pueden utilizar resinas. Alternativamente, una resina polimérica puede ser utilizada. Por ejemplo, resina TosoHaas Amberchrome CG1000sd puede ser utilizada, que es una resina de polímero de estireno. Ciano o resinas poliméricas con una amplia variedad de longitudes de cadena de alquilo también se pueden utilizar. Además, la columna de la RP-HPLC se puede lavar con un disolvente tal como dimetilo-L. La columna de Fuente de RP es otro ejemplo de una columna de RP-HPLC.

Un tampón de elución adecuado que contiene un agente de emparejamiento iónico y un modificador orgánico tal como metanol, isopropanol, tetrahidrofurán, acetonitrilo o dimetilo-L, puede usarse para eluir el polipéptido de insulina de la columna de RP-HPLC. Los agentes de apareamiento iónico más comúnmente usados incluyen, pero no se limitan a, ácido acético, ácido fórmico, ácido perclórico, ácido fosfórico, ácido trifluoroacético, ácido heptafluorobutírico, trietilamina, tetrametiloamoníaco, tetrabutiloamoníaco, y acetato de trietilamoníaco. La elución puede llevarse a cabo usando uno o más gradientes o condiciones isocráticas, con condiciones de gradiente preferidas para reducir el tiempo de separación y para disminuir la anchura del pico. Otro método implica el uso de dos gradientes con diferentes rangos de concentración de disolvente. Ejemplos de tampones de elución adecuados para su uso en el presente documento pueden incluir, pero no se limitan a, acetato de amoníaco y soluciones de acetonitrilo.

**Técnicas de purificación cromatografía de interacción hidrófoba** Cromatografía de interacción hidrófoba (HIC) se puede realizar en el polipéptido de insulina. Véase, en general HIDROPHOBIC INTERACTION CHROMATOGRAPHY HANDBOOK: PRINCIPLES AND METHODS (Cat n° 18-1020-90, Amersham Biosciences (Piscataway, NJ). Matrices HIC adecuados pueden incluir, pero no se limitan a, matrices sustituidos por alquilo o arilo, tales como butilo, hexilo, octilo o matrices sustituidos por fenilo incluyendo agarosa, agarosa reticulada, sefarosa, celulosa, sílice, dextrano, poliestireno, matrices poli(metacrilato), y resinas de modo mixto, incluyendo, pero no limitado a, una resina de polietilamina o un matriz poli(metacrilato) sustituido por butilo o fenilo. Fuentes

comercialmente disponibles para la cromatografía de columna de interacción hidrófoba incluyen, pero no se limitan a columnas HITRAP®, HIPREP®, y HILOAD® (Amersham Biosciences, Piscataway, NJ).

En resumen, antes de la carga, la columna de HIC se puede equilibrar usando tampones estándar conocidos por personas de experiencia ordinaria en la técnica, tal como una solución de cloruro de ácido/sodio acético o HEPES que contienen sulfato de amoníaco. El sulfato de amoníaco se puede utilizar como el tampón para la carga de la columna de HIC. Después de cargar el polipéptido de insulina, la columna puede después lavarse usando tampones y condiciones estándar para eliminar los materiales no deseados, pero conservando el polipéptido de insulina en la columna de HIC. El polipéptido de insulina puede eluirse con aproximadamente 3 a aproximadamente 10 volúmenes de columna de un tampón estándar, tales como un tampón HEPES que contiene EDTA y menor concentración de sulfato de amoníaco que el tampón de equilibrio, o un tampón de ácido acético/cloruro de sodio, entre otros. Un gradiente decreciente lineal de sal usando, por ejemplo, un gradiente de fosfato de potasio, también se puede usar para eluir las moléculas de insulina. El eluyente puede entonces ser concentrado, por ejemplo, por filtración tales como diafiltración o ultrafiltración. La diafiltración se puede utilizar para eliminar la sal utilizada para eluir el polipéptido de insulina.

**Otras técnicas de purificación** Otra etapa de aislamiento usando, por ejemplo, filtración en gel (GEL FILTRACIÓN: PRINCIPLES AND METHODS (Cat n° 18-1022-18, Amersham Biosciences, Piscataway, NJ), cromatografía de hidroxapatita (matrices adecuados incluyen, pero no se limitan a, HA-Ultrogel, alta resolución (Calbiochem), cerámica de hidroxapatita de CHT (BioRad), Bio - Hidroxapatita de Gel HTP (BioRad)), HPLC, adsorción de lecho expandido, ultrafiltración, diafiltración, liofilización, y similares, se pueden realizar en la primera mezcla de polipéptido de insulina o cualquier mezcla posterior de los mismos; para eliminar cualquier exceso de sales y para reemplazar el tampón con un tampón adecuado para la siguiente etapa de aislamiento o incluso formulación del producto farmacológico final.

El rendimiento de polipéptido de insulina, incluyendo polipéptido de insulina sustancialmente purificada, puede monitorizarse en cada etapa descrita en el presente documento usando técnicas conocidas por los expertos normales en la técnica. Tales técnicas se pueden usar también para evaluar el rendimiento de polipéptido de insulina sustancialmente purificada después de la última etapa de aislamiento. Por ejemplo, el rendimiento de polipéptido de insulina puede ser monitoreado usando cualquiera de varias columnas de cromatografía líquida de alta presión de fase inversa, que tienen una variedad de longitudes de cadena de alquilo tal como RP-HPLC de ciano, C18RP-HPLC; así como HPLC de intercambio catiónico y HPLC de filtración en gel.

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

Pureza puede determinarse usando técnicas estándar, tales como SDS-PAGE, o mediante la medición de polipéptido de insulina mediante transferencia Western y ensayos ELISA. Por ejemplo, los anticuerpos policlonales se pueden generar contra proteínas aisladas de fermentación de levadura de control negativo y la recuperación de intercambio catiónico. Los anticuerpos también se pueden usar para sondear la presencia de proteínas contaminantes de la célula huésped.

El material de RP-HPLC Vydac C4 (Vydac) consiste en partículas de gel de sílice, las superficies de las cuales llevan cadenas C4. La separación de polipéptido de insulina de las impurezas proteínicas se basa en las diferencias en la fuerza de las interacciones hidrófobas. La elución se realizó con un gradiente de acetonitrilo en ácido trifluoroacético diluido. La HPLC preparativa se realiza usando una columna de acero inoxidable (llena de 2/8 a 3/2 litro de Vydac C4 de gel de sílice). El eluato de Hidroxapatita Ultrogel se acidificó mediante la adición de ácido trifluoroacético y se cargó en la columna de la Vydac C4. Para el lavado y la elución se utiliza un gradiente de acetonitrilo en ácido trifluoroacético diluido. Se recogen las fracciones y se neutralizaron inmediatamente con tampón de fosfato. Las fracciones de polipéptido de insulina que están dentro de los límites IPC se combinaron.

Material de DEAF Sepharose (Pharmacia) consiste en grupos de dietilaminoetil (DEAE) que están unidos covalentemente a la superficie de perlas de Sepharose. La unión de polipéptido de insulina 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 que estas sustancias han sido lavadas, impurezas se eliminan por lavado de la columna con tampón de acetato a un pH bajo. A continuación, la columna se lava con tampón de fosfato neutro y polipéptido de insulina se eluye con un tampón con un aumento de la fuerza iónica. La columna está repleta de flujo rápido DEAE Sepharose. El volumen de la columna se ajusta para asegurar una carga de polipéptido de insulina en el intervalo de

3-10 mg de gel de polipéptido de insulina/ml. La columna se lava con agua y tampón de equilibrado (fosfato sódico/potásico). Las fracciones reunidas del eluato de HPLC se cargan y se lavan la columna con tampón de equilibrado. A continuación, la columna se lava con tampón de lavado (tampón de acetato de sodio) seguido de lavado con tampón de equilibrado. Posteriormente, el polipéptido de insulina se eluye de la columna con tampón de elución (cloruro sódico, fosfato sódico/potásico) y se recoge en una fracción única de acuerdo con el perfil de elución patrón. El eluato de la columna de DEAE Sepharose se ajustó a la conductividad especificada. La sustancia de fármaco resultante se filtra de manera estéril en botellas de Teflon y se almacenó a -70°C.

Métodos adicionales que pueden ser empleados incluyen, pero no se limitan a, etapas para eliminar endotoxinas. Las endotoxinas son lipopolisacáridos (LPS) que se encuentran en la membrana externa de las células huésped gram negativas, tales como, por ejemplo, *Escherichia coli*. Los métodos para reducir los niveles de endotoxinas son conocidos para un experto normal en la técnica e incluyen, pero no se limitan a, técnicas de purificación utilizando soportes de sílice, polvo de vidrio o de hidroxapatita, fase reversa, de afinidad, de exclusión de tamaño, cromatografía de intercambio aniónico, cromatografía de interacción hidrófoba, una combinación de estos métodos, y similares. Modificaciones o métodos adicionales pueden ser necesarios para eliminar los contaminantes tales como proteínas comigratorias del polipéptido de interés. Los métodos para medir los niveles de endotoxinas son conocidos para un experto normal en la técnica e incluyen, pero no se limitan a ensayos, limulus lisado de amebocitos (LAL). El ensayo EndosafeTM-PTS es un sistema de tubo sencillo colorimétrico que utiliza cartuchos precargados con el reactivo LAL, sustrato cromogénico, y el control de endotoxina estándar junto con un espectrofotómetro de mano. Los métodos alternativos incluyen, pero no se limitan a, un método cinético LAL que es turbidimétrico y utiliza un formato de 96 pocillos.

Una amplia variedad de métodos y procedimientos se puede utilizar para evaluar el rendimiento y la pureza de una proteína de la insulina que comprende uno o más aminoácidos no naturales codificados, incluyendo pero no limitado al ensayo de Bradford, SDS-PAGE, SDS-PAGE teñido con plata, SDS-PAGE teñido con Coomassie, espectrometría de masas (incluyendo pero no limitado a, MALDITOF) y otros métodos para la caracterización de proteínas conocidas para un experto normal en la técnica.

Métodos adicionales incluyen, pero no se limitan a: SDS-PAGE junto con métodos de tinción de proteínas, inmunotransferencia, asistida por matriz de desorción láser/espectrometría de masa de ionización (MALDI-MS), cromatografía líquida/espectrometría de masas, enfoque isoeléctrico, intercambio aniónico analítico, cromatofoco y dicroísmo circular.

Un método *in vivo*, denominado incorporación de presión selectiva, se desarrolló para explotar la promiscuidad de sintetasas de tipo silvestre. 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 suministra a la célula con un aminoácido natural particular, se desconecta, se cultiva en medio mínimo que contienen concentraciones limitadas del aminoácido natural, mientras que la transcripción del gen diana es reprimida. En el inicio de una fase de crecimiento estacionario, el aminoácido natural se agota y se reemplaza con el análogo de aminoácido no natural. La inducción de la expresión de los resultados de proteínas recombinantes en 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 las proteínas, y exhiben 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); trifluorometionina se ha utilizado para sustituir la metionina en la lisozima del bacteriófago T4 para estudiar su interacción con ligandos de chitoooligosacáridos por <sup>19</sup>F RMN, véase, por ejemplo, H. Duetzel, E. Daub, V. Robinson y J.F. Honek, Biochemistry, 36: 3404 (1997); y trifluoroleucina se ha incorporado en lugar de leucina, resultando en una mayor estabilidad térmica y química de una proteína de leucine-cremallera. Véase, por ejemplo, Y. Tang, G. Ghirlanda, W.A. Petka, T. Nakajima, WF DeGrado y DA Tirrell, Angew. Chem. Int. Ed. Engl. 40: 1494-(2001). Por otra parte, selenometionina y tellurometionina se incorporan en diversas proteínas recombinantes para facilitar la solución de fases en cristalografía de rayos X. Véase, 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. Lebiada y M. Hatada, Nat Struct Biol. 1: 283 (1994); N. Budisa, B. Steipe, P. Demange, C. Eckerskom, J. Kellermann y R. Huber, Eur. J. Biochem. 230: 788 (1995); y, N. Budisa, W. Karnbrock, S. Steinbacher, A. Humm, L. Prade, T. Neufeind, L. Moroder y R. Huber, J. Mol. Biol. 270: 616 (1997). análogos de metionina con funcionalidades de alqueno o alquino también se han incorporado de manera eficiente, lo que permite una modificación adicional de las proteínas por medios químicos. Véase, por ejemplo, J.C. van Hest y D.A. Tirrell, FEBS Lett. 428: 68 (1998); J.C. van Hest, KL Kiick y DA Tirrell, J. Am. Chem. Soc. 122: 1282 (2000); y, K.L. Kiick y D.A. Tirrell, Tetrahedron, 56: 9487 (2000); Patente de EE.UU. n° 6.586.207; La Publicación de Patente de EE.UU. 2002/0042097.

El éxito de este método depende del reconocimiento de los análogos de aminoácidos no naturales por sintetasas de aminoacilo, que, en general, requieren una alta selectividad para asegurar la fidelidad de la traducción de proteínas. Una manera de ampliar el alcance de este método es para relajar la especificidad de sustrato de sintetasas de aminoacilo, que se ha logrado en un número limitado de casos. Por ejemplo, la sustitución de Ala<sup>294</sup> por Gly en sintetasa feniloalanilo-ARNt de *Escherichia coli* (PheRS) aumenta el tamaño del bolsillo de unión del sustrato, y da como resultado la acilación de ARNtPhe por p-Cl-feniloalanina (p-Cl-Phe). Véase, M. Ibba, P. Kast y H. Hennecke, Biochemistry 33: 7107 (1994). Una cepa de *Escherichia coli* que alberga este PheRS mutante permite la

incorporación de p-Cl-feniloalanina o p-Br-feniloalanina en lugar de fenilo-alanina. Véase, 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). Del mismo modo, se demostró que una mutación de punto Phe130Ser cerca del sitio de unión de aminoácidos de sintetasa tirosilo-ARNt de *Escherichia coli* permitía que azatirosina se incorporase más eficientemente 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* consiste en modificar sintetetasas que tienen mecanismos de corrección de pruebas. Estas sintetetasas no pueden discriminar y por lo tanto activar aminoácidos que son estructuralmente similares a los aminoácidos naturales afines. Este error se corrige en un sitio separado, que desacila el aminoácido de carga errónea de la ARNt para mantener la fidelidad de la traducción de proteínas. Si la actividad correctora de la sintetetasa está deshabilitada, análogos estructurales que están misactivadas pueden escapar de la función de edición e incorporarse. Este enfoque se ha demostrado recientemente con la sintetetasa valilo-ARNt (ValRS). Véase, 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 pueden misaminoacilar ARNtVal con Cys, Thr, o aminobutirato (Abu); estos aminoácidos no afines se hidrolizan posteriormente por el dominio de edición. Después de la mutagénesis aleatoria del cromosoma de *Escherichia coli*, se seleccionó una cepa de *Escherichia coli* mutante que tiene una mutación en el sitio de edición de ValRS. Este ValRS de edición deficiente carga ARNtVal con Cys de forma incorrecta. Debido a que Abu se asemeja estéricamente a Cys (grupo -SH de Cys se reemplaza con -CH<sub>3</sub> en Abu), los ValRS mutantes también incorporan Abu en proteínas cuando esta cepa mutante de *Escherichia coli* se cultiva en presencia de Abu. El análisis de espectrometría de masas muestra que alrededor del 24% de valinas se sustituyen por Abu en cada posición de valina en la proteína nativa.

La síntesis en fase sólida y métodos semisintéticos también han permitido la síntesis de una serie de proteínas que contienen nuevos aminoácidos. Por ejemplo, véase las siguientes publicaciones y referencias citadas, 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); Hofmann, K., Bohn, H. *Studies on polypeptides. XXXVI. The effect of pyrazole-imidazole replacements on the S-protein activating potency of an S-peptide fragment.* J. Am. Chem. 88 (24): 5914-5919 (1966); 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).

La modificación química se ha utilizado para introducir una variedad de cadenas laterales no naturales, incluyendo cofactores, marcadores 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. y ML Bender. Una nueva enzima que contiene un sitio activo sintéticamente formado. TIOsubtilisin. 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, 242 (4881): 1038-1040 (1988).

De forma alternativa, los métodos biosintéticos que emplean aminoacilo-ARNt químicamente modificados se han utilizado para incorporar varias sondas biofísicas en proteínas sintetizadas *in vitro*. Véase las siguientes publicaciones y referencias citadas: Brunner, J. *New Photolabeling and crosslinking methods.* Annu. Rev Biochem 62: 483-514 (1993); y, Krieg, U.C., Walter, P., Hohanson, A.E. *Photocrosslinking of the signal sequence of nascent preprolactin of the 54-kilodalton polypeptide of the signal recognition particle.* Proc Natl. Acad. Sci. 83-(22): 8604-8608 (1986).

Anteriormente, se ha demostrado que los aminoácidos no naturales se pueden incorporar por sitio específico en proteínas *in vitro* por la adición de ARNt supresores químicamente aminoacilados a reacciones de síntesis de proteínas programadas con un gen que contiene una mutación sin sentido ámbra deseada. Mediante el uso de estos enfoques, se puede sustituir un número de los veinte aminoácidos comunes con homólogos estructurales cercanos, por ejemplo, fluorofenilo-alanina para fenilo-alanina, utilizando cepas auxotróficas para un aminoácido particular. Véase, por ejemplo, Noren, C.J., Anthony-Cahill, Griffith, M.C., Schultz, PG *A general method for site-specific incorporation of unnatural amino acids into proteins.* Science 244: 182-188 (1989); M.W. Nowak, et al., 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 et al., 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, 301-336 (1992); y, Mendel D., Cornish, V.W. y Schultz, P.G. *Site-Directed Mutagenesis with an Expanded Genetic Code,*

Annu Rev Biophys Biomol Struct 24, 435-62 (1995).

Por ejemplo, un supresor ARNt se preparó que reconoció el codón de parada UAG y fue aminoacilado químicamente con un aminoácido no natural. mutagénesis de sitio dirigido convencional se utilizó para introducir el codón de parada TAG, en el sitio de interés en el gen de la proteína. Véase, por ejemplo, Sayers, J.R., Schmidt, W. Eckstein, F. 5'-3' *Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis*, Nucleic Acids Res, 16 (3): 791-802 (1988). Cuando el supresor acilado ARNt y el gen mutante se combinan en un sistema de transcripción/traducción *in vitro*, el aminoácido no natural se incorpora en respuesta al codón UAG que dio una proteína que contiene ese aminoácido en la posición especificada. Los experimentos utilizando [<sup>3</sup>H]-Phe y experimentos con ácidos α-hidroxi demostraron que sólo 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. Véase, por ejemplo, Noren, et al., *supra*; Kobayashi et al., (2003) *Nature Structural Biology* 10(6): 425-432; y, Eliman, J.A., Mendel, D., Schultz, P.G.. *Site-specific incorporation of novel backbone structures into proteins*, Science, 255 (5041): 197-200 (1992).

Un ARNt puede ser aminoacilado con un aminoácido deseado por cualquier método o técnica, incluyendo, pero no limitado a, aminoacilación química o enzimática.

Aminoacilación se puede realizar mediante sintetetasas de ARNt de aminoacilo o por otras moléculas enzimáticas, incluyendo, pero no limitado a, ribozimas. El término "ribozima" es intercambiable con "ARN catalítico." Cech y colaboradores (Cech, 1987, *Science*, 236: 1532-1539; McCorkle et al., 1987, *Concepts Biochem* 64: 221-226) demostraron la presencia de ARN de origen natural que pueden actuar como catalizadores (ribozimas). Sin embargo, aunque sólo se han mostrado estos catalizadores de ARN naturales para actuar sobre sustratos de ácido ribonucleico para la escisión y empalme, el reciente desarrollo de la evolución artificial de ribozimas ha ampliado el repertorio de la catálisis a varias reacciones químicas. Los estudios han identificado moléculas de ARN que pueden catalizar enlaces de aminoacilo-RNA por sí sólo (2')-3'-termini (Illangakekare et al., 1995 *Science* 267: 643-647), y una molécula de ARN que puede transferir un aminoácido de una molécula de ARN a otra (Lohse et al., 1996, *Nature* 381: 442-444).

La Publicación de Solicitud de Patente de EE.UU. 2003/0228593, describe métodos para construir ribozimas y su uso en la aminoacilación de ARNt con aminoácidos naturalmente codificados y no naturalmente codificados. Formas inmovilizadas por sustrato de moléculas enzimáticas que pueden aminoacilar ARNt, incluyendo pero no limitado a, ribozimas, pueden permitir la purificación por afinidad eficaz de los productos aminoacilados. Ejemplos de sustratos adecuados incluyen agarosa, sefarosa, y perlas magnéticas. La producción y el uso de una forma inmovilizada de sustrato de ribozima para aminoacilación se describe en *Chemistry and Biology* 2003, 10: 1077-1084 y Solicitud de Patente de EE.UU. 2003/0228593.

Métodos de aminoacilación química incluyen, pero no se limitan a, los introducidos por Hecht y colaboradores (Hecht, S.M. *Ace 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.; Eliman, 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. et al. *Nature* 1992, 356, 537; Gallivan, J.P.; Lester, H.A.; Dougherty, D.A. *Chem Biol* 1997, 4, 740; Turcatti, et al. *J. Biol Chem* 1996, 271, 19991; Nowak, M.W. et al. *Science*, 1995, 268, 439; Saks, M.E. et al. *J. Biol Chem* 1996, 271, 23169; Hohsaka, T. et al. *J. Am Chem Soc* 1999, 121, 34), para evitar el uso de sintetetasas en aminoacilación. Tales métodos u otros métodos de aminoacilación químicos se pueden utilizar para aminoacilar moléculas de ARNt.

Los métodos para generar ARN catalítico pueden implicar la generación de depósitos separados de secuencias de ribozima aleatorias, la realización de la evolución dirigida de los depósitos, la detección de los depósitos para la actividad de aminoacilación deseable, y la selección de secuencias de los ribozimas que muestran actividad de aminoacilación deseada.

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

Las ribozimas pueden ser generadas por selección *in vitro* usando un r24mini parcialmente aleatorizado conjugado con ARNt<sup>Asn</sup><sub>CCCG</sub>, seguido por ingeniería sistemática de una secuencia de consenso encontrada en los clones activos. Una ribozima ejemplar obtenida por este método se denomina "ribozima Fx3" y se describe en *Sol. Pub. de EE.UU. n° 2003/0228593*, actúa como un catalizador versátil para la síntesis de diversos aminoacilo-ARNt cargados con aminoácidos no naturales afines.

Inmovilización sobre un sustrato se puede usar para permitir la purificación por afinidad eficaz de los ARNt aminoacilados. Los ejemplos de sustratos adecuados incluyen, pero no se limitan a, agarosa, sefarosa, y perlas magnéticas. Las ribozimas se pueden inmovilizar en resinas mediante el aprovechamiento de la estructura química de ARN, tales como el 3'-cis-diol en la ribosa de ARN pueden ser oxidados con peryodato para dar el dialdehído correspondiente para facilitar la inmovilización del ARN en la resina. Varios tipos de resinas se pueden utilizar incluyendo resinas de hidrazida de bajo costo en el que la aminación reductora hace la interacción entre la resina y la ribozima un enlace irreversible. Síntesis de aminoacilo-ARNt puede facilitarse significativamente por esta técnica de aminoacilación en columna. Kourouklis *et al.* *Methods* 2005; 36: 239-4 describe un sistema de aminoacilación a base de columna.

El aislamiento de los ARNt aminoacilados se pueden lograr de una variedad de maneras. Un método adecuado consiste en eluir los ARNt aminoacilados de una columna con un tampón tal como una solución de acetato de sodio con 10 mM de EDTA, un tampón que contiene 50 mM N-(2-hidroxietilo)piperazina-N'-(3-ácido propanosulfónico), 12,5 mM KCl, pH 7,0, 10 mM de EDTA, o únicamente un agua tamponada con EDTA (pH 7,0).

Los ARNt aminoacilados se pueden añadir a las reacciones de traslado con el fin de incorporar el aminoácido con el que el ARNt se aminoacila en una posición de elección en un polipéptido fabricado por la reacción de traslado. Ejemplos de sistemas de traslado en los que los ARNt aminoacilados de la presente divulgación se pueden usar incluyen, pero no se limitan a los lisados celulares. Los lisados celulares proporcionan componentes de la reacción necesarios para la traducción *in vitro* de un polipéptido a partir de un ARNm de entrada. Ejemplos de tales componentes de reacción incluyen, pero no se limitan a las proteínas ribosómicas, ARNr, aminoácidos, ARNt, GTP, ATP, iniciación de la traducción y factores de elongación y factores adicionales asociados con la traducción. Además, los sistemas de traducción pueden ser trasladados por lotes o traslado compartimentado. Los sistemas de traslado por lotes combinan componentes de la reacción en un solo compartimiento, mientras que los sistemas de traslado compartimentados separan los componentes de la reacción de traslado de los productos de reacción que pueden inhibir la eficacia del traslado. Tales sistemas de traslado están disponibles comercialmente.

Además, puede usarse un sistema de transcripción/traslado acoplado. Sistemas de transcripción/traslado acoplados permiten la transcripción de un ADN de entrada en un ARNm correspondiente, que a su vez se traslada por los componentes de la reacción. Un ejemplo de una transcripción/traslado acoplado comercialmente disponible es el Sistema de Traslado Rápido (RTS, Roche Inc.). El sistema incluye una mezcla que contiene lisado de *E. coli* para proporcionar componentes de traslado tales como ribosomas y factores de traducción. Además, una polimerasa de ARN está incluida para la transcripción del ADN de entrada en una plantilla de ARNm para el uso en el traslado. RTS pueden utilizar compartimentación de los componentes de la reacción por medio de una membrana interpuesta entre compartimientos de reacción, incluyendo un compartimiento de suministro/residuos y un compartimiento de transcripción/traslado.

Aminoacilación de ARNt puede ser realizada por otros agentes, incluyendo, pero no limitado a, transferasas, polimerasas, anticuerpos catalíticos, proteínas multifuncionales, y similares.

Stephan en *Scientist* 2005 Oct 10; páginas 30-33 describe métodos adicionales para incorporar aminoácidos no naturalmente codificados en proteínas. Lu *et al.* en *Mol Cell*. 2001 Oct; 8(4): 759-69 describen un método en el que una proteína se liga químicamente a un péptido sintético que contiene aminoácidos no naturales (expresado de ligación de proteínas).

Las técnicas de microinyección también se han utilizado 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.U. 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). Un oocito de *Xenopus* se coinyecta con dos especies de ARN realizadas *in vitro*: un ARNm que codifica la proteína diana con un codón de parada UAG en la posición de aminoácidos de interés y un supresor ARNt ámbra aminoacilado con el aminoácido no natural deseado. La maquinaria de traslado del oocito después inserta el aminoácido no natural en la posición especificada por UAG. Este método ha permitido estudios estructura-función *in vivo* de proteínas integrales de membrana, que generalmente no son susceptibles de sistemas de expresión *in vitro*. Los ejemplos incluyen la incorporación de un aminoácido fluorescente en receptor taqicinina neuroquinina-2 para medir distancias mediante transferencia de energía por resonancia de fluorescencia, véase, por ejemplo, G. Tureatti, 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 los residuos expuestos a superficie en los canales iónicos, véase, por ejemplo, J.P. Gallivan, H.A. Lester y D.A. Dougherty, *Chem. Biol.* 4: 739 (1997); el uso de análogos de tirosina enjaulados para monitorizar cambios conformacionales en un canal iónico 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 aminoácidos alfa hidroxilo para cambiar cadenas principales de los canales iónicos para sondear sus mecanismos de activación periódica. Véase, 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 de incorporar aminoácidos no naturales directamente en proteínas *in vivo* ofrece una amplia variedad

de ventajas, incluyendo pero no limitado a, altos rendimientos de proteínas mutantes, facilidad técnica, el potencial para estudiar las proteínas mutantes en células o posiblemente en los organismos vivos y el uso de estas proteínas mutantes en tratamientos terapéuticos y usos diagnósticos. La capacidad de incluir aminoácidos no naturales con tamaños diversos, acideces, nucleofilidades, hidrofobicidades y otras propiedades en proteínas pueden expandir en gran medida nuestra capacidad para manipular racional y sistemáticamente las estructuras de las proteínas, tanto para sondear la función de proteínas y crear nuevas proteínas u organismos con propiedades novedosas.

En un intento de incorporar por sitio específico para-F-Phe, un supresor ámbar de levadura ARNtPheCUA /par de sintetasa fenilalanilo-ARNt se utilizó en una cepa *Escherichia coli* auxotrófica 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 de insulina de la presente invención utilizando un sistema de traslado libre de células (*in vitro*). Sistemas de traslado pueden ser celulares o libres de células, y pueden ser procariotas o eucariotas. Los sistemas de traslado celulares incluyen, pero no se limitan a, preparaciones de células completas, tales como células permeabilizadas o cultivos de células en los que una secuencia de ácido nucleico deseado puede ser transcrito a ARNm y el ARNm trasladado. Los sistemas de traslado libres de células están disponibles comercialmente y muchos tipos y sistemas diferentes son bien conocidos. Ejemplos de sistemas libres de células incluyen, pero no se limitan a, los lisados procariotas tales como lisados de *Escherichia coli*, y los lisados eucarióticos, tales como extractos de germen de trigo, lisados de células de insectos, lisados de reticulocitos de conejo, lisados de ovocitos de conejo y los lisados celulares humanos. Extractos o lisados eucarióticos pueden ser preferidos cuando la proteína resultante está glicosilada, fosforilada o modificada de otra manera, porque muchas de estas modificaciones sólo son posibles en los sistemas eucariotas. Algunos de estos extractos y lisados están disponibles comercialmente (Promega, Madison, Wis.; Stratagene; La Jolla, Calif.; Amersham; Arlington Heights, Ill.; GIBCO/BRL, Grand Island, NY). Extractos membranosos, tales como los extractos pancreáticos caninos que contienen membranas microsomales, están también disponibles que son útiles para el traslado de las proteínas secretoras. En estos sistemas, los cuales pueden incluir ARNm como una plantilla (traslado *in vitro*) o ADN como una plantilla (transcripción *in vitro* combinada y traslado), la síntesis *in vitro* se dirige por los ribosomas. Un esfuerzo considerable se ha aplicado al desarrollo de sistemas de expresión de proteínas libres de células. Véase, por ejemplo, Kim, D.M. y J.R. Swartz, *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 Estados Unidos nº 6.337.191; Publicación de Patente de Estados Unidos nº 2002/0081660; WO 00/55353; WO 90/05785. Otro enfoque que se puede aplicar a la expresión de polipéptidos de insulina que comprenden un aminoácido no naturalmente codificado incluye la técnica de fusión ARNm-péptido. Véase, por ejemplo, R. Roberts y J. Szostak, *Proc. Natl Acad. Sci. (EE.UU.)* 94: 12297-12302(1997); A. Frankel, et al., *Chemistry & Biology* 10: 1043-1050 (2003). En este enfoque, un molde de ARNm ligado a puromicina se traslada en péptido en el ribosoma. Si una o más moléculas de ARNt se ha modificado, los aminoácidos no naturales también se pueden incorporar en el péptido. Después de leerse el último codón de ARNm, la puromicina captura el término C del péptido. Si se encuentra que el conjugado de ARNm-péptido resultante tiene propiedades interesantes en un ensayo *in vitro*, su identidad puede ser revelada fácilmente de la secuencia ARNm. De esta manera, se puede seleccionar bibliotecas de polipéptidos de insulina que comprenden uno o más aminoácidos no naturalmente codificados para identificar polipéptidos que tienen propiedades deseadas. Más recientemente, traslados de ribosoma *in vitro* con componentes purificados se han reportado que permiten la síntesis de péptidos sustituidos con aminoácidos no naturalmente codificados. Véase, por ejemplo, A. Forster *et al.*, *Proc. Natl Acad. Sci. (EE.UU.)* 100: 6353 (2003).

