

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 537 647**

51 Int. Cl.:

C12N 15/52 (2006.01)

C07K 14/34 (2006.01)

C12R 1/15 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.12.2005 E 10157889 (6)**

97 Fecha y número de publicación de la concesión europea: **11.03.2015 EP 2210946**

54 Título: **Mutante del gen proB de las bacterias corineformes**

30 Prioridad:

22.12.2004 DE 102004061696

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

10.06.2015

73 Titular/es:

**EVONIK DEGUSSA GMBH (100.0%)
Rellinghauser Strasse 1-11
45128 Essen, DE**

72 Inventor/es:

**HANS, STEPHAN;
BATHE, BRIGITTE, DR. y
THIERBACH, GEORG, DR.**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 537 647 T3

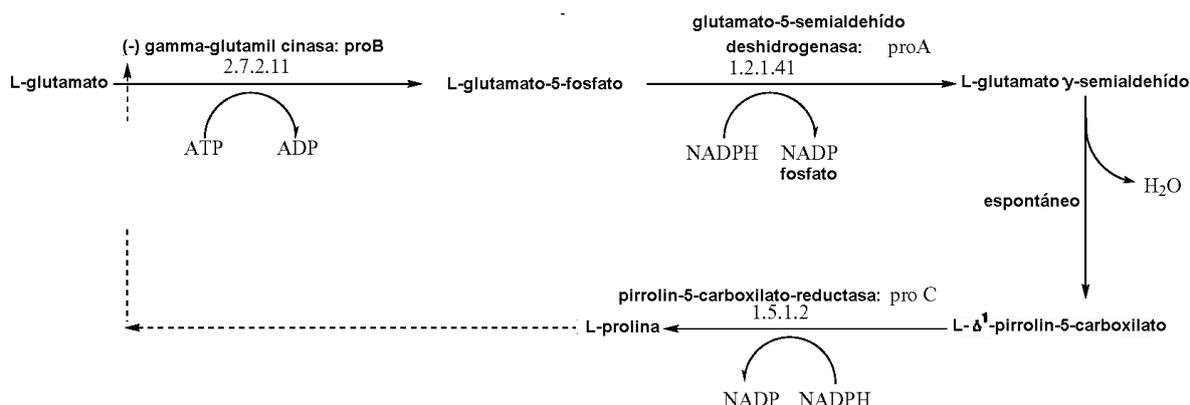
Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar 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

Mutante del gen proB de las bacterias corineformes

La presente invención se refiere a organismos hospedantes no humanos que comprenden un polinucleótido que codifica una enzima que tiene actividad de γ -glutamil cinasa. Más particularmente, la presente invención caracteriza organismos hospedantes no humanos que comprenden un polinucleótido que codifica aquellas enzimas que tienen un aminoácido proteinogénico distinto de glicina en la posición 149 de la secuencia de aminoácidos o en una posición comparable.

Los organismos hospedantes no humanos que comprenden un polinucleótido que codifica tales enzimas que tienen actividad de γ -glutamil cinasa se emplean preferiblemente en la producción por fermentación de L-prolina. En el esquema 1 se representa la ruta biosintética de la L-prolina, partiendo de glutamato.



Esquema 1

Es sabido que los aminoácidos se pueden producir fermentando cepas de, por ejemplo, bacterias corineformes, en particular *Corynebacterium glutamicum*. Debido a la gran importancia, se están realizando constantemente esfuerzos para mejorar los procesos de producción. Las mejoras en el procedimiento pueden implicar medidas de relacionadas con la tecnología de la fermentación, tales como, por ejemplo, la agitación y el suministro de oxígeno, o la composición de los medios nutrientes, tal como, por ejemplo, la concentración de azúcar durante la fermentación, o el tratamiento hasta obtener la forma del producto, por ejemplo por medio de cromatografía de intercambio iónico, o las propiedades intrínsecas de comportamiento del microorganismo propiamente dicho.

Las propiedades de comportamiento de estos microorganismos se mejoran aplicando métodos de mutagénesis, selección y elección de mutantes. Esto da como resultado cepas que son resistentes a antimetabolitos o son auxótrofas para metabolitos importantes para la regulación y que producen aminoácidos. Un antimetabolito conocido es el análogo de prolina 3,4-deshidro-DL-prolina (DHP).

Desde hace varios años se han empleado igualmente métodos de ADN recombinante para mejorar las cepas de *Corynebacterium* que producen L-aminoácidos, amplificando genes de la biosíntesis de aminoácidos individuales e investigando el efecto sobre la producción de aminoácidos.

La secuencia nucleotídica del genoma de *Corynebacterium glutamicum* se describe, entre otros, en el documento EP-A-1108790, y también se ha depositado en la base de datos del National Center for Biotechnology Information (NCBI) de la National Library of Medicine (Bethesda, MD, USA) con los números de acceso NC_003450.2 y BX927148.1 a BX927157.1.

Sleator et al. dan a conocer la posibilidad de provocar sobreproducción en la biosíntesis de prolina mutando el gen proB de *Listeria monocytogenes* (Appl. Environ. Microbiol. 2001, 67, 4560-5). Las mutaciones especificadas allí y declaradas como exitosas implican genes de proB que codifican γ -glutamil cinasas que tienen las siguientes mutaciones: V121I, A144V y E146K. Además, se hace mención al hecho de que las regiones en las que se producen estas mutaciones corresponden muy bien a la región también identificada en otros organismos como una región diana para mutaciones ventajosas.

Por lo tanto, fue el objeto de la presente invención hacer que estén disponibles organismos hospedantes no humanos que comprenden un polinucleótido que codifica otras variantes proteicas mutadas y, cuando sea apropiado, mejoradas de una γ -glutamil cinasa, que se pueden emplear ventajosamente en un proceso técnico para la producción de L-prolina por fermentación.

Este objeto y otros objetos que no se especifican con detalle pero que surgen de manera obvia a partir de la técnica anterior se logran partiendo de organismos hospedantes no humanos que comprenden un polinucleótido que

codifica las γ -glutamil cinasas de la reivindicación 1. Las reivindicaciones 2 a 8 se centran en organismos hospedantes no humanos que comprenden un polinucleótido que codifica enzimas preferidas de este tipo. Finalmente, las reivindicaciones 12 a 14 se refieren a un proceso de producción según la invención para la L-prolina con la ayuda de organismos hospedantes no humanos que comprenden un polinucleótido que codifica las enzimas mencionadas.

Proporcionando organismos hospedantes no humanos que comprenden un polinucleótido que codifica una γ -glutamil cinasa (producto del gen proB) que tiene un aminoácido proteínogénico distinto de glicina en la posición 149 de los aminoácidos o en una posición comparable, sorprendentemente, aunque no menos ventajosamente, se logra de forma particular el objeto expuesto. En comparación con las enzimas de tipo salvaje, las γ -glutamil cinasas que tienen una mutación apropiada ayudan a producir L-prolina de manera mejorada en un proceso de producción por fermentación. Usando los métodos según la invención, es posible mejorar el comportamiento de los organismos hospedantes no humanos o del proceso de fermentación con respecto a uno o más de los parámetros seleccionados del grupo de concentración de producto (producto por volumen), rendimiento de producto (producto formado por fuente de carbono consumida) y formación de producto (producto formado por volumen y tiempo), o cualesquiera otros parámetros del proceso y sus combinaciones en al menos 0,5%, al menos 1%, al menos 1,5% o al menos 2%, basándose en la cepa de partida o cepa progenitora o el proceso de fermentación con el uso de organismos hospedantes no humanos que comprenden un polinucleótido que codifica dichas enzimas.

Se da preferencia a proporcionar organismos hospedantes no humanos que comprenden un polinucleótido que codifica γ -glutamil cinasas en las que un aminoácido, preferiblemente el L-aminoácido, seleccionado del grupo que consiste en Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Ala, Val, Gln, His, Pro, Leu, Tyr, Trp, Cys o Phe está presente en la posición del aminoácido como se define según la invención o en una posición comparable. (Los aminoácidos mencionados, incluyendo glicina, también se denominan en la técnica como aminoácidos proteínogénicos). Se da una preferencia muy particular a la sustitución de glicina con ácido L-aspártico en la posición 149 (G149D) en la mencionada enzima. Se sabe que las enzimas intrínsecas al hospedante no humano, denominadas aminopeptidasas, son capaces de escindir el aminoácido N-terminal metionina, separándolo de la proteína formada. Además, se sabe que la escisión de uno (1) o dos (2) y de no más de tres (3) aminoácidos del término C de la proteína altera la actividad enzimática sólo de forma insignificante como mucho, si es que lo hace. Sin embargo, la enzima puede aumentar en longitud debido a la anexión de ciertas proteínas de fusión (véase más abajo).

