

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 403 199**

51 Int. Cl.:

**C12N 9/10** (2006.01)

**C12N 9/90** (2006.01)

**C12P 13/08** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.07.2008 E 08791407 (3)**

97 Fecha y número de publicación de la concesión europea: **23.01.2013 EP 2180052**

54 Título: **Método para producir L-lisina**

30 Prioridad:

**23.07.2007 JP 2007190795**

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

**16.05.2013**

73 Titular/es:

**AJINOMOTO CO., INC. (100.0%)  
15-1, KYOBASHI 1-CHOME CHUO-KU  
TOKYO 104-8315, JP**

72 Inventor/es:

**DOI, HIDETAKA y  
UEDA, TAKUJI**

74 Agente/Representante:

**FÚSTER OLAGUIBEL, Gustavo Nicolás**

ES 2 403 199 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

Método para producir L-lisina.

5 **Campo técnico**

La presente invención se refiere a un método para producir L-lisina utilizando *Escherichia coli*. L-Lisina es un aminoácido esencial y se utiliza como componente de fármacos y diversas mezclas nutricionales tales como aditivos de piensos para animales.

10

**Técnica anterior**

Los L-aminoácidos tales como la L-lisina se producen industrialmente mediante fermentación usando bacterias productoras de aminoácidos tales como bacterias corineformes y bacterias *Escherichia* que tienen capacidad para producir tales L-aminoácidos. Como tales bacterias productoras de aminoácidos se usan cepas aisladas de la naturaleza, cepas mutantes artificiales de tales cepas, cepas recombinantes en las que está potenciada la actividad de las enzimas de biosíntesis de L-aminoácidos mediante recombinación génica, etc., con el fin de obtener productividad mejorada. Los ejemplos de métodos para producir L-lisina incluyen los métodos descritos en los documentos de patente 1 a 4.

15

20

Como métodos para mejorar la capacidad para producir aminoácidos tales como L-lisina, además del método de aumentar la expresión de una enzima de una ruta biosintética característica para un aminoácido diana, se ha desarrollado un método de modificación de la ruta de la cadena respiratoria para mejorar la eficacia energética (documento de patente 5) y un método de amplificación de un gen de nicotinamida nucleótido transhidrogenasa para aumentar la capacidad productora de nicotinamida adenina dinucleótido fosfato (documento de patente 6).

25

Además, como medio basado en la modificación de una ruta común a la biosíntesis de diversos aminoácidos, se conocen bacterias productoras de L-aminoácidos en las que está modificada la ruta anaplerótica, tal como la bacteria corineforme productora de L-lisina en la que está aumentada la actividad piruvato (documento de patente 7), la bacteria *Escherichia* productora de L-lisina deficiente en piruvato cinasa (documento de patente 8) y la bacteria corineforme productora de L-lisina que es deficiente en maleato quinona oxidoreductasa (documento de patente 9).

30

Como precursor de L-lisina, está el ácido meso- $\alpha,\epsilon$ -diaminopimélico (también denominado en lo sucesivo "meso-DAP"). Meso-DAP es un precursor de L-lisina, y al mismo tiempo, es una sustancia indispensable para el crecimiento de bacterias como componente constituyente de las paredes celulares. En cuanto a la síntesis de meso-DAP en bacterias *Escherichia*, se sabe que meso-DAP se sintetiza a partir de un precursor del mismo, el ácido 2,3,4,5-tetrahidropiridina-2,6-dicarboxílico (también denominado en lo sucesivo "THDP"), mediante las funciones de cuatro enzimas, 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa (también denominada en lo sucesivo "DapD", documento no de patente 1), succinildiaminopimelato transaminasa (también denominada en lo sucesivo "DapC", documento no de patente 2), succinildiaminopimelato desuccinilasa (también denominada en lo sucesivo "DapE", documento no de patente 3) y diaminopimelato epimerasa (también denominada en lo sucesivo "DapF", documento no de patente 4), que pertenecen a la ruta de síntesis de meso-DAP. Se ha esclarecido que las bacterias corineformes tienen otra ruta de síntesis de meso-DAP que utiliza THDP como precursor y sintetizan meso-DAP a partir de THDP mediante una reacción de una etapa usando meso-DAP deshidrogenasa (también denominada en lo sucesivo "diaminopimelato deshidrogenasa" o "DDH"), y se sabe que la expresión de DDH es útil para la producción de meso-DAP (documento de patente 10). Además, se dio a conocer que, en un procedimiento de investigación de la etapa limitante de la velocidad de la biosíntesis de L-lisina en bacterias *Escherichia*, se introdujo un gen que codifica para DDH de bacteria corineforme en una bacteria *Escherichia* productora de L-lisina, en lugar de potenciar una actividad de una enzima que pertenece a la ruta de síntesis de meso-DAP, para obtener la producción de L-lisina (documento de patente 13). Sin embargo, no se ha esperado que, cuando se expresa DDH en una bacteria *Escherichia*, la disminución de una actividad de una enzima que pertenece a la ruta de síntesis de meso-DAP será eficaz para la producción de L-lisina.

35

40

45

50

55 Documento de patente 1: patente japonesa abierta a consulta por el público (KOKAI) n.º 10-165180

Documento de patente 2: patente japonesa abierta a consulta por el público n.º 11-192088

Documento de patente 3: patente japonesa abierta a consulta por el público n.º 2000-253879

60

Documento de patente 4: patente japonesa abierta a consulta por el público n.º 2001-057896

Documento de patente 5: patente japonesa abierta a consulta por el público n.º 2002-17363

65 Documento de patente 6: patente japonesa n.º 2817400

Documento de patente 7: patente japonesa abierta a consulta por el público basada en la solicitud PCT en lengua extranjera (KOHYO) n.º 2002-508921

Documento de patente 8: publicación de patente internacional WO03/008600

Documento de patente 9: solicitud publicada de patente estadounidense n.º 2003/0044943

Documento de patente 10: patente japonesa abierta a consulta por el público n.º 61-289887

Documento de patente 11: publicación de patente internacional WO2006/093322

Documento de patente 12: solicitud publicada de patente estadounidense n.º 2006/0160191

Documento de patente 13: patente estadounidense n.º 6.040.160

Documento no de patente 1: Richaud, C. *et al.*, J. Biol. Chem., 259 (23):14824-14828, 1984

Documento no de patente 2: Heimberg, H. *et al.*, Gene, 90 (1): 69-78, 1990

Documento no de patente 3: Bouvier, J. *et al.*, J. Bacteriol., 174 (16):5265-71, 1992

Documento no de patente 4: Wiseman, J.S. *et al.*, J. Biol. Chem., 259 (14):8907-14, 1984

## Descripción de la invención

### Objeto que ha de lograrse mediante la invención

Un objeto de la presente invención es proporcionar una *Escherichia coli* que tiene capacidad productora de L-lisina mejorada y un método para producir L-lisina utilizando tal *Escherichia coli*.

### Medios para lograr el objeto

Los inventores de la presente invención realizaron diversas investigaciones con el fin de lograr el objeto mencionado anteriormente, y como resultado, encontraron que la capacidad productora de L-lisina de *Escherichia coli* podría mejorarse modificando *Escherichia coli* para disminuir la actividad de una enzima que pertenece a la ruta sintética de meso-DAP e introduciendo un gen que codifica para diaminopimelato deshidrogenasa. La presente invención se llevó a cabo partiendo de la base de este hallazgo.

Por tanto, la presente invención es tal como sigue.

(1) Una *Escherichia coli* que tiene una capacidad productora de L-lisina, que se ha modificado para disminuir la actividad o actividades de una o más clases de enzimas de la ruta de síntesis del ácido meso- $\alpha,\epsilon$ -diaminopimélico y en la que se ha introducido un gen que codifica para diaminopimelato deshidrogenasa.

(2) La *Escherichia coli* mencionada anteriormente, en la que las enzimas de la ruta de síntesis del ácido meso- $\alpha,\epsilon$ -diaminopimélico se seleccionan de 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa, succinildiaminopimelato transaminasa, succinildiaminopimelato desuccinilasa y diaminopimelato epimerasa.

(3) La *Escherichia coli* mencionada anteriormente, en la que la 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa, succinildiaminopimelato transaminasa, succinildiaminopimelato desuccinilasa y diaminopimelato epimerasa se codifican por el gen *dapD*, gen *dapC*, gen *dapE* y gen *dapF*, respectivamente.

(4) La *Escherichia coli* mencionada anteriormente, en la que las actividades de las enzimas de la ruta de síntesis del ácido meso- $\alpha,\epsilon$ -diaminopimélico se disminuyen mediante la disminución de la expresión de los genes o mediante la alteración de los genes.

(5) La *Escherichia coli* mencionada anteriormente, que se ha modificado para disminuir al menos la actividad de 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa.

(6) La *Escherichia coli* mencionada anteriormente, en la que el gen que codifica para diaminopimelato deshidrogenasa es el gen *ddh* de una bacteria corineforme.

(7) La *Escherichia coli* mencionada anteriormente, en la que la 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa es una proteína definida en (A) o (B) siguientes:

(A) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 2,

(B) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 2 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa.

5 (8) La *Escherichia coli* mencionada anteriormente, en la que la succinildiaminopimelato transaminasa es una proteína definida en (C) o (D) siguientes:

(C) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 4,

10 (D) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 4 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de succinildiaminopimelato transaminasa.

15 (9) La *Escherichia coli* mencionada anteriormente, en la que la succinildiaminopimelato desuccinilasa es una proteína definida en (E) o (F) siguientes:

(E) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 6,

20 (F) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 6 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de succinildiaminopimelato desuccinilasa.

25 (10) La *Escherichia coli* mencionada anteriormente, en la que la diaminopimelato epimerasa es una proteína definida en (G) o (H) siguientes:

(G) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 8,

30 (H) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 8 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de diaminopimelato epimerasa.

(11) La *Escherichia coli* mencionada anteriormente, en la que el gen *dapD* es un ADN definido en (a) o (b) siguientes:

35 (a) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 1, o

(b) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 1 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa.

(12) La *Escherichia coli* mencionada anteriormente, en la que el gen *dapC* es un ADN definido en (c) o (d) siguientes:

45 (c) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 3, o

(d) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 3 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de succinildiaminopimelato transaminasa.

50 (13) La *Escherichia coli* mencionada anteriormente, en la que el gen *dapE* es un ADN definido en (e) o (f) siguientes:

(e) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 5, o

55 (f) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 5 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de succinildiaminopimelato desuccinilasa.

(14) La *Escherichia coli* mencionada anteriormente, en la que el gen *dapF* es un ADN definido en (g) o (h) siguientes:

60 (g) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 7, o

65 (h) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 7 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de diaminopimelato epimerasa.

(15) La *Escherichia coli* mencionada anteriormente, en la que la diaminopimelato deshidrogenasa es una proteína definida en (I) o (J) siguientes:

5 (I) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 10, 12, 14, 16 ó 18,

(J) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 10, 12, 14, 16 ó 18 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de diaminopimelato deshidrogenasa.

10 (16) La *Escherichia coli* mencionada anteriormente, en la que el gen *ddh* es un ADN definido en (i) o (j) siguientes:

(i) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 9, 11, 13, 15 ó 17, o

15 (j) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 9, 11, 13, 15 ó 17 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de diaminopimelato deshidrogenasa.

20 (17) La *Escherichia coli* mencionada anteriormente, que tiene además una dihidrodipicolinato sintasa que está desensibilizada para la retroinhibición por L-lisina y una aspartocinasa que está desensibilizada para la retroinhibición por L-lisina, y tiene actividad potenciada de dihidrodipicolinato reductasa.

(18) Un método para producir L-lisina, que comprende hacer crecer en cultivo la *Escherichia coli* mencionada anteriormente en un medio y recoger L-lisina del medio.

25 **Mejor modo de llevar a cabo la invención**

A continuación en el presente documento, se explicará la presente invención en detalle.

30 <1> *Escherichia coli* de la presente invención

La *Escherichia coli* de la presente invención (también denominada en lo sucesivo la “bacteria de la presente invención”) es una *Escherichia coli* que tiene una capacidad productora de L-lisina, que se ha modificado para disminuir la actividad o actividades de una o más enzimas de la ruta de síntesis del ácido meso- $\alpha,\epsilon$ -diaminopimélico, y en la que se ha introducido un gen que codifica para diaminopimelato deshidrogenasa.

35 Además de las características mencionadas anteriormente, la bacteria de la presente invención tiene preferiblemente una dihidrodipicolinato sintasa que está desensibilizada para la retroinhibición por L-lisina y una aspartocinasa que está desensibilizada para la retroinhibición por L-lisina, y tiene una actividad potenciada de dihidrodipicolinato reductasa.

40 La bacteria de la presente invención puede obtenerse usando *Escherichia coli* que tiene capacidad productora de L-lisina como cepa original, modificándola para disminuir una actividad de una enzima que pertenece a la ruta de síntesis de meso-DAP, e introduciendo adicionalmente un gen que codifica para diaminopimelato deshidrogenasa. El orden de la modificación para disminuir una actividad de una enzima que pertenece a la ruta de síntesis de meso-DAP y la introducción de un gen que codifica para diaminopimelato deshidrogenasa no está limitado particularmente. Además, la capacidad productora de L-lisina puede transmitirse entre o tras la modificación y la introducción del gen.

50 La bacteria como cepa original de *Escherichia coli* usada para obtener la bacteria de la presente invención no está limitada particularmente. Sin embargo, los ejemplos específicos incluyen, por ejemplo, los descritos en el trabajo de Neidhardt *et al.* (Neidhardt F.C. *et al.*, *Escherichia coli* and *Salmonella typhimurium*, American Society for Microbiology, Washington D.C., 1029, tabla 1). Los ejemplos específicos incluyen *Escherichia coli* W3110 (ATCC 27325), *Escherichia coli* MG1655 (ATCC 47076) etc., derivadas de la cepa de tipo natural prototipo, cepa K12.

55 Estas cepas están disponibles de, por ejemplo, American Type Culture Collection (Colección Americana de Cultivos Tipo) (Dirección: 12301 Parklawn Drive, Rockville, Maryland 20852, Estados Unidos de América). Es decir, se facilitan números de registro para cada una de las cepas y las cepas pueden ordenarse usando esos números de registro. Los números de registro de las cepas se enumeran en el catálogo de la American Type Culture Collection.

60 <1-1> Transmisión de la capacidad productora de L-lisina y *Escherichia coli* que tiene capacidad productora de L-lisina

Los ejemplos de bacterias *Escherichia coli* productoras de L-lisina que tienen resistencia a un análogo de L-lisina. Los análogos de L-Lisina inhiben el crecimiento de *Escherichia coli*, pero esta inhibición está completa o parcialmente desensibilizada cuando la L-lisina está presente en el medio. Los ejemplos de análogos de L-lisina incluyen, pero sin limitarse a, oxalislina, hidroxamato de lisina, S-(2-aminoetil)-L-cisteína (AEC),  $\gamma$ -metil-lisina,  $\alpha$ -clorocaprolactama, etc. Pueden obtenerse mutantes que tienen resistencia a estos análogos de lisina sometiendo la

*Escherichia coli* a una mutagénesis artificial convencional. Los ejemplos específicos de cepas bacterianas útiles para producir L-lisina incluyen *Escherichia coli* AJ11442 (FERM BP- 1543, NRRL B-12185; véase la patente estadounidense n.º 4.346.170) y *Escherichia coli* VL611. En estos microorganismos, la aspartocinasa está desensibilizada para la retroinhibición por L-lisina.

La cepa WC196 puede usarse como una bacteria productora de L-lisina de *Escherichia coli*. Esta cepa bacteriana se obtuvo confiriendo resistencia a AEC a la cepa W3110, que se deriva de *Escherichia coli* K-12. La cepa resultante de designó cepa *Escherichia coli* AJ13069 y se depositó en el National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (actualmente National Institute of Advanced Industrial Science and Technology, International Patent Organism Depository, Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japón) el 6 de diciembre de 1994 y se le asignó un número de registro de FERM P-14690. Entonces, el depósito se convirtió en un depósito internacional según las disposiciones del Tratado de Budapest el 29 de septiembre de 1995 y se le asignó un número de registro de FERM BP-5252 (patente estadounidense n.º 5.827.698).

Los ejemplos de cepas originales que pueden usarse para derivar bacterias productoras de L-lisina también incluyen cepas en las que está aumentada la expresión de uno o más genes que codifican para una enzima biosintética de L-lisina. Los ejemplos de tales genes incluyen, pero sin limitarse, genes que codifican para dihidrodipicolinato sintasa (*dapA*), aspartocinasa (*lysC*), dihidrodipicolinato reductasa (*dapB*), diaminopimelato descarboxilasa (*lysA*), fosfoenolpiruvato carboxilasa (*ppc*), aspartato semialdehído deshidrogenasa (*asd*) y aspartasa (*aspA*) (documento EP 1253195 A). Además, las cepas originales pueden tener un nivel aumentado de expresión del gen implicado en la eficacia energética (*cyo*) (documento EP 1170376 A), el gen que codifica para nicotinamida nucleótido transhidrogenasa (*pntAB*) (patente estadounidense n.º 5.830.716), el gen *ybjE* (documento WO2005/073390), el gen que codifica para glutamato deshidrogenasa (*gdhA*, Gene, 23:199-209 (1983)) o una combinación de los mismos. Las abreviaturas para los genes de las enzimas se muestran en los paréntesis.

