



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 343 010**

51 Int. Cl.:
C12P 13/08 (2006.01)
C12N 15/31 (2006.01)
C07K 14/245 (2006.01)
C07K 14/25 (2006.01)
C07K 14/225 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04803729 .5**
96 Fecha de presentación : **10.12.2004**
97 Número de publicación de la solicitud: **1697530**
97 Fecha de publicación de la solicitud: **06.09.2006**

54 Título: **Proceso para preparación de L-aminoácidos que utiliza cepas de la familia Enterobacteriáceas.**

30 Prioridad: **24.12.2003 DE 103 61 268**

45 Fecha de publicación de la mención BOPI:
21.07.2010

45 Fecha de la publicación del folleto de la patente:
21.07.2010

73 Titular/es: **Evonik Degussa GmbH**
Rellinghauser Strasse 1-11
45128 Essen, DE

72 Inventor/es: **Dusch, Nicole**

74 Agente: **Lehmann Novo, María Isabel**

ES 2 343 010 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).

ES 2 343 010 T3

DESCRIPCIÓN

Proceso para preparación de L-aminoácidos que utiliza cepas de la familia Enterobacteriáceas.

5 Esta invención se refiere a un proceso para preparación de L-lisina, L-valina y L-treonina, que utiliza cepas recombinantes de microorganismos de la familia Enterobacteriáceas en la cual el marco de lectura abierto (ORF) designado yaaU está sobreexpresado, y a dichos microorganismos.

Técnica anterior

10 Los L-aminoácidos, en particular L-treonina, se utilizan en medicina humana y en la industria farmacéutica, en la industria alimentaria y, muy particularmente, en nutrición animal.

15 Es sabido que los L-aminoácidos se pueden preparar por fermentación de cepas de Enterobacteriáceas, en particular *Escherichia coli* (*E. coli*) y *Serratia marcescens*. Debido a su gran importancia, están realizándose continuamente esfuerzos para mejorar los métodos de preparación. Las mejoras metodológicas pueden referirse a medidas relacionadas con la tecnología de la fermentación, tales como la agitación o el suministro de oxígeno, o a la composición de los medios nutrientes, tales como la concentración de azúcar durante la fermentación, o el acabado de la forma del producto, por ejemplo por medio de cromatografía de intercambio iónico, o las propiedades intrínsecas de eficiencia del microorganismo propiamente dicho.

20 Para la mejora de las propiedades de eficiencia de estos microorganismos se utilizan métodos de mutagénesis, selección y elección de mutantes. Esto da por tanto como resultado cepas que son resistentes a antimetabolitos, tales como el análogo de treonina ácido α -amino- β -hidroxivalérico (AHV), o auxótrofas para metabolitos de importancia reguladora, y producen L-aminoácidos tales como L-treonina.

30 Desde hace unos cuantos años, se han utilizado también métodos de DNA recombinante para la mejora de las cepas productoras de L-aminoácidos de la familia Enterobacteriáceas por amplificación de genes individuales de biosíntesis de aminoácidos e investigación del efecto sobre la producción. Información recopilada sobre la biología celular y biología molecular de *Escherichia coli* y *Salmonella* puede encontrarse en Neidhardt (ed): *Escherichia coli* and *Salmonella*, Cellular and Molecular Biology, 2nd edition, ASM Press, Washington, D.C., USA, (1996).

35 Los documentos D1-D3 describen las secuencias de aminoácidos de los polipéptidos de acuerdo con SEQ ID NO: 4, 6 y 8 de la presente solicitud, respectivamente.

D1 DATABASE UNI-PROT YAAU_ECOLI, 1 de noviembre de 1997 (1997-11-01),

XP 002321166

40 Acceso a la base de datos No. P31679

Resumen

45 abarca una proteína hipotética yaaU de transporte de metabolitos de *Escherichia coli*

D2 DATABASE UNI-PROT Q83Q5_SHIFL

1 de junio de 2003 (2003-06-01),

50 XP002321167

Acceso a la base de datos No. Q83SQ5

55 Resumen

comprende una proteína supuesta de transporte de *Shigella flexneri*

60 D3 DATABASE UNI-PROT Q8ZRW7_SALTY

1 de octubre de 2003 (2003-10-01),

XP002321168

65 Acceso a la base de datos no. Q8ZRW7

ES 2 343 010 T3

Resumen

comprende una proteína supuesta de transporte de la familia MFS de *Salmonella typhimurium*.

5 Los documentos D4 WO 03/076637A y (D5) WO 03/008600A describen métodos para la producción de L-treonina por medio de sobreexpresión o delección de ciertos genes.

10 El problema a resolver en (D4) está considerado como la provisión de un método para preparación de L-treonina por fermentación de microorganismos de la familia Enterobacteriáceas que comprenden un gen pepB sobreexpresado.

Objeto de la invención

15 El objeto de esta invención es proporcionar nuevas medidas para mejorar la preparación de L-aminoácidos, en particular L-treonina.

Descripción de la invención

20 La invención se refiere a microorganismos recombinantes de la familia Enterobacteriáceas que contienen un ORF yaaU sobreexpresado, que codifica un polipéptido que está apostillado como un transportador de azúcar supuesto, y que producen L-lisina, L-valina o L-treonina, de una manera mejorada.

25 En cada caso, se utilizan como punto de partida para la comparación los microorganismos que no son recombinantes para el ORF yaaU, y que no contienen cantidad alguna de ORF yaaU sobreexpresado.

30 Estos microorganismos recombinantes incluyen, en particular, microorganismos de la familia Enterobacteriáceas en los cuales un polinucleótido que codifica un polipéptido cuya secuencia de aminoácidos es al menos 90%, en particular al menos 95%, preferiblemente al menos 98%, al menos 99%, de modo particularmente preferible 99,7% y de modo muy particular preferible 100%, idéntico a una secuencia de aminoácidos seleccionada del grupo SEQ ID No. 4, SEQ ID No. 6 y SEQ ID No. 8 está sobreexpresado.

Se da preferencia a secuencias de aminoácidos que son idénticas a las de SEQ ID No. 4, SEQ ID No. 6 o SEQ ID No. 8.

35 Dichos microorganismos contienen polinucleótidos sobreexpresados seleccionados del grupo:

- a) polinucleótido que tiene la secuencia de nucleótidos SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7;
- 40 b) polinucleótido que tiene una secuencia de nucleótidos que corresponde a SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7 dentro de los límites de la degeneración del código genético;
- c) secuencia de polinucleótidos que tiene una secuencia que se hibrida, en condiciones severas, con la secuencia que es complementaria a la secuencia SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7;
- 45 d) polinucleótido que tiene una secuencia SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7 que contiene mutantes sentido funcionalmente neutros,

codificando los polinucleótidos un transportador de azúcar supuesto.

50 La invención se refiere también a un proceso para preparar por fermentación L-lisina, L-valina o L-treonina, utilizando microorganismos recombinantes de la familia Enterobacteriáceas que, en particular, producen ya dichos L-aminoácidos y en los cuales al menos el marco de lectura abierto (ORF) que tiene la designación yaaU, o secuencias de nucleótidos que codifican su producto génico, está o están sobreexpresados.

55 Se da preferencia a la utilización de los microorganismos que se describen.

60 Cuando se mencionan L-aminoácidos o aminoácidos en lo que sigue, esto significa uno o más aminoácidos, con inclusión de sus sales, seleccionados del grupo L-asparagina, L-treonina, L-serina, L-glutamato, L-glicina, L-alanina, L-cisteína, L-valina, L-metionina, L-prolina, L-isoleucina, L-leucina, L-tirosina, L-fenil-alanina, L-histidina, L-lisina, L-triptófano, L-arginina and L-homoserina. Se prefiere particularmente L-treonina.

65 En este contexto, el término “intensificación” describe el aumento, en un microorganismo, de la actividad o concentración intracelular de una o más enzimas o proteínas que están codificadas por el DNA correspondiente, estando incrementado, por ejemplo, el número de copias del gen o genes, o del ORF o los ORFs, al menos en una (1) copia, haciéndose uso de un promotor fuerte enlazado operativamente al gen o de un gen o alelo u ORF que codifica una enzima o proteína correspondiente que tiene una actividad alta, y combinándose estas medidas en caso apropiado.

ES 2 343 010 T3

Un segmento de una secuencia de nucleótidos que codifica, o puede codificar, una proteína y/o un polipéptido o ácido ribonucleico al cual ha sido incapaz de asignar función alguna la técnica anterior se designa como un marco de lectura abierto (ORF). Después que se ha asignado una función al segmento de secuencia nucleotídica en cuestión, se hace referencia generalmente a este segmento como un gen. Los alelos se entienden generalmente como formas alternativas de un gen dado. Las formas se distinguen por diferencias en la secuencia de nucleótidos.

En general, la proteína, o el ácido ribonucleico, codificado por una secuencia de nucleótidos, es decir, un ORF, un gen o un alelo, se designa como un producto génico.

Las medidas de intensificación, en particular de sobreexpresión, aumentan generalmente la actividad o concentración de la proteína correspondiente al menos en un 10%, 25%, 50%, 75%, 100%, 150%, 200%, 300%, 400% o 500%, como máximo hasta 1000% o 2000%, basada en la de la proteína de tipo salvaje o en la actividad o concentración de la proteína en la cepa o microorganismo parental que no es recombinante para la enzima o proteína correspondiente. El microorganismo o cepa parental no recombinante se entiende como el microorganismo sobre el cual se realizan las medidas de acuerdo con la invención.

La invención se refiere a un proceso para preparación de L-lisina, L-valina o L-treonina por fermentación de microorganismos recombinantes de la familia Enterobacteriáceas, caracterizado porque

- a) dichos microorganismos reivindicados productores de L-aminoácidos se cultivan en un medio en condiciones en las cuales el L-aminoácido deseado está enriquecido en el medio o en las células, y
- b) dicho L-aminoácido se aísla, manteniéndose opcionalmente los constituyentes del caldo de fermentación y/o la biomasa en su totalidad o en porciones (desde ≥ 0 a 100%) en el producto aislado o eliminándose por completo.

Los microorganismos que tienen un marco de lectura abierto (ORF) sobreexpresado designado yaaU, y que son en particular recombinantes, son análogamente parte de la materia que constituye el objeto de la presente invención, pueden producir dichos L-aminoácidos a partir de glucosa, sacarosa, lactosa, fructosa, maltosa, melazas, en caso apropiado almidón y en caso apropiado celulosa, o a partir de glicerol y etanol. Los microorganismos son representativos de la familia Enterobacteriáceas y se seleccionan de los géneros *Escherichia*, *Erwinia*, *Providencia* y *Serratia*. Se prefieren los géneros *Escherichia* y *Serratia*. La especie *Escherichia coli* puede mencionarse, en particular, en el caso del género *Escherichia*, mientras que la especie *Serratia marcescens* puede mencionarse, en particular, en conexión con el género *Serratia*.

En general, los microorganismos recombinantes se generan por medio de transformación, transducción o conjugación, o una combinación de estos métodos, con un vector que contiene el ORF deseado, el gen deseado, un alelo de este ORF o gen, o partes de los mismos, y/o un promotor que aumenta la expresión del ORF o gen. Este promotor puede ser el promotor que se ha producido por intensificación de la mutación de la secuencia reguladora endógena localizada aguas arriba del gen u ORF; alternativamente, un promotor eficiente se ha fusionado al gen u ORF.