También se pueden utilizar sistemas de traslado reconstituidos. Las mezclas de factores de traslado purificados también se han utilizado con éxito para trasladar ARNm en proteínas, así como combinaciones de lisados o lisados suplementados con factores de traslado purificados, tales como la factor de iniciación 1 (IF-1), IF-2, IF-3 ( $\alpha$  o  $\beta$ ), factor de elongación T (EF-Tu), o factores de terminación. Sistemas libres de células también pueden ser sistemas acoplados de transcripción/traslado en donde ADN se introduce en el sistema, transcritos en ARNm y ARNm trasladado como se describe en *Current Protocols in Molecular Biology* (F.M. Ausubel *et al.* editores, Wiley Interscience, 1993). ARN transcrito en el sistema de transcripción eucariota puede estar en forma de ARN heteronuclear (ARNhn) o tapas 5'-extremo (guanósina 7-metilo) y ARNm maduro de cola poli A 3'-extremo, que puede ser una ventaja en ciertos sistemas de traslado. Por ejemplo, los ARNm tapados se trasladan con una alta eficiencia en el sistema de lisado de reticulocitos.

#### **Polímeros macromoleculares acoplados a polipéptidos de insulina**

Diversas modificaciones a los polipéptidos de aminoácidos no naturales descritos en el presente documento se pueden efectuar usando las composiciones, procedimientos, técnicas y estrategias descritas en este documento. Estas modificaciones incluyen la incorporación de funcionalidad adicional en el componente de aminoácido no natural del polipéptido, incluyendo, pero no limitado a, una etiqueta; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotoentrecruzador; un radionucleido; un compuesto citotóxico; una droga; una etiqueta de afinidad; un marcador 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 metales; un cofactor;

un ácido graso; un carbohidrato; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido; un dendrímero soluble en agua; una ciclodextrina; un ácido ribonucleico inhibitorio; un biomaterial; una nanopartícula; una etiqueta de espín; un fluoróforo, un resto que contiene metal; un resto radioactivo; un grupo funcional novela; un grupo que covalentemente o no covalentemente interactúa con otras moléculas; un resto excitable de radiación actínica; un resto fotoisomerizable; biotina; un derivado de biotina; un análogo de biotina; un resto de la incorporación de un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar de enlace de carbono; un agente redoxactivo; un amino tioácido; un resto tóxico; un resto marcado isotópicamente; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo de electrones densos; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; un marcador detectable; una molécula pequeña; un punto cuántico; un nanotransmisor; un radionucleótido; un radiotransmisor; un agente de captura de neutrones; o cualquier combinación del otro compuesto o sustancia deseable anteriormente. Como un ejemplo ilustrativo, no limitante de las composiciones, procedimientos, técnicas y estrategias descritas en el presente documento, la siguiente descripción se centrará en la adición de polímeros macromoleculares con el polipéptido de aminoácidos no naturales con el entendimiento de que las composiciones, procedimientos, técnicas y estrategias descritas son también aplicables (con las modificaciones apropiadas, si es necesario y para los que un experto en la técnica podría hacer con las descripciones del presente documento) a la adición de otras funcionalidades, incluyendo pero no limitado a los enumerados anteriormente. polipéptidos de insulina de la invención comprenden un polímero soluble en agua unido por un enlace de oxima covalente con la cadena A de polipéptido de insulina como se define en las reivindicaciones.

Una amplia variedad de polímeros macromoleculares y otras moléculas pueden estar enlazadas a polipéptidos de insulina de la presente invención para modular las propiedades biológicas del polipéptido de insulina, y/o proporcionar nuevas propiedades biológicas a la molécula de insulina. Estos polímeros macromoleculares se pueden vincular con el polipéptido de insulina a través de un aminoácido codificado de forma natural, a través de un aminoácido no naturalmente codificado, o cualquier sustituyente funcional de un aminoácido natural o no natural, o cualquier sustituyente o grupo funcional añadido a un aminoácido natural o no natural. El peso molecular del polímero puede ser un amplio intervalo, incluyendo, pero no limitado 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, incluyendo pero no limitado 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 el presente documento se divulgan preparaciones sustancialmente homogéneas de polímero: conjugados de proteína. "Sustancialmente homogéneo" tal como se usa en la presente memoria significa que moléculas de conjugado polímero:proteína son mayores que la mitad de la proteína total. El conjugado de polímero:proteína tiene actividad biológica y las presentes preparaciones de polipéptido de insulina PEGiladas "sustancialmente homogéneas" proporcionadas en este documento son aquellas que son lo suficientemente homogéneas para mostrar las ventajas de una preparación homogénea, por ejemplo, la facilidad en la aplicación clínica en la previsibilidad de farmacocinética lote a lote.

También se puede optar para preparar una mezcla de moléculas de conjugado de polímero:proteína, y la ventaja proporcionada en este documento es que se puede seleccionar la proporción de conjugado de monopolímero:proteína que se ha de incluir en la mezcla. Por lo tanto, si se desea, se puede preparar una mezcla de diversas proteínas con diversos números de restos de polímero unidos (es decir, di, tn, tetra, etc.) y combinar dichos conjugados con el conjugado de monopolímero:proteína preparado utilizando los métodos de la presente divulgación, y tener una mezcla con una proporción predeterminada de conjugados de monopolímero:proteína.

El polímero seleccionado puede ser soluble en agua de modo que la proteína a la que está unido no precipita en un entorno acuoso, tal como un entorno fisiológico. El polímero puede ser ramificado o no ramificado. Para el uso terapéutico de la preparación de producto final, el polímero será farmacéuticamente aceptable.

Ejemplos de polímeros incluyen, pero no se limitan a éteres de polialquilo y análogos tapados por alcoxi de los mismos (por ejemplo, glicol de polioxietileno, polioxietileno/propileno, y análogos tapados por metoxi o etoxi de los mismos, especialmente glicol de polioxietileno, este último también se conoce como polietilenglicol o PEG); polivinilpirrolidonas; éteres de poliviniloalquilo; polioxazolininas, polialquilo oxazolininas y oxazolininas de polihidroxi alquilo; poli acrilamidas, acrilamidas de polialquilo, y acrilamidas de polihidroxi alquilo (por ejemplo, polihidroxi propilometacrilamida y derivados de los mismos); acrilatos de polihidroxi alquilo; ácidos polisialícos y análogos de los mismos, secuencias de péptidos hidrófilos; polisacáridos y sus derivados, incluyendo derivados de dextrano y dextrano, por ejemplo, carboximetilodextrano, sulfatos de dextrano, aminodextrano; celulosa y sus



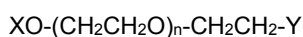
derivados, por ejemplo, carboximetilcelulosa, hidroxialquilcelulosas; quitina y sus derivados, por ejemplo, quitosano, succinilo de quitosano, carboximetilquitina, carboximetilquitosano; ácido hialurónico y sus derivados; almidones; alginatos; sulfato de condroitina; albúmina; pululano y carboximetilo pululano; poliaminoácidos y derivados de los mismos, por ejemplo, ácidos poliglutámicos, polilisoínas, ácidos poliaspárticos, amidas de ácido poliaspártico; copolímeros de anhídrido maleico tales como: copolímero de anhídrido maleico de estireno, copolímero de anhídrido maleico-éter de diviniloetilo; alcoholes de polivinilo; copolímeros de los mismos; terpolímeros de los mismos; mezclas de los mismos; y derivados de los anteriores.

La proporción de moléculas de polietilenglicol a moléculas de proteína variará, como lo harán sus concentraciones en la mezcla de reacción. En general, la relación óptima (en términos de eficacia de la reacción en que no hay un mínimo exceso de proteína sin reaccionar o polímero) se puede determinar por el peso molecular del polietilenglicol seleccionado y del número de grupos reactivos disponibles. En lo que corresponde al peso molecular, típicamente cuando más alto sea el peso molecular del polímero, menor será el número de moléculas de polímero que puede estar unido a la proteína. De manera similar, la ramificación del polímero debe tenerse en cuenta cuando se optimizan estos parámetros. Generalmente, cuanto mayor sea el peso molecular (o cuantas más ramas) mayor será la proporción de polímero:proteína.

Tal como se usa en este documento, y cuando se contempla PEG:conjugados de polipéptido de insulina, el término "cantidad terapéuticamente eficaz" se refiere a una cantidad que da el beneficio deseado a un paciente. La cantidad variará de un individuo a otro y dependerá de un número de factores, incluyendo la condición física general del paciente y la causa subyacente de la condición a ser tratada. La cantidad de polipéptido de insulina utilizado para la terapia da una tasa aceptable de cambio y mantiene la respuesta a un nivel beneficioso deseado. Una cantidad terapéuticamente eficaz de las presentes composiciones puede ser determinada fácilmente por un experto normal en la técnica usando materiales y procedimientos disponibles públicamente.

El polímero de agua soluble puede ser cualquier forma estructural incluyendo pero no limitado a lineal, bifurcado o ramificada. Típicamente, el polímero soluble en agua es un poli(alquilenglicol), tal como poli(etilenglicol)(PEG), pero otros polímeros solubles en agua también se pueden emplear. A modo de ejemplo, PEG se utiliza para describir ciertas realizaciones de esta invención.

PEG es un polímero bien conocido, soluble en agua que está disponible comercialmente o se puede preparar por polimerización de apertura de anillo de etilenglicol de acuerdo con procedimientos conocidos por los expertos normales en la técnica (Sandler y Karo, Polymer Synthesis, Academic Press, Nueva York, Vol. 3, páginas 138-161). El término "PEG" se usa ampliamente para abarcar cualquier molécula de polietilenglicol, sin tener en cuenta al tamaño o a la modificación en un extremo del PEG, y se puede representar como ligado al polipéptido de insulina por la fórmula:



donde n es de 2 a 10,000 y X es H o una modificación terminal, incluyendo pero no limitado a, C<sub>1-4</sub>alquilo, un grupo protector, o un grupo funcional terminal.

En algunos casos, un PEG usado en la invención termina en un extremo con hidroxilo o metoxi, es decir, X es H o CH<sub>3</sub> ("metoxi PEG"). Como alternativa, el PEG puede terminar con un grupo reactivo, formando de este modo un polímero bifuncional. Grupos reactivos típicos pueden incluir aquellos grupos reactivos que se utilizan comúnmente para reaccionar con los grupos funcionales que se encuentran en los 20 aminoácidos comunes (incluyendo pero no limitado a, grupos maleimidas, carbonatos activados (incluyendo pero no limitado a, éster p-nitrofenilo), ésteres activados (incluyendo pero no limitado a, N-hidroxisuccinimida, éster p-nitrofenilo) y aldehídos), así como grupos funcionales que son inertes a los 20 aminoácidos comunes, pero que reaccionan específicamente con grupos funcionales complementarios presentes en aminoácidos no naturalmente codificados (incluyendo pero no limitados a, grupos azida, grupos alquino). Se hace notar que el otro extremo del PEG, que se muestra en la fórmula anterior por Y, se unirá directa o indirectamente a un polipéptido de insulina a través de un aminoácido no naturalmente codificado en estado natural. Por ejemplo, Y puede ser una amida, carbamato o enlace de urea a un grupo amina (incluyendo pero no limitado a la amina épsilon de la lisina o el término N) del polipéptido. Alternativamente, Y puede ser un enlace maleimida a un grupo tiol (incluyendo pero no limitado al grupo tiol de la cisteína). Alternativamente, Y puede ser un enlace a un residuo no comúnmente accesible a través de los 20 aminoácidos comunes. Por ejemplo, un grupo azida en el PEG puede hacerse reaccionar con un grupo alquino en el polipéptido de insulina para formar un producto de cicloadición de Huisgen [3 + 2]. Alternativamente, un grupo alquino en el PEG puede hacerse reaccionar con un grupo azida presente en un aminoácido no naturalmente codificado para formar un producto similar. En algunas realizaciones, un nucleófilo fuerte (incluyendo pero no limitado a, hidrazina, hidrazida, hidroxilamina, semicarbazida) puede hacerse reaccionar con un grupo aldehído o cetona presente en un aminoácido no naturalmente codificado para formar una hidrazona, oxima o semicarbazona, según corresponda, que en algunos casos puede ser reducido aún más por tratamiento con un agente reductor apropiado. Como alternativa, el nucleófilo fuerte se puede incorporar en el polipéptido de insulina a través de un aminoácido no naturalmente codificado y se utiliza para reaccionar preferentemente con un grupo cetona o aldehído presente en el polímero soluble en agua.

5 Cualquier masa molecular para un PEG se puede utilizar como prácticamente deseado, incluyendo pero no limitado a, de aproximadamente 100 Daltons (Da) a 100,000 Da o más según se desee (incluyendo pero no limitado a, 0,1-50 kDa o 10-40 kDa). El peso molecular de PEG puede ser un amplio intervalo, incluyendo, pero no limitado a, entre aproximadamente 100 Da y aproximadamente 100,000 Da o más. PEG puede estar entre aproximadamente 100 Da y aproximadamente 100,000 Da, incluyendo pero no limitado 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 PEG está entre aproximadamente 100 Da y aproximadamente 50,000 Da. En algunas realizaciones, el PEG está entre aproximadamente 100 Da y aproximadamente 40,000 Da. En algunas realizaciones, el PEG está entre aproximadamente 1.000 Da y aproximadamente 40,000 Da. En algunas realizaciones, el PEG está entre aproximadamente 5.000 Da y aproximadamente 40,000 Da. En algunas realizaciones, el PEG está entre aproximadamente 10,000 Da y aproximadamente 40,000 Da. PEG de cadena ramificada, incluyendo, pero no limitado a, moléculas de PEG con cada cadena que tiene un PM en el intervalo de 1-100 kDa (incluyendo pero no limitado a, 1-50 kDa o 5-20 kDa) también puede ser utilizado. El peso molecular de cada cadena de PEG de cadena ramificada puede ser, incluyendo, pero no limitado a, entre aproximadamente 1.000 Da y aproximadamente 100.000 Da o más. El peso molecular de cada cadena de PEG de cadena ramificada puede estar entre aproximadamente 1.000 Da y aproximadamente 100.000 Da, incluyendo pero no limitado 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 de cada cadena de PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular de cada cadena de PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular de cada cadena de PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular de cada cadena de PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 20.000 Da. Una amplia gama de moléculas de PEG se describen en, incluyendo pero no limitado al catálogo de Shearwater Polimers, Inc., catálogo de Nektar Therapeutics.

30 Por lo general, al menos un terminal de la molécula de PEG está disponible para la reacción con el aminoácido no naturalmente codificado. Por ejemplo, derivados de PEG que llevan alquino y azida para la reacción con cadenas laterales de aminoácidos se pueden utilizar para unir PEG a aminoácidos no naturalmente codificados como se describe en el presente documento. Si el aminoácido no naturalmente codificado comprende una azida, a continuación, el PEG contendrá típicamente un resto alquino para efectuar la formación del producto de cicloadición [3 + 2] o especies de PEG activado (es decir, éster, carbonato) que contienen un grupo fosfina para efectuar la formación del enlace amida. Alternativamente, si el aminoácido no naturalmente codificado comprende un alquino, entonces el PEG típicamente contendrá un resto azida para efectuar la formación del producto de cicloadición [3 + 2] Huisgen. Si el aminoácido no naturalmente codificado comprende un grupo carbonilo, el PEG comprenderá típicamente un nucleófilo potente (incluyendo pero no limitado a, una hidrazida, hidrazina, hidroxilamina, o funcionalidad semicarbazida) con el fin de efectuar la formación de los enlaces correspondientes hidrazona, oxima, y semicarbazona, respectivamente. En otras alternativas, una inversa de la orientación de los grupos reactivos descritos anteriormente se puede utilizar, es decir, un resto azida en el aminoácido no naturalmente codificado se puede hacer reaccionar con un derivado de PEG que contiene un alquino.

45 En algunas realizaciones, la variante de polipéptido de insulina con un derivado de PEG contiene una funcionalidad química que es reactivo con la funcionalidad química presente en la cadena lateral del aminoácido no naturalmente codificado.

50 La invención proporciona en algunas realizaciones derivados poliméricos que contienen azida y acetileno que comprenden una cadena principal de polímero soluble en agua que tiene un peso molecular medio de aproximadamente 800 Da a aproximadamente 100,000 Da. La cadena principal de polímero del polímero soluble en agua puede ser poli(etilenglicol). Sin embargo, se debe entender que una amplia variedad de polímeros solubles en agua incluyendo, pero no limitado a, poli(etilenglicol) y otros polímeros relacionados, incluyendo poli(dextrano) y poli(propiloenglicol), también son adecuados para uso en la práctica de esta invención y que el uso del término PEG o poli(etilenglicol) pretende abarcar e incluir todas dichas moléculas. El término PEG incluye, pero no se limita a, poli(etilenglicol) en cualquiera de sus formas, incluyendo PEG bifuncional, PEG de brazos múltiples, PEG derivatizado, PEG bifurcado, PEG ramificado, pendent PEG (es decir, PEG o polímeros relacionados que tienen uno o más grupos funcionales colgantes de la cadena principal del polímero), o PEG con enlaces degradables.

60 PEG es normalmente 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, lo que quiere decir que el PEG es capaz de coexistir con tejidos vivos u organismos sin causar daños. Más específicamente, el PEG es sustancialmente no inmunogénico, lo que quiere decir que el PEG no tiende a producir una respuesta inmune en el cuerpo. Cuando está unido 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

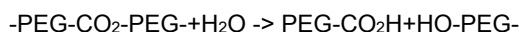
eliminar cualquier respuesta inmune de forma que un organismo puede tolerar la presencia del agente. Conjugados de PEG tienden a no producir una respuesta inmune sustancial o provocar coagulación u otros efectos indeseables. PEG que tiene la fórmula -- CH<sub>2</sub>CH<sub>2</sub>O--(CH<sub>2</sub>CH<sub>2</sub>O)<sub>n</sub> -- CH<sub>2</sub>CH<sub>2</sub>--, donde n es de aproximadamente 3 a aproximadamente 4.000, típicamente de aproximadamente 20 a aproximadamente 2.000, es adecuado para uso en la presente invención. PEG que tiene un peso molecular de aproximadamente 800 Da a aproximadamente 100.000 Da son en algunas realizaciones de la presente invención particularmente útiles como la cadena principal polimérica. El peso molecular del PEG puede ser de un intervalo amplio, incluyendo, pero no limitado a, entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. El peso molecular del PEG puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, incluyendo pero no limitado 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 PEG está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del PEG está entre aproximadamente 100 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da.

La cadena principal del polímero puede ser lineal o ramificada. Cadenas principales de polímeros ramificados son generalmente conocidos en la técnica. Típicamente, un polímero ramificado tiene un resto de núcleo ramificado central y una pluralidad de cadenas poliméricas lineales unidas al núcleo de rama central. PEG se utiliza comúnmente en formas ramificadas que se pueden preparar por adición de óxido de etileno a diversos polioles, tales como glicerol, oligómeros de glicerol, pentaeritritol y sorbitol. El resto de rama central también puede derivarse de varios aminoácidos, tales como lisina. El poli(etilenglicol)ramificado se puede representar en forma general como R(-PEG-OH)<sub>m</sub> en la que R se deriva de un resto del núcleo, tal como glicerol, oligómeros de glicerol, o pentaeritritol, y m representa el número de brazos. moléculas de brazos múltiples de PEG, tales como los descritos en la Pat. de EE.UU. Nos. 5.932.462; 5.643.575; 5.229.490; 4.289.872; Sol. de Pat. EE.UU. WO 96/21469; y WO 93/21259, también se pueden utilizar como la cadena principal del polímero.

PEG ramificado también puede ser en forma de un PEG en horquilla representado por PEG(--YCHZ<sub>2</sub>), donde Y es un grupo de enlace y Z es un grupo terminal activado unido a CH por una cadena de átomos de longitud definida.

Sin embargo, otra forma ramificada, el PEG colgante, tiene grupos reactivos, tales como carboxilo, a lo largo del esqueleto de PEG en lugar de al final de las cadenas de PEG.

Además de estas formas de PEG, el polímero se puede preparar también con uniones débiles o degradables en la cadena principal. Por ejemplo, PEG puede prepararse con uniones de éster en la cadena principal del polímero que están sujetas a la hidrólisis. Como se muestra a continuación, esta hidrólisis da como resultado la escisión del polímero en fragmentos de peso molecular inferior:



Se entiende por los expertos en la técnica que el término poli(etilenglicol) o PEG representa o incluye todas las formas conocidas en la técnica, incluyendo pero no limitado a los descritos en el presente documento.

Muchos otros polímeros también son adecuados para uso en la presente invención. En algunas realizaciones, cadenas principales de polímeros que son solubles en agua, con de 2 a aproximadamente 300 extremos, son particularmente útiles en la invención. Ejemplos de polímeros adecuados incluyen, pero no se limitan a, otros poli(alquilenglicoles), tales como poli(propiloenglicol) ("PPG"), copolímeros de los mismos (incluyendo pero no limitados a copolímeros de etilenglicol y propiloenglicol), terpolímeros de los mismos, mezclas de los mismos, y similares. Aunque el peso molecular de cada cadena de la cadena principal del polímero puede variar, es típicamente en el intervalo de aproximadamente 800 Da a aproximadamente 100.000 Da, a menudo desde aproximadamente 6.000 Da a aproximadamente 80.000 Da. El peso molecular de cada cadena de la cadena principal del polímero puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, incluyendo pero no limitado 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 de cada cadena de la cadena principal del polímero está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular de cada cadena de la cadena principal del polímero está entre aproximadamente 100 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular de cada cadena principal del polímero está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular de cada cadena de la cadena principal del polímero está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular de cada cadena de la cadena principal del polímero está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da.

Los expertos en la técnica reconocerán que la lista anterior para esqueletos sustancialmente solubles en agua es de ningún modo exhaustiva y es meramente ilustrativa, y que todos los materiales poliméricos que tienen las cualidades descritas anteriormente se contemplan como adecuadas para su uso en la presente invención.

En algunas realizaciones de la presente invención los derivados de polímeros son "multifuncionales", lo que significa que la cadena principal del polímero tiene al menos dos extremos, y posiblemente aproximadamente 300 extremos, funcionalizados o activados con un grupo funcional. Derivados de polímeros multifuncionales incluyen, pero no se limitan a, polímeros lineales que tienen dos extremos, cada terminal está unido a un grupo funcional que puede ser el mismo o diferente.

En una realización, el derivado de polímero tiene la estructura:

$$X-A-POLY-B-N=N=N$$

donde:

N=N=N es un resto azida;

B es un resto de unión, que puede estar presente o ausente;

POLY es un polímero no antigénico hidrosoluble;

A es un resto de unión, que puede estar presente o ausente y que puede ser el mismo que B o diferentes; y

X es un segundo grupo funcional.

Ejemplos de un resto de unión para A y B incluyen, pero no se limitan a, un grupo de alquilo funcionalizado por múltiples que contiene hasta 18, y puede contener entre 1-10 átomos de carbono. Un heteroátomo tal como nitrógeno, oxígeno o azufre puede ser incluido con la cadena de alquilo. La cadena de alquilo puede también ser ramificada en un heteroátomo. Otros ejemplos de un resto de unión para A y B incluyen, pero no se limitan a, un grupo arilo funcionalizado por múltiples, que contiene hasta 10 y puede contener 5-6 átomos de carbono. El grupo arilo puede estar sustituido con uno más átomos de carbono, nitrógeno, oxígeno o átomos de azufre. Otros ejemplos de grupos de unión adecuados incluyen los grupos de unión descritos en la patente de los Estados Unidos. Nos. 5.932.462; 5.643.575; y Publicación de Sol. de Pat. de EE.UU. 2003/0143596. Personas de experiencia ordinaria en la técnica reconocerán que la lista anterior de restos de unión es de ningún modo exhaustiva y es meramente ilustrativa, y que se contemplan todos los restos de unión que tienen las cualidades descritas anteriormente para ser adecuadas para uso en la presente invención.

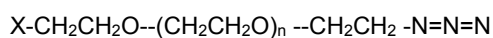
Los ejemplos de grupos funcionales adecuados para uso como X incluyen, pero no se limitan a, hidroxilo, hidroxilo protegido, alcoxilo, éster activo, tal como ésteres de N-hidroxisuccinimidilo y ésteres de 1-benzotriazolilo, carbonato activo, tal como carbonatos de N-hidroxisuccinimidilo y carbonatos 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, vinylsulfona, ditiopiridina, vinilpiridina, yodoacetamida, epóxido, glioxales, dionas, mesilatos, tosilatos, tresilato, alqueno, cetona, y azida. Como se entiende por los expertos ordinarios en la técnica, el resto X seleccionado debe ser compatible con el grupo azida de modo que no se produzca la reacción con el grupo azida. Los derivados poliméricos que contienen azida pueden ser homobifuncionales, lo que significa que el segundo grupo funcional (es decir, X) es también un resto azida, o heterobifuncional, lo que significa que el segundo grupo funcional es un grupo funcional diferente.

El término "protegido" se refiere a la presencia de un grupo protector o un resto que impide la reacción del grupo funcional químicamente reactivo en determinadas condiciones de reacción. El grupo protector variará dependiendo del tipo de grupo químicamente reactivo que se protege. Por ejemplo, si el grupo químicamente reactivo es una amina o una hidrazida, el grupo protector se puede seleccionar del grupo de terc-butiloxicarbonilo (tBoc) y 9-fluorenilometoxicarbonilo (Fmoc). Si el grupo químicamente reactivo es un tiol, el grupo protector puede ser ortopiridilodisulfuro. 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. Otros grupos protectores conocidos en la técnica también se pueden usar en la presente invención.

Los ejemplos específicos de grupos funcionales terminales en la bibliografía incluyen, pero no se limitan a, carbonato de N-succinimidilo (véase, por ejemplo, la patente de los Estados Unidos. Nos. 5.281.698, 5.468.478), amina (véase, por ejemplo, Buckmann et al. Makromol. Chem. 182: 1379 (1981), Zalipsky et al. Eur. Polym. J. 19: 1177 (1983)), hidrazida (véase, por ejemplo, Andrez et al. Makromol Chem 179: 301 (1978)), propionato de succinimidilo y butanoato de succinimidilo (véase, por ejemplo, Olson et al. en poli(etilenglicol) Chemistry & Biological Applications, pp 170-181, Harris & Zalipsky Eds, ACS, Washington, DC, 1997; véase también la patente

de los Estados Unidos. nº 5.672.662), succinato de succinimidilo (véase, por ejemplo, Abuchowski et al. *Cancer Biochem Biophys.* 7: 175 (1984) y Joppich et al. *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 (véase, por ejemplo, Pitha *et al.* *Eur. J Biochem.* 94:11 (1979), Elling *et al.*, *Biotech Appl Biochem* 13: 354 (1991), oxicarboniloimidazol (véase, por ejemplo, Beauchamp, et al., *Anal Biochem* 131: 25 (1983), Tondelli et al. *J. Controlled Release* 1: 251 (1985)), carbonato de p-nitrofenilo (véase, por ejemplo, Veronese, et al., *Appl. Biochem Biotech.*, 11: 141 (1985); y Sartore et al., *Appl Biochem Biotech*, 27:45 (1991)), aldehído (véase, por ejemplo, Harris et al. *J. Polym. Sci. Chem. Ed.* 22: 341 (1984), la patente de los Estados Unidos. nº 5.824.784, la patente de los Estados Unidos. nº 5.252.714), maleimida (véase, por ejemplo, Goodson et al. *Biotechnology (NY)* 8:343 (1990), Romani et al. en *Chemistry of Peptides and Proteins* 2:29 (1984)), y Kogan, *Synthetic Comm.* 22: 2417 (1992)), ortopiridilodisulfuro (véase, por ejemplo, Woghiren, et al. *Bioconj. Chem.* 4: 314 (1993)), acrilol (véase, por ejemplo, Sawhney et al., *Macromolecules*, 26: 581 (1993)), vinilsulfona (véase, por ejemplo, la patente de EE.UU. nº 5.900.461).

15 En ciertas realizaciones de la presente invención, los derivados poliméricos de la invención comprenden una cadena principal polimérica que tiene la estructura:

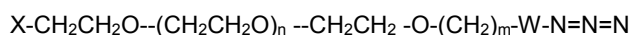


20 en la que:

X es un grupo funcional como se describe anteriormente; y

n es aproximadamente 20 a aproximadamente 4.000.

25 En otra realización, los derivados de polímeros de la invención comprenden una cadena principal de polímero que tiene la estructura:



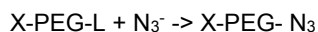
30 donde:

W es un resto enlazador alifático o aromático que comprende entre 1-10 átomos de carbono;

35 n es aproximadamente 20 a aproximadamente 4.000; y

X es un grupo funcional como se describe anteriormente. m está entre 1 y 10.

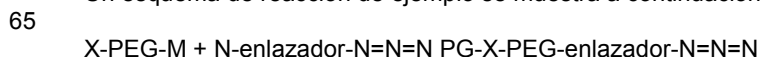
40 Los derivados de PEG que contienen azida de la invención pueden prepararse por una variedad de métodos conocidos en la técnica y/o descritos en este documento. En un método, mostrado a continuación, una cadena principal de polímero soluble en agua que tiene un peso molecular medio de aproximadamente 800 Da a aproximadamente 100.000 Da, el esqueleto del polímero que tiene un primer extremo unido a un primer grupo funcional y un segundo terminal unido a un grupo saliente adecuado, se hace reaccionar con un anión azida (que puede estar emparejado con cualquiera de un número de contraiones adecuados, incluyendo sodio, potasio, terc-butilamoníaco y así sucesivamente). El grupo saliente se somete a un desplazamiento nucleófilo y se sustituye por el resto azida, proporcionando el polímero de PEG deseado que contiene azida.



50 Como se muestra, una cadena principal de polímero adecuado para uso en la presente invención tiene la fórmula XPEG-L, en la que PEG es poli(etilenglicol) y X es un grupo funcional que no reacciona con los grupos azida y L es un grupo saliente adecuado. Los ejemplos de grupos funcionales adecuados incluyen, pero no se limitan a, hidroxilo, hidroxilo protegido, acetal, alquenilo, amina, aminooxi, amina protegida, hidrazida protegida, tiol protegido, ácido carboxílico, ácido carboxílico protegido, maleimida, ditiopiridina, y vinilpiridina, y la cetona. Ejemplos de grupos salientes adecuados incluyen, pero no se limitan a, cloruro, bromuro, yoduro, mesilato, tresilato, y tosilato.

55 En otro procedimiento para la preparación de los derivados de polímeros que contiene azida de la presente invención, un agente de unión teniendo una funcionalidad de azida se pone en contacto con una cadena principal de polímero soluble en agua que tiene un peso molecular medio de aproximadamente 800 Da a aproximadamente 100.000 Da, en donde el agente de enlace tiene una funcionalidad química que reaccionará selectivamente con una funcionalidad química en el polímero de PEG, para formar un producto derivado de polímero que contiene azida en el que la azida se separa de la cadena principal del polímero por un grupo de unión.

60 Un esquema de reacción de ejemplo se muestra a continuación:



donde:

5 PEG es poli(etilenglicol) y X es un grupo de terminación tal como alcoxi o un grupo funcional como se describió anteriormente; y

M es un grupo funcional que no es reactivo con la funcionalidad azida pero que reaccionará eficaz y selectivamente con el grupo funcional N.

10 Los ejemplos de grupos funcionales adecuados incluyen, pero no se limitan a, siendo M un ácido carboxílico, carbonato o éster activo si N es una amina; siendo M una cetona si N es una hidrazida o un resto aminooxi; siendo M un grupo saliente si N es un nucleófilo.