La presente invención también engloba organismos hospedantes no humanos que comprenden un polinucleótido que codifica la secuencia de aminoácidos de SEC ID NO: 2, preferiblemente 4, o una secuencia que es al menos 90% idéntica a ella, comprendiendo las secuencias la sustitución de aminoácidos de glicina por otro aminoácido, en particular cualquiera de los aminoácidos preferidos mencionados, en la posición 149 o en una posición comparable. De este modo, la invención también engloba organismos no humanos que comprenden un polinucleótido que codifica aquellas enzimas que tienen los grados mencionados anteriormente de identidad a nivel de aminoácidos en comparación con SEC ID NO: 2, preferiblemente 4. Igualmente, estas enzimas se pueden originar a partir de fuentes naturales. Como alternativa, se pueden haber modificado mediante tecnología de ADN recombinante, de tal manera que el experto puede predecir la actividad enzimática a retener o esencialmente retenida (véase, por ejemplo, Sambrook et al, "Molecular Cloning, A Laboratory Handbook", 2ª edición 1989, CSH Press, Cold Spring Harbor, Ausubel et al. "Current Protocols in Molecular Biology", John Wiley & Sons, NY 2001). De este modo, los aminoácidos que no están presentes en el sitio activo y cuya sustitución por un aminoácido "del mismo tipo" no se espera que dé como resultado, a primera vista, una estructura tridimensional sustancialmente alterada se pueden sustituir por un aminoácido "del mismo tipo". Por ejemplo, se puede esperar que ciertos aminoácidos con cadenas laterales no polares (aminoácidos del mismo tipo) se puedan sustituir, por ejemplo isoleucina por valina, sin que esto tenga una influencia (sustancial) sobre la función biológica o enzimática de la enzima según la invención, o sobre la actividad enzimática. El experto puede alcanzar, basándose en su conocimiento, conclusiones correspondientes también para la sustitución de otros tipos de aminoácidos (por ejemplo la sustitución de aminoácidos básicos por otros aminoácidos básicos, o de aminoácidos con cadenas laterales polares no cargadas por otros aminoácidos de este grupo).

Igualmente se da preferencia a organismos hospedantes no humanos que comprenden un polinucleótido que codifica una γ -glutamil cinasa que tiene al menos la secuencia de aminoácidos correspondiente a las posiciones 145 a 154, o a posiciones comparables, más preferiblemente 130 a 169, y preferiblemente de forma muy particular 110 a 189 de SEC ID NO: 2, preferiblemente 4, siendo posible que la longitud de la secuencia de aminoácidos de γ -glutamil cinasa corresponda a los números indicados anteriormente.

En una realización adicionalmente preferida, los organismos hospedantes no humanos que comprenden un polinucleótido que codifica las enzimas según la invención comprenden adicionalmente en esas enzimas al menos una sección de aminoácidos heteróloga, por medio de la cual estos polipéptidos se caracterizan como proteínas de fusión. Los ejemplos de componentes heterólogos de la proteína de fusión codificada por polinucleótidos comprendidos en los organismos hospedantes no humanos según la invención pueden ser marcadores (por ejemplo marcador de His o marcador de Flag) que se pueden emplear en la purificación de las proteínas de fusión codificada por polinucleótidos comprendidos en los organismos hospedantes no humanos según la invención. En otras realizaciones, los componentes heterólogos pueden tener una actividad enzimática separada. En tal caso, los

dos componentes enzimáticos están conectados preferiblemente mediante un enlazador, tal como un enlazador de glicina o de glicina-serina flexible de 6-10 aminoácidos de longitud, a fin de asegurar la funcionalidad de los componentes. El término "heterólogo", como se usa aquí, puede significar, por un lado, que los componentes de la proteína de fusión no se producen de forma natural juntos en una forma enlazada covalentemente, y, por otro lado, que los componentes derivan de diferentes especies. Las proteínas de fusión se preparan habitualmente por medio de tecnología de ADN recombinante (véase Sambrook et al., en el lugar citado).

En una realización adicional, la presente invención se refiere a organismos hospedantes no humanos que comprenden una secuencia nucleotídica (secuencia de ácidos nucleicos) que codifica una γ -glutamil cinasa como se especifica anteriormente. En consecuencia, la invención también se refiere a organismos hospedantes no humanos que comprenden secuencias nucleotídicas replicables que codifican la enzima γ -glutamil cinasa, siendo posible que las secuencias de aminoácidos correspondientes codificadas por estas secuencias nucleotídicas tengan cualquier aminoácido proteínogénico, excepto glicina, en la posición 149 o en una posición comparable.

Las secuencias nucleotídicas comprendidas en los organismos hospedantes no humanos de la invención codifican preferiblemente una γ -glutamil cinasa, comprendiendo la secuencia de aminoácidos codificada correspondiente un aminoácido seleccionado del grupo que consiste en Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Ala, Val, Gln, His, Pro, Leu, Tyr, Trp, Cys o Phe en la posición 149 o en una posición comparable. Se prefiere de forma muy particular organismos hospedantes no humanos que comprenden una secuencia nucleotídica que codifica una γ -glutamil cinasa que tiene un ácido L-aspártico en la posición 149 o en una posición comparable.

La invención se refiere igualmente a organismos hospedantes no humanos que comprenden secuencias de ácidos nucleicos replicables que codifican una γ -glutamil cinasa que tiene la sustitución de aminoácidos según la invención en la posición 149 o en una posición comparable, codificando estas secuencias,

d) una γ -glutamil cinasa según la invención, que tiene la secuencia de aminoácidos de SEC ID NO: 2, preferiblemente SEC ID NO: 4, al menos en las posiciones 145 a 154 o en posiciones comparables y que se hibrida en condiciones experimentales restrictivas a la secuencia nucleotídica complementaria a SEC ID NO: 3.

La invención también se refiere a organismos hospedantes no humanos que comprenden una secuencia nucleotídica como se representa en SEC ID NO: 3. Un ejemplo de una secuencia nucleotídica que es al menos 70% idéntica a aquella de SEC ID NO: 3 se muestra en SEC ID NO: 5. SEC ID NO: 6 muestra la secuencia de aminoácidos de γ -glutamil cinasa, codificada por SEC ID NO: 5.

La invención se refiere igualmente a organismos hospedantes no humanos que comprenden secuencias de ácidos nucleicos replicables que codifican γ -glutamil cinasas que incluyen preferiblemente al menos la secuencia de aminoácidos correspondiente a las posiciones 130-169 o en posiciones comparables, y muy particularmente de forma preferible las posiciones 110 a 189 de SEC ID NO: 2, preferiblemente 4, y que se hibrida en condiciones experimentales restrictivas con la secuencia nucleotídica complementaria a SEC ID NO: 3. El término "complementaria" significa, según la invención, que la complementariedad se extiende sin saltos a lo largo de toda la región de la molécula de ácido nucleico según la invención. En otras palabras, según la invención, se da preferencia a una complementariedad que se extiende 100% a lo largo de toda la región de la secuencia según la invención, es decir, desde el término 5' representado al término 3' representado, en particular las regiones codificantes (cfs). En otras realizaciones preferidas, la complementariedad se extiende a lo largo de una región de al menos 19, preferiblemente al menos 21, nucleótidos sucesivos que preferiblemente no codifican el sitio activo de actividad enzimática.

Se da preferencia a aquellas secuencias de ácidos nucleicos derivadas de las bacterias corineformes, preferiblemente *Corynebacterium glutamicum*. Las secuencias de ácidos nucleicos de genes o alelos, que están presentes en la población de una especie, también se denominan en la técnica como genes o alelos endógenos.

La presente invención se refiere además a organismos hospedantes no humanos recombinantes (rec) que tienen las secuencias nucleotídicas según la invención. Los organismos hospedantes no humanos adecuados son cualesquiera realizaciones consideradas para este fin por el experto.

Los ejemplos de organismos hospedantes no humanos que se pueden mencionar a este respecto son levaduras tales como *Hansenula polymorpha*, *Pichia sp.*, *Saccharomyces cerevisiae*, procariotas tales como *E. coli*, *Bacillus subtilis*, bacterias corineformes tales como *Corynebacterium glutamicum*, o eucariotas tales como células de mamíferos no humanos, células de insectos o células vegetales. Estos organismos acumulan, cuando sea apropiado ya antes de las medidas de la presente invención, L-prolina en sus células o en el medio de fermentación que las rodea. De este modo, en esta realización preferida, el hospedante no humano según la invención es una célula recombinante que se ha transformado o transfectado con una secuencia de ácidos nucleicos según la invención o un vector según la invención (véase más abajo), o se ha proporcionado con ellos por medio de conjugación (los términos "transformación", "transfección" y "conjugación" se usan de forma sinónima según la presente invención). La transformación y transfección, respectivamente, se pueden llevar a cabo según métodos conocidos, por ejemplo por medio de coprecipitación con fosfato de calcio, lipofección, electroporación, bombardeo con partículas o

infección vírica. La célula según la invención puede contener el ácido nucleico recombinante extracromosómicamente, o en una forma integrada cromosómicamente. En otras palabras, la transfección/transformación puede ser una estable o una transitoria. Los procesos para la clonación son bien conocidos por el experto (Sambrook, J.; Fritsch, E. F. y Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2ª ed., Cold Spring Harbor Laboratory Press, Nueva York). Las cepas de *E. coli* se pueden utilizar para la preparación recombinante y métodos de mutagénesis, que incluyen entre otras: *E. coli* XL1 Blue, NM 522, JM101, JM109, JM105, RR1, DH5 α , TOP10, HB101, BL21 codon plus, BL21 (DE3) codon plus, BL21, BL21 (DE3), MM294. Igualmente, los plásmidos usados entre otros para la clonación del constructo génico que tiene el ácido nucleico según la invención en el organismo hospedante no humano son conocidos por el experto (véase también el documento PCT/EP03/07148; véase más abajo).