La secuencia de nucleótidos del gen *lysC* de *Escherichia coli* se muestra en SEQ ID NO: 21 y la secuencia de aminoácidos codificada de aspartocinasa se muestra en SEQ ID NO: 22. La secuencia de nucleótidos del gen *dapA* de *Escherichia coli* se muestra en SEQ ID NO: 23 y la secuencia de aminoácidos codificada de dihidrodipicolinato sintasa se muestra en SEQ ID NO: 24. Además, la secuencia de nucleótidos del gen *dapB* de *Escherichia coli* se muestra en SEQ ID NO: 25 y la secuencia de aminoácidos codificada de dihidrodipicolinato reductasa se muestra en SEQ ID NO: 26.

Se sabe que la dihidrodipicolinato sintasa de tipo natural derivada de *Escherichia coli* se somete a la retroinhibición por L-lisina y se sabe que la aspartocinasa de tipo natural derivada de *Escherichia coli* se somete a supresión y retroinhibición por L-lisina. Por tanto, cuando se usan los genes *dapA* y *lysC*, estos genes son preferiblemente genes que codifican para enzimas mutantes que están desensibilizadas para la retroinhibición por L-lisina.

Los ejemplos de ADN que codifican para una dihidrodipicolinato sintetasa mutante que está desensibilizada para la retroinhibición por L-lisina incluyen un ADN que codifica para una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 24 en la que el residuo de histidina en la posición 118 está sustituido por un residuo de tirosina. Los ejemplos de ADN que codifica para una aspartocinasa mutante que está desensibilizada para la retroinhibición por L-lisina incluyen un ADN que codifica para una AKIII que tiene la secuencia de aminoácidos de SEQ ID NO: 22 en la que el residuo de treonina en la posición 352, el residuo de glicina en la posición 323 y el residuo de metionina en la posición 318 están sustituidos por residuos de isoleucina, asparagina e isoleucina, respectivamente (patente estadounidense n.º 5.661.012 y patente estadounidense n.º 6.040.160). Tales ADN mutantes pueden obtenerse mediante mutagénesis específica de sitio usando PCR o similares.

Los plásmidos con amplia variedad de huéspedes RSFD80, pCAB1 y pCABD2 derivados de RSF1010 se conocen como plásmidos que contienen un gen *dapA* mutante que codifica para una dihidrodipicolinato sintasa de *Escherichia coli* mutante y un gen *lysC* que codifica para una aspartocinasa de *Escherichia coli* mutante (patente estadounidense n.º 6.040.160). La cepa de *Escherichia coli* JM109 transformada con RSFD80 se denominó AJ12396 (patente estadounidense n.º 6.040.160) y la cepa se depositó en el National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry (actualmente National Institute of Advanced Industrial Science and Technology, International Patent Organism Depository, Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japón) el 28 de octubre de 1993 y se le asignó un número de registro de FERM P-13936, y luego el depósito se convirtió en un depósito internacional según las disposiciones del Tratado de Budapest el 1 de noviembre de 1994 y se le asignó un número de registro de FERM BP-4859. RSFD80 puede obtenerse a partir de la cepa AJ12396 mediante un método convencional. pCAB1 se preparó insertando adicionalmente el gen *dapB* de *Escherichia coli* en RSFD80 descrito anteriormente. Además, pCABD2 se preparó insertando adicionalmente el gen *ddh* de la cepa de *Brevibacterium lactofermentum* (*Corynebacterium glutamicum*) 2256 (ATCC 13869) en pCAB1 descrito anteriormente (patente estadounidense n.º 6.040.160).

Los ejemplos de bacterias productoras de L-lisina o cepas originales para derivarlas también incluyen cepas en las que está disminuida o se ha vuelto deficiente la actividad de una enzima que cataliza una reacción que se ramifica

desde la ruta de biosíntesis de L-lisina y produce un compuesto distinto de L-lisina. Los ejemplos de tales enzimas que catalizan una reacción que se ramifica desde la ruta de biosíntesis de L-lisina y produce un compuesto distinto de L-lisina incluyen homoserina deshidrogenasa, lisina descarboxilasa (patente estadounidense n.º 5.827.698) y enzima málica (documento WO2005/010175). Es preferible que las expresiones de ambos genes *cadA* y *ldcC* que codifican para lisina descarboxilasa estén disminuidas con el fin de disminuir o eliminar la actividad de lisina descarboxilasa (documento WO2006/038695).

En la bacteria usada para la presente invención, con el fin de potenciar la capacidad de asimilación de glicerol, puede atenuarse la expresión del gen *glpR* (documento EP 1715056) o puede potenciarse la expresión de genes del metabolismo de glicerol (documento EP 1715055 A) tales como los genes *glpA*, *glpB*, *glpC*, *glpD*, *glpE*, *glpF*, *glpG*, *glpK*, *glpQ*, *glpT*, *glpX*, *tpiA*, *gldA*, *dhaK*, *dhaL*, *dhaM*, *dhaR*, *fsa* y *talC*.

<1-2> Construcción de *Escherichia coli* de la presente invención

A continuación en el presente documento se explicará la modificación para reducir la actividad de una enzima que pertenece a la ruta de síntesis de meso-DAP y la introducción de un gen que codifica para diaminopimelato deshidrogenasa.

La ruta de síntesis de meso-DAP de *Escherichia coli* es una ruta para generar meso-DAP (meso-2,6-diaminopimelato, meso- $\alpha,\epsilon$ -diaminopimelato o meso-diaminoheptanodioato) a partir de (S)-2,3,4,5-tetrahidropiridina-2,6-dicarboxilato ((S)-2,3,4,5-tetrahidrodipicolinato), y está catalizada por las enzimas para las siguientes reacciones de cuatro etapas. Estas reacciones son reacciones reversibles. En *Escherichia coli*, la ruta de síntesis de meso-DAP también se denomina ruta DapDCEF.

1) DapD (2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa, EC 2.3.1.117)

Succinil-CoA + (S)-2,3,4,5-tetrahidropiridina-2,6-dicarboxilato + H<sub>2</sub>O -> CoA + N-succinil-L-2-amino-6-oxoheptanodioato

DapD está codificada por el gen *dapD*. La secuencia del gen *dapD* de *Escherichia coli* se muestra en SEQ ID NO: 1 y la secuencia de aminoácidos de DapD se muestra en SEQ ID NO: 2.

La actividad enzimática de DapD puede medirse haciendo referencia al método de S.A, Simms *et al.* (J. Biol. Chem., 10 de marzo de 1984; 259(5):2734-2791).

2) DapC (succinildiaminopimelato transaminasa, también denominada SDAP aminotransferasa, EC 2.6.1.17)

N-Succinil-LL-2,6-diaminoheptanodioato + 2-oxoglutarato -> N-succinil-L-2-amino-6-oxoheptanodioato + L-glutamato

DapC está codificada por el gen *dapC*. La secuencia del gen *dapC* de *Escherichia coli* se muestra en SEQ ID NO: 3 y la secuencia de aminoácidos se muestra en SEQ ID NO: 4.

La actividad enzimática de DapC puede medirse mediante el método de Thilo, M. *et al.* (J. Bacteriol., julio de 2000; 182 (13):3626-3631).

3) DapE (succinildiaminopimelato desuccinilasa, también denominada enzima de desuccinilación de SDAP, EC 3.5.1.18)

N-Succinil-LL-2,6-diaminoheptanodioato + H<sub>2</sub>O -> succinato + LL-2,6-diaminoheptanodioato

DapE está codificada por el gen *dapE*. La secuencia del gen *dapE* de *Escherichia coli* se muestra en SEQ ID NO: 5 y la secuencia de aminoácidos se muestra en SEQ ID NO: 6.

La actividad enzimática de DapE puede medirse mediante el método de Lin, Y.K. *et al.* (J. Biol. Chem., 5 de febrero de 1988; 263(4):1622-1627).

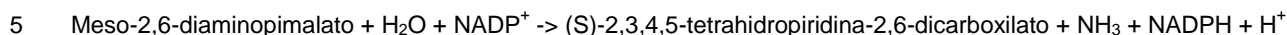
4) DapF (diaminopimelato epimerasa, EC 5.1.1.7)

LL-2,6-Diaminoheptanodioato -> meso-2,6-diaminopimelato

DapF está codificada por el gen *dapF*. La secuencia del gen *dapF* de *Escherichia coli* se muestra en SEQ ID NO: 7 y la secuencia de aminoácidos se muestra en SEQ ID NO: 8.

La actividad enzimática de DapF puede medirse haciendo referencia al método de Wiseman, J.S. *et al.* (J. Biol. Chem., 25 de julio de 1984; 259(14):8907-8914).

Además, en la presente invención, DDH (diaminopimelato deshidrogenasa, EC 1.4.1.16) es una enzima que genera de manera reversible (S)-2,3,4,5-tetrahidropiridina-2,6-dicarboxilato a partir de meso-2,6-diaminopimalato y cataliza la siguiente reacción.



La actividad enzimática de DDH puede medirse haciendo referencia al método de Misono, H. *et al.* (J. Biol. Chem., 255, 10599-10605, 1980).

10 Aunque las bacterias *Escherichia* no tienen ningún gen *ddh* que codifica para DDH, puede usarse un gen *ddh* de una bacteria corineforme tal como los de *Corynebacterium glutamicum* (SEQ ID NO: 9), *Brevibacterium lactofermentum* (SEQ ID NO: 11) y *Corynebacterium efficiens* (SEQ ID NO: 13).

15 El gen *ddh* de *Corynebacterium glutamicum*, ATCC 13032 (NCg12528), está registrado como Genbank NP\_601818.2. GI:23308957, y el gen *ddh* de *Corynebacterium efficiens* (CE2498) está registrado como NP\_739108.1. GI: 25029054.

20 Además de los de las bacterias corineformes, puede usarse el gen *ddh* de *Herminiimonas arsenicoxydans* (SEQ ID NO: 15) y el gen *ddh* de *Bacteroides thetaiotaomicron* (SEQ ID NO: 17). El gen *ddh* de *Herminiimonas arsenicoxydans* está registrado como Genbank YP\_001100730.1 GI:134095655 y el gen *ddh* de *Bacteroides thetaiotaomicron* está registrado como NP\_810892.1. GI:29347389.

25 Los genes mencionados anteriormente y los genes de la enzima de biosíntesis de L-lisina mencionados anteriormente no se limitan a genes que se corresponden exactamente con la información génica descrita anteriormente ni con genes que tienen secuencias conocidas, y pueden ser genes que tienen una mutación conservadora tal como homólogos de los mismos y genes modificados artificialmente, siempre que no se transmitan las funciones de las proteínas codificadas. Es decir, los genes pueden ser un gen que codifica para una secuencia de aminoácidos de una proteína conocida que incluye sustituciones, deleciones, inserciones, adiciones o similares de uno o varios residuos de aminoácidos en una o varias posiciones.

30 Aunque el número representado por el término “uno o varios” usado en el presente documento puede diferir dependiendo de las posiciones en la estructura tridimensional de la proteína o de los tipos de residuos de aminoácidos, específicamente, puede ser preferiblemente de 1 a 20, más preferiblemente de 1 a 10, todavía más preferiblemente de 1 a 5. La mutación conservadora es normalmente una sustitución conservadora. La sustitución conservadora es una sustitución en la que la sustitución tiene lugar mutuamente entre Phe, Trp y Tyr, si el sitio de sustitución es un aminoácido aromático; entre Leu, Ile y Val, si el sitio de sustitución es un aminoácido hidrófobo; entre Gln y Asn, si es un aminoácido polar; entre Lys, Arg y His, si es un aminoácido básico; entre Asp y Glu, si es un aminoácido ácido; y entre Ser y Thr, si es un aminoácido que tiene un grupo hidroxilo. Los ejemplos específicos de sustitución considerada sustitución conservadora incluyen: sustitución de Ser o Thr por Ala; sustitución de Gln, His o Lys por Arg; sustitución de Glu, Gln, Lys, His o Asp por Asn; sustitución de Asn, Glu o Gln por Asp; sustitución de Ser o Ala por Cys; sustitución de Asn, Glu, Lys, His, Asp o Arg por Gln; sustitución de Gly, Asn, Gln, Lys o Asp por Glu; sustitución de Pro por Gly; sustitución de Asn, Lys, Gln, Arg o Tyr por His; sustitución de Leu, Met, Val o Phe por Ile; sustitución de Ile, Met, Val o Phe por Leu; sustitución de Asn, Glu, Gln, His o Arg por Lys; sustitución de Ile, Leu, Val o Phe por Met; sustitución de Trp, Tyr, Met, Ile o Leu por Phe; sustitución de Thr o Ala por Ser; sustitución de Ser o Ala por Thr; sustitución de Phe o Tyr por Trp; sustitución de His, Phe o Trp por Tyr; y sustitución de Met, Ile o Leu por Val. La mutación para tal sustitución, deleción, inserción, adición, inversión o similares de residuos de aminoácidos tal como se describió anteriormente también incluye una mutación que se produce de manera natural basándose en la diferencia individual, la diferencia en especies de microorganismos de los que se derivan los genes (mutantes o variantes) etc. Un gen de este tipo puede obtenerse modificando una secuencia de nucleótidos de un gen conocido mediante, por ejemplo, mutagénesis específica de sitio, de modo que la sustitución, deleción, inserción o adición de un residuo o residuos de aminoácido se incluye en un sitio específico de la proteína codificada.

55 Además, un gen que tiene la mutación conservadora mencionada anteriormente puede ser un gen que codifica para una proteína que muestra una homología del 80% o más, preferiblemente del 90% o más, más preferiblemente del 95% o más, de manera particularmente preferible del 97% o más, con la secuencia entera de la proteína codificada y que tiene una función equivalente a la de una proteína de tipo natural correspondiente. En esta memoria descriptiva, el término “homología” puede usarse para referirse a “identidad”.

60 Además, pueden sustituirse codones de las secuencias génicas por los que usa fácilmente un huésped en el que se introducen los genes.

Los genes que tienen una mutación conservadora pueden ser los obtenidos mediante un método habitualmente usado para mutagénesis tal como tratamiento con un mutagén.

65 Además, los genes también pueden ser un ADN que puede hibridarse con una sonda que puede prepararse a partir de una secuencia génica conocida, por ejemplo, las secuencias génicas mencionadas anteriormente y



secuencias complementarias a ellas, en condiciones rigurosas y que codifica para una proteína que tiene una función equivalente a la de un producto génico conocido correspondiente. Las "condiciones rigurosas" se refieren a condiciones en las que se forma un híbrido denominado específico y no se forma un híbrido no específico. Los ejemplos de las condiciones rigurosas incluyen, por ejemplo, condiciones en las que ADN que muestran alta homología entre sí, por ejemplo, ADN que muestran una homología de, por ejemplo, no menos del 80%, preferiblemente no menos del 90%, más preferiblemente no menos del 95%, de manera particularmente preferible no menos del 97%, se hibridan entre sí, y ADN que tienen una homología inferior al nivel anterior no se hibridan entre sí, y condiciones de lavado en hibridación de tipo Southern convencional, es decir, condiciones de lavado una vez, preferiblemente dos veces o tres veces, a concentraciones de sal y temperatura de 1 x SSC, SDS al 0,1% a 60°C, preferiblemente 0,1 x SSC, SDS al 0,1% a 60°C, más preferiblemente 0,1 x SSC, SDS al 0,1% a 68°C.

Como sonda, también puede usarse una parte de secuencias complementarias de los genes. Una sonda de este tipo puede producirse mediante PCR usando oligonucleótidos preparados basándose en una secuencia génica conocida como cebadores y un fragmento de ADN que incluye la secuencia de nucleótidos del gen como molde. Cuando se usa un fragmento de ADN que tiene una longitud de aproximadamente 300 pb como sonda, las condiciones de lavado tras la hibridación en las condiciones mencionadas anteriormente pueden mostrarse a modo de ejemplo mediante 2 x SSC, SDS al 0,1% a 50°C.

La expresión "que va a modificarse de modo que disminuya una actividad enzimática de la ruta de síntesis de meso-DAP" significa que va a modificarse de modo que se elimine completamente o disminuya la actividad de una enzima que pertenece a la ruta de síntesis de meso-DAP (también denominada en lo sucesivo "ruta DapDCEF"), específicamente, al menos una de las cuatro enzimas, DapD, DapC, DapE y DapF, en comparación con la de una cepa no modificada de *Escherichia coli*, por ejemplo, una cepa de tipo natural.

La enzima de la que se disminuye la actividad puede ser cualquiera de DapD, DapC, DapE y DapF, y puede consistir en dos o más clases de ellas. Resulta preferible disminuir una actividad de una enzima que funciona en la región anterior de la ruta DapDCEF, y es particularmente preferible realizar una modificación de modo que al menos disminuya la actividad DapD.

La disminución de una actividad enzimática de la ruta DapDCEF significa preferiblemente que, por ejemplo, cada actividad enzimática en la ruta DapDCEF disminuye hasta el 50% o menos, preferiblemente el 30% o menos, más preferiblemente el 10% o menos, por célula, en comparación con la de una cepa no modificada, por ejemplo, una cepa de tipo natural.

Los ejemplos de la cepa como objeto de la comparación incluyen *Escherichia coli* W3110 (ATCC 27325), *Escherichia coli* MG1655 (ATCC 47076) etc., derivadas de la cepa de tipo natural prototipo, cepa K12, como cepa de tipo natural.