Ejemplos de cepas adecuadas, que producen en particular L-treonina, del género *Escherichia*, en particular de la especie *Escherichia coli* son

- *Escherichia coli* H4581 (EP 0 301 572)
- *Escherichia coli* KY10935 (Bioscience Biotechnology and Biochemistry 61(11):1877-1882 (1997))
- *Escherichia coli* VNIIGenetica MG442 (US-A-4.278.765)
- *Escherichia coli* VNIIGenetica M1 (US-A-4.321.325)
- *Escherichia coli* VNIIGenetica 472T23 (US-A- 5.631.157)
- *Escherichia coli* BKIIM B-3996 (US-A-5.175.107)
- *Escherichia coli* cat 13 (WO 98/04715)
- *Escherichia coli* KCCM-10132 (WO 00/09660)

Ejemplos de cepas productoras de L-treonina del género *Serratia*, en particular de la especie *Serratia marcescens*, son

- *Serratia marcescens* HNr21 (Applied and Environmental Microbiology 38(6): 1045-1051 (1979))

ES 2 343 010 T3

- *Serratia marcescens* TLR156 (Gene 57(2-3): 151-158 (1987))
- *Serratia marcescens* T-2000 (Applied Biochemistry and Biotechnology 37(3): 255- 265 (1992))

5

Las cepas productoras de L-treonina de la familia Enterobacteriáceas poseen preferiblemente, entre otras cosas, una o más de las características genéticas o fenotípicas seleccionadas del grupo: resistencia a ácido α -amino- β -hidroxivalérico, resistencia a tialisina, resistencia a etionina; resistencia a α -metilserina, resistencia a ácido diaminosuccínico, resistencia a ácido α -aminobutírico, resistencia a borrelidina, resistencia a ácido ciclopentanocarboxílico, resistencia a rifampicina, resistencia a análogos de valina tales como valina hidroxamato, resistencia a análogos de purina, tales como 6-dimetilaminopurina, requerimiento de L-metionina, posible requerimiento parcial y compensable de L-isoleucina, requerimiento de ácido mesodiaminopimélico, auxotrofia en relación con dipéptidos que contienen treonina, resistencia a L-treonina, resistencia a refinado de treonina, resistencia a L-homoserina, resistencia a L-lisina, resistencia a L-metionina, resistencia a ácido L-glutámico, resistencia a L-aspartato, resistencia a L-leucina, resistencia a L-fenilalanina, resistencia a L-serina, resistencia a L-cisteína, resistencia a L-valina, sensibilidad a fluoropiruvato, deficiencia en treonina-deshidrogenasa, posible aptitud para utilizar sacarosa, intensificación del operón treonina, intensificación de homoserina deshidrogenasa I-aspartato-quinasa I, preferiblemente de la forma resistente a la realimentación, intensificación de homoserina-quinasa, intensificación de treonina-sintasa, intensificación de aspartato-quinasa, posiblemente de la forma resistente a la realimentación, intensificación de aspartato-semialdehído-deshidrogenasa, intensificación de fosfoenolpiruvato-carboxilasa, posiblemente de la forma resistente a la realimentación, intensificación de fosfoenolpiruvato-sintasa, intensificación de transhidrogenasa, intensificación del producto del gen RhtB, intensificación del producto del gen RhtC, intensificación del producto del gen YfiK, intensificación de una piruvato-carboxilasa y atenuación de la formación de ácido acético.

25 Se ha encontrado que, después de la sobreexpresión del gen o el marco de lectura abierto (ORF) yaaU, o sus alelos, los microorganismos de la familia Enterobacteriáceas producen L-lisina, L-valina o L-treonina, de una manera incrementada.

30 Las secuencias de nucleótidos de los genes o marcos de lectura abiertos (ORFs) de *Escherichia coli* pertenecen a la técnica anterior y pueden obtenerse de la secuencia del genoma de *Escherichia coli* publicada por Blattner *et al.* (Science 277: 1453-1462 (1997)). Se sabe que las enzimas endógenas (metionina-aminopeptidasa) son capaces de escindir el aminoácido N-terminal metionina.

35 Las secuencias de nucleótidos para el ORF yaaU de *Shigella flexneri* y *Salmonella typhimurium*, que pertenecen análogamente a la familia Enterobacteriáceas, han sido también descritas.

El ORF yaaU de *Escherichia coli* K12 se describe, entre otras cosas, por los datos siguientes:

40 Designación:	marco de lectura abierto
Función:	apostillado como una proteína supuesta de transporte, un transportador de membrana de la MFS (superfamilia facilitadora principal). Los sustratos transportados varían notablemente; la familia MFS contiene transportadores de monosacáridos, disacáridos y oligosacáridos, transportadores de potasio y aminoácidos, transportadores de compuestos intermedios del ciclo de los ácidos tricarboxílicos, así como transportadores que bombean antibióticos al exterior de las células.
45 Descripción:	el marco de lectura abierto yaaU codifica una proteína de 48,7 KDa; el punto isoelectrónico es 8,8; cuando está localizado en el cromosoma, yaaU está presente, por ejemplo en el caso de <i>Escherichia coli</i> K12 MG1655, en la región intergénica de los genes carB, que codifican la cadena larga de carbamoil-fosfato-sintasa, y kefC, que codifica una proteína del sistema de flujo de potasio regulado por (K+)/H(+) glutatión);
50 Referencia:	Blattner <i>et al.</i> , Science 277 (5331): 1453-1474 (1997)
55 No. de Acceso:	AE000114
60 Nombre de gen alternativo:	B0045

Las secuencias de ácido nucleico pueden obtenerse de las bases de datos pertenecientes al National Center for Biotechnology Information (NCBI) de la National Library de Medicine (Bethesda, MD, EE.UU.), la base de datos de secuencias de ácido nucleico de los European Molecular Biology Laboratories (EMBL, Heidelberg, Alemania o Cambridge, Reino Unido) o la base de datos de DNA japonesa (DDBJ, Mishima, Japón).

Para mayor claridad, la secuencia de nucleótidos conocida del ORF yaaU de *Escherichia coli* se da en SEQ ID No. 3 y las secuencias conocidas para el ORF yaaU de *Shigella flexneri* (AE015041) y *Salmonella typhimurium*

ES 2 343 010 T3

(AE008697) se dan bajo SEQ ID No. 5 y, respectivamente, SEQ ID No. 7. Las secuencias de aminoácidos de las proteínas codificadas por estos marcos de lectura se representan como SEQ ID No. 4, SEQ ID No. 6 y, respectivamente, SEQ ID No. 8.

5 Los marcos de lectura abiertos descritos en los pasajes indicados pueden utilizarse de acuerdo con la invención. Adicionalmente, es posible utilizar alelos de los genes o marcos de lectura abiertos, que son resultado de la degeneración del código genético o como consecuencia de mutaciones sentido funcionalmente neutras. Se da preferencia a la utilización de genes endógenos o marcos de lectura abiertos endógenos.

10 “Genes endógenos” o “secuencias nucleotídicas endógenas” deben entenderse como los genes o marcos de lectura abiertos o alelos o secuencias de nucleótidos que están presentes en una población de especie.

15 Los alelos adecuados del ORF yaaU, que contienen mutaciones sentido funcionalmente neutras incluyen, entre otras, aquéllas que conducen a, como máximo, 50 o como máximo 40 o como máximo 30 o como máximo 20, preferiblemente a, como máximo, 10 o a como máximo 5, y de modo muy particularmente preferible a como máximo 3 o a como máximo 2, o a al menos una, sustitución conservadora de aminoácido en la proteína que codifican los mismos.

20 En el caso de los aminoácidos aromáticos, se dice que las sustituciones son conservadoras cuando fenilalanina, triptófano y tirosina se sustituyen uno por otro. En el caso de los aminoácidos hidrófobos, se dice que las sustituciones son conservadoras cuando leucina, isoleucina y valina se sustituyen uno por otro. En el caso de los aminoácidos polares, se dice que las sustituciones son conservadoras cuando se sustituyen glutamina y asparagina uno por otro. En el caso de los aminoácidos básicos, se dice que las sustituciones son conservadoras cuando se sustituyen arginina, lisina e histidina uno por otro. En el caso de los aminoácidos ácidos, se dice que las sustituciones son conservadoras cuando se sustituyen ácido aspártico y ácido glutámico uno por otro. En el caso de los aminoácidos que contienen grupos hidroxilo, se dice que las sustituciones son conservadoras cuando se sustituyen serina y treonina uno por otro.

25 De igual manera, es asimismo posible utilizar secuencias de nucleótidos que codifican variantes de dichas proteínas, variantes que contienen adicionalmente una extensión o truncación de al menos un (1) aminoácido en el término N o el término C. Esta extensión o truncación asciende a no más de 50, 40, 30, 20, 10, 5, 3 ó 2 aminoácidos o residuos de aminoácidos.

30 Los alelos adecuados incluyen también aquéllos que codifican proteínas en las cuales al menos un (1) aminoácido ha sido insertado o delecionado. El número máximo de tales cambios, denominados indels, puede afectar a 2, 3, 5, 10, 20, pero en ningún caso más de 30 aminoácidos.

35 Los alelos adecuados incluyen adicionalmente aquéllos que pueden obtenerse por medio de hibridación, en particular en condiciones severas, utilizando SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7 o partes de las mismas, en particular las regiones codificantes o las secuencias que son complementarias a ellas.

40 La persona experta puede encontrar instrucciones para la identificación de secuencias de DNA por medio de hibridación, entre otros lugares, en el manual “The DIG System Users Guide for Filter Hybridization” suministrado por Boehringer Mannheim GmbH (Mannheim, Alemania, 1993) y Liebl *et al.* (International Journal de Systematic Bacteriology 41: 255-260 (1991)). La hibridación tiene lugar en condiciones severas, es decir, que los únicos híbridos formados son aquéllos en los cuales la sonda y la secuencia diana, es decir los polinucleótidos tratados con la sonda, son idénticos al menos en un 80%. Es sabido que la severidad de la hibridación, con inclusión de los pasos de lavado, está influenciada y/o determinada por variación de la composición de los tampones, la temperatura y la concentración de sales. En general, la reacción de hibridación se lleva a cabo a una severidad que es relativamente baja comparada con la de los pasos de lavado (Hybaid Hybridization Guide, Hybaid Limited, Teddington, Reino Unido, 1996).

50 Por ejemplo, un tampón correspondiente al tampón 5x SSC puede utilizarse para la reacción de hibridación a una temperatura de aprox. 50°C-68°C. En estas condiciones, las sondas pueden hibridarse también con polinucleótidos que poseen menos de 70% de identidad con la secuencia de la sonda. Estos híbridos son menos estables y se eliminan por lavado en condiciones severas. Esto puede conseguirse, por ejemplo, por disminución de la concentración de sal hasta por debajo de 2 x SSC y, en caso apropiado, subsiguientemente a 0,5 x SSC (The DIG System User’s Guide for Filter Hybridization, 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, o aprox. 66°C-68°C. Se prefieren intervalos de temperatura de aprox. 64°C-68°C, o aprox. 66°C-68°C. Es posible, en caso apropiado, reducir la concentración de sales hasta una concentración correspondiente a 0,2 x SSC o 0,1 x SSC. Por medio del aumento gradual de la temperatura de hibridación, en escalones de aprox. 1-2°C, desde 50°C a 68°C, es posible aislar fragmentos polinucleotídicos que, por ejemplo, poseen al menos 80%, o al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o al menos 99%, de identidad con la secuencia de la sonda empleada o con las secuencias de nucleótidos que se muestran en SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7. Instrucciones adicionales para la hibridación pueden obtenerse comercialmente en forma de lo que se denominan kits (tales como DIG width Easy Hyb, de Roche Diagnostics GmbH, Mannheim, Alemania, No. de Catálogo 1603558). Las secuencias de nucleótidos que se obtienen de este modo codifican polipéptidos que poseen al menos 90%, en particular al menos 95%, preferiblemente al menos 98% o al menos 99%, de modo muy particularmente preferible 99,7%, de identidad con las secuencias de aminoácidos representadas en SEQ ID No. 4, SEQ ID No. 6 o SEQ ID No. 8.

ES 2 343 010 T3

A fin de lograr la intensificación, es posible, por ejemplo, aumentar la expresión de los genes o marcos de lectura abiertos o alelos, o aumentar las propiedades catalíticas de la proteína. Ambas medidas pueden combinarse, en caso apropiado.

5 Para conseguir la sobreexpresión, puede aumentarse el número de copias de los genes o marcos de lectura abiertos correspondientes, por ejemplo, o pueden mutarse la región promotora y la región reguladora o el sitio de fijación de ribosoma que está localizado aguas arriba del gen estructural. Casetes de expresión que se incorporan aguas arriba del gen estructural actúan del mismo modo. Es asimismo posible aumentar la expresión durante el curso de la producción fermentativa de L-treonina por incorporación de promotores inducibles; adicionalmente, la utilización de promotores para la expresión de genes que permite una expresión cronológica diferente de los genes puede ser también ventajosa. La expresión se mejora análogamente por medio de medidas para ampliar la duración de vida del mRNA. Adicionalmente, la actividad enzimática puede mejorarse también impidiendo que la enzima proteínica se descomponga. Los ORFs, genes o constructos génicos pueden o bien estar presentes en plásmidos que tienen números de copias diferentes, o estar integrados, y amplificados, en el cromosoma. Alternativamente, la sobreexpresión de los genes en cuestión puede conseguirse también por alteración de la composición de los medios y la realización del cultivo.