Un hospedante no humano del género *Corynebacterium*, del cual se puede hacer mención particular, es la especie *Corynebacterium glutamicum* que es conocido en la técnica. Los ejemplos de cepas de tipo salvaje conocidas del género *Corynebacterium* son:

- Corynebacterium glutamicum* ATCC13032
- Corynebacterium acetoglutamicum* ATCC15806
- Corynebacterium acetoacidophilum* ATCC13870
- Corynebacterium efficiens* DSM 44549
- Corynebacterium melassecola* ATCC17965
- Corynebacterium thermoaminogenes* FERM BP-1539
- Brevibacterium flavum* ATCC14067
- Brevibacterium lactofermentum* ATCC13869 y
- Brevibacterium divaricatum* ATCC14020.

La información relativa a la clasificación taxonómica de cepas de este grupo de bacterias se puede encontrar, entre otros, en Kämpfer y Kroppenstedt (*Canadian Journal of Microbiology* 42, 989-1005 (1996)) y en el documento US-A-5.250.434. Desde hace ya algunos años (Liebl et al., *International Journal of Systematic Bacteriology* 41(2), 255-260 (1991)), las bacterias corineformes con los nombres de especies "*Brevibacterium flavum*", "*Brevibacterium lactofermentum*" y "*Brevibacterium divaricatum*" se han clasificado bajo la especie *Corynebacterium glutamicum*. Igualmente, las bacterias corineformes con el nombre de especie "*Corynebacterium melassecola*" pertenecen a la especie *Corynebacterium glutamicum*.

Los ejemplos de cepas de bacterias corineformes productoras de L-prolina son las cepas:

- Brevibacterium lactofermentum* NRRL B-11421,
- Brevibacterium flavum* NRRL B-11422,
- Corynebacterium glutamicum* NRRL B-11423,
- Microbacterium ammoniaphilum* NRRL B-11424,
- Corynebacterium glutamicum* ATCC 21157,
- Corynebacterium glutamicum* ATCC 21158,
- Corynebacterium glutamicum* ATCC 21159,
- Corynebacterium glutamicum* ATCC 21355,
- Corynebacterium acetophilum* FERM-P 4045,
- Corynebacterium acetoacidophilum* FERM-P 4962,
- Arthrobacter citreus* FERM-P 4963, y
- Microbacterium ammoniaphilum* FERM-P 4964,

todas las cuales se describen en los documentos US 4.224.409 y US 4.444.885.

En los organismos hospedantes no humanos según la invención, las secuencias de ácidos nucleicos según la invención están preferiblemente enlazadas de forma operativa dentro de un vector a una secuencia de control de la

expresión para que se puedan transcribir y, cuando sea apropiado, traducir en una célula hospedante no humana adecuada. Las secuencias de control de la expresión comprenden habitualmente un promotor y, cuando sea apropiado, otras secuencias reguladoras tales como operadores o potenciadores. Además, también es posible que estén presentes secuencias de iniciación de la traducción. El experto en la técnica conoce secuencias de control de la expresión adecuadas para células hospedantes no humanas procariontas o eucariotas (véase, por ejemplo, Sambrook et al., en el lugar citado). El vector recombinante puede además incluir también elementos habituales tales como un origen de la replicación y un gen marcador de selección. Los ejemplos de vectores recombinantes adecuados son plásmidos, cósmidos, fagos, o virus (véase, por ejemplo, Sambrook et al., más arriba). Los plásmidos adecuados son en principio cualesquiera realizaciones disponibles para este fin al experto. Tales plásmidos se pueden encontrar, por ejemplo, en la publicación de Studier y colaboradores (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendorff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol.* 185, 61-89) o en los folletos de Novagen, Promega, New England Biolabs, Clontech o Gibco BRL. Otros plásmidos y vectores preferidos se pueden encontrar en: Glover, D. M. (1985), *DNA cloning: a practical approach*, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. y Denhardt, D. T (eds) (1988), *Vectors: a survey of molecular cloning vectors and their uses*, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), *Systems for heterologous gene expression*, *Methods Enzymol.* 185, 3-7; Sambrook, J.; Fritsch, E. F. y Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2ª ed., Cold Spring Harbor Laboratory Press, Nueva York.

Los plásmidos que se pueden usar para clonar en el organismo hospedante no humano los constructos génicos que tienen las secuencias de ácidos nucleicos contempladas son o se basan en, de una manera muy particularmente preferida: pUC18/19 (Roche Biochemicals), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pCR (Invitrogen), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) o pET (Novagen). Otros plásmidos preferidos son pBR322 (DSM3879), pACYC184 (DSM4439) y pSC101 (DSM6202), todos los cuales se pueden obtener de la DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Brunswick, Alemania. Los ejemplos de promotores preferidos son el promotor T7, el promotor lac, el promotor tac, el promotor trp, el promotor rha y el promotor ara.

Preferiblemente, las γ -glutamil cinasas codificadas por polinucleótidos comprendidos en los organismos hospedantes no humanos según la invención, o los ácidos nucleicos que las codifican, se sobreexpresan en bacterias corineformes, preferiblemente del género *Corynebacterium*, particularmente de forma preferible de la especie *Corynebacterium glutamicum*.

Sobreexpresión significa un incremento en la concentración intracelular o actividad de las γ -glutamil cinasas codificadas por polinucleótidos comprendidos en los organismos hospedantes no humanos según la invención.

Las medidas de la sobreexpresión incrementan la actividad o concentración de la proteína correspondiente codificada por polinucleótidos comprendidos en los organismos hospedantes no humanos según la invención en al menos 10%, 25%, 50%, 75%, 100%, 150%, 200%, 300%, 400% o 500%, hasta un máximo de 1000% o 2000%, basado en la actividad o concentración de la proteína en el microorganismo de partida.

La sobreexpresión se logra según la invención mutando la región promotora y la región reguladora o el sitio de unión a ribosoma, que está situado en dirección 5' del gen estructural. Los casetes de expresión incorporados en dirección 5' del gen estructural actúan de la misma manera. Además, los promotores inducibles hacen posible incrementar la expresión durante la producción de L-prolina por fermentación. Las medidas para extender la duración del ARNm mejoran igualmente la expresión. Además, la actividad enzimática se amplifica evitando la degradación de la proteína enzimática. Los genes o constructos génicos pueden estar presentes en plásmidos o en el cromosoma en una forma integrada y amplificada.

Los plásmidos que se replican en bacterias corineformes son útiles para la invención. Numerosos vectores plasmídicos conocidos, tales como, por ejemplo, pZ1 (Menkel et al., *Applied and Environmental Microbiology* (1989) 64: 549-554), pEKEx1 (Eikmanns et al., *Gene* 102:93-98 (1991)) o pHS2-1 (Sonnen et al., *Gene* 107:69-74 (1991)) se basan en los plásmidos crípticos pHM1519, pBL1 o pGA1. De la misma manera, se pueden usar otros vectores plasmídicos, tales como, por ejemplo, los basados en pCG4 (documento US-A 4.489.160), o pNG2 (Serwold-Davis et al., *FEMS Microbiology Letters* 66, 119-124 (1990)) o pAG1 (documento US-A 5.158.891). En Tauch et al. (*Journal of Biotechnology* 104(1-3), 27-40 (2003)) se puede encontrar un repaso sobre vectores plasmídicos de *Corynebacterium glutamicum*.

Los métodos de mutagénesis descritos en la técnica anterior se usan para generar la sustitución de aminoácidos según la invención en γ -glutamil cinasas (por ejemplo SEC ID NO: 2) y otras mutaciones de proB según la invención, caracterizados por una sustitución de aminoácidos en la posición 149. Los métodos de mutagénesis adecuados son cualesquiera métodos disponibles para este fin para el experto.

Para la mutagénesis, se pueden usar métodos de mutagénesis in vivo clásicos que usan sustancias mutagénicas tales como, por ejemplo, N-metil-N'-nitro-N-nitrosoguanidina (MNNG) o luz ultravioleta. Las células mutagenizadas se aplican subsiguientemente, cuando sea apropiado, a un agar mínimo que contiene 3,4-deshidro-DL-prolina a concentraciones de aprox. 0,5-1 g/l, aprox. 1-2 g/l o aprox. 2-3 g/l. Los mutantes individuales se aíslan, y la

secuencia nucleotídica del gen o alelo proB se determina, cuando sea apropiado, después de un proceso de clonación previo.