Una modificación de este tipo en que disminuye una actividad enzimática en la ruta DapDCEF se logra específicamente delecionando una parte de o todo el gen de un cromosoma que codifica para una enzima de la ruta DapDCEF, específicamente, la región codificante del gen *dapD*, *dapC*, *dapE* o *dapF*, o modificando una secuencia de control de la expresión tal como la secuencia promotora y de Shine-Dalgarno (SD). Además, la expresión de un gen también puede disminuirse mediante la modificación de una región de no traducción distinta de la secuencia de control de la expresión. Además, puede delecionarse todo el gen que incluye las secuencias a ambos lados del gen en un cromosoma. Además, también puede lograrse mediante la introducción de una mutación para una sustitución de aminoácido (mutación de aminoácido), un codón de terminación (mutación sin sentido) o una mutación de cambio de marco que añade o deleciona uno o dos nucleótidos, dando lugar a una región codificante que codifica para una enzima en un cromosoma (Journal of Biological Chemistry, 272:8611-8617 (1997); Proceedings of the National Academy of Sciences, USA, 95 5511-5515 (1998); Journal of Biological Chemistry, 266, 20833-20839 (1991)).

En la presente invención, es preferible disminuir una actividad enzimática intracelular delecionando una parte de o toda la secuencia reguladora de la expresión del gen en un cromosoma tal como una región promotora, o una parte o toda la región codificante o una región no codificante, o insertando otras secuencias en una de esas secuencias usando recombinación homóloga. Sin embargo, la modificación puede ser una modificación producida por una mutagénesis habitual basada en irradiación con rayos X o ultravioleta o por el uso de un mutágeno tal como N-metil-N'-nitro-N-nitrosoguanidina, siempre que la modificación disminuya una actividad enzimática de la ruta DapDCEF.

Modificación de una secuencia de control de la expresión se realiza preferiblemente para uno o más nucleótidos, más preferiblemente dos o más nucleótidos, de manera particularmente preferible tres o más nucleótidos. Cuando se realiza la deleción para una región codificante, la región que va a delecionarse puede ser una región N-terminal, una región interna o una región C-terminal, o incluso toda la región codificante, siempre que se disminuya o elimine la función de la proteína enzimática que va a producirse. La deleción de una región más larga puede inactivar habitualmente de manera más segura un gen. Además, se prefiere que los marcos de lectura que se ubican en el sentido de 5' y en el sentido de 3' de la región que va a delecionarse no sean iguales.

Cuando se inserta otra secuencia en una región codificante, la secuencia puede insertarse en cualquier región del gen, y la inserción de una secuencia más larga puede inactivar habitualmente de manera más segura el gen que codifica para una enzima. Se prefiere que los marcos de lectura que se ubican en el sentido de 5' y en el sentido de 3' del sitio de inserción no sean iguales. La otra secuencia no está limitada particularmente siempre que se escoja una secuencia que disminuya o elimine la función de la proteína enzimática, y los ejemplos incluyen, por ejemplo, un transposón que porta un gen de resistencia a antibióticos o un gen útil para la producción de L-lisina, etc.

Un gen en un cromosoma puede modificarse tal como se describió anteriormente, por ejemplo, preparando una versión de tipo deleción del gen en el que se deleciona una secuencia parcial del gen de modo que el gen de tipo deleción no produzca una proteína que no funciona normalmente, y transformando una bacteria con un ADN que contiene el gen de tipo deleción para producir recombinación homóloga entre el gen de tipo deleción y el gen nativo en el cromosoma, y de ese modo sustituir el gen de tipo deleción por el gen en el cromosoma. La proteína codificada por el gen de tipo deleción tiene una conformación diferente de la de la proteína enzimática de tipo natural, si acaso se produce, y por tanto la función está disminuida o eliminada. Ya se ha establecido una alteración génica de este tipo basada en la sustitución génica que utiliza recombinación homóloga y existe un método denominado integración conducida por Red ("*Red-driven integration*") (Datsenko, K.A, y Wanner, B.L., Proc. Natl. Acad. Sci. USA, 97:6640-6645 (2000)), un método de uso de un ADN lineal tal como un método que utiliza la integración conducida por Red en combinación con un sistema de escisión derivado del fago  $\lambda$  (Cho, E.H., Gumpert, R.I., Gardner, J.F., J. Bacteriol., 184: 5200-5203 (2002)), un método de uso de un plásmido que contiene un origen de replicación sensible a temperatura o un plásmido que puede realizar transferencia conjugativa, un método de utilización de un vector suicida que no tiene origen de replicación en un huésped (patente estadounidense n.º 6.303.383, patente japonesa abierta a consulta por el público n.º 05-007491), etc.

La disminución de la cantidad de expresión de un gen puede confirmarse comparando la cantidad de ARNm transcrito a partir del gen con la de una cepa de tipo natural o no modificada. La cantidad de expresión puede confirmarse mediante hibridación de tipo Northern, RT-PCR (Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA, 2001)), etc.

La disminución de la cantidad de una proteína codificada por un gen puede confirmarse mediante inmunotransferencia de tipo Western usando anticuerpos (Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA, 2001).

Con el fin de introducir un gen que codifica para DDH (*ddh*) en *Escherichia coli*, por ejemplo, puede transformarse *Escherichia coli* con un gen *ddh* mediante el uso de un vector tal como un plásmido y un fago. Los ejemplos de tales vectores incluyen pUC19, pUC18, pBR322, pHSG299, pHSG298, pHSG399, pHSG398, RSF1010, pMW119, pMW118, pMW219, pMW218, etc. Aunque puede usarse un promotor inherente del gen *ddh*, siempre que el gen pueda expresarse en *Escherichia coli* con él, también puede usarse un promotor que funciona eficazmente en *Escherichia coli*. Los ejemplos de tales promotores incluyen promotor lac, promotor trp, promotor trc, promotor tac, promotor PR y promotor PL del fago  $\lambda$ , promotor tet, etc.

El gen *ddh* también puede incorporarse en el cromosoma de *Escherichia coli* mediante un método que usa transducción, un transposón (Berg, D.E. y Berg, C.M., Bio/Technol., 1, 417 (1983)), el fago Mu (patente japonesa abierta a consulta por el público n.º 2-109985) o recombinación homóloga (Experiments in Molecular Genetics, Cold Spring Harbor Lab. (1972)). Además, el número de copias del gen *ddh* puede aumentarse transfiriendo los genes *ddh* incorporados en el cromosoma.

La incorporación del gen *ddh* puede confirmarse mediante, por ejemplo, hibridación de tipo Southern. Además, puede confirmarse si *Escherichia coli* en la que se ha introducido el gen *ddh* tiene la actividad en DDH o no, por ejemplo, midiendo la actividad de DDH según el método descrito en Haruo Misono, Fermentation and Industry, 45, 964 (1987). Además, también puede detectarse DDH mediante inmunotransferencia de tipo Western usando un anticuerpo.

#### <2> Método para producir L-lisina

El método para producir L-lisina de la presente invención comprende hacer crecer en cultivo la bacteria de la presente invención en un medio para producir y acumular L-lisina en el medio o las células, y recoger L-lisina del medio o las células.

Como el medio que va a usarse, pueden usarse medios usados de manera convencional en la producción de L-lisina mediante fermentación usando microorganismos. Es decir, pueden usarse medios convencionales que contienen una fuente de carbono, una fuente de nitrógeno, iones inorgánicos y opcionalmente otros componentes orgánicos según se requiera. Como fuente de carbono, pueden usarse sacáridos tales como glucosa, sacarosa, lactosa, galactosa, fructosa e hidrolizado de almidón; alcoholes tales como glicerol y sorbitol; y ácidos orgánicos tales como ácido fumárico, ácido cítrico y ácido succínico. Como fuente de nitrógeno, pueden usarse sales de amonio inorgánicas tales como sulfato de amonio, cloruro de amonio y fosfato de amonio, nitrógeno orgánico tal como hidrolizado de soja, gas amoniacal, amoniacal acuoso, etc. En cuanto a las fuentes de nutrientes traza orgánicas, es

deseable que el medio contenga sustancias requeridas tales como vitamina B<sub>1</sub> y L-homoserina, extracto de levadura o similares en cantidades apropiadas. Además de los anteriores, se añaden fosfato de potasio, sulfato de magnesio, iones hierro, iones manganeso etc. en pequeñas cantidades, según se requiera. Además, el medio usado en la presente invención puede ser o bien un medio natural o bien un medio sintético, siempre que se use un medio que contenga una fuente de carbono, una fuente de nitrógeno e iones inorgánicos, y que contenga otros componentes traza orgánicos según se requiera.

En la presente invención, se usa glicerol de manera particularmente preferible como fuente de carbono. Aunque el glicerol puede ser glicerol como reactivo, es deseable usar glicerol producido industrialmente que contiene impurezas. Por ejemplo, es deseable usar glicerol producido industrialmente mediante la reacción de esterificación para la producción de combustible biodiésel (Mu Y, *et al*, *Biotechnol Lett.*, 28, 1755-91759 (2006); Haas M.J., *et al.*, *Bioresour. Technol.*, 97, 4, 671-8678 (2006)).

El glicerol contenido en el medio usado en la presente invención puede ser una única fuente de carbono, o también puede usarse un medio mixto al que se añaden otras fuentes de carbono además de glicerol. Otras fuentes de carbono preferibles son sacáridos tales como glucosa, fructosa, sacarosa, lactosa, galactosa, melaza residual, hidrolizado de almidón, disolución de azúcar obtenida mediante hidrólisis de biomasa, alcoholes tales como etanol, y ácidos orgánicos tales como ácido fumárico, ácido cítrico y ácido succínico. Cuando se usa un medio mixto, es deseable que el glicerol esté contenido en el medio en una razón del 50% o más, preferiblemente del 60% o más, más preferiblemente del 70% o más, todavía más preferiblemente del 80% o más, de manera especialmente preferible del 90% o más, basado en la fuente de carbono total contenida en el medio.

El cultivo se realiza preferiblemente durante de 1 a 7 días en condiciones aerobias. La temperatura de cultivo es preferiblemente de 24 a 37°C y el pH durante el cultivo es preferiblemente de entre 5 y 9. Para ajustar el pH, pueden usarse sustancias alcalinas o ácidas orgánicas o inorgánicas, gas amoniacal, etc. Para recoger la L-lisina del medio de fermentación, puede usarse una combinación de métodos conocidos, tal como mediante el uso de una resina de intercambio iónico y precipitación. Cuando se acumula la L-lisina en las células, pueden usarse ondas supersónicas, por ejemplo, o similares para alterar las células y puede recogerse la L-lisina mediante el uso de una resina de intercambio iónico o similar, del sobrenadante obtenido retirando las células de la suspensión con células alteradas mediante centrifugación.

La producción también puede realizarse mediante un método en el que se realiza la fermentación controlando que el pH del medio durante el cultivo sea de 6,5 a 9,0 y que el pH del medio al final del cultivo sea de 7,2 a 9,0 y controlando que la presión en el tanque de fermentación durante la fermentación sea positiva, o suministrando dióxido de carbono o un gas mixto que contiene dióxido de carbono al medio de modo que estén presentes iones bicarbonato y/o iones carbonato en el medio de cultivo en una cantidad de al menos 2 g/l durante el cultivo, y estos iones bicarbonato y/o iones carbonato sirven como contraiones de cationes que consisten principalmente en aminoácidos básicos, y entonces se recoge lisina (véase la patente japonesa abierta a consulta por el público n.º 2002-065287, solicitud de patente publicada estadounidense n.º 2002/025564).

## Ejemplos

A continuación en el presente documento, la presente invención se explicará todavía más específicamente con referencia a los ejemplos.

Ejemplo 1: Construcción de la bacteria productora de L-lisina con alteración de *dapD*

<1-1> Construcción de la cepa con alteración del gen *dapD*

En primer lugar, se construyó una cepa con alteración de *dapD* usando la cepa de tipo natural *Escherichia coli*, la cepa MG1655.

Usando el plásmido pMW118 ( $\lambda$ .attL-Km<sup>r</sup>- $\lambda$ .attR) (véase la publicación de patente internacional WO2006/093322) como molde y oligonucleótidos sintéticos de SEQ ID NO: 19 y 20 que tienen secuencias correspondientes a ambos extremos de las secuencias de los sitios de unión del fago  $\lambda$ , *attL* y *attR*, en los extremos 3' y secuencias correspondientes a partes del gen *dapD* como el gen diana en los extremos 5' como cebadores, se realizó una PCR para construir la cepa MG1655 $\Delta$ dapD::att-Km según el método  $\lambda$ -red descrito en la solicitud publicada de patente estadounidense n.º 2006/0160191 y en el documento WO2005/010175. Se obtuvo un recombinante resistente a Km según el método  $\lambda$ -red haciendo crecer en cultivo la bacteria a 37°C en un medio de L-agar que contiene Km (kanamicina, 50 mg/l) como cultivo en placa y seleccionando un recombinante resistente a Km.

<1-2> Transducción de la cepa WC196LC/pCABD2 de la bacteria productora de L-lisina con  $\Delta$ dapD::att-kan

Se obtuvo el lisado P1 a partir de la cepa MG1655 $\Delta$ dapD::att-Km obtenida en <1-1> de una manera convencional y se usó la cepa WC196 $\Delta$ cadA $\Delta$ ldcC/pCABD2 de la bacteria productora de L-lisina construida mediante el método

descrito en la solicitud publicada de patente estadounidense n.º 2006/0160191 como huésped para construir la cepa WC196 $\Delta$ cadA $\Delta$ ldcChdapD::att-Km/pCABD2 según el método de transducción de P1. Se obtuvo la cepa WC196 $\Delta$ cadA $\Delta$ ldc a partir de la cepa WC196 de *Escherichia coli* mediante la alteración de genes de lisina descarboxilasa, *cadA* y *ldc* según el método usando el método de integración de conducido por Red (Datsenko K.A., Wanner, B.L., 2000, Proc. Natl. Acad. Sci. USA, 97, 6640-6645) y el sistema de escisión derivado del fago  $\lambda$  (Cho, E.H., Gumpert, R.L., Gardner, J.F., J. Bacteriol., 184:5200-5203 (2002)) en combinación (véase el documento WO2005/010175). Una cepa obtenida introduciendo pCABD2 en esta cepa es la cepa WC196 $\Delta$ cadA $\Delta$ ldcC/pCABD2.

Se obtuvo la cepa transducida objetivo realizando un cultivo en placa a 37°C en un medio de L-agar que contenía Km (kanamicina, 50 mg/l) y Sm (estreptomocina, 20 mg/l), y seleccionando una cepa recombinante resistente a Km y resistente a Sm.

Además, se hicieron crecer en cultivo cada una de estas cepas a 37°C en un medio de L que contenía 20 mg/l de estreptomocina hasta que la DO600 final alcanzó aproximadamente 0,6, entonces se añadió el cultivo un volumen igual de disolución de glicerol al 40% y se agitó la mezcla, se dividió en volúmenes apropiados y se almacenaron a -80°C. Estas mezclas se denominarán disoluciones madre de glicerol.

Ejemplo 2: Evaluación de la capacidad de producción de L-lisina de bacterias productoras de L-lisina con alteración de *dapD*

Se descongelaron las disoluciones madre de glicerol de las cepas obtenidas en el ejemplo 1 y se aplicaron uniformemente a una L-placa que contenía 20 mg/l de estreptomocina en un volumen de 100  $\mu$ l cada una y se realizó el cultivo a 37°C durante 24 horas. Se suspendieron las células en una cantidad de aproximadamente 1/8 de las células obtenidas en una placa en 0,5 ml de solución salina fisiológica y se midió la turbidez de la suspensión a 600 nm usando un espectrofotómetro (U-2000, Hitachi). Se inoculó la suspensión que contenía las células obtenidas en 20 ml de un medio de fermentación (medio MS, la composición se muestra más adelante) que contenía 20 mg/l de estreptomocina en un matraz Sakaguchi de 500 ml en un volumen tal que la turbidez de la mezcla llegó a ser de 0,15 a 600 nm, y se realizó el cultivo a 37°C y 114 rpm durante 24 horas en una máquina de cultivo con agitación a vaivén. Tras el cultivo, se midieron las cantidades de L-lisina acumulada y la glucosa que quedaba en el medio usando en analizador Biotec AS210 (SAKURA SEIKI). También se midió el glicerol acumulado en el medio usando el analizador Biotec BF-5 (Oji Scientific Instruments).

Composición del medio de fermentación, g/l

Glicerol o glucosa	40
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	24
KH <sub>2</sub> PO <sub>4</sub>	11,0
MgSO <sub>4</sub> • 7H <sub>2</sub> O	1,0
FeSO <sub>4</sub> • 7H <sub>2</sub> O	0,01
MnSO <sub>4</sub> • 5H <sub>2</sub> O	0,01
Extracto de levadura	2,0

Se ajustó el medio a pH 7,0 con KOH, se sometió a la autoclave a 115°C durante 10 minutos (siempre que el glicerol o la glucosa y MgSO<sub>4</sub>•7H<sub>2</sub>O se esterilizaran por separado) y luego se añadieron 30 g/l de CaCO<sub>3</sub> (Farmacopea Japonesa) (se sometió a esterilización con aire caliente a 180°C durante 2 horas). Se añadió estreptomocina en una cantidad de 20 mg/l.