15 La purificación del producto crudo se puede realizar por métodos conocidos que incluyen, pero no se limitan a, la precipitación del producto seguido de cromatografía, si es necesario.

20 Un ejemplo más específico se muestra a continuación en el caso de PEG diamina, en la que una de las aminas está protegida por un resto de grupo protector tal como terc-butilo-Boc y la diamina PEG monoprottegida resultante se hace reaccionar con un resto de unión que lleva la funcionalidad azida:



25 En este ejemplo, el grupo amina puede acoplarse al grupo ácido carboxílico usando una variedad de agentes de activación tales como cloruro de tionilo o reactivos de carbodiimida y N-hidroxisuccinimida o N-hidroxibenzotriazol para crear un enlace amida entre la monoamina derivada de PEG y el resto enlazador que lleva azida. Después de la formación con éxito del enlace amida, el derivado que contiene azida protegido por N-terc-butilo-Boc resultante se puede utilizar directamente para modificar moléculas bioactivas o puede elaborarse adicionalmente para instalar otros grupos funcionales útiles. Por ejemplo, el grupo N-t-Boc se puede hidrolizar por tratamiento con ácido fuerte para generar omega-amino-PEG-azida. La amina resultante se puede utilizar como un mango sintético para instalar otra funcionalidad, como grupos maleimida, disulfuros activados, ésteres activados y así sucesivamente para la creación de valiosos reactivos heterobifuncionales.

30

35 Derivados heterobifuncionales son particularmente útiles cuando se desea unir diferentes moléculas a cada extremo del polímero. Por ejemplo, el PEG omega-N-amino-N-azido permitiría la unión de una molécula que tiene un grupo electrófilo activado, tal como un aldehído, cetona, éster activado, carbonato activado y así sucesivamente, a un extremo del PEG y una molécula que tiene un grupo acetileno al otro extremo del PEG.

En otra realización de la invención, el derivado de polímero tiene la estructura:



donde:

45 R puede ser H o un grupo alquilo, alqueno, alquiloxi, o grupo arilo o arilo sustituido;

B es un resto de unión, que puede estar presente o ausente;

POLY es un polímero no antigénico hidrosoluble;

50 A es un resto de unión, que puede estar presente o ausente y que puede ser el mismo que B o diferentes; y

X es un segundo grupo funcional.

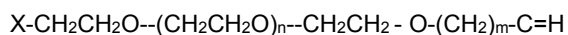
55 Ejemplos de un resto de unión para A y B incluyen, pero no se limitan a, un grupo alquilo funcionalizado por múltiples que contiene hasta 18, y pueden contener entre 1-10 átomos de carbono. Un heteroátomo tal como nitrógeno, oxígeno o azufre puede ser incluido con la cadena de alquilo. La cadena de alquilo puede también ser ramificada en un heteroátomo. Otros ejemplos de un resto de unión para A y B incluyen, pero no se limitan a, un grupo arilo funcionalizado por múltiples, que contiene hasta 10 y puede contener 5-6 átomos de carbono. El grupo arilo puede estar sustituido con átomos de uno más de carbono, nitrógeno, oxígeno, o átomos de azufre. Otros ejemplos grupos de unión adecuados incluyen aquellos grupos de unión descritos en la patente de los Estados Unidos. Nos. 5.932.462 y 5.643.575 y Publicación de Sol. de Pat. de EE.UU. 2003/0143596. Personas de experiencia ordinaria en la técnica reconocerán que la lista anterior de restos de unión es de ningún modo exhaustiva y pretende ser meramente ilustrativa, y que una amplia variedad de restos de unión que tienen las cualidades descritas anteriormente se contemplan para ser útiles en la presente invención.

60

65 Los ejemplos de grupos funcionales adecuados para uso como X incluyen hidroxilo, hidroxilo protegido, alcoxi,

éster activo, tal como ésteres de N-hidroxisuccinimidilo y ésteres 1-benzotriazolilo, carbonato activo, tales como carbonatos N-hidroxisuccinimidilo y carbonatos 1-benzotriazolilo, acetal, aldehído, hidratos de aldehído, alqueno, acrilato, metacrilato, acrilamida, sulfona activa, amina, aminooxi, amina protegida, hidrazida, hidrazida protegida, tior protegido, ácido carboxílico, ácido carboxílico protegido, isocianato, isotiocianato, maleimida, vinilsulfona, ditiopiridina, vinilpiridina, yodoacetamida, epóxido, glioxales, dionas, mesilatos, tosيلات, y tresilato, alqueno, cetona, y acetileno. Como se entenderá, el resto X seleccionado debe ser compatible con el grupo acetileno de modo que no se produzca la reacción con el grupo acetileno. Los derivados poliméricos que contienen acetileno pueden ser homobifuncionales, es decir, que el segundo grupo funcional (es decir, X) es también un resto acetileno, o heterobifuncional, lo que significa que el segundo grupo funcional es un grupo funcional diferente.

En otra realización de la presente invención, los derivados de polímeros comprenden una cadena principal de polímero que tiene la estructura:



en donde:

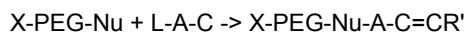
X es un grupo funcional como se describe anteriormente;

n es aproximadamente 20 a aproximadamente 4.000; y

m es entre 1 y 10.

Los ejemplos específicos de cada uno de los polímeros de PEG heterobifuncionales se muestran a continuación.

Los derivados de PEG que contienen acetileno de la invención se pueden preparar usando métodos conocidos por personas de experiencia ordinaria en la técnica y/o descritos en este documento. En un método, una cadena principal de polímero soluble en agua que tiene un peso molecular medio de aproximadamente 800 Da a aproximadamente 100.000 Da, el esqueleto del polímero que tiene un primer extremo unido a un primer grupo funcional y un segundo terminal unido a un grupo nucleófilo adecuado, se hace reaccionar con un compuesto que lleva tanto una funcionalidad de acetileno y un grupo saliente que es adecuado para la reacción con el grupo nucleófilo en el PEG. Cuando el polímero de PEG que lleva el resto nucleófilo y la molécula que lleva el grupo saliente se combinan, el grupo saliente se somete a un desplazamiento nucleófilo y se sustituye por el resto nucleófilo, proporcionando el polímero que contiene acetileno deseado.



Como se muestra, una cadena principal de polímero preferido para uso en la reacción tiene la fórmula X-PEG-Nu, en la que PEG es poli(etilenglicol), Nu es un resto nucleófilo y X es un grupo funcional que no reacciona con Nu, L o la funcionalidad de acetileno.

Ejemplos de Nu incluyen, pero no se limitan a, amina, alcoxi, ariloxi, sulfhidrilo, imino, carboxilato, hidrazida, grupos aminoxi que reaccionarían principalmente a través de un mecanismo de tipo SN<sub>2</sub>. Ejemplos adicionales de grupos Nu incluyen aquellos grupos funcionales que reaccionarían principalmente a través de una reacción de adición nucleófila. Ejemplos de grupos L incluyen cloruro, bromuro, yoduro, mesilato, tresilato, y tosيلات y otros grupos que se espera que se sometan a desplazamiento nucleófilo, así como cetonas, aldehídos, tioésteres, olefinas, grupos carbonilo insaturados Alfa-Beta, carbonatos y otros grupos electrófilos que se espera que se sometan a la adición por nucleófilos.

En otra realización de la presente invención, A es un enlazador alifático de entre 1-10 átomos de carbono o un anillo de arilo sustituido de entre 6-14 átomos de carbono. X es un grupo funcional que no reacciona con los grupos azida y L es un grupo saliente adecuado.

En otro procedimiento para la preparación de los derivados de polímeros que contiene acetileno de la invención, un polímero de PEG que tiene un peso molecular medio de aproximadamente 800 Da a aproximadamente 100.000 Da, teniendo un grupo funcional protegido o un agente de protección terminal en un extremo y un grupo saliente adecuado en el otro extremo se pone en contacto por un anión de acetileno.

Un esquema de reacción de ejemplo se muestra a continuación:



donde:

PEG es poli(etilenglicol) y X es un grupo de terminación tal como alcoxi o un grupo funcional como se describe anteriormente; y

R' es H, un alquilo, alcoxi, arilo o un grupo ariloxi o un alquilo sustituido, alcoxilo, arilo o un grupo ariloxi.

5 En el ejemplo anterior, el grupo saliente L debe ser suficientemente reactivo para experimentar un desplazamiento de tipo SN<sub>2</sub> cuando entra en contacto con una concentración suficiente del anión de acetileno. Las condiciones de reacción requeridas para llevar a cabo el desplazamiento de SN<sub>2</sub> de grupos salientes por aniones de acetileno son conocidos por personas de experiencia ordinaria en la técnica.

10 La purificación del producto crudo por lo general se puede lograr por métodos conocidos en la técnica incluyendo, pero no se limitan a, la precipitación del producto seguido de cromatografía, si es necesario.

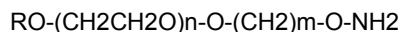
15 Polímeros solubles en agua pueden estar unidos a los polipéptidos de insulina de la invención. Los polímeros solubles en agua pueden estar unidos a través de un aminoácido no naturalmente codificado incorporado en el polipéptido de insulina o cualquier grupo funcional o sustituyente de un aminoácido no naturalmente codificado o codificado de forma natural, o cualquier grupo funcional o sustituyente añadido a un aminoácido no naturalmente codificado o codificado de forma natural. Alternativamente, los polímeros solubles en agua están ligados a un polipéptido de insulina que incorporan un aminoácido no naturalmente codificado a través de un aminoácido en estado natural (incluyendo pero no limitado a, cisteína, lisina o el grupo amino del residuo N-terminal). En algunos casos, los polipéptidos de insulina de la invención comprenden 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 aminoácidos no naturales, en el que uno o más aminoácidos no naturalmente codificados están relacionados con el agua de polímero soluble (incluyendo pero no limitado a, PEG y/o oligosacáridos). En algunos casos, los polipéptidos de insulina de la invención comprenden además 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, o más aminoácidos naturalmente codificados ligados a polímeros solubles en agua. En algunos casos, los polipéptidos de insulina de la invención comprenden uno o más aminoácidos no naturalmente codificados ligados a polímeros solubles en agua y uno o más aminoácidos en estado natural ligados a polímeros solubles en agua. En algunas realizaciones, los polímeros solubles en agua usados en la presente invención potencian la vida media en suero del polipéptido de insulina con respecto a la forma no conjugada.

30 El número de polímeros solubles en agua unido a un polipéptido de insulina (es decir, el grado de PEGilación o glucosilación) de la presente invención se puede ajustar para proporcionar una característica alterada (incluyendo pero no limitado a, aumentada o disminuida) farmacológica, farmacocinética o farmacodinámica tales como vida media *in vivo*. En algunas realizaciones, se aumenta la vida media de la insulina al menos aproximadamente 10, 20, 30, 40, 50, 60, 70, 80, 90 por ciento, 2 veces, 5 veces, 6 veces, 7 veces, 8 veces, 9 veces, 10 veces, 11 veces, 12 veces, 13 veces, 14 veces, 15 veces, 16 veces, 17 veces, 18 veces, 19 veces, 20 veces, 25 veces, 30 veces, 35 veces, 40 veces, 50 veces, o al menos aproximadamente 100 veces más de un polipéptido no modificado.

**Derivados de PEG que contienen un grupo nucleófilo fuerte (es decir, hidroxilamina, hidrazina, hidrazida o semicarbazida)**

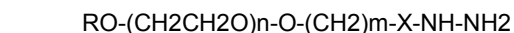
40 En una realización de la presente invención, un polipéptido de insulina que comprende un aminoácido no naturalmente codificado que contiene carbonilo se modifica con un derivado de PEG que contiene una hidrazina terminal, hidroxilamina, hidrazida o resto de semicarbazida que se enlaza directamente a la cadena principal de PEG.

45 En algunas realizaciones, el derivado de PEG hidroxilamina-terminal tendrá la estructura:



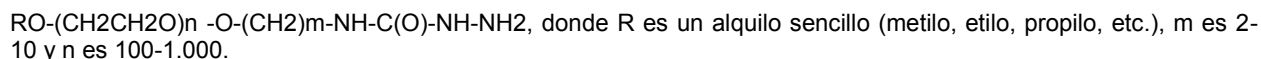
50 donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000 (es decir, el peso molecular promedio es de entre 5-40 kDa).

En algunas realizaciones, el derivado de PEG que contiene hidrazina o hidrazida que tendrá la estructura:



donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000 y X es opcionalmente un grupo carbonilo (C=O) que puede estar presente o ausente.

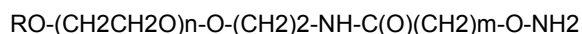
60 En algunas realizaciones, el derivado de PEG que contiene semicarbazida tendrá la estructura:



65 En otra realización de la invención, un polipéptido de insulina que comprende un aminoácido que contiene carbonilo se modifica con un derivado de PEG que contiene una hidroxilamina terminal, hidrazida, hidrazina, o resto de semicarbazida que está vinculada a la cadena principal de PEG por medio de un enlace amida.

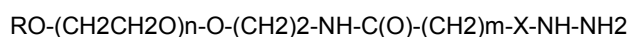


En algunas realizaciones, los derivados de PEG hidroxilamina-terminal tienen la estructura:



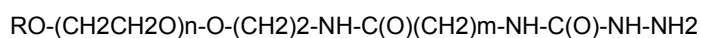
donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000 (es decir, el peso molecular promedio es de entre 5-40 kDa).

En algunas realizaciones, los derivados de PEG que contienen hidrazina o hidrazida que tienen la estructura:



donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10, n es 100-1.000 y X es opcionalmente un grupo carbonilo (C=O) que puede estar presente o ausente.

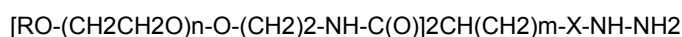
En algunas realizaciones, los derivados de PEG que contienen semicarbazida tienen la estructura:



donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000.

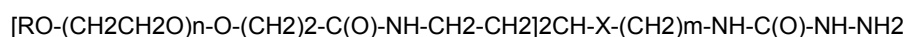
En otra realización de la invención, un polipéptido de insulina que comprende un aminoácido que contiene carbonilo se modifica con un derivado de PEG ramificado que contiene una hidrazina terminal, hidroxilamina, hidrazida o un resto de semicarbazida, con cada cadena del PEG ramificado que tiene un MW en el intervalo de 10-40 kDa y, puede ser de 5-20 kDa.

En otra realización de la invención, un polipéptido de insulina que comprende un aminoácido no naturalmente codificado se modifica con un derivado de PEG que tiene una estructura ramificada. Por ejemplo, en algunas formas de realización, el derivado de PEG terminal de hidrazina o hidrazida tendrá la siguiente estructura:



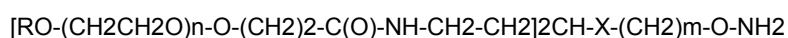
donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000, y X es opcionalmente un grupo carbonilo (C=O) que puede estar presente o ausente.

En algunas realizaciones, los derivados de PEG que contienen un grupo semicarbazida tendrá la estructura:



donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), X es opcionalmente NH, O, S, C(O) o no está presente, m es 2-10 y n es 100-1.000.

En algunas realizaciones, los derivados de PEG que contienen un grupo hidroxilamina tendrá la estructura:



donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), X es opcionalmente NH, O, S, C(O) o no presente, m es 2-10 y n es 100-1.000.

El grado y sitios en los que el polímero soluble en agua está unido al polipéptido de insulina pueden modular la unión del polipéptido de la insulina al receptor de polipéptido de insulina. En algunas realizaciones, los enlaces están dispuestos de tal manera que el polipéptido de insulina se une al receptor polipéptido de insulina con una Kd de aproximadamente 400 nM o inferior, con una KD de 150 nM o inferior, y en algunos casos con una KD de 100 nM o inferior, tal como se mide mediante un ensayo de unión en equilibrio, tal como el descrito en Spencer *et al.*, J. Biol. Chem., 263: 7862-7867 (1988).

Métodos y química para la activación de polímeros, así como para la conjugación de péptidos se describen en la literatura y son conocidos en la técnica. Métodos utilizados comúnmente para la activación de polímeros incluyen, pero no se limitan a, la activación de grupos funcionales con bromuro de cianógeno, peryodato, glutaraldehído, biepóxidos, epíclorhidrina, divinilsulfona, carbodiimida, haluros de sulfonilo, triclorotriazina, etc. (véase, R. F. Taylor, (1991), PROTEIN IMMOBILISATION. FUNDAMENTAL AND APPLICATIONS, Marcel Dekker, N.Y.; S. S. Wong, (1992), CHEMISTRY OF PROTEIN CONJUGATION AND CROSSLINKING, CRC Press, Boca Raton; G. T. Hermanson *et al.*, (1993), IMMOBILIZED AFFINITY LIGAND TECHNIQUES, Academic Press, N.Y.; Dunn, R.L., *et al.*, Eds. POLYMERIC DRUGS AND DRUG DELIVERY SYSTEMS, ACS Symposium Series Vol. 469, American Chemical Society, Washington, DC, 1991).

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

Los métodos para la activación de polímeros también se pueden encontrar en el documento WO 94/17039, patente de Estados Unidos. n° 5.324.844, WO 94/18247, WO 94/04193, patente de los Estados Unidos. n° 5.219.564, patente de Estados Unidos. n° 5.122.614, WO 90/13540, patente de Estados Unidos. n° 5.281.698, y WO 93/15189, y para la conjugación entre polímeros activados y enzimas incluyendo pero no limitado a factor de coagulación VIII (documento WO 94/15625), hemoglobina (documento WO 94/09027), molécula que transporta oxígeno (Pat. de EE.UU. n° 4.412.989), ribonucleasa y dismutasa de superóxido (Veronese *et al.*, *Sol. Biochem. Biotech.* 11: 141-52 (1985)).

PEGilación (es decir, la adición de cualquier polímero soluble en agua) de polipéptidos de insulina que contienen un aminoácido no naturalmente codificado, tal como p-azido-L-feniloalanina, se lleva a cabo por cualquier método conveniente. Por ejemplo, el polipéptido de insulina está PEGilado con un derivado de mPEG terminado con alquino.

Brevemente, se añade un exceso de mPEG(5000)-O-CH<sub>2</sub>-C≡CH sólido, con agitación, a una solución acuosa de polipéptido de insulina que contiene p-azido-L-Phe a temperatura ambiente. Típicamente, la solución acuosa está tamponada con un tampón que tiene un pKa cerca del pH en el que la reacción se debe llevar a cabo (generalmente aproximadamente pH 4-10). Los ejemplos de tampones adecuados para la pegilación a pH 7,5, por ejemplo, incluyen, pero no se limitan a, HEPES, fosfato, borato, TRIS-HCL, EPPS, y TES. El pH se controla continuamente y se ajusta si es necesario. La reacción típicamente se permite que continúe durante entre aproximadamente 148 horas.

Los productos de reacción se someten a continuación a cromatografía de interacción hidrófoba para separar las variantes de polipéptido de insulina PEGiladas de mPEG(5000)-O-CH<sub>2</sub>-C≡CH libre y cualesquiera complejos de peso highmolecular del polipéptido de insulina pegilado que pueden formarse cuando PEG desbloqueado se activa en ambos extremos de la molécula, reticulando de este modo las moléculas de variante de polipéptido de insulina. Las condiciones durante la cromatografía de interacción hidrófoba son tales que el mPEG(5000)-O-CH<sub>2</sub>-C≡CH libre fluye a través de la columna, mientras que cualesquiera complejos variantes de polipéptido de insulina PEGilados reticulados eluyen después de las formas deseadas, que contienen una molécula de variante de polipéptido de insulina conjugada con uno o más grupos de PEG. Las condiciones adecuadas varían dependiendo de los tamaños relativos de los complejos entrecruzados en comparación con los conjugados deseados y se determinan fácilmente por los expertos normales en la técnica. El eluyente que contiene los conjugados deseados se concentra por ultrafiltración y se desaló por diafiltración.

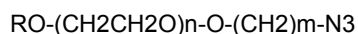
Si es necesario, el polipéptido de insulina PEGilado obtenido de la cromatografía hidrófoba se puede purificar adicionalmente por uno o más procedimientos conocidos por personas de experiencia ordinaria en la técnica que incluyen, pero no se limitan a, cromatografía de afinidad; cromatografía de intercambio aniónico o catiónico (utilizando, incluyendo, pero no limitado a, DEAE SEPHAROSE); cromatografía sobre sílice; HPLC de fase inversa; filtración en gel (usando, incluyendo pero no limitado a, SEPHADEX G-75); cromatografía de interacción hidrófoba; cromatografía de exclusión de tamaño, cromatografía de quelato metálico; ultrafiltración/diafiltración; precipitación con etanol; precipitación con sulfato de amoníaco; cromatografía de desplazamiento; procedimientos electroforéticos (incluyendo pero no limitado a enfoque isoelectrico preparativo), solubilidad diferencial (incluyendo pero no limitado a, precipitación con sulfato de amoníaco), o extracción. El peso molecular aparente puede ser estimado por GPC por comparación con patrones de proteínas globulares (Preneta, AZ en *PROTEIN PURIFICATION METHODS, A PRACTICAL APPROACH* (Harris & Angal, Eds.) IRL Press 1989, 293-306). La pureza del conjugado de insulina-PEG puede ser evaluada por la degradación proteolítica (incluyendo pero no limitado a, escisión con tripsina), seguido por análisis de espectrometría de masas. Pepinsky R.B., *et al.*, *J. Pharmacol. & Exp. Ther.* 297(3):1059-66 (2001).

Un polímero soluble en agua unido a un aminoácido de un polipéptido de insulina de la invención se puede derivatizar o sustituir sin limitación adicional.

#### **Derivados de PEG que contiene azida**

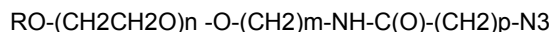
[534] En otra realización de la invención, un polipéptido de insulina se modifica con un derivado de PEG que contiene un resto azida que reaccionará con un resto alquino presente en la cadena lateral del aminoácido no naturalmente codificado. En general, los derivados de PEG tienen un peso molecular medio que varía de 1-100 kDa y, en algunas realizaciones, desde 10-40 kDa.

[535] En algunas realizaciones, el derivado de PEG azida-terminal tendrá la estructura:



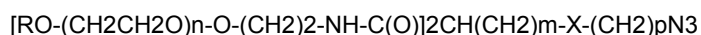
donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000 (es decir, el peso molecular promedio es de entre 5-40 kDa).

[536] En otra realización, el derivado de PEG azida-terminal tendrá la estructura:



donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10, p es 2-10 y n es 100-1.000 (es decir, el peso molecular promedio es de entre 5-40 kDa).

En otra realización de la invención, un polipéptido de insulina que comprende un aminoácido que contiene alquino se modifica con un derivado de PEG ramificado que contiene un resto de azida terminal, con cada cadena del PEG ramificado que tiene un PM en el intervalo de 10-40 kDa y puede ser de 520 kDa. Por ejemplo, en algunas formas de realización, el derivado de PEG azida-terminal tendrá la siguiente estructura:

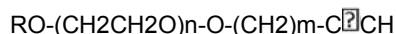


donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10, p es de 2-10, y n es 100-1.000, y X es opcionalmente un grupo O, N, S o carbonilo (C=O), en cada caso, que puede estar presente o ausente.

Derivados PEG que contiene alquino

En otra realización de la invención, un polipéptido de insulina se modifica con un derivado de PEG que contiene un resto alquino que reaccionará con un resto azida presente en la cadena lateral del aminoácido no naturalmente codificado.

En algunas realizaciones, el derivado de PEG alquino-terminal tendrá la siguiente estructura:

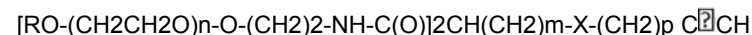


donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000 (es decir, el peso molecular promedio es de entre 5-40 kDa).

En otra realización de la invención, un polipéptido de insulina que comprende un aminoácido no naturalmente codificado que contiene alquino se modifica con un derivado de PEG que contiene un resto azida terminal o alquino terminal que está vinculado a la cadena principal de PEG por medio de un enlace amida.

En algunas realizaciones, el derivado de PEG alquino-terminal tendrá la siguiente estructura:  $\text{RO}-(\text{CH}_2\text{CH}_2\text{O})_n-\text{O}-(\text{CH}_2)_m-\text{NH}-\text{C}(\text{O})-(\text{CH}_2)_p-\text{C}\equiv\text{CH}$  donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10, p es 2-10 y n es 100-1.000.

[542] En otra realización de la invención, un polipéptido de insulina que comprende un aminoácido que contiene azida se modifica con un derivado de PEG ramificado que contiene un resto alquino terminal, con cada cadena del PEG ramificado que tiene un MW en el intervalo de 1,040 kDa y puede ser de 520 kDa. Por ejemplo, en algunas formas de realización, el derivado de PEG alquino-terminal tendrá la siguiente estructura:

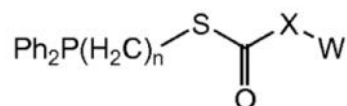


donde R es un alquilo sencillo (metilo, etilo, propilo, etc.), m es 2-10, p es 2-10, y n es 100-1.000, y X es opcionalmente un grupo O, N, S o carbonilo (C=O), o no está presente.

#### **Derivados de PEG que contienen fosfina**

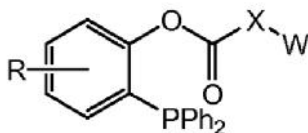
En otra realización de la invención, un polipéptido de insulina se modifica con un derivado de PEG que contiene un grupo funcional activado (incluyendo pero no limitado a, éster, carbonato) que comprende además un grupo arilo-fosfina que reaccionará con un resto azida presente en la cadena lateral del aminoácido no naturalmente codificado. En general, los derivados de PEG tienen un peso molecular medio que varía de 1-100 kDa y, en algunas realizaciones, desde 10-40 kDa.

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



en donde n es 1-10; X puede ser O, N, S o no está presente, Ph es fenilo, y W es un polímero soluble en agua.

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



en la que X puede ser O, N, S o no está presente, Ph es fenilo, W es un polímero soluble en agua y R puede ser H, alquilo, arilo, alquilo sustituido y grupos arilo sustituidos. Los grupos R ejemplares incluyen pero no se limitan a CH<sub>2</sub>, C(CH<sub>3</sub>), -OR', -NR'R'', -SR', -halógeno, -C(O)R', -CONR'R'', -S(O)<sub>2</sub>R', -S(O)<sub>2</sub>NR'R'', -CN y -NO<sub>2</sub>. R', R'', R''' y R'''' cada uno independientemente se refieren a hidrógeno, heteroalquilo sustituido o no sustituido, arilo sustituido o no sustituido, incluyendo pero no limitado a, arilo sustituido con 1-3 halógenos, alquilo sustituido o no sustituido, alcoxi o tioalcoxi, o grupos arilalquilo. Cuando un compuesto de la invención incluye más de un grupo R, por ejemplo, cada uno de los grupos R se selecciona independientemente como son cada uno grupos R', R'', R''' y R'''' cuando más de uno de estos grupos está presente. Cuando R' y R'' están unidos al mismo átomo de nitrógeno, se pueden combinar con el átomo de nitrógeno para formar un anillo de 5, 6, o 7 miembros. Por ejemplo, -NR'R'' está destinado a incluir, pero no limitarse a, 1-pirrolidinilo y 4-morfolinilo. De la discusión anterior de sustituyentes, un experto en la técnica entenderá que el término 'alquilo' se entiende que incluye grupos que incluyen átomos de carbono unidos a grupos distintos de grupos hidrógeno, tales como haloalquilo (incluyendo pero no limitados a, -CF<sub>3</sub> y -CH<sub>2</sub>CF<sub>3</sub>) y acilo (incluyendo, pero no limitado a, -C(O)CH<sub>3</sub>, -C(O)CF<sub>3</sub>, -C(O)CH<sub>2</sub>OCH<sub>3</sub>, y similares).

#### **Otros derivados de PEG y técnicas de PEGilación generales**

Otras moléculas de PEG ejemplares que pueden estar enlazadas a polipéptidos de insulina, así como los métodos de PEGilación incluyen, pero no se limitan a, los descritos en, por ejemplo, la Publicación de Patente de Estados Unidos n° 2004/0001838; 2002/0052009; 2003/0162949; 2004/OCH3637; 2003/0228274; 2003/0220447; 2003/0158333; 2003/0143596; 2003/OH4647; 2003/0105275; 2003/0105224; 2003/0023023; 2002/0156047; 2002/0099133; 2002/0086939; 2002/0082345; 2002/0072573; 2002/0052430; 2002/0040076; 2002/0037949; 2002/0002250; 2001/0056171; 2001/0044526; 2001/0021763; Patente de Estados Unidos n° 6.646.110; 5.824.778; 5.476.653; 5.219.564; 5.629.384; 5.736.625; 4.902.502; 5.281.698; 5.122.614; 5.473.034; 5.516.673; 5.382.657; 6.552.167; 6.610.281; 6.515.100; 6.461.603; 6.436.386; 6.214.966; 5.990.237; 5.900.461; 5.739.208; 5.672.662; 5.446.090; 5.808.096; 5.612.460; 5.324.844; 5.252.714; 6.420.339; 6.201.072; 6.451.346; 6.306.821; 5.559.213; 5.747.646; 5.834.594; 5.849.860; 5.980.948; 6.004.573; 6.129.912; WO 97/32607, EP 229108, EP 402378, WO 92/16555, WO 94/04193, WO 94/14758, WO 94/17039, WO 94/18247, WO 94/28024, WO 95/00162, WO 95/11924, WO95/13090, WO 95/33490, WO 96/00080, WO 97/18832, WO 98/41562, WO 98/48837, WO 99/32134, WO 99/32139, WO 99/32140, WO 96/40791, WO 98/32466, WO 95/06058, EP 439 508, WO 97/03106, WO 96/21469, WO 95/13312, EP 921 131, WO 98/05363, EP 809 996, WO 96/41813, WO 96/07670, EP 605 963, EP 510 356, EP 400 472, EP 183 503 y EP 154 316. Cualquiera de las moléculas PEG descritas en el presente documento pueden usarse en cualquier forma, incluyendo pero no limitado a la cadena sencilla, cadena ramificada, cadena de brazos múltiples, monofuncional, bifuncional, multifuncional, o cualquier combinación de los mismos.

Derivados de polímeros y de PEG adicionales incluyendo, pero no limitado a, derivados de PEG de hidroxilamina (aminooxi), se describen en las siguientes solicitudes de patente: Publicación de Patente de Estados Unidos n° 2006/0194256, la Publicación de Patente de Estados Unidos n° 2006/0217532, la patente US n° de publicación 2006/0217289, la patente provisional de EE.UU. n° 60/755.338; Patente de Estados Unidos provisional n° 60/755.711; Patente de Estados Unidos provisional n° 60/755.018; Solicitud de Patente Internacional n° PCT/USO6/49397; WO 2006/069246; Patente de Estados Unidos provisional n° 60/743.041; Patente de Estados Unidos provisional n° 60/743.040; Solicitud de Patente Internacional n° PCT/USO6/47822; Patente de Estados Unidos provisional n° 60/882.819; Patente de Estados Unidos provisional n° 60/882.500; y patente de Estados Unidos provisional n° 60/870.594.

#### **Proteínas de fusión heterólogas Fc**

Los compuestos de insulina descritos anteriormente pueden ser fusionados directamente o mediante un enlazador peptídico a la parte Fc de una inmunoglobulina. Las inmunoglobulinas son moléculas que contienen cadenas polipeptídicas unidas por enlaces de disulfuro, que tienen típicamente dos cadenas ligeras y dos cadenas pesadas. En cada cadena, un dominio (V) tiene una secuencia de aminoácidos variable en función de la especificidad del anticuerpo de la molécula. Los otros dominios (C) tienen una secuencia bastante constante común a moléculas de la misma clase.