Además, es posible usar para la mutagénesis métodos in vitro tales como, por ejemplo, el tratamiento con hidroxilamina (Miller, J. H.: A Short Course in Bacterial Genetics. A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1992) u oligonucleótidos mutagénicos (T. A. Brown: Gentechnologie für Einsteiger [ingeniería genética para principiantes], Spektrum Akademischer Verlag, Heidelberg, 1993; libro de texto por Knippers ("Molekulare Genetik" [genética molecular], 6ª edición, Georg Thieme Verlag, Stuttgart, Alemania, 1995); Winnacker ("Gene und Klone" [genes y clones], VCH Verlagsgesellschaft, Weinheim, Alemania, 1990); Hagemann ("Allgemeine Genetik" [genética general], Gustav Fischer Verlag, Stuttgart, Alemania, 1986)) o la reacción en cadena de la polimerasa (PCR) como se describe en el manual de Newton y Graham (PCR, Spektrum Akademischer Verlag, Heidelberg, Alemania, 1994). Incluyen, en particular, métodos de mutagénesis de saturación, mutagénesis aleatoria, de recombinación in vitro y mutagénesis dirigida al sitio (Eigen, M. y Gardiner, W. (1984), Evolutionary molecular engineering based on RNA replication, Pure Appl. Chem. 56, 967-978; Chen, K. y Arnold, F. (1991), Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. Bio/Technology 9, 1073-1077; Horwitz, M. y Loeb, L. (1986), Promoters Selected From Random DNA-Sequences, Proc Natl Acad Sci USA 83, 7405-7409; Dube, D. y L. Loeb (1989), Mutants Generated By The Insertion Of Random Oligonucleotides Into The Active-Site Of The Beta-Lactamase Gene, Biochemistry 28, 5703-5707; Stemmer, P.C. (1994), Rapid evolution of a protein in vitro by DNA shuffling, Nature 370, 389-391 y Stemmer, P.C. (1994), DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. Proc Natl Acad Sci USA 91, 10747-10751).

El uso de métodos in vitro implica amplificar el gen proB descrito en la técnica anterior con la ayuda de la reacción en cadena de la polimerasa, partiendo de ADN total aislado de una cepa de tipo salvaje, clonar dicho gen, cuando sea apropiado, en vectores plasmídicos adecuados, y después someter al ADN a un proceso de mutagénesis. El experto puede encontrar instrucciones sobre la amplificación de secuencias de ADN con la ayuda de la reacción en cadena de la polimerasa (PCR), entre otros, en el manual de Gait: Oligonucleotide Synthesis: A Practical Approach (IRL Press, Oxford, UK, 1984) y en Newton y Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Alemania, 1994). De la misma manera, también es posible usar métodos de mutagénesis in vitro, como se describe, por ejemplo, en el manual bien conocido de Sambrook et al. (Molecular Cloning, A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, USA, 1989). Los métodos correspondientes también están comercialmente disponibles en forma de "kits", tales como, por ejemplo, el "QuikChange Site-Directed Mutagenesis Kit" de Stratagene (La Jolla, USA), descrito por Papworth et al. (Strategies 9(3), 3-4 (1996)). Los mutantes de proB adecuados se seleccionan subsiguientemente y se investigan mediante los métodos descritos anteriormente.

Para la producción de L-prolina, puede ser ventajoso, además de usar organismos hospedantes no humanos que comprenden un polinucleótido que codifica las γ -glutamil cinasas según la invención, que han sido sobreexpresadas o amplificadas, para amplificar, en particular para sobreexpresar, además de dichas cinasas, al mismo tiempo una o más enzimas de la biosíntesis de la prolina. Habitualmente se da preferencia al uso de genes endógenos.

"Genes endógenos" o "secuencias nucleotídicas endógenas" significan los genes o secuencias nucleotídicas y alelos presentes en la población de una especie.

En este contexto, el término "amplificación" describe el incremento en la actividad intracelular o concentración de una o más enzimas o proteínas en un microorganismo, que son codificadas por el ADN correspondiente, por ejemplo incrementando el número de copias del gen o genes, usando un promotor potente, o usando un gen o alelo que codifica una enzima o proteína correspondiente con actividad elevada y, cuando sea apropiado, combinando estas medidas.

Las medidas de la amplificación, en particular la sobreexpresión, incrementan la actividad o concentración de la enzima o proteína correspondiente habitualmente en al menos 10%, 25%, 50%, 75%, 100%, 150%, 200%, 300%, 400% o 500%, hasta un máximo de 1000% o 2000%, basado en la de la proteína de tipo salvaje o en la actividad o concentración de la proteína en el microorganismo de partida.

De este modo, para producir L-prolina es posible, además de usar la variante del gen proB según la invención, para amplificar, en particular sobreexpresar, uno o más de los genes seleccionados del grupo que consiste en

- el gen gdh que codifica glutamato deshidrogenasa (EC 1.4.1.4),
- el gen proA que codifica γ -glutamil-fosfato reductasa (EC 1.2.1.41),
- el gen proC que codifica la pirrolin-5-carboxilato reductasa (EC 1.5.1.2), y
- el gen ocd que codifica ornitina ciclodesaminasa (EC 4.3.1.12).

Para la producción de L-prolina, adicionalmente puede ser ventajoso, además de usar mutantes según la invención del gen proB, al mismo tiempo atenuar, en particular reducir la expresión de uno o más de los genes endógenos seleccionados del grupo que consiste en

- el gen ilvA que codifica treonina desaminasa (EC 4.2.1.16),
- 5 • el gen putA que codifica prolina deshidrogenasa/pirrolin-5-carboxilato deshidrogenasa (EC 1.5.99.8),
- el gen sucA que codifica 2-cetoglutarato deshidrogenasa (EC 1.2.4.2),
- el gen sucB que codifica dihidrolipoamida succiniltransferasa (EC 2.3.1.61), y
- el gen argD que codifica acetilornitina aminotransferasa (EC 2.6.1.11).

10 A este respecto, el término “atenuación” describe la reducción o eliminación de la actividad intracelular o concentración de una o más enzimas o proteínas que son codificadas por el ADN correspondiente en un microorganismo usando, por ejemplo, un promotor débil, o usando un gen o alelo que codifica una enzima correspondiente con baja actividad, o inactivando el gen o enzima o proteína correspondiente, y, cuando sea apropiado, combinando estas medidas.

15 La atenuación se puede lograr reduciendo o eliminando la expresión de los genes o las propiedades catalíticas o reguladoras de las proteínas enzimáticas. Ambas medidas se pueden combinar cuando sea apropiado.

La expresión génica se puede reducir mediante un proceso de cultivo adecuado o mediante modificación genética (mutación) de las estructuras de las señales de la expresión génica. Los ejemplos de las estructuras de señales de la expresión génica son genes represores, genes activadores, operadores, promotores, atenuadores, sitios de unión a ribosoma, el codón de partida, y terminadores. La información sobre esto la puede encontrar el experto en, por ejemplo, la solicitud de patente WO 96/15246, en Boyd y Murphy (Journal of Bacteriology 170: 5949-5952 (1988)), en Voskuil y Chambliss (Nucleic Acids Research 26: 3584-3590 (1998)), en Pátek et al. (Microbiology 142: 1297-309 (1996) y Journal of Biotechnology 104: 311-323 (2003)), y en libros de texto conocidos de genética y biología molecular, tales como, por ejemplo, el libro de texto de Knippers (“Molekulare Genetik”, 6ª edición, Auflage, Georg Thieme Verlag, Stuttgart, Alemania, 1995), o el de Winnacker (“Gene und Klone”, VCH Verlagsgesellschaft, Weinheim, Alemania, 1990).

Un ejemplo de la regulación específica de la expresión génica es la clonación del gen a atenuar para ponerlo bajo el control de un promotor inducible añadiendo cantidades medidas de IPTG (isopropil-β-D-tiogalactopiranosido), tal como, por ejemplo, el promotor trc o el promotor tac. Los vectores útiles aquí son, por ejemplo, el vector de expresión de Escherichia coli pXK99E (documento WO 0226787; depositado según el Tratado de Budapest el 31 de julio de 2001 en DH5alpha/pXK99E como DSM14440 con el Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSMZ, Brunswick, Alemania)) o pVWEx2 (Wendisch, tesis doctoral, Berichte des Forschungszentrums Jülich, Jül-3397, ISSN 0994-2952, Jülich, Alemania (1997)), que permiten que el gen clonado sea expresado en Corynebacterium glutamicum de una manera dependiente de IPTG.

Este método se empleó, por ejemplo, en la patente WO 02/26787 para la expresión regulada del gen deaD integrando el vector pXK99EdeaD en el genoma de Corynebacterium glutamicum, y por Simic et al. (Applied and Environmental Microbiology 68: 3321-3327 (2002)) para la expresión regulada del gen glyA mediante integración del vector pK18mobglyA' en Corynebacterium glutamicum.