Los resultados se muestran en la tabla 1 (DO significa la cantidad de células indicada en lo que se refiere a la absorbancia a 660 nm medida para el cultivo diluido 26 veces, Lys (g/l) significa la cantidad de L-lisina acumulada en el matraz, glucosa (g/l) y glicerol (g/l) significan las cantidades de glucosa y glicerol que quedan en el medio, respectivamente, y rendimiento (%) significa el rendimiento de L-lisina basándose en el sustrato). Tal como se observa a partir de los resultados mostrados en la tabla 1, la cepa WC196 $\Delta$ cadA $\Delta$ ldcC $\Delta$ dapD::att-Km/pCABD2 acumuló L-lisina en una cantidad mayor en comparación con la obtenida con la cepa WC196 $\Delta$ cadA $\Delta$ ldcC/pCABD2 en la que el gen *dapD* no estaba alterado.

Tabla 1

Medio de glucosa MS

Cepa	D.O. (x26)	Lys (g/l)	Glucosa (g/l)	Rendimiento (%)
Cepa WC196 $\Delta$ cadA $\Delta$ ldcC/pCABD2	0,361	8,28	18,67	38,82
Cepa WC196 $\Delta$ cadA $\Delta$ ldcC $\Delta$ dapD::att-Km/pCABD2	0,377	8,75	18,13	40,01

## ES 2 403 199 T3

Medio de glicerol MS

Cepa	D.O. (x26)	Lys (g/l)	Glicerol (g/l)	Rendimiento (%)
Cepa WC196 $\Delta$ cadA $\Delta$ ldcC/pCABD2	0,248	2,79	28,64	24,57
Cepa WC196 $\Delta$ cadA $\Delta$ ldcC $\Delta$ dapD::att-Km/pCABD2	0,320	5,10	23,69	31,12

Explicación de la lista de secuencias

- 5
- SEQ ID NO: 1: Secuencia de nucleótidos de *dapD* de *E. coli*
- SEQ ID NO: 2: Secuencia de aminoácidos de DapD de *E. coli*
- 10
- SEQ ID NO: 3: Secuencia de nucleótidos de *dapC* de *E. coli*
- SEQ ID NO: 4: Secuencia de aminoácidos de DapC de *E. coli*
- SEQ ID NO: 5: Secuencia de nucleótidos de *dapE* de *E. coli*
- 15
- SEQ ID NO: 6: Secuencia de aminoácidos de DapE de *E. coli*
- SEQ ID NO: 7: Secuencia de nucleótidos de *dapF* de *E. coli*
- 20
- SEQ ID NO: 8: Secuencia de aminoácidos de DapF de *E. coli*
- SEQ ID NO: 9: Secuencia de nucleótidos del gen *ddh* de *C. glutamicum*
- SEQ ID NO: 10: Secuencia de aminoácidos de DDH de *C. glutamicum*
- 25
- SEQ ID NO: 11: Secuencia de nucleótidos de gen *ddh* de *B. lactofermentum*
- SEQ ID NO: 12: Secuencia de aminoácidos de DDH de *B. lactofermentum*
- 30
- SEQ ID NO: 13: Secuencia de nucleótidos de gen *ddh* de *C. efficiens*
- SEQ ID NO: 14: Secuencia de aminoácidos de DDH de *C. efficiens*
- SEQ ID NO: 15: Secuencia de nucleótidos de gen *ddh* de *H. arsenicoxydans*
- 35
- SEQ ID NO: 16: Secuencia de aminoácidos de DDH de *H. arsenicoxydans*
- SEQ ID NO: 17: Secuencia de nucleótidos de gen *ddh* de *B. thetaiotaomicron*
- 40
- SEQ ID NO: 18: Secuencia de aminoácidos de DDH de *B. thetaiotaomicron*
- SEQ ID NO 19: Cebador para la delección del gen *dapD*
- SEQ ID NO: 20: Cebador para la delección del gen *dapD*
- 45
- SEQ ID NO: 21: Secuencia de nucleótidos de *lysC* de *E. coli*
- SEQ ID NO: 22: Secuencia de aminoácidos de LysC de *E. coli*
- 50
- SEQ 10 NO: 23: Secuencia de nucleótidos de *dapA* de *E. coli*
- SEQ ID NO: 24: Secuencia de aminoácidos de DapA de *E. coli*
- SEQ ID NO: 25: Secuencia de nucleótidos de *dapB* de *E. coli*
- 55
- SEQ ID NO: 26: Secuencia de aminoácidos de DapB de *E. coli*

Aplicabilidad industrial

Según la presente invención, en la producción de L-lisina mediante fermentación usando *Escherichia coli*, puede mejorarse la cantidad de producción y/o el rendimiento de fermentación de L-lisina. Además, la presente invención puede usarse para obtener una bacteria productora de L-lisina de *Escherichia coli*.

5 LISTA DE SECUENCIAS

<110> Ajinomoto Co., Inc.

<120> Método para producir L-lisina

10

<130> C919-C8129

<150> JP2007-190795

<151> 2007-07-23

15

<160> 26

<170> PatentIn versión 3.5

20

<210> 1

<211> 825

<212> ADN

<213> *Escherichia coli*

25

<220>

<221> CDS

<222> (1)..(825)

30

<400> 1

ES 2 403 199 T3

```

atg cag cag tta cag aac att att gaa acc gct ttt gaa cgc cgt gcc      48
Met Gln Gln Leu Gln Asn Ile Ile Glu Thr Ala Phe Glu Arg Arg Ala
1          5          10          15
gag atc acg cca gcc aat gca gac acc gtt acc cgc gaa gcg gta aat      96
Glu Ile Thr Pro Ala Asn Ala Asp Thr Val Thr Arg Glu Ala Val Asn
          20          25          30
cag gtg atc gcc ctg ctg gat tcc ggc gca ctg cgt gta gcg gaa aaa      144
Gln Val Ile Ala Leu Leu Asp Ser Gly Ala Leu Arg Val Ala Glu Lys
          35          40          45
att gac ggt cag tgg gtg acg cat cag tgg ttg aaa aaa gcg gtg ctg      192
Ile Asp Gly Gln Trp Val Thr His Gln Trp Leu Lys Lys Ala Val Leu
          50          55          60
ctc tct ttc cgt att aat gat aat cag gtg atc gaa ggg gca gaa agc      240
Leu Ser Phe Arg Ile Asn Asp Asn Gln Val Ile Glu Gly Ala Glu Ser
65          70          75          80
cgc tac ttc gac aaa gtg ccg atg aaa ttc gcc gac tac gac gaa gca      288
Arg Tyr Phe Asp Lys Val Pro Met Lys Phe Ala Asp Tyr Asp Glu Ala
          85          90          95
cgt ttc cag aaa gaa ggc ttc cgc gtt gtg cca cca gcg gcg gta cgt      336
Arg Phe Gln Lys Glu Gly Phe Arg Val Val Pro Pro Ala Ala Val Arg
          100          105          110
cag ggt gcg ttt att gcc cgt aac acc gtg ctg atg ccg tct tac gtc      384
Gln Gly Ala Phe Ile Ala Arg Asn Thr Val Leu Met Pro Ser Tyr Val
          115          120          125
aac atc ggc gca tat gtt gat gaa ggc acc atg gtt gat acc tgg gcg      432
Asn Ile Gly Ala Tyr Val Asp Glu Gly Thr Met Val Asp Thr Trp Ala
          130          135          140
acc gtc ggt tct tgt gcg cag att ggt aaa aac gtc cac ctt tcc ggt      480
Thr Val Gly Ser Cys Ala Gln Ile Gly Lys Asn Val His Leu Ser Gly
145          150          155          160
ggc gtg ggc atc ggc ggc gtg ctg gaa ccg ctg cag gct aac cca acc      528
Gly Val Gly Ile Gly Gly Val Leu Glu Pro Leu Gln Ala Asn Pro Thr
          165          170          175
atc att gaa gat aat tgc ttc atc ggc gcg cgc tct gaa gtg gtt gaa      576
Ile Ile Glu Asp Asn Cys Phe Ile Gly Ala Arg Ser Glu Val Val Glu
          180          185          190

ggg gtg att gtc gaa gaa ggt tcc gtc att tcc atg ggc gta tac att      624
Gly Val Ile Val Glu Glu Gly Ser Val Ile Ser Met Gly Val Tyr Ile
          195          200          205
ggt cag agc acc cgt att tac gac cgt gaa acc ggc gaa atc cac tac      672
Gly Gln Ser Thr Arg Ile Tyr Asp Arg Glu Thr Gly Glu Ile His Tyr
          210          215          220
ggt cgc gtt ccg gcg ggg tct gtg gtt gtt tca ggt aat ctg ccg tca      720
Gly Arg Val Pro Ala Gly Ser Val Val Val Ser Gly Asn Leu Pro Ser
225          230          235          240
aaa gat ggc aaa tac agc ctc tac tgt gcg gtt atc gtt aag aaa gtt      768
Lys Asp Gly Lys Tyr Ser Leu Tyr Cys Ala Val Ile Val Lys Lys Val
          245          250          255
gac gcg aaa act cgc ggc aaa gtc ggc att aac gaa ctg ctg cgt acc      816
Asp Ala Lys Thr Arg Gly Lys Val Gly Ile Asn Glu Leu Leu Arg Thr
          260          265          270
atc gac taa
Ile Asp
825

```

5 <210> 2  
 <211> 274  
 <212> PRT  
 <213> *Escherichia coli*

10 <400> 2

ES 2 403 199 T3

Met Gln Gln Leu Gln Asn Ile Ile Glu Thr Ala Phe Glu Arg Arg Ala  
 1 5 10 15  
 Glu Ile Thr Pro Ala Asn Ala Asp Thr Val Thr Arg Glu Ala Val Asn  
 20 25 30  
 Gln Val Ile Ala Leu Leu Asp Ser Gly Ala Leu Arg Val Ala Glu Lys  
 35 40 45  
 Ile Asp Gly Gln Trp Val Thr His Gln Trp Leu Lys Lys Ala Val Leu  
 50 55 60  
 Leu Ser Phe Arg Ile Asn Asp Asn Gln Val Ile Glu Gly Ala Glu Ser  
 65 70 75 80  
 Arg Tyr Phe Asp Lys Val Pro Met Lys Phe Ala Asp Tyr Asp Glu Ala  
 85 90 95  
 Arg Phe Gln Lys Glu Gly Phe Arg Val Val Pro Pro Ala Ala Val Arg  
 100 105 110  
 Gln Gly Ala Phe Ile Ala Arg Asn Thr Val Leu Met Pro Ser Tyr Val  
 115 120 125  
 Asn Ile Gly Ala Tyr Val Asp Glu Gly Thr Met Val Asp Thr Trp Ala  
 130 135 140  
 Thr Val Gly Ser Cys Ala Gln Ile Gly Lys Asn Val His Leu Ser Gly  
 145 150 155 160  
 Gly Val Gly Ile Gly Gly Val Leu Glu Pro Leu Gln Ala Asn Pro Thr  
 165 170 175  
 Ile Ile Glu Asp Asn Cys Phe Ile Gly Ala Arg Ser Glu Val Val Glu  
 180 185 190  
 Gly Val Ile Val Glu Glu Gly Ser Val Ile Ser Met Gly Val Tyr Ile  
 195 200 205  
 Gly Gln Ser Thr Arg Ile Tyr Asp Arg Glu Thr Gly Glu Ile His Tyr  
 210 215 220  
 Gly Arg Val Pro Ala Gly Ser Val Val Val Ser Gly Asn Leu Pro Ser  
 225 230 235 240  
 Lys Asp Gly Lys Tyr Ser Leu Tyr Cys Ala Val Ile Val Lys Lys Val  
 245 250 255  
 Asp Ala Lys Thr Arg Gly Lys Val Gly Ile Asn Glu Leu Leu Arg Thr  
 260 265 270  
 Ile Asp

<210> 3  
 <211> 1221  
 5 <212> ADN  
 <213> *Escherichia coli*

<220>  
 <221> CDS  
 10 <222> (1)..(1221)

<400> 3



ES 2 403 199 T3

atg gca att gaa caa aca gca att aca cgc gcg act ttc gat gaa gtg	48
Met Ala Ile Glu Gln Thr Ala Ile Thr Arg Ala Thr Phe Asp Glu Val	
1 5 10 15	
atc ctg ccg att tat gct ccg gca gag ttt att ccg gta aaa ggt cag	96
Ile Leu Pro Ile Tyr Ala Pro Ala Glu Phe Ile Pro Val Lys Gly Gln	
20 25 30	
ggc agc cga atc tgg gat cag caa ggc aag gag tat gtc gat ttc gcg	144
Gly Ser Arg Ile Trp Asp Gln Gln Gly Lys Glu Tyr Val Asp Phe Ala	
35 40 45	
ggt ggc att gca gtt acg gcg ttg ggc cat tgc cat cct gcg ctg gtg	192
Gly Gly Ile Ala Val Thr Ala Leu Gly His Cys His Pro Ala Leu Val	
50 55 60	
aac gcg tta aaa acc cag ggc gaa act ctg tgg cat atc agt aac gtt	240
Asn Ala Leu Lys Thr Gln Gly Glu Thr Leu Trp His Ile Ser Asn Val	
65 70 75 80	
ttc acc aat gaa ccg gcg ctg cgt ctt ggg cgt aaa ctg att gag gca	288
Phe Thr Asn Glu Pro Ala Leu Arg Leu Gly Arg Lys Leu Ile Glu Ala	
85 90 95	
acg ttt gcc gaa cgc gtg gtg ttt atg aac tcc ggc acg gaa gct aac	336
Thr Phe Ala Glu Arg Val Val Phe Met Asn Ser Gly Thr Glu Ala Asn	
100 105 110	
gaa acc gcc ttt aaa ctg gca cgc cat tac gcc tgt gtg cgt cat agc	384
Glu Thr Ala Phe Lys Leu Ala Arg His Tyr Ala Cys Val Arg His Ser	
115 120 125	
ccg ttc aaa acc aaa att att gcc ttc cat aac gct ttt cat ggt cgc	432
Pro Phe Lys Thr Lys Ile Ile Ala Phe His Asn Ala Phe His Gly Arg	
130 135 140	
tcg ctg ttt acc gtt tcg gtg ggt ggg cag cca aaa tat tcc gac ggc	480
Ser Leu Phe Thr Val Ser Val Gly Gly Gln Pro Lys Tyr Ser Asp Gly	
145 150 155 160	
ttt ggg ccg aaa ccg gca gac atc atc cac gtt ccc ttt aac gat ctc	528
Phe Gly Pro Lys Pro Ala Asp Ile Ile His Val Pro Phe Asn Asp Leu	
165 170 175	
cat gca gtg aaa gcg gtg atg gat gat cac acc tgt gcg gtg gtg gtt	576
His Ala Val Lys Ala Val Met Asp Asp His Thr Cys Ala Val Val Val	
180 185 190	
gag ccg atc cag ggc gag ggc ggt gtg acg gca gcg acg cca gag ttt	624
Glu Pro Ile Gln Gly Glu Gly Gly Val Thr Ala Ala Thr Pro Glu Phe	
195 200 205	
ttg cag ggc ttg cgc gag ctg tgc gat caa cat cag gca tta ttg gtg	672
Leu Gln Gly Leu Arg Glu Leu Cys Asp Gln His Gln Ala Leu Leu Val	
210 215 220	
ttt gat gaa gtg cag tgc ggg atg ggg cgg acc ggc gat ttg ttt gct	720
Phe Asp Glu Val Gln Cys Gly Met Gly Arg Thr Gly Asp Leu Phe Ala	
225 230 235 240	
tac atg cac tac ggc gtt acg ccg gat att ctg acc tct gcg aaa gcg	768
Tyr Met His Tyr Gly Val Thr Pro Asp Ile Leu Thr Ser Ala Lys Ala	
245 250 255	

ES 2 403 199 T3

tta ggc ggc ggc ttc ccg att agc gcc atg ctg acc acg gcg gaa att	816
Leu Gly Gly Gly Phe Pro Ile Ser Ala Met Leu Thr Thr Ala Glu Ile	
260 265 270	
gct tct gcg ttt cat cct ggt tct cac ggt tcc acc tac ggc ggt aat	864
Ala Ser Ala Phe His Pro Gly Ser His Gly Ser Thr Tyr Gly Gly Asn	
275 280 285	
cct ctg gcc tgt gca gta gcg ggg gcg gcg ttt gat atc atc aat acc	912
Pro Leu Ala Cys Ala Val Ala Gly Ala Ala Phe Asp Ile Ile Asn Thr	
290 295 300	
cct gaa gtg ctg gaa ggc att cag gcg aaa cgc cag cgt ttt gtt gac	960
Pro Glu Val Leu Glu Gly Ile Gln Ala Lys Arg Gln Arg Phe Val Asp	
305 310 315 320	
cat ctg cag aag atc gat cag cag tac gat gta ttt agc gat att cgc	1008
His Leu Gln Lys Ile Asp Gln Gln Tyr Asp Val Phe Ser Asp Ile Arg	
325 330 335	
ggt atg ggg ctg ttg att ggc gca gag ctg aaa cca cag tac aaa ggt	1056
Gly Met Gly Leu Leu Ile Gly Ala Glu Leu Lys Pro Gln Tyr Lys Gly	
340 345 350	
cgg gcg cgt gat ttc ctg tat gcg ggc gca gag gct ggc gta atg gtg	1104
Arg Ala Arg Asp Phe Leu Tyr Ala Gly Ala Glu Ala Gly Val Met Val	
355 360 365	
ctg aat gcc gga ccg gat gtg atg cgt ttt gca ccg tcg ctg gtg gtg	1152
Leu Asn Ala Gly Pro Asp Val Met Arg Phe Ala Pro Ser Leu Val Val	
370 375 380	
gaa gat gcg gat atc gat gaa ggg atg caa cgt ttc gcc cac gcg gtg	1200
Glu Asp Ala Asp Ile Asp Glu Gly Met Gln Arg Phe Ala His Ala Val	
385 390 395 400	
gcg aag gtg gtt ggg gcg taa	1221
Ala Lys Val Val Gly Ala	
405	