Métodos para sobreexpresión se describen adecuadamente en la técnica anterior, por ejemplo en Makrides *et al.* (Microbiological Reviews 60(3), 512-538 (1996)). La utilización de vectores aumenta el número de copias al menos en una (1) copia. Los vectores utilizados pueden ser plásmidos como ha sido descrito, por ejemplo, en US 5.538.873. Los vectores utilizados pueden ser también fagos, por ejemplo el fago Mu, como ha sido descrito en EP 0332448, o el fago lambda (λ). El número de copias puede aumentarse también por incorporación de una copia adicional en otro sitio del cromosoma, por ejemplo, en el sitio att del fago λ (Yu y Court, Gene 223, 77-81 (1998)). US 5.939.307 informa que fue posible aumentar la expresión por incorporación de casetes o promotores de expresión, tales como el promotor tac, el promotor trp, el promotor lpp, o el promotor P_L o el promotor P_R del fago lambda (λ), aguas arriba, por ejemplo, del operón cromosómico treonina. De igual manera, es posible utilizar los promotores de fago T7, los promotores "gearbox" o el promotor nar. Tales casetes de expresión o promotores pueden utilizarse también, como ha sido descrito en EP 0 593 792, para sobreexpresar genes unidos a plásmidos. La utilización del alelo lacI^Q hace posible a su vez controlar la expresión de genes unidos a plásmidos (Glascock y Weickert, Gene 223, 221-231 (1998)). Es adicionalmente posible aumentar la actividad de los promotores por modificación de su secuencia por medio de una o más sustituciones de nucleótidos, por medio de una o más inserciones y/o por medio de una o más deleciones. Puede conseguirse una expresión génica cronológica diferente, por ejemplo, como ha sido descrito en Walker *et al.* (Journal de Bacteriology 181: 1269-80 (1999)), por utilización del promotor fis dependiente de la fase de crecimiento.

La persona experta puede encontrar instrucciones generales a este respecto, entre otros lugares, en Chang y Cohen (Journal de Bacteriology 134: 1141-1156 (1978)), Hartley y Gregori (Gene 13: 347-353 (1981)), Amann and Brosius (Gene 40: 183-190 (1985)), de Broer *et al.* (Proceedings de the National Academy de Sciences de the United States de America 80: 21-25 (1983)), LaVallie *et al.* (BIO/TECHNOLOGY 11: 187-193 (1993)), en PCT/US97/13359, Llosa *et al.* (Plasmid 26: 222-224 (1991)), Quandt y Klipp (Gene 80: 161-169 (1989)), Hamilton *et al.* (Journal de Bacteriology 171: 4617-4622 (1989)), Jensen y Hammer (Biotechnology and Bioengineering 58: 191-195 (1998)) y libros de texto conocidos de genética y biología molecular.

Pueden utilizarse vectores plasmídicos que se pueden replicar en Enterobacteriáceas, tales como vectores de clonación derivados de pACYC184 (Bartolomé *et al.*; Gene 102: 75-78 (1991)), pTrc99A (Amann *et al.*; Gene 69: 301-315 (1988)) o derivados de pSC101 (Vocke y Bastia; Proceedings de the National Academy de Sciences USA 80(21): 6557-6561 (1983)). En un proceso de acuerdo con la invención, es posible utilizar una cepa que se transforma con un vector plasmídico que lleva al menos una secuencia de nucleótidos, o alelo, que codifica el ORF yaaU o su producto génico.

El término "transformación" debe entenderse con el significado de la captura de un ácido nucleico aislado por un hospedador (microorganismo).

Es asimismo posible utilizar intercambios de secuencia (Hamilton *et al.*, Journal de Bacteriology 171: 4617-4622 (1989)), conjugación o transducción para transferir mutaciones, que afectan a la expresión de los genes o marcos de lectura abiertos dados, en cepas diferentes.

Explicaciones más detalladas de los conceptos de genética y biología molecular pueden encontrarse en libros de texto conocidos de genética y biología molecular tales como el libro de texto de Birge (Bacterial and Bacteriophage Genetics, 4ª edición, Springer Verlag, Nueva York (EE.UU.), 2000) o el libro de texto de Berg, Tymoczko y Stryer (Biochemistry, 5ª edición, Freeman and Company, Nueva York (EE.UU.), 2002) o el manual de Sambrook *et al.* (Molecular Cloning, A Laboratory Manual, (Serie de 3 volúmenes), Cold Spring Harbor Laboratory Press, Cold Spring Harbor (EE.UU.), 2001).

Adicionalmente, cuando se utilizan cepas de la familia Enterobacteriáceas para producir L-lisina, L-valina o L-treonina, puede ser ventajoso, además de sobreexpresar el marco de lectura abierto yaaU, sobreexpresar una o más enzimas del camino conocido de la biosíntesis de la treonina, o enzimas del metabolismo anaplerótico, o enzimas para producir nicotinamida-adenina-dinucleótido-fosfato reducido, o enzimas de la glucólisis, o enzimas PTS o enzimas de metabolismo del azufre. Se prefiere generalmente la utilización de genes endógenos.

ES 2 343 010 T3

Así, es posible, por ejemplo, sobreexpresar simultáneamente, uno o más de los genes seleccionados del grupo

- 5 • al menos un gen del operón thrABC que codifica aspartato-quinasa, homoserina-deshidrogenasa, homoserina-quinasa y treonina-sintasa (US-A-4.278.765),
- el gen pyc de *Corynebacterium glutamicum* que codifica piruvato-carboxilasa (WO 99/18228),
- el gen pps que codifica fosfoenolpiruvato-sintasa (Molecular and General Genetics 231(2), 332-336 (1992)),
- 10 • el ppc gen que codifica fosfoenolpiruvato carboxilasa (WO 02/064808),
- los genes pntA y pntB que codifican las subunidades de piridina-transhidrogenasa (European Journal de Biochemistry 158, 647-653 (1986)),
- 15 • el gen rhtB que codifica la proteína que confiere resistencia a homoserina (EP-A-0 994 190),
- el gen rhtC que codifica la proteína que confiere resistencia a treonina (EP-A-1 013 765),
- 20 • el gen thrE de *Corynebacterium glutamicum* que codifica la proteína del portador de exportación de treonina (WO 01/92545),
- el gdhA gen que codifica glutamato-deshidrogenasa (Nucleic Acids Research 11, 5257-5266 (1983); Gene 23, 199-209 (1983)),
- 25 • el gen pgm que codifica fosfoglucomutasa (WO 03/004598),
- el gen fba que codifica fructosa-bifosfato-aldolasa (WO 03/004664),
- 30 • el gen ptsH del operón ptsHIcrr que codifica la proteína fosfohistidina-hexosa-fosfotransferasa del sistema PTS de fosfotransferasa (WO 03/004674),
- el gen ptsI del operón ptsHIcrr que codifica la enzima I del sistema PTS de fosfotransferasa (WO 03/004674),
- 35 • el gen crr del operón ptsHIcrr que codifica el componente IIA específico de glucosa del sistema PTS de fosfo-transferasa (WO 03/004674),
- el gen ptsG que codifica el componente IIBC específico de glucosa (WO 03/004670),
- 40 • el gen lrp que codifica el regulador del regulón leucina (WO 03/004665),
- el gen fadR que codifica el regulador del regulón fad (WO 03/038106),
- el gen iclR que codifica el regulador del metabolismo central intermedio (WO 03/038106),
- 45 • el gen ahpC del operón ahpCF que codifica la subunidad pequeña de alquil-hidroperóxido-reductasa (WO 03/004663),
- el gen ahpF del operón ahpCF que codifica la subunidad grande de alquil-hidroperóxido-reductasa (WO 03/004663),
- 50 • el gen cysK que codifica cisteína-sintasa A (WO 03/006666),
- el gen cysB que codifica el regulador del regulón cys (WO 03/006666),
- 55 • el gen cysJ del operón cysJIH que codifica la flavoproteína NADPH- sulfito-reductasa (WO 03/006666),
- el gen cysI del operón cysJIH que codifica la hemoproteína NADPH-sulfito-reductasa (WO 03/006666),
- 60 • el gen cysH del operón cysJIH que codifica adenilil-sulfato-reductasa (WO 03/006666),
- el gen rseA del operón rseABC que codifica una proteína de membrana que posee actividad anti-sigmaE (WO 03/008612),
- 65 • el gen rseC del operón rseABC que codifica un regulador global del factor sigmaE (WO 03/008612),

ES 2 343 010 T3

- el gen *sucA* del operón *sucABCD* que codifica la subunidad descarboxilasa de 2-cetoglutarato-deshidrogenasa (WO 03/008614),
- el gen *sucB* del operón *sucABCD* que codifica la subunidad dihidrolipoil-transsuccinasa E2 de 2-cetoglutarato-deshidrogenasa (WO 03/008614),
- el gen *sucC* del operón *sucABCD* que codifica la subunidad β de succinil-CoA-sintetasa (WO 03/008615),
- el gen *sucD* del operón *sucABCD* que codifica la subunidad α de succinil-CoA-sintetasa (WO 03/008615),
- el gen *aceE* que codifica el componente E1 del complejo piruvato-deshidrogenasa (WO 03/076635),
- el gen *aceF* que codifica el componente E2 del complejo piruvato-deshidrogenasa (WO 03/076635),
- el gen *rseB* que codifica el regulador de la actividad del factor sigmaE (*Molecular Microbiology* 24(2), 355-371 (1997)),
- el producto génico del marco de lectura abierto (ORF) *yodA* de *Escherichia coli* (Número de Acceso AE000288 del National Center for Biotechnology Information (NCBI, Bethesda, MD, EE.UU.), DE10361192.4).

Adicionalmente, para el propósito de producción de L-aminoácidos, en particular L-treonina, puede ser ventajoso, además de intensificar el marco de lectura abierto *yaaU*, atenuar, en particular eliminar o reducir la expresión de uno o más de los genes seleccionados del grupo

- el gen *tdh* que codifica treonina-deshidrogenasa (*Journal de Bacteriology* 169, 4716-4721 (1987)),
- el gen *mdh* que codifica malato-deshidrogenasa (E.C.1.1.1.37) (*Archives in Microbiology* 149: 36-42 (1987)),
- el producto génico del marco de lectura abierto (ORF) *yjFA* de *Escherichia coli* (Número de Acceso AAC77180 del National Center for Biotechnology Information (NCBI, Bethesda, MD, EE.UU. WO 02/29080)),
- el producto génico del marco de lectura abierto (ORF) *ytfP* de *Escherichia coli* (Número de Acceso AAC77179 del National Center for Biotechnology Information (NCBI, Bethesda, MD, EE.UU. WO 02/29080)),
- el gen *pckA* que codifica la enzima fosfoenolpiruvato-carboxiquinasa (WO 02/29080),
- el gen *poxB* que codifica piruvato-oxidasa (WO 02/36797),
- el gen *dgsA* (WO 02/081721), que es conocido también como el gen *mlc*, que codifica el regulador DgsA del sistema de fosfotransferasa,
- el gen *fruR* (WO 02/081698), que es conocido también como el gen *cra*, que codifica el represor fructosa,
- el gen *rpoS* (WO 01/05939), que es conocido también como el gen *katF*, que codifica el factor sigma³⁸, y
- el gen *aspA* que codifica aspartato-amonio-liasa (WO 03/008603).

En este contexto, el término “atenuación” describe la disminución o desactivación, en un microorganismo, de la actividad o concentración intracelular de una o más enzimas o proteínas que son codificadas por el DNA correspondiente, por ejemplo utilizando un promotor más débil que en el microorganismo o cepa parental no recombinante para la enzima o proteína correspondiente, o un gen o alelo que codifica una enzima o proteína correspondiente que tiene actividad menor, o por desactivación de la enzima o proteína correspondiente, o el marco de lectura abierto o gen y en caso apropiado, combinación de estas medidas.