Otro método para reducir específicamente la expresión génica es la técnica antisentido, que implica introducir en las células diana oligorribonucleótidos u oligodesoxirribonucleótidos o vectores cortos para sintetizar ARN antisentido más largo. Allí, el ARN antisentido se puede unir a secciones complementarias de los ARNm específicos y reducir su estabilidad o bloquear la capacidad de traducción. Un ejemplo de esto lo puede encontrar el experto en Srivastava et al. (Applied Environmental Microbiology 2000 Oct; 66 (10): 4366 - 4371).

Las mutaciones que dan como resultado un cambio o reducción en las propiedades catalíticas de proteínas enzimáticas son conocidas de la técnica anterior; los ejemplos que se pueden mencionar son los estudios de Qiu y Goodman (Journal of Biological Chemistry 272: 8611-8617 (1997)), Sugimoto et al. (Bioscience Biotechnology and Biochemistry 61: 1760-1762 (1997)) y Möckel (Tesis doctoral, Berichte des Forschungszentrums Jülich, Jül-2906, ISSN09442952, Jülich, Alemania (1994)). Se pueden encontrar visiones generales en libros de texto conocidos de genética y biología molecular, por ejemplo el de Hagemann (“Allgemeine Genetik”, Gustav Fischer Verlag, Stuttgart, Alemania, 1986).

Las mutaciones que reciben consideración son transiciones, transversiones, inserciones y supresiones. Dependiendo del efecto de la sustitución de aminoácidos sobre la actividad enzimática, se hace referencia a mutaciones de pérdida de sentido o a mutaciones sin sentido. Una mutación sin sentido conduce a que al menos un codón de parada esté localizado en la región codificante del gen, y que en consecuencia se termine prematuramente la traducción. Las inserciones o supresiones de al menos un par de bases en un gen conducen a mutaciones de desplazamiento del marco, como resultado de lo cual se incorporan aminoácidos incorrectos o la traducción se

termina prematuramente. Las supresiones de uno o más codones conducen típicamente a una pérdida total de actividad enzimática. Las instrucciones para generar tales mutaciones pertenecen a la técnica anterior y se pueden encontrar en libros de texto conocidos de genética y biología molecular, tal como el libro de texto de Knippers ("Molekulare Genetik", 6ª edición, Georg Thieme Verlag, Stuttgart, Alemania, 1995), el de Winnacker ("Gene und Klone", VCH Verlagsgesellschaft, Weinheim, Alemania, 1990) o el de Hagemann ("Allgemeine Genetik", Gustav Fischer Verlag, Stuttgart, 1986).

Como resultado del uso de las medidas para lograr la atenuación, normalmente la actividad o concentración de la proteína correspondiente se reduce hasta una cantidad de 0 a 75%, de 0 a 50%, de 0 a 25%, de 0 a 10%, de 0 a 5% o de 0 a 1%, de la actividad o concentración de la proteína de tipo salvaje, o de la actividad o concentración de la proteína en el microorganismo de partida.

Los microorganismos preparados según la invención, que son igualmente una materia objeto de la presente invención, se pueden cultivar de forma continua o discontinua, en un proceso por lotes, o un proceso por lotes alimentados, o un proceso por lotes alimentados repetidos, con el fin de producir L-prolina. Un sumario de los métodos de cultivo conocidos se describe en el libro de texto de Chmiel (Bioprozesstechnik 1. Einführung in die Bioverfahrenstechnik [Tecnología de Bioprocesos 1. Introducción de Tecnología de Bioprocesos] (Gustav Fischer Verlag, Stuttgart, 1991)), o en el libro de texto de Storhas (Bioreaktoren und periphere Einrichtungen [Biorreactores y Equipo Periférico] (Vieweg Verlag, Brunswick/Wiesbaden, 1994)).

El medio de cultivo a usar debe satisfacer de forma adecuada los requisitos de las cepas particulares. En el manual "Manual of Methods for General Bacteriology", publicado por la American Society for Bacteriology (Washington D.C., USA, 1981) se dan las descripciones de medios para cultivar diversos organismos.

La fuente de carbono empleada puede ser azúcares e hidratos de carbono, tales como, por ejemplo, glucosa, sacarosa, lactosa, fructosa, maltosa, molasas, almidón y celulosa, aceites y grasas, tales como, por ejemplo, aceite de haba de soja, aceite de girasol, aceite de cacahuete y aceite de coco, ácidos grasos tales como, por ejemplo, ácido palmítico, ácido esteárico y ácido linoleico, alcoholes tales como, por ejemplo, glicerol y etanol, alcoholes de azúcares tales como, por ejemplo, ribitol o manitol, y ácidos orgánicos tales como, por ejemplo, ácido acético. Estas sustancias se pueden usar individualmente o como mezclas.

La fuente de nitrógeno empleada puede ser compuestos orgánicos que contienen nitrógeno, tales como peptonas, extracto de levadura, extracto de carne, extracto de malta, licor de maíz fermentado, harina de haba de soja y urea, o compuestos inorgánicos tales como sulfato de amonio, cloruro de amonio, fosfato de amonio, carbonato de amonio y nitrato de amonio. Las fuentes de nitrógeno se pueden usar individualmente o como mezclas.

La fuente de fósforo empleada puede ser ácido fosfórico, dihidrogenofosfato de potasio o hidrogenofosfato dipotásico, o las sales correspondientes que contienen sodio. El medio de cultivo debe contener además sales de metales, por ejemplo sulfato de magnesio o sulfato de hierro, que son necesarias para el crecimiento. Finalmente, se pueden usar sustancias esenciales para el crecimiento, tales como aminoácidos y vitaminas, además de las sustancias mencionadas anteriormente. Además de esto, se pueden añadir al medio de cultivo precursores adecuados. Las sustancias añadidas mencionadas se pueden añadir al cultivo en forma de una mezcla de una sola vez, o se pueden alimentar de manera adecuada durante el cultivo.

Los compuestos básicos, tales como hidróxido de sodio, hidróxido de potasio, amoníaco o amoníaco acuoso, o compuestos ácidos tales como ácido fosfórico o ácido sulfúrico se emplean de manera adecuada para controlar el pH del cultivo. Es posible usar antiespumantes tales como, por ejemplo, ésteres poliglicólicos de ácidos grasos, para controlar la formación de espuma. Se pueden añadir al medio sustancias adecuadas que actúan selectivamente, tales como, por ejemplo, antibióticos, a fin de mantener la estabilidad de los plásmidos. A fin de mantener las condiciones aerobias, se hace pasar en el cultivo oxígeno o mezclas gaseosas que contienen oxígeno, tales como, por ejemplo, aire. La temperatura del cultivo es normalmente de 20°C a 45°C, y preferiblemente de 25°C a 40°C. El cultivo se continúa hasta que se ha formado un máximo de L-prolina, o hasta que el rendimiento o productividad ha alcanzado un nivel óptimo deseado. Este objetivo se logra normalmente en 10 horas a 160 horas.

Los métodos para determinar L-prolina se describen en la técnica anterior. El análisis puede tener lugar, por ejemplo, por medio de cromatografía de intercambio aniónico, seguido de derivatización con ninhidrina y de la detección a una longitud de onda apropiada, como se describe en Spackman et al. (Analytical Chemistry, 30 (1958), 1190).

En consecuencia, otra realización de la presente invención constituye un procedimiento para producir L-prolina

a) fermentando organismos hospedantes no humanos según la invención, y

b) aislando o recogiendo L-prolina, cuando sea apropiado con componentes del caldo de fermentación y/o de la biomasa.

Se da preferencia al empleo de un procedimiento para producir L-prolina que comprende las siguientes etapas:

- a) fermentar bacterias corineformes según la invención,
- b) concentrar L-prolina en el caldo de fermentación o en las células de las bacterias corineformes,
- c) aislar o recoger L-prolina del caldo de fermentación, cuando sea apropiado
- 5 d) con componentes del caldo de fermentación y/o de la biomasa (de > 0 a < 100% en peso de biomasa, preferiblemente de 10 a 80% en peso, más preferiblemente 20-60% en peso).

La L-prolina producida de esta manera se puede recolectar y aislar y, cuando sea apropiado, purificar, como lo determina el experto.

El procedimiento según la invención se usa para la producción de L-prolina por fermentación.

10 La expresión "posición comparable" significa, según la invención, una posición que, comparando la secuencia de partida con la secuencia comparativa con aplicación de un programa de comparación de secuencias (BLAST, Altschul et al. J. Mol. Biol. 1990, 215, 403-10) en la posición contemplada de la secuencia de partida, proporciona una posición de aminoácidos en la secuencia comparativa que difiere de la posición a comparar en no más de ± 5 , más preferiblemente ± 4 , adicionalmente de forma preferible ± 3 , todavía más preferiblemente ± 2 , muy preferiblemente ± 1 , y especialmente cero posiciones.