<210> 4  
 <211> 406  
 5 <212> PRT  
 <213> *Escherichia coli*

<400> 4

Met Ala Ile Glu Gln Thr Ala Ile Thr Arg Ala Thr Phe Asp Glu Val	
1 5 10 15	
Ile Leu Pro Ile Tyr Ala Pro Ala Glu Phe Ile Pro Val Lys Gly Gln	
20 25 30	
Gly Ser Arg Ile Trp Asp Gln Gln Gly Lys Glu Tyr Val Asp Phe Ala	
35 40 45	
Gly Gly Ile Ala Val Thr Ala Leu Gly His Cys His Pro Ala Leu Val	
50 55 60	
Asn Ala Leu Lys Thr Gln Gly Glu Thr Leu Trp His Ile Ser Asn Val	
65 70 75 80	
Phe Thr Asn Glu Pro Ala Leu Arg Leu Gly Arg Lys Leu Ile Glu Ala	
85 90 95	
Thr Phe Ala Glu Arg Val Val Phe Met Asn Ser Gly Thr Glu Ala Asn	
100 105 110	
Glu Thr Ala Phe Lys Leu Ala Arg His Tyr Ala Cys Val Arg His Ser	
115 120 125	
Pro Phe Lys Thr Lys Ile Ile Ala Phe His Asn Ala Phe His Gly Arg	
130 135 140	
Ser Leu Phe Thr Val Ser Val Gly Gly Gln Pro Lys Tyr Ser Asp Gly	
145 150 155 160	
Phe Gly Pro Lys Pro Ala Asp Ile Ile His Val Pro Phe Asn Asp Leu	
165 170 175	
His Ala Val Lys Ala Val Met Asp Asp His Thr Cys Ala Val Val Val	

ES 2 403 199 T3

			180					185				190			
Glu	Pro	Ile	Gln	Gly	Glu	Gly	Gly	Val	Thr	Ala	Ala	Thr	Pro	Glu	Phe
		195					200					205			
Leu	Gln	Gly	Leu	Arg	Glu	Leu	Cys	Asp	Gln	His	Gln	Ala	Leu	Leu	Val
	210					215					220				
Phe	Asp	Glu	Val	Gln	Cys	Gly	Met	Gly	Arg	Thr	Gly	Asp	Leu	Phe	Ala
225					230					235					240
Tyr	Met	His	Tyr	Gly	Val	Thr	Pro	Asp	Ile	Leu	Thr	Ser	Ala	Lys	Ala
				245					250					255	
Leu	Gly	Gly	Gly	Phe	Pro	Ile	Ser	Ala	Met	Leu	Thr	Thr	Ala	Glu	Ile
			260					265					270		
Ala	Ser	Ala	Phe	His	Pro	Gly	Ser	His	Gly	Ser	Thr	Tyr	Gly	Gly	Asn
		275					280					285			
Pro	Leu	Ala	Cys	Ala	Val	Ala	Gly	Ala	Ala	Phe	Asp	Ile	Ile	Asn	Thr
	290					295					300				
Pro	Glu	Val	Leu	Glu	Gly	Ile	Gln	Ala	Lys	Arg	Gln	Arg	Phe	Val	Asp
305					310					315					320
His	Leu	Gln	Lys	Ile	Asp	Gln	Gln	Tyr	Asp	Val	Phe	Ser	Asp	Ile	Arg
				325					330					335	
Gly	Met	Gly	Leu	Leu	Ile	Gly	Ala	Glu	Leu	Lys	Pro	Gln	Tyr	Lys	Gly
			340					345					350		
Arg	Ala	Arg	Asp	Phe	Leu	Tyr	Ala	Gly	Ala	Glu	Ala	Gly	Val	Met	Val
		355					360					365			
Leu	Asn	Ala	Gly	Pro	Asp	Val	Met	Arg	Phe	Ala	Pro	Ser	Leu	Val	Val
	370					375					380				
Glu	Asp	Ala	Asp	Ile	Asp	Glu	Gly	Met	Gln	Arg	Phe	Ala	His	Ala	Val
385					390					395					400
Ala	Lys	Val	Val	Gly	Ala										
				405											

<210> 5  
 <211> 1128  
 5 <212> ADN  
 <213> *Escherichia coli*

<220>  
 <221> CDS  
 10 <222> (1)..(1128)

<400> 5

atg	tcg	tgc	ccg	gtt	att	gag	ctg	aca	caa	cag	ctt	att	cgc	cgc	cct	48
Met	Ser	Cys	Pro	Val	Ile	Glu	Leu	Thr	Gln	Gln	Leu	Ile	Arg	Arg	Pro	
1				5					10					15		
tcc	ctg	agt	cct	gat	gat	gca	gga	tgc	cag	gct	ttg	ttg	att	gaa	cgt	96
Ser	Leu	Ser	Pro	Asp	Asp	Ala	Gly	Cys	Gln	Ala	Leu	Leu	Ile	Glu	Arg	
			20					25					30			
ttg	cag	gcg	atc	ggc	ttt	acc	ggt	gaa	cgc	atg	gac	ttt	gcc	gat	acg	144
Leu	Gln	Ala	Ile	Gly	Phe	Thr	Val	Glu	Arg	Met	Asp	Phe	Ala	Asp	Thr	
		35				40					45					
cag	aat	ttt	tgg	gca	tgg	cgt	ggg	cag	ggt	gaa	acg	tta	gcc	ttt	gcc	192
Gln	Asn	Phe	Trp	Ala	Trp	Arg	Gly	Gln	Gly	Glu	Thr	Leu	Ala	Phe	Ala	
	50					55					60					
ggg	cat	acc	gac	gtg	gtg	ccg	cct	ggc	gac	gcc	gat	cgt	tgg	atc	aat	240
Gly	His	Thr	Asp	Val	Val	Pro	Pro	Gly	Asp	Ala	Asp	Arg	Trp	Ile	Asn	
65				70					75					80		
ccc	ccg	ttt	gaa	ccc	acc	att	cgt	gac	ggc	atg	tta	ttc	ggg	cgc	ggc	288
Pro	Pro	Phe	Glu	Pro	Thr	Ile	Arg	Asp	Gly	Met	Leu	Phe	Gly	Arg	Gly	
				85				90					95			
gcg	gca	gat	atg	aaa	ggc	tcg	ctg	gcg	gcg	atg	gtg	gtg	gcg	gca	gaa	336
Ala	Ala	Asp	Met	Lys	Gly	Ser	Leu	Ala	Ala	Met	Val	Val	Ala	Ala	Glu	

ES 2 403 199 T3

				100					105				110				
cgt	ttt	gtc	gca	caa	cat	ccc	aac	cat	acg	ggg	cga	ctg	gca	ttt	ctg		384
Arg	Phe	Val	Ala	Gln	His	Pro	Asn	His	Thr	Gly	Arg	Leu	Ala	Phe	Leu		
		115							120				125				
atc	acc	tct	gat	gaa	gaa	gcc	agt	gcc	cac	aac	ggt	acg	gta	aaa	gtc		432
Ile	Thr	Ser	Asp	Glu	Glu	Ala	Ser	Ala	His	Asn	Gly	Thr	Val	Lys	Val		
		130						135					140				
gtc	gaa	gcg	tta	atg	gca	cgt	aat	gag	cgt	ctc	gat	tac	tgc	ctg	gtt		480
Val	Glu	Ala	Leu	Met	Ala	Arg	Asn	Glu	Arg	Leu	Asp	Tyr	Cys	Leu	Val		
	145					150					155				160		
ggc	gaa	ccg	tcg	agt	atc	gaa	gtg	gta	ggt	gat	gtg	gtg	aaa	aat	ggt		528
Gly	Glu	Pro	Ser	Ser	Ile	Glu	Val	Val	Gly	Asp	Val	Val	Lys	Asn	Gly		
				165						170				175			
cgt	cgc	gga	tca	tta	acc	tgc	aac	ctt	acc	att	cat	ggc	gtt	cag	ggg		576
Arg	Arg	Gly	Ser	Leu	Thr	Cys	Asn	Leu	Thr	Ile	His	Gly	Val	Gln	Gly		
			180						185					190			
cat	gtt	gcc	tac	cca	cat	ctg	gct	gac	aat	ccg	gta	cat	cgc	gca	gca		624
His	Val	Ala	Tyr	Pro	His	Leu	Ala	Asp	Asn	Pro	Val	His	Arg	Ala	Ala		
		195						200						205			
cct	ttc	ctt	aat	gaa	tta	gtg	gct	att	gag	tgg	gat	cag	ggc	aat	gaa		672
Pro	Phe	Leu	Asn	Glu	Leu	Val	Ala	Ile	Glu	Trp	Asp	Gln	Gly	Asn	Glu		
		210					215					220					
ttc	ttc	ccg	gcg	acc	agt	atg	cag	att	gcc	aat	att	cag	gcg	gga	acg		720
Phe	Phe	Pro	Ala	Thr	Ser	Met	Gln	Ile	Ala	Asn	Ile	Gln	Ala	Gly	Thr		
		225				230						235			240		
ggc	agt	aac	aac	gtt	att	ccg	ggt	gaa	ctg	ttt	gtg	cag	ttt	aac	ttc		768
Gly	Ser	Asn	Asn	Val	Ile	Pro	Gly	Glu	Leu	Phe	Val	Gln	Phe	Asn	Phe		
				245						250				255			
cgc	ttc	agc	acc	gaa	ctg	act	gat	gag	atg	atc	aaa	gcg	cag	gtg	ctt		816
Arg	Phe	Ser	Thr	Glu	Leu	Thr	Asp	Glu	Met	Ile	Lys	Ala	Gln	Val	Leu		
			260						265					270			
gcc	ctg	ctt	gaa	aaa	cat	caa	ctg	cgc	tat	acg	gtg	gat	tgg	tgg	ctt		864
Ala	Leu	Leu	Glu	Lys	His	Gln	Leu	Arg	Tyr	Thr	Val	Asp	Trp	Trp	Leu		
		275					280						285				
tcc	ggg	cag	cca	ttt	ttg	acc	gcg	cgc	ggt	aaa	ctg	gtg	gat	gcg	gtc		912
Ser	Gly	Gln	Pro	Phe	Leu	Thr	Ala	Arg	Gly	Lys	Leu	Val	Asp	Ala	Val		
		290				295						300					
gtt	aac	gcg	gtt	gag	cac	tat	aat	gaa	att	aaa	ccg	cag	cta	ctg	acc		960
Val	Asn	Ala	Val	Glu	His	Tyr	Asn	Glu	Ile	Lys	Pro	Gln	Leu	Leu	Thr		
				310						315				320			
aca	ggc	gga	acg	tcc	gac	ggg	cgc	ttt	att	gcc	cgc	atg	ggg	gcg	cag		1008
Thr	Gly	Gly	Thr	Ser	Asp	Gly	Arg	Phe	Ile	Ala	Arg	Met	Gly	Ala	Gln		
				325					330					335			
gtg	gtg	gaa	ctc	ggg	ccg	gtc	aat	gcc	act	att	cat	aaa	att	aat	gaa		1056
Val	Val	Glu	Leu	Gly	Pro	Val	Asn	Ala	Thr	Ile	His	Lys	Ile	Asn	Glu		
		340							345					350			
tgt	gtg	aac	gct	gcc	gac	ctg	cag	cta	ctt	gcc	cgt	atg	tat	caa	cgt		1104
Cys	Val	Asn	Ala	Ala	Asp	Leu	Gln	Leu	Leu	Ala	Arg	Met	Tyr	Gln	Arg		
		355					360						365				
atc	atg	gaa	cag	ctc	gtc	gcc	tga										1128
Ile	Met	Glu	Gln	Leu	Val	Ala											
		370				375											

<210> 6  
 <211> 375  
 5 <212> PRT  
 <213> *Escherichia coli*

<400> 6

10 Met Ser Cys Pro Val Ile Glu Leu Thr Gln Gln Leu Ile Arg Arg Pro

ES 2 403 199 T3

```

1           5           10           15
Ser Leu Ser Pro Asp Asp Ala Gly Cys Gln Ala Leu Leu Ile Glu Arg
      20           25           30
Leu Gln Ala Ile Gly Phe Thr Val Glu Arg Met Asp Phe Ala Asp Thr
      35           40           45
Gln Asn Phe Trp Ala Trp Arg Gly Gln Gly Glu Thr Leu Ala Phe Ala
      50           55           60
Gly His Thr Asp Val Val Pro Pro Gly Asp Ala Asp Arg Trp Ile Asn
      65           70           75
Pro Pro Phe Glu Pro Thr Ile Arg Asp Gly Met Leu Phe Gly Arg Gly
      85           90           95
Ala Ala Asp Met Lys Gly Ser Leu Ala Ala Met Val Val Ala Ala Glu
      100          105          110
Arg Phe Val Ala Gln His Pro Asn His Thr Gly Arg Leu Ala Phe Leu
      115          120          125
Ile Thr Ser Asp Glu Glu Ala Ser Ala His Asn Gly Thr Val Lys Val
      130          135          140
Val Glu Ala Leu Met Ala Arg Asn Glu Arg Leu Asp Tyr Cys Leu Val
      145          150          155
Gly Glu Pro Ser Ser Ile Glu Val Val Gly Asp Val Val Lys Asn Gly
      165          170          175
Arg Arg Gly Ser Leu Thr Cys Asn Leu Thr Ile His Gly Val Gln Gly
      180          185          190
His Val Ala Tyr Pro His Leu Ala Asp Asn Pro Val His Arg Ala Ala
      195          200          205
Pro Phe Leu Asn Glu Leu Val Ala Ile Glu Trp Asp Gln Gly Asn Glu
      210          215          220
Phe Phe Pro Ala Thr Ser Met Gln Ile Ala Asn Ile Gln Ala Gly Thr
      225          230          235
Gly Ser Asn Asn Val Ile Pro Gly Glu Leu Phe Val Gln Phe Asn Phe
      245          250          255
Arg Phe Ser Thr Glu Leu Thr Asp Glu Met Ile Lys Ala Gln Val Leu
      260          265          270
Ala Leu Leu Glu Lys His Gln Leu Arg Tyr Thr Val Asp Trp Trp Leu
      275          280          285
Ser Gly Gln Pro Phe Leu Thr Ala Arg Gly Lys Leu Val Asp Ala Val
      290          295          300
Val Asn Ala Val Glu His Tyr Asn Glu Ile Lys Pro Gln Leu Leu Thr
      305          310          315
Thr Gly Gly Thr Ser Asp Gly Arg Phe Ile Ala Arg Met Gly Ala Gln
      325          330          335
Val Val Glu Leu Gly Pro Val Asn Ala Thr Ile His Lys Ile Asn Glu
      340          345          350
Cys Val Asn Ala Ala Asp Leu Gln Leu Leu Ala Arg Met Tyr Gln Arg
      355          360          365
Ile Met Glu Gln Leu Val Ala
      370          375

```

<210> 7  
 <211> 825  
 5 <212> ADN  
 <213> *Escherichia coli*

<220>  
 <221> CDS  
 10 <222> (1)..(825)

<400> 7

```

atg cag ttc tcg aaa atg cat ggc ctt ggc aac gat ttt atg gtc gtc
Met Gln Phe Ser Lys Met His Gly Leu Gly Asn Asp Phe Met Val Val