Tal como se usa en el presente documento, la porción Fc de una inmunoglobulina tiene el significado dado comúnmente al término en el campo de la inmunología. Específicamente, este término se refiere a un fragmento de anticuerpo que se obtiene eliminando las dos regiones de unión a antígeno (los fragmentos Fab) del anticuerpo. Una forma de eliminar los fragmentos Fab consiste en digerir la inmunoglobulina con la proteasa de papaína. Por lo tanto, la porción de Fc se forma a partir de fragmentos de tamaño aproximadamente iguales de la región constante de ambas cadenas pesadas, que se asocian a través de interacciones no covalentes y enlaces de disulfuro. La porción Fc puede incluir las regiones bisagra y extenderse a través de los dominios CH<sub>2</sub> y CH<sub>3</sub> al término C del anticuerpo. Regiones bisagra representativas para las inmunoglobulinas humanas y de ratón se pueden encontrar en Antibody Engineering, A Practical Guide, Borrebaeck, C. A. K., ed., W. H. Freeman and Co., 1992. La porción de Fc puede incluir además uno o más sitios de glicosilación. Las secuencias de aminoácidos de proteínas Fc numerosas representativas que contienen una región bisagra, los dominios CH<sub>2</sub> y CH<sub>3</sub>, y un sitio de N-glicosilación son bien conocidos en la técnica.

Hay cinco tipos de regiones de inmunoglobulina humana Fc con diferentes funciones efectoras y propiedades farmacocinéticas: IgG, IgA, IgM, IgD, e IgE. IgG es la inmunoglobulina más abundante en suero. IgG también tiene la vida media más larga en el suero de cualquier inmunoglobulina (23 días). A diferencia de otras inmunoglobulinas, IgG se recircula de manera eficiente después de la unión a un receptor de Fc. Hay cuatro subclases de IgG G1, G2, G3, G4 y, cada uno de los cuales tiene diferentes funciones efectoras. G1, G2, y G3 pueden unir Clq y fijar el complemento mientras que G4 no puede. Aunque G3 es capaz de unir Clq más eficientemente que G1, G1 es más eficaz en la mediación de la lisis celular de complemento dirigido. G2 fija el complemento de manera muy ineficiente. El sitio de unión de Clq en IgG se encuentra en la región carboxi terminal del dominio CH<sub>2</sub>.

Todas las subclases de IgG son capaces de unir a los receptores de Fc (CD16, CD32, CD64) siendo G1 y G3 más eficaces que G2 y G4. La región de unión al receptor Fc de IgG está formada por residuos localizados tanto en la bisagra como en las regiones carboxi terminales del dominio CH<sub>2</sub>.

IgA puede existir tanto en una forma monomérica como dimérica unida por una cadena J. IgA es la segunda Ig más abundante en suero, pero tiene una vida media de sólo 6 días. IgA tiene tres funciones efectoras. Se une a un receptor específico de IgA sobre macrófagos y eosinófilos, lo que conduce a fagocitosis y desgranulación, respectivamente, también se puede fijar el complemento a través de una ruta alternativa desconocida.

IgM se expresa como un pentámero o un hexámero, ambos de los cuales se mantienen unidos por una cadena J. IgM tiene una vida media en suero de 5 días. Se une débilmente a Clq a través de un sitio de unión localizado en su dominio CH<sub>3</sub>. IgD tiene una vida media de 3 días en el suero. No está claro qué funciones efectoras son atribuibles a esta Ig. IgE es una Ig monomérica y tiene una vida media en suero de 2,5 días. IgE se une a dos receptores de Fc que acciona la desgranulación y resulta en la liberación de agentes proinflamatorios.

Dependiendo del efecto deseado *in vivo*, las proteínas de fusión heterólogas de la presente invención pueden contener cualquiera de los isotipos descritos anteriormente o pueden contener regiones de Fc mutadas en las que las funciones de unión al complemento y/o receptores de Fc han sido alterados. Por lo tanto, las proteínas de fusión heterólogas de la presente invención pueden contener toda la porción de Fc de una inmunoglobulina, fragmentos de la porción Fc de una inmunoglobulina, o análogos de los mismos fusionados a un compuesto de interferón beta.

Las proteínas de fusión de la presente invención pueden consistir en proteínas de cadena sencilla o polipéptidos como múltiples cadenas. Dos o más proteínas de fusión Fc se pueden producir de tal manera que interactúan a través de enlaces de disulfuro que, naturalmente, se forman entre las regiones Fc. Estos multímeros pueden ser homogéneos con respecto al compuesto de interferón beta o pueden contener diferentes compuestos de interferón beta fusionados en el término N de la porción Fc de la proteína de fusión.

Independientemente de la estructura final de la proteína de fusión, la región Fc o de tipo Fc pueden servir para prolongar la vida media plasmática *in vivo* del compuesto del interferón beta fusionado al término N. Además, el componente de interferón beta de un compuesto de proteína de fusión debe conservar al menos una actividad biológica de interferón beta. Un aumento de la vida media terapéutica o de circulación se puede demostrar utilizando el método descrito en este documento o conocido en la técnica, en el que la vida media de la proteína de fusión se compara con la vida media del compuesto beta interferón solo. La actividad biológica se puede determinar por métodos *in vitro* e *in vivo* conocidos en la técnica.

Dado que la región Fc de IgG producida por proteólisis tiene la misma vida media *in vivo* que la molécula de IgG intacta y los fragmentos Fab son degradados rápidamente, se cree que la secuencia relevante para prolongar la vida media reside en los dominios CH<sub>2</sub> y/o CH<sub>3</sub>. Además, se ha demostrado en la literatura que las tasas catabólicas de las variantes de IgG que no se unen al receptor de Fc de afinidad alta o Clq son indistinguibles de la tasa de aclaramiento del anticuerpo de tipo silvestre de los padres, lo que indica que el sitio catabólico es distinto de los sitios involucrados en receptor de Fc o de unión a Clq. [Wawrzynszak et al., (1992) Molecular Immunology. 29: 221]. Estudios de mutagénesis de sitio dirigido utilizando una región IgG1 Fc murina sugirieron que el sitio de la región IgG1 Fc que controla la velocidad catabólica está localizada en la interfase de dominio CH<sub>2</sub>-CH<sub>3</sub>. Las regiones Fc pueden ser modificadas en el sitio catabólico para optimizar la vida media de las proteínas de fusión. La región Fc

utilizada para las proteínas de fusión de la presente invención puede derivarse de una IgG1 o una región IgG4 Fc, y puede contener las regiones tanto CH<sub>3</sub> como CH<sub>2</sub>, lo que incluye la región bisagra.

### **Proteínas de fusión de albúmina heteróloga**

La insulina descrita en este documento puede fusionarse directamente o mediante un conector peptídico, el polímero soluble en agua, o enlazador profármaco a la albúmina o un análogo, fragmento o derivado del mismo. Generalmente, las proteínas de albúmina que son parte de las proteínas de fusión de la presente invención pueden derivarse a partir de albúmina clonada a partir de cualquier especie, incluyendo humano. La albúmina sérica humana (HSA) se compone de una única cadena polipeptídica no glicosilada de 585 aminoácidos con un peso molecular de fórmula de 66.500. Se conoce la secuencia de aminoácidos de HSA humano [véase Meloun, *et al.* (1975) FEBS Letters 58: 136; Behrens, *et al.* (1975) Fed. Proc. 34: 591; Lawn, *et al.* (1981) Nucleic Acids Research 9: 6102-6114; Minghetti, *et al.* (1986) J. Biol. Chem. 261: 6747]. Se ha descrito una variedad de variantes polimórficas, así como análogos y fragmentos de albúmina. [Véase Weitkamp, *et al.*, (1973) Ann. Tararear. Gineta. 37: 219]. Por ejemplo, en EP 322094, varias formas más cortas de HSA. Algunos de estos fragmentos de HSA se describen, incluyendo HSA (1373), HSA (1388), HSA (1389), HSA (1369) y HSA (1419) y los fragmentos entre 1369 y 1419. El documento EP 399666 da a conocer fragmentos de albúmina que incluyen HSA (1177) y HSA (1200) y los fragmentos entre HSA (1177) y HSA (1200).

Se entiende que las proteínas de fusión heterólogas de la presente invención incluyen compuestos de insulina que están acoplados a cualquier proteína de albúmina incluyendo fragmentos, análogos y derivados en los que tal proteína de fusión es biológicamente activa y tiene una vida media en plasma más larga que el compuesto de insulina sola. Por lo tanto, la porción de la albúmina de la proteína de fusión no necesita necesariamente una vida media de plasma igual a la de la albúmina humana nativa. Los fragmentos, análogos y derivados son conocidos o pueden ser generados que tienen vidas medias más largas o tienen vidas medias intermedias a la de la albúmina humana nativa y el compuesto de insulina de interés.

Las proteínas de fusión heterólogas de la presente invención abarcan proteínas que tienen sustituciones conservadoras de aminoácidos en el compuesto de insulina y/o el Fc o porción de albúmina de la proteína de fusión. Una "sustitución conservadora" es la sustitución de un aminoácido con otro aminoácido que tiene la misma carga electrónica neta y aproximadamente el mismo tamaño y forma. Los aminoácidos con cadenas laterales alifáticas de aminoácidos sustituidos alifáticos o tienen aproximadamente el mismo tamaño cuando el número total de carbono y heteroátomos en sus cadenas laterales difiere en no más de aproximadamente cuatro. Tienen aproximadamente la misma forma cuando el número de ramificaciones en sus cadenas laterales difiere en no más de uno. Los aminoácidos con fenilo o grupos fenilo sustituidos en sus cadenas laterales se considera que tienen aproximadamente el mismo tamaño y forma. Salvo que se especifique otra cosa en la presente memoria, las sustituciones conservadoras se hacen preferiblemente con aminoácidos naturales.

Proteínas de albúmina y de inmunoglobulina de tipo silvestre pueden obtenerse de una variedad de fuentes. Por ejemplo, estas proteínas se pueden obtener de una biblioteca de ADNc preparada a partir de tejido o células que expresan el ARNm de interés a un nivel detectable. Bibliotecas se pueden cribar con sondas diseñadas usando el ADN publicado o la secuencia de proteínas para la proteína particular de interés. Por ejemplo, regiones constantes de inmunoglobulina ligera o cadena pesada se describen en Adams, *et al.* (1980) Biochemistry 19: 2711-2719; Goughet, *et al.* (1980) Biochemistry 19: 2702-2710; Dolby, *et al.* (1980) Proc. Natl. Acad. Sci. Estados Unidos 77: 6027-6031; Rice *et al.* (1982) Proc. Natl. Acad. Sci. Estados Unidos 79: 7862-7862; Falkner, *et al.* (1982) Nature 298: 286-288; y Morrison, *et al.* (1984) Ann. Rev. Immunol. 2: 239-256. Algunas referencias que describen secuencias de proteínas y de ADN de albúmina incluyen Meloun, *et al.* (1975) FEBS Letters 58: 136; Behrens, *et al.* (1975) Fed. Proc. 34: 591; Lawn, *et al.* (1981) Nucleic Acids Research 9: 6102-6114; y Minghetti, *et al.* (1986) J. Biol. Chem. 261: 6747.

### **Caracterización de las proteínas de fusión heterólogas de la presente invención**

Métodos numerosos existen para caracterizar las proteínas de fusión de la presente invención. Algunos de estos métodos incluyen, pero no se limitan a: SDS-PAGE junto con métodos de tinción de proteínas o inmunotransferencia usando anticuerpos anti-IgG o anti-HSA. Otros métodos incluyen espectrometría de masa de desorción/ionización de láser asistida por matriz (MALDI-MS), cromatografía líquida/espectrometría de masas, enfoque isoelectrico, intercambio aniónico analítico, cromatofluorescencia y difracción circular, por ejemplo.

### **Mejora de afinidad para albúmina de suero**

Varias moléculas también pueden fusionarse con los polipéptidos de insulina de la invención para modular la vida media de polipéptidos de insulina en el suero. En algunas realizaciones, las moléculas están unidas o fusionadas a los polipéptidos de insulina de la invención para mejorar la afinidad para la albúmina sérica endógena en un animal.

Por ejemplo, en algunos casos, se realiza una fusión recombinante de un polipéptido de insulina y una secuencia de unión a la albúmina. Secuencias de unión a albúmina ejemplares incluyen, pero no están limitadas al dominio de unión a albúmina de la proteína G estreptocócica (véase, por ejemplo, Makrides *et al.*, J. Pharmacol Exp Ther 277:

534-542 (1996) y Sjolander et al., J. Immunol. Methods 201: 115-123 (1997)), o péptidos de unión a albúmina tales como los descritos en, por ejemplo, Dennis, et al., J. Biol. Chem. 277: 35035-35043 (2002).

5 En otras realizaciones, los polipéptidos de insulina de la presente invención están acilados con ácidos grasos. En algunos casos, los ácidos grasos promueven la unión a la albúmina sérica. Véase, por ejemplo, Kurtzhals, *et al.*, Biochem. J. 312: 725-731 (1995).

10 En otras realizaciones, los polipéptidos de insulina de la invención se fusionan directamente con albúmina de suero (incluyendo pero no limitados a la albúmina de suero humano). Los expertos en la técnica reconocerán que una amplia variedad de otras moléculas también puede estar enlazada a la insulina en la presente invención para modular la unión a la albúmina sérica u otros componentes del suero.

### **Glicosilación de polipéptidos de insulina**

15 La invención incluye polipéptidos de insulina que incorporan uno o más aminoácidos no naturalmente codificados que llevan residuos de sacárido. Los residuos de sacárido pueden ser naturales (incluyendo pero no limitado a, n-acetilo-glucosamina) o no naturales (incluyendo pero no limitado a, 3-fluorogalactosa). Los sacáridos pueden estar unidos a los aminoácidos no naturalmente codificados ya sea por un enlace glicosídico enlazado por N o O (incluyendo pero no limitado a, N-acetilo-galactosa-L-serina) o un enlace no natural (incluyendo pero no limitado a, una oxima o el correspondiente glicósido enlazado por C o S).

20 Los restos de sacárido (incluyendo pero no limitado a, glicosilo) pueden añadirse a polipéptidos de insulina *in vivo* o *in vitro*. En algunas realizaciones de la invención, un polipéptido de insulina que comprende un aminoácido no naturalmente codificado que contiene carbonilo se modifica con un sacárido derivatizado con un grupo aminoóxi para generar el polipéptido glicosilado correspondiente unido a través de un enlace oxima. Una vez unido al aminoácido no naturalmente codificado, el sacárido puede ser elaborado adicionalmente por tratamiento con glicosiltransferasas y otras enzimas para generar un oligosacárido unido al polipéptido de insulina. Véase, por ejemplo, H. Liu, *et al.* J. Am. Chem. Soc. 125: 1702-1703 (2003).

25 En algunas realizaciones de la invención, un polipéptido de insulina que comprende un que contiene carbonilo aminoácido no naturalmente codificado se modifica directamente con un glicano con estructura definida preparado como un derivado de aminoóxi. Alguien de experiencia ordinaria en la técnica reconocerá que otras funcionalidades, incluyendo azida, alquino, hidrazida, hidrazina, y semicarbazida, se pueden usar para enlazar el sacárido al aminoácido no naturalmente codificado.

30 En algunas realizaciones de la invención, un polipéptido de insulina que comprende un aminoácido no naturalmente codificado que contiene azida o alquino a continuación se puede modificar por, incluyendo pero no limitado a, una reacción de cicloadición Huisgen [3 + 2] con, incluyendo pero no limitado a, alquino o derivados de azida, respectivamente. Este método permite que proteínas se modifiquen con extremadamente alta selectividad.

### **Dímeros y multímeros de insulina**

35 La presente invención también prevé combinaciones de insulina y análogos de insulina tales como homodímeros, heterodímeros, homomultímeros o heteromultímeros (es decir, trímeros, tetrámeros, etc.) donde la insulina que contiene uno o más aminoácidos no naturalmente codificados está enlazada a otra variante de insulina o insulina o cualquier otro polipéptido que no es insulina o variante de la insulina, ya sea directamente a la cadena principal del polipéptido o mediante un enlazador. Debido a su mayor peso molecular en comparación con los monómeros, el dímero de insulina o conjugados de multímeros pueden presentar propiedades nuevas o deseables, incluyendo pero no limitado a diferente vida media farmacológica, farmacocinética, farmacodinámica, terapéutica modulada, o vida media modulada en plasma con relación a la insulina monomérica. Para ejemplos de análogos de insulina monoméricos véase, por ejemplo, Balschmidt, P., *et al.*, Patente de Estados Unidos. nº 5.164.366, expedida el 17 de Nov., 1992; Brange, J., *et al.*, Patente de Estados Unidos. nº 5.618.913, expedida el 8 de abril 1997; Chance, R.E., *et al.*, Patente de Estados Unidos. nº 5.514.646, concedida el 7 de mayo de 1996; y Ertl, J., et al., número de publicación EPO 885.961, 23 de diciembre de 1998. Algunas realizaciones de la presente invención proporcionan análogos de insulina monoméricos que contienen uno o más residuos de aminoácidos no naturalmente codificados y en algunas realizaciones estos incluyen análogos de insulina monoméricos en los que la posición B28 es Asp, Lys, Ile, Leu, Val o Ala y el residuo aminoácido en la posición B29 es Lys o Pro; análogo de insulina monomérico con Lys (B28) Pro (B29)-insulina humana; análogo de insulina monomérico Asp (B28)-insulina humana; y análogo de insulina monomérico Lys (B3) Ile (B28)-insulina humana. En algunas realizaciones, los dímeros de insulina de la invención modulan la transducción de señal del receptor de insulina. En otras realizaciones, los dímeros de insulina o multímeros de la presente invención actúan como un antagonista del receptor de insulina, agonista, o modulador.

40 En algunas realizaciones, una o más de las moléculas de insulina presentes en una insulina contiene dímero o multímero comprende un aminoácido no naturalmente codificado unido a un polímero soluble en agua.

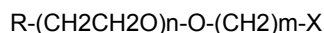
65 En algunas realizaciones, los polipéptidos de insulina están enlazados directamente, incluyendo, pero no limitado a,

a través de un enlace amida Asn-Lys o enlace disulfuro Cys-Cys. En algunas realizaciones, los polipéptidos de insulina, y/o la molécula no de insulina enlazada, comprenderá diferentes aminoácidos no naturalmente codificados para facilitar la dimerización, incluyendo, pero no limitado a, un alquino en un aminoácido no naturalmente codificado de un primer polipéptido de insulina y una azida en un segundo aminoácido no naturalmente codificado de una segunda molécula se puede conjugar a través de una cicloadición de Huisgen [3 + 2]. Alternativamente, la insulina, y/o la molécula no dependiente de insulina ligada que comprende un aminoácido que contiene cetona no naturalmente codificado puede conjugarse con un segundo polipéptido que comprende aminoácidos no naturalmente codificados que contienen hidroxilamina y se hacen reaccionar los polipéptidos a través de la formación de la oxima correspondiente.

De forma alternativa, los dos polipéptidos de insulina, y/o la molécula no de insulina enlazada, están enlazados a través de un enlazador. Cualquier enlazador hetero o homobifuncional se pueden utilizar para unir las dos moléculas, y/o las moléculas enlazadas no de insulina, que pueden tener la misma o diferente secuencia primaria. En algunos casos, el enlazador utilizado para enlazar la insulina, y/o las moléculas enlazadas no de insulina pueden ser un reactivo de PEG bifuncional. El enlazador puede tener una amplia gama de peso molecular o longitud molecular. Enlazadores de peso molecular mayores o menores se pueden usar para proporcionar una relación espacial deseada o conformación entre la insulina y la entidad enlazada o entre la insulina y su receptor, o entre la entidad vinculada y su pareja de unión, si existe. Los enlazadores que tienen una longitud molecular más larga o más corta pueden también utilizarse para proporcionar un espacio deseado o la flexibilidad entre la insulina y la entidad enlazada, o entre la entidad vinculada y su pareja de unión, si existe.

En algunas realizaciones, la invención proporciona enlazadores bifuncionales solubles en agua que tienen una estructura de mancuerna que incluye: a) una azida, un alquino, una hidrazina, una hidrazida, una hidroxilamina, o un resto que contiene carbonilo en al menos un primer extremo de una cadena principal polimérica; y b) al menos un segundo grupo funcional en un segundo extremo de la cadena principal del polímero. El segundo grupo funcional puede ser el mismo o diferente que el primer grupo funcional. El segundo grupo funcional, en algunas realizaciones, no es reactivo con el primer grupo funcional. La invención proporciona, en algunas realizaciones, los 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, la invención proporciona multímeros que comprenden uno o más polipéptidos de insulina, formados por reacciones con polímeros activados solubles en agua que tienen la estructura:



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

#### **Medición de actividad de polipéptido de insulina y afinidad de polipéptido de insulina para el receptor de insulina**

La actividad del polipéptido de la insulina se puede determinar usando ensayos estándar o conocidos *in vitro* o *in vivo*. Polipéptidos de insulina pueden ser analizados para la actividad biológica mediante procedimientos adecuados conocidos en la técnica. Tales ensayos incluyen, pero no se limitan a la activación de genes responsivos a interferon, ensayos de unión a receptor, ensayos de actividad antiviral, ensayos de inhibición del efecto citopático, (Familletti et al., Meth Enzymol 78: 387-394), ensayos antiproliferativos, (Aebersold y Sample, Meth Enzymol. 119: 579-582), ensayos inmunomodulatorios (Patente de Estados Unidos nº 4.914.033; 4.753.795.), y ensayos que monitorizan la inducción de moléculas de MHC (por ejemplo, Hokland et al., Meth Enzymol 119: 688-693), Como se describe en Meager, J. Immunol. Meth., 261: 21-36 (2002).

La captación de glucosa en adipocitos 3T3-1 puede evaluarse utilizando el siguiente método. Las células 3T3-L1 se obtuvieron de la American Type Culture Collection (ATCC, Rockville, Md.). Las células se cultivan en medio de crecimiento (GM) que contiene 10% de suero bovino fetal rico en hierro en medio de Eagle modificado por Dulbecco. Para la diferenciación de adipocitos estándar, 2 días después de que las células alcanzaron la confluencia (denominado día 0), las células se exponen a medio de diferenciación (DM) que contenía suero bovino fetal al 10%, 10 µg/ml de insulina, 1 µM dexametasona, y 0,5 µM de isobutilmetiloxantina, durante 48 horas. Después, las células se mantienen en medio de diferenciación puesto que contenía suero bovino fetal al 10%, y 10 µg/ml de insulina. La potencia *in vitro* se puede medir con los ensayos de captación de glucosa que son conocidos por los expertos normales en la técnica. La potencia *in vitro* se puede definir como la medida de la captación de glucosa de un



compuesto de insulina en un ensayo a base celular y es una medida de la potencia biológica del compuesto de insulina. Se puede expresar como el CE50 que es la concentración eficaz de compuesto que da como resultado 50% de actividad en un solo experimento de dosis-respuesta.

5 La captación de Ensayo de Transporte de Glucosa--Insulina Dependiente--Hexosa, como se ensayó por la acumulación de 0,1 mM 2-deoxi-D-[14C]glucosa, se mide del siguiente modo: 3T3-L1 adipocitos en placas de 12 pocillos se lavan dos veces con tampón KRP (NaCl 136 mM NaCl, 4,7 mM KCl, 10 mM de NaPO<sub>4</sub>, 0,9 mM de CaCl<sub>2</sub>, 0,9 mM de MgSO<sub>4</sub>, pH 7,4) se calentó a 37°C y que contiene 0,2% de BSA, se incubaron en medio L15 de Leibovitz que contiene 0,2% de BSA durante 2 horas a 37°C en el aire ambiente, se lavan dos veces de nuevo con KRP que contiene, 0,2% de tampón BSA, y se incubaron en KRP, 0,2% de tampón BSA en ausencia (Me<sub>2</sub>SO solamente) o presencia de wortmanina durante 30 minutos a 37°C en el aire ambiente. Se añade la insulina a continuación, a una concentración final de 100 nM durante 15 minutos, y la captación de 2-deoxi-D-[14C]glucosa se mide durante los últimos 4 minutos. La captación no específica, medida en presencia de 10 µM de citocalasina B, se resta de todos los valores. Las concentraciones de proteína se determinan con el ensayo de ácido bicinónico Pierce. La captación se mide de forma rutinaria por triplicado o cuadruplicado para cada experimento. El efecto del pretratamiento agudo y crónico de los adipocitos 3T3-L1 con FGF-21 en presencia de insulina puede ser investigado.

20 Ensayo de transporte de glucosa--insulina Independiente--fibroblastos de 3T3-L1 se sembraron en placas de 96 pocillos y se diferenciaron en células grasas (adipocitos) durante 2 semanas. Después de la diferenciación, se privan en medio libre de suero y se trataron con diversos polipéptidos de insulina de la presente invención durante 24 horas. Tras el tratamiento, las células se lavaron dos veces con tampón de KRBH, que contiene 0,1% de BSA. La captación de glucosa se lleva a cabo en presencia de glucosa marcada en tampón KPBH. Esto permite la evaluación cualitativa de una variedad de polipéptidos de insulina y análogos producidos por medio de la presente invención, y los que se han pegilado, ya que se sabe que la pegilación causa una disminución en la eficiencia de la molécula nativa, y compara la eficacia de diferentes insulinas. Además, los polipéptidos de insulina de la presente invención podrían mostrar que inducen la captación de glucosa en un modelo de tejido ex vivo.

30 En el modelo de transporte ex vivo de la glucosa, el ensayo de transporte de glucosa se describe del siguiente modo: Soluciones madre de tampón Krebs-Henseleit--Reserva 1: NaCl (1,16 M); KCl (0,046 M); KH<sub>2</sub>PO<sub>4</sub> (0,0116 M); NaHCO<sub>3</sub> (0,0253 M). Reserva 2: CaCl<sub>2</sub> (0,025 M); MgSO<sub>4</sub> (2H<sub>2</sub>O) (0,0116 M). BSA: Usar ICN Cohn Fracción V, BSA libre de ácido graso directamente sin dializante. Preparación de medios: Añadir 50 ml de reserva Krebs 1 a 395 ml de dH<sub>2</sub>O y gas con 95% de O<sub>2</sub>/5% de CO<sub>2</sub> durante 1 hora. Añadir 50 ml de solución madre 2 y llevar a 500 ml con dH<sub>2</sub>O destilada. Añadir 500 mg de BSA libre de ácido graso ICN. La preincubación e incubación de medios: 32 mM de Manitol, 8 mM de Glucosa. Medios de lavado: 40 mM de Manitol, 2 mM de piruvato. Medios de transporte: 39 mM de Manitol, 1 mM 2-DG; 32 mM de Manitol, 8 mM de 3-O-MG. Solución de insulina: (insulina porcina [Lilly] 100.000.000 µU/ml) a una concentración final de 2.000 µU/ml o 13,3 nM. Preparación de Medios de Marcador Radiactivo: actividades específicas utilizadas: 2DG=1,5 mCi/ml; 3O MG=437 µCi/ml; o, Manitol=8 µCi/ml. Las ratas se anestesiaron con 0,1 cc Nembutal por 100 g de peso corporal. El tejido muscular se escinde y se enjuaga en 0,9% de solución salina luego se coloca en medios de preincubación (2 ml) a 29°C durante 1 hora. El tejido muscular se transfiere a medio de incubación (2 ml; igual a preincubación excepto incluyendo insulina o compuesto de ensayo) y se incubó durante 30 minutos (depende de las condiciones experimentales). El tejido muscular luego se transfiere a medios de lavado (2 ml) durante 10 minutos a 29°C, luego se transfirieron a medios de etiqueta (1,5 ml) durante 10 min (30 mg) o 20 min (2DG). El tejido muscular se recortó, se pesó y se colocó en tubos de polipropileno en hielo seco. Se añade 1 ml de 1 N KOH a los tubos que luego se colocan en un baño de agua a 70°C durante 10-15 minutos, con agitación por vórtex de los tubos de cada dos minutos. Los tubos se enfriaron en hielo y se añade 1 ml de 1 N HCl, a continuación se mezcló bien. 200 µl de sobrenadante se pone entonces en viales de centelleo duplicados y se contó en un contador de centelleo en comparación con estándares radiactivos conocidos.

50 Para la contracción, los músculos se incuban primero durante 1 hora en medios de preincubación/incubación. Después de 1 hora, un músculo de cada par (un par por rata) se fija al aparato de estimulación y el otro músculo se transfiere a un nuevo matraz de medio de incubación. El músculo contraído es estimulado por trenes de 200 mseg de 70 Hz con cada impulso en un tren de 0,1 mseg. Los trenes se administran a 1/seg a 10-15V durante 2x10 minutos con un descanso de 1 minuto. Al final del período de estimulación, el músculo es retirado del aparato de estimulación y se coloca en medio de lavado durante 10 minutos, seguido por los medios de comunicación de la etiqueta como se describe anteriormente.

60 Independientemente de los medios que se utilizan para crear los presentes análogos de insulina, los análogos son objeto de ensayos para la actividad biológica. Ensayos de timidina tritiados pueden llevarse a cabo para determinar el grado de división celular. Otros ensayos biológicos, sin embargo, pueden ser utilizados para determinar la actividad deseada. Polipéptidos de insulina pueden ser analizados por su actividad antiviral y/o actividad antiproliferativa. Ensayos antiproliferativos son conocidos por los expertos ordinarios de la técnica. Basu et al. en Bioconjugate Chem (2006) 17: 618-630 describen un ensayo de antiproliferación utilizando células A549 y MTT para medir la proliferación. Los ensayos biológicos tales como el ensayo de la capacidad de inhibir la replicación viral, también proporciona una indicación de la actividad de la insulina. Ensayos conocidos por un experto ordinario de la técnica pueden ser también utilizados para evaluar la actividad biológica y posibles efectos secundarios de los

polipéptidos de insulina de la invención.

Cantidades medias de insulina, los polipéptidos de insulina, y/o análogos de insulina de la presente invención pueden variar y en particular deberían basarse en las recomendaciones y prescripción de un médico calificado. La cantidad exacta de la insulina, los polipéptidos de insulina, y/o análogos de insulina de la presente invención es una cuestión de preferencia sujeta a factores tales como el tipo exacto y/o la gravedad de la afección a tratar, la condición del paciente a tratar, así como los otros ingredientes de la composición. La invención también proporciona la administración de una cantidad terapéuticamente eficaz de otro agente activo. La cantidad a ser dada puede ser determinada fácilmente por un experto ordinario en la técnica basándose en la terapia con insulina, las terapias de insulina disponibles, y/u otros análogos de insulina.

Las composiciones farmacéuticas de la descripción pueden ser fabricadas de una manera convencional.

## 15 EJEMPLOS

Los siguientes ejemplos se ofrecen para ilustrar, pero no limitan la invención reivindicada.

### 20 Ejemplo 1

Este ejemplo describe uno de los muchos conjuntos posibles de criterios para la selección de los sitios de la incorporación de aminoácidos no naturalmente codificados en la insulina.

La Figura 1 muestra la estructura y secuencia de la insulina (NP\_000198 - número de acceso de Proteína Humana), y la tabla abajo incluye la SEQ ID NO: 13 con la secuencia de longitud completa de la insulina humana, SEQ ID NO: 14 con la secuencia de la insulina humana de longitud completa sin el líder, SEQ ID NO: 1-12 contiene secuencias de cadena A y de cadena B para la insulina, lispro, aspart, glulisina, detemir y glargina. Polipéptidos de insulina fueron generados por la sustitución de un aminoácido codificado de forma natural con un aminoácido no naturalmente codificado. Cada polipéptido tenía uno de los aminoácidos sustituidos con para-acetilfenilalanina o con para-aminofenilalanina. Los polipéptidos generados carecían de la secuencia líder y eran polipéptidos de insulina de cadena A/B (SEQ ID NO. 1-14). Cada uno de los polipéptidos generados tenían una sustitución de aminoácido no naturalmente codificado en una de las siguientes posiciones 1, 5, 8, 9, 10, 12, 14, 15, 18, 19, y 21 de la SEQ ID NO: 1 y 1, 2, 3, 4, 17, 20, 21, 22, 25, 28, y 29 de la SEQ ID NO: 2.