15 Las instrucciones con respecto a la hibridación las puede encontrar el experto, entre otros, en el manual "The DIG System Users Guide for Filter Hybridization" de Boehringer Mannheim GmbH (Mannheim, Alemania, 1993), y en Liebl et al. (International Journal of Systematic Bacteriology 41: 255-260 (1991)). La hibridación se lleva a cabo en condiciones restrictivas, es decir, sólo se forman híbridos en los que la sonda, por ejemplo la secuencia nucleotídica complementaria a SEC ID NO: 3, y la secuencia diana, es decir, los polinucleótidos tratados con la sonda, son al menos 70% idénticas. Se sabe que la restricción de la hibridación, incluyendo aquella de las etapas de lavado, está influida o determinada por la variación de la composición del tampón, la temperatura y la concentración de sal. La reacción de hibridación se lleva a cabo generalmente a una restricción relativamente baja en comparación con las etapas de lavado (Hybaid Hybridisation Guide, Hybaid Limited, Teddington, UK, 1996).

25 Por ejemplo, para la reacción de hibridación se puede usar un tampón que corresponde a un tampón de SSC 5x a una temperatura de aprox. 50°C-68°C. En este caso, las sondas también se pueden hibridar con polinucleótidos que son menos de 70% idénticos a la secuencia de la sonda. Tales híbridos son menos estables y se eliminan mediante lavado en condiciones restrictivas. Esto se puede lograr, por ejemplo, reduciendo la concentración de sal hasta 2x SSC y, cuando sea apropiado, subsiguientemente hasta 0,5x SSC (The DIG System User's Guide for Filter Hybridisation, Boehringer Mannheim, Mannheim, Alemania, 1995), ajustándose la temperatura a aprox. 50°C-68°C, aprox. 52°C-68°C, aprox. 54°C-68°C, aprox. 56°C-68°C, aprox. 58°C-68°C, aprox. 60°C-68°C, aprox. 62°C-68°C, aprox. 64°C-68°C, aprox. 66°C-68°C. Se da preferencia a llevar a cabo las etapas de lavado a temperaturas de aprox. 62°C-68°C, particularmente de forma preferible aprox. 64°C-68°C o aprox. 66°C-68°C. Es posible, cuando sea apropiado, reducir la concentración de sal hasta una concentración que corresponde a 0,2x SSC o 0,1x SSC. Incrementando gradualmente la temperatura de la hibridación en etapas de aprox. 1-2°C desde 50°C hasta 68°C, es posible aislar fragmentos polinucleotídicos que tienen al menos 70% o al menos 80% o al menos 90% a 95% o al menos 96% a 98% o al menos 99% de identidad con la secuencia de la sonda empleada. Las instrucciones adicionales con respecto a la hibridación están comercialmente disponibles en forma de "kits" (por ejemplo, DIG Easy Hyb de Roche Diagnostics GmbH, Mannheim, Alemania, Número de Catálogo 1603558).

40 Según la invención, los organismos hospedantes no humanos que comprenden un polinucleótido que codifica polipéptidos reivindicados (secuencias de aminoácidos) también comprenden aquellas secuencias que tienen una identidad (a nivel de aminoácidos) de más de 90% (con respecto a los polipéptidos), de forma preferible más de 91%, 92%, 93% o 94%, más preferiblemente más de 95% o 96%, y particularmente de forma preferible más de 97%, 98% o 99% con cualquiera de estas secuencias, en tanto que se retenga la acción o el propósito de tal secuencia. El término "homología" (o identidad), como se usa aquí, se puede definir por la ecuación $H(\%) = [1 - V/X] \times 100$, en la que H es la homología, X es el número total de nucleobases/aminoácidos de la secuencia comparativa, y V es el número de diferentes nucleobases/aminoácidos de la secuencia a considerar, basándose en la secuencia comparativa. En cualquier caso, la expresión secuencias de ácidos nucleicos que codifican polipéptidos incluye cualesquiera secuencias que parecen posibles según la degeneración del código genético.

50 El porcentaje de identidad para las secuencias de aminoácidos indicadas en la presente memoria descriptiva por los números SEC ID se puede determinar fácilmente por la persona experta usando métodos conocidos en la técnica anterior. Un programa adecuado que se puede emplear según la invención es BLASTP (Altschul et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):3389-3402). La secuencias de ácidos nucleicos según la invención puede ser una molécula de ADN o una molécula de ARN. Se da preferencia a que la secuencia de ácidos nucleicos sea una molécula de ADN o una molécula de ARNm. Según la invención, la molécula de ADN puede ser además una molécula de ADN genómico o una molécula de ADN aislado. La invención engloba además realizaciones en las que la molécula de ADN es una molécula de PNA u otro derivado de una molécula de ADN.

ES 2 537 647 T3

Los microorganismos mencionados en esta solicitud, que se indican por un número DSMZ, se pueden obtener de la Deutsche Sammlung für Mikroorganismen und Zellkulturen, Mascheroder Weg 4, Brunswick (Alemania).

LISTADO DE SECUENCIAS

- 5 <110> Degussa AG
<120> Mutante del gen proB de las bacterias corineformes
<130> 040071 AM
<160> 6
<170> PatentIn version 3.3
- 10 <210> 1
<211> 1110
<212> ADN
<213> bacteria corineforme
- 15 <220>
<221> CDS
<222> (1)..(1107)
- 20 <220>
<221> misc feature
<222> (445)..(447)
<223> Xaa todos los aminoácidos proteinogénicos distintos de glicina
- 25 <400> 1

ES 2 537 647 T3

atg	cgt	gag	cgc	atc	tcc	aac	gct	aag	cga	gtg	gtg	gtg	aaa	att	ggt	48
Met	Arg	Glu	Arg	Ile	Ser	Asn	Ala	Lys	Arg	Val	Val	Val	Lys	Ile	Gly	
1				5				10					15			
tcg	tcc	tca	ttg	act	aac	gat	gag	gac	gga	cac	acc	gtc	gat	ccc	aac	96
Ser	Ser	Ser	Leu	Thr	Asn	Asp	Glu	Asp	Gly	His	Thr	Val	Asp	Pro	Asn	
			20					25					30			
cgc	atc	aac	act	att	gtc	aat	gcc	ttg	caa	gca	cgc	atg	gaa	gct	ggc	144
Arg	Ile	Asn	Thr	Ile	Val	Asn	Ala	Leu	Gln	Ala	Arg	Met	Glu	Ala	Gly	
		35					40					45				
tcg	gac	ctc	atc	ggt	gtg	tcc	tct	ggc	gca	gtg	gcc	gcg	gga	atg	gcc	192
Ser	Asp	Leu	Ile	Val	Val	Ser	Ser	Gly	Ala	Val	Ala	Ala	Gly	Met	Ala	
	50					55					60					
ccg	ctt	gga	ttg	agc	acc	cgg	ccc	acg	gaa	ttg	gca	gtc	aag	cag	gct	240
Pro	Leu	Gly	Leu	Ser	Thr	Arg	Pro	Thr	Glu	Leu	Ala	Val	Lys	Gln	Ala	
65				70					75					80		
gca	gca	gca	gtg	ggg	caa	ggt	cac	ctc	atg	cac	cag	tgg	gga	cgt	tct	288
Ala	Ala	Ala	Val	Gly	Gln	Val	His	Leu	Met	His	Gln	Trp	Gly	Arg	Ser	
			85					90						95		
ttt	gcc	cgg	tat	ggt	cgc	ccc	atc	ggc	cag	gtg	ctt	ctt	acc	gca	gct	336
Phe	Ala	Arg	Tyr	Gly	Arg	Pro	Ile	Gly	Gln	Val	Leu	Leu	Thr	Ala	Ala	
			100					105					110			
gat	gca	gga	aag	cgt	gat	cgt	gcg	agg	aat	gcg	cag	cgt	acc	atc	gac	384
Asp	Ala	Gly	Lys	Arg	Asp	Arg	Ala	Arg	Asn	Ala	Gln	Arg	Thr	Ile	Asp	
		115					120					125				

ES 2 537 647 T3

aag ctg cgc att ttg ggc gcg gtt cct atc gtc aat gaa aat gac acc 432
 Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
 130 135 140

gtg gca acc acc nnn gtg aat ttt ggt gac aac gac cga ctt gct gca 480
 Val Ala Thr Thr Xaa Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
 145 150 155 160

att gtg gcg cac ctg gtg tcg gct gat gct ttg gtg ctg ctc agt gac 528
 Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
 165 170 175

gtg gat gga ctt ttt gat aaa aac cct act gat ccc acc gcg aag ttt 576
 Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
 180 185 190

att tcc gag gtt cgt gac ggc aat gat ttg aaa ggt gtc att gcc ggc 624
 Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
 195 200 205

gac ggc gga aaa gtg ggc acc ggt ggc atg gca tca aag gtg tct gct 672
 Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
 210 215 220

gca cgt ttg gct tcc cga agt ggc gtg cct gtg ctg ttg acc tct gcg 720
 Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
 225 230 235 240

gca aac att ggc cca gca ctg gaa gac gcc cag gtg ggc act gta ttc 768
 Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
 245 250 255

cac ccc aag gac aac cgc ctc tcc gcg tgg aag ttc tgg gct ttg tat 816
 His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
 260 265 270

gcc gca gat act gca gga aag atc cga ctc gat gac ggc gcg gtg gaa 864
 Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
 275 280 285

gca gtg acc tcc ggt ggt aaa tct ttg ctg gct gtg ggc att act gaa 912
 Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
 290 295 300

atc att ggt gat ttc cag cag ggt gag atc gtg gag atc ttg gga cct 960
 Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
 305 310 315 320

gcc ggc caa atc atc ggg cga ggc gag gtg tcc tac gat tct gat acc 1008
 Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
 325 330 335

ttg caa tca atg gtt ggt atg caa acg cag gac ctt cca gat ggc atg 1056
 Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
 340 345 350

cag cgc ccg gta gtg cat gca gat tat ctg tcc aac tac gcc agc cgc 1104
 Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
 355 360 365

<210> 2

<211> 369

ES 2 537 647 T3

<212> PRT

<213> bacteria corineforme

<220>

5 <221> misc feature

<222> (149)..(149)

<223> "Xaa" en la localización 149 representa Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Ala, Val, Gln, His, Pro, Leu, Tyr, Trp, Cys, o Phe.

