

PRECURSOR DE INSULINA RECOMBINANTE

Campo de la invención

La invención se refiere al campo de los precursores de insulina recombinante para maximizar la
5 expresión, el rendimiento y la pureza final de un precursor de insulina recombinante.

Antecedentes

La insulina forma parte de la superfamilia de la insulina, que consiste en insulinas, polipéptidos
relacionados con la insulina y el factor de crecimiento similar a la insulina. Estos polipéptidos no
10 solo se encuentran en vertebrados, sino también en hongos y protistas. De hecho, los orígenes
moleculares se remontan, como mínimo, a los eucariotas unicelulares más simples.

La insulina madura que se encuentra en los mamíferos es un heterodímero que consiste en dos
cadenas polipeptídicas, una cadena A y una cadena B, que están unidas entre sí por dos enlaces
disulfuro intermoleculares en la insulina humana. Los enlaces disulfuro intermoleculares se
15 forman entre el residuo de aminoácido 7 de la cadena A y el residuo de aminoácido 7 de la
cadena B, y entre el residuo de aminoácido 20 de la cadena A y el residuo de aminoácido 19 de la
cadena B. Un tercer enlace disulfuro intramolecular se forma dentro de la cadena A. En la
insulina humana, este enlace disulfuro intramolecular se forma entre los residuos de aminoácidos
6 y 11 de la cadena A. La insulina tiene una masa molecular de aproximadamente 5900 Da y
20 consiste en 51-53 aminoácidos en total, con variaciones menores entre diferentes especies
animales. Por ejemplo, la cadena A y la cadena B de la insulina humana consisten en 21 y 30
aminoácidos, respectivamente. La insulina porcina es casi idéntica a la insulina humana, siendo
la única diferencia el residuo 30 de la cadena B, en donde la insulina humana contiene una
treonina y la insulina porcina contiene una alanina.

25 En la mayoría de los vertebrados, la insulina se produce como un precursor de insulina inactivo
denominado preproinsulina en las células beta de los islotes de Langerhans del páncreas. La
preproinsulina es el producto traduccional inicial del ARNm de la insulina. Como se muestra en
la Fig. 1, la preproinsulina comprende un péptido señal N-terminal o péptido líder (péptido L), y
la cadena A y la cadena B unidas entre sí por una cadena C o péptido conector (péptido C). Si
30 bien la cadena A y la cadena B están altamente conservadas entre especies, el péptido C muestra
mayor variación de secuencia y puede variar en longitud, p. ej., entre 26 y 38 residuos. La
preproinsulina se transporta cotraduccionalmente al retículo endoplasmático, donde el péptido L
se elimina para producir el precursor de insulina inactivo denominado proinsulina. Durante el
plegamiento, se forman los tres enlaces disulfuro y se elimina el péptido C, dando como

resultado la insulina madura. La insulina madura se empaca luego en gránulos como hexámeros y puede secretarse ante la estimulación de la célula beta.

La secreción de insulina ocurre cuando los niveles de glucosa en sangre aumentan y las células beta captan y metabolizan grandes cantidades de glucosa. También puede estimularse por ácidos grasos, aminoácidos, hormonas y cetoácidos secretados por el tracto gastrointestinal. Por otro lado, la secreción de insulina se inhibe cuando los niveles de glucosa en sangre disminuyen, por la somatostatina y por la activación simpática del sistema nervioso. Sin embargo, pequeñas cantidades de insulina se secretan de manera continua, incluso en ayuno. La insulina secretada se une a receptores específicos en las membranas externas de las células diana del hígado, el músculo y el tejido adiposo, lo que provoca la translocación de los transportadores de glucosa hacia la membrana celular desde el interior de la célula, permitiendo que la glucosa sea captada por la célula. La insulina también promueve la lipogénesis, la glucogénesis y la síntesis de proteínas en el músculo esquelético y los tejidos grasos a través de la vía del receptor tirosina quinasa.

Cuando la insulina no regula o ya no puede regular correctamente los niveles de glucosa en sangre, puede desarrollarse diabetes mellitus (o diabetes, en forma abreviada). La diabetes tipo I se caracteriza por la destrucción de las células beta. Por lo tanto, se produce menos insulina o no se produce, lo que lleva a una deficiencia de insulina. En la mayoría de los casos, la diabetes tipo I se debe a una respuesta autoinmune y a un ataque mediado por células T de las células beta. La diabetes tipo II se caracteriza por resistencia a la insulina, que puede combinarse con una secreción de insulina relativamente reducida. El tipo II es el tipo más común de diabetes y se debe principalmente a factores del estilo de vida y a la genética. El tratamiento de la diabetes tipo II está dirigido, por lo tanto, a un cambio en el estilo de vida, la incorporación de medicamentos destinados a mejorar la sensibilidad a la insulina o a reducir la producción de glucosa por el hígado, y a la administración de insulina. Por otro lado, la diabetes tipo I es principalmente hereditaria y se trata mediante la administración de insulina.

Ya en 1910 se propuso que la diabetes se debía a una hormona que faltaba en el páncreas. En 1921, se identificó que esta hormona faltante era la insulina y, desde entonces, la insulina se ha utilizado en el tratamiento de la diabetes. Inicialmente, se utilizó insulina aislada de cerdos y vacas, pero la invención de la clonación de ADN condujo a la primera insulina recombinante producida en bacterias *Escherichia coli* en 1978.

Hoy en día, *E. coli* sigue siendo un sistema de expresión ampliamente utilizado para la producción de insulina recombinante. En general, las bacterias se transforman con un vector que codifica para una molécula precursora inactiva de insulina y las células se cultivan en

- fermentadores. Tras la inducción, el precursor se expresa y típicamente se acumula dentro de las bacterias como cuerpos de inclusión insolubles. Tras la recolección de las células, estas se lisan para obtener los cuerpos de inclusión que comprenden predominantemente el precursor. Los cuerpos de inclusión se aíslan y solubilizan posteriormente, y el precursor se repliega en su
- 5 estructura nativa, donde se forman los enlaces disulfuro. A continuación, el precursor plegado se purifica para eliminar las impurezas de la célula huésped y las formas mal plegadas de la molécula. Finalmente, el precursor se digiere con tripsina y/o carboxipeptidasa B para escindir el péptido L N-terminal y el péptido C, dando como resultado la insulina recombinante madura. Como parte del proceso de fabricación recombinante, se necesitarán pasos adicionales de
- 10 purificación de la insulina madura para alcanzar la calidad de producto deseada. El número de personas diagnosticadas con diabetes aumentó de 108 millones en 1980 a 537 millones en 2021 en todo el mundo, y se espera que este número continúe aumentando. Además, los animales pueden padecer diabetes, siendo más común en perros y gatos. Al igual que en los humanos, la prevalencia de la diabetes está aumentando en los animales. Para satisfacer la
- 15 demanda de insulina para el tratamiento tanto en humanos como en animales, el proceso de producción de insulina recombinante se mejora continuamente para reducir los costos, aumentar la efectividad y aumentar los rendimientos. WO 2001/049742 describe la síntesis en *Saccharomyces cerevisiae* de un precursor de insulina humana que comprende péptidos C que contienen al menos un residuo de aminoácido aromático
- 20 Phe, Trp o Tyr y que en general no son más largos de 15 aminoácidos. Estos precursores resultaron en un mayor rendimiento de producción en comparación con los precursores que tienen un péptido C nativo. WO 2001/025278 describe precursores de insulina humana que comprenden péptidos C de entre 2 y 50 aminoácidos de longitud que mejoran el rendimiento de producción del precursor.
- 25 Min et al., 2011, J. Biotechnol. 151:350-356 describe una expresión, plegamiento y tasa de reacción enzimática mejorados de la insulina humana recombinante cuando se seleccionaron ciertos péptidos L para el precursor de insulina humana. WO 2017/040363 describe un proceso para obtener insulina con enlaces disulfuro correctamente formados.
- 30 Sin embargo, aún existe la necesidad de mejorar aún más la producción recombinante de insulina.

Breve descripción de la invención

En un primer aspecto de la invención, se proporciona un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde:

A es una cadena A de insulina,

B es una cadena B de insulina,

5 C es un péptido C que conecta la cadena A y la cadena B, y

L es un péptido L N-terminal,

en donde C es un péptido C que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 y/o en donde L es un péptido L que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2.

10 Otro aspecto de la invención es una secuencia de ADN que codifica para el precursor de insulina recombinante de acuerdo con la invención y/o sus modalidades, preferiblemente en donde el péptido C está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 9 y/o en donde el péptido L está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 10.

15 Se encontró que el precursor de insulina recombinante de acuerdo con la invención proporciona un rendimiento deseable, tal como un rendimiento mejorado de insulina.

El rendimiento (p. ej., expresado como título de fermentación, p. ej., en mmol/l o g/l) de uno o más precursores de insulina recombinante de acuerdo con la invención se comparó con el rendimiento de un precursor de insulina recombinante de referencia (véase, la SEQ ID NO: 5). El

20 precursor de insulina recombinante de referencia que consiste en la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 5 es sustancialmente de acuerdo con un precursor de insulina descrito en WO 2017/040363 (véase, la Figura 2 de WO 2017/040363). Una diferencia entre la SEQ ID NO: 5 y el precursor de insulina mostrado en WO 2017/040363 comprende que el precursor de insulina recombinante de referencia de SEQ ID NO: 5 comprende un residuo de metionina (M) adicional en la posición N-terminal en comparación con el precursor divulgado en WO 2017/040363. Este residuo de metionina se escinde durante la fermentación en *Escherichia coli* por la metionina aminopeptidasa. WO 2017/040363 describe un proceso para obtener insulina con enlaces disulfuro correctamente formados utilizando un precursor de insulina. La escisión del residuo de metionina adicional en la posición N-terminal del precursor de insulina de

25 WO2017/040363 ya se completa antes de iniciar el proceso descrito en WO2017/040363.

30 Además, el precursor de insulina recombinante de referencia de acuerdo con la SEQ ID NO: 5 comprende en el residuo B30 de la cadena B una alanina (A), mientras que la secuencia de aminoácidos de la Figura 2 de WO 2017/040363 comprende una treonina (T) en el residuo B30 de la cadena B.