En general, las medidas de atenuación reducen la actividad o concentración de la proteína correspondiente desde 0 a 75%, desde 0 a 50%, desde 0 a 25%, desde 0 a 10% o desde 0 a 5% de la actividad o concentración de la proteína de tipo salvaje o de la actividad o concentración de la proteína para la cepa parental o microorganismo que no es recombinante para la enzima o proteína correspondiente. La cepa o microorganismo parental que no es recombinante se entiende como el microorganismo sobre el cual se realizan las medidas de acuerdo con la invención.

Con objeto de conseguir una atenuación, por ejemplo la expresión de los genes o marcos de lectura abiertos, o las propiedades catalíticas de las proteínas enzimáticas, pueden reducirse o anularse. En caso apropiado, pueden combinarse ambas medidas.

La expresión de los genes puede reducirse por realización del cultivo de una manera adecuada, por modificación genética (mutación) de las estructuras señal para la expresión génica o bien por la técnica del RNA antisentido. Es-

estructuras señal para la expresión génica son tales como genes represores, genes activadores, operadores, promotores, atenuadores, sitios de fijación de ribosoma, el codón de partida y los terminadores. Los expertos encontrarán información a este respecto, inter alia, por ejemplo en Jensen y Hammer (*Biotechnology and Bioengineering* 58, 191-195 (1998)), en Carrier y Keasling (*Biotechnology Progress* 15, 58-64 (1999)), en Franch y Gerdes (*Current Opinion in Microbiology* 3, 159-164 (2000)) y en libros de texto bien conocidos de genética y biología molecular, tales como el libro de texto de Knippers (“*Molekulare Genetik [Molecular Genetics]*”, 6ª edición, Georg Thieme Verlag, Stuttgart, Alemania, 1995) o el de Winnacker (“*Gene und Klone [Genes and Clones]*”, VCH Verlagsgesellschaft, Weinheim, Alemania, 1990).

Mutaciones que dan como resultado una modificación o reducción en las propiedades catalíticas de las proteínas enzimáticas son conocidas por la técnica anterior. Ejemplos que pueden mencionarse son los artículos de Qiu y Goodman (*Journal de Biological Chemistry* 272: 8611-8617 (1997)), Yano *et al.* (*Proceedings de the National Academy de Sciences de the United States de America* 95: 5511-5515 (1998)) y Wentz y Schachmann (*Journal de Biological Chemistry* 266: 20833-20839 (1991)). Pueden encontrarse revisiones en libros de textos bien conocidos de genética y biología molecular, por ejemplo el publicado por Hagemann (“*Allgemeine Genetik [General Genetics]*”, Gustav Fischer Verlag, Stuttgart, 1986).

Mutaciones que entran en consideración son transiciones, transversiones, inserciones y deleciones de al menos un (1) par de bases o nucleótido. Dependiendo del efecto de la sustitución de aminoácido causada por la mutación sobre la actividad enzimática, se hace referencia a mutaciones de sentido erróneo o mutaciones sin sentido. Una mutación de sentido erróneo da como resultado un cambio de un aminoácido dado en una proteína por un aminoácido diferente, siendo dicho cambio de aminoácido particularmente no conservador. Esto deteriora por tanto la aptitud o actividad funcional de la proteína y reduce la misma a un valor de 0 a 75%, 0 a 50%, 0 a 25%, 0 a 10% o 0 a 5%. Una mutación sin sentido da como resultado un codón de parada en la región codificante del gen y por tanto una terminación prematura de la traducción. Inserciones o deleciones de al menos un par de bases en un gen conducen a mutaciones de desplazamiento de marco, que dan a su vez como resultado la incorporación de aminoácidos incorrectos o la terminación prematura de la traducción. Si se forma un codón de parada en la región codificante como consecuencia de la mutación, esto conduce también a la terminación prematura de la traducción. Deleciones de al menos uno (1) o más codones conducen también típicamente a la pérdida completa de la actividad enzimática.

Instrucciones acerca de la generación de estas mutaciones pertenecen a la técnica anterior y pueden encontrarse en libros de texto bien conocidos de genética y biología molecular, tales como el libro de texto de Knippers (“*Molekulare Genetik [Molecular Genetics]*”, 6ª edición, Georg Thieme Verlag, Stuttgart, Alemania, 1995), el de Winnacker (“*Gene und Klone [Genes and Clones]*”, VCH Verlagsgesellschaft, Weinheim, Alemania, 1990) o el de Hagemann (“*Allgemeine Genetik [General Genetics]*”, Gustav Fischer Verlag, Stuttgart, 1986).

Mutaciones adecuadas en los genes pueden incorporarse en cepas adecuadas por cambio de genes o alelos. Un método habitual es el método descrito por Hamilton *et al.* (*Journal de Bacteriology* 171: 4617-4622 (1989)) de cambio de gen utilizando un derivado de pSC101, pMAK705, que se replica condicionalmente. Pueden utilizarse también otros métodos descritos en la técnica anterior, por ejemplo el de Martínez-Morales *et al.* (*Journal de Bacteriology* 181: 7143-7148 (1999)) o el de Boyd *et al.* (*Journal de Bacteriology* 182: 842-847 (2000)).

Es asimismo posible transferir mutaciones en los genes relevantes, o mutaciones que afectan a la expresión de los genes o marcos de lectura abiertos relevantes, a cepas diferentes por conjugación o transducción.

Adicionalmente, para el propósito de producción de L-lisina, L-valina o L-treonina, puede ser ventajoso, además de intensificar el marco de lectura abierto yaaU, eliminar reacciones secundarias indeseables (Nakayama: “*Breeding de Amino Acid Producing Microorganisms*”, en: *Overproduction de Microbial Products*, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, Londres, Reino Unido, 1982).

Los microorganismos que se preparan de acuerdo con la invención pueden cultivarse en un proceso de lotes, en un proceso de lote alimentado (“fed batch”), en un proceso de lote alimentado repetido, o en un proceso continuo (DE 102004028859.3 o US 5.763.230). Métodos de cultivo conocidos se proporcionan en el libro de texto de Chmiel (*Bioprozesstechnik 1. Einführung in die Bioverfahrenstechnik [Bioprocess Technology 1. Introduction to bioprocess technology]* (Gustav Fischer Verlag, Stuttgart, 1991)) o en el libro de texto de Storhas (*Bioreaktoren und periphere Einrichtungen [Bioreactors and peripheral installations]* (Vieweg Verlag, Brunswick/Wiesbaden, 1994)).

El medio de cultivo a utilizar debe satisfacer los requerimientos de las cepas particulares de manera apropiada. El manual de la American Society for Bacteriology “*Manual de Methods for General Bacteriology*” (Washington DC, EE.UU., 1981) contiene descripciones de medios de cultivo para diversos microorganismos.

Como fuente de carbono se pueden utilizar azúcares y carbohidratos, tales como glucosa, sacarosa, lactosa, fructosa, maltosa, melazas, almidón y en caso apropiado celulosa, aceites y grasas, tales como aceite de soja, aceite de girasol, aceite de cacahuete y grasa de coco, ácidos grasos, tales como ácido palmítico, ácido esteárico y ácido linoleico, alcoholes, tales como glicerol y etanol, y ácidos orgánicos, tales como ácido acético. Estas sustancias pueden utilizarse individualmente o en forma de mezcla.

ES 2 343 010 T3

Como fuente de nitrógeno se pueden utilizar compuestos orgánicos nitrogenados tales como peptonas, extracto de levadura, extracto de carne, extracto de malta, licor de maceración de maíz, harina 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 pueden utilizarse individualmente o en forma de mezcla.

5 Como fuente de fósforo se pueden utilizar ácido fosfórico, dihidrogenofosfato de potasio o hidrogenofosfato dipotásico, o las sales de sodio correspondientes. El medio de cultivo tiene que contener además sales metálicas, tales como sulfato de magnesio o sulfato de hierro, que son necesarias para el crecimiento. Finalmente, pueden utilizarse promotores esenciales del crecimiento tales como aminoácidos y vitaminas, además de las sustancias arriba mencionadas. Pueden añadirse también al medio de cultivo precursores adecuados. Dichos ingredientes pueden añadirse al cultivo en forma de una mezcla única, o dosificarse adecuadamente durante el cultivo.

15 La fermentación se lleva a cabo generalmente a un pH de 5,5 a 9,0, particularmente de 6,0 a 8,0. Para controlar el pH del cultivo se utilizan de manera adecuada compuestos básicos tales como hidróxido de sodio, hidróxido de potasio, amoníaco o agua amoniacoal, o compuestos ácidos tales como ácido fosfórico o ácido sulfúrico. Para controlar la formación de espuma pueden utilizarse antiespumantes tales como poliglicolésteres de ácidos grasos. La estabilidad de los plásmidos puede mantenerse por adición al medio de sustancias de acción selectiva adecuadas, por ejemplo antibióticos. Para el mantenimiento de condiciones aerobias se introducen en el cultivo oxígeno o mezclas gaseosas que contienen oxígeno, por ejemplo aire. La temperatura del cultivo es normalmente de 25°C a 45°C y preferiblemente 20 30°C a 40°C. El cultivo se continúa hasta que la formación de L-aminoácidos o L-treonina ha alcanzado un máximo. Este objetivo se alcanza normalmente dentro de 10 horas a 160 horas.

Los L-aminoácidos pueden analizarse por cromatografía de intercambio aniónico seguida por derivatización con ninhidrina, como ha sido descrito por Spackman *et al.* (Analytical Chemistry 30: 1190-1206 (1958)), o por medio de HPLC en fase inversa, como ha sido descrito por Lindroth *et al.* (Analytical Chemistry 51: 1167-1174 (1979)).

El proceso de acuerdo con la invención puede utilizarse para preparar L-aminoácidos tales como L-treonina, L-valina y L-lisina, particularmente L-treonina, por fermentación.

30 El microorganismo siguiente ha sido depositado en la Deutsche Sammlung für Mikroorganismen und Zellkulturen [German collection de microorganisms and cell cultures] (DSMZ, Brunswick, Alemania) de acuerdo con el Convenio de Budapest:

- *Escherichia coli*, cepa *E. coli* MG442 como DSM 16574.

35

La presente invención se ilustra a continuación con mayor detalle con ayuda de ejemplos de implementación.

40 El medio mínimo (M9) y el medio completo (LB) utilizados para *Escherichia coli* han sido descritos por J.H. Miller, (A short course in bacterial genetics (1992), Cold Spring Harbor Laboratory Press). El aislamiento de DNA plasmídico a partir de *Escherichia coli*, así como todas las técnicas para restricción, ligación y tratamiento con fosfatasa Klenow y fosfatasa alcalina se llevan a cabo como se describe en Sambrook *et al.* (Molecular Cloning - A Laboratory Manual (1989), Cold Spring Harbor Laboratory Press). A no ser que se indique otra cosa, la transformación de *Escherichia coli* se lleva a cabo como se describe en Chung *et al.* (Proceedings de the National Academy de Sciences de the United States de America 86: 2172-2175 (1989)).

45

La temperatura de incubación en la preparación de cepas y transformantes es 37°C.

50 Ejemplo 1

Construcción del plásmido de expresión pTrc99A_{yaaU}

55 El ORF yaaU de *E. coli* K12 se amplifica utilizando la reacción en cadena de polimerasa (PCR) y oligonucleótidos sintéticos. Los iniciadores para la PCR se sintetizan (MWG Biotech, Ebersberg, Alemania) sobre la base de la secuencia de nucleótidos del ORF yaaU en *E. coli* K12 MG1655 (Número de Acceso AE000114), Blattner *et al.* (Science 277: 1453-1474 (1997)). Las secuencias de los iniciadores se modifican a fin de formar sitios de reconocimiento para las enzimas de restricción. La secuencia de reconocimiento de EcoRI se selecciona para el iniciador yaaU-ex1, y la secuencia de reconocimiento de BamHI se selecciona para el iniciador yaaU-ex2, subrayándose estas secuencias en las secuencias de nucleótidos que se muestran a continuación:

60

yaaU-ex1:

5'-GATCTGAATTCTAAGGAATAACCATGCAACCGTC-3' (SEQ ID No. 1)

65

ES 2 343 010 T3

yaaU-ex2:

5'-GATCTAGGATCCCAATTTACCCATTCTCTGC-3' (SEQ ID No. 2)

5 El DNA cromosómico de *E. coli* K12 MG1655 utilizado para la PCR se aísla utilizando “Qiagen Genomic-tips 100/G” (QIAGEN, Hilden, Alemania) de acuerdo con las instrucciones del fabricante. Un fragmento de DNA de tamaño aproximado 1371 pb (SEQ ID No. 3) puede amplificarse en condiciones estándar de PCR (Innis *et al.* (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press) using Vent DNA polymerase (New England Biolabs GmbH, Frankfurt, Alemania) y los iniciadores específicos.