La Figura 2 muestra la estructura de la insulina humana que se marcó usando el software PyMOL (DeLano Scientific; Palo Alto, CA) y algunos aminoácidos que corresponden a los sustituidos con para-acetilfenilalanina en polipéptidos de insulina de la invención. La Figura 6 muestra la homología de secuencia entre la insulina y análogos de insulina conocidos.

Otro conjunto de criterios para la selección de los sitios preferidos de incorporación de aminoácidos no naturalmente codificados incluye el uso y la comparación de las estructuras cristalinas del Banco de Datos de Proteína, u otros bancos de datos, se utilizan para modelar la estructura de la insulina y los residuos se identifican que 1) no interferirían con la unión al receptor de FGF o heparina, y 2) no estarían presentes en el interior de la proteína. En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en, pero no se limitan a, una o más de las siguientes posiciones de la insulina: 1, 5, 8, 9, 10, 12, 14, 15, 18, 19, y 21 de la SEQ ID NO: 1 y 1, 2, 3, 4, 17, 20, 21, 22, 25, 28, y 29 de la SEQ ID NO: 2. En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en, pero no se limitan a, una o más de las siguientes posiciones de insulina: 1, 5, 8, 9, 10, 12, 14, 15, 18, 19, y 21 de la SEQ ID NO: 1 (o los aminoácidos correspondientes en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en, pero no se limitan a, una o más de las siguientes posiciones de la insulina: 1, 2, 3, 4, 17, 20, 21, 22, 25, 28, y 29 de la SEQ ID NO: 2 (o los aminoácidos correspondientes en SEQ ID NOs: 4, 6, 8, 10, 12). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en, pero no se limitan a, una o más de las siguientes posiciones de la insulina: 1, 5, 8, 9, 10, y 12 de la SEQ ID NO: 1 (o los aminoácidos correspondientes en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en, pero no se limitan a, una o más de las siguientes posiciones de la insulina: 1, 2, 3, 4, 17, y 20 de la SEQ ID NO: 2 (o los aminoácidos correspondientes en SEQ ID NOs: 4, 6, 8, 10, 12). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en, pero no se limitan a, una o más de las siguientes posiciones de insulina: 14, 15, 18, 19, y 21 de la SEQ ID NO: 1 (o los aminoácidos correspondientes en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en, pero no se limitan a, una o más de las siguientes posiciones de la insulina: 21, 22, 25, 28, y 29 de la SEQ ID NO: 2 (o los aminoácidos correspondientes en SEQ ID NOs: 4, 6, 8, 10, 12).

Se utilizaron los siguientes criterios para evaluar cada posición de insulina y análogos de insulina para la introducción de un aminoácido no naturalmente codificado: el residuo (a) no debe interferir con la unión del receptor de insulina basado en el análisis estructural, b) no debe verse afectado por mutagénesis de escaneo de alanina o

homólogos (c) debe ser expuesto en la superficie y exhibir mínimas interacciones de van der Waals o enlaces de hidrógeno con residuos circundantes, (d) debe eliminarse o ser variable en variantes de la insulina, (e) resultaría en cambios conservadores tras la sustitución con un aminoácido no naturalmente codificado y (f) se puede encontrar en cualquiera de las regiones altamente flexibles o regiones estructuralmente rígidas. Además, otros cálculos se pueden realizar en la molécula de insulina, utilizando el programa Cx (Pintar *et al.* (2002) *Bioinformatics*, 18, pp 980) para evaluar la extensión de la protuberancia para cada átomo de proteína.

En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en una o más de las siguientes posiciones en la insulina: antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína) (SEQ ID NO: 1 o los aminoácidos correspondientes en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorpora en una o más de las siguientes posiciones en la insulina: antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32 (es decir, en el extremo carboxilo de la proteína) (SEQ ID NO: 2 o los aminoácidos correspondientes en SEQ ID NOs: 4, 6, 8, 10, 12). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en una o más de las siguientes posiciones en la insulina: 8, 9, 10, 14 (SEQ ID NO: 1 o los aminoácidos correspondientes en SEQ ID NOs: 3, 5, 7, 9, 11). En algunas realizaciones, uno o más aminoácidos no naturalmente codificados se incorporan en una o más de las siguientes posiciones en la insulina: 1, 17, 25, 28 (SEQ ID NO: 2 o los aminoácidos correspondientes en SEQ ID NOs: 4, 6, 8, 10, 12).

### Ejemplo 2

Este ejemplo detalla la clonación y expresión de un polipéptido de insulina incluyendo un aminoácido no naturalmente codificado en *E. coli*.

Métodos para la clonación de la insulina son conocidos por personas de experiencia ordinaria en la técnica. Las secuencias de polipéptidos y polinucleótidos para la insulina y la clonación de la insulina en células huésped se detallan en la Patente de Estados Unidos nº 5.962.267; Patente de EE.UU. nº 4.751,180, 7.105.314, 6.630.348, 6.777 y 7.091.032 patente de Annibali titulado "Expression of a human insulin precursor in *P. pastoris*", así como Números de Publicación de Patente de Estados Unidos 20080268519, 20080255045, 20080227205, 20080207877, 20080207877, 20080213828, 20080261311, y 20080227195.

ADNc que codifican las formas lispro de la insulina se muestran como SEQ ID NOs: 34, 35, 36, y 37. Este polipéptido maduro se muestra como SEQ ID NOs: 3 y 4.

Un sistema de traslado introducido que comprende un ARNt ortogonal (O-ARNt) y una sintetasa de ARNt de aminoácido ortogonal (O-RS) se utiliza para expresar insulina o análogos de insulina que contienen un aminoácido no naturalmente codificado. La O-RS aminoacila de forma preferente el O-ARNt con un aminoácido no naturalmente codificado. A su vez el sistema de traslado inserta el aminoácido no naturalmente codificado en la insulina o análogo de insulina, en respuesta a un codón selector codificado. Secuencias de O-RS y O-ARNt adecuadas se describen en el documento WO 2006/068802 titulado "Compositions of Aminoacyl-tRNA Synthetase and Uses Thereof (E9; SEQ ID NO: 15) y WO 2007/021297, titulada "Compositions of tRNA and Uses Thereof (F13, SEQ ID NO: 16).

5	SEQ ID NO: 17	M. jannaschii; ARNmt <sup>Tyr</sup> <sub>CUA</sub>	ARNt
	SEQ ID NO: 18	HLAD03; un supresor ámbar optimizado ARNt	ARNt
	SEQ ID NO: 19	HL325A: un supresor de ARNt de cambio de marco AGGA optimizado	ARNt
10	SEQ ID NO: 20	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -azido-L-fenilalanina p-Az-PheRS (6)	RS
	SEQ ID NO: 21	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -benzoilo-L-fenilalanina p-BpaRS (1)	RS
15	SEQ ID NO: 22	sintetasa aminoacilo ARNt para la incorporación de propargilo-fenilalanina Propargyl/-P Hers	RS
	SEQ ID NO: 23	sintetasa aminoacilo ARNt para la incorporación de propargilo fenilalanina Propargyl/-P Hers	RS
20	SEQ ID NO: 24	sintetasa aminoacilo ARNt para la incorporación de propargilo fenilalanina propargilo-PheRS	RS
	SEQ ID NO: 25	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -azido fenilalanina p-AZ-PheRS (1)	RS
25	SEQ ID NO: 26	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -azido fenilalanina p-AZ-PheRS (3)	RS
	SEQ ID NO: 27	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -azido fenilalanina p-AZ-PheRS (4)	RS
30	SEQ ID NO: 28	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -azido fenilalanina p-AZ-PheRS (2)	RS
	SEQ ID NO: 29	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -acetilo fenilalanina p-AZ-PheRS (LW1)	RS
35	SEQ ID NO: 30	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -acetilo fenilalanina (LW5)	R <sup>8</sup>
	SEQ ID NO: 31	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -acetilo fenilalanina (LW6)	R <sup>8</sup>
40	SEQ ID NO: 32	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -azido fenilalanina (AzPheRS-5)	R <sup>8</sup>
	SEQ ID NO: 33	sintetasa aminoacilo ARNt para la incorporación de <i>p</i> -azido fenilalanina (AzPheRS-6)	R <sup>8</sup>
45			

50 La transformación de *E. coli* con plásmidos que contienen insulina modificada o gen análogo de insulina y sintetasa ARNt/par ARNt de aminoacilo ortogonal (específico para el aminoácido no naturalmente codificado deseado) permite la incorporación de sitio específico de aminoácido no naturalmente codificado en el polipéptido de insulina.

55 La insulina madura de tipo silvestre se amplifica por PCR a partir de una reacción de síntesis de ADNc utilizando protocolos estándar y se clonaron en pET30 (NcoI-BamHI). Antes de o tras la confirmación de secuencia, la insulina que incluye una secuencia N-terminal HHHHHHSGG se subclona en un vector de supresión que contiene un tirosilo supresor de ámbar ARNt Tyr/CUA de *Methanococcus jannaschii* (Mj ARNt Tyr/CUA) bajo control constitutivo de un promotor sintético derivado de la secuencia del promotor de lipoproteína de *E. coli* (Miller, J.H., Gene, 1986), así como el tirosil-ARNt-sintetasa ortogonal (MjTyrRS) bajo el control del promotor GlnRS *E. coli*. La expresión de la insulina está bajo el control del promotor T7. Mutaciones ámbar se introducen utilizando protocolos de mutación de cambio rápido estándar (Stratagene, La Jolla, California). Las construcciones se verifican por secuencia.

60 Ensayos de compuestos de insulina de efecto largo se pueden hacer usando el modelo de rata diabética STZ (PCO 08-400-209).

65 Supresión con para-acetilo-feniloalanina (pAcF)

Plásmidos (insulina pVK6-LisPro) se transformaron en la cepa W3110 B2 de *E. coli* en la que la expresión de la polimerasa T7 estaba bajo el control de un promotor inducible por arabinosa. Se diluyeron cultivos de bacterias durante la noche 1: 100 en matraces de agitación que contenían medios de cultivo 2X YT y se cultivaron a 37°C a una OD<sub>600</sub> de ~0,8. La expresión de proteínas fue inducida por la adición de arabinosa (0,2% final), y *para*-acetilofeniloalanina (pAcF) a una concentración final de 4 mM. Los cultivos se incubaron a 37°C durante 5 horas. Las células se sedimentaron y se resuspendieron en tampón de lisis B-PER (Pierce) 100ul/OD/ml + 10ug/ml de DNasa y se incubaron a 37°C durante 30 min. El material celular se eliminó mediante centrifugación y el sobrenadante se eliminó. El sedimento se resuspendió en una cantidad igual de tampón de carga de proteína de SDS-PAGE. Todas las muestras se cargaron en un gel PAGE a 4-12% con MES y DTT. Los métodos para la purificación de la insulina son conocidos por personas de experiencia ordinaria en la técnica y están confirmados por SDS-PAGE, análisis de Transferencia Western, o espectrometría de masas de trampa de iones de electropulverización-ionización y similares.

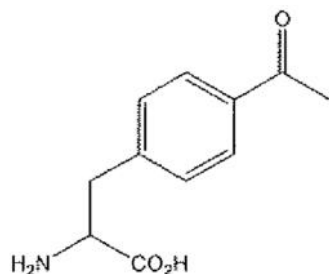
El plásmido y la secuencia utilizados se muestran en la Figura 5, y la expresión de insulina etiquetada His N-terminal y la supresión a los 7 sitios ámbar se muestra como la Figura 6. El polipéptido de proinsulina está marcado con una flecha. La Figura 6 muestra las muestras de gránulo B-PER - a la izquierda de Carril 1: Marcador; Carril 1: VK6-proinsulina-pAF-Q15; Carril 2: VK6-proinsulina-pAF-Y14; Carril 3: VK6-proinsulina-pAF-R22; Carril 4: VK6-proinsulina-pAF-F1; Carril 5: VK6-proinsulina-pAF-G1; Carril 6: VK6-proinsulina-pAF-S9, 0,2% de arabinosa; Carril 7: VK6-proinsulina-pAF-K28. Los números de posición indicados para las sustituciones de aminoácidos en los carriles 1, 2, 5, y 6 se basan en insulina LisPro cadena A, SEQ ID NO: 3, y los números de posición indicados para las sustituciones de aminoácidos en los carriles 3, 4, y 7 se basan en insulina LisPro cadena B, SEQ ID NO: 4.

Proteínas de insulina mutantes etiquetados por His pueden purificarse usando métodos conocidos por personas de experiencia ordinaria en la técnica. La Resina de quelación de níquel ProBond (Invitrogen, Carlsbad, CA) puede utilizarse a través de los procedimientos de purificación de proteínas etiquetados por His estándar proporcionados por el fabricante. Mediciones funcionales de las proteínas se pueden realizar a través de métodos conocidos en la técnica y métodos proporcionados en esta solicitud, y alternativamente, ELISA en células vivas puede ser desarrollado para evaluar polipeptidos de insulina de la invención.

### Ejemplo 3

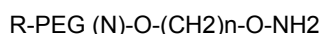
Este ejemplo detalla la introducción de un aminoácido que contiene carbonilo y posterior reacción con un PEG que contiene aminooxi.

Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora un aminoácido no naturalmente codificado que contiene cetona que se hace reaccionar posteriormente con un PEG que contiene aminooxi de aproximadamente 5.000 MW. Cada uno de los residuos antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 1 o los correspondientes aminoácidos de SEQ ID NOs: 3, 5, 7, 9, 11) y cada uno de los residuos antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 2 o los aminoácidos correspondientes en SEQ ID NOs: 4, 6, 8, 10, 12) está sustituido por separado con un aminoácido no naturalmente codificado que tiene la siguiente estructura:



Las secuencias utilizadas para la incorporación de sitio específico de *p*-acetilofenilalanina en insulina son SEQ ID NO: 1 y 2 (cadenas A y B de la insulina), y SEQ ID NO: 16 o 17 (mutARNt, *M. jannaschii*), y 15, 29, 30 o 31 TyrRS LW1, 5, o 6) descrito en el Ejemplo 2 anterior.

Una vez modificado, la variante de polipéptido de insulina que comprende el aminoácido que contiene carbonilo se hace reaccionar con un derivado de PEG que contiene aminooxi de la forma:

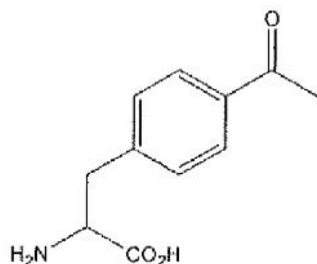


donde R es metilo, n es 3 y N es de aproximadamente 5.000 MW. La insulina purificada que contiene p-acetilo-  
feniloalanina disuelto a 10 mg/ml en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de Hepes  
(Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato de sodio (Sigma Chemical, St. Louis, MO) pH 4,5,  
se hace reaccionar con un exceso de 10 a 100 veces de PEG que contiene aminooxi y, a continuación, se agitó  
durante 10-16 horas a temperatura ambiente (Jeneks, W.J. Am. Chem. Soc. 1959, 81, pp 475). La insulina PEG se  
diluye a continuación en tampón apropiado para la purificación y análisis inmediato.

#### Ejemplo 4

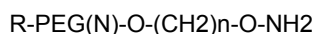
Este ejemplo detalla la introducción de un aminoácido que contiene carbonilo y posterior reacción con un PEG que  
contiene aminooxi.

Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora un aminoácido no  
naturalmente codificado que contiene cetona que se hace reaccionar posteriormente con un PEG que contiene  
aminooxi de aproximadamente 20.000 MW. Cada uno de los residuos antes de la posición 1 (es decir, en el N-  
término), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, (es decir, en el extremo carboxilo de la  
proteína de la SEQ ID NO: 1 o los aminoácidos correspondientes en SEQ ID NOs: 3, 5, 7, 9, 11) y cada uno de los  
residuos antes de la posición 1 (es decir, al término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,  
20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 2 o  
los aminoácidos correspondientes en la SEQ ID NO: 4, 6, 8, 10, 12) está sustituido por separado con un aminoácido  
no naturalmente codificado que tiene la siguiente estructura:



Las secuencias utilizadas para la incorporación de sitio específico de p-aminofenilalanina en insulina son SEQ 1D  
NO: 1 y 2 (cadenas A y B de la insulina), y SEQ ID NO: 16 o 17 (mutARnt, M. jannaschii), y secuencias descritas  
anteriormente e incorporadas para la incorporación de sitio específico de p-amino-feniloalanina.

Una vez modificada, la variante de polipéptido de insulina que comprende el aminoácido que contiene carbonilo se  
hace reaccionar con un derivado de PEG que contiene amino de la forma:



donde R es metilo, n es 3 y N es de aproximadamente 20.000 MW. La insulina purificada que contiene p-  
aminofenilalanina disuelto a 10 mg/ml, en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de Hepes  
(Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato sódico (Sigma Chemical, St. Louis, MO) pH 4,5, se  
hace reaccionar con un exceso de 10 a 100 veces de PEG que contiene aminooxi, y después se agitó durante 10 -  
16 horas a temperatura ambiente (Jencks, W. J. Am. Chem. Soc. 1959, 81, pp 475). La insulina PEG se diluye a  
continuación en tampón apropiado para la purificación y análisis inmediato,

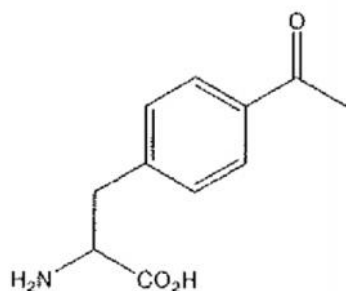
#### Ejemplo 5

Este ejemplo detalla la introducción de un aminoácido que contiene carbonilo y posterior reacción con PEG que  
contiene aminooxi.

Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora un aminoácido no  
naturalmente codificado que contiene cetona que se hace reaccionar posteriormente con un PEG que contiene  
aminooxi de aproximadamente 20.000 MW. Cada uno de los residuos antes de la posición 1 (es decir, en el término  
N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33,  
34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64,  
65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95,  
96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la  
proteína de SEQ ID NO: 13) está sustituido por separado con aminoácido no naturalmente codificado que tiene la  
siguiente estructura:

5

10

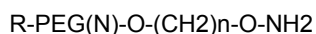


15

Las secuencias utilizadas para la incorporación de sitio específico de *p*-aminofenilalanina en insulina son SEQ ID NO: 13 y SEQ ID NO: 16 o 17 (mutARNt, *M. jannaschii*), y secuencias descritas anteriormente e incorporadas para la incorporación específica de sitio de *p* amino-fenilalanina.

20

Una vez modificada, la variante de polipéptido de insulina que comprende el aminoácido que contiene carbonilo se hace reaccionar con un derivado de PEG que contiene aminooxi de la forma:



25

donde R es metilo, n es 3 y N es de aproximadamente 20.000 MW. La insulina purificada que contiene *p*-amino-feniloalanina disuelto a 10 mg/mL en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de HEPES (Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato de sodio (Sigma Chemical, St. Louis, MO) pH 4,5, se hace reaccionar con un 10 a 100 veces el exceso de PEG que contiene amino, y después se agitó durante 10-16 horas a temperatura ambiente (Jenekcs, W. J. Am. Chem. Soc. 1959, 81, pp 475). La insulina PEG se diluye a continuación en tampón apropiado para la purificación y análisis inmediato.

30

#### Ejemplo 6

35

Este ejemplo detalla la introducción de un carbonilo que contiene aminoácido y posterior reacción con un PEG que contiene amino.

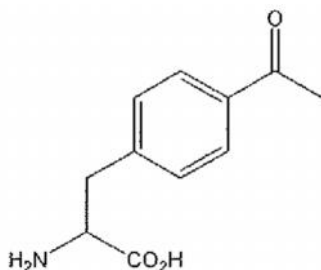
40

Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora un aminoácido no naturalmente codificado que contiene cetona que se hace reaccionar posteriormente con un PEG que contiene aminooxi de aproximadamente 20.000 MW. Cada uno de los residuos antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 14) está sustituido por separado con un aminoácido no naturalmente codificado que tiene la siguiente estructura:

45

50

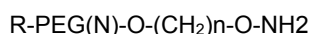
55



60

Las secuencias utilizadas para la incorporación de sitio específico de *p*-amino-feniloalanina en insulina son SEQ ID NO: 14 y SEQ ID NO: 16 o 17 (mutARNt, *M. jannaschii*), y secuencias descritas anteriormente e incorporadas para la incorporación de sitio específico de *p*-amino-feniloalanina.

65

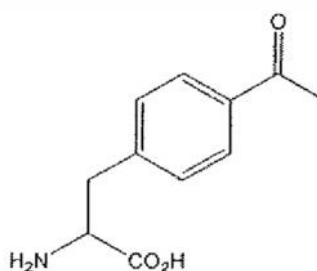


donde R es metilo, n es 3 y N es de aproximadamente 20.000 MW. La insulina purificada que contiene *p*-aminofenilalanina disuelto a 10 mg/ml, en 25 min de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de Hepes (Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato de sodio (Sigma Chemical, St. Louis, MO) pH 4,5, se hace reaccionar con 10 a 100 veces el exceso de PEG que contienen aminooxi, y después se agitó durante 10 - 16 horas a temperatura ambiente (Jencks, W. J. Am. Chem. Soc. 1959, 81, pp 475). La insulina PEG se diluye a continuación en tampón apropiado para la purificación y análisis inmediato.

#### Ejemplo 7

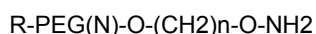
Este ejemplo detalla introducción de un aminoácido que contiene carbonilo y posterior reacción con un PEG que contiene aminooxi.

Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora un aminoácido no naturalmente codificado que contiene cetona que se hace reaccionar posteriormente con un PEG que contiene aminooxi de aproximadamente 30.000 MW. Cada uno de los residuos antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o las posiciones correspondientes en la SEQ ID NOs: 112 y 14) está sustituido por separado con un aminoácido no naturalmente codificado que tiene la siguiente estructura:



Las secuencias utilizadas para la incorporación de sitio específico de *p*-amino-feniloalanina en insulina son SEQ ID NO: 13 (o SEQ ID NO: 1,2, o 14), y SEQ ID NO: 16 o 17 (mutARNt, M. jannaschii), y las secuencias descritas anteriormente e incorporadas para la incorporación de sitio específico de *p*-amino-feniloalanina.

Una vez modificada, la variante de polipéptido de insulina que comprende el aminoácido que contiene carbonilo se hace reaccionar con un amino oxieontaining derivado de PEG de la forma:



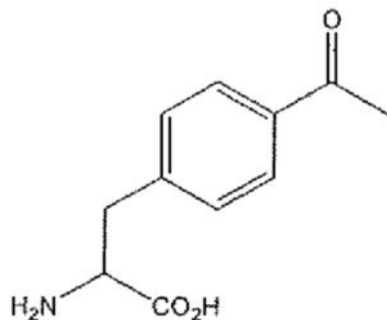
donde R es metilo, n es 3 y N es de aproximadamente 30.000 MW. La insulina purificada que contiene *p*-amino-feniloalanina disuelto a 10 mg/mL en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de Hepes (Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 min de acetato de sodio (Sigma Chemical, St. Louis, MO), pH 4,5, se hace reaccionar con un exceso de 10 a 100 veces de PEG que contiene aminooxi, y después se agitó durante 10 - 16 horas a temperatura ambiente (Jencks, W. J. Am. Chem. Soc. 1959, 81, pp 475). La insulina PEG se diluye a continuación en tampón apropiado para la purificación y análisis inmediato.

#### Ejemplo 8

Este ejemplo detalla la introducción de un aminoácido que contiene carbonilo y posterior reacción con un PEG que contiene aminooxi.

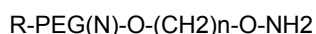
Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora un aminoácido no naturalmente codificado que contiene cetona que se hace reaccionar posteriormente con un PEG que contiene aminooxi de aproximadamente 40.000 MW. Cada uno de los residuos antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o las posiciones correspondientes en SEQ ID NOs: 112 y 14) está sustituido por separado con un aminoácido no naturalmente codificado que tiene la siguiente estructura:





20 Las secuencias utilizadas para la incorporación de sitio específico de *p*-amino-feniloalanina en insulina son SEQ ID NO: 13 (o SEQ ID NO: 1, 2, o 14), y SEQ ID NO: 16 o 17 (mutARnt, M. jannaschii), y secuencias descritas anteriormente e incorporadas para la incorporación de sitio específico de *p*-aminofeniloalanina.

Una vez modificada, la variante de polipéptido de insulina que comprende el carbonilo que contiene aminoácido se hace reaccionar con un derivado de PEG que contiene amino de la forma:

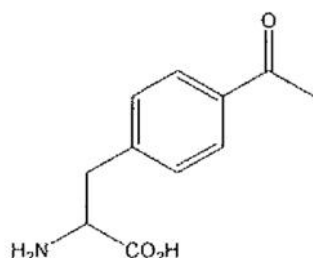


30 donde R es metilo, n es 3 y N es de aproximadamente 40.000 MW. La insulina purificada que contiene *p*-aminofeniloalanina disuelta a 10 mg/ml en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de Hepes (Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato sódico (Sigma Chemical, St. Louis, MO) pH 4,5, se hace reaccionar con un exceso de 10 a 100 veces de PEG que contiene aminoóxi, y después se agitó durante 10-16 horas a temperatura ambiente (Jencks, W. J, Am. Chem. Soc. 1959, 81, pp 475). La insulina PEG se diluye a continuación en tampón apropiado para la purificación y análisis inmediato.

#### Ejemplo 9

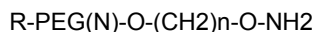
35 Este ejemplo detalla la introducción de un aminoácido que contiene carbonilo y posterior reacción con un PEG que contiene aminoóxi.

40 Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora un aminoácido no naturalmente codificado que contiene cetona que se hace reaccionar posteriormente con un PEG que contiene aminoóxi de aproximadamente 10.000 MW. Cada uno de los residuos antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o las posiciones correspondientes en la SEQ ID NOs: 112 y 14) está sustituido por separado con un aminoácido no naturalmente codificado que tiene la siguiente estructura:



60 Las secuencias utilizadas para la incorporación de sitio específico de *p*-aminofeniloalanina en insulina son SEQ ID NO: 13 (o posiciones correspondientes en SEQ ID NO: 1,2, o 14), y SEQ ID NO: 16 o 17 (mutARnt, M. jannaschii), y secuencias descritas anteriormente e incorporadas para la incorporación de sitio específico de *p*-aminofeniloalanina.

65 Una vez modificada, la variante de polipéptido de insulina que comprende el aminoácido que contiene carbonilo se hace reaccionar con un derivado de PEG que contiene aminoóxi de la forma:-



donde R es metilo, n es 3 y N es de aproximadamente 10.000 MW. La insulina purificada que contiene *p*-aminofeniloalanina disuelta a 10 mg/ml en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM HEPES (Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato de sodio (Sigma Chemical, St. Louis, MO) pH 4,5, se hace reaccionar con un exceso de 10 a 100 veces de PEG que contiene aminooxi, y después se agitó durante 10-16 horas a temperatura ambiente (Jencks. W. J. Am. Chem. Soc. 1959, XL, pp 475). La insulina de PEG se diluye a continuación en tampón apropiado para la purificación y análisis inmediato.

#### 10 Ejemplo 10

Conjugación con un PEG que consiste en un grupo de hidroxilamina enlazado al PEG mediante un enlace de amida.

Un reactivo de PEG que tiene la siguiente estructura está acoplada a un aminoácido no naturalmente codificado que contiene cetona usando el procedimiento descrito en los Ejemplos 39:



donde R=metilo, n=4 y N es de aproximadamente 5,000 MW - 40.000 MW. La reacción, purificación, y análisis de las condiciones son como se ha descrito y son conocidas en la técnica.

#### 25 Ejemplo 11

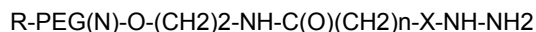
Este ejemplo detalla la introducción de dos aminoácidos no naturalmente codificados distintos en los polipéptidos de insulina y polipéptidos análogos de insulina.

Este ejemplo demuestra un método para la generación de un polipéptido de insulina que incorpora aminoácidos no naturalmente codificados que comprenden una funcionalidad de cetona en dos posiciones entre los siguientes restos: antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84., 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o las posiciones correspondientes en SEQ ID NOs: 1-12 y 14). El polipéptido de insulina se prepara como se ha descrito anteriormente, excepto que el codón selector es introducido en dos sitios distintos dentro del ácido nucleico.

#### 40 Ejemplo 12

Este ejemplo detalla la conjugación de polipéptido de insulina o análogo de polipéptido de insulina a un PEG que contiene hidrazida y posterior reducción in situ.

un polipéptido de insulina que incorpora un aminoácido que contiene carbonilo se prepara de acuerdo con el procedimiento descrito anteriormente. Una vez modificado, un PEG que contiene hidrazida que tiene la siguiente estructura se conjuga con el polipéptido de insulina:



donde R=metilo, n=2 y N=5000; 10.000, 20.000; 30,000; o 40.000 MW y X es un grupo de carbonilo (C=O). La insulina purificada que contiene *p*-acetilofeniloalanina se disuelve a entre 0,110 mg/ml en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de Hepes (Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato de sodio (Sigma Chemical, St. Louis, MO) pH 4,5, se hace reaccionar con un exceso de 1 a 100 veces de PEG que contiene hidrazida, y la hidrazona correspondiente se reduce in situ mediante la adición de reserva 1 M NaCNBH3 (Sigma Chemical, St. Louis, MO), disuelta en H2O, a una concentración final de 10-50 mM. Las reacciones se llevan a cabo en la oscuridad a 4°C a TA durante 18-24 horas. Las reacciones se detuvieron mediante la adición de 1 M Tris (Sigma Chemical, St. Louis, MO) a aproximadamente pH 7,6 a una concentración final de Tris de 50 min o se diluyeron en tampón apropiado para la purificación inmediata.

#### 60 Ejemplo 13

Este ejemplo detalla la conjugación de polipéptido de insulina o análogo de polipéptido de insulina a un PEG que contiene hidrazida y posterior reducción in situ.

Un polipéptido de insulina que incorpora un aminoácido que contiene carbonilo se prepara de acuerdo con el procedimiento descrito anteriormente. Una vez modificado, un PEG que contiene hidrazida que tiene la siguiente estructura se conjuga con el polipéptido de insulina:

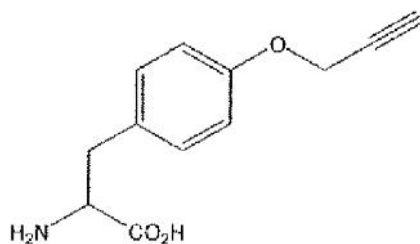
R-PEG(N)-O-(CH<sub>2</sub>)<sub>2</sub>-NHC(O)(CH<sub>2</sub>)<sub>n</sub>-X-NH-NH<sub>2</sub>

donde R=metilo, n=2 y N = 20.000 MW y X es un grupo carbonilo (C=O). La *p*-acetilfeniloalanina purificada que contiene insulina se disuelve a entre 0,110 mg/ml en 25 mM de MES (Sigma Chemical, St. Louis, MO) pH 6,0, 25 mM de Hepes (Sigma Chemical, St. Louis, MO) pH 7,0, o en 10 mM de acetato de sodio (Sigma Chemical, St. Louis, MO) pH 4,5, se hace reaccionar con un exceso de 1 a 100 veces de PEG que contiene hidrazida, y la hidrazona correspondiente se reduce in situ mediante la adición de reserva 1 M NaCNBH<sub>3</sub> (Sigma Chemical, St. Louis, MO), disuelta en H<sub>2</sub>O, a una concentración final de 10-50 mM. Las reacciones se llevan a cabo en la oscuridad a 4°C a temperatura ambiente durante 18-24 horas. Las reacciones se detuvieron mediante la adición de 1 M Tris (Sigma Chemical, St. Louis, MO) a aproximadamente pH 7,6 a una concentración final de Tris de 50 mM o se diluyeron en tampón apropiado para la purificación inmediata.