El precursor de insulina recombinante que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 6 es un precursor de insulina recombinante de acuerdo con la invención y/o sus modalidades, y contiene la cadena A y la cadena B de la insulina, que comprenden una secuencia de aminoácidos de acuerdo con las SEQ ID NO: 3 y 4, respectivamente, y el péptido L de referencia del precursor de insulina recombinante de referencia (SEQ ID NO: 5), que es sustancialmente similar al péptido L descrito en WO2017/040363; SEQ ID NO: 5 incluye, además, un residuo de metionina N-terminal. A diferencia del precursor de insulina recombinante de referencia (SEQ ID NO: 5), el precursor de insulina recombinante de acuerdo con la invención de la SEQ ID NO: 6 contiene un nuevo péptido C recombinante de acuerdo con la SEQ ID NO:

1. Sorprendentemente, el título de fermentación (en mmol/l) del precursor de insulina de acuerdo con la SEQ ID NO: 6 fue mayor que el título de fermentación (en mmol/l) del precursor de insulina recombinante de referencia de acuerdo con la SEQ ID NO: 5. Como se muestra en la Tabla 1 del Ejemplo 1, el título de fermentación (en mmol/l) del precursor de insulina de acuerdo con la SEQ ID NO: 6 fue aproximadamente un 43 % mayor que el título de fermentación (en mmol/l) del precursor de insulina de referencia de acuerdo con la SEQ ID NO: 5 cuando los precursores se expresaron en células huésped en las mismas condiciones de fermentación. Por lo tanto, reemplazando el péptido C humano nativo del precursor de referencia por el péptido C de acuerdo con la SEQ ID NO: 1 aumenta significativamente el título de fermentación del precursor. Otro precursor de insulina recombinante de acuerdo con la invención y/o sus modalidades es un precursor de insulina recombinante que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 7. Dicho precursor de insulina recombinante contiene la cadena A y la cadena B de la insulina, que comprenden una secuencia de aminoácidos de acuerdo con las SEQ ID NO: 3 y 4, respectivamente, en combinación con un nuevo péptido L recombinante de acuerdo con la SEQ ID NO: 2, y un péptido C modificado derivado de la secuencia de aminoácidos del péptido C humano. El título de fermentación (en mmol/l) del precursor de insulina de acuerdo con la SEQ ID NO: 7 fue mayor que el título de fermentación (en mmol/l) del precursor de insulina recombinante de referencia (SEQ ID NO: 5). Como se muestra en la Tabla 1 del Ejemplo 1, el título de fermentación (en mmol/l) del precursor de insulina de acuerdo con la SEQ ID NO: 7 fue aproximadamente un 39 % mayor que el título de fermentación (en mmol/l) del precursor de insulina recombinante de referencia (SEQ ID NO: 5). Por lo tanto, reemplazando el péptido L humano nativo del precursor de referencia por el péptido L de acuerdo con la SEQ ID NO: 2 aumenta el título de fermentación del precursor de insulina.

Otro precursor de insulina recombinante de acuerdo con la invención y/o sus modalidades es un precursor de insulina recombinante que comprende una secuencia de aminoácidos de acuerdo con

- la SEQ ID NO: 8 y contiene la cadena A y la cadena B de insulina que comprenden una secuencia de aminoácidos de acuerdo con las SEQ ID NO: 3 y 4, respectivamente, y el péptido C de acuerdo con la SEQ ID NO: 1. Además, la SEQ ID NO: 8 contiene un péptido L de acuerdo con la SEQ ID NO: 2. El título de fermentación (en mmol/l) del precursor de insulina de acuerdo con la SEQ ID NO: 8 fue incluso mayor que el título de fermentación (en mmol/l) del precursor de insulina de acuerdo con la SEQ ID NO: 5, SEQ ID NO: 6 o SEQ ID NO: 7. Como se muestra en la Tabla 1 del Ejemplo 1, el título de fermentación (en mmol/l) del precursor de insulina de acuerdo con la SEQ ID NO: 8 fue incluso aproximadamente un 118 % mayor que el título de fermentación (en mmol/l) del precursor de insulina recombinante de referencia (SEQ ID NO: 5).
- Por lo tanto, reemplazando el péptido C humano nativo por el péptido C de acuerdo con la SEQ ID NO: 1 y adicionalmente el péptido L de referencia por el péptido L de acuerdo con la SEQ ID NO: 2 aumenta el título de fermentación del precursor de insulina de manera sinérgica.
- Un aspecto adicional de la invención se refiere a un vector que comprende una secuencia de ADN que codifica un precursor de insulina recombinante de acuerdo con la invención.
- Un aspecto adicional de la invención se refiere a una célula huésped que comprende (un vector que comprende) una secuencia de ADN que codifica un precursor de insulina recombinante de acuerdo con la invención.
- Un aspecto adicional se refiere al método para producir insulina recombinante, en donde se expresa un precursor de insulina recombinante de acuerdo con la invención, en donde el método comprende:
- a) expresar el precursor de insulina recombinante mediante fermentación de un microorganismo,
 - b) lisar el microorganismo para obtener el precursor de insulina recombinante,
 - c) replegar el precursor de insulina recombinante, y
 - d) escindir el péptido L y el péptido C del precursor de insulina recombinante mediante una reacción enzimática para obtener insulina recombinante.
- Un aspecto adicional se refiere al uso de un precursor de insulina recombinante de acuerdo con la invención para la producción de insulina recombinante.
- Un aspecto final se refiere a una composición que comprende insulina recombinante, en donde la composición comprende una cantidad detectable de un péptido C y/o péptido L como se proporciona en el presente documento.

Descripción de las figuras

La Fig. 1 muestra un diagrama de una molécula de preproinsulina, una molécula de proinsulina y una molécula de insulina.

La Fig. 2 muestra un perfil de cromatografía líquida-espectrometría de masas de una molécula de insulina madura resultante de la expresión de un constructo de ADN que codifica residuos de asparagina mediante el codón “AAC” en *E. coli*.

La Fig. 3 muestra un perfil de cromatografía líquida-espectrometría de masas de una molécula de insulina madura resultante de la expresión de un constructo de ADN que codifica residuos de serina mediante el codón “TCT” en *E. coli*.

10 Definiciones

A los efectos de la presente invención, los siguientes términos se definen a continuación.

El término “Péptido” se refiere a una secuencia corta de aminoácidos, típicamente de 2-50 aminoácidos de longitud.

El término “péptido C” se refiere al péptido conector o fracción conectora “C” de la secuencia B-C-A de una molécula precursora de insulina. El péptido C conecta el extremo C-terminal de la cadena B y el extremo N-terminal de la cadena A de la insulina.

El término “péptido L” se refiere al péptido líder que está presente como una secuencia líder N-terminal en la forma precursora de la proteína insulina. El péptido L de la insulina de origen natural permite que la proteína se transloque al retículo endoplasmático. El péptido L se escinde normalmente durante este proceso de translocación.

El término “Precursor” se refiere a una forma inmadura de una proteína, que aún debe procesarse antes de obtener la forma madura final de la proteína. En el caso de la insulina, el precursor aún comprende el péptido C y/o el péptido L. El péptido C y el péptido L deben escindirse para producir insulina madura.

El término “Proinsulina” se refiere a una forma inmadura de insulina que consiste en un péptido de cadena B, que está fusionado al extremo N-terminal de un péptido C que a su vez está fusionado en su extremo C-terminal al extremo N-terminal de un péptido de cadena A. La proinsulina es un tipo de precursor de insulina.

El término “Preproinsulina” se refiere a una forma inmadura de insulina que consiste en un péptido L N-terminal que está fusionado en su extremo C-terminal al extremo N-terminal de un péptido de cadena B, que está fusionado al extremo N-terminal de un péptido C que a su vez está fusionado en su extremo C-terminal al extremo N-terminal de un péptido de cadena A. La preproinsulina se representa, por lo tanto, mediante la fórmula L-B-C-A. La preproinsulina es un tipo de precursor de insulina.

El término “vector” en el presente documento se refiere a un constructo utilizado para introducir un ácido nucleico, p. ej., ADN, en una célula huésped. Después de que el vector se introduce en la célula huésped, el ácido nucleico, p. ej., ADN, transportado por el vector puede entonces ser expresado por la célula huésped, dando como resultado la producción de una proteína codificada por el ácido nucleico, p. ej., ADN. El vector puede ser, por ejemplo, un virus, pero típicamente es un plásmido, tal como un plásmido pET o un plásmido pGEX.

El término “célula huésped” en el presente documento se refiere a una célula utilizada para expresar un precursor de insulina recombinante. Típicamente, la célula huésped es un microorganismo tal como una célula de levadura o una bacteria. Las células huésped adecuadas son *Escherichia coli* (*E. coli*) y *Saccharomyces cerevisiae* (*S. cerevisiae*).

El término “cultivo de células huésped” se refiere a una población de células huésped cultivadas bajo condiciones controladas, tales como en un medio que comprende uno o más nutrientes y que tiene un pH adecuado para la célula huésped particular y en una temperatura y una concentración de O₂ y/o CO₂ apropiada para la propagación de la célula huésped particular.

El término “título de fermentación” se refiere a la concentración de una molécula de interés en una solución. En el presente documento, el título de fermentación es la concentración del precursor de insulina recombinante en el medio utilizado para cultivar la célula huésped que expresa el precursor. El título de fermentación puede expresarse, por ejemplo, en g/ml o mmol/l (mM) y puede utilizarse, p. ej., como medida del rendimiento de una proteína.

Como se utiliza en el presente documento, una secuencia de aminoácidos es 100 % “idéntica” o tiene 100 % de “identidad de secuencia” con una segunda secuencia de aminoácidos cuando los residuos de aminoácidos de ambas secuencias, al alinearse, son idénticos, cuando se determina utilizando el programa informático “BLAST®” seleccionando el subprograma: “Global Alignment” usando la configuración estándar, que puede encontrarse en

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>. En consecuencia, una secuencia de aminoácidos (o secuencia de ácido nucleico) puede ser 50 % “idéntica” a una segunda secuencia de aminoácidos (o segunda secuencia de ácido nucleico) cuando el 50 % de los residuos de aminoácidos de las dos secuencias de aminoácidos (o ácidos nucleicos de las dos secuencias de ácido nucleico) son idénticos cuando se determina utilizando el programa informático “BLAST®” seleccionando el subprograma: “Global Alignment” usando la configuración estándar. La comparación de secuencias se realiza sobre un bloque contiguo de residuos de aminoácidos comprendido por una proteína dada, p. ej., una proteína, o una porción del polipéptido que se compara. En una modalidad particular, también se contemplan sustituciones, deleciones o inserciones seleccionadas que alteran o no alteran la correspondencia entre las dos secuencias de

aminoácidos. De manera similar, una secuencia de ácido nucleico es 100 % “idéntica” o tiene 100 % de “identidad de secuencia” con una segunda secuencia de ácido nucleico cuando los residuos de ácido nucleico de ambas secuencias, al alinearse, son idénticos, cuando se determina utilizando el programa informático “BLAST®” seleccionando el subprograma: “Global Alignment”, que puede encontrarse en <https://blast.ncbi.nlm.nih.gov/Blast.cgi> y usando la configuración estándar.