15 El fragmento amplificado yaaU se liga al vector pCR-Blunt II-TOPO (Zero TOPO TA Cloning Kit, Invitrogen, Groningen, Holanda) de acuerdo con las instrucciones del fabricante, y se transforma en la cepa TOP10 de *E. coli*. Se seleccionan células que albergan plásmidos en agar LB que contiene 50 µg de kanamicina/ml. Después que se ha aislado el DNA plasmídico, el vector se escinde con las enzimas EcoRV y EcoRI y, después que se ha comprobado la escisión en un gel de agarosa al 0,8%, se designa pCRBluntyaaU.

20 El vector pCRBluntyaaU se escinde luego con las enzimas EcoRI y BamHI y el fragmento yaaU se separa en un gel de agarosa al 0,8%; se aísla luego del gel (QIAquick Gel Extraction Kit, QIAGEN, Hilden, Alemania) y se liga al vector pTrc99A (Pharmacia Biotech, Uppsala, Suecia) que se ha digerido con las enzimas BamHI y EcoRI. La cepa de *E. coli* XL1Blue MRF' (Stratagene, La Jolla, EE.UU.) se transforma con la mezcla de ligación y las células que albergan los plásmidos se seleccionan en agar LB que contiene 50 µg de ampicilina/ml.

25 El hecho de que la clonación ha tenido éxito puede demostrarse, después de aislamiento del DNA plasmídico, por realización de una escisión de control utilizando las enzimas EcoRI/BamHI y EcoRV.

El plásmido se designa pTrc99AyaaU (Figura 1).

30 Ejemplo 2

Preparación de L-treonina utilizando la cepa MG442/pTrc99AyaaU

35 La cepa de *E. coli* MG442 productora de L-treonina se describe en la memoria descriptiva de patente US-A-4.278.765 y ha sido depositada en la Colección Nacional Rusa de Microorganismos Industriales (VKPM, Moscú, Rusia) como CMIM B-1628 y en la Deutsche Sammlung für Mikroorganismen und Zellkulturen [German collection of microorganisms and cell cultures] (DSMZ, Brunswick, Alemania), de acuerdo con el Tratado de Budapest, como DSM 16574).

40 La cepa MG442 se transforma con el plásmido de expresión pTrc99AyaaU descrito en el Ejemplo 1, y con el vector pTrc99A, y las células que albergan el plásmido se seleccionan en agar LB que contiene 50 µg de ampicilina/ml. Esto da como resultado las cepas MG442/pTrc99AyaaU y MG442/pTrc99A. Las colonias individuales seleccionadas se propagan luego adicionalmente en medio mínimo que tiene la composición siguiente: 3,5 g de Na₂HPO₄*2H₂O/l, 1,5 g de KH₂PO₄/l, 1 g de NH₄Cl/l, 0,1 g de MgSO₄*7H₂O/l, 2 g de glucosa/l, 20 g de agar/l, 50 mg de ampicilina/l.

45 La formación de L-treonina se comprueba en cultivos de lotes de 10 ml que están contenidos en matraces Erlenmeyer de 100 ml. Para ello, se inocula un medio de precultivo de 10 ml de la composición siguiente: 2 g de extracto de levadura/10 g de (NH₄)₂SO₄/l, 1 g de KH₂PO₄/l, 0,5 g de MgSO₄*7H₂O/l, 15 g de CaCO₃/l, 20 g de glucosa/l, 50 mg de ampicilina/l, y se incuba, a 37°C y 180 rpm durante 16 horas, en una incubadora Kühner AG ESR (Birsfelden, Suiza). En cada caso, se inoculan 250 µl de este cultivo preliminar en 10 ml de medio de producción (25 g de (NH₄)₂SO₄/l, 2 g de KH₂PO₄/l, 1 g de MgSO₄*7H₂O/l, 0,03 g de FeSO₄*7H₂O/l, 0,018 g de MnSO₄*1H₂O/l, 30 g de CaCO₃/l, 20 g de glucosa/l, 50 mg de ampicilina/l) y se incuban a 37°C durante 48 horas. La formación de L-treonina por la cepa MG442 de partida se comprueba de igual manera excepto que no se añade cantidad alguna de ampicilina al medio. Después de la incubación, se determina la densidad óptica (DO) de la suspensión de cultivo a una longitud de onda de medición de 660 nm utilizando un fotómetro Dr. Lange LP2W (Düsseldorf, Alemania).

55 Se utiliza luego un analizador de aminoácidos Eppendorf-BioTronik (Hamburgo, Alemania) para determinar, por medio de cromatografía de intercambio iónico y reacción post-columna que implica detección con ninhidrina, la concentración de la L-treonina resultante en el sobrenadante de cultivo, que se ha esterilizado por filtración.

60

65

ES 2 343 010 T3

El resultado del experimento se muestra en la Tabla 1.

TABLA 1

Cepa	DO (660 nm)	L-treonina, g/l
MG442	5,6	1,4
MG442/pTrc99A	3	1,3
MG442/pTrc99AyaaU	4,1	2,7

Breve descripción de la figura

Figura 1: Mapa del plásmido pTrc99AyaaU que contiene el gen yaaU

Las especificaciones de longitud deben considerarse como aproximadas. Las abreviaturas y designaciones empleadas tienen los significados siguientes:

- bla: gen que codifica resistencia a ampicilina
- lac Iq: gen para la proteína represora del promotor trc
- trc: región promotora de trc, inducible por IPTG
- yaaU: región codificante del gen yaaU
- 5S: región de rRNA 5S
- rrnBT: región terminadora del rRNA

Las abreviaturas para las enzimas de restricción tienen el significado siguiente:

- BamHI: endonucleasa de restricción de *Bacillus amyloliquefaciens* H
- EcoRI: endonucleasa de restricción de *Escherichia coli* RY13
- EcoRV: endonucleasa de restricción de *Escherichia coli* B946.

ES 2 343 010 T3

REIVINDICACIONES

- 5 1. Un microorganismo recombinante de la familia Enterobacteriáceas que contiene un ORF yaaU sobreexpresado que codifica un polipéptido que está apostillado como un transportador de azúcar supuesto, en el cual la secuencia de aminoácidos del polipéptido es idéntica al menos en un 90% a una secuencia de aminoácidos seleccionada del grupo SEQ ID NO. 4, SEQ ID No. 6 y SEQ ID No. 8.
- 10 2. Un microorganismo de acuerdo con la reivindicación 1, **caracterizado** porque contiene un polinucleótido sobreexpresado yaaU que se selecciona del grupo:
- a) polinucleótido que tiene la secuencia de nucleótidos de SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7;
 - 15 b) polinucleótido que tiene una secuencia de nucleótidos que corresponde a SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7 dentro de los límites de la degeneración del código genético;
 - c) secuencia de polinucleótidos que tiene una secuencia que se hibrida, en condiciones severas, con la secuencia que es complementaria a la secuencia SEQ ID No. 3, SEQ ID No. 5 o SEQ ID No. 7.
- 20 3. Un microorganismo de acuerdo con la reivindicación 1, **caracterizado** porque el polipéptido posee una secuencia de aminoácidos que es idéntica al menos en un 95% a una de las secuencias seleccionadas del grupo SEQ ID No. 4, SEQ ID No. 6 y SEQ ID No. 8.
- 25 4. Un microorganismo de acuerdo con la reivindicación 1, **caracterizado** porque el polipéptido posee la secuencia de aminoácidos que es idéntica en un 100% a la de SEQ ID No. 4, SEQ ID No. 6 o SEQ ID No. 8.
- 30 5. Un microorganismo de acuerdo con las reivindicaciones 1 a 4, **caracterizado** porque se produce por transformación, transducción o conjugación, o una combinación de estos métodos, con un vector que contiene dicho ORF yaaU, un alelo de este ORF, o partes del mismo, y/o un promotor.
- 35 6. Un microorganismo de acuerdo con la reivindicación 1 ó 5, en el cual el número de copias del ORF yaaU o los alelos se ha incrementado al menos en 1.
- 40 7. El microorganismo de acuerdo con la reivindicación 6, **caracterizado** porque el aumento en el número de copias de dicho ORF yaaU en al menos 1 se consigue por integración de dicho ORF o los alelos del mismo en el cromosoma del microorganismo.
8. El microorganismo de acuerdo con la reivindicación 6, **caracterizado** porque el aumento en el número de copias de dicho ORF yaaU en al menos 1 se consigue por medio de un vector que se replica extracromosómicamente.
9. El microorganismo de acuerdo con la reivindicación 1 ó 2, **caracterizado** porque:
- 45 a) la región promotora y reguladora o el sitio de fijación de ribosoma está mutada(o) aguas arriba de dicho ORF yaaU, o
 - b) están incorporados casetes de expresión o promotores aguas arriba del ORF yaaU.
- 50 10. El microorganismo de acuerdo con la reivindicación 1 ó 2, **caracterizado** porque dicho polinucleótido yaaU se halla bajo el control de un promotor que intensifica la expresión del gen.
- 55 11. El microorganismo de acuerdo con las reivindicaciones 1 a 10, **caracterizado** porque la intensificación del ORF yaaU aumenta la concentración o actividad del producto (proteína) del gen yaaU al menos en un 10%, basada en la actividad o concentración del producto del gen en la cepa o microorganismo parental no recombinante para el ORF yaaU.
- 60 12. El microorganismo de acuerdo con las reivindicaciones 1 a 11, **caracterizado** porque otros genes del camino metabólico para la biosíntesis del L-aminoácido deseado están presentes también en forma sobreexpresada.
- 65 13. El microorganismo de acuerdo con la reivindicación 12, **caracterizado** porque el microorganismo se selecciona de los géneros Escherichia, Erwinia, Providencia y Serratia.
14. El microorganismo de acuerdo con las reivindicaciones 1 a 13, **caracterizado** porque produce L-treonina.
15. Un proceso para preparar L-aminoácidos seleccionados del grupo L-treonina, L-valina y L-lisina por fermentación de microorganismos recombinantes de la familia Enterobacteriáceas, **caracterizado** porque

ES 2 343 010 T3

- a) dichos microorganismos productores de L-aminoácidos de acuerdo con las reivindicaciones 1-14, se cultivan en un medio en condiciones en las cuales dicho L-aminoácido se enriquece en el medio o en las células, y
- 5 b) dicho L-aminoácido se aísla, quedando los constituyentes del caldo de fermentación, y/o la biomasa restante en su totalidad o en porciones (desde ≥ 0 a 100%) en el producto aislado o eliminándose por completo.