#### Ejemplo 14

Este ejemplo detalla la introducción de un aminoácido que contiene alquino en un polipéptido de insulina o polipéptido de análogo de insulina y derivatización con mPEG-azida.

Los siguientes residuos, antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o las correspondientes posiciones en SEQ ID NOs: 112 y 14), están sustituidos cada uno con el siguiente aminoácido no naturalmente codificado:



Las secuencias utilizadas para la incorporación de sitio específico de *p*-propargilo-tirosina en insulina son SEQ ID NO: 13 (o posiciones correspondientes en SEQ ID NO: 1, 2, o 14), SEQ ID NO: 16 o 17 (mutARNt, M. jannaschii), y 22, 23 o 24 como se ha descrito anteriormente. El polipéptido de insulina que contiene la tirosina de propargilo se expresa en *E. coli* y se purificó usando las condiciones descritas anteriormente.

La insulina que contiene propargilo-tirosina purificada disuelta a entre 0,1-10 mg/ml en tampón PB (100 mM de fosfato de sodio, 0,15 M de NaCl, pH=8) y un exceso de 10 a 1000 veces de un PEG que contiene azida se añade a la mezcla de reacción. Una cantidad catalítica de CuSO<sub>4</sub> y el alambre Cu se añadió luego a la mezcla de reacción. Después se incubó la mezcla (incluyendo pero no limitado a, aproximadamente 4 horas a temperatura ambiente o 37°C, o durante la noche a 4°C), se añade H<sub>2</sub>O y la mezcla se filtra a través de una membrana de diálisis. La muestra puede ser analizada para la adición, incluyendo, pero no limitado a, por procedimientos similares a los descritos en el Ejemplo 3. En este ejemplo, el PEG tendrá la siguiente estructura:

R-PEG(N)-O-(CH<sub>2</sub>)<sub>2</sub>-NH-C(O)(CH<sub>2</sub>)<sub>n</sub>-N<sub>3</sub>

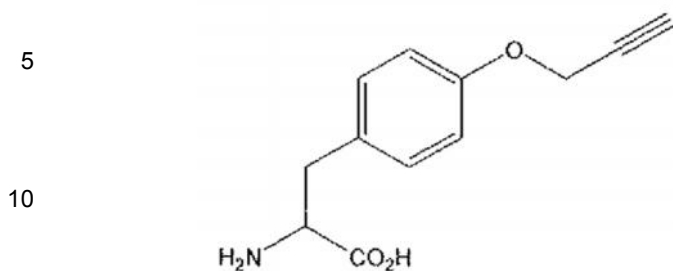
donde R es metilo, n es 4 y N = 5.000; 10.000, 20.000; 30.000; o 40.000 MW.

#### Ejemplo 15

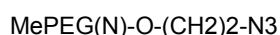
Este ejemplo detalla la sustitución de un aminoácido grande, hidrófobo en un polipéptido de insulina con tirosina de propargilo,

Un residuo Phe, Trp o Tyr presente dentro de una de las siguientes regiones de insulina: antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o las posiciones correspondientes en la SEQ ID NOs: 1-12 y 14),

está sustituido con el siguiente aminoácido no naturalmente codificado como se describió anteriormente:



15 Una vez modificado, un PEG está unido a la variante de polipéptido de insulina que comprende el aminoácido que contiene alquino. El PEG tendrá la siguiente estructura:



20 y procedimientos de acoplamiento seguirían los de los Ejemplos anteriores. Esto generará una variante de polipéptido de insulina que comprende un aminoácido no naturalmente codificado que es aproximadamente isostérico con uno de los grandes aminoácidos hidrófobos naturales y que está modificado con un derivado de PEG en un sitio distinto dentro del polipéptido.

#### 25 Ejemplo 16

Este ejemplo detalla la generación de un homodímero de polipéptido de insulina, heterodímero, homomultímero o heteromultímero separados por uno o más enlazadores de PEG. Multímeros de polipéptido de insulina se pueden formar entre proinsulinas o entre polipéptidos de insulina maduros de cadena A y B de la invención.

30 La variante de polipéptido de insulina que contiene alquino producida en el ejemplo anterior se hace reaccionar con un derivado de PEG bifuncional de la Forma:



35 donde n es 4 y el PEG tiene un peso molecular medio de aproximadamente 5.000; 10.000; 20.000; 30.000; o 40.000 MW para generar el correspondiente homodímero de polipéptido de insulina donde las dos moléculas de insulina están separadas físicamente por PEG. De una manera análoga un polipéptido de insulina puede estar acoplado a uno o más de otros polipéptidos para formar heterodímeros, homomultímeros o heteromultímeros. El acoplamiento, purificación y análisis se realizaron como en los ejemplos anteriores.

#### 40 Ejemplo 7

45 Este ejemplo detalla la generación de un homodímero de polipéptido de insulina, heterodímero, homomultímero o heteromultímero separados por uno o más enlazadores de PEG. Multímeros de polipéptido de insulina se pueden formar entre cadenas A y otras cadenas A o cadenas B y otras cadenas B.

50 La variante de polipéptido de insulina que contiene alquino producido en el ejemplo anterior se hace reaccionar con un derivado de PEG bifuncional de la forma:



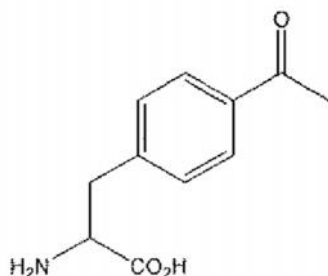
55 donde n es 4 y el PEG tiene un peso molecular medio de aproximadamente 5.000; 10.000; 20.000; 30.000; o 40.000 MW para generar el correspondiente homodímero de polipéptido de insulina donde las dos moléculas de insulina están separadas físicamente por PEG. De una manera análoga un polipéptido de insulina puede estar acoplado a uno o más de otros polipéptidos para formar heterodímeros, homomultímeros, o heteromultímeros. El acoplamiento, purificación y análisis se realizaron como en los ejemplos anteriores.

#### 60 Ejemplo 18

Este ejemplo detalla el acoplamiento de un resto sacárido a un polipéptido de insulina.

Un resto del siguiente está sustituido con el aminoácido no naturalmente codificado a continuación: antes de la posición 1 (es decir, en el término N), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86,

87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111 (es decir, en el extremo carboxilo de la proteína de SEQ ID NO: 13 o las posiciones correspondientes en SEQ ID NOs: 1-12 y 14) como se ha descrito anteriormente.

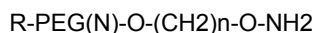


Una vez modificada, la variante de polipéptido de insulina que comprende el aminoácido que contiene carbonilo se hace reaccionar con un análogo de aminooxi  $\beta$ -enlazado de n-acetilo-glucosamina (GlcNAc). La variante de polipéptido de insulina (10 mg/ml) y el sacárido de aminooxi (21 mM) se mezclan en 100 mM de tampón de acetato de sodio acuoso (pH 5,5) y se incubaron a 37°C durante 7 a 26 horas. Un segundo sacárido está acoplado al primero enzimáticamente incubando el polipéptido de insulina conjugado por sacárido (5 mg/ml) con UDP-galactosa (16 mM) y  $\beta$ -1,4-galactosiltransferasa (0,4 unidades/ml) en tampón de 150 mM de HEPES (pH 7,4) durante 48 horas a temperatura ambiente (Schanbachcr *et al.* J. Biol. Chem. 1970, 245, 5057-5061).

#### 25 Ejemplo 19

Este ejemplo detalla la generación de un antagonista del polipéptido de insulina pegilado.

Un residuo, incluyendo pero no limitado a, aquellos implicados en la unión de receptor de insulina está sustituido con el siguiente aminoácido no naturalmente codificado como se describió anteriormente. Una vez modificada, la variante de polipéptido de insulina que comprende el aminoácido que contiene carbonilo se hace reaccionar con un derivado de PEG que contienen amino de la forma:



donde R es metilo, n es 4 y N es 5.000; 10.000; 20.000; 30.000; o 40.000 MW para generar un antagonista de polipéptido de insulina que comprende un aminoácido no naturalmente codificado que está modificado con un derivado de PEG en un solo sitio dentro del polipéptido. Acoplamiento, purificación y análisis se realizaron como se describe anteriormente.

#### 40 Ejemplo 20

La generación de un homodímero, heterodímero, homomultímero, o heteromultímero de polipéptido de insulina en los que las moléculas de insulina están enlazados directamente.

Una variante de polipéptido de insulina que comprende el aminoácido que contiene alquino se puede acoplar directamente a otra variante de polipéptido de insulina que comprende el aminoácido que contiene azido. De una manera análoga un polipéptido de insulina puede acoplarse a uno o más de otros polipéptidos para formar heterodímeros, homomultímeros, o heteromultímeros. Más descripción con respecto a los multímeros que se pueden formar se proporciona anteriormente en los Ejemplos 16 y 17 y el acoplamiento, la purificación y análisis se realizaron como se describe anteriormente.

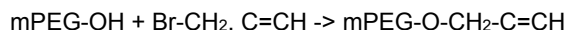
#### 50 Ejemplo 21



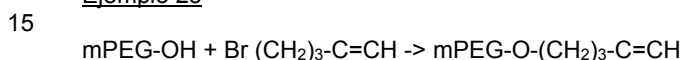
A

B

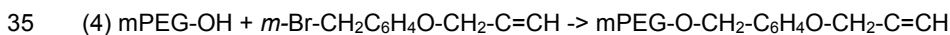
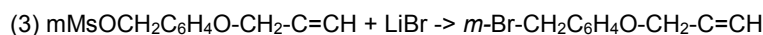
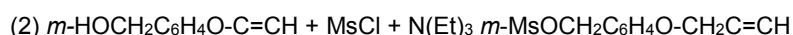
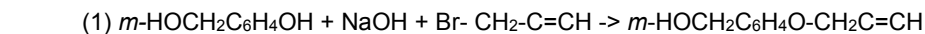
El polialquilenglicol (POH) se hace reaccionar con el haluro de alquilo (A) para formar el éter (B). En estos compuestos, n es un número entero de uno a nueve y R' puede ser un grupo alquilo o heteroalquilo de cadena lineal o ramificada, saturada o insaturada C1, a C20. R' también puede ser un alquilo o heteroalquilo cíclico saturado o insaturado C3 a C7 o un grupo arilo o heteroarilo cíclico, sustituido o no sustituido, o un alcarilo sustituido o no sustituido (el alquilo es un alquilo saturado o insaturado C1 a C20) o grupo heteroalcarilo. Típicamente, PEG-OH es polietilenglicol (PEG) o polietilenglicol de monometoxi (mPEG) que tiene un peso molecular de 800 a 40.000 Daltons (Da).

Ejemplo 22

5 mPEG-OH con un peso molecular de 20.000 Da (mPEG-OH 20 kDa; 2,0 g, 0,1 mmol, Sunbio) se trató con el NaH (12 mg, 0,5 mmol) en THF (3,5 ml). Una solución de bromuro de propargilo, se disuelve como una solución de 80% en peso en xileno (0,56 ml, 5 mmol, 50 equiv., Aldrich), y una cantidad catalítica de KI a continuación, se añadieron a la solución y la mezcla resultante se calentó a reflujo durante 2 horas. Después se añadió agua (1 ml) y el disolvente se eliminó a vacío. Al residuo se añadió CH<sub>2</sub>Cl<sub>2</sub> (25 ml) y la capa orgánica se separó, se secó sobre Na<sub>2</sub>SO<sub>4</sub> anhidro, y el volumen se redujo a aproximadamente 2 ml. Esta solución de CH<sub>2</sub>Cl<sub>2</sub> se añadió a éter dietílico (150 ml) gota a gota. Se recogió el precipitado resultante, se lavó con varias porciones de éter dietílico frío, y se secó para proporcionar propargilo-O-PEG.

Ejemplo 23

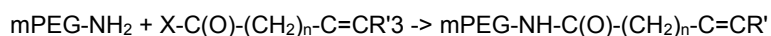
El mPEG-OH con un peso molecular de 20.000 Da (mPEG-OH 20 kDa; 2,0 g, 0,1 mmol, Sunbio) se trató con NaH (12 mg, 0,5 mmol) en THF (35 ml). Cincuenta equivalentes de 5-bromo-1-pentino (0,53 ml, 5 mmol, Aldrich) y una cantidad catalítica de KI se añadieron a la mezcla a continuación. La mezcla resultante se calentó a reflujo durante 16 horas. Después se añadió agua (1 ml) y el disolvente se eliminó a vacío. Al residuo se añadió CH<sub>2</sub>Cl<sub>2</sub> (25 ml) y la capa orgánica se separó, se secó sobre Na<sub>2</sub>SO<sub>4</sub> anhidro, y el volumen se redujo a aproximadamente 2 ml. Esta solución de CH<sub>2</sub>Cl<sub>2</sub> se añadió a éter dietílico (150 ml) gota a gota. Se recogió el precipitado resultante, se lavó con varias porciones de éter dietílico frío, y se secó para proporcionar el alquino correspondiente. 5-cloro-1-pentino puede usarse en una reacción similar.

Ejemplo 24

A una solución de 3-hidroxibencilalcohol (2,4 g, 20 mmol) en THF (50 ml) y agua (2,5 mL) primero se añadió hidróxido de sodio en polvo (1,5 g, 37,5 mmol) y luego una solución de bromuro de propargilo, se disuelve como una solución 80% en peso en xileno (3,36 ml, 30 mmol). La mezcla de reacción se calentó a reflujo durante 6 horas. Se añadió a la mezcla de 10% de ácido cítrico (2,5 ml) y el disolvente se eliminó bajo vacío. El residuo se extrajo con acetato de etilo (3 x 15 mL) y las capas orgánicas combinadas se lavaron con solución de NaCl saturado (10 ml), se secaron sobre MgSO<sub>4</sub> y se concentraron para dar el alcohol 3-propargilobencil.

45 Cloruro de metanosulfonilo (2,5 g, 15,7 mmol) y trietilamina (2,8 mL, 20 mmol) se añadieron a una solución del compuesto 3 (2,0 g, 11,0 mmol) en CH<sub>2</sub>Cl<sub>2</sub> a 0°C y la reacción se colocó en el refrigerador durante 16 horas. Una elaboración usual produjo el mesilato como un aceite amarillo pálido. Este aceite (2,4 g, 9,2 mmol) se disolvió en THF (20 ml) y LiBr (2,0 g, 23,0 mmol) se añadió. La mezcla de reacción se calentó a reflujo durante 1 hora y luego se enfrió a temperatura ambiente. A la mezcla se añadió agua (2,5 ml) y el disolvente se eliminó a vacío. El residuo se extrajo con acetato de etilo (3 x 15 ml) y las capas orgánicas combinadas se lavaron con solución saturada de NaCl (10 ml), se secó sobre Na<sub>2</sub>SO<sub>4</sub> anhidro, y se concentró para dar el bromuro deseado.

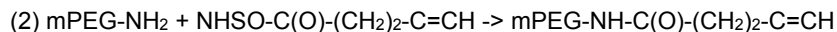
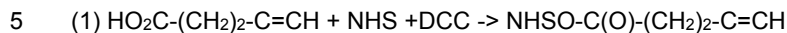
mPEG-OH 20 kDa (1,0 g, 0,05 mmol, Sunbio) se disolvió en THF (20 ml) y la solución se enfrió en un baño de hielo. Se añadió NaH (6 mg, 0,25 mmol) con agitación vigorosa durante un período de varios minutos seguido de la adición del bromuro obtenido desde arriba (2,55 g, 11,4 mmol) y una cantidad catalítica de KI. El baño de refrigeración se retiró y la mezcla resultante se calentó a reflujo durante 12 horas. Se añadió agua (1,0 mL) a la mezcla y el disolvente se eliminó a vacío. Para el residuo se añadió CH<sub>2</sub>Cl<sub>2</sub> (25 ml) y la capa orgánica se separó, se secó sobre Na<sub>2</sub>SO<sub>4</sub> anhidro, y el volumen se redujo a aproximadamente 2 ml. La adición gota a gota a una solución en éter (150 ml) dio como resultado un precipitado blanco, que se recogió para dar el derivado de PEG.

Ejemplo 25

65 Los polímeros de poli(etilenglicol) terminales que contienen alquino también se pueden obtener por acoplamiento de un polímero de poli(etilenglicol) que contiene un grupo funcional terminal a una molécula reactiva que contiene la funcionalidad de alquino como se muestra arriba. n está entre 1- y 10. R' puede ser H o un grupo alquilo pequeño de

C1 a C4.

Ejemplo 26

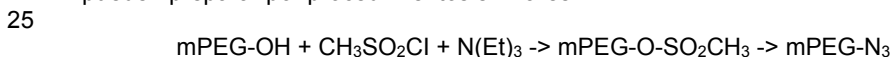


10 Ácido 4-pentinoico (2,943 g, 3,0 mmol) se disolvió en  $\text{CH}_2\text{Cl}_2$  (25 ml). N-hidroxisuccinimida (3,80 g, 3,3 mmol) y DCC (4,66 g, 3,0 mmol) se añadieron y la solución se agitó durante la noche a temperatura ambiente. El éster de NHS 7 en bruto resultante se usó en la siguiente reacción sin purificación adicional.

15 mPEG-NH<sub>2</sub> con un peso molecular de 5.000 Da (mPEG-NH<sub>2</sub>, 1 g, Sunbio) se disolvió en THF (50 ml) y la mezcla se enfrió a 4°C. Éster NHS 7 (400 mg, 0,4 mmol) se añadió en porciones con agitación vigorosa. La mezcla se dejó en agitación durante 3 horas mientras se calentaba a temperatura ambiente. Después se añadió agua (2 mL) y el disolvente se eliminó bajo vacío. Al residuo se añadió  $\text{CH}_2\text{Cl}_2$  (50 ml) y la capa orgánica se separó, se secó sobre  $\text{Na}_2\text{SO}_4$  anhidro, y el volumen se redujo a aproximadamente 2 ml. Esta solución de  $\text{CH}_2\text{Cl}_2$  se añadió a éter (150 ml) gota a gota. Se recogió el precipitado resultante y se secó a vacío.

20 Ejemplo 27

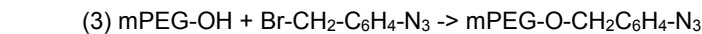
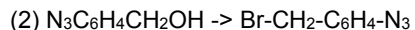
Este ejemplo representa la preparación del éster de sulfonilo de metano de poli(etilenglicol), que también se puede denominar como el metanosulfonato o mesilato de poli(etilenglicol). El tosilato correspondiente y los haluros se pueden preparar por procedimientos similares.



30 La mPEG-OH (MW=3.400, 25 g, 10 mmol) en 150 ml. de tolueno se destiló azeotropicamente durante 2 horas bajo nitrógeno y la solución se enfrió a temperatura ambiente. se añadieron 40 ml, de  $\text{CH}_2\text{Cl}_2$  seco y 2,1 ml, de trietilamina seca (15 mmol) a la solución, la solución se enfrió en un baño de hielo y se añadió 1,2 ml de cloruro de mdimetilsulfonilo destilado (15 mmol) gota a gota. La solución se agitó a temperatura ambiente bajo nitrógeno durante la noche, y la reacción se inactivó mediante la adición de 2 ml de etanol absoluto. La mezcla se evaporó bajo vacío para eliminar los disolventes, principalmente aquellos distintos de tolueno, se filtró, se concentró de nuevo al vacío y después se precipitó en 100 ml, de éter dietílico. El filtrado se lavó con varias porciones de éter dietílico frío y se secó en vacío para proporcionar el mesilato.

40 El mesilato (20 g, 8 mmol) se disolvió en 75 ml de THF y la solución se enfrió a 4° C. A la solución enfriada se le añadió azida de sodio (1,56 g, 24 mmol). La reacción se calentó a reflujo bajo nitrógeno durante 2 horas. Los disolventes se evaporaron y el residuo se diluyó con  $\text{CH}_2\text{Cl}_2$  (50 ml). La fracción orgánica se lavó con solución de NaCl y se secó sobre  $\text{MgSO}_4$  anhidro. El volumen se redujo a 20 ml y el producto se precipitó mediante la adición a 150 ml de éter seco frío.

Ejemplo 28



55 4-azidobencilo alcohol puede producirse usando el método descrito en Patente de EE.UU. 5.998.595. Cloruro de mdimetilsulfonilo (2,5 g, 15,7 mmol) y trietilamina (2,8 mL, 20 mmol) se añadieron a una solución de alcohol 4-azidobencilo (1,75 g, 11,0 mmol) en  $\text{CH}_2\text{Cl}_2$  a 0°C y la reacción se colocó en el refrigerador durante 16 horas. Una elaboración usual produjo el mesilato como un aceite amarillo pálido. Este aceite (9,2 mmol) se disolvió en THF (20 ml) y LiBr (2,0 g, 23,0 mmol) se añadió. La mezcla de reacción se calentó a reflujo durante 1 hora y luego se enfrió a temperatura ambiente. A la mezcla se añadió agua (2,5 ml) y el disolvente se eliminó a vacío. El residuo se extrajo con acetato de etilo (3 x 15 ml) y las capas orgánicas combinadas se lavaron con solución saturada de NaCl (10 ml), se secó sobre  $\text{Na}_2\text{SO}_4$  anhidro, y se concentró para dar el bromuro deseado.

60 mPEG-OH 20 kDa (2,0 g, 0,1 mmol, Sunbio) se trató con NaH (12 mg, 0,5 mmol) en THF (3,5 ml) y el bromuro (3,32 g, 15 mmol) se añadió a la mezcla junto con una cantidad catalítica de K1. La mezcla resultante se calentó a reflujo durante 12 horas. Se añadió agua (1,0 mL) a la mezcla y el disolvente se eliminó a vacío. Al residuo se añadió  $\text{CH}_2\text{Cl}_2$  (25 ml) y la capa orgánica se separó, se secó sobre  $\text{Na}_2\text{SO}_4$  anhidro, y el volumen se redujo a aproximadamente 2 ml. La adición gota a gota a una solución en éter (150 ml) dio como resultado un precipitado, que se recogió para producir mPEG-O-CH<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>N<sub>3</sub>.

Ejemplo 29

5 NH<sub>2</sub>-PEG-O-CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H (MW 3.400 Da, 2,0 g) se disolvió en una solución acuosa saturada de NaHCO<sub>3</sub> (10 ml) y la solución se enfrió a 0°C. Se añadió propionato de 3-azido-1-N-hidroxisuccinimido (5 equiv.) con agitación vigorosa. Después de 3 horas, 20 ml. de H<sub>2</sub>O se añadió y la mezcla se agitó durante 45 minutos adicionales a temperatura ambiente. El pH se ajustó a 3 con 0,5 N H<sub>2</sub>SO<sub>4</sub> y NaCl se añadió a una concentración de aproximadamente 15% en peso. La mezcla de reacción se extrajo con CH<sub>2</sub>Cl<sub>2</sub> (100 ml, x 3), se secó sobre Na<sub>2</sub>SO<sub>4</sub> y se concentró. Después de la precipitación con éter dietílico frío, el producto se recogió por filtración y se secó bajo vacío para producir el derivado de PEG omegacarboxiazida.

Ejemplo 30

15 mPEG-OMs + HC=CLi → mPEG-O-CH<sub>2</sub>-CH<sub>2</sub>-C=C-H

[673] A una solución de acetiluro de litio (4 equiv.), preparada como se conoce en la técnica y se enfrió a -78°C en THF, se añade gota a gota una solución de mPEG-OMs disueltos en THF con agitación vigorosa. Después de 3 horas, se permite que la reacción se calentara a temperatura ambiente y se inactivó con la adición de 1 ml de butanol. Después se añaden 20 ml de H<sub>2</sub>O y la mezcla se agitó durante 45 minutos adicionales a temperatura ambiente. La pH se ajustó a 3 con 0,5 N H<sub>2</sub>SO<sub>4</sub> y NaCl se añadió a una concentración de aproximadamente 15% en peso. La mezcla de reacción se extrajo con CH<sub>2</sub>Cl<sub>2</sub> (100 ml x 3), se secó sobre Na<sub>2</sub>SO<sub>4</sub> y se concentró. Después de la precipitación con éter dietílico frío, el producto se recogió por filtración y se secó bajo vacío para producir el 1-(but-3-iniloxi)metoxipoliethylenglicol (mPEG).

Ejemplo 31

Aminoácidos que contienen azida y acetileno se pueden incorporar con selección de sitios en proteínas usando los métodos descritos en L. Wang, et al., (2001), Science 292: 498-500, J. W., Chin et al., Science 301: 964-7 (2003)), J. W. Chin et al., (2002), Revista de la Sociedad Química Americana 124: 9026-9027; J. W. Chin, & P. G Schultz, (2002), Chem Bio Chem 3 (11): 1135-1137; J. W. Chin, et al., (2002), PNAS Estados Unidos de América 99: 11020-11024; y, L. Wang, y P. G. Schultz, (2002), Chem. Comm., 1: 1-11. Una vez que se incorporaron los aminoácidos, la reacción de cicloadiación se lleva a cabo con la proteína de 0,01 mM en tampón de fosfato (PB), pH 8, en presencia de un derivado de 2 mM de PEG, 1 mM de CuSO<sub>4</sub> y ~1 mg Cu-alambre durante 4 horas a 37°C.

Ejemplo 32

Este ejemplo describe la síntesis de p-acetilo-D,L-fenilalanina (pAF) y derivados de m-PEG-hidroxilamina.

40 El pAP racémico es sintetizado usando el procedimiento previamente descrito en Zhang, Z., Smith, B.A.C., Wang, L., Brock, A., Cho, C. & Schultz, P. CL, Biochemistry, (2003) 42, 6735-6746.

45 Para sintetizar el derivado de m-PEG-hidroxilamina, los siguientes procedimientos se han completado. A una solución de (N-t-Boc-aminooxi) ácido acético (0,382 g, 2,0 mmol) y 1,3-Diisopropilcarbodiimida (0,16 ml, 1,0 mmol) en diclorodimetilo (DCM, 70mL), que se agitó a temperatura ambiente (TA) durante 1 hora, metoxipoliethylenglicol se añade amina (m-PEG-NH<sub>2</sub>, 7,5 g, 0,25 mmol, Mt. 30 K, de BioVectra) y diisopropiloetilamina (0,1 ml, 0,5 mmol), La reacción se agitó a TA durante 48 horas, y luego se concentró a aproximadamente 100 mL. La mezcla se añadió gota a gota a éter frío (800 ml). El producto t-Boc-prottegido precipitó y se recogió por filtración, se lavó con éter 3x100mL. Se purificó más por redisolución en DCM (100 ML) y precipitación en éter (800 ml) dos veces. El producto se secó en vacío, produciendo 7,2 g (96%), confirmado por la prueba de RMN y prueba de Nihidrina.

50 El deBoc del producto protegido (7,0 g) obtenido anteriormente se lleva a cabo en 50% de TFA/DCM (40 ml) a 0°C durante 1 hora y luego a TA durante 1,5 horas. Después de eliminar la mayor parte de TFA en vacío, la sal de TFA del derivado de hidroxilamina se convirtió en la sal HCl mediante la adición de 4N HCl en dioxano (1mL) al residuo. El precipitado se disuelve en DCM (50 ml) y se volvió a precipitar en éter (800 ml). El producto final (6,8 g, 97%) se recoge por filtración, se lavó con éter 3x 100ml, se secó en vacío, almacenado bajo nitrógeno. Otros derivados de hidroxilamina PEG (5K, 20K) son sintetizados usando el mismo procedimiento.

Ejemplos 33

60 Estudios *in vivo* de insulina pegilada

65 PEG-insulina, la insulina no modificada y solución de tampón se administran a ratones o ratas. Los resultados mostrarán una actividad superior y la vida media prolongada de la insulina PEGilada de la presente invención en comparación con la insulina no modificada. Del mismo modo, la insulina modificada, la insulina no modificada, y solución de tampón se administran a ratones o ratas.



*Análisis farmacéutica*

5 Un polipéptido de insulina de la invención se administra por vía intravenosa o subcutánea a ratones. Los animales se sangraron antes y en puntos de tiempo después de la dosificación. Se recoge plasma de cada muestra y se analiza por radioinmunoensayo. La vida media de eliminación puede ser calculada y comparada entre los polipéptidos de insulina que comprende un aminoácido no naturalmente codificado e insulina de tipo silvestre o varios polipéptidos análogos de insulina de la invención. Del mismo modo, los polipéptidos de insulina de la invención pueden ser administradas a monos cynomolgus. Los animales se sangraron antes y en puntos de tiempo después de la dosificación. Se recoge plasma de cada muestra y se analiza por radioinmunoensayo.

15 Los polipéptidos de la invención se pueden administrar a ratas macho ZDF (ratas gordas diabéticas; 8 semanas de edad al inicio del estudio, Charles River-GMI). Las ratas son alimentadas con pienso de Purina 5008 ad libitum. Los siguientes grupos de prueba se configuran: Salina; Insulina 4U/día; polipéptidos de insulina de la invención, 8U/día agudo (grupo de dosificación aguda se dosifica una vez y se sangraron a T=0, 2, 4, 8, y 24 horas después de la dosis); polipéptidos de insulina de la invención, 6U/día; polipéptidos de insulina de la invención, 4U/día; polipéptidos de insulina de la invención, 3U/día; polipéptidos de insulina de la invención, 2U/día; Salina magra; Insulina magra 4U/día, y los grupos que recibieron 6U/día, 4U/día, y 2U/día de polipéptidos de insulina en ratas no diabéticas. Grupos magros representan ratas ZDF delgadas no diabéticas.

20 Se inyectan los compuestos s.c. se mantienen una vez por día, y luego grupos secundarios con las mismas cantidades de administración con inyecciones cada dos días. Las ratas de control se inyectaron con vehículo (PBS; 0,1 ml). Después de 7 días de dosificación, los animales se sometieron a una prueba de tolerancia a la glucosa oral. La sangre para la glucosa y los triglicéridos se recogen por el clip de sangrado de la cola sin anestesia. Los polipéptidos de insulina pueden reducir los niveles de glucosa en plasma de una manera dependiente de la dosis. También ratas ZDF magras se pueden ensayar para la hipoglucemia después de la exposición a los polipéptidos de insulina de la invención en comparación con ratas tratadas con insulina de tipo silvestre.

*Modelo de obesidad ob/ob*

30 El modelo de ratón ob/ob es un modelo animal para hiperglucemia, resistencia a la insulina y la obesidad. Los niveles de glucosa en plasma después del tratamiento con polipéptido de insulina en comparación con grupos de vehículos y control de la insulina pueden ser medidos en ratones ob/ob. En este modelo de obesidad, los grupos de prueba de ratones machos ob/ob (7 semanas de edad) se inyectaron con vehículo solo (PBS), insulina (4 U/día), o polipéptidos de insulina como se describió anteriormente, por vía subcutánea (0,1 ml, una vez al día y en otros grupos de prueba una vez cada dos días y en otros grupos de prueba una vez por semana) durante catorce días. La sangre se recoge por clip de cola de sangrado en los días 1, 3, y 7, 9, 11, 14, una hora después de la primera inyección del compuesto, y los niveles de glucosa en plasma se miden usando el protocolo estándar. Polipéptidos de insulina de la invención estimulan la absorción de glucosa si reducen los niveles de glucosa en plasma en comparación con el grupo de control del vehículo. Los niveles de triglicéridos se pueden comparar después del tratamiento con polipéptidos de insulina de la invención en comparación con otras moléculas. El polipéptido puede ser administrado a los ratones a través de múltiples dosis, infusión continua, o una dosis única, etc.