Descripción detallada de la invención

La invención se define en el presente documento y en las reivindicaciones adjuntas. La materia que no está comprendida dentro del alcance de las reivindicaciones no forma parte de la presente invención reivindicada.

Se contempla que cualquier producto, método, uso o composición descrito en el presente documento puede implementarse con respecto a cualquier otro producto, método, uso o composición descrito en el presente documento. Las modalidades divulgadas en el contexto de productos, métodos, usos o composiciones de la invención pueden emplearse con respecto a cualquier otro producto, método, uso o composición descrito en el presente documento. Por lo tanto, una modalidad relativa a un producto, método, uso o composición puede aplicarse a otros productos, métodos, usos o composiciones de la invención también.

Secuencias de aminoácidos

La invención se refiere a un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A. En el presente documento, L representa un péptido L N-terminal, B representa una cadena B de insulina, C representa un péptido C que conecta la cadena A y la cadena B y A representa una cadena A de insulina.

C es un péptido conector (o péptido C) que conecta el extremo N-terminal de la cadena A y el extremo C-terminal de la cadena B de la insulina. Si bien la cadena A y la cadena B están ampliamente conservadas entre especies, el péptido C muestra mayor variación tanto en secuencia como en longitud. El péptido C humano nativo tiene, por ejemplo, 31 residuos de longitud, mientras que el péptido C porcino nativo tiene 29 residuos de longitud y el péptido C bovino nativo tiene 26 residuos de longitud.

Para apoyar el proceso de replegamiento del precursor de insulina recombinante, el péptido C del precursor de insulina recombinante comprende 2-30 residuos de aminoácidos, preferiblemente 8-20 residuos, más preferiblemente 10-15 residuos de aminoácidos.

La tripsina o las enzimas similares a la tripsina catalizan la escisión de los enlaces amida de proteínas en el lado C-terminal de los residuos de arginina y lisina y se utilizan en este proceso

para la conversión de la proteína precursora correctamente plegada en insulina o análogo de insulina correctamente plegado. Para facilitar la escisión del péptido C utilizando tripsina después de la expresión de un precursor de insulina recombinante en un microorganismo, se prefiere que el residuo C-terminal del péptido C sea un residuo de arginina. En una modalidad preferida, el péptido C de acuerdo con la invención consiste en 30 residuos de aminoácidos (péptido C modificado, SEQ ID NO: 7 y Tabla 2) o 11 aminoácidos (péptido C recombinante novedoso; SEQ ID NO: 1).

Para facilitar la escisión del péptido C, p. ej., utilizando tripsina, después de la expresión de un precursor de insulina recombinante en un microorganismo, se prefiere que el residuo C-terminal del péptido B y/o péptido C comprenda un residuo de aminoácido adecuado para permitir la escisión del péptido, tal como uno o más residuos de arginina.

Por lo tanto, en una modalidad preferida el péptido C es un péptido C que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1, o un péptido C que consiste en una secuencia de aminoácidos que tiene al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 % de identidad de secuencia con la SEQ ID NO: 1. En una modalidad más preferida, el péptido C consiste en una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1. Se encontró que un precursor de insulina recombinante que comprende el péptido C como se proporciona en el presente documento resultó en títulos de fermentación deseables del precursor de insulina.

L es un péptido líder N-terminal (o péptido L) conectado al extremo N-terminal de la cadena B. El péptido L en la insulina humana de origen natural tiene una longitud de 24 residuos. Cuando la insulina se expresa en una célula de mamífero, el péptido L actúa como un péptido señal para transportar la preproinsulina al retículo endoplasmático, donde el péptido L se escinde. Sin embargo, cuando la insulina recombinante se expresa en un microorganismo, el péptido L sirve para proteger la insulina recombinante de la degradación interna y/o modificación de la insulina recombinante durante la síntesis en el microorganismo, y para mejorar la expresión de proteínas del precursor de insulina recombinante en el microorganismo. Cuando la insulina recombinante se expresa en un microorganismo, el péptido L se escinde después de la expresión del precursor de insulina recombinante. Para facilitar la escisión del péptido L, p. ej., utilizando tripsina o un agente similar a la tripsina, después de la expresión de un precursor de insulina recombinante en un microorganismo, se prefiere que el residuo C-terminal del péptido L comprenda un residuo de aminoácido adecuado para permitir la escisión del péptido, tal como uno o más residuos de arginina.

Preferiblemente, el péptido L comprende 10-25 residuos de aminoácidos. Más preferiblemente, el péptido L comprende 12-24 residuos de aminoácidos, y aún más preferiblemente comprende 15-20 residuos. En una modalidad preferida, el péptido L de acuerdo con la invención consiste en 16 residuos de aminoácidos (péptido L recombinante novedoso; SEQ ID NO: 2).

- 5 Para obtener un título de fermentación relativamente alto del precursor de insulina recombinante tras la expresión del precursor de insulina recombinante en un microorganismo, el péptido L proporcionado en el presente documento tiene preferiblemente una carga neta positiva a pH 6,0. En otras palabras, el péptido L tiene un pI > 6,0. Preferiblemente, el péptido L tiene un pI > 6,5, y más preferiblemente un pI > 7,0. Los aminoácidos lisina, arginina e histidina tienen un pI de >
- 10 7,5. Por lo tanto, preferiblemente el péptido L al menos 40 %, al menos 45 %, al menos 50 %, al menos 55 %, al menos 60 %, al menos 65 % o al menos 70 % de los residuos del péptido L se eligen entre lisina, arginina e histidina. Por ejemplo, el péptido L contiene 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 residuos elegidos entre lisina, arginina e histidina.

- Por lo tanto, en una modalidad preferida el péptido L es un péptido L que comprende una
- 15 secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2, o un péptido L que consiste en una secuencia de aminoácidos que tiene al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 % de identidad de secuencia con la SEQ ID NO: 2. En una modalidad más preferida, el péptido L consiste en una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2. Se encontró que un precursor de insulina
- 20 recombinante que comprende el péptido L como se proporciona en el presente documento resultó en títulos de fermentación deseables del precursor de insulina.

- Se encontró, además, que un precursor de insulina recombinante que tiene una secuencia de aminoácidos de la SEQ ID NO: 8 resultó en títulos de fermentación particularmente deseables del precursor de insulina. Por lo tanto, en una modalidad, el precursor de insulina recombinante de
- 25 acuerdo con la invención comprende un péptido C que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 1 y un péptido L que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 2. Preferiblemente, el péptido C
- 30 consiste en una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 y el péptido L consiste en una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2.

La insulina es una proteína altamente conservada, con solo variaciones menores entre diferentes especies animales. Por ejemplo, la cadena A y la cadena B de la insulina humana consisten en 21

aminoácidos (numerados A1-A21) y la cadena B de la insulina humana consiste en 30 aminoácidos (numerados B1-B30).

A es una cadena A de insulina y puede ser cualquier tipo de cadena A, tal como una cadena A de insulina humana, ovina, canina, felina, porcina o bovina. En una modalidad preferida, A se deriva de una cadena A de insulina porcina. La cadena A de la insulina porcina corresponde a la cadena A de la insulina humana. La cadena A puede comprender una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 3, mientras que la cadena A bovina difiere de la cadena A humana en solo dos aminoácidos. La insulina humana y porcina contienen un residuo de treonina e isoleucina en las posiciones de residuo A8 y A10 (que se refieren al 8^{vo} y 10^{mo} residuo de aminoácido, p. ej., de acuerdo con la numeración de la SEQ ID NO: 3), respectivamente, mientras que la insulina bovina contiene un residuo de alanina y valina en las posiciones respectivas. En una modalidad preferida, la cadena A se deriva de una cadena A de insulina porcina.

Además, también se han introducido modificaciones de aminoácidos para obtener diferentes variantes sintéticas de insulina. Los ejemplos de variantes de insulina son la insulina lispro y la insulina glargina. Un ejemplo de una modificación de aminoácido para obtener insulina sintética es, por ejemplo, la sustitución del residuo de asparagina en la posición A21 de la insulina humana regular por glicina en la insulina glargina, un tipo de insulina de acción lenta que permanece en niveles relativamente bajos en la sangre durante un período de 24 horas tras la inyección subcutánea. Esta modificación en una proteína de insulina también está contemplada por la presente invención.

Por lo tanto, en una modalidad de la invención y/o sus modalidades, la cadena A es una cadena A que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 3, o una cadena A que tiene al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 % de identidad de secuencia con la SEQ ID NO: 3. La cadena A de acuerdo con la SEQ ID NO: 3 puede modificarse mediante, p. ej., sustitución, delección o inserción de aminoácidos en una posición seleccionada entre A1, A2, A3, A4, A5, A8, A9, A10, A12, A13, A14, A15, A16, A17, A18, A19 y/o A21 para obtener la cadena A comprendida en el precursor de insulina recombinante. En una modalidad de la invención y/o sus modalidades, la cadena A de acuerdo con la SEQ ID NO: 3 se modifica en al menos 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 posiciones de residuo. Una persona normalmente versada en la materia conoce las secuencias de aminoácidos de las cadenas A de insulina adecuadas que pueden utilizarse en la invención proporcionada en el presente documento.

Adicionalmente, los aminoácidos de la cadena A del precursor de insulina recombinante pueden experimentar modificaciones postraduccionales tales como acetilación (conjugación de un grupo acetilo a un aminoácido), PEGilación (conjugación de un grupo polietilenglicol (PEG) a un aminoácido), lipidación (conjugación de un lípido a un aminoácido) o glicosilación (conjugación de una fracción de carbohidrato a un aminoácido), ya sea directamente o a través de un enlazador, o desamidación de residuos de asparagina. En una modalidad de la invención y/o sus modalidades, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o 21 de la cadena A del precursor de insulina recombinante está/están modificados postraduccionalmente.