16. El proceso de acuerdo con la reivindicación 15, **caracterizado** porque, para el propósito de preparación de L-treonina, se fermentan microorganismos de la familia Enterobacteriáceas en los cuales uno o más de los genes seleccionados del grupo:

- 15 a) al menos un gen del operón thrABC que codifica aspartato-quinasa, homoserina-deshidrogenasa, homoserina-quinasa y treonina-sintasa,
- b) el gen pyc de *Corynebacterium glutamicum* que codifica piruvato-carboxilasa,
- c) el gen pps que codifica fosfoenolpiruvato-sintasa,
- 20 d) el ppc gen que codifica fosfoenolpiruvato carboxilasa,
- e) los genes pntA y pntB que codifican las subunidades de piridina-transhidrogenasa,
- f) el gen rhtB que codifica la proteína que confiere resistencia a homoserina,
- 25 g) el gen rhtC que codifica la proteína que confiere resistencia a treonina,
- h) el gen thrE de *Corynebacterium glutamicum* que codifica la proteína del portador de exportación de treonina,
- 30 i) el gdhA gen que codifica glutamato-deshidrogenasa,
- j) el gen pgm que codifica fosfoglucomutasa,
- 35 k) el gen fba que codifica fructosa-bifosfato-aldolasa,
- l) el gen ptsH que codifica la proteína fosfohistidina-hexosa-fosfotransferasa,
- m) el gen ptsI que codifica la enzima I del sistema de fosfotransferasa,
- 40 n) el gen crr que codifica el componente IIA específico de glucosa,
- o) el gen ptsG que codifica el componente IIBC específico de glucosa,
- 45 p) el gen lrp que codifica el regulador del regulón leucina,
- q) el gen fadR que codifica el regulador del regulón fad
- r) el gen iclR que codifica el regulador del metabolismo central intermedio,
- 50 s) el gen ahpF del operón ahpCF que codifica la subunidad pequeña de alquil-hidroperóxido-reductasa,
- t) el gen ahpF del operón ahpCF que codifica la subunidad grande de alquil-hidroperóxido-reductasa,
- 55 u) el gen cysK que codifica cisteína-sintasa A,
- v) el gen cysB que codifica el regulador del regulón cys,
- w) el gen cysJ que codifica la flavoproteína NADPH-sulfito-reductasa,
- 60 x) el gen cysI que codifica la hemoproteína NADPH-sulfito-reductasa,
- y) el gen cysH que codifica adenilil-sulfato-reductasa,
- 65 z) el gen rseA que codifica una proteína de membrana que posee actividad anti-sigmaE,
- a1) el gen rseC que codifica un regulador global del factor sigmaE,

ES 2 343 010 T3

- b1) el gen *sucA* que codifica la subunidad descarboxilasa de 2-cetoglutarato-deshidrogenasa,
c1) el gen *sucB* que codifica la subunidad dihidrolipoil-transsuccinasa E2 de 2-cetoglutarato-deshidrogenasa,
5 d1) el gen *sucC* que codifica la subunidad β de succinil-CoA-sintetasa,
e1) el gen *sucD* que codifica la subunidad α de succinil-CoA-sintetasa,
f1) el gen *aceE* que codifica el componente E1 del complejo piruvato-deshidrogenasa,
10 g1) el gen *aceF* que codifica el componente E2 del complejo piruvato-deshidrogenasa,
h1) el gen *rseB* que codifica el regulador de la actividad del factor sigmaE, y
15 i1) el producto génico del marco de lectura abierto (ORF) *yodA* de *Escherichia coli*,

está(n) adicionalmente sobreexpresados al mismo tiempo.

20 17. El proceso de acuerdo con las reivindicaciones 15 a 16, **caracterizado** porque se hace uso de microorganismos en los cuales los caminos metabólicos que reducen la formación de dicho L-aminoácido están al menos parcialmente atenuados.

25 18. El proceso de acuerdo con la reivindicación 17, **caracterizado** porque, para el propósito de preparación de L-treonina, se fermentan microorganismos de la familia Enterobacteriáceas en los cuales uno o más de los genes seleccionados del grupo:

- a) el gen *tdh* que codifica treonina-deshidrogenasa,
30 b) el gen *mdh* que codifica malato-deshidrogenasa,
c) el producto génico del marco de lectura abierto (ORF) *yjfA* de *Escherichia coli*,
d) el producto génico del marco de lectura abierto (ORF) *ytfP* de *Escherichia coli*,
35 e) el gen *pckA* que codifica fosfoenolpiruvato-carboxiquinasa,
f) el gen *poxB* que codifica piruvato-oxidasa,
40 g) el gen *dgsA* que codifica el regulador DgsA del sistema de fosfotransferasa,
h) el gen *fruR* que codifica el represor fructosa,
i) el gen *rpoS* que codifica el factor sigma³⁸, y
45 j) el gen *aspA* que codifica aspartato-amonio-liasa,

está/están atenuados adicionalmente al mismo tiempo, en particular eliminados, o su expresión está reducida.

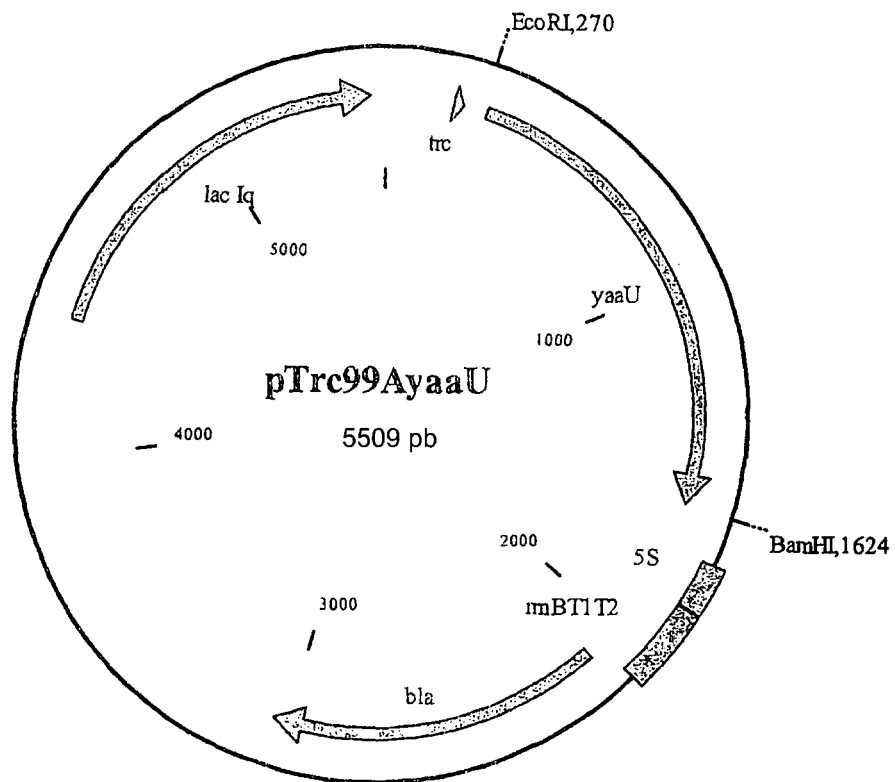
50

55

60

65

Fig. 1: Mapa del plásmido pTrc99AyaU



ES 2 343 010 T3

LISTA DE SECUENCIAS

- <110> Degussa AG
- 5 <120> Proceso para preparación de L-aminoácidos que utiliza cepas de la familia Enterobacteriáceas
- <130> 030321 BT
- 10 <160> 8
- <170> PatentIn versión 3.2
- 15 <210> 1
<211> 34
<212> DNA
<213> Secuencia artificial
- 20 <220>
<221> Iniciador
<222> (1)..(34)
<223> yaaU-ex1
- 25 <400> 1
- 30 gatctgaatt ctaaggaata accatgcaac cgtc 34
- <210> 2
<211> 32
<212> DNA
<213> Secuencia artificial
- 35 <220>
<221> Iniciador
<222> (1)..(32)
<223> yaaU-ex2
- 40 <400> 2
- gatctaggat cccaatttac cccattetct gc 32
- 50 <210> 3
<211> 1372
<212> DNA
<213> *Escherichia coli*
- <220>
<221> producto PCR
<222> (1)..(1372)
<223> producto PCR yaaU
- 60 <220>
<221> Iniciador_unión
<222> (1)..(34)
- 65

ES 2 343 010 T3

<223> Iniciador yaaU-ex1

<220>

5 <221> CDS

<222> (25)..(1356)

<223> región codificante yaaU

10 <220>

<221> Iniciador_unión

<222> (1341)..(1372)

15 <223> Iniciador yaaU-ex2

<400> 3

20	gatctgaatt cttaaggaat aacc atg caa ccg tcc aga aac ttt gac gat Met Gln Pro Ser Arg Asn Phe Asp Asp	51
	1 5	
25	ctc aaa ttc tcc tct att cac cgc cgc att ttg ctg tgg gga agc ggt Leu Lys Phe Ser Ser Ile His Arg Arg Ile Leu Leu Trp Gly Ser Gly 10 15 20 25	99
30	ggt ccg ttt ctg gat ggt tat gta ctg gta atg att ggc gtg gcg ctg Gly Pro Phe Leu Asp Gly Tyr Val Leu Val Met Ile Gly Val Ala Leu 30 35 40	147
35	gag caa ctg acg ccg gcg ctg aaa ctg gac gct gac tgg att ggc ttg Glu Gln Leu Thr Pro Ala Leu Lys Leu Asp Ala Asp Trp Ile Gly Leu 45 50 55	195
40	ctg ggc gcg gga acg ctc gcc ggg ctg ttc gtt ggc aca tcg ctg ttt Leu Gly Ala Gly Thr Leu Ala Gly Leu Phe Val Gly Thr Ser Leu Phe 60 65 70	243
45	ggt tat att tcc gat aaa gtc gga cgg cgc aaa atg ttc ctc att gat Gly Tyr Ile Ser Asp Lys Val Gly Arg Arg Lys Met Phe Leu Ile Asp 75 80 85	291
50	atc atc gcc atc ggc gtg ata tcg gtg gcg acg atg ttt gtt tca tcc Ile Ile Ala Ile Gly Val Ile Ser Val Ala Thr Met Phe Val Ser Ser 90 95 100 105	339
55	ccc gtc gaa ctg ttg gtg atg cgg gta ctt atc ggc att gtc atc ggt Pro Val Glu Leu Leu Val Met Arg Val Leu Ile Gly Ile Val Ile Gly 110 115 120	387
60	gca gat tat ccc atc gcc acc tca atg atc acc gag ttc tcc agt acc Ala Asp Tyr Pro Ile Ala Thr Ser Met Ile Thr Glu Phe Ser Ser Thr 125 130 135	435
65	cgt cag cgg gcg ttt tcc atc agc ttt att gcc gcg atg tgg tat gtc Arg Gln Arg Ala Phe Ser Ile Ser Phe Ile Ala Ala Met Trp Tyr Val 140 145 150	483
70	ggc gcg acc tgt gcc gat ctg gtc ggc tac tgg ctt tat gat gtg gaa Gly Ala Thr Cys Ala Asp Leu Val Gly Tyr Trp Leu Tyr Asp Val Glu 155 160 165	531
75	ggc ggc tgg cgc tgg atg ctg ggt agc gcg gcg atc ccc tgt ttg ttg Gly Gly Trp Arg Trp Met Leu Gly Ser Ala Ala Ile Pro Cys Leu Leu 170 175 180 185	579
80	att ttg att ggt cga ttc gaa ctg cct gaa tct ccc cgc tgg tta tta Ile Leu Ile Gly Arg Phe Glu Leu Pro Glu Ser Pro Arg Trp Leu Leu 190 195 200	627
85	cgc aaa ggg cga gta aaa gag tgc gaa gag atg atg atc aaa ctg ttt Arg Lys Gly Arg Val Lys Glu Cys Glu Glu Met Met Ile Lys Leu Phe 205 210 215	675
90	ggc gaa ccg gtg gct ttc gat gaa gag cag ccg cag caa acc cgt ttt Gly Glu Pro Val Ala Phe Asp Glu Glu Gln Pro Gln Gln Thr Arg Phe 220 225 230	723