Ejemplo 34

45 El sistema de expresión de insulina usando Novagen (promotor T7 inducible; se describe en detalle en el manual del sistema pET, versión 9), vector de expresión pET30a y cepa de expresión BL21 (DE3).

50 2 ml de LB/kanamicina (10 µg/ml) se inoculan con un barrido de placa de BL21 (DE3) transformada con el análogo deseado. Esto disminuye los efectos causados por variabilidad de colonia a colonia en los niveles de expresión. Este cultivo se hace crecer durante la noche a 37°C con agitación vigorosa y el día siguiente, 10 ml de cultivo de LB/kanamicina se inocula con 1 ml del cultivo durante la noche (OD600 ~ 0,4-0,5). La ml restante del cultivo durante la noche se puede congelar como reserva de glicerol.

55 10 ml del cultivo crecido se pone a 37°C y 250 rpm durante 30-45 min hasta que OD600 alcanza 0,8-0,9. Esto es inducido con 1mM de IPTG (con 1mL que ser un control de cultivo no inducido) y se cosecha generalmente 3-4 horas posinducción y se analizó en SDS-PAGE.

60 También es posible hacer un transcurso de tiempo de la expresión (por ejemplo, 1,2,4,6 horas posinducción y O/N) para determinar la tasa de acumulación, estabilidad de la proteína, etc.

65 Análisis de gel: en el punto de tiempo deseado posinducción 1mL se recoge del cultivo, las células se centrifugan, se resuspendieron en 100 µl de 2X SDS-PAGE, se sonicaron para reducir la viscosidad y 10 µl se ejecutan en SDS-PAGE. Si se desea, se puede comparar con el control o controles no inducidos y/o control positivo conocido o estándar y el nivel de expresión se puede estimar (por ejemplo, buena expresión podría ser @> 100 µg/ml). El análisis de transferencia Western también se puede utilizar. También es posible dejar a un lado 4 ml de los cultivos,

preparar cuerpos de inclusión (si expresan análogos insolubles) y obtener análisis de espectrometría de masas en ellos para confirmar la identidad de la proteína sobreexpresada.

5 Para la expresión de proteínas a gran escala, > 250 ml de LB/kanamicina (10 tg/ml) se inoculan con 250µL de reserva de glicerol congelado y se hicieron crecer durante la noche. Al día siguiente, 10 X 1L LB/cultivos de kanamicina se inoculan con 25 mL del cultivo de una noche (OD<sub>600</sub> ~ 0,1).

10 Los cultivos de 1L se hacen crecer a 37°C y 250 rpm ~ 2h hasta OD<sub>600</sub> alcanza 0,8-0,9. Esto es inducido luego con 1 mM de IPTG y se recogieron 4 h posinducción o la siguiente mañana (cosecha puede utilizar centrifugación durante 15 min a 4.000 rpm). Los granulados se enjuagan con 50 mM Tris-HCl, pH 8,0 (50 ml por granulado + 50 ml para enjuagar la botella) si se desea para reducir la endotoxina y facilitar la purificación. Los granulados se agruparon y se centrifugaron de nuevo.

### 15 Ejemplo 35

#### *Estudio de la expresión Pichia - Prep de ADN, electroporación, protocolos de expresión*

20 Este ejemplo proporciona un protocolo para la preparación de polipéptidos de insulina de la presente invención en *Pichia*. SEQ ID NOs: 34, 35, 36, y 37 se utilizaron, y la Figura 8 muestra un plásmido utilizado para la clonación en *Pichia* y este u otros plásmidos modificados pueden usarse para obtener la expresión de proteínas de polipéptidos de insulina en *Pichia*, las modificaciones realizadas en el plásmido usando métodos conocidos en la técnica.

25 En el día 1 del protocolo, hay una digestión durante la noche, típicamente usando enzima 2U por µg ADN para ser digerido y 10mL de cultivo YPhyD se inocula durante la noche en un matraz de 50mL, agitándose a 260rpm a 30°C de la reserva de glicerol.

#### *Preparación de ADN*

30 ADN se precipitó mediante la adición primero de 1/10º volumen estéril 3M NaOAc y después de 0,7 volúmenes estériles IPA y después la muestra es vigorosamente mezclada y la precipitación se continuó durante una noche a -20°C o a -70°C hasta que se congele. El ADN se sedimentó por centrifugación (Centrífuga sobre mesa 14.000 rpm/10 minutos), el sobrenadante se eliminó, y el sedimento se lavó usando 500 µl de ETOH estéril al 70%. Girar (centrífuga sobre mesa 14.000 rpm/10 minutos) y decantar el sobrenadante y granulado secado al aire durante 15-20 minutos. El sedimento de ADN se resuspende con agua estéril a 1µg/µl y transformar *Pichia* con 10µg de ADN.

#### 35 *Electroporación*

40 Mediante el uso de cultivo durante la noche con OD<sub>600</sub>, diluir en YPhyD a OD<sub>600</sub>=0,2. Agitar el cultivo a 260 rpm a 30°C hasta que OD<sub>600</sub> alcanza 0,8-1,0. Recoger las células por centrifugación (4000 rpm/5 minutos). Decantar el medio, lavar las células en 20 ml, de agua estéril helada, decantar de nuevo y repetir. Después del lavado, lavar el granulado en 20 ml de 1 M de sorbitol helado estéril, decantar, y resuspender el granulado celular se lavó en 600 µl de 1 M de sorbitol frío, entonces esto puede ser almacenado en hielo.

45 A partir de las células lavadas, mezclar 50 µL con 10 µg de ADN linealizado en 1,5 mL de tubo Eppendorf estéril, mezclar suavemente e incubar en hielo durante 25 minutos. Transferir la mezcla de células/ADN a 0,2 cm de cubeta preenfriada utilizando puntas de pipeta de largo. Electroporar células utilizando la unidad BioRad GenePulsar II con los siguientes ajustes: 2000 V, 200 ohmios, 25 µFd (utilizar un solo pulso) e inmediatamente añadir 0,5 ml de medio YPhyD a la cubeta y se mezcla mediante pipeteo. Transferir todo el contenido al tubo de fondo redondo estéril y agitar suavemente (200 rpm) durante 30 minutos a 30°C. Sembrar y extender células uniformemente e Incubar las placas, invertidas, durante 3 días a 30°C.

50 Después de tres días de incubación, recoger colonias con un bucle e inocular 10 ml de medios BYPhyD de comunicación en un matraz de 50 mL e incubar durante 3 días a 30°C. Contar las colonias en las placas de 20µl y registrar el número promedio y luego las células de cosecha, primero mediante la preparación de 2 conjuntos de crioviales etiquetados con el nombre de cepa y el número de clon, la insulina (es decir, proteína expresada), y la fecha. Transferencia del cultivo a 15 ml de tubo cónico, tomar OD<sub>600</sub> de cada cultivo, diluir el cultivo 1:50 o 1:20 en medio YPhyD. Guardar un alícuota del cultivo para la reserva de glicerol. Entonces levadura en gránulos a 4000 rpm durante 5 min a TA, transferir el sobrenadante a un nuevo tubo cónico de 15 ml con la etiqueta, y se almacena a -20 o -80°C hasta que sea necesario para los datos analíticos.

#### 60 *Análisis de la expresión de proteínas*

Ejecutar muestras en 4-12% de gel NuPAGE TB (Novex). Reactivos de SDS-PAGE utilizados de Invitrogen, analizar mediante transferencia Western o análisis de gel teñido.

#### 65 *Formulaciones de los medios*

**Dextrosa de fitona de levadura tamponada (BYPhyD)**

5	Extracto de levadura	10 g/L
	Peptona de fitona	20 g/L
	1 M de tampón de fosfato de potasio (pH 6)	100 ml/L
10	10X YNB	100 mL/L
	20% de dextrosa	100 mL/L

**Dextrosa de fitona de levadura (YPhyD)**

	Extracto de levadura	10 g/L
20	Peptona de fitona	20 g/L
	20% de dextrosa	100 ml/L

**10X YNB (13,4% Base nitrogenada de levadura con sulfato de amoníaco sin aminoácidos)**

25	Base nitrogenada de levadura	134 g/L
----	------------------------------	---------

Ejemplo 36*Producción de Insulina A21G*

30 En este ejemplo, 4,0L de cultivo se fermentó para producir 13,4 g de pasta celular húmeda y una preparación de cuerpos de inclusión se realizó con y sin TritonX100. Cuerpos de inclusión húmedos 2,07 g fueron producidos de esta manera, y a continuación se produjo la solubilización y replegamiento. Los cuerpos de inclusión se resuspendieron con 200 ml de H<sub>2</sub>O por gramo de cuerpos de inclusión húmedos (IBS) a una concentración final de 3

35 mM y se añadió cisteína a la resuspensión. IB entonces se solubilizan por aumento del pH a 11,5 por hora a TA. A continuación se dejó que se produjera el replegamiento dejando caer el pH del material solubilizado a 10,6 ± 0,1 y se almacenó a 2-8°C durante 72 horas y estos resultados se muestran en la Figura 10. La reacción de replegamiento se detuvo mediante la adición de HCl a un pH final del 3.0, 0.45 μM se filtró y se almacenó a 2-8°C hasta su posterior procesamiento.

40 La proteína replegada se purifica (como se muestra en la Figura 11) al aumentar el pH de replegamiento se inactivó a 8,0 con base Tris y se cargó directamente en una columna de Q HP. La conductividad de la carga en el ejemplo mostrado era > 3.5mS/cm. Las condiciones de operación fueron (A) 20 mM Tris, 8,0; (B) 20 mM Tris, 8,0; 200 mM de NaCl y hubo 0-100%B sobre 30CV. La proinsulina correctamente replegada se agrupó y se recuperó 79 mg de proinsulina.

45

50 Ultrafiltración/diafiltración (UF/DF) tuvo lugar y la precipitación se realizó con cinc de 25 mM, la proteína precipitada se volvió a suspender a una concentración de 2 mg/mL con 20 mM NaOAc, 4,0, 30% de ACN, 5 mM de EDTA y 20K PEG se añadió a una relación molar final de 10: 1 PEG a la proteína y se dejó incubar durante 48-72 horas a 28°C.

La reacción de PEG se diluyó 1:10 en 0,5X PEG tampón A, 0,22μM se filtró y se ejecutó sobre una columna de SP 650S. Las condiciones de operación fueron (A) 10 mM de NaOAc, 4,0, 1 mM de EDTA; (B) 10 mM de NaOAc, 4,0, 1 mM de EDTA, 0,4 M de NaCl; 0-50%B sobre muestras 20CV y PEG formuladas en 10 mM de Nacitrato, 6,5; 150mM de NaCl y esto se muestra en la Figura 12.

55

Se utilizaron estos métodos para producir una variedad de polipéptidos de insulina con aminoácidos no naturales y una gama de 0,122mg para las cantidades de proteína final de las variantes purificadas y PEGiladas. Se encontró que ACN ayudaba a solubilizar la mezcla PEG/proteína en la reacción de PEG y precipitación de cinc a pl facilitó la concentración en presencia de CAN.

60

Ejemplo 37*Ensayos de polipéptidos de insulina en el modelo de rata diabética STZ*

65 Las ratas se trataron con una sola inyección de dos polipéptidos de insulina diferentes de la presente invención: A14pAF-PEG-A21N insulina LysPro y A14pAF-PEG-A21G insulina LysPro, cada uno en cuatro niveles diferentes;

568 nmol/kg; 94 nmol/kg; 56,8 nmol/kg; y 9,4 nmol/kg (3, 0,5, 0,3, 0,05 mg/kg de contenido de insulina).

La glucosa en sangre se midió a continuación en 1, 2, 4, 8, 12, 24, 48, 72, 96, 120 h después de la dosis y el área media bajo la curva, las mediciones de glucosa en sangre a partir de 0-120 horas para cada una de estas, incluyendo un control de la inyección de único vehículo, se muestran en la Figura 9.-

### Ejemplo 38

Ensayo clínico en humanos de la seguridad y/o la eficacia de la insulina PEGilada que comprende un aminoácido no naturalmente codificado.

Objetivo: Observar la seguridad y farmacocinética de administración subcutánea de insulina pegilada humana recombinante que comprende un aminoácido no naturalmente codificado.

Pacientes Dieciocho voluntarios sanos que oscilan entre los 20-40 años de edad y un peso de entre 60-90 kg se inscribieron en el estudio. Los sujetos no tendrán valores de laboratorio clínicamente significativos anormales para hematología o química sérica, y un examen de toxicología de orina negativo, examen de VIH, y el antígeno de superficie de hepatitis B. No deben tener ninguna evidencia de lo siguiente: hipertensión; antecedentes de ninguna enfermedad hematológica primaria; historia de hepática significativa, renal, cardiovascular, gastrointestinal, genitourinaria, metabólica, enfermedad neurológica; antecedentes de anemia o trastorno convulsivo; una sensibilidad conocida a productos bacterianos o derivados de mamíferos, PEG, o albúmina de suero humano; consumidor habitual de bebidas que contienen cafeína; participación en cualquier otro ensayo clínico o que hubieran transfundido sangre donada o dentro de los 30 días de ingreso en el estudio; que han tenido una exposición a la insulina dentro de los tres meses de entrada en el estudio; que hubieran tenido una enfermedad dentro de los siete días de la entrada en el estudio; y tuvieran anomalías significativas en el examen físico antes del estudio o las evaluaciones clínicas de laboratorio dentro de los 14 días de la entrada en el estudio. Todos los sujetos son evaluables para la seguridad y todas las extracciones de sangre para el análisis farmacocinético se recogen como estaba previsto, todos los estudios se llevan a cabo con la aprobación del comité de ética institucional y el consentimiento del paciente.

Diseño del estudio: Este será un estudio de fase I, de un centro, etiqueta abierta, estudio de dos períodos cruzado aleatorizado en voluntarios sanos de sexo masculino, Dieciocho sujetos son asignados al azar a uno de dos grupos de secuencia de tratamiento (nueve sujetos/grupo). La insulina se administra durante dos períodos de dosificación separados como una inyección bolus s.c., en el muslo superior usando dosis equivalentes de la insulina PEGilada que comprende un aminoácido no naturalmente codificado y el producto comercialmente disponible elegido. La dosis y la frecuencia de administración del producto comercialmente disponible es como se indica en la etiqueta del paquete. La dosificación adicional, la frecuencia de dosificación, u otro parámetro deseado, utilizando los productos disponibles comercialmente pueden ser añadidos al estudio mediante la inclusión de grupos adicionales de sujetos. Cada período de dosificación está separado por un período de lavado de 14 días. Los sujetos están confinados al centro de estudio por lo menos 12 horas antes y 72 horas después de la dosificación para cada uno de los dos períodos de dosificación, pero no entre los períodos de dosificación. Grupos adicionales de los sujetos pueden añadirse si se ha de examinar la dosificación adicional, la frecuencia, u otro parámetro, para la insulina PEGilada también. La formulación experimental de la insulina es la insulina PEGilada que comprende un aminoácido no naturalmente codificado.

Muestras de sangre: la sangre de serie se extrae por punción de la vena directa antes y después de la administración de la insulina. muestras de sangre venosa (5 ml) para la determinación de las concentraciones de insulina en suero se obtienen a aproximadamente 30, 20, y 10 minutos antes de la dosificación (3 muestras de línea de base) y a aproximadamente los siguientes tiempos después de la dosificación: 30 minutos y a 1, 2, 5, 8, 12, 15, 18, 24, 30, 36, 48, 60 y 72 horas. Cada muestra de suero se dividió en dos alícuotas. Todas las muestras de suero se almacenaron a 20°C. Las muestras de suero se envían en hielo seco. Pruebas clínicas de laboratorio en ayunas (hematología, química sérica y análisis de orina se llevan a cabo inmediatamente antes de la dosis inicial el día 1, la mañana del día 4, inmediatamente antes de la dosificación en el día 16, y la mañana del día 19.

Métodos bioanalíticos: Un kit de ELISA se utiliza para la determinación de las concentraciones de insulina en suero.

Determinaciones de seguridad: Los signos vitales se registran inmediatamente antes de cada dosificación (Días 1 y 16), y a las 6, 24, 48, y 72 horas después de cada dosificación. Determinaciones de seguridad se basan en la incidencia y el tipo de eventos adversos y los cambios en las pruebas de laboratorio clínico de la línea de base. Además, se evaluaron los cambios desde antes del estudio en las mediciones de signos vitales, incluyendo la presión arterial, y los resultados del examen físico.

Análisis de datos Valores de concentración en suero después de la dosis de se corrigen para las concentraciones de insulina de línea base de dosis previa restando de cada uno de los valores posteriores a la dosis la concentración de insulina basal media determinada a partir de un promedio de los niveles de insulina a partir de las tres muestras recogidas a los 30, 20, y 10 minutos antes de la dosificación. Concentraciones previas a la dosis de insulina en

5 suero no se incluyen en el cálculo del valor medio si están por debajo del nivel de cuantificación del ensayo. Parámetros farmacocinéticos se determinan a partir de datos de concentración en suero corregidos por las concentraciones de insulina de línea base. Parámetros farmacocinéticos se calculan por métodos independientes en un modelo de sistema informático Digital Equipment Corporación VAX8600 utilizando la última versión del software BIOAVL. Los siguientes parámetros farmacocinéticos se determinan: concentración sérica máxima (Cmax); tiempo para alcanzar la concentración máxima en suero (tmax); área bajo la curva concentración/tiempo (AUC) desde el tiempo cero hasta el último tiempo de muestreo de sangre (AUC0-72) calculado con el uso de la regla trapezoidal lineal; y vida media de eliminación terminal (t1/2), calculada a partir de la constante de velocidad de eliminación. La constante de velocidad de eliminación se estimó por regresión lineal de puntos de datos consecutivos en la región lineal terminal del diagrama log-lineal concentración/tiempo. La media, la desviación estándar (SD), y el coeficiente de variación (CV) de los parámetros farmacocinéticos se calculan para cada tratamiento. La relación de los medios de parámetros (formulación preservada/formulación no preservada) se calcula.

15 Resultados de seguridad: La incidencia de acontecimientos adversos se distribuye por igual en todos los grupos de tratamiento. No hay cambios clínicamente significativos de las pruebas de línea de base o estudios clínicos de laboratorio previos al estudio o presiones sanguíneas, y no hay cambios notables de pre-estudio en los resultados del examen físico y las mediciones de signos vitales. Los perfiles de seguridad para los dos grupos de tratamiento deben ser similares.

20 Resultados farmacocinéticos: Insulina en suero media de los perfiles concentración-tiempo (sin corregir para los niveles de insulina de línea de base) en los 18 sujetos después de recibir la insulina PEGilada que comprende un aminoácido no naturalmente codificado en cada punto de tiempo medido. Todos los sujetos deben tener las concentraciones de insulina basal predosis dentro del intervalo fisiológico normal. Los parámetros farmacocinéticos se determinaron a partir de los datos en suero corregidos por concentraciones de insulina de línea de base antes de la dosis media y se determinan la Cmax y tmax. El tmax medio para cualquier comparador clínico elegido es significativamente más corto que el ARNtx para la insulina PEGilada que comprende el aminoácido no naturalmente codificado. Valores de vida media terminales son significativamente más cortos para el comparador preclínico probado en comparación con la vida media terminal para la insulina PEGilada que comprende un aminoácido no naturalmente codificado.

30 Aunque el presente estudio se realizó en sujetos masculinos sanos, las características de absorción similares y perfiles de seguridad se podría anticipar en otras poblaciones de pacientes; tales como pacientes masculinos o femeninos con la diabetes, hombres o mujeres con cáncer o insuficiencia renal crónica, pacientes con insuficiencia renal pediátricos, pacientes en programas de predepósito, o pacientes programados para cirugía electiva.

35 En conclusión, las dosis únicas administradas por vía subcutánea de insulina pegilada que comprenden aminoácidos no naturalmente codificados serán seguras y bien toleradas por los sujetos masculinos sanos. En base a una incidencia comparativa de acontecimientos adversos, valores de laboratorio clínico, signos vitales, y los resultados del examen físico, los perfiles de seguridad de las formas comercialmente disponibles de insulina y la insulina PEGilada que comprenden aminoácidos no naturalmente codificados serán equivalentes. La insulina PEGilada que comprende aminoácidos no naturalmente codificados potencialmente proporciona gran utilidad clínica a los pacientes y los profesionales sanitarios.

45

Tabla 1: Secuencias Citadas

SEQ ID NO:	Nombre de secuencia	Secuencia
1	Cadena A de insulina, secuencia de aminoácidos	GTVEQCCTSICSLYQLENYCN
2	Cadena B de insulina, secuencia de aminoácidos	FVNQHLCGSHLVEALYLVCGERGFFYTPKT
3	Cadena A de insulina lispro, secuencia de aminoácidos (Humalog)	GTVEQCCTSICSLYQLENYCN
4	Cadena B de insulina lispro, secuencia de aminoácidos (Humalog)	FVNQHLCGSHLVEALYLVCGERGFFYTPKT
5	Cadena A de insulina aspart, secuencia de aminoácidos (Novolog)	GTVEQCCTSICSLYQLENYCN
6	Cadena B de insulina aspart, secuencia de aminoácidos (Novolog)	FVNQHLCGSHLVEALYLVCGERGFFYTDKT
7	Cadena A de insulina glulisina, secuencia de aminoácidos (Apidra)	GIVEQCCTSICSLYQLENYCN
8	Cadena B de insulina glulisina, secuencia de aminoácidos (Apidra)	EVkQI-ILCGSHLVEALYLVCGERGFFYTD <sup>e</sup> T
9	Cadena A de insulina detemir, secuencia de aminoácidos (Levemir)	GTVEQCCTSICSLYQLENYCG
10	Cadena B de insulina detemir, secuencia de aminoácidos (Levemir)	FVNQIILCGSHLVEALYLVCGERGFFYTDK
11	Cadena A de insulina glargina, secuencia de aminoácidos (Lantus)	GIVEQCCTSICSLYQLENYCG
12	Cadena B de insulina glargina, secuencia de aminoácidos (Lantus)	FVNQHLCGSI-ILVEALYLVCGERGFFYTPKTRR
13	Secuencia de aminoácido de longitud total con líder	MALWMRLLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPIJALEGSLQKRGIVEQCCTSTCSLYQLENYCN
14	Secuencia de aminoácido de longitud total sin líder	FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVEL000POAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN

Secuencias de ADN de insulina lispro para la expresión en Pichia

LisPro & Longitud en aa	Configuración	Longitud en bp	Sec. ADN
INS1 69 SEQ ID NO: 34	<b>KR- EEAEAEAPK- B1-30-<u>RRLQKR- A1-21</u></b>	213	<b><u>aagcgtgaggaggetgaggetgaggetgagcctaag</u></b> <u>ttcgtcaaccaacacttgtgcggttcacttgg</u> <u>ttgaggcccttacttggtttgcggtgagcgtggtttctctacaccaag</u> <u>cctact cgtcgtttacaaaagcgt</u> <u>ggatcgttgagcaatgetgcacete</u> <u>tatctgctcctgttaccattggagaactactgcaac taatga</u>
INS2 59 SEQ ID NO: 35	<b>KR-B1-30- <u>RRLQKR-A1-21</u></b>	183	<b><u>Aagcgtttcgtcaaccaacacttgtgcggttctcacttgg</u></b> <u>ttgaggcccttacttggtttgcggtgagcgtggtttctt</u> <u>ctacaccaagcctact</u> <u>cgtcgtttacaaaagcgt</u> <u>ggatc gttgagcaatgetgcacete</u> <u>tatctgctcctgttaccattggagaactactgcaac taatga</u>
INS3 55 SEQ ID NO: 36	<b>KR-B1-30-KR-<u>A1- 21</u></b>	171	<b><u>aagcgtttcgtcaaccaacacttgtgcggttcactt</u></b> <u>ggttgaggcccttacttgg ttgcggtgagcgtggtttctct</u> <u>acaccaagcctactaagcgt ggatcgttgagcaatget</u> <u>gcaceteatctgctcctg taccattggagaactactgcaac</u> <u>taatga</u>
INS4 65 SEQ ID NO: 37	<b>KR- EEAEAEAPK- B1-30-KR-<u>A1-21</u></b>	201	<b><u>aagcgtgaggaggetgaggetgaggetgagcctaag</u></b> <u>ttcgtc aaccaacacttgtgcggttcacttggttgaggcccttact</u> <u>ttggttgcggtgagcgtggtttctctacaccaagcctactaagcgt</u> <u>ggatcgttgagcaatgetgcaceteatctgctcctgttaccattgga</u> <u>gaactactgcaac taatga</u>

Lista de Secuencias

[0707]

5

<110> Miao, Zhenwei Krawitz, Denise Kravnov, Vadim Tian, Feng Sim, Bee-Cheng Ho, Lillian Hays Putnam, Anna-Maria A Knudsen, Nick DeGuzman, Michael

<120> Polipéptidos de insulina modificados y sus usos

<130> AMBX-0146.00PCT

<160> 37

<170> PatentIn version 3.4

<210> 1

<211> 21

<212> PRT

<213> Homo sapiens

<400> 1

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu  
 1 5 10 15

Glu Asn Tyr Cys Asn 20

<210> 2

<211> 30

<212> PRT

<213> Homo sapiens

<400> 2

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr  
 20 25 30

<210> 3

<211> 21

<212> PRT

<213> Artificial

<220>

<223> Una cadena de Insulina Lispro, secuencia de amino acido (Humalog)

<400> 3

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu  
 1 5 10 15

Glu Asn Tyr Cys Asn  
 20

<210> 4



5 <400> 7  
 Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu  
 1 5 10 15  
 Glu Asn Tyr Cys Asn  
 20

10 <210> 8  
 <211> 30  
 <212> PRT  
 <213> Artificial

15 <220>  
 <223> Una cadena de Insulina gulsina, secuencia de amino acido (Apiedra)

20 <400> 8  
 Phe Val Lys Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15  
 Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Asp Glu Thr  
 20 25 30

25 <210> 9  
 <211> 21  
 <212> PRT  
 <213> Artificial

30 <220>  
 <223> Una cadena de Insulina detemir, secuencia de amino acido (Levermir)

35 <400> 9  
 Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu  
 1 5 10 15  
 Glu Asn Tyr Cys Gly  
 20

40 <210> 10  
 <211> 20  
 <212> PRT  
 <213> Artificial

45 <220>  
 <223> Una cadena de Insulina detemir, secuencia de amino acido (Levermir)

50 <400> 10  
 Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15  
 Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Asp Lys  
 20 25

55

60

65



5  
 Gln Val Gln Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu  
 65 70 75 80

10  
 Ala Leu Gln Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys Cys  
 85 90 95

15  
 Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn  
 100 105 110

20  
 <210> 14  
 <211> 88  
 <212> PRT  
 <213> Homo sapiens  
 <400> 14

25  
 Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr  
 1 5 10 15

30  
 Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg  
 20 25 30

35  
 Glu Ala Gln Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly Gly Pro  
 35 40 45

40  
 Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys  
 50 55 60

45  
 Arg Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln  
 65 70 75 80

50  
 Leu Glu Asn Tyr Cys Asn  
 85

55  
 <210> 15  
 <211> 306  
 <212> PRT  
 <213> Artificial  
 <220>  
 <221> Derivado sintético demutante de Metanococos jannaschii sintético  
 <400> 15

60  
 Met Asp Gln Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

65  
 Glu Glu Gln Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Val  
 20 25 30

ES 2 632 504 T3

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

Tyr Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu His Gly Leu Asp Lys  
 100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Glu Val Asn Gly Ile His  
 145 150 155 160

Tyr Glu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
 165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
 180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
 195 200 205

Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
 210 215 220

Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
 225 230 235 240

Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
 245 250 255

Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
 260 265 270

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
 275 280 285

ES 2 632 504 T3

	Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys	
5	290 295 300	
	Arg Leu	
	305	
10	<210> 16 <211> 77 <212> DNA <213> Artificial	
15	<220> <223> tRNA derivado mutante de Metanococo jannaschii tRNA	
	<400> 16	
20	ccggcggtag ttcagcaggg cagaacggcg gactctaaat ccgcatggcg ctggttcaaa 60	
	tccagcccgc cggacca 77	
25	<210> 17 <211> 77 <212> DNA <213> Metanococo jannaschii	
30	<400> 17	
35	ccggcggtag ttcagcaggg cagaacggcg gactctaaat ccgcatggcg ctggttcaaa 60	
	tccagcccgc cggacca 77	
40	<210> 18 <211> 88 <212> DNA <213> Artificial	
45	<220> <223> Un amber supresor optimizado tRNA	
	<400> 18	
50	cccagggtag ccaagctcgg ccaacggoga cggactctaa atccgttctc gtaggagttc 60	
	gagggttoga atcccttccc tgggacca 88	
55	<210> 19 <211> 89 <212> DNA <213> Artificial	
60	<220> <223> Un optimizado AGGA marco de lectura supresor tRNA	
	<400> 19	
65		

ES 2 632 504 T3

5  
gcgagggtag ccaagctcgg ccaacggcga cggacttct aatcgttct cgtaggagt 60  
cgaggggtcg aatccctccc ctgcacca 89

10

<210> 20  
<211> 306  
<212> PRT  
<213> Artificial

15

<220>  
<223> Aminoacido tRNA sintético por la incorporación de p-acido-L-fenilalina B

20

<400> 20

25

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Gly  
 20 25 30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Thr Phe Gln Leu Asp Lys  
 100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Thr Tyr Tyr  
 145 150 155 160

Tyr Leu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
 165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
 180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
 195 200 205

Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala

5  
 210 215 220

10 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
 225 230 235 240

15 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
 245 250 255

20 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
 260 265 270

25 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
 275 280 285

30 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
 290 295 300

35 Arg Leu  
 305

40 <210> 21  
 <211> 306  
 <212> PRT  
 <213> Artificial

45 <220>  
 <223> Aminoacido tRNA sintético por la incorporación de p-acido-L-fenilalina

<400> 21

50

55

60

65



ES 2 632 504 T3

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Gly  
 20 25 30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Ser Phe Gln Leu Asp Lys

ES 2 632 504 T3

5                                    100                                    105                                    110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
115                                    120

10 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
130                                    135                                    140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Thr Ser His  
145                                    150                                    155                                    160

15 Tyr Leu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
165                                    170

20 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
180                                    185

25 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
195                                    200                                    205

Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
210                                    215                                    220

30 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
225                                    230                                    235

35 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
245                                    250                                    255

Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
260                                    265                                    270

40 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
275                                    280                                    285

45 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
290                                    295                                    300

Arg Leu  
305

50 <210> 22  
<211> 305  
<212> PRT  
<213> Artificial

55 <220>  
<221> Aminoacido tRNA sintético por la incorporación de p-acido-L-fenilalina

60

65

ES 2 632 504 T3

<400> 22

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Ala  
 20 25 30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Pro Phe Gln Leu Asp Lys  
 100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Ala Ile Tyr  
 145 150 155 160

Leu Ala Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile His  
 165 170 175

Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His Asn  
 180 185 190

Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser Lys  
 195 200 205

Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala Lys  
 210 215 220

Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro Ile  
 225 230 235 240

5 Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys Arg  
 245 250 255

Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu Leu  
 260 265 270

10 Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys Asn  
 275 280 285

15 Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys Arg  
 290 295 300

Leu  
 305

20 <210> 23  
 <211> 305  
 <212> PRT  
 <213> Artificial

25 <220>  
 <223> Aminoacido tRNA sintético por la incorporación de p-acido-L-fenilalina

<400> 23

30 Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Ala  
 20 25 30

35 Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

40 Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

45 Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Pro Phe Gln Leu Asp Lys  
 100 105 110