10 <400> 2

```

Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
1           5           10           15

Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
          20           25           30

Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
          35           40           45

Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
          50           55           60

Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65           70           75           80

Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
          85           90           95

Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
          100          105          110

Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
          115          120          125

Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
130          135          140

Val Ala Thr Thr Xaa Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
145          150          155          160

Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
          165          170          175

```

ES 2 537 647 T3

Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
180 185 190

Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
195 200 205

Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
210 215 220

Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
225 230 235 240

Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
245 250 255

His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
260 265 270

Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
275 280 285

Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
290 295 300

Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
305 310 315 320

Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
325 330 335

Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
340 345 350

Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
355 360 365

Ala

<210> 3

<211> 1110

<212> ADN

5 <213> bacteria corineforme

<220>

<221 > CDS

ES 2 537 647 T3

<222> (1)..(1107)

<220>

<221> mutación

5 <222> (446)..(446)

<223> adenina

<400> 3

```

atg cgt gag cgc atc tcc aac gct aag cga gtg gtg gtg aaa att ggt      48
Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
1          5          10          15

tcg tcc tca ttg act aac gat gag gac gga cac acc gtc gat ccc aac      96
Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
          20          25          30

cgc atc aac act att gtc aat gcc ttg caa gca cgc atg gaa gct ggc      144
Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
          35          40          45

tcg gac ctc atc gtt gtg tcc tct ggc gca gtg gcc gcg gga atg gcc      192
Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
          50          55          60

ccg ctt gga ttg agc acc cgg ccc acg gaa ttg gca gtc aag cag gct      240
Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65          70          75          80

gca gca gca gtg ggg caa gtt cac ctc atg cac cag tgg gga cgt tct      288
Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
          85          90          95

ttt gcc cgg tat ggt cgc ccc atc ggc cag gtg ctt ctt acc gca gct      336
Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
          100          105          110

gat gca gga aag cgt gat cgt gcg agg aat gcg cag cgt acc atc gac      384
Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
          115          120          125

aag ctg cgc att ttg ggc gcg gtt cct atc gtc aat gaa aat gac acc      432
Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
          130          135          140

gtg gca acc acc gat gtg aat ttt ggt gac aac gac cga ctt gct gca      480
Val Ala Thr Thr Asp Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
145          150          155          160

att gtg gcg cac ctg gtg tcg gct gat gct ttg gtg ctg ctc agt gac      528
Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
          165          170          175

gtg gat gga ctt ttt gat aaa aac cct act gat ccc acc gcg aag ttt      576
Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
          180          185          190

att tcc gag gtt cgt gac ggc aat gat ttg aaa ggt gtc att gcc ggc      624
Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
          195          200          205

gac ggc gga aaa gtg ggc acc ggt ggc atg gca tca aag gtg tct gct      672
Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
          210          215          220

```

ES 2 537 647 T3

gca cgt ttg gct tcc cga agt ggc gtg cct gtg ctg ttg acc tct gcg 720
 Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
 225 230 235 240

gca aac att ggc cca gca ctg gaa gac gcc cag gtg ggc act gta ttc 768
 Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
 245 250 255

cac ccc aag gac aac cgc ctc tcc gcg tgg aag ttc tgg gct ttg tat 816
 His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
 260 265 270

gcc gca gat act gca gga aag atc cga ctc gat gac ggc gcg gtg gaa 864
 Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
 275 280 285

gca gtg acc tcc ggt ggt aaa tct ttg ctg gct gtg ggc att act gaa 912
 Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
 290 295 300

atc att ggt gat ttc cag cag ggt gag atc gtg gag atc ttg gga cct 960
 Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
 305 310 315 320

gcc ggc caa atc atc ggg cga ggc gag gtg tcc tac gat tct gat acc 1008
 Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
 325 330 335

ttg caa tca atg gtt ggt atg caa acg cag gac ctt cca gat ggc atg 1056
 Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
 340 345 350

cag cgc ccg gta gtg cat gca gat tat ctg tcc aac tac gcc agc cgc 1104
 Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
 355 360 365

gcg taa 1110
 Ala

<210> 4

<211 > 369

5 <212> PRT

<213> bacteria corineforme

<400> 4

Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
 1 5 10 15

Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
 20 25 30

Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
 35 40 45

ES 2 537 647 T3

Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
50 55 60

Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65 70 75 80

Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
85 90 95

Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
100 105 110

Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
115 120 125

Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
130 135 140

Val Ala Thr Thr Asp Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
145 150 155 160

Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
165 170 175

Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
180 185 190

Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
195 200 205

Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
210 215 220

Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
225 230 235 240

Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
245 250 255

His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
260 265 270

Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
275 280 285

Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
290 295 300

ES 2 537 647 T3

Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
305 310 315 320

Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
325 330 335

Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
340 345 350

Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
355 360 365

Ala

<210> 5

<211> 1110

5 <212> ADN

<213> bacteria corineforme

<220>

<221> CDS

10 <222> (1)..(1107)

<220>

<221> mutación

<222> (446)..(446)

15 <223> adenina

<400> 5

ES 2 537 647 T3

atg	cgt	gaa	cgc	atc	tcc	aac	gct	aag	cga	gtg	gtg	gtg	aaa	att	ggt	48
Met	Arg	Glu	Arg	Ile	Ser	Asn	Ala	Lys	Arg	Val	Val	Val	Lys	Ile	Gly	
1				5				10					15			
tcg	tcc	tca	ttg	act	aac	gat	gag	gac	gga	cac	acc	gtc	gat	ccc	aac	96
Ser	Ser	Ser	Leu	Thr	Asn	Asp	Glu	Asp	Gly	His	Thr	Val	Asp	Pro	Asn	
			20					25					30			
cgc	atc	aac	act	att	gtc	aat	gcc	ttg	caa	gca	cgc	atg	gaa	gct	ggc	144
Arg	Ile	Asn	Thr	Ile	Val	Asn	Ala	Leu	Gln	Ala	Arg	Met	Glu	Ala	Gly	
		35					40					45				
tcg	gac	ctc	atc	gtt	gtg	tcc	tct	ggc	gca	gtg	gcc	gcg	gga	atg	gcc	192
Ser	Asp	Leu	Ile	Val	Val	Ser	Ser	Gly	Ala	Val	Ala	Ala	Gly	Met	Ala	
	50					55					60					
ccg	ctt	gga	ttg	agc	acc	cgg	ccc	acg	gaa	ttg	gca	gtc	aag	cag	gct	240
Pro	Leu	Gly	Leu	Ser	Thr	Arg	Pro	Thr	Glu	Leu	Ala	Val	Lys	Gln	Ala	
65				70					75					80		
gca	gca	gca	gtg	ggg	caa	gtt	cac	ctc	atg	cac	cag	tgg	gga	cgt	tct	288
Ala	Ala	Ala	Val	Gly	Gln	Val	His	Leu	Met	His	Gln	Trp	Gly	Arg	Ser	
				85				90						95		