B es una cadena B de insulina y puede ser cualquier tipo de cadena B, tal como una cadena B de insulina humana, ovina, canina, felina, porcina o bovina. En una modalidad preferida, la cadena B se deriva de una cadena B de insulina porcina. Al igual que la cadena A, la cadena B de la insulina está altamente conservada entre especies animales. La insulina porcina y la insulina bovina solo difieren de la insulina humana en el residuo B30, donde la insulina humana contiene un residuo de treonina, mientras que la insulina porcina y bovina contienen un residuo de alanina. La sustitución de aminoácidos de la cadena B de la insulina puede afectar la actividad de la insulina. Por ejemplo, extender la cadena B con dos residuos de arginina contribuye al perfil de acción lenta de la insulina glargina. Por otro lado, sustituir el residuo de prolina en la posición B28 de la insulina humana por lisina y el residuo de lisina en la posición B29 de la insulina humana por prolina resulta en insulina de acción rápida que causa un pico de insulina en sangre en un plazo de 30 minutos tras la inyección subcutánea y que se degrada en un plazo de 4 horas desde la inyección.

La cadena B como se divulga y proporciona en el presente documento de acuerdo con la SEQ ID NO: 4 comprende los dos residuos de arginina en el extremo C-terminal de la cadena B, es decir, en la posición B31 y B32. Se entiende que la presente invención también contempla cadenas B de insulina que no comprenden dichos dos residuos de arginina.

Por lo tanto, en una modalidad de la invención y/o sus modalidades, la cadena B es una cadena B que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 4, o una cadena B que tiene al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 % de identidad de secuencia con la SEQ ID NO: 4. La cadena B de acuerdo con la SEQ ID NO: 4 puede modificarse mediante, p. ej., sustitución, delección o inserción de aminoácidos en una posición seleccionada entre B1, B2, B3, B4, B5, B9, B10, B13, B14, B15, B16, B17, B18, B20, B21, B22, B23, B26, B27, B28, B29 y/o B30 para obtener la cadena B comprendida en el precursor de insulina recombinante. En una modalidad de la invención y/o sus modalidades, la cadena B de acuerdo con la SEQ ID NO: 4 se modifica en al

menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 posiciones de residuo. Una persona normalmente versada en la materia conoce las secuencias de aminoácidos de las cadenas B de insulina adecuadas que pueden utilizarse en la invención proporcionada en el presente documento.

- 5 Los aminoácidos de la cadena B del precursor de insulina recombinante pueden experimentar modificaciones postraduccionales similares a las descritas previamente para los aminoácidos de la cadena A del precursor de insulina recombinante. En una modalidad de la invención y/o sus modalidades, 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 o 30 aminoácidos de la cadena B del precursor de insulina recombinante está/están
10 modificados postraduccionamente.

La cadena A y la cadena B del precursor de insulina recombinante pueden derivarse de diferentes especies. La cadena A puede ser, por ejemplo, una cadena A de insulina humana, mientras que la cadena B es una cadena B de insulina porcina, o viceversa. En el presente documento también se contemplan combinaciones adicionales de cadenas A y cadenas B de diferentes especies.

- 15 Opcionalmente, el precursor de insulina recombinante puede incluir uno o más péptidos de extensión y/o péptidos espaciadores entre el extremo C-terminal del péptido líder y el extremo N-terminal del péptido de cadena B.

En una modalidad adicional de la invención y/o sus modalidades, el precursor de insulina recombinante comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 6, 7 u 8
20 o comprende una secuencia de aminoácidos con una identidad de secuencia de al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 %, al menos 99 % con el precursor de insulina recombinante de acuerdo con la SEQ ID NO: 6, 7 y/o 8. Preferiblemente, el precursor de insulina recombinante es una proteína que tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 6, 7 u 8.

- 25 El precursor de insulina recombinante que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 6, 7 u 8 contiene una cadena A de acuerdo con la SEQ ID NO: 3 y una cadena B de acuerdo con la SEQ ID NO: 4. Se contempla en el presente documento que una persona normalmente versada en la materia puede derivar la cadena A y/o la cadena B del precursor de insulina recombinante de diferentes especies (p. ej., humana, rumiante o mascota) y
30 proporcionar un precursor de insulina recombinante que comprende un péptido C y/o péptido L de acuerdo con la invención que comprende, además, una cadena A y/o la cadena B de diferentes especies. Por lo tanto, la invención también proporciona precursores de insulina recombinante que comprenden una identidad de secuencia de al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 %, al menos 99 %

con el precursor de insulina recombinante de acuerdo con la SEQ ID NO: 6, 7 y/o 8, o que comprenden el péptido C de acuerdo con la SEQ ID NO: 1 y el péptido L de acuerdo con la SEQ ID NO: 2, y que comprenden una cadena A y/o cadena B de diferentes especies.

- Además, una proteína que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 6 comprende el péptido C de acuerdo con la SEQ ID NO: 1 y el péptido L de referencia de WO2017/040363 que comprende, además, una metionina N-terminal. La proteína que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 7 comprende el péptido L de acuerdo con la SEQ ID NO: 2 y un péptido C modificado derivado de la secuencia del péptido C humano. La proteína que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 8 comprende tanto el péptido C de acuerdo con la SEQ ID NO: 1 como el péptido L de acuerdo con la SEQ ID NO: 2. Como se ejemplifica en el presente documento, la expresión de los precursores de insulina recombinante que comprenden una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 6, 7 u 8 resultó en un título de fermentación mayor del precursor de insulina recombinante que el precursor de insulina recombinante de referencia de acuerdo con la SEQ ID NO: 5 (derivado de WO2017/040363, en donde la secuencia de aminoácidos de la SEQ ID NO: 5, además, contiene una metionina N-terminal).

Secuencias de ADN

En otro aspecto, la invención se refiere a una secuencia de ADN que codifica un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde:

- A es una cadena A de insulina,
B es una cadena B de insulina,
C es un péptido C que conecta la cadena A y la cadena B, y
L es un péptido L N-terminal.

- La cadena A, la cadena B, el péptido C y el péptido L como se divulgan en el presente documento pueden estar codificados por cualquier secuencia de ADN que resulte en una secuencia de aminoácidos que forme dicha cadena A de insulina, una cadena B de insulina, un péptido C, preferiblemente de acuerdo con la SEQ ID NO: 1, y/o un péptido L, preferiblemente de acuerdo con la SEQ ID NO: 2, respectivamente.

- En una modalidad, la invención se refiere a una secuencia de ADN que codifica un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde:

A es una cadena A de insulina,
B es una cadena B de insulina,
C es un péptido C que conecta la cadena A y la cadena B, y
L es un péptido L N-terminal;

en donde C es un péptido C que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 1 y/o L es un péptido L que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 2.

Es sabido que codones similares se transcriben con eficiencia variable en diferentes organismos.

Dado que el precursor de insulina recombinante se produce principalmente en un microorganismo, la secuencia de ADN en algunas modalidades preferidas de la invención está optimizada por codones para su expresión en un microorganismo. La optimización puede lograrse, por ejemplo, utilizando software, tal como GeneOptimizer y DNA2.0.

La secuencia de ADN del gen que codifica para la insulina nativa, tal como la insulina humana nativa o porcina nativa, incluyendo la cadena A, la cadena B, y el péptido C ha sido descrita en la técnica, por ejemplo, por *Bell, G., Pictet, R., Rutter, W. et al. Sequence of the human insulin gene. Nature 284, 26-32 (1980)* o por *Xu-Guang Han et al. Cloning and characterization of porcine insulin gene, Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, Volume 129, Issue 1, 2001, Pages 87-95*. Por lo tanto, una persona normalmente versada en la materia está familiarizado con las secuencias de ADN que codifican la insulina nativa, o cualquiera de la cadena A nativa, la cadena B nativa, el péptido C nativo o el péptido L nativo, tal como se encuentran en varios mamíferos. Un ejemplo de una secuencia de ADN para un péptido C humano nativo.

En el presente documento se proporciona una secuencia de ADN para expresar un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde:

A es una cadena A de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 11;

B es una cadena B de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 12;

C es un péptido C que conecta la cadena A y la cadena B; y

L es un péptido L N-terminal, en donde el péptido C está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 9 y/o en donde el péptido L está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 10.

Preferiblemente, la cadena A del precursor de insulina recombinante está codificada por una secuencia de ADN de acuerdo con la SEQ ID NO: 11, o una secuencia de ADN que tiene, con preferencia creciente, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos

80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % de identidad de secuencia con la SEQ ID NO: 11. Preferiblemente, la cadena B del precursor de insulina recombinante está codificada por una secuencia de ADN de acuerdo con la SEQ ID NO: 12, o una secuencia de ADN que tiene, con preferencia creciente, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % de identidad de secuencia con la SEQ ID NO: 12. Preferiblemente, el péptido C del precursor de insulina recombinante está codificado por una secuencia de ADN de acuerdo con la SEQ ID NO: 9, o una secuencia de ADN que tiene, con preferencia creciente, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % de identidad de secuencia con la SEQ ID NO: 9. Preferiblemente, el péptido L del precursor de insulina recombinante está codificado por una secuencia de ADN de acuerdo con la SEQ ID NO: 10, o una secuencia de ADN que tiene, con preferencia creciente, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % de identidad de secuencia con la SEQ ID NO: 10.

La insulina es una proteína de mamífero que normalmente no es expresada por el microorganismo típicamente utilizado como células huésped para producir insulina recombinante. Aunque típicamente un aminoácido está codificado por diferentes codones, la preferencia de codones difiere entre las células de mamífero y los microorganismos, y la expresión de una secuencia de ADN de mamífero en un microorganismo típicamente resulta en rendimientos relativamente bajos del péptido o proteína codificado. La optimización de codones de la secuencia de ADN para su expresión en un microorganismo mejora el rendimiento del precursor recombinante y la fidelidad de la expresión.

La insulina recombinante o la molécula de insulina madura producida tras la expresión de una secuencia de ADN por un microorganismo y la posterior digestión de la molécula para eliminar el péptido L y el péptido C puede analizarse utilizando técnicas de análisis conocidas en la técnica, tales como la cromatografía líquida-espectrometría de masas (LC-MS).

Como se muestra en la Fig. 2, los picos del perfil LC-MS entre 21 y 23 segundos (indicados por flechas) se reducen notablemente cuando se utilizó el codón “AAC” (en lugar de “AAT”) para codificar los residuos de asparagina presentes en las posiciones de residuo A18 y A21 de la cadena A de la molécula de insulina cuando la secuencia de ADN que codifica el precursor de insulina recombinante se expresó en *E. coli*. Esto indica una fidelidad de expresión mejorada y una reducción de la incorporación errónea de aminoácidos. Por lo tanto, en una modalidad de la invención y/o sus modalidades, un residuo de asparagina de la cadena A del precursor de insulina recombinante está codificado por el codón “AAC”. Preferiblemente, los residuos de asparagina

en las posiciones de residuo A18 y A21 de la cadena A de la insulina están codificados por el codón “AAC”. Cuando más de un residuo de asparagina está presente en la cadena A del precursor de insulina recombinante, por ejemplo, 2, 3, 4 o todos los residuos de asparagina de la cadena A del precursor de insulina recombinante están codificados por el codón “AAC”.