50 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125

55

60

65

5 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
130 135 140

10 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Ile Pro Tyr  
145 150 155 160

15 Leu Pro Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile His  
165 170 175

20 Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His Asn  
180 185 190

25 Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser Lys  
195 200 205

30 Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala Lys  
210 215 220

35 Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro Ile  
225 230 235 240

40 Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys Arg  
245 250 255

45 Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu Leu  
260 265 270

50 Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys Asn  
275 280 285

55 Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys Arg  
290 295 300

60 Leu  
305

<210> 24  
<211> 305  
<212> PRT  
<213> Artificial

<220>  
<223> Aminoacido tRNA sintético por la incorporación de p-acido-L-fenilalina

<400> 24

65 Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Ala  
 20 25 30  
 Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45  
 Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60  
 Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80  
 Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95  
 Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Lys Phe Gln Leu Asp Lys  
 100 105 110  
 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125  
 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140  
 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Ala Ile Tyr  
 145 150 155 160  
 Leu Ala Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile His  
 165 170 175  
 Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His Asn  
 180 185 190  
 Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser Lys  
 195 200 205  
 Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala Lys  
 210 215 220  
 Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro Ile  
 225 230 235 240  
 Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys Arg  
 245 250 255  
 Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu Leu  
 260 265 270

5

10

Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys Asn  
 275 280 285

15

Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys Arg  
 290 295 300

Leu  
 305

20

<210> 25  
 <211> 306  
 <212> PRT  
 <213> Artificial

25

<220>  
 <221> Aminoacido tRNA sintético por la incorporación de p-acido-ferritarina

<400> 25

30

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Thr  
 20 25 30

35

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

40

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

45

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

50

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Asn Phe Gln Leu Asp Lys  
 100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125

55

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140

60

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Pro Leu His  
 145 150 155 160

65

5 Tyr Gln Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
165 170 175

10 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
180 185 190

15 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
195 200 205

20 Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
210 215 220

25 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
225 230 235 240

30 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
245 250 255

35 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
260 265 270

40 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
275 280 285

45 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
290 295 300

50 Arg Leu  
305

<210> 28  
<211> 305  
<212> PRT  
<213> Artificial

<220>  
<223> Aminoacido tRNA sintético por la incorporación de p-acido-ferritelarina

<400> 28

55 Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
1 5 10 15

60 Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Thr  
20 25 30

65 Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
35 40 45



Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
50 55 60

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Ser Phe Gln Leu Asp Lys  
100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Pro Leu His  
145 150 155 160

Tyr Gln Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
195 200 205

Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
210 215 220

Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
225 230 235 240

Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
245 250 255

Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
260 265 270

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
275 280 285

Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
290 295 300

5

Arg Leu  
305

10

<210> 27  
<211> 306  
<212> PRT  
<213> Artificial

15

<220>  
<221> Aminoacido tRNA sintético por la incorporación de p-acido-ferritelarina  
<400> 27

20

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
1 5 10 15

25

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Leu  
20 25 30

30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
35 40 45

35

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
50 55 60

40

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
65 70 75 80

45

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
85 90 95

50

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Thr Phe Gln Leu Asp Lys  
100 105 110

55

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
115 120 125

60

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
130 135 140

65

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Pro Val His  
145 150 155 160

Tyr Gln Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
180 185 190

ES 2 632 504 T3

5           Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
                  195                   200                   205

10       Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
                  210                   215                   220

15       Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
                  225                   230                   235                   240

20       Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
                          245                   250                   255

25       Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
                          260                   265                   270

30       Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
                          275                   280                   285

35       Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
                          290                   295                   300

40       Arg Leu  
                  305

45       <210> 28  
          <211> 306  
          <212> PRT  
          <213> Artificial

50       <220>  
          <221> Aminoacido tRNA sintético por la incorporación de p-acido-ferritelarina

55       <400> 28

60       Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
          1                   5                   10                   15

65       Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Thr  
                  20                   25                   30

70       Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
                  35                   40                   45

75       Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
                  50                   55                   60

80       Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
          65                   70                   75                   80

5 Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
85 90 95

10 Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Ser Phe Gln Leu Asp Lys  
100 105 110

15 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
115 120 125

20 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
130 135 140

25 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Pro Ser His  
145 150 155 160

30 Tyr Gln Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
165 170 175

35 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
180 185 190

40 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
195 200 205

45 Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
210 215 220

50 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
225 230 235 240

55 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
245 250 255

60 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
260 265 270

65 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
275 280 285

70 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
290 295 300

75 Arg Leu  
305

<210> 29  
<211> 306

60

65

<212> PRT  
 <213> Artificial  
 5  
 <220>  
 <223> Aminoacido tRNA sintetico por incorporación de p-acetil-fenilalanina  
 <400> 29  
 10  
 Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15  
 15  
 Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Thr  
 20 25 30  
 20  
 Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45  
 25  
 Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60  
 30  
 Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80  
 35  
 Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95  
 40  
 Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Ser Phe Gln Leu Asp Lys  
 100 105 110  
 45  
 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125  
 50  
 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140  
 55  
 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Pro Ser His  
 145 150 155 160  
 60  
 Tyr Gln Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
 165 170 175  
 65  
 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
 180 185 190  
 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
 195 200 205  
 Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala

ES 2 632 504 T3

5

210 215 220

10 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
 225 230 235 240

15 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
 245 250 255

20 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
 260 265 270

25 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
 275 280 285

30 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
 290 295 300

35 Arg Leu  
 305

35

<210> 30  
 <211> 306  
 <212> PRT  
 <213> Artificial

40

<220>  
 <223> Aminoacido tRNA sintetico por incorporación de p-acetil-fenilalanina

45

<400> 30

50

55

60

65

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Leu  
 20 25 30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu Phe Gln Leu Asp Lys

				100					105					110			
5	Asp	Tyr	Thr	Leu	Asn	Val	Tyr	Arg	Leu	Ala	Leu	Lys	Thr	Thr	Leu	Lys	
			115					120					125				
	Arg	Ala	Arg	Arg	Ser	Met	Glu	Leu	Ile	Ala	Arg	Glu	Asp	Glu	Asn	Pro	
		130					135					140					
10	Lys	Val	Ala	Glu	Val	Ile	Tyr	Pro	Ile	Met	Gln	Val	Asn	Gly	Thr	His	
		145				150					155					160	
	Tyr	Arg	Gly	Val	Asp	Val	Ala	Val	Gly	Gly	Met	Glu	Gln	Arg	Lys	Ile	
15					165					170					175		
	His	Met	Leu	Ala	Arg	Gln	Leu	Leu	Pro	Lys	Lys	Val	Val	Cys	Ile	His	
			180						185					190			
20	Asn	Pro	Val	Leu	Thr	Gly	Leu	Asp	Gly	Glu	Gly	Lys	Met	Ser	Ser	Ser	
			195					200					205				
	Lys	Gly	Asn	Phe	Ile	Ala	Val	Asp	Asp	Ser	Pro	Glu	Glu	Ile	Arg	Ala	
25		210					215					220					
	Lys	Ile	Lys	Lys	Ala	Tyr	Cys	Pro	Ala	Gly	Val	Val	Glu	Gly	Asn	Pro	
		225				230					235					240	
30	Ile	Met	Glu	Ile	Ala	Lys	Tyr	Phe	Leu	Glu	Tyr	Pro	Leu	Thr	Ile	Lys	
					245					250					255		
	Arg	Pro	Glu	Lys	Phe	Gly	Gly	Asp	Leu	Thr	Val	Asn	Ser	Tyr	Glu	Glu	
35				260					265					270			
	Leu	Glu	Ser	Leu	Phe	Lys	Asn	Lys	Glu	Leu	His	Pro	Met	Asp	Leu	Lys	
			275					280					285				
40	Asn	Ala	Val	Ala	Gln	Gln	Leu	Ile	Lys	Ile	Leu	Glu	Pro	Ile	Arg	Lys	
		290					295					300					
	Arg	Leu															
45		305															

<210> 31  
 <211> 306  
 <212> PRT  
 <213> Artificial

<220>  
 <221> Aminoacido tRNA sintetica por incorporación de p-acetil-fenilalanina

55  
 60  
 65



<400> 31

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Ala  
 20 25 30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu Phe Gln Leu Asp Lys  
 100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Gly Gly His  
 145 150 155 160

Tyr Leu Gly Val Asp Val Ile Val Gly Gly Met Glu Gln Arg Lys Ile  
 165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
 180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
 195 200 205

Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
 210 215 220

Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
 225 230 235 240

Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
 245 250 255  
 5  
 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
 260 265 270  
 10  
 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
 275 280 285  
 15  
 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
 290 295 300  
 Arg Leu  
 305  
 20  
 <210> 32  
 <211> 306  
 <212> PRT  
 <213> Artificial  
 25  
 <220>  
 <223> Aminoacido tRNA sintetico por incorporación de p-acetil-fenilalanina  
 <400> 32  
 30  
 Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15  
 35  
 Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Ala  
 20 25 30  
 40  
 Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45  
 45  
 Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60  
 45  
 Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80  
 50  
 Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95  
 55  
 Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Arg Phe Gln Leu Asp Lys  
 100 105 110  
 60  
 65  
 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125

ES 2 632 504 T3

5 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
130 135 140

10 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Glu Val Asn Val Ile His  
145 150 155 160

15 Tyr Asp Gly Val Asp Val Ala Val Gly Gly Met Glu Glu Arg Lys Ile  
165 170 175

20 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
180 185 190

25 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
195 200 205

30 Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
210 215 220

35 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
225 230 235 240

40 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
245 250 255

45 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
260 265 270

50 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
275 280 285

55 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
290 295 300

Arg Leu  
305

50  
<210> 33  
<211> 306  
<212> PRT  
<213> Artificial

50  
<220>  
<221> Aminoacido tRNA sintético for la incorporación de p-acido-fenilalanina

55  
<400> 33

55 Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
1 5 10 15

60

65

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Gly  
 20 25 30  
 Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45  
 Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60  
 Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp  
 65 70 75 80  
 Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95  
 Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Thr Phe Gln Leu Asp Lys  
 100 105 110  
 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125  
 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140  
 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Thr Tyr Tyr  
 145 150 155 160  
 Tyr Leu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
 165 170 175  
 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
 180 185 190  
 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
 195 200 205  
 Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
 210 215 220  
 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
 225 230 235 240  
 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
 245 250 255  
 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
 260 265 270

5

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
 275 280 285

10

Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
 290 295 300

15

Arg Leu  
 305

20

<210> 34  
 <211> 213  
 <212> DNA  
 <213> Artificial

25

<220>  
 <223> Secuencia de ADN de insulina LysPro por expresión en Pichia  
 <400> 34

aagcgtgagg aggcgtgaggc tgagcgtgag octaagttcg tcaaccaaca cttgtgoggt 60  
 tctcaacttgg ttgaggccct ttactttggt ttgcggtgagc gtggtttctt ctacaaccaag 120  
 octactcgtc gtttacaaaa gctggttata gttgagcaat gctgcacctc tatctgctcc 180  
 ttgtaccaat tggagaacta ctgcaactaa tga 213

35

<210> 35  
 <211> 183  
 <212> DNA  
 <213> Artificial

40

<220>  
 <223> Secuencia de ADN de insulina LysPro por expresión en Pichia  
 <400> 35

aagcgtttcg tcaaccaaca cttgtgoggt tctcaacttgg ttgaggccct ttactttggt 60  
 tgccggtgagc gtggtttctt ctacaaccaag octactcgtc gtttacaaaa gctggttata 120  
 gttgagcaat gctgcacctc tatctgctcc ttgtaccaat tggagaacta ctgcaactaa 180  
 tga 183

55

<210> 36  
 <211> 171  
 <212> DNA  
 <213> Artificial

60

<220>  
 <223> Secuencia de ADN de insulina LysPro por expresión en Pichia  
 <400> 36

65

ES 2 632 504 T3

5        aagcgtttcg tcaaccaaca cttgtgcggt tctcacttgg ttgaggccct ttacttggtt        60  
       tgcggtgagc gtggtttcct ctacaccaag cctactaagc gtggtatcgt tgagcaatgc        120  
 10        tgcacctcta tctgctcctt gtaccaattg gagaactact gcaactaatg a                171

<210> 37  
 <211> 201  
 <212> DNA  
 <213> Artificial

15

<220>  
 <223>        Secuencia de ADN de insulina LysPro por expresión en Pichia

20

<400> 37

25        aagcgtgagg aggctgaggc tgaggctgag cctaagttcg tcaaccaaca cttgtgcggt        60  
       tctcacttgg ttgaggccct ttacttggtt tgcggtgagc gtggtttcct ctacaccaag        120  
       cctactaagc gtggtatcgt tgagcaatgc tgcacctcta tctgctcctt gtaccaattg        180  
       gagaactact gcaactaatg a                201

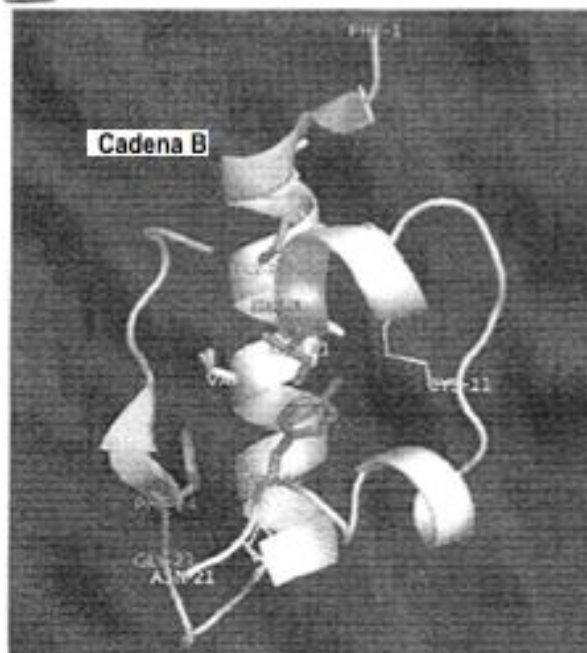
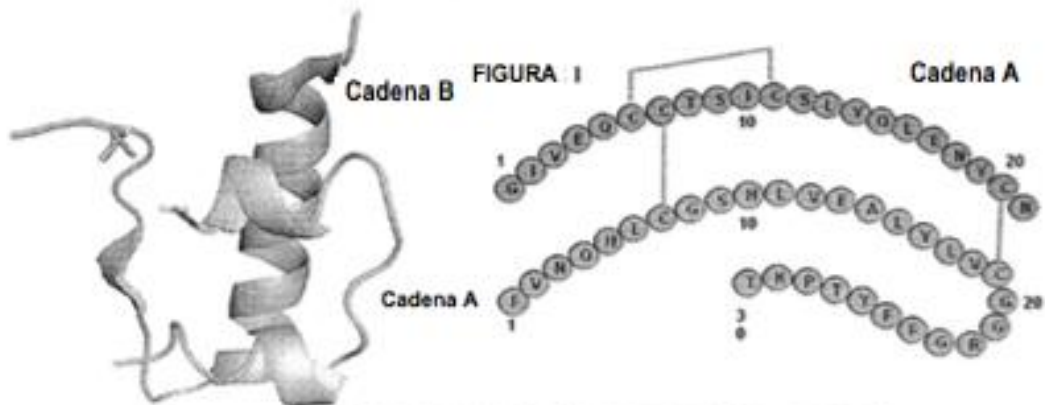
30

**Reivindicaciones**

- 5 1. Un polipéptido de insulina que comprende un polímero soluble en agua unido por un enlace covalente de oxima con el cadena A de polipéptido de insulina en la posición 14 de la secuencia de aminoácidos de cadena A seleccionado del grupo SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, y SEQ ID NO: 11, en el que el aminoácido en la posición 14 de la cadena A es un aminoácido no naturalmente codificado.
- 10 2. El polipéptido de insulina de la reivindicación 1 en la que el polímero soluble en agua es una molécula de poli(etilenglicol).
- 15 3. El polipéptido de insulina de la reivindicación 2 en el que la molécula de poli(etilenglicol) tiene un peso molecular entre 1 y 100 kDa.
4. El polipéptido de insulina de la reivindicación 2 en el que la molécula de poli(etilenglicol) es ramificado o es lineal.
5. El polipéptido de insulina de la reivindicación 1 en el que el polímero soluble en agua es una molécula lineal poli(etilenglicol) que tiene un peso molecular de 20 kDa.

20

Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS	
Solicitante:	AMBRX, INC	
Nº de Sol.	No asignado	
Ab:	John W. Wallen, III	Teléfono: (858) 875-2403
Nº de expediente:	AMBX-0146.00PCT	



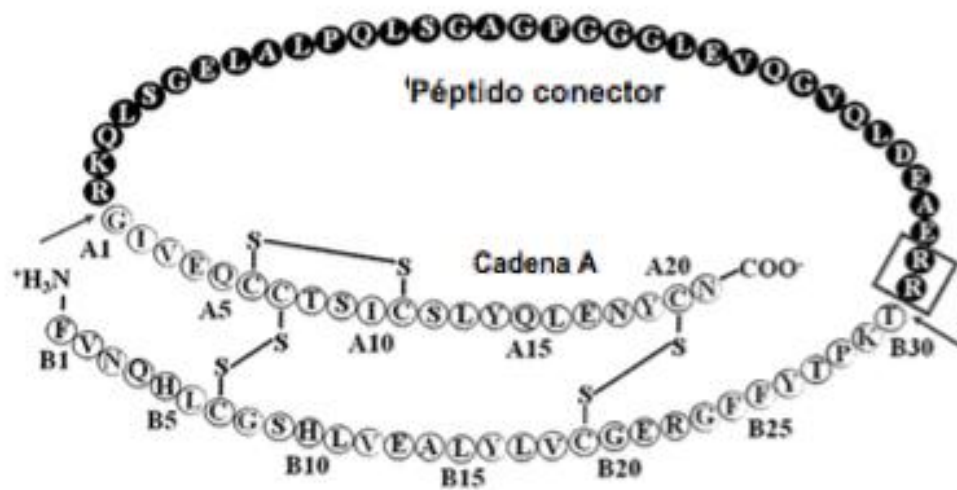




Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS		
Solicitante:	AMBRX, INC		
Nº de Scl.	No asignado		
Ab:	John W. Wallen, III	Teléfono: (858) 675-2403	
Nº de expediente	AMBX-0146.00PCT		

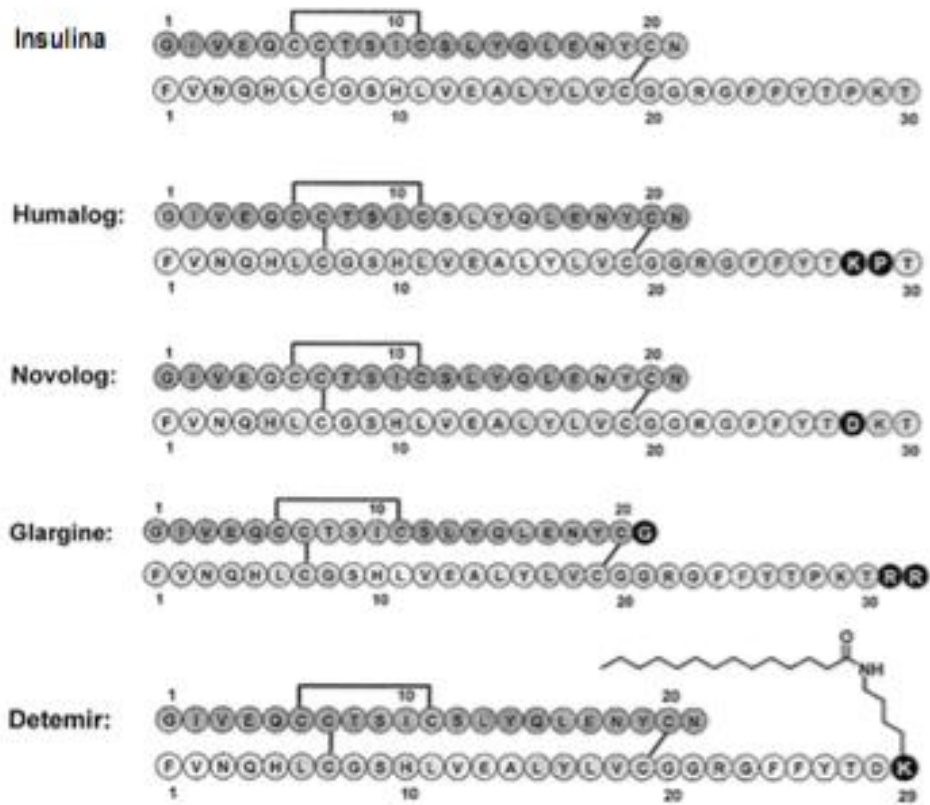
FIGURA 3

ESTRUCTURA DE PROINSULINA PRIMARIA



Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS	
Solicitante:	AMBRX, INC	
Nº de Sol.:	No asignado	
Ab:	John W. Wallen, III	Teléfono: (858) 875-2403
Nº de expediente:	AMBX-0146.00PCT	

FIGURA 4



Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS		
Solicitante:	AMBRX, INC		
Nº de Sol.:	No asignado		
Ab:	John W. Wallen, III	Teléfono: (858) 875-2403	
Nº de expediente:	AMBX-0146.00PCT		

FIGURA 5

Expresión pro-Insulina y supresión de TAG en *E.coli*

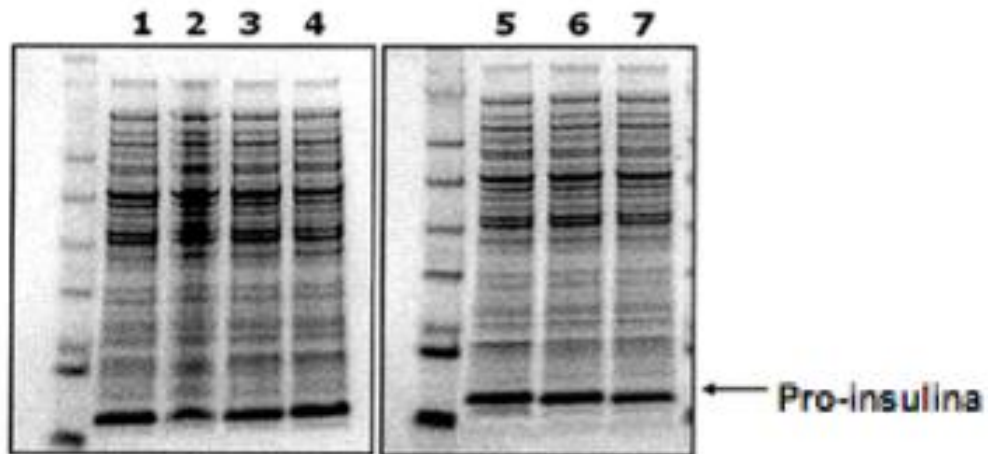


```

atgattgaagggtgggtogtcttctgtaaccagcaccctgtgagggtccccacctggaggagct
N I E G G R F V N Q H L C G S N L V E A
ctgtacctgggtgtgagggtgaaggtgggtctctctacacccaagcagcccccggcgtgagga
L Y L V C G E R G F F Y T K P T R R E A
gaggsacctgcaggtgggtcaggtggagctggggcgtggcccgggtgagggcagcctgcag
E D L Q V G Q V E L G G G P G A G S L Q
ccgctggccctggagggttccctgcagagcaggtggccttctggaccactgctgtaccagc
P L A L E G S L Q R G I V E Q C C T S
atctgctccctgtaccagctggaaacctctggagc
I C S L Y Q L E N Y C G
    
```

Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS	
Solicitante:	AMBRX, INC	
Nº de Sol.	No asignado	
Ab:	John W. Wallen, III	Teléfono: (858) 875-2403
Nº de expediente	AMBX-0146.00PCT	

FIGURA 6



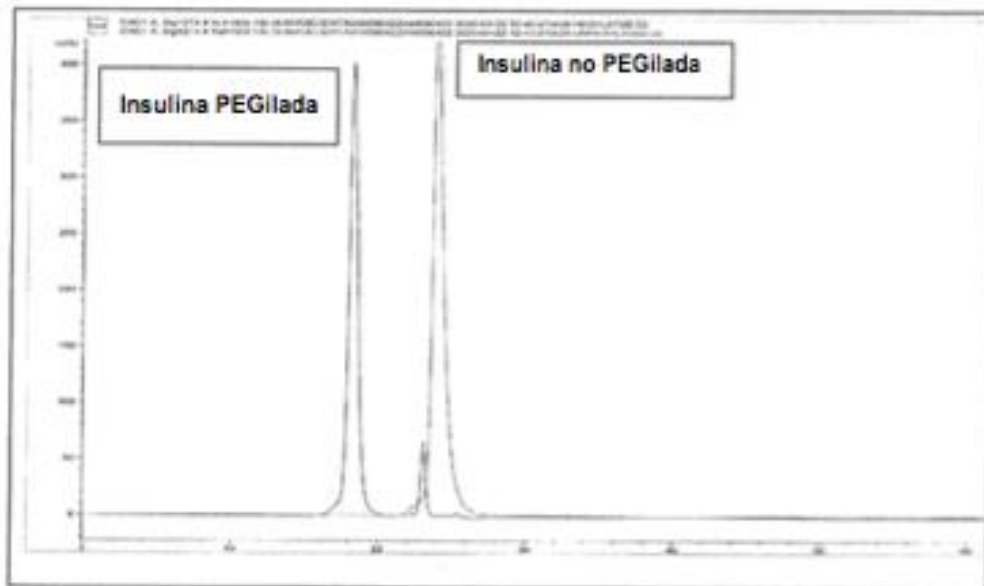
1. Pro-INS-A(Q15pAF)
2. Pro-INS-A(Y14pAF)
3. Pro-INS-B(R22pAF)
4. Pro-INS-B(F1pAF)
5. Pro-INS-A(G1pAF)
6. Pro-INS-A(S9pAF)
7. Pro-INS-B(K28pAF)

Muestras totalmente seguras de cultivos de matraz de agitación LB, 5 h posinducción (células W3110-B2, transformadas con derivados pVK6-LisPro e inducidas con 0.2% L-arabinosa a 37 °C)

Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS		
Solicitante:	AMBRX, INC		
Nº de Sol.	No asignado		
Ab:	John W. Wallen, III	Teléfono:	(858) 875-2403
Nº de expediente	AMBX-0146.00PCT		

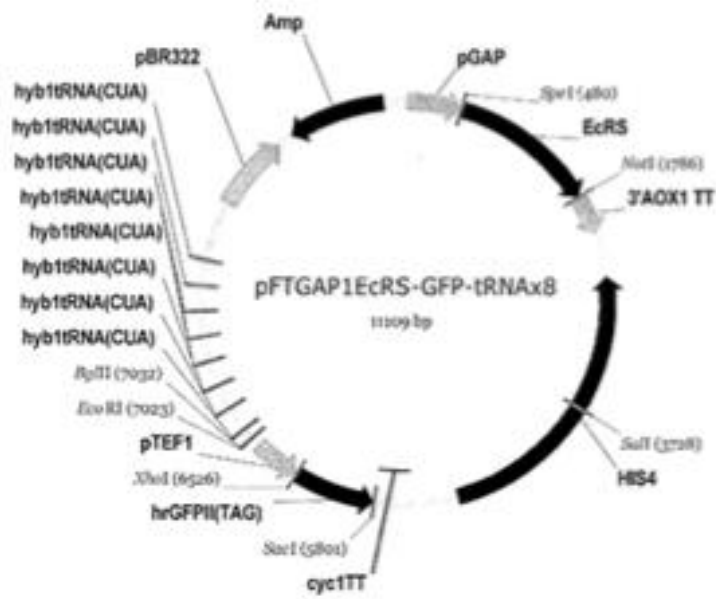
FIGURA 7

SEC-HPLC de INS pegilado y no pegilado



Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS	
Solicitante:	AMBRX, INC	
Nº de Sol.	No asignado	
Ab:	John W. Wallen, III	Teléfono: (858) 875-2403
Nº de expediente	AMBX-0146.00PCT	

FIGURA 8



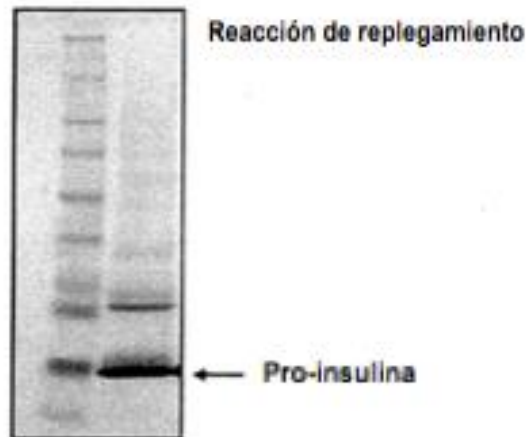




Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS	
Solicitante:	AMBRX, INC	
Nº de Sol.	No asignado	
Ab:	John W. Wallen, III	Teléfono: (858) 875-2403
Nº de expediente	AMBX-0146.00PCT	

**FIGURA 10**

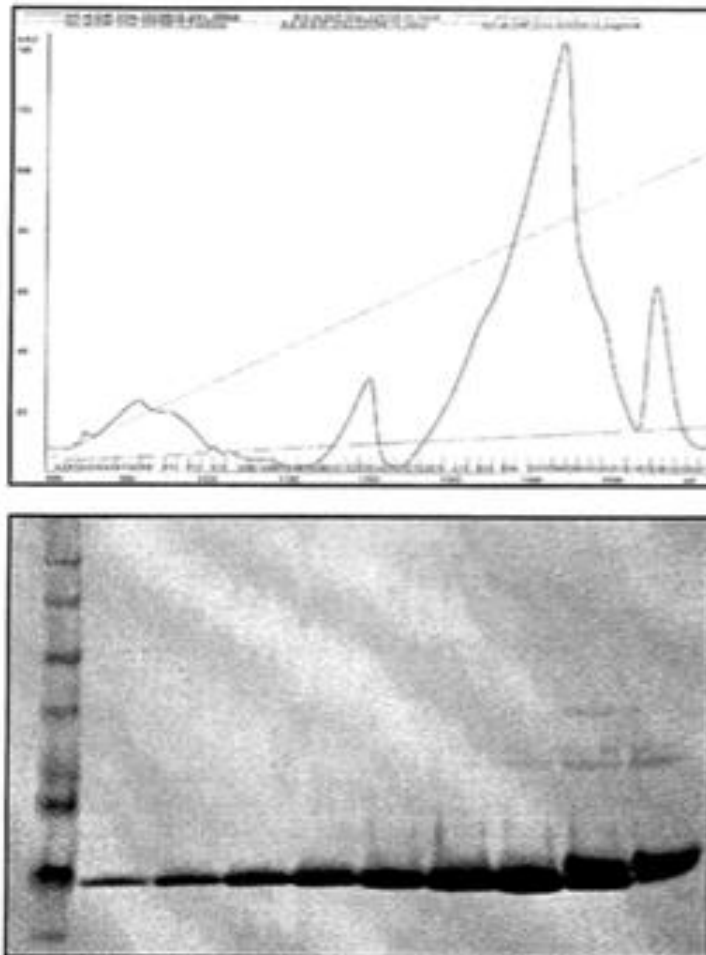
**Reacción de replegamiento de insulina A21G**



Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS		
Solicitante:	AMBRX, INC		
Nº de Sol.	No asignado		
Ab:	John W. Wallen, III	Teléfono:	(858) 875-2403
Nº de expediente	AMBX-0146.00PCT		

FIGURA 11

Purificación QHP de proinsulina replegada



Título:	POLIPEPTIDOS DE INSULINA MODIFICADOS Y SUS USOS	
Solicitante:	AMBRX, INC	
Nº de Sol.:	No asignado	
Ab:	John W. Wallen, III	Teléfono: (858) 875-2403
Nº de expediente:	AMBX-0146.00PCT	

FIGURA 12

Reacción de PEGilación y purificación