```

48

ES 2 403 199 T3

```

1           5           10           15
gac gcg gta acg cag aat gtc ttt ttt tca ccg gag ctg att cgt cgc           96
Asp Ala Val Thr Gln Asn Val Phe Phe Ser Pro Glu Leu Ile Arg Arg
      20           25           30
ctg gct gat cgg cac ctg ggg gta ggg ttt gac caa ctg ctg gtg gtt           144
Leu Ala Asp Arg His Leu Gly Val Gly Phe Asp Gln Leu Leu Val Val
      35           40           45
gag ccg ccg tat gat cct gaa ctg gat ttt cac tat cgc att ttc aat           192
Glu Pro Pro Tyr Asp Pro Glu Leu Asp Phe His Tyr Arg Ile Phe Asn
      50           55           60
gct gat ggc agt gaa gtg gcg cag tgc ggc aac ggt gcg cgc tgc ttt           240
Ala Asp Gly Ser Glu Val Ala Gln Cys Gly Asn Gly Ala Arg Cys Phe
      65           70           75           80
gcc cgt ttt gtg cgt ctg aaa gga ctg acc aat aag cgt gat atc cgc           288
Ala Arg Phe Val Arg Leu Lys Gly Leu Thr Asn Lys Arg Asp Ile Arg
      85           90           95
gtc agc acc gcc aac ggg cgg atg gtt ctg acc gtc acc gat gat gat           336
Val Ser Thr Ala Asn Gly Arg Met Val Leu Thr Val Thr Asp Asp Asp
      100           105           110
ctg gtc cgc gta aat atg ggc gaa ccc aac ttc gaa cct tcc gcc gtg           384
Leu Val Arg Val Asn Met Gly Glu Pro Asn Phe Glu Pro Ser Ala Val
      115           120           125
ccg ttt cgc gct aac aaa gcg gaa aag acc tat att atg cgc gcc gcc           432
Pro Phe Arg Ala Asn Lys Ala Glu Lys Thr Tyr Ile Met Arg Ala Ala
      130           135           140
gag cag aca atc tta tgc ggc gtg gtg tgc atg gga aat ccg cat tgc           480
Glu Gln Thr Ile Leu Cys Gly Val Val Ser Met Gly Asn Pro His Cys
      145           150           155           160
gtg att cag gtc gat gat gtc gat acc gcg gcg gta gaa acg ctt ggt           528
Val Ile Gln Val Asp Asp Val Asp Thr Ala Ala Val Glu Thr Leu Gly
      165           170           175
cct gtt ctg gaa agc cac gag cgt ttt ccg gag cgc gcc aat atc ggt           576
Pro Val Leu Glu Ser His Glu Arg Phe Pro Glu Arg Ala Asn Ile Gly
      180           185           190
ttt atg caa gtg gtt aag cgc gag cat att cgt tta cgc gtt tat gag           624
Phe Met Gln Val Val Lys Arg Glu His Ile Arg Leu Arg Val Tyr Glu
      195           200           205
cgt ggg gca gga gaa acc cag gcc tgc ggc agc ggc gcg tgt gcg gcg           672
Arg Gly Ala Gly Glu Thr Gln Ala Cys Gly Ser Gly Ala Cys Ala Ala
      210           215           220
gtt gca gta ggg att cag caa ggt ttg ctg gcc gaa gaa gta cgc gtg           720
Val Ala Val Gly Ile Gln Gln Gly Leu Leu Ala Glu Glu Val Arg Val
      225           230           235           240
gaa ctc ccc ggc ggt cgt ctt gat atc gcc tgg aaa ggt ccg ggt cac           768
Glu Leu Pro Gly Gly Arg Leu Asp Ile Ala Trp Lys Gly Pro Gly His
      245           250           255
ccg tta tat atg act ggc ccg gcg gta cat gtc tac gac gga ttt att           816
Pro Leu Tyr Met Thr Gly Pro Ala Val His Val Tyr Asp Gly Phe Ile
      260           265           270
cat cta tga
His Leu

```

<210> 8  
 <211> 274  
 <212> PRT  
 <213> *Escherichia coli*

5

<400> 8

Met Gln Phe Ser Lys Met His Gly Leu Gly Asn Asp Phe Met Val Val

ES 2 403 199 T3

```

1           5           10           15
Asp Ala Val Thr Gln Asn Val Phe Phe Ser Pro Glu Leu Ile Arg Arg
      20           25           30
Leu Ala Asp Arg His Leu Gly Val Gly Phe Asp Gln Leu Leu Val Val
      35           40           45
Glu Pro Pro Tyr Asp Pro Glu Leu Asp Phe His Tyr Arg Ile Phe Asn
      50           55           60
Ala Asp Gly Ser Glu Val Ala Gln Cys Gly Asn Gly Ala Arg Cys Phe
      65           70           75           80
Ala Arg Phe Val Arg Leu Lys Gly Leu Thr Asn Lys Arg Asp Ile Arg
      85           90           95
Val Ser Thr Ala Asn Gly Arg Met Val Leu Thr Val Thr Asp Asp Asp
      100          105          110
Leu Val Arg Val Asn Met Gly Glu Pro Asn Phe Glu Pro Ser Ala Val
      115          120          125
Pro Phe Arg Ala Asn Lys Ala Glu Lys Thr Tyr Ile Met Arg Ala Ala
      130          135          140
Glu Gln Thr Ile Leu Cys Gly Val Val Ser Met Gly Asn Pro His Cys
      145          150          155          160
Val Ile Gln Val Asp Asp Val Asp Thr Ala Ala Val Glu Thr Leu Gly
      165          170          175
Pro Val Leu Glu Ser His Glu Arg Phe Pro Glu Arg Ala Asn Ile Gly
      180          185          190
Phe Met Gln Val Val Lys Arg Glu His Ile Arg Leu Arg Val Tyr Glu
      195          200          205
Arg Gly Ala Gly Glu Thr Gln Ala Cys Gly Ser Gly Ala Cys Ala Ala
      210          215          220
Val Ala Val Gly Ile Gln Gln Gly Leu Leu Ala Glu Glu Val Arg Val
      225          230          235          240
Glu Leu Pro Gly Gly Arg Leu Asp Ile Ala Trp Lys Gly Pro Gly His
      245          250          255
Pro Leu Tyr Met Thr Gly Pro Ala Val His Val Tyr Asp Gly Phe Ile
      260          265          270
His Leu

```

<210> 9

<211> 963

5 <212> ADN

<213> *Corynebacterium glutamicum*

<220>

<221> CDS

10 <222> (1)..(963)

<400> 9

```

atg acc aac atc cgc gta gct atc gtg ggc tac gga aac ctg gga cgc      48
Met Thr Asn Ile Arg Val Ala Ile Val Gly Tyr Gly Asn Leu Gly Arg
1           5           10           15
agc gtc gaa aag ctt att gcc aag cag ccc gac atg gac ctt gta gga      96
Ser Val Glu Lys Leu Ile Ala Lys Gln Pro Asp Met Asp Leu Val Gly
      20           25           30
atc ttc tcg cgc cgg gcc acc ctc gac aca aag acg cca gtc ttt gat      144
Ile Phe Ser Arg Arg Ala Thr Leu Asp Thr Lys Thr Pro Val Phe Asp
      35           40           45
gtc gcc gac gtg gac aag cac gcc gac gac gtg gac gtg ctg ttc ctg      192
Val Ala Asp Val Asp Lys His Ala Asp Asp Val Asp Val Leu Phe Leu
      50           55           60
tgc atg ggc tcc gcc acc gac atc cct gag cag gca cca aag ttc gcg      240
Cys Met Gly Ser Ala Thr Asp Ile Pro Glu Gln Ala Pro Lys Phe Ala

```

ES 2 403 199 T3

65					70					75				80	
cag ttc gcc tgc acc gta gac acc tac gac aac cac cgc gac atc cca															288
Gln Phe Ala Cys Thr Val Asp Thr Tyr Asp Asn His Arg Asp Ile Pro															
				85					90				95		
cgc cac cgc cag gtc atg aac gaa gcc gcc acc gca gcc ggc aac gtt															336
Arg His Arg Gln Val Met Asn Glu Ala Ala Thr Ala Ala Gly Asn Val															
			100					105					110		
gca ctg gtc tct acc ggc tgg gat cca gga atg ttc tcc atc aac cgc															384
Ala Leu Val Ser Thr Gly Trp Asp Pro Gly Met Phe Ser Ile Asn Arg															
			115					120					125		
gtc tac gca gcg gca gtc tta gcc gag cac cag cag cac acc ttc tgg															432
Val Tyr Ala Ala Ala Val Leu Ala Glu His Gln Gln His Thr Phe Trp															
			130					135					140		
ggc cca ggt ttg tca cag ggc cac tcc gat gct ttg cga cgc atc cct															480
Gly Pro Gly Leu Ser Gln Gly His Ser Asp Ala Leu Arg Arg Ile Pro															
			145					150					155		160
ggc gtt caa aag gca gtc cag tac acc ctc cca tcc gaa gac gcc ctg															528
Gly Val Gln Lys Ala Val Gln Tyr Thr Leu Pro Ser Glu Asp Ala Leu															
			165					170					175		
gaa aag gcc cgc cgc ggc gaa gcc ggc gac ctt acc gga aag caa acc															576
Glu Lys Ala Arg Arg Gly Glu Ala Gly Asp Leu Thr Gly Lys Gln Thr															
			180					185					190		
cac aag cgc caa tgc ttc gtg gtt gcc gac gcg gcc gat cac gag cgc															624
His Lys Arg Gln Cys Phe Val Ala Asp Ala Ala Asp His Glu Arg															
			195					200					205		
atc gaa aac gac atc cgc acc atg cct gat tac ttc gtt ggc tac gaa															672
Ile Glu Asn Asp Ile Arg Thr Met Pro Asp Tyr Phe Val Gly Tyr Glu															
			210					215					220		
gtc gaa gtc aac ttc atc gac gaa gca acc ttc gac tcc gag cac acc															720
Val Glu Val Asn Phe Ile Asp Glu Ala Thr Phe Asp Ser Glu His Thr															
			225					230					235		240
ggc atg cca cac ggt ggc cac gtg att acc acc ggc gac acc ggt ggc															768
Gly Met Pro His Gly Gly His Val Ile Thr Thr Gly Asp Thr Gly Gly															
			245					250					255		
ttc aac cac acc gtg gaa tac atc ctc aag ctg gac cga aac cca gat															816
Phe Asn His Thr Val Glu Tyr Ile Leu Lys Leu Asp Arg Asn Pro Asp															
			260					265					270		
ttc acc gct tcc tca cag atc gct ttc ggt cgc gca gct cac cgc atg															864
Phe Thr Ala Ser Ser Gln Ile Ala Phe Gly Arg Ala Ala His Arg Met															
			275					280					285		
aag cag cag ggc caa agc gga gct ttc acc gtc ctc gaa gtt gct cca															912
Lys Gln Gln Gly Gln Ser Gly Ala Phe Thr Val Leu Glu Val Ala Pro															
			290					295					300		
tac ctg ctc tcc cca gag aac ttg gac gat ctg atc gca cgc gac gtc															960
Tyr Leu Leu Ser Pro Glu Asn Leu Asp Asp Leu Ile Ala Arg Asp Val															
			305					310					315		320
taa															963

<210> 10

<211> 320

5 <212> PRT

<213> *Corynebacterium glutamicum*

<400> 10

Met Thr Asn Ile Arg Val Ala Ile Val Gly Tyr Gly Asn Leu Gly Arg															
1			5					10					15		
Ser Val Glu Lys Leu Ile Ala Lys Gln Pro Asp Met Asp Leu Val Gly															
			20					25					30		
Ile Phe Ser Arg Arg Ala Thr Leu Asp Thr Lys Thr Pro Val Phe Asp															
			35					40					45		

10



Val Ala Asp Val Asp Lys His Ala Asp Asp Val Asp Val Leu Phe Leu  
 50 55 60  
 Cys Met Gly Ser Ala Thr Asp Ile Pro Glu Gln Ala Pro Lys Phe Ala  
 65 70 75 80  
 Gln Phe Ala Cys Thr Val Asp Thr Tyr Asp Asn His Arg Asp Ile Pro  
 85 90 95  
 Arg His Arg Gln Val Met Asn Glu Ala Ala Thr Ala Ala Gly Asn Val  
 100 105 110  
 Ala Leu Val Ser Thr Gly Trp Asp Pro Gly Met Phe Ser Ile Asn Arg  
 115 120 125  
 Val Tyr Ala Ala Ala Val Leu Ala Glu His Gln Gln His Thr Phe Trp  
 130 135 140  
 Gly Pro Gly Leu Ser Gln Gly His Ser Asp Ala Leu Arg Arg Ile Pro  
 145 150 155 160  
 Gly Val Gln Lys Ala Val Gln Tyr Thr Leu Pro Ser Glu Asp Ala Leu  
 165 170 175  
 Glu Lys Ala Arg Arg Gly Glu Ala Gly Asp Leu Thr Gly Lys Gln Thr  
 180 185 190  
 His Lys Arg Gln Cys Phe Val Val Ala Asp Ala Ala Asp His Glu Arg  
 195 200 205  
 Ile Glu Asn Asp Ile Arg Thr Met Pro Asp Tyr Phe Val Gly Tyr Glu  
 210 215 220  
 Val Glu Val Asn Phe Ile Asp Glu Ala Thr Phe Asp Ser Glu His Thr  
 225 230 235 240  
 Gly Met Pro His Gly Gly His Val Ile Thr Thr Gly Asp Thr Gly Gly  
 245 250 255  
 Phe Asn His Thr Val Glu Tyr Ile Leu Lys Leu Asp Arg Asn Pro Asp  
 260 265 270  
 Phe Thr Ala Ser Ser Gln Ile Ala Phe Gly Arg Ala Ala His Arg Met  
 275 280 285  
 Lys Gln Gln Gly Gln Ser Gly Ala Phe Thr Val Leu Glu Val Ala Pro  
 290 295 300  
 Tyr Leu Leu Ser Pro Glu Asn Leu Asp Asp Leu Ile Ala Arg Asp Val  
 305 310 315 320

<210> 11  
 <211> 963  
 5 <212> ADN  
 <213> *Brevibacterium lactofermentum*

<220>  
 <221> CDS  
 10 <222> (1)..(963)

<400> 11

atg acc aac atc cgc gta gct atc gta ggc tac gga aac ctg gga cgc 48  
 Met Thr Asn Ile Arg Val Ala Ile Val Gly Tyr Gly Asn Leu Gly Arg  
 1 5 10 15  
 agc gtc gaa aag ctt att gcc aag cag ccc gac atg gac ctt gta gga 96  
 Ser Val Glu Lys Leu Ile Ala Lys Gln Pro Asp Met Asp Leu Val Gly  
 20 25 30  
 atc ttc tcg cgc cgg gcc acc ctc gac aca aag acg cca gtc ttt gat 144  
 Ile Phe Ser Arg Arg Ala Thr Leu Asp Thr Lys Thr Pro Val Phe Asp  
 35 40 45  
 gtc gcc gac gtg gac aag cac gcc gac gac gtg gac gtg ctg ttc ctg 192  
 Val Ala Asp Val Asp Lys His Ala Asp Asp Val Asp Val Leu Phe Leu  
 50 55 60  
 tgc atg ggc tcc gcc acc gac atc cct gag cag gca cca aag ttc gcg 240  
 Cys Met Gly Ser Ala Thr Asp Ile Pro Glu Gln Ala Pro Lys Phe Ala  
 65 70 75 80



	50					55					60					
	Cys	Met	Gly	Ser	Ala	Thr	Asp	Ile	Pro	Glu	Gln	Ala	Pro	Lys	Phe	Ala
	65					70					75					80
	Gln	Phe	Ala	Cys	Thr	Val	Asp	Thr	Tyr	Asp	Asn	His	Arg	Asp	Ile	Pro
					85					90					95	
	Arg	His	Arg	Gln	Val	Met	Asn	Glu	Ala	Ala	Thr	Ala	Ala	Gly	Asn	Val
				100					105					110		
	Ala	Leu	Val	Ser	Thr	Gly	Trp	Asp	Pro	Gly	Met	Phe	Ser	Ile	Asn	Arg
			115					120					125			
	Val	Tyr	Ala	Ala	Ala	Val	Leu	Ala	Glu	His	Gln	Gln	His	Thr	Phe	Trp
		130					135					140				
	Gly	Pro	Gly	Leu	Ser	Gln	Gly	His	Ser	Asp	Ala	Leu	Arg	Arg	Ile	Pro
	145					150					155					160
	Gly	Val	Gln	Lys	Ala	Val	Gln	Tyr	Thr	Leu	Pro	Ser	Glu	Glu	Ala	Leu
				165						170					175	
	Glu	Lys	Ala	Arg	Arg	Gly	Glu	Ala	Gly	Asp	Leu	Thr	Gly	Lys	Gln	Thr
				180					185					190		
	His	Lys	Arg	Gln	Cys	Phe	Val	Val	Ala	Asp	Ala	Ala	Asp	His	Glu	Arg
			195				200						205			
	Ile	Glu	Asn	Asp	Ile	Arg	Thr	Met	Pro	Asp	Tyr	Phe	Val	Gly	Tyr	Glu
		210					215						220			
	Val	Glu	Val	Asn	Phe	Ile	Asp	Glu	Ala	Thr	Phe	Asp	Ala	Glu	His	Thr
	225					230						235				240
	Gly	Met	Pro	His	Gly	Gly	His	Val	Ile	Thr	Thr	Gly	Asp	Thr	Gly	Gly
				245						250					255	
	Phe	Asn	His	Thr	Val	Glu	Tyr	Ile	Leu	Lys	Leu	Asp	Arg	Asn	Pro	Asp
				260					265					270		
	Phe	Thr	Ala	Ser	Ser	Gln	Ile	Ala	Phe	Gly	Arg	Ala	Ala	His	Arg	Met
			275					280						285		
	Lys	Gln	Gln	Gly	Gln	Ser	Gly	Ala	Phe	Thr	Val	Leu	Glu	Val	Ala	Pro
		290					295					300				
	Tyr	Leu	Leu	Ser	Pro	Glu	Asn	Leu	Asp	Asp	Leu	Ile	Ala	Arg	Asp	Val
	305					310					315					320

<210> 13  
 <211> 963  
 5 <212> ADN  
 <213> *Corynebacterium efficiens*

<220>  
 <221> CDS  
 10 <222> (1)..(963)