- 5 Como se muestra en la Fig. 3, los picos del perfil LC-MS entre 20,5 y 21,5 segundos se reducen notablemente cuando se utilizó el codón “TCT” (en lugar de “AGC”) para codificar los residuos de serina presentes en las posiciones de residuo A9 y A12 de la cadena A de la molécula de insulina cuando la secuencia de ADN que codifica el precursor de insulina recombinante se expresó en *E. coli*. Nuevamente, esto indica una fidelidad de expresión mejorada y una reducción
- 10 en la incorporación errónea de aminoácidos. Por lo tanto, en una modalidad de la invención y/o sus modalidades, un residuo de serina de la cadena A del precursor de insulina recombinante está codificado por el codón “TCT”. Preferiblemente, los residuos de serina en las posiciones de residuo A9 y A12 de la cadena A de la insulina están codificados por el codón “TCT”. Cuando más de un residuo de serina está presente en la cadena A del precursor de insulina recombinante,
- 15 por ejemplo 2, 3, 4 o todos los residuos de serina de la cadena A del precursor de insulina recombinante están codificados por el codón “TCT”.

- En algunas modalidades de la invención y/o sus modalidades, uno o más residuos de asparagina de la cadena A del precursor de insulina recombinante están codificados por el codón “AAC” y uno o más residuos de serina de la cadena A del precursor de insulina recombinante están
- 20 codificados por el codón “TCT”. En una modalidad preferida, la cadena A comprende que los residuos de asparagina en las posiciones de residuo A18 y A21 de la cadena A de la insulina están codificados por el codón “AAC” y los residuos de serina en las posiciones de residuo A9 y A12 de la cadena A de la insulina están codificados por el codón “TCT”, tal como se ejemplifica en la secuencia de aminoácidos de la cadena A de la SEQ ID NO: 3 y la secuencia de ADN
- 25 correspondiente de la SEQ ID NO: 11.

- Dado que la secuencia de aminoácidos de la cadena A está ubicada en el extremo C-terminal de la secuencia de aminoácidos del precursor de insulina recombinante, la secuencia de ADN que codifica la cadena A y/o el precursor de insulina recombinante de acuerdo con la invención puede comprender, además, una secuencia de codón de terminación. Una persona normalmente versada
- 30 en la materia conoce la función y las secuencias de ácido nucleico de tales codones de terminación, que por lo tanto están contemplados en el presente documento. Un codón de terminación como se divulga en el presente documento puede ser, en un ejemplo no limitante, “TAA” y puede situarse después del residuo de ácido nucleico 63 de la SEQ ID NO: 11 (por lo tanto, “TAA” será los residuos de ácido nucleico 64-66 de dicha SEQ ID NO: 11).

En una modalidad adicional, las secuencias de ADN divulgadas en el presente documento han sido optimizadas para su expresión en un microorganismo, preferiblemente *E. coli*. Se entiende que la optimización por codones puede utilizarse para optimizar las secuencias de ADN para su expresión en microorganismos distintos de *E. coli*. Por lo tanto, se entiende que tales secuencias de ADN optimizadas, preferiblemente cualquiera o más de las secuencias de ADN de la SEQ ID NO: 9-15, también están contempladas por la presente divulgación.

En una modalidad de la invención y/o sus modalidades, la secuencia de ADN que codifica el precursor de insulina recombinante de acuerdo con la invención comprende una secuencia de ADN de acuerdo con la SEQ ID NO: 13, 14 o 15, o una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 13, 14 o 15. Más preferiblemente, el precursor de insulina recombinante de acuerdo con la invención está codificado por una secuencia de ADN que tiene, con preferencia creciente, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 99 % de identidad de secuencia con la SEQ ID NO: 13, 14 o 15. En una modalidad preferida, el precursor de insulina recombinante de acuerdo con la invención está codificado por una secuencia de ADN de la SEQ ID NO: 13, 14 o 15.

La secuencia de ADN que codifica el precursor de insulina recombinante puede incluirse en un vector para expresar el precursor de insulina recombinante. Por lo tanto, la invención se refiere en otro aspecto a un vector que comprende una secuencia de ADN para expresar un precursor de insulina recombinante como se proporciona en el presente documento. En algunas modalidades se proporciona un vector que comprende una secuencia de ADN para expresar un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde:

- A es una cadena A de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 11;
- B es una cadena B de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 12;
- C es un péptido C que conecta la cadena A y la cadena B, preferiblemente en donde el péptido C está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 9; y
- L es un péptido L N-terminal, preferiblemente codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 10.

La secuencia de ADN se clona típicamente aguas abajo de y operativamente ligada a una secuencia promotora, tal como un promotor T7 o un promotor tac. La secuencia de ADN puede clonarse en el vector cortando el vector con enzimas de restricción que reconocen sitios de

restricción presentes en el vector, tales como XhoI, Bam, NdeI, BglI, HindIII. Una secuencia de ADN flanqueada por secuencias de ácido nucleico compatibles con el sitio de restricción respectivo puede entonces insertarse en el vector, y el vector cerrarse utilizando la enzima ligasa.

Típicamente, el vector comprende uno o más genes de resistencia a antibióticos que confieren resistencia contra, p. ej., kanamicina o ampicilina, para facilitar la selección de una célula huésped que ha incorporado el vector después de la transformación. El promotor puede ser un promotor inducible que está, por ejemplo, bajo el control de LacI, previniendo así la expresión de la secuencia de ADN hasta la inducción por, p. ej., isopropil β -d-1-tiogalactopiranosido (IPTG). Preferiblemente, el vector es un plásmido. Más preferiblemente, el vector es un plásmido que comprende un sitio de clonación múltiple. Un sitio de clonación múltiple contiene varios sitios de restricción. De esta manera, se amplía la elección de posibles enzimas de restricción que pueden utilizarse para insertar el ADN en el vector.

El vector que codifica el precursor de insulina recombinante puede expresarse en una célula huésped. Por lo tanto, la invención se refiere en otro aspecto con una célula huésped que comprende una secuencia de ADN o que comprende un vector que comprende una secuencia de ADN para expresar un precursor de insulina recombinante como se proporciona en el presente documento. Por lo tanto, en una modalidad se proporciona una célula huésped que comprende una secuencia de ADN de acuerdo con la fórmula L-B-C-A, en donde:

A es una cadena A de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 11;

B es una cadena B de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 12;

C es un péptido C que conecta la cadena A y la cadena B, preferiblemente en donde el péptido C está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 9; y

L es un péptido L N-terminal, preferiblemente codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 10.

Típicamente, la célula huésped es un microorganismo, tal como una bacteria o una célula de levadura. En una modalidad de la invención y/o sus modalidades, la célula huésped es una célula de *Escherichia coli* o una célula de *Saccharomyces cerevisiae*. En una modalidad preferida, la célula huésped es una célula de *Escherichia coli*.

Para producir mayores cantidades del precursor de insulina recombinante, el precursor de insulina recombinante puede expresarse en una población de células huésped para obtener el precursor de insulina. Por lo tanto, la invención se refiere en otro aspecto a un cultivo de células

huésped que comprende una secuencia de ADN para expresar un precursor de insulina recombinante de acuerdo con la invención.

Típicamente, las condiciones bajo las cuales se cultiva el cultivo de células huésped están controladas. Las células huésped se cultivan, por ejemplo, en un medio que comprende uno o
 5 más nutrientes y que tiene un pH adecuado para la célula huésped particular y en una temperatura y a una concentración de O₂ y/o CO₂ apropiada para la propagación de la célula huésped particular.

Métodos para producir insulina

10 Los métodos y procesos para la producción de insulina o análogos de insulina incluyen en general varios pasos. Procesos similares para la producción de insulina madura son conocidos por una persona normalmente versada en la materia y han sido utilizados previamente, tal como en WO2017/040363A1. El precursor de insulina recombinante puede utilizarse en métodos para producir insulina recombinante (madura). Por lo tanto, la invención se refiere en otro aspecto a
 15 un método para producir insulina recombinante (madura), en donde se expresa un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde:

A es una cadena A de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 11;

20 B es una cadena B de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 12;

C es un péptido C que conecta la cadena A y la cadena B, preferiblemente en donde el péptido C está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 9; y

25 L es un péptido L N-terminal, preferiblemente codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 10;

el método comprende los pasos de:

- a) expresar el precursor de insulina recombinante en una célula huésped,
- b) lisar la célula huésped para obtener el precursor de insulina recombinante,
- c) replegar el precursor de insulina recombinante, y
- 30 d) escindir el péptido L y el péptido C del precursor de insulina recombinante mediante una reacción enzimática para obtener insulina recombinante.

La expresión del precursor de insulina recombinante en un microorganismo puede realizarse mediante métodos conocidos en la técnica, tales como la fermentación de una célula huésped, preferiblemente en donde la célula huésped es un microorganismo.

En algunas modalidades, el paso de clonar una secuencia de ADN que codifica el precursor de insulina recombinante en un vector precede al paso a) del método para producir insulina recombinante como se divulga en el presente documento.

- En algunas modalidades, el método comprende, además, un paso para purificar la insulina recombinante obtenida por el método de producción divulgado en el presente documento. Una persona normalmente versada en la materia conoce los métodos adecuados para la purificación de la insulina recombinante.

En una modalidad de la invención y/o sus modalidades, el microorganismo es *Escherichia coli* o *Saccharomyces cerevisiae*.

- En otro aspecto, la invención se refiere a un uso de un precursor de insulina recombinante como se proporciona en el presente documento en la producción de insulina recombinante.

- Finalmente, la invención se refiere a una composición que comprende insulina y una cantidad detectable de un péptido C y/o un péptido L como se proporciona en el presente documento, preferiblemente en donde el péptido C comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 o una secuencia que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 1 y/o el péptido L comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2 o una secuencia que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 2. En algunas modalidades, el péptido C y/o el péptido L pueden estar conectados a la cadena A y/o la cadena B de una molécula de insulina. Esto puede ocurrir, por ejemplo, cuando el precursor de insulina no se digiere completamente. Alternativamente, el péptido C y/o el péptido L pueden estar presentes en la composición de insulina en forma libre. Esto puede ocurrir, por ejemplo, cuando el precursor de insulina se digiere completamente, pero la purificación de la insulina digerida es subóptima. Por lo tanto, en una modalidad, una composición de insulina comprende una cantidad detectable de un péptido C, un péptido L, o tanto un péptido C como un péptido L. La presencia del péptido C y/o el péptido L puede detectarse utilizando técnicas analíticas estándar, tales como la cromatografía líquida-espectrometría de masas en tándem (LC-MS/MS). Por ejemplo, Foulon et al., 2022, JMSACL, 25:19-26 describe un método para cuantificar la insulina y el péptido C mediante LC-MS/MS sin el uso de anticuerpos, utilizando muestras de suero. La persona normalmente versada en la materia sabrá, sin embargo, aplicar métodos similares en una muestra obtenida de una composición de insulina.