ES 2 343 010 T3

```

cgc gat ctg ttt aat cgc cgc cat ttt cct ttt gtt ctg ttt gtt gcc      771
Arg Asp Leu Phe Asn Arg Arg His Phe Pro Phe Val Leu Phe Val Ala
 235                240                245

5   gcc atc tgg acc tgc cag gtg atc cca atg ttc gcc att tac acc ttt      819
Ala Ile Trp Thr Cys Gln Val Ile Pro Met Phe Ala Ile Tyr Thr Phe
 250                255                260                265

10  ggc ccg caa atc gtt ggt ttg ttg gga ttg ggg gtt ggc aaa aac gcg      867
Gly Pro Gln Ile Val Gly Leu Leu Gly Leu Gly Val Gly Lys Asn Ala
                270                275                280

15  gca cta ggg aat gtg gtg att agc ctg ttc ttt atg ctc ggc tgt att      915
Ala Leu Gly Asn Val Val Ile Ser Leu Phe Phe Met Leu Gly Cys Ile
                285                290                295

20  ccg ccg atg ctg tgg tta aac act gcc gga cgg cgt cca ttg ttg att      963
Pro Pro Met Leu Trp Leu Asn Thr Ala Gly Arg Arg Pro Leu Leu Ile
 300                305                310

25  ggc agc ttt gcc atg atg acg ctg gcg ctg gcg gtt ttg ggg cta atc      1011
Gly Ser Phe Ala Met Met Thr Leu Ala Leu Ala Val Leu Gly Leu Ile
 315                320                325

30  ccg gat atg ggg atc tgg ctg gta gtg atg gcc ttt gcg gtg tat gcc      1059
Pro Asp Met Gly Ile Trp Leu Val Val Met Ala Phe Ala Val Tyr Ala
 330                335                340                345

35  ttt ttc tct ggc ggg ccg ggt aat ttg cag tgg ctc tat cct aat gaa      1107
Phe Phe Ser Gly Gly Pro Gly Asn Leu Gln Trp Leu Tyr Pro Asn Glu
                350                355                360

40  ctc ttc ccg aca gat atc cgc gcc tct gcc gtg ggc gtg att atg tcc      1155
Leu Phe Pro Thr Asp Ile Arg Ala Ser Ala Val Gly Val Ile Met Ser
                365                370                375

45  tta agt cgt att ggc acc att gtt tcg acc tgg gca cta ccg atc ttt      1203
Leu Ser Arg Ile Gly Thr Ile Val Ser Thr Trp Ala Leu Pro Ile Phe
                380                385                390

50  atc aat aat tac ggt atc agt aac acg atg cta atg ggg gcg ggt atc      1251
Ile Asn Asn Tyr Gly Ile Ser Asn Thr Met Leu Met Gly Ala Gly Ile
                395                400                405

55  tcg ctg ttt ggc ttg ttg att tcc gta gcg ttt gcc ccg gag act cga      1299
Ser Leu Phe Gly Leu Leu Ile Ser Val Ala Phe Ala Pro Glu Thr Arg
 410                415                420                425

60  ggg atg tca ctg gcg cag acc agc aat atg acg atc cgc ggg cag aga      1347
Gly Met Ser Leu Ala Gln Thr Ser Asn Met Thr Ile Arg Gly Gln Arg
                430                435                440

65  atg ggg taa attgggatcc tagatc      1372
Met Gly

```

<220> 4

<221> 443

55 <222> PRT

<223> *Escherichia coli*

60

65

ES 2 343 010 T3

<400> 4

5 Met Gln Pro Ser Arg Asn Phe Asp Asp Leu Lys Phe Ser Ser Ile His
 1 5 10 15
 Arg Arg Ile Leu Leu Trp Gly Ser Gly Gly Pro Phe Leu Asp Gly
 20 25 30
 10 Val Leu Val Met Ile Gly Val Ala Leu Glu Gln Leu Thr Pro Ala
 35 40 45
 15 Lys Leu Asp Ala Asp Trp Ile Gly Leu Leu Gly Ala Gly Thr Leu
 50 55 60
 20 Gly Leu Phe Val Gly Thr Ser Leu Phe Gly Tyr Ile Ser Asp Lys
 65 70 75
 Gly Arg Arg Lys Met Phe Leu Ile Asp Ile Ile Ala Ile Gly Val
 85 90 95
 25 Ser Val Ala Thr Met Phe Val Ser Ser Pro Val Glu Leu Leu Val
 100 105 110
 30 Arg Val Leu Ile Gly Ile Val Ile Gly Ala Asp Tyr Pro Ile Ala
 115 120 125
 35 Ser Met Ile Thr Glu Phe Ser Ser Thr Arg Gln Arg Ala Phe Ser
 130 135 140
 Ser Phe Ile Ala Ala Met Trp Tyr Val Gly Ala Thr Cys Ala Asp
 145 150 155
 40 Val Gly Tyr Trp Leu Tyr Asp Val Glu Gly Gly Trp Arg Trp Met
 165 170 175
 45 Gly Ser Ala Ala Ile Pro Cys Leu Leu Ile Leu Ile Gly Arg Phe
 180 185 190
 50 Leu Pro Glu Ser Pro Arg Trp Leu Leu Arg Lys Gly Arg Val Lys
 195 200 205
 Cys Glu Glu Met Met Ile Lys Leu Phe Gly Glu Pro Val Ala Phe
 210 215 220
 55 Glu Glu Gln Pro Gln Gln Thr Arg Phe Arg Asp Leu Phe Asn Arg
 225 230 235
 60 His Phe Pro Phe Val Leu Phe Val Ala Ala Ile Trp Thr Cys Gln
 245 250 255
 65 Ile Pro Met Phe Ala Ile Tyr Thr Phe Gly Pro Gln Ile Val Gly
 260 265 270

ES 2 343 010 T3

agc ctg ttc ttt atg ctc ggc tgt att ccg ccg atg ctg tgg cta aac 912
 Ser Leu Phe Phe Met Leu Gly Cys Ile Pro Pro Met Leu Trp Leu Asn
 290 295 300

5
 act gcc gga cgg cgt cca ttg ttg att ggc agc ttt gcc atg atg acg 960
 Thr Ala Gly Arg Arg Pro Leu Leu Ile Gly Ser Phe Ala Met Met Thr
 305 310 315 320

10
 ctg gcg ttg gcg gtt ttg ggg cta atc ccg gat atg ggg atc tgg ctg 1008
 Leu Ala Leu Ala Val Leu Gly Leu Ile Pro Asp Met Gly Ile Trp Leu
 325 330 335

15
 gta gtg atg gcc ttt gcg gtg tat gcc ttt ttc tct gcc ggg ccg ggt 1056
 Val Val Met Ala Phe Ala Val Tyr Ala Phe Phe Ser Gly Gly Pro Gly
 340 345 350

20
 aat ttg cag tgg ctc tat cct aat gaa ctc ttc ccg acg gat atc cgc 1104
 Asn Leu Gln Trp Leu Tyr Pro Asn Glu Leu Phe Pro Thr Asp Ile Arg
 355 360 365

25
 gcc tct gcc gtg ggc gtg att atg tcc tta agt cgt att gcc acc att 1152
 Ala Ser Ala Val Gly Val Ile Met Ser Leu Ser Arg Ile Gly Thr Ile
 370 375 380

30
 gtt tcg acc tgg gca ctg ccg ata ttt atc aat aat tac ggt atc agt 1200
 Val Ser Thr Trp Ala Leu Pro Ile Phe Ile Asn Asn Tyr Gly Ile Ser
 385 390 395 400

35
 aac acg atg cta atg ggg gcg ggt atc tcg ctg ttt gcc ttg ttg att 1248
 Asn Thr Met Leu Met Gly Ala Gly Ile Ser Leu Phe Gly Leu Leu Ile
 405 410 415

40
 tcc gta gcg ttt gcc ccg gag act cga gga atg tca ctg gcg cag acc 1296
 Ser Val Ala Phe Ala Pro Glu Thr Arg Gly Met Ser Leu Ala Gln Thr
 420 425 430

45
 agc aat atg acg atc cgc ggg cag aga atg gga tta ttt gtt cag att 1344
 Ser Asn Met Thr Ile Arg Gly Gln Arg Met Gly Leu Phe Val Gln Ile
 435 440 445

50
 tct ctc ttt tct gaa tca ata tta ttg act ata agc cgc gtg aat ata 1392
 Ser Leu Phe Ser Glu Ser Ile Leu Leu Thr Ile Ser Arg Val Asn Ile
 450 455 460

55
 tga 1395

<210> 6
 <211> 464
 <212> PRT
 <213> *Shigella flexneri*
 <400> 6

Met Gln Pro Ser Arg Asn Phe Asp Asp Leu Lys Phe Ser Ser Ile His
 1 5 10 15

Arg Arg Ile Leu Leu Trp Gly Ser Gly Gly Pro Phe Leu Asp Gly Tyr
 20 25 30

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

ES 2 343 010 T3

Lys Leu Asp Ala Asp Trp Ile Gly Leu Leu Gly Ala Gly Thr Leu Ala
 50 55 60

5

Gly Leu Phe Val Gly Thr Ser Leu Phe Gly Tyr Ile Ser Asp Lys Val
 65 70 75 80

10

Gly Arg Arg Lys Met Phe Leu Ile Asp Ile Ile Ala Ile Gly Val Ile
 85 90 95

15

Ser Val Ala Thr Met Cys Val Ser Ser Pro Val Glu Leu Leu Val Met
 100 105 110

20

Arg Val Leu Ile Gly Ile Val Ile Gly Ala Asp Tyr Pro Ile Ala Thr
 115 120 125

25

Ser Met Ile Thr Glu Phe Ser Ser Thr Arg Gln Arg Ala Phe Ser Ile
 130 135 140

30

Val Gly Tyr Trp Leu Tyr Asp Val Glu Gly Gly Trp Arg Trp Met Leu
 165 170 175

35

Gly Ser Ala Ala Ile Pro Cys Leu Leu Ile Leu Ile Gly Arg Phe Glu
 180 185 190

40

Leu Pro Glu Ser Pro Arg Trp Leu Leu Arg Lys Gly Arg Val Lys Glu
 195 200 205

45

Cys Glu Glu Met Met Ile Lys Leu Phe Gly Glu Pro Val Ala Phe Asp
 210 215 220

50

Glu Glu Gln Pro Gln Gln Thr Arg Phe Arg Asp Leu Phe Asn Arg Arg
 225 230 235 240

55

His Phe Pro Phe Val Leu Phe Val Ala Ala Ile Trp Thr Cys Gln Val
 245 250 255

60

Ile Pro Met Phe Ala Ile Tyr Thr Phe Gly Pro Gln Ile Val Gly Ser
 260 265 270

65

Leu Gly Leu Gly Val Gly Lys Asn Ala Ala Leu Gly Asn Val Val Ile
 275 280 285

Ser Leu Phe Phe Met Leu Gly Cys Ile Pro Pro Met Leu Trp Leu Asn
 290 295 300

ES 2 343 010 T3

	Thr	Ala	Gly	Arg	Arg	Pro	Leu	Leu	Ile	Gly	Ser	Phe	Ala	Met	Met	Thr	
	305					310					315					320	
5	Leu	Ala	Leu	Ala	Val	Leu	Gly	Leu	Ile	Pro	Asp	Met	Gly	Ile	Trp	Leu	
					325					330					335		
10	Val	Val	Met	Ala	Phe	Ala	Val	Tyr	Ala	Phe	Phe	Ser	Gly	Gly	Pro	Gly	
				340					345					350			
15	Asn	Leu	Gln	Trp	Leu	Tyr	Pro	Asn	Glu	Leu	Phe	Pro	Thr	Asp	Ile	Arg	
			355					360					365				
20	Ala	Ser	Ala	Val	Gly	Val	Ile	Met	Ser	Leu	Ser	Arg	Ile	Gly	Thr	Ile	
		370					375					380					
25	Val	Ser	Thr	Trp	Ala	Leu	Pro	Ile	Phe	Ile	Asn	Asn	Tyr	Gly	Ile	Ser	
	385					390					395					400	
30	Asn	Thr	Met	Leu	Met	Gly	Ala	Gly	Ile	Ser	Leu	Phe	Gly	Leu	Leu	Ile	
					405					410					415		
35	Ser	Val	Ala	Phe	Ala	Pro	Glu	Thr	Arg	Gly	Met	Ser	Leu	Ala	Gln	Thr	
			420						425					430			
40	Ser	Asn	Met	Thr	Ile	Arg	Gly	Gln	Arg	Met	Gly	Leu	Phe	Val	Gln	Ile	
			435				440						445				
45	Ser	Leu	Phe	Ser	Glu	Ser	Ile	Leu	Leu	Thr	Ile	Ser	Arg	Val	Asn	Ile	
		450					455					460					
50	<210> 7																
55	<211> 1341																
	<212> DNA																
	<213> <i>Salmonella typhimurium</i>																
60	<220>																
	<221> CDS																
	<222> (1)..(1341)																
	<223> región codificante yaaU																
65	<400> 7																
70	atg	cag	cag	ccc	agg	aac	ttt	gat	gac	ctt	aaa	ttc	tcc	tcc	att	cat	48
75	Met	Gln	Gln	Pro	Arg	Asn	Phe	Asp	Asp	Leu	Lys	Phe	Ser	Ser	Ile	His	
	1			5						10					15		
80	cgc	aga	att	atg	ctg	tgg	gga	agc	ggc	ggc	cgc	ttc	ctt	gat	ggc	tac	96
85	Arg	Arg	Ile	Met	Leu	Trp	Gly	Ser	Gly	Gly	Pro	Phe	Leu	Asp	Gly	Tyr	
				20					25					30			
90	gtg	ctg	gtg	atc	att	ggc	gtg	gcg	ctg	gaa	caa	ctg	acg	cgc	cta	ttg	144
95	Val	Leu	Val	Ile	Ile	Gly	Val	Ala	Leu	Glu	Gln	Leu	Thr	Pro	Leu	Leu	
			35				40						45				
100	cac	ctt	gat	gct	gaa	tgg	att	ggc	gcg	ctc	ggc	gcg	gct	acg	ctt	gcc	192