<400> 13

atg	tcg	aag	atc	cgc	gca	gca	atc	ggt	ggt	tat	gga	aat	ctg	ggg	aag	48
Met	Ser	Lys	Ile	Arg	Ala	Ala	Ile	Val	Gly	Tyr	Gly	Asn	Leu	Gly	Lys	
1				5					10					15		
agc	gtc	gag	aag	ctc	atc	gtc	cag	caa	ccg	gac	atg	gaa	ctg	gtg	ggg	96
Ser	Val	Glu	Lys	Leu	Ile	Val	Gln	Gln	Pro	Asp	Met	Glu	Leu	Val	Gly	
			20				25						30			
atc	ttc	tcc	cgc	cgc	gac	acc	ctg	gac	acc	gac	acc	ccc	gtg	ttc	aac	144
Ile	Phe	Ser	Arg	Arg	Asp	Thr	Leu	Asp	Thr	Asp	Thr	Pro	Val	Phe	Asn	
		35					40					45				
gtc	gcc	gag	acg	gag	aag	cac	acc	ggc	gat	ggt	gat	ctc	ctc	ttc	ctc	192
Val	Ala	Glu	Thr	Glu	Lys	His	Thr	Gly	Asp	Val	Asp	Leu	Leu	Phe	Leu	
	50					55						60				
tgc	atg	ggt	tcc	gcc	act	gac	atc	ccg	gag	cag	gcc	ccg	ggt	ttt	gcg	240
Cys	Met	Gly	Ser	Ala	Thr	Asp	Ile	Pro	Glu	Gln	Ala	Pro	Gly	Phe	Ala	
65				70						75				80		
gca	ttc	gcc	tgc	acc	gtg	gac	acc	tat	gac	aac	cac	cgg	gac	atc	ccg	288

ES 2 403 199 T3

Ala	Phe	Ala	Cys	Thr	Val	Asp	Thr	Tyr	Asp	Asn	His	Arg	Asp	Ile	Pro	
				85					90					95		
cgt	cac	cgt	cag	gtg	atg	gat	gag	gcc	gcc	cgt	gcc	gcc	ggc	aat	gtc	336
Arg	His	Arg	Gln	Val	Met	Asp	Glu	Ala	Ala	Arg	Ala	Ala	Gly	Asn	Val	
			100					105					110			
tct	gtt	gtc	gcc	acc	ggg	tgg	gat	ccg	ggg	atg	ttc	tcc	atc	aac	cgc	384
Ser	Val	Val	Ala	Thr	Gly	Trp	Asp	Pro	Gly	Met	Phe	Ser	Ile	Asn	Arg	
		115						120					125			
gtg	tac	ggc	gca	gcc	ctg	ctc	gcc	gat	cac	cag	cag	cac	acc	ttc	tgg	432
Val	Tyr	Gly	Ala	Ala	Leu	Leu	Ala	Asp	His	Gln	Gln	His	Thr	Phe	Trp	
	130						135					140				
gga	ccg	ggg	ctg	tcc	cag	ggc	cac	tcc	gat	gcc	ttg	cga	cgc	atc	gac	480
Gly	Pro	Gly	Leu	Ser	Gln	Gly	His	Ser	Asp	Ala	Leu	Arg	Arg	Ile	Asp	
145					150					155					160	
ggc	gtc	gag	aag	gcc	gtc	cag	tac	acc	ctg	cct	tcc	gag	gat	gcc	ctg	528
Gly	Val	Glu	Lys	Ala	Val	Gln	Tyr	Thr	Leu	Pro	Ser	Glu	Asp	Ala	Leu	
				165					170					175		
gag	aag	gca	cgc	cgc	ggg	gag	gct	gag	ggg	ctg	acc	ggg	aaa	cag	acc	576
Glu	Lys	Ala	Arg	Arg	Gly	Glu	Ala	Glu	Gly	Leu	Thr	Gly	Lys	Gln	Thr	
		180						185						190		
cac	aag	cgt	cag	tgt	ttc	gtg	gtc	gcc	ccg	gag	tcc	gag	cac	gag	cgc	624
His	Lys	Arg	Gln	Cys	Phe	Val	Val	Ala	Pro	Glu	Ser	Glu	His	Glu	Arg	
		195						200						205		
atc	gag	aat	gag	atc	cgc	acc	atg	gct	gac	tac	ttc	gtc	ggc	tat	gag	672
Ile	Glu	Asn	Glu	Ile	Arg	Thr	Met	Ala	Asp	Tyr	Phe	Val	Gly	Tyr	Glu	
	210						215					220				
gtg	gag	gtc	aac	ttc	atc	gat	gag	gct	acc	ttc	gat	tcc	gag	cac	acc	720
Val	Glu	Val	Asn	Phe	Ile	Asp	Glu	Ala	Thr	Phe	Asp	Ser	Glu	His	Thr	
					230						235				240	
gga	atg	ccc	cac	ggc	ggg	cat	gtg	atc	acc	acc	ggg	gac	acc	ggc	ggg	768
Gly	Met	Pro	His	Gly	Gly	His	Val	Ile	Thr	Thr	Gly	Asp	Thr	Gly	Gly	
				245						250				255		
ttc	cac	cac	act	gtg	gag	tac	acc	ctg	aag	ctg	gat	cgc	aac	cct	gac	816
Phe	His	His	Thr	Val	Glu	Tyr	Thr	Leu	Lys	Leu	Asp	Arg	Asn	Pro	Asp	
			260					265						270		
ttc	acc	gcc	tcc	tcc	cag	att	gcg	ttc	gga	cgt	gct	gcc	tac	cga	ctg	864
Phe	Thr	Ala	Ser	Ser	Gln	Ile	Ala	Phe	Gly	Arg	Ala	Ala	Tyr	Arg	Leu	
		275					280						285			
aag	gaa	gca	ggg	cag	gct	ggg	gca	ttc	acc	ggt	ctg	gag	gtc	gct	ccc	912
Lys	Glu	Ala	Gly	Gln	Ala	Gly	Ala	Phe	Thr	Val	Leu	Glu	Val	Ala	Pro	
	290					295					300					
tac	ctc	ctg	tcc	ccg	aca	cca	ctc	gat	gac	ctg	atc	gcc	cgc	gac	gtc	960
Tyr	Leu	Leu	Ser	Pro	Thr	Pro	Leu	Asp	Asp	Leu	Ile	Ala	Arg	Asp	Val	
					310					315					320	
tag																963

<210> 14  
 <211> 320  
 5 <212> PRT  
 <213> *Corynebacterium efficiens*

<400> 14

Met	Ser	Lys	Ile	Arg	Ala	Ala	Ile	Val	Gly	Tyr	Gly	Asn	Leu	Gly	Lys	
1				5					10					15		
Ser	Val	Glu	Lys	Leu	Ile	Val	Gln	Gln	Pro	Asp	Met	Glu	Leu	Val	Gly	
			20					25					30			
Ile	Phe	Ser	Arg	Arg	Asp	Thr	Leu	Asp	Thr	Asp	Thr	Pro	Val	Phe	Asn	
		35					40					45				
Val	Ala	Glu	Thr	Glu	Lys	His	Thr	Gly	Asp	Val	Asp	Leu	Leu	Phe	Leu	
	50					55					60					

10

ES 2 403 199 T3

Cys Met Gly Ser Ala Thr Asp Ile Pro Glu Gln Ala Pro Gly Phe Ala  
 65 70 75 80  
 Ala Phe Ala Cys Thr Val Asp Thr Tyr Asp Asn His Arg Asp Ile Pro  
 85 90 95  
 Arg His Arg Gln Val Met Asp Glu Ala Ala Arg Ala Ala Gly Asn Val  
 100 105 110  
 Ser Val Val Ala Thr Gly Trp Asp Pro Gly Met Phe Ser Ile Asn Arg  
 115 120 125  
 Val Tyr Gly Ala Ala Leu Leu Ala Asp His Gln Gln His Thr Phe Trp  
 130 135 140  
 Gly Pro Gly Leu Ser Gln Gly His Ser Asp Ala Leu Arg Arg Ile Asp  
 145 150 155 160  
 Gly Val Glu Lys Ala Val Gln Tyr Thr Leu Pro Ser Glu Asp Ala Leu  
 165 170 175  
 Glu Lys Ala Arg Arg Gly Glu Ala Glu Gly Leu Thr Gly Lys Gln Thr  
 180 185 190  
 His Lys Arg Gln Cys Phe Val Val Ala Pro Glu Ser Glu His Glu Arg  
 195 200 205  
 Ile Glu Asn Glu Ile Arg Thr Met Ala Asp Tyr Phe Val Gly Tyr Glu  
 210 215 220  
 Val Glu Val Asn Phe Ile Asp Glu Ala Thr Phe Asp Ser Glu His Thr  
 225 230 235 240  
 Gly Met Pro His Gly Gly His Val Ile Thr Thr Gly Asp Thr Gly Gly  
 245 250 255  
 Phe His His Thr Val Glu Tyr Thr Leu Lys Leu Asp Arg Asn Pro Asp  
 260 265 270  
 Phe Thr Ala Ser Ser Gln Ile Ala Phe Gly Arg Ala Ala Tyr Arg Leu  
 275 280 285  
 Lys Glu Ala Gly Gln Ala Gly Ala Phe Thr Val Leu Glu Val Ala Pro  
 290 295 300  
 Tyr Leu Leu Ser Pro Thr Pro Leu Asp Asp Leu Ile Ala Arg Asp Val  
 305 310 315 320

<210> 15  
 <211> 987  
 5 <212> ADN  
 <213> *Herminiimonas arsenicoxydans*

<220>  
 <221> CDS  
 10 <222> (1)..(987)

<400> 15  
 atg gat gaa aaa atc cgt att ggg gtt gcc gga tac ggt aac ctc ggt 48  
 Met Asp Glu Lys Ile Arg Ile Gly Val Ala Gly Tyr Gly Asn Leu Gly  
 1 5 10 15  
 cgc ggg gtc gaa atg gcg atc gca cgc aac cct gat atg caa ctg gtc 96  
 Arg Gly Val Glu Met Ala Ile Ala Arg Asn Pro Asp Met Gln Leu Val  
 20 25 30  
 ggc gtt ttc agc cgg cgc gat ccc gct agc atc gag ctt ttg acc cag 144  
 Gly Val Phe Ser Arg Arg Asp Pro Ala Ser Ile Glu Leu Leu Thr Gln  
 35 40 45  
 gct gta ccc gtc cat aaa ttc gac gat atc gaa cag ttt cgc gac cag 192  
 Ala Val Pro Val His Lys Phe Asp Asp Ile Glu Gln Phe Arg Asp Gln  
 50 55 60  
 att gac gtg ctc att ctg tgc ggc gga tca aag aac gat ctg ccc gaa 240  
 Ile Asp Val Leu Ile Leu Cys Gly Gly Ser Lys Asn Asp Leu Pro Glu  
 65 70 75 80  
 caa ggc cca gca ttg gct acc ttg ttc aac acg att gac agc ttt gat 288  
 Gln Gly Pro Ala Leu Ala Thr Leu Phe Asn Thr Ile Asp Ser Phe Asp

ES 2 403 199 T3

				85					90				95						
aca	cac	aat	aaa	att	ccg	gaa	tat	ttc	gcc	gcg	atg	gac	agt	gcc	tca				336
Thr	His	Asn	Lys	Ile	Pro	Glu	Tyr	Phe	Ala	Ala	Met	Asp	Ser	Ala	Ser				
			100					105					110						
tgc	ggt	ggc	gag	cgc	aca	tcc	atc	atc	tcg	gta	ggc	tggt	gat	ccg	ggc				384
Cys	Val	Gly	Glu	Arg	Thr	Ser	Ile	Ile	Ser	Val	Gly	Trp	Asp	Pro	Gly				
		115					120					125							
gtg	ttc	tcg	ctc	aat	cgt	ctg	ttc	ggc	gaa	gcc	att	ttg	cca	gaa	ggc				432
Val	Phe	Ser	Leu	Asn	Arg	Leu	Phe	Gly	Glu	Ala	Ile	Leu	Pro	Glu	Gly				
	130					135				140									
gaa	act	tat	acc	ttc	tgg	ggt	aaa	ggt	cta	agt	cag	ggg	cac	tcg	gac				480
Glu	Thr	Tyr	Thr	Phe	Trp	Gly	Lys	Gly	Leu	Ser	Gln	Gly	His	Ser	Asp				
	145				150				155			160							
gca	ata	cgc	cgc	gtg	ccg	ggc	gtc	aag	gcc	ggt	gtg	cag	tac	acg	atc				528
Ala	Ile	Arg	Arg	Val	Pro	Gly	Val	Lys	Ala	Gly	Val	Gln	Tyr	Thr	Ile				
			165				170					175							
ccc	tcg	gct	gag	gcg	atg	gaa	ctg	gta	cgc	agt	ggc	agc	cag	ccc	caa				576
Pro	Ser	Ala	Glu	Ala	Met	Glu	Leu	Val	Arg	Ser	Gly	Ser	Gln	Pro	Gln				
		180					185					190							
ctg	tcg	aca	cgc	gag	aaa	cat	act	cgc	gag	tgt	cat	ggt	gta	ctc	gaa				624
Leu	Ser	Thr	Arg	Glu	Lys	His	Thr	Arg	Glu	Cys	His	Val	Val	Leu	Glu				
		195				200					205								
gcc	ggt	gcc	gat	gcc	aaa	gcg	gtg	gaa	cat	gcc	atc	gtc	agc	atg	ccg				672
Ala	Gly	Ala	Asp	Ala	Lys	Ala	Val	Glu	His	Ala	Ile	Val	Ser	Met	Pro				
	210					215				220									
gac	tat	ttc	gcc	gac	tat	gac	acg	acc	gtg	cac	ttc	atc	agc	gag	gaa				720
Asp	Tyr	Phe	Ala	Asp	Tyr	Asp	Thr	Thr	Val	His	Phe	Ile	Ser	Glu	Glu				
	225			230				235						240					
gag	tta	cga	agc	aat	cac	tct	gcc	atg	cca	cac	gga	ggc	ttc	ggt	atc				768
Glu	Leu	Arg	Ser	Asn	His	Ser	Ala	Met	Pro	His	Gly	Gly	Phe	Val	Ile				
			245				250					255							
cgc	agc	ggc	cag	aca	gga	gat	ggc	agc	aaa	cag	gta	ata	gag	tat	tcg				816
Arg	Ser	Gly	Gln	Thr	Gly	Asp	Gly	Ser	Lys	Gln	Val	Ile	Glu	Tyr	Ser				
		260				265					270								
ctc	aag	ctc	ggt	agc	aat	cct	gaa	ttt	aca	gcc	gcg	gta	cta	gtc	gct				864
Leu	Lys	Leu	Gly	Ser	Asn	Pro	Glu	Phe	Thr	Ala	Ala	Val	Leu	Val	Ala				
		275				280						285							
tac	gcc	cgc	gct	gct	ttc	cgc	ctc	cac	aaa	aaa	gga	gtg	cac	gga	gcg				912
Tyr	Ala	Arg	Ala	Ala	Phe	Arg	Leu	His	Lys	Lys	Gly	Val	His	Gly	Ala				
	290			295						300									
cac	agc	gtg	ctg	gat	atc	gct	ccg	ggc	ctt	ctg	tca	ccc	aaa	agt	ccg				960
His	Ser	Val	Leu	Asp	Ile	Ala	Pro	Gly	Leu	Leu	Ser	Pro	Lys	Ser	Pro				
	305			310				315				320							
gct	cag	ttg	cgc	aaa	gag	ttg	ctc	tga											987
Ala	Gln	Leu	Arg	Lys	Glu	Leu	Leu												
				325															

<210> 16  
 <211> 328  
 5 <212> PRT  
 <213> *Herminiimonas arsenicoxydans*

<400> 16

Met	Asp	Glu	Lys	Ile	Arg	Ile	Gly	Val	Ala	Gly	Tyr	Gly	Asn	Leu	Gly				
1				5					10				15						
Arg	Gly	Val	Glu	Met	Ala	Ile	Ala	Arg	Asn	Pro	Asp	Met	Gln	Leu	Val				
			20				25					30							
Gly	Val	Phe	Ser	Arg	Arg	Asp	Pro	Ala	Ser	Ile	Glu	Leu	Leu	Thr	Gln				
		35				40					45								
Ala	Val	Pro	Val	His	Lys	Phe	Asp	Asp	Ile	Glu	Gln	Phe	Arg	Asp	Gln				