En una modalidad de la invención y/o sus modalidades, la composición que comprende insulina comprende el péptido C y/o el péptido L en una concentración de $< 10 \mu\text{g/ml}$, preferiblemente menos de $9 \mu\text{g/ml}$, $< 8 \mu\text{g/ml}$, $< 7,5 \mu\text{g/ml}$, $< 7 \mu\text{g/ml}$, $< 6 \mu\text{g/ml}$, $< 5 \mu\text{g/ml}$, $< 4 \mu\text{g/ml}$, $< 3 \mu\text{g/ml}$, $< 2,5 \mu\text{g/ml}$, $< 2 \mu\text{g/ml}$, $< 1 \mu\text{g/ml}$, $< 0,5 \mu\text{g/ml}$, $< 0,25 \mu\text{g/ml}$ o $< 0,1 \mu\text{g/ml}$. Por ejemplo,

la composición que comprende insulina comprende el péptido C de acuerdo con la SEQ ID NO: 1 y/o el péptido L de acuerdo con la SEQ ID NO: 2 en una concentración de 1-5000 ng/ml, tal como 50-4000 ng/ml, 100-3000 ng/ml, 250-2500 ng/ml, 500-2000 ng/ml, 750-1500 ng/ml o 1000-1250 ng/ml. En una modalidad alternativa, la composición de insulina comprende < 10

5 ppm del péptido C de acuerdo con la SEQ ID NO: 1 y/o el péptido L de acuerdo con la SEQ ID NO: 2 o < 3 ppm del péptido C de acuerdo con la SEQ ID NO: 1 y/o el péptido L de acuerdo con la SEQ ID NO: 2 o < 3000 ng/ml del péptido C y/o el péptido L. La concentración del péptido C y la concentración del péptido L pueden ser independientes entre sí.

Otra modalidad de la invención es una composición de insulina recombinante producida por los métodos anteriores y un vehículo farmacéuticamente aceptable.

Otra modalidad de la invención es un método para tratar la diabetes que comprende administrar a un animal que lo necesite una cantidad terapéutica de la composición de insulina recombinante anterior.

Otra modalidad es una composición de insulina recombinante para su uso en el tratamiento de la diabetes.

15 La descripción anterior de las modalidades específicas revelará completamente la naturaleza general de la invención de tal manera que otros pueden, aplicando conocimiento dentro de la habilidad de la materia (incluyendo el contenido de las referencias citadas en el presente documento), modificar y/o adaptar fácilmente para diversas aplicaciones, tales como modalidades específicas, sin experimentación indebida, sin apartarse del concepto general de la presente invención. Por lo tanto, tales adaptaciones y modificaciones están destinadas a estar dentro del significado y el rango de equivalentes de las modalidades divulgadas, basándose en la enseñanza y la orientación presentadas en el presente documento.

Todas las referencias citadas en el presente documento, incluyendo artículos de revistas o resúmenes, solicitudes de patente publicadas o correspondientes, patentes u otras referencias, se incorporan por referencia en el presente documento en su totalidad, incluyendo todos los datos, tablas, figuras y texto presentados en las referencias citadas. Adicionalmente, el contenido completo de las referencias citadas dentro de las referencias citadas en el presente documento también se incorpora en su totalidad por referencia.

30 Se entiende que la fraseología o terminología en el presente documento tiene por objeto la descripción y no la limitación, de tal manera que la terminología o fraseología de la presente memoria descriptiva debe ser interpretada por la persona normalmente versada en la materia a la luz de las enseñanzas y la orientación presentadas en el presente documento, en combinación con el conocimiento de una persona normalmente versada en la materia.

Habiendo descrito ahora la invención en términos generales, la misma se comprenderá más fácilmente mediante referencia al siguiente ejemplo que se proporciona a modo de ilustración y no está destinado a ser limitante de la presente invención. Aspectos y modalidades adicionales serán evidentes para las personas normalmente versadas en la materia.

5 Ejemplos

Ejemplo 1

Producción del precursor de insulina e insulina

Optimización por codones

Las secuencias de ADN que codifican los precursores de insulina de acuerdo con la SEQ ID NO: 5-8 fueron optimizadas por codones para su expresión en *E. coli* utilizando GeneART y DNA2.0, dando como resultado secuencias de ADN de acuerdo con la SEQ ID NO: 13-15.

Además, la incorporación del codón AAC para todos los residuos de asparagina y el codón TCT para todos los residuos de serina redujo la formación de impurezas (incorporaciones erróneas de aminoácidos) después de la digestión como se muestra en la Fig. 2 y 3, respectivamente.

15 *Construcción de los plásmidos*

El vector pET29 como vector de expresión se escindió con las enzimas de restricción NdeI y XhoI, y se sometió a electroforesis en un gel de agarosa al 1 % para aislar un segmento de ADN de 5,2 kb.

Una secuencia de ADN de acuerdo con las SEQ ID NO: 13, 14 o 15 flanqueada por un sitio de reconocimiento de la enzima de restricción NdeI y un sitio de restricción de la enzima de restricción XhoI se unió entonces al vector escindido utilizando la ADN ligasa T4 para formar el plásmido que comprende la secuencia de ADN de acuerdo con las SEQ ID NO: 13, 14 o 15.

E. coli BL21(DE3) se transformó mediante el método de choque térmico y se sembró en una placa de agar. Tras la incubación, se seleccionaron las células transformadas resistentes a la kanamicina. El ADN del plásmido se aisló de un transformante individual y se confirmó que el ADN deseado había sido insertado correctamente mediante un análisis por escisión con enzimas de restricción.

Producción del precursor

Las células *E. coli* ITP3 competentes se transformaron mediante el método de choque térmico con un plásmido que comprende la secuencia de ADN de acuerdo con la SEQ ID NO: 13, 14 o 15 con el fin de producir un precursor de insulina de acuerdo con la SEQ ID NO: 5, 6, 7 u 8, respectivamente. SEQ ID NO: 5 corresponde sustancialmente al precursor de insulina descrito en WO2017/040363, excepto que en comparación con la secuencia de WO2017/040363, la SEQ ID

NO: 5 comprende, además, un residuo de metionina presente en la primera posición (posición N-terminal) de la SEQ ID NO: 5.

Las células *E. coli* ITP3 transformadas se cultivaron durante 18 h a 37 °C en matraces de agitación que contenían 800 ml de caldo LB suplementado con 100 µg/l de kanamicina hasta que la densidad celular alcanzó aproximadamente 2,1 (A600 nm). Tras el cultivo, se realizó un análisis por SDS-PAGE reducido en un gel de gradiente de poliacrilamida al 4-12 % precolado, y se utilizó Invitrogen Simply Blue SafeStain para la tinción. El nivel de expresión de cada proinsulina fusionada se cuantificó utilizando ChemiDoc.

Las células se recolectaron por centrifugación a aproximadamente 7000 rpm durante 20 min y los pesos húmedos de las células se midieron con una balanza electrónica (Sartorius, Alemania). Las células sedimentadas se suspendieron en 20 ml de Tris-HCl 50 mM (pH 7,5), EDTA 10 mM, lisozima al 0,02 %, se lisaron por sonicación y se centrifugaron a aproximadamente 10.000 rpm durante 20 min.

El sedimento que contenía los cuerpos de inclusión (IB) que comprenden el precursor de insulina se lavó resuspendiéndolo en Tris-HCl 20 mM (pH 7,5), Triton X-100 al 1 %, tampón de urea 2 M, seguido de centrifugación a 10.000 rpm durante 20 min. Finalmente, el sedimento se lavó dos veces con agua desionizada.

La Tabla 1 muestra el título de fermentación de los precursores de insulina de acuerdo con las SEQ ID NO: 5-8 producidos tras la expresión de las secuencias de ADN de acuerdo con las SEQ ID NO: 13-15, respectivamente, en *E. coli*. El péptido L N-terminal está subrayado y el péptido C se indica en negrita.

El precursor de insulina recombinante de acuerdo con la SEQ ID NO: 6 contiene la cadena A y la cadena B de insulina de acuerdo con las SEQ ID NO: 3 y 4, respectivamente, y el péptido L de referencia de WO2017/040363 incluyendo el residuo de metionina N-terminal para la fermentación. En contraste con la SEQ ID NO: 5, SEQ ID NO: 6 contiene un péptido C de acuerdo con la SEQ ID NO: 1. Sorprendentemente, el título de fermentación del precursor de insulina de acuerdo con la SEQ ID NO: 6 fue mayor que el título de fermentación del precursor de insulina de referencia de acuerdo con la SEQ ID NO: 5. Como se muestra en la Tabla 1 del Ejemplo 1, el título de fermentación en mmol/l del precursor de insulina de acuerdo con la SEQ ID NO: 6 fue aproximadamente un 43 % mayor que el título de fermentación en mmol/l del precursor de insulina de referencia de acuerdo con la SEQ ID NO: 5 cuando los precursores se expresaron en células huésped en las mismas condiciones de fermentación. Por lo tanto, reemplazando el péptido C humano nativo del precursor de referencia de acuerdo con el péptido

C de acuerdo con la SEQ ID NO: 1 aumenta significativamente el título de fermentación del precursor.

Además, el precursor de insulina recombinante de acuerdo con la SEQ ID NO: 7 contiene la cadena A y la cadena B de insulina de acuerdo con las SEQ ID NO: 3 y 4, respectivamente, en combinación con un péptido L de acuerdo con la SEQ ID NO: 2. Además, la SEQ ID NO: 7 contiene un péptido C modificado derivado de la secuencia del péptido C humano. Como se muestra en la Tabla 1 del Ejemplo 1, el título de fermentación en mmol/l del precursor de insulina de acuerdo con la SEQ ID NO: 7 fue aproximadamente un 39 % mayor que el título de fermentación del precursor de insulina de referencia de acuerdo con la SEQ ID NO: 5 y que comprende el péptido L de referencia y el péptido C humano nativo bajo las mismas condiciones de fermentación.