ES 2 343 010 T3

	His	Leu	Asp	Ala	Glu	Trp	Ile	Gly	Ala	Leu	Gly	Ala	Ala	Thr	Leu	Ala	
	50						55				60						
5	gga	ctg	ttt	atc	ggc	acc	tcg	ttg	ttt	ggc	tat	att	tgc	gat	aaa	gtc	240
	Gly	Leu	Phe	Ile	Gly	Thr	Ser	Leu	Phe	Gly	Tyr	Ile	Cys	Asp	Lys	Val	
	65				70					75					80		
10	ggg	cgg	cgt	aaa	atg	ttc	ctg	ctt	gat	att	atc	gct	atc	ggc	gtt	atc	288
	Gly	Arg	Arg	Lys	Met	Phe	Leu	Leu	Asp	Ile	Ile	Ala	Ile	Gly	Val	Ile	
				85					90						95		
15	tcc	gtg	gcg	acc	atg	ttt	gtt	tcc	acg	cgc	ctt	gaa	tta	ctg	gtg	atg	336
	Ser	Val	Ala	Thr	Met	Phe	Val	Ser	Thr	Pro	Leu	Glu	Leu	Leu	Val	Met	
				100					105						110		
20	cgc	ggt	tta	att	ggc	att	ggt	atc	ggc	gcg	gat	tac	cct	atc	gct	act	384
	Arg	Val	Leu	Ile	Gly	Ile	Val	Ile	Gly	Ala	Asp	Tyr	Pro	Ile	Ala	Thr	
			115					120					125				
25	tcc	atg	att	acg	gag	ttt	tcc	aac	acc	cgc	cag	cgc	gcg	ttc	tcc	atc	432
	Ser	Met	Ile	Thr	Glu	Phe	Ser	Asn	Thr	Arg	Gln	Arg	Ala	Phe	Ser	Ile	
		130						135					140				
30	ggt	ttt	atc	gcc	gcg	atg	tgg	tac	gtc	ggc	gcg	acc	tgc	gcc	aac	ctg	480
	Gly	Phe	Ile	Ala	Ala	Met	Trp	Tyr	Val	Gly	Ala	Thr	Cys	Ala	Asn	Leu	
	145					150					155					160	
35	gtg	ggc	tac	tgg	ctg	tat	gac	atg	gaa	ggc	ggc	tgg	cgt	tgg	atg	ctt	528
	Val	Gly	Tyr	Trp	Leu	Tyr	Asp	Met	Glu	Gly	Gly	Trp	Arg	Trp	Met	Leu	
				165						170					175		
40	ggc	agc	gcc	ttt	atc	ccg	tgc	tta	atc	att	ttg	att	ggc	cgc	ttt	gac	576
	Gly	Ser	Ala	Phe	Ile	Pro	Cys	Leu	Ile	Ile	Leu	Ile	Gly	Arg	Phe	Asp	
			180						185					190			
45	ctg	ccg	gaa	tcg	cca	cgc	tgg	cta	ttg	cga	aaa	gga	cgg	gta	aaa	gag	624
	Leu	Pro	Glu	Ser	Pro	Arg	Trp	Leu	Leu	Arg	Lys	Gly	Arg	Val	Lys	Glu	
			195					200					205				
50	tgc	gaa	caa	atg	atg	atc	aaa	ctg	ttc	ggc	gaa	ccg	gtc	tgc	ttt	gac	672
	Cys	Glu	Gln	Met	Met	Ile	Lys	Leu	Phe	Gly	Glu	Pro	Val	Cys	Phe	Asp	
		210				215						220					
55	gat	gag	ctg	ccg	cag	gag	acc	cgc	ttc	ctg	cag	ctg	ttt	aac	cgc	cgt	720
	Asp	Glu	Leu	Pro	Gln	Glu	Thr	Arg	Phe	Leu	Gln	Leu	Phe	Asn	Arg	Arg	
	225					230					235				240		
60	cat	ttc	ccg	ttt	gta	ctg	ttc	ggt	gcg	gct	atc	tgg	acc	tgt	cag	gtg	768
	His	Phe	Pro	Phe	Val	Leu	Phe	Val	Ala	Ala	Ile	Trp	Thr	Cys	Gln	Val	
				245						250					255		
65	atc	ccg	atg	ttc	gct	atc	tat	acc	ttt	ggc	ccg	cag	atc	gtc	ggg	tta	816
	Ile	Pro	Met	Phe	Ala	Ile	Tyr	Thr	Phe	Gly	Pro	Gln	Ile	Val	Gly	Leu	
			260						265					270			
70	ctc	ggc	tgg	gag	cag	ggg	cgc	aac	gcg	gcg	ctt	ggc	aat	gtg	gtg	att	864
	Leu	Gly	Trp	Glu	Gln	Gly	Arg	Asn	Ala	Ala	Leu	Gly	Asn	Val	Val	Ile	
			275					280					285				
75	agc	ctc	ttt	ttt	atg	ctg	ggc	tgt	ata	ccg	gcg	atg	ttc	tgg	cta	aac	912
	Ser	Leu	Phe	Phe	Met	Leu	Gly	Cys	Ile	Pro	Ala	Met	Phe	Trp	Leu	Asn	
		290					295					300					
80	agt	att	ggc	cgt	cgc	cct	ctg	ctt	att	ggc	agc	ttc	gcc	atg	atg	act	960
	Ser	Ile	Gly	Arg	Arg	Pro	Leu	Leu	Ile	Gly	Ser	Phe	Ala	Met	Met	Thr	

ES 2 343 010 T3

	305	310	315	320	
5	atc gcg ctg gcg ctg ctg ggg ctg gtg tca aat ctg ggc atc ata ctg Ile Ala Leu Ala Leu Leu Gly Leu Val Ser Asn Leu Gly Ile Ile Leu				1008
	325		330	335	
10	gtc gtg gtg gcg ttt gcg gta tac gcc ttc ttt tcc ggc gga ccg ggc Val Val Val Ala Phe Ala Val Tyr Ala Phe Phe Ser Gly Gly Pro Gly				1056
	340		345	350	
15	ata ctg caa tgg ctt tat cct aat gaa ctc ttc ccg acg gat att cgc Ile Leu Gln Trp Leu Tyr Pro Asn Glu Leu Phe Pro Thr Asp Ile Arg				1104
	355		360	365	
20	gct tcc gcc gtg ggg gtc atc atg tca tta agc cgc atc gga acc att Ala Ser Ala Val Gly Val Ile Met Ser Leu Ser Arg Ile Gly Thr Ile				1152
	370		375	380	
25	gtc tcc acc tgg gca tta ccg att ttt att act cgt tac ggt att aat Val Ser Thr Trp Ala Leu Pro Ile Phe Ile Thr Arg Tyr Gly Ile Asn				1200
	385	390	395	400	
30	aac gtc atg ctg att ggc gcg ttg att tcc tta gtc ggc ctt ggc gta Asn Val Met Leu Ile Gly Ala Leu Ile Ser Leu Val Gly Leu Gly Val				1248
	405		410	415	
35	tcc gtg atg ttt gcc ccg gaa acc cgc ggc ttg acg cta aca cag acc Ser Val Met Phe Ala Pro Glu Thr Arg Gly Leu Thr Leu Thr Gln Thr				1296
	420		425	430	
40	gga aac atg act ttg cgc gga acg cca tcc gat aac cct cgt taa Gly Asn Met Thr Leu Arg Gly Thr Pro Ser Asp Asn Pro Arg				1341
	435		440	445	
45	<210> 8				
	<211> 446				
	<212> PRT				
	<213> <i>Salmonella typhimurium</i>				
50	<400> 8				
	Met Gln Gln Pro Arg Asn Phe Asp Asp Leu Lys Phe Ser Ser Ile His				
	1	5	10	15	
55	Arg Arg Ile Met Leu Trp Gly Ser Gly Gly Pro Phe Leu Asp Gly Tyr				
	20	25	30		
60	Val Leu Val Ile Ile Gly Val Ala Leu Glu Gln Leu Thr Pro Leu Leu				
	35	40	45		
65	His Leu Asp Ala Glu Trp Ile Gly Ala Leu Gly Ala Ala Thr Leu Ala				
	50	55	60		
70	Gly Leu Phe Ile Gly Thr Ser Leu Phe Gly Tyr Ile Cys Asp Lys Val				
	65	70	75	80	
75	Gly Arg Arg Lys Met Phe Leu Leu Asp Ile Ile Ala Ile Gly Val Ile				
	85	90	95		

ES 2 343 010 T3

Ser Val Ala Thr Met Phe Val Ser Thr Pro Leu Glu Leu Leu Val Met
 100 105 110

5

Arg Val Leu Ile Gly Ile Val Ile Gly Ala Asp Tyr Pro Ile Ala Thr
 115 120 125

10

Ser Met Ile Thr Glu Phe Ser Asn Thr Arg Gln Arg Ala Phe Ser Ile
 130 135 140

15

Gly Phe Ile Ala Ala Met Trp Tyr Val Gly Ala Thr Cys Ala Asn Leu
 145 150 155 160

20

Val Gly Tyr Trp Leu Tyr Asp Met Glu Gly Gly Trp Arg Trp Met Leu
 165 170 175

25

Gly Ser Ala Phe Ile Pro Cys Leu Ile Ile Leu Ile Gly Arg Phe Asp
 180 185 190

30

Leu Pro Glu Ser Pro Arg Trp Leu Leu Arg Lys Gly Arg Val Lys Glu
 195 200 205

35

Cys Glu Gln Met Met Ile Lys Leu Phe Gly Glu Pro Val Cys Phe Asp
 210 215 220

40

Asp Glu Leu Pro Gln Glu Thr Arg Phe Leu Gln Leu Phe Asn Arg Arg
 225 230 235 240

45

His Phe Pro Phe Val Leu Phe Val Ala Ala Ile Trp Thr Cys Gln Val
 245 250 255

50

Ile Pro Met Phe Ala Ile Tyr Thr Phe Gly Pro Gln Ile Val Gly Leu
 260 265 270

55

Leu Gly Trp Glu Gln Gly Arg Asn Ala Ala Leu Gly Asn Val Val Ile
 275 280 285

60

Ser Leu Phe Phe Met Leu Gly Cys Ile Pro Ala Met Phe Trp Leu Asn
 290 295 300

65

Ser Ile Gly Arg Arg Pro Leu Leu Ile Gly Ser Phe Ala Met Met Thr
 305 310 315 320

70

Ile Ala Leu Ala Leu Leu Gly Leu Val Ser Asn Leu Gly Ile Ile Leu
 325 330 335

75

Val Val Val Ala Phe Ala Val Tyr Ala Phe Phe Ser Gly Gly Pro Gly
 340 345 350

ES 2 343 010 T3

Ile Leu Gln Trp Leu Tyr Pro Asn Glu Leu Phe Pro Thr Asp Ile Arg
355 360 365

5 Ala Ser Ala Val Gly Val Ile Met Ser Leu Ser Arg Ile Gly Thr Ile
370 375 380

10 Val Ser Thr Trp Ala Leu Pro Ile Phe Ile Thr Arg Tyr Gly Ile Asn
385 390 395 400

15 Asn Val Met Leu Ile Gly Ala Leu Ile Ser Leu Val Gly Leu Gly Val
405 410 415

20 Ser Val Met Phe Ala Pro Glu Thr Arg Gly Leu Thr Leu Thr Gln Thr
420 425 430

25

30

35

40

45

50

55

60

65