ES 2 537 647 T3

ttt gcc cgg tat ggt cgc ccc atc ggc cag gtg ctt ctt acc gca gct	336
Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala	
100 105 110	
gat gca gga aag cgt gat cgt gcg agg aat gcg cag cgt acc atc gac	384
Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp	
115 120 125	
aag ctg cgc att ttg ggc gcg gtt cct atc gtc aat gaa aat gac acc	432
Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr	
130 135 140	
gtg gca acc acc gat gtg aat ttt ggt gac aac gac cga ctt gct gca	480
Val Ala Thr Thr Asp Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala	
145 150 155 160	
att gtg gcg cac ctg gtg tcg gct gac gct ttg gtg ctg ctc agt gac	528
Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp	
165 170 175	
gtg gat gga ctt ttt gat aag aac cct act gat ccc acc gcg aag ttt	576
Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe	
180 185 190	
att tcc gag gtt cgt gac ggc aat gat ttg aaa ggt gtc att gcc ggc	624
Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly	
195 200 205	
gac ggc gga aaa gtg ggc acc ggc ggc atg gca tca aag gtg tct gct	672
Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala	
210 215 220	
gca cgt ttg gct tcc cga agt ggc gtg cct gtg ctg ttg acc tct gcg	720
Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala	
225 230 235 240	
gca aac att ggc cca gca ctg gaa gac gcc cag gtg ggc act gta ttc	768
Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe	
245 250 255	
cac ccc aag gac aac cgc ctc tcc gcg tgg aag ttc tgg gct ttg tat	816
His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr	
260 265 270	
gcc gca gat act gca gga aag atc cga ctt gat gat ggc gcg gtg gaa	864
Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu	
275 280 285	
gca gtg acc tcc ggt ggt aaa tct ttg ctg gct gtg ggc att act gag	912
Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu	
290 295 300	
atc att ggt gat ttc caa cag ggt gag atc gtg gag atc ttg gga cct	960
Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro	
305 310 315 320	
gcc ggc caa atc atc ggg cga ggc gag gtg tcc tac gat tct gat acc	1008
Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr	
325 330 335	

ES 2 537 647 T3

```

ttg caa tca atg gtt ggc atg caa acg cag gac ctt cca gat ggc atg      1056
Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
          340                      345                      350

cag cgc ccg gta gtg cat gca gat tat ctg tcc aac tac gcc agc cgc      1104
Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
          355                      360                      365

gcg taa
Ala
                                                    1110

```

<210> 6

<211> 369

5 <212> PRT

<213> bacteria corineforme

<400> 6

```

Met Arg Glu Arg Ile Ser Asn Ala Lys Arg Val Val Val Lys Ile Gly
1          5          10          15

Ser Ser Ser Leu Thr Asn Asp Glu Asp Gly His Thr Val Asp Pro Asn
          20          25          30

Arg Ile Asn Thr Ile Val Asn Ala Leu Gln Ala Arg Met Glu Ala Gly
          35          40          45

Ser Asp Leu Ile Val Val Ser Ser Gly Ala Val Ala Ala Gly Met Ala
          50          55          60

Pro Leu Gly Leu Ser Thr Arg Pro Thr Glu Leu Ala Val Lys Gln Ala
65          70          75          80

Ala Ala Ala Val Gly Gln Val His Leu Met His Gln Trp Gly Arg Ser
          85          90          95

Phe Ala Arg Tyr Gly Arg Pro Ile Gly Gln Val Leu Leu Thr Ala Ala
          100          105          110

Asp Ala Gly Lys Arg Asp Arg Ala Arg Asn Ala Gln Arg Thr Ile Asp
          115          120          125

```

10

ES 2 537 647 T3

Lys Leu Arg Ile Leu Gly Ala Val Pro Ile Val Asn Glu Asn Asp Thr
 130 135 140

Val Ala Thr Thr Asp Val Asn Phe Gly Asp Asn Asp Arg Leu Ala Ala
 145 150 155 160

Ile Val Ala His Leu Val Ser Ala Asp Ala Leu Val Leu Leu Ser Asp
 165 170 175

Val Asp Gly Leu Phe Asp Lys Asn Pro Thr Asp Pro Thr Ala Lys Phe
 180 185 190

Ile Ser Glu Val Arg Asp Gly Asn Asp Leu Lys Gly Val Ile Ala Gly
 195 200 205

Asp Gly Gly Lys Val Gly Thr Gly Gly Met Ala Ser Lys Val Ser Ala
 210 215 220

Ala Arg Leu Ala Ser Arg Ser Gly Val Pro Val Leu Leu Thr Ser Ala
 225 230 235 240

Ala Asn Ile Gly Pro Ala Leu Glu Asp Ala Gln Val Gly Thr Val Phe
 245 250 255

His Pro Lys Asp Asn Arg Leu Ser Ala Trp Lys Phe Trp Ala Leu Tyr
 260 265 270

Ala Ala Asp Thr Ala Gly Lys Ile Arg Leu Asp Asp Gly Ala Val Glu
 275 280 285

Ala Val Thr Ser Gly Gly Lys Ser Leu Leu Ala Val Gly Ile Thr Glu
 290 295 300

Ile Ile Gly Asp Phe Gln Gln Gly Glu Ile Val Glu Ile Leu Gly Pro
 305 310 315 320

Ala Gly Gln Ile Ile Gly Arg Gly Glu Val Ser Tyr Asp Ser Asp Thr
 325 330 335

Leu Gln Ser Met Val Gly Met Gln Thr Gln Asp Leu Pro Asp Gly Met
 340 345 350

Gln Arg Pro Val Val His Ala Asp Tyr Leu Ser Asn Tyr Ala Ser Arg
 355 360 365

Ala

REIVINDICACIONES

1. El organismo hospedante no humano que comprende un polinucleótido que codifica un polipéptido que tiene una secuencia de aminoácidos que es al menos un 90% idéntica a la secuencia de aminoácidos expuesta en SEC ID NO: 2, con lo que en las secuencias de aminoácidos la glicina en la posición 149 se sustituye por un aminoácido
5 proteínogénico, con lo que el polipéptido tiene actividad de γ -glutamil cinasa, con lo que se sobreexpresa el polinucleótido utilizando un promotor potente de modo que la actividad o concentración de la proteína correspondiente se incrementa en al menos un 10% basado en la actividad o concentración de la proteína en el microorganismo de partida.
2. El organismo hospedante según la reivindicación 1, en el que el intercambio de aminoácidos se selecciona de
10 SEC ID NO: 2, en los que la glicina en la posición 149 se sustituye por un aminoácido proteínogénico.
3. El organismo hospedante según las reivindicaciones 1 ó 2, en el que la glicina en la posición 149 de SEC ID NO: 2 se sustituye por ácido L-aspártico.
4. El organismo hospedante de la reivindicación 3, en el que dicho polinucleótido comprende la secuencia nucleotídica de SEC ID NO: 3.
- 15 5. El organismo hospedante no humano que comprende un polinucleótido que codifica un polipéptido que tiene desde las posiciones 145 a 154 la secuencia de aminoácidos de las posiciones 145 a 154 de SEC ID NO: 2, en los que dicho polinucleótido se hibrida en condiciones restrictivas que comprenden una etapa de lavado a 0,5 SSC a 50-68°C al complemento de SEC ID NO: 3, en los que el polipéptido tiene actividad de γ -glutamil cinasa.
- 20 6. El organismo hospedante según la reivindicación 5, en el que dicho polipéptido codificado tiene desde las posiciones 145 a 154 la secuencia de aminoácidos de las posiciones 145 a 154 de SEC ID NO: 4.
7. El organismo hospedante según la reivindicación 5, en el que dicho polipéptido codificado tiene desde las posiciones 110 a 189 la secuencia de aminoácidos de las posiciones 110 a 189 de SEC ID NO: 2.
8. El organismo hospedante según la reivindicación 6, en el que dicho polipéptido codificado tiene desde las posiciones 110 a 189 la secuencia de aminoácidos de las posiciones 110 a 189 de SEC ID NO: 4.
- 25 9. El organismo hospedante según las reivindicaciones 1 a 8, en el que el organismo hospedante se escoge a partir del grupo que comprende levaduras, E. coli, B. subtilis, bacterias coreniformes.
10. El organismo hospedante de acuerdo con la reivindicación 9, donde dicho organismo hospedante es una bacteria coreniforme.
- 30 11. El organismo hospedante según la reivindicación 10, en el que dicho organismo hospedante es Corynebacterium glutamicum.
12. Un procedimiento para la preparación de L-prolina, que comprende
 - a) fermentar el organismo hospedante según una o más de las reivindicaciones 1 a 12
 - b) enriquecer dicho L-aminoácido en el medio o en las células de las bacterias, y
 - 35 c) aislar o recoger dicho aminoácido, opcionalmente con componentes procedentes del caldo de fermentación y/o biomasa.
13. El procedimiento según la reivindicación 12, en el que se fermentan los organismos hospedantes que comprenden uno o más genes sobreexpresados escogidos del grupo que comprende:
 - a) el gen gdh que codifica glutamato deshidrogenasa
 - b) el gen proA que codifica γ -glutamil-fosfato reductasa
 - 40 c) el gen proC que codifica la pirrolin-5-carboxilato reductasa y
 - d) el gen ocd que codifica ornitina ciclodesaminasa.
14. El procedimiento de acuerdo con la reivindicación 12, en el que se fermenta un organismo hospedante que comprende uno o más genes atenuados escogidos del grupo que comprende:
 - a) el gen ilvA que codifica treonina desaminasa
 - 45 b) el gen putA que codifica prolina deshidrogenasa/pirrolin-5-carboxilato deshidrogenasa
 - c) el gen sucA que codifica 2-cetoglutarato deshidrogenasa

- d) el gen *sucB* que codifica dihidrolipoamida succiniltransferasa, y
- e) el gen *argD* que codifica acetilornitina aminotransferasa.