Al igual que el precursor de insulina recombinante de acuerdo con la SEQ ID NO: 6, el precursor de insulina recombinante de acuerdo con la SEQ ID NO: 8 contiene la cadena A y la cadena B de insulina de acuerdo con las SEQ ID NO: 3 y 4, respectivamente, en combinación con el péptido C de acuerdo con la SEQ ID NO: 1. Además, la SEQ ID NO: 8 contiene el péptido L de acuerdo con la SEQ ID NO: 2. El título de fermentación del precursor de insulina de acuerdo con la SEQ ID NO: 8 fue incluso mayor que el título de fermentación del precursor de insulina de acuerdo con la SEQ ID NO: 5 o la SEQ ID NO: 6. Como se muestra en la Tabla 1 del Ejemplo 1, el título de fermentación en mmol/l del precursor de insulina de acuerdo con la SEQ ID NO: 8 fue incluso aproximadamente un 118 % mayor que el título de fermentación en mmol/l del precursor de insulina de referencia de acuerdo con la SEQ ID NO: 5. Por lo tanto, reemplazando el péptido C humano nativo por el péptido C de acuerdo con la SEQ ID NO: 1 y adicionalmente el péptido L de referencia por el péptido L de acuerdo con la SEQ ID NO: 2 aumenta el título de fermentación del precursor de insulina de manera sinérgica.

Tabla 1. Títulos de fermentación (en g/l y en mmol/l) de los precursores de insulina recombinante expresados en E. coli. Los respectivos péptidos L están subrayados y los respectivos péptidos C se indican en negrita.

<u>SEQ ID NO</u>	<u>Secuencia</u>	<u>Peso molecular (kDa)</u>	<u>Título de fermentación (g/l)</u>	<u>Título de fermentación (mmol/l)</u>
5	<u>MSQSIGVVFVDYNSHTRF</u> <u>VNQHLGSHLVEALYL</u> <u>VG</u> RGFFYTPKARREAED <u>LQVGQVELGGPGAGSL</u> <u>QPLALEGSLQKRGIVEQC</u> <u>CTSICSLYQLENYCN</u>	11,32	6,0	0,53

6	<u>MSQSIGVFVQDYNSTRF</u> VNQHLCGSHLVEALYL VCGERGFFYTPKARRYA PE DGPGEQRGIVEQCCTS ICS LYQLENYCN	9,23	7,0	0,758
7	<u>MSPHHKHHKHHSEDLRF</u> VNQHLCGSHLVEALYL VCGERGFFYTPKARREA EN PQAGAVELGGGLGGLQ A LALEGPQRGIVEQCCT SI CSLYQLENYCN	10,96	8,1	0,739
8	<u>MSPHHKHHKHHSEDLRF</u> VNQHLCGSHLVEALYL VCGERGFFYTPKARRYA PE DGPGEQRGIVEQCCTS ICS LYQLENYCN	9,28	10,7	1,153

Solubilización de los cuerpos de inclusión

Se proporciona una suspensión de proteína de cuerpos de inclusión que comprende el precursor de insulina a una concentración de aproximadamente 30 g/l en agua. Tras cargar los IB en el agua, se añade etanolamina como líquido puro a una concentración final de aproximadamente 315 mM, seguido de la adición de urea 8 M a una concentración final de aproximadamente 4,0 M. Finalmente, se añade una solución acuosa de ditioneitol (DTT) 1 M para alcanzar una concentración de aproximadamente 2,5 mM, y la mezcla se agita durante un mínimo de 30 minutos para completar la solubilización de los IB.

10 *Replegamiento del precursor*

El precursor de insulina se diluye aproximadamente 10 veces hasta una concentración objetivo del precursor de insulina de aproximadamente 1,6 g/l. Primero, se prepara la cantidad objetivo de solución diluyente de replegamiento (etanolamina 10 mM, glicol de hexileno al 10 % (v/v)), luego se añade una cantidad determinada de diclorhidrato de cistamina 1 M (p. ej., aproximadamente 384 μ M) y se mezcla hasta obtener una mezcla homogénea. A continuación, la solución de proteína solubilizada se transfiere a la solución diluyente de replegamiento con un nivel mínimo de agitación necesario para asegurar la mezcla de la solución mientras se mantiene una temperatura de la solución de 10 °C ($\pm$ 2 °C). La reacción de replegamiento se agita a 10 °C hasta que la tasa de conversión al precursor de insulina correctamente plegada cae por debajo del 5 % por hora, según se mide mediante el ensayo POROS HPLC, punto en el cual la reacción se neutraliza. Después de la finalización del replegamiento, la reacción se detiene/ralentiza acidificando la solución de replegamiento, con ácido clorhídrico 2N, hasta un pH de aproximadamente 9,2 a 10 °C. La temperatura se mantiene a 10 °C durante el ajuste del pH.

Cromatografía de intercambio aniónico (AEX)

Los restos celulares y la Proteína de la Célula Huésped (HCP) precipitada se separan antes de la carga en la columna AEX. En preparación para la clarificación, la temperatura de la solución posterior al replegamiento se incrementa de 10 °C a 20 °C durante un lapso de 60 - 120 minutos.

- 5 La clarificación se realiza a través de dos bancos de filtros de profundidad y un banco de filtros de membrana de 0,22 µm, todos en serie. Los filtros de profundidad consisten en filtros CUNO EXT 60ZA05A (banco nro. 1, área total de 56 m²) en serie con filtros CUNO EXT 90ZA08A (banco nro. 2, también de 56 m²). Adicionalmente en serie con los filtros de profundidad hay filtros de 0,22 µm (área total de 9 m²). Ambos bancos de filtros de profundidad se enjuagan
- 10 juntos con 54 l/m² de agua, o aproximadamente 3000 l, antes del uso. Una vez que los filtros de profundidad se enjuagaron, los filtros de 0,22 µm se enjuagan (>20 l/m²) haciendo fluir agua a través de los bancos de filtros de profundidad nro. 1 y nro. 2 y luego a través de los filtros de 0,22 µm hacia el drenaje debido a las restricciones del equipo. Se realiza un desplazamiento de aire del agua de enjuague en los filtros de profundidad antes de introducir el producto. Los filtros se
- 15 operan al principio bajo flujo constante y luego, a medida que los filtros comienzan a ensuciarse, la tasa de flujo se reduce para evitar que la presión exceda las 40 psig. Después de que la solución posterior al replegamiento se ha procesado completamente, se realiza un lavado con agua de recuperación (25-30 l/m²) para maximizar la recuperación del precursor correctamente plegado en una solución posterior al replegamiento clarificada.
- 20 El pH de la solución posterior al replegamiento clarificada se ajusta a aproximadamente 9,4 con hidróxido de sodio y se diluye con agua, si es necesario, para lograr una conductividad de menos de 2,5 mS/cm. La solución resultante se carga en una columna empacada con DEAE Sepharose Fast Flow, que se ha equilibrado con una solución de equilibrado que comprende aproximadamente borato de sodio 50 mM y cloruro de sodio 2,5 mM a pH 9,4. Las tasas de flujo
- 25 durante los pasos de carga y posterior elución se ajustan para mantener un tiempo de residencia de aproximadamente cinco a siete minutos, y se usa un factor de carga de aproximadamente 23 g del precursor de insulina correctamente plegado por l de resina de columna. Después de la carga, la columna se lava con aproximadamente cinco volúmenes de columna (CV) de solución de equilibrado (borato de sodio 50 mM, cloruro de sodio 2,5 mM a pH 9,4), y luego con
- 30 aproximadamente seis CV de solución de elución (aproximadamente borato de sodio 50 mM y cloruro de sodio 160 mM a pH 9,0). El pico principal observado durante la elución con la solución de elución de cloruro de sodio 160 mM se recolecta para proporcionar un conjunto posterior a AEX que contiene el precursor de insulina correctamente plegado.

Citraconilación y digestión triptica

El propósito del paso de citraconilación es disminuir la generación de escisiones erróneas producidas durante el paso de digestión triptica. La citraconilación se logra a través de la reacción de anhídrido citracónico con el precursor de insulina bajo condiciones básicas. El anhídrido citracónico reacciona con cualquier amina primaria, “bloqueando” así el extremo N-terminal, así como todos los residuos de lisina que se encuentran en la molécula. En este estado “protegido” la molécula procede a través del paso de digestión triptica, donde se añade tripsina porcina recombinante para escindir el péptido L N-terminal y el péptido C interno de la molécula del precursor de insulina. Después de que la digestión está completa, la proteína se desprotege a través de hidrólisis ácida para producir el producto de insulina deseado para purificación adicional a través de procesamiento posterior subsiguiente.

Antes de comenzar la reacción de protección, se determina la concentración del precursor de insulina correctamente plegado en el conjunto posterior a AEX. A continuación, se añaden tres inyecciones en bolo de anhídrido citracónico puro al conjunto AEX a temperatura ambiente para lograr la proporción correcta de anhídrido a precursor de insulina correctamente plegado.

Después de que la adición de anhídrido está completa, el pH se ajusta a 8,5 (si es necesario) con solución de HCl o NaOH y se permite que la reacción de protección proceda por $2 \pm 0,25$ horas. Después de que la reacción de protección está completa, el pH se verifica y, si ha cambiado, se titula de nuevo a 8,5 usando una solución de HCl o NaOH. La solución de tripsina se añade luego para lograr una proporción de tripsina a precursor de insulina correctamente plegado de aproximadamente 1:11.000 en masa. La digestión procede por 12 ± 1 horas. Después de que la digestión está completa, la reacción se detiene primero añadiendo ácido acético para lograr una concentración en el conjunto de acetato 150 mM, y luego ajustando el pH del conjunto de reacción a pH 2,4 usando HCl. La reacción de desprotección (en un volumen de aproximadamente 4100 l) luego continúa por $4 \pm 0,25$ horas para producir insulina correctamente plegada.

Tabla 2: Las secuencias de aminoácidos de las SEQ ID NO: 1-8 y las correspondientes secuencias de ADN de las SEQ ID NO: 9-15. La secuencia 5 divulga la secuencia de aminoácidos de un precursor de insulina que comprende el péptido L de referencia y el péptido C humano nativo. En las secuencias 6-8 y 13-15, las respectivas secuencias de aminoácidos de los péptidos L y las secuencias de ADN que codifican las mismas están subrayadas, y las respectivas secuencias de aminoácidos de los péptidos C y las secuencias de ADN que codifican las mismas se indican en negrita.

	SEQ ID NO	Tipo de secuencia	Secuencia
Péptido C recombinante	1	Aminoácidos	YAPEDGPGEQR
	9	ADN	TATGCGCCGGAAGATGGCCCGGGTGAACAGCGT
Péptido L recombinante	2	Aminoácidos	MSPHHKHHKHSEDLR
	10	ADN	ATGAGCCCCGCATCACAAACATCACAAACATCACAGCAAGATCTGCGT
Cadena A	3	Aminoácidos	GIVEQCCTSICSLYQLENYCN
	11	ADN	GGTATTGTTGAACAGTGTGTACCTCTATTTGTCTCTGTATCAGCTGGAAAACATTGCAAC
Cadena B	4	Aminoácidos	FVNQHLCGSHLVEALYLVCGERGFFYTPKARR
	12	ADN	TTTGTTAATCAGCATCTGTGTGGTAGCCATCTGGTTGAAGCACTGTATCTGGTTTGTGGTGAACGCGGTTTTTTTATACCCCGAAAGCGCGTCTG
Precursor de insulina recombinante de referencia	5	Aminoácidos	<u>MSQSIGVFVQDYN</u> SHTRFVNQHLCGSHLVEALYLVCGERGFFYTPKARREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN
Precursor de insulina recombinante	6	Aminoácidos	<u>MSQSIGVFVQDYN</u> SHTRFVNQHLCGSHLVEALYLVCGERGFFYTPKARRYAPEDGPGEQRGIVEQCCTSICSLYQLENYCN
	13	ADN	<u>ATGAGCCAATCGATTGGGGTTTTCTGTT</u> CAGGACTACAACAGCCACACACGCTTTGTTAATCAGCATCTGTGTGGTAGCCATCTGGTTGAAGCACTGTATCTGGTTTGTGGTGAACGCGGTTTTTTTATACCCCGAAAGCGCGTCTGTTATGCGCCGGAAGATGGCCCGGGTGAACAGCGTGGTATTGTTGAACAGTGTGTACCTCTATTTGTTCTCTGTATCAGCTGGAAAACATTGCAAC
Precursor de insulina recombinante	7	Aminoácidos	<u>MSPHHKHHKHSEDLR</u> FVNQHLCGSHLVEALYLVCGERGFFYTPKARREAENPQAGAVELGGGLGLQALALEGPPQRGIVEQCCTSICSLYQLENYCN
	14	ADN	<u>ATGAGCCCCGCATCACAAACATCACAAACATCACAGCAAGATCTGCGT</u> TTTGTGTTAATCAGCATCTGTGTGGTAGCCATCTGGTTGAAGCACTGTATCTGGTTTGTGGTGAACGCGGTTTTTTTATACCCGAAAGCGCGTCTGTAAGCAGAAAACCCGCAGGCTGGTGCTGTTGAACTGGGTGGTGGTCTGGGTGGTCTGCAGGCGCTGGCACTGGAAGGTCCGCCGCAGCGTGGTATTGTTGAACAGTGTGTGTA CCTCTATTTGTTCTCTGTATCAGCTGGAAAACATTGCAAC

	SEQ ID NO	Tipo de secuencia	Secuencia
Precursor de insulina recombinante	8	Aminoácidos	<u>MSPHHKHHKHHSEDLRFVNQHL</u> CGSHLVEALY LVCGERGFFYTPKARRYAPEDGPGEQRGIVEQC CTSICSLYQLENYCN
	15	ADN	<u>ATGAGCCCCGCATCACAAACATCACAAACATCA</u> <u>CAGCGAAGATCTGCGTTTTGTTAATCAGCATC</u> TGTGTGGTAGCCATCTGGTTGAAGCACTGTAT CTGGTTTGTGGTGAACGCGGTTTTTTTTATACC CCGAAAGCGCGTCGTTATGCGCCGGAAGATGG CCCGGGTGAACAGCGTGGTATTGTTGAACAGT GTTGTACCTCTATTTGTTCTCTGTATCAGCTGG AAACTATTGCAAC

LO QUE SE REIVINDICA ES:

1. Un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde
A es una cadena A de insulina,
B es una cadena B de insulina,
5 C es un péptido C que conecta la cadena A y la cadena B, y
L es un péptido L N-terminal,
en donde C es un péptido C que comprende una secuencia de aminoácidos de acuerdo con la
SEQ ID NO: 1 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de
secuencia con la SEQ ID NO: 1.
- 10 2. Un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde
A es una cadena A de insulina,
B es una cadena B de insulina,
C es un péptido C que conecta la cadena A y la cadena B, y
15 L es un péptido L N-terminal,
en donde L es un péptido L que comprende una secuencia de aminoácidos de acuerdo con la SEQ
ID NO: 2 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con
la SEQ ID NO: 2.
- 20 3. El precursor de insulina recombinante de acuerdo con la reivindicación 1 o 2, en donde C
es un péptido C que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1
o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ
ID NO: 1 y L comprende un péptido L que comprende una secuencia de aminoácidos de acuerdo
con la SEQ ID NO: 2 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de
25 secuencia con la SEQ ID NO: 2.
4. El precursor de insulina recombinante de acuerdo con cualquiera de las reivindicaciones
anteriores, el péptido C consiste en una secuencia de aminoácidos de acuerdo con la SEQ ID NO:
1 y el péptido L consiste en una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2.
- 30 5. El precursor de insulina recombinante de acuerdo con cualquiera de las reivindicaciones
anteriores, en donde la cadena A comprende una secuencia de aminoácidos de acuerdo con la
SEQ ID NO: 3 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de
secuencia con la SEQ ID NO: 3.

6. El precursor de insulina recombinante de acuerdo con cualquiera de las reivindicaciones anteriores, en donde la cadena B comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 4 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de
5 secuencia con la SEQ ID NO: 4.
7. El precursor de insulina recombinante de acuerdo con cualquiera de las reivindicaciones anteriores, en donde la cadena A se deriva de insulina porcina y/o la cadena B se deriva de insulina porcina.
- 10 8. El precursor de insulina recombinante de acuerdo con cualquiera de las reivindicaciones anteriores, en donde el precursor de insulina recombinante comprende una secuencia de aminoácidos de acuerdo con las SEQ ID NO: 6, 7 u 8 o una secuencia que tiene al menos 60 % de identidad de secuencia con las SEQ ID NO: 6, 7 y/o 8.
- 15 9. El precursor de insulina recombinante de acuerdo con cualquiera de las reivindicaciones anteriores, en donde el precursor de insulina recombinante consiste en una secuencia de aminoácidos de acuerdo con las SEQ ID NO: 6, 7 u 8.
- 20 10. Una secuencia de ADN que codifica un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde
A es una cadena A de insulina,
B es una cadena B de insulina,
C es un péptido C que conecta la cadena A y la cadena B, y
25 L es un péptido L N-terminal,
en donde C es un péptido C que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de
secuencia con la SEQ ID NO: 1 y/o L es un péptido L que comprende una secuencia de
aminoácidos de acuerdo con la SEQ ID NO: 2 o una secuencia de aminoácidos que tiene al
30 menos 60 % de identidad de secuencia con la SEQ ID NO: 2.
11. La secuencia de ADN de acuerdo con la reivindicación 8, en donde la secuencia de ADN está optimizada por codones para la expresión en un microorganismo

12. La secuencia de ADN de acuerdo con la reivindicación 9, en donde la secuencia de ADN está optimizada por codones para la expresión en *E. coli*.
13. La secuencia de ADN de acuerdo con cualquiera de las reivindicaciones 10-12, en donde
 - 5 A es una cadena A de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 11;
B es una cadena B de insulina, preferiblemente codificada por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 12;
C es un péptido C que conecta la cadena A y la cadena B, preferiblemente en donde el
 - 10 péptido C está codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 9; y
L es un péptido L N-terminal, preferiblemente codificado por una secuencia de ADN que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 10.
14. Un vector que comprende una secuencia de ADN de acuerdo con cualquiera de las reivindicaciones 10-13.
15. Una célula huésped que comprende una secuencia de ADN de acuerdo con cualquiera de las reivindicaciones 10-13 o un vector de acuerdo con la reivindicación 14.
16. La célula huésped de acuerdo con la reivindicación 14, en donde la célula huésped es una célula de *E. coli* y/o una célula de *S. cerevisiae*.
17. Un método para producir insulina recombinante, en donde se expresa un precursor de
 - 25 insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde
A es una cadena A de insulina,
B es una cadena B de insulina,
C es un péptido C que conecta la cadena A y la cadena B, y
L es un péptido L N-terminal,
en donde C es un péptido C que comprende una secuencia de aminoácidos de acuerdo con la
 - 30 SEQ ID NO: 1 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 1 y/o L es un péptido L que comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 2, y el método comprende:

- a) expresar el precursor de insulina recombinante en una célula huésped,
 - b) lisar la célula huésped para obtener el precursor de insulina recombinante,
 - c) replegar el precursor de insulina recombinante, y
 - d) escindir el péptido L y el péptido C del precursor de insulina recombinante mediante una
5 reacción enzimática para obtener insulina recombinante.
18. El método de acuerdo con la reivindicación 17, en donde el paso a) es precedido por un paso de clonar una secuencia de ADN que codifica el precursor de insulina recombinante en un vector.
- 10 19. El método de acuerdo con la reivindicación 17 o 18, en donde el método comprende, además, un paso de purificar la insulina recombinante.
20. Uso de un precursor de insulina recombinante de acuerdo con cualquiera de las
15 reivindicaciones 1-9 para la producción de insulina recombinante.
21. Una composición que comprende insulina recombinante, en donde la composición comprende una cantidad detectable de un péptido C y/o péptido L, en donde el péptido C comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 o una secuencia de
20 aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 1 y/o el péptido L comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2 o una secuencia de aminoácidos que tiene al menos 60 % de identidad de secuencia con la SEQ ID NO: 2.
- 25 22. Una composición de insulina recombinante que comprende una insulina recombinante producida por cualquiera de los métodos de las reivindicaciones 17-19 y un vehículo farmacéuticamente aceptable.
23. Un método para tratar la diabetes que comprende administrar a un animal que lo necesite
30 una cantidad terapéutica de la composición de insulina recombinante de la reivindicación 22.
24. La composición de insulina recombinante de la reivindicación 22 para uso en el tratamiento de la diabetes.

RESUMEN

La invención se refiere a precursores de insulina recombinante. Para mejorar la producción recombinante de insulina, se proporciona un precursor de insulina recombinante de acuerdo con la fórmula L-B-C-A, en donde A es una cadena A de insulina, B es una cadena B de insulina, C es un péptido C que conecta la cadena A y la cadena B, y L es un péptido L N-terminal. El

5 péptido C comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 1 y/o el péptido L comprende una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2.