10

ES 2 403 199 T3

```

      50              55              60
Ile Asp Val Leu Ile Leu Cys Gly Gly Ser Lys Asn Asp Leu Pro Glu
65              70              75              80
Gln Gly Pro Ala Leu Ala Thr Leu Phe Asn Thr Ile Asp Ser Phe Asp
      85              90              95
Thr His Asn Lys Ile Pro Glu Tyr Phe Ala Ala Met Asp Ser Ala Ser
      100              105              110
Cys Val Gly Glu Arg Thr Ser Ile Ile Ser Val Gly Trp Asp Pro Gly
      115              120              125
Val Phe Ser Leu Asn Arg Leu Phe Gly Glu Ala Ile Leu Pro Glu Gly
      130              135              140
Glu Thr Tyr Thr Phe Trp Gly Lys Gly Leu Ser Gln Gly His Ser Asp
145              150              155              160
Ala Ile Arg Arg Val Pro Gly Val Lys Ala Gly Val Gln Tyr Thr Ile
      165              170              175
Pro Ser Ala Glu Ala Met Glu Leu Val Arg Ser Gly Ser Gln Pro Gln
      180              185              190
Leu Ser Thr Arg Glu Lys His Thr Arg Glu Cys His Val Val Leu Glu
      195              200              205
Ala Gly Ala Asp Ala Lys Ala Val Glu His Ala Ile Val Ser Met Pro
      210              215              220
Asp Tyr Phe Ala Asp Tyr Asp Thr Thr Val His Phe Ile Ser Glu Glu
225              230              235              240
Glu Leu Arg Ser Asn His Ser Ala Met Pro His Gly Gly Phe Val Ile
      245              250              255
Arg Ser Gly Gln Thr Gly Asp Gly Ser Lys Gln Val Ile Glu Tyr Ser
      260              265              270
Leu Lys Leu Gly Ser Asn Pro Glu Phe Thr Ala Ala Val Leu Val Ala
      275              280              285
Tyr Ala Arg Ala Ala Phe Arg Leu His Lys Lys Gly Val His Gly Ala
      290              295              300
His Ser Val Leu Asp Ile Ala Pro Gly Leu Leu Ser Pro Lys Ser Pro
305              310              315              320
Ala Gln Leu Arg Lys Glu Leu Leu
      325

```

<210> 17  
 <211> 900  
 5 <212> ADN  
 <213> *Bacteroides thetaiotaomicron*

<220>  
 <221> CDS  
 10 <222> (1)..(900)

```

<400> 17
atg aaa aaa gta aga gca gct att gtc ggt tac ggc aat atc gga cac      48
Met Lys Lys Val Arg Ala Ala Ile Val Gly Tyr Gly Asn Ile Gly His
1              5              10              15
tat gta ctt gaa gcg cta cag gca gcg cct gat ttc gaa ata gcc gga      96
Tyr Val Leu Glu Ala Leu Gln Ala Ala Pro Asp Phe Glu Ile Ala Gly
      20              25              30
gta gtt cgt cgt gca gga gca gag aac aag ccg gaa gag ttg gca aac      144
Val Val Arg Arg Ala Gly Ala Glu Asn Lys Pro Glu Glu Leu Ala Asn
      35              40              45
tat gca gta gtc aaa gat atc aaa gag ctg gaa gga gtg gaa gtg gcc      192
Tyr Ala Val Val Lys Asp Ile Lys Glu Leu Glu Gly Val Glu Val Ala
      50              55              60
atc ctc tgc aca ccg acc cgc agc gtt gag aaa tac gcg aaa gaa tat      240
Ile Leu Cys Thr Pro Thr Arg Ser Val Glu Lys Tyr Ala Lys Glu Tyr

```

ES 2 403 199 T3

65					70					75				80	
ttg gca atg gga atc aac acg gtg gac agc ttc gac atc cac aca ggc															288
Leu Ala Met Gly Ile Asn Thr Val Asp Ser Phe Asp Ile His Thr Gly															
			85					90					95		
atc gtt gac ctg cgc cgc acg ctg gat gcc acc gcc aaa gag cac aaa															336
Ile Val Asp Leu Arg Arg Thr Leu Asp Ala Thr Ala Lys Glu His Lys															
			100					105					110		
gcc gta tcc atc atc tcc gca gga tgg gat ccg gga agc gac tcg atc															384
Ala Val Ser Ile Ile Ser Ala Gly Trp Asp Pro Gly Ser Asp Ser Ile								120					125		
			115												
gta cgc acc atg ctc gaa gca atc gct ccg aaa ggc atc act tac acc															432
Val Arg Thr Met Leu Glu Ala Ile Ala Pro Lys Gly Ile Thr Tyr Thr															
			130					135					140		
aac ttc ggt ccc ggc atg agt atg gga cac acc gta gcc gta aaa gcg															480
Asn Phe Gly Pro Gly Met Ser Met Gly His Thr Val Ala Val Lys Ala															
			145					150					155		160
atc gac gga gtg aaa gcc gcc ctc tcc atg acg atc cct act gga acg															528
Ile Asp Gly Val Lys Ala Ala Leu Ser Met Thr Ile Pro Thr Gly Thr															
			165					170					175		
gga atc cac cgc cgc atg gta tat atc gaa ctg aaa gac gga tac aag															576
Gly Ile His Arg Arg Met Val Tyr Ile Glu Leu Lys Asp Gly Tyr Lys															
			180					185					190		
ttt gaa gaa gta gcc gca gcc atc aag gca gac cct tac ttc gtg aac															624
Phe Glu Glu Val Ala Ala Ala Ile Lys Ala Asp Pro Tyr Phe Val Asn															
			195					200					205		
gac gag aca cat gta aaa ctt gtg ccc agc gta gac gca ctg ctc gat															672
Asp Glu Thr His Val Lys Leu Val Pro Ser Val Asp Ala Leu Leu Asp															
			210					215					220		
atg gga cac ggt gta aat ctg act cgc aag gga gtt tcc ggt aaa acg															720
Met Gly His Gly Val Asn Leu Thr Arg Lys Gly Val Ser Gly Lys Thr															
			225					230					235		240
cag aat cag ctg ttc gag ttc aat atg cgc atc aac aat ccg gca ctg															768
Gln Asn Gln Leu Phe Glu Phe Asn Met Arg Ile Asn Asn Pro Ala Leu															
			245					250					255		
acc gca cag gta ctt gtg tgc gtg gca cgc gct tcc atg aag cag caa															816
Thr Ala Gln Val Leu Val Cys Val Ala Arg Ala Ser Met Lys Gln Gln															
			260					265					270		
ccg ggg tgc tac acc atg gta gaa gtc ccc gtc atc gac ctc ctt ccg															864
Pro Gly Cys Tyr Thr Met Val Glu Val Pro Val Ile Asp Leu Leu Pro															
			275					280					285		
ggc gac cgc gaa gag tgg atc gga cac ctg gta taa															900
Gly Asp Arg Glu Glu Trp Ile Gly His Leu Val															
			290					295							

<210> 18  
 <211> 299  
 5 <212> PRT  
 <213> *Bacteroides thetaiotaomicron*

<400> 18

Met Lys Lys Val Arg Ala Ala Ile Val Gly Tyr Gly Asn Ile Gly His															
1			5					10					15		
Tyr Val Leu Glu Ala Leu Gln Ala Ala Pro Asp Phe Glu Ile Ala Gly															
			20					25					30		
Val Val Arg Arg Ala Gly Ala Glu Asn Lys Pro Glu Glu Leu Ala Asn															
			35					40					45		
Tyr Ala Val Val Lys Asp Ile Lys Glu Leu Glu Gly Val Glu Val Ala															
			50					55					60		
Ile Leu Cys Thr Pro Thr Arg Ser Val Glu Lys Tyr Ala Lys Glu Tyr															
			65					70					75		80

10



ES 2 403 199 T3

Leu Ala Met Gly Ile Asn Thr Val Asp Ser Phe Asp Ile His Thr Gly  
 85 90 95  
 Ile Val Asp Leu Arg Arg Thr Leu Asp Ala Thr Ala Lys Glu His Lys  
 100 105 110  
 Ala Val Ser Ile Ile Ser Ala Gly Trp Asp Pro Gly Ser Asp Ser Ile  
 115 120 125  
 Val Arg Thr Met Leu Glu Ala Ile Ala Pro Lys Gly Ile Thr Tyr Thr  
 130 135 140  
 Asn Phe Gly Pro Gly Met Ser Met Gly His Thr Val Ala Val Lys Ala  
 145 150 155 160  
 Ile Asp Gly Val Lys Ala Ala Leu Ser Met Thr Ile Pro Thr Gly Thr  
 165 170 175  
 Gly Ile His Arg Arg Met Val Tyr Ile Glu Leu Lys Asp Gly Tyr Lys  
 180 185 190  
 Phe Glu Glu Val Ala Ala Ala Ile Lys Ala Asp Pro Tyr Phe Val Asn  
 195 200 205  
 Asp Glu Thr His Val Lys Leu Val Pro Ser Val Asp Ala Leu Leu Asp  
 210 215 220  
 Met Gly His Gly Val Asn Leu Thr Arg Lys Gly Val Ser Gly Lys Thr  
 225 230 235 240  
 Gln Asn Gln Leu Phe Glu Phe Asn Met Arg Ile Asn Asn Pro Ala Leu  
 245 250 255  
 Thr Ala Gln Val Leu Val Cys Val Ala Arg Ala Ser Met Lys Gln Gln  
 260 265 270  
 Pro Gly Cys Tyr Thr Met Val Glu Val Pro Val Ile Asp Leu Leu Pro  
 275 280 285  
 Gly Asp Arg Glu Glu Trp Ile Gly His Leu Val  
 290 295

<210> 19  
 <211> 69  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> cebador

10 <400> 19

atgcagcagt tacagaacat tattgaaacc gcttttgaac gccgtgccga tctagacgct 60  
 caagttagt 69

15 <210> 20  
 <211> 70  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> cebador

<400> 20

ttagtcgatg gtacgcagca gttcgttaat gccgactttg ccgcgagttt agatcttgaa 60  
 gcctgctttt 70

30 <210> 21  
 <211> 1350  
 <212> ADN  
 <213> *Escherichia coli*

<220>  
 <221> CDS  
 <222> (1)..(1350)

ES 2 403 199 T3

<400> 21

atg tct gaa att gtt gtc tcc aaa ttt ggc ggt acc agc gta gct gat	48
Met Ser Glu Ile Val Val Ser Lys Phe Gly Gly Thr Ser Val Ala Asp	
1 5 10 15	
ttt gac gcc atg aac cgc agc gct gat att gtg ctt tct gat gcc aac	96
Phe Asp Ala Met Asn Arg Ser Ala Asp Ile Val Leu Ser Asp Ala Asn	
20 25 30	
gtg cgt tta gtt gtc ctc tcg gct tct gct ggt atc act aat ctg ctg	144
Val Arg Leu Val Val Leu Ser Ala Ser Ala Gly Ile Thr Asn Leu Leu	
35 40 45	
gtc gct tta gct gaa gga ctg gaa cct ggc gag cga ttc gaa aaa ctc	192
Val Ala Leu Ala Glu Gly Leu Glu Pro Gly Glu Arg Phe Glu Lys Leu	
50 55 60	
gac gct atc cgc aac atc cag ttt gcc att ctg gaa cgt ctg cgt tac	240
Asp Ala Ile Arg Asn Ile Gln Phe Ala Ile Leu Glu Arg Leu Arg Tyr	
65 70 75 80	
ccg aac gtt atc cgt gaa gag att gaa cgt ctg ctg gag aac att act	288
Pro Asn Val Ile Arg Glu Glu Ile Glu Arg Leu Leu Glu Asn Ile Thr	
85 90 95	
gtt ctg gca gaa gcg gcg gcg ctg gca acg tct ccg gcg ctg aca gat	336
Val Leu Ala Glu Ala Ala Ala Leu Ala Thr Ser Pro Ala Leu Thr Asp	
100 105 110	
gag ctg gtc agc cac ggc gag ctg atg tcg acc ctg ctg ttt gtt gag	384
Glu Leu Val Ser His Gly Glu Leu Met Ser Thr Leu Leu Phe Val Glu	
115 120 125	
atc ctg cgc gaa cgc gat gtt cag gca cag tgg ttt gat gta cgt aaa	432
Ile Leu Arg Glu Arg Asp Val Gln Ala Gln Trp Phe Asp Val Arg Lys	
130 135 140	
gtg atg cgt acc aac gac cga ttt ggt cgt gca gag cca gat ata gcc	480
Val Met Arg Thr Asn Asp Arg Phe Gly Arg Ala Glu Pro Asp Ile Ala	
145 150 155 160	
gcg ctg gcg gaa ctg gcc gcg ctg cag ctg ctc cca cgt ctc aat gaa	528
Ala Leu Ala Glu Leu Ala Ala Leu Gln Leu Leu Pro Arg Leu Asn Glu	
165 170 175	
ggc tta gtg atc acc cag gga ttt atc ggt agc gaa aat aaa ggt cgt	576
Gly Leu Val Ile Thr Gln Gly Phe Ile Gly Ser Glu Asn Lys Gly Arg	
180 185 190	
aca acg acg ctt ggc cgt gga ggc agc gat tat acg gca gcc ttg ctg	624
Thr Thr Thr Leu Gly Arg Gly Gly Ser Asp Tyr Thr Ala Ala Leu Leu	
195 200 205	
gcg gag gct tta cac gca tct cgt gtt gat atc tgg acc gac gtc ccg	672
Ala Glu Ala Leu His Ala Ser Arg Val Asp Ile Trp Thr Asp Val Pro	
210 215 220	
ggc atc tac acc acc gat cca cgc gta gtt tcc gca gca aaa cgc att	720
Gly Ile Tyr Thr Thr Asp Pro Arg Val Val Ser Ala Ala Lys Arg Ile	
225 230 235 240	
gat gaa atc gcg ttt gcc gaa gcg gca gag atg gca act ttt ggt gca	768
Asp Glu Ile Ala Phe Ala Glu Ala Ala Glu Met Ala Thr Phe Gly Ala	
245 250 255	
aaa gta ctg cat ccg gca acg ttg cta ccc gca gta cgc agc gat atc	816
Lys Val Leu His Pro Ala Thr Leu Leu Pro Ala Val Arg Ser Asp Ile	
260 265 270	
ccg gtc ttt gtc ggc tcc agc aaa gac cca cgc gca ggt ggt acg ctg	864
Pro Val Phe Val Gly Ser Ser Lys Asp Pro Arg Ala Gly Gly Thr Leu	
275 280 285	
gtg tgc aat aaa act gaa aat ccg ccg ctg ttc cgc gct ctg gcg ctt	912
Val Cys Asn Lys Thr Glu Asn Pro Pro Leu Phe Arg Ala Leu Ala Leu	

ES 2 403 199 T3

290	295	300	
cgt cgc aat cag act ctg ctc act ttg cac agc ctg aat atg ctg cat			960
Arg Arg Asn Gln Thr Leu Leu Thr Leu His Ser Leu Asn Met Leu His			
305	310	315	320
tct cgc ggt ttc ctc gcg gaa gtt ttc ggc atc ctc gcg cgg cat aat			1008
Ser Arg Gly Phe Leu Ala Glu Val Phe Gly Ile Leu Ala Arg His Asn			
	325	330	335
att tcg gta gac tta atc acc acg tca gaa gtg agc gtg gca tta acc			1056
Ile Ser Val Asp Leu Ile Thr Thr Ser Glu Val Ser Val Ala Leu Thr			
	340	345	350
ctt gat acc acc ggt tca acc tcc act ggc gat acg ttg ctg acg caa			1104
Leu Asp Thr Thr Gly Ser Thr Ser Thr Gly Asp Thr Leu Leu Thr Gln			
	355	360	365
tct ctg ctg atg gag ctt tcc gca ctg tgt cgg gtg gag gtg gaa gaa			1152
Ser Leu Leu Met Glu Leu Ser Ala Leu Cys Arg Val Glu Val Glu Glu			
	370	375	380
ggc ctg gcg ctg gtc gcg ttg att ggc aat gac ctg tca aaa gcc tgc			1200
Gly Leu Ala Leu Val Ala Leu Ile Gly Asn Asp Leu Ser Lys Ala Cys			
	385	390	400
ggc gtt ggc aaa gag gta ttc ggc gta ctg gaa ccg ttc aac att cgc			1248
Gly Val Gly Lys Glu Val Phe Gly Val Leu Glu Pro Phe Asn Ile Arg			
	405	410	415
atg att tgt tat ggc gca tcc agc cat aac ctg tgc ttc ctg gtg ccc			1296
Met Ile Cys Tyr Gly Ala Ser Ser His Asn Leu Cys Phe Leu Val Pro			
	420	425	430
ggc gaa gat gcc gag cag gtg gtg caa aaa ctg cat agt aat ttg ttt			1344
Gly Glu Asp Ala Glu Gln Val Val Gln Lys Leu His Ser Asn Leu Phe			
	435	440	445
gag taa			1350
Glu			

<210> 22  
 <211> 449  
 5 <212> PRT  
 <213> *Escherichia coli*

<400> 22

Met Ser Glu Ile Val Val Ser Lys Phe Gly Gly Thr Ser Val Ala Asp	
1	5
Phe Asp Ala Met Asn Arg Ser Ala Asp Ile Val Leu Ser Asp Ala Asn	
	20
Val Arg Leu Val Val Leu Ser Ala Ser Ala Gly Ile Thr Asn Leu Leu	
	35
Val Ala Leu Ala Glu Gly Leu Glu Pro Gly Glu Arg Phe Glu Lys Leu	
	50
Asp Ala Ile Arg Asn Ile Gln Phe Ala Ile Leu Glu Arg Leu Arg Tyr	
65	70
Pro Asn Val Ile Arg Glu Glu Ile Glu Arg Leu Leu Glu Asn Ile Thr	
	85
Val Leu Ala Glu Ala Ala Ala Leu Ala Thr Ser Pro Ala Leu Thr Asp	
	100
Glu Leu Val Ser His Gly Glu Leu Met Ser Thr Leu Leu Phe Val Glu	
	115
Ile Leu Arg Glu Arg Asp Val Gln Ala Gln Trp Phe Asp Val Arg Lys	
	130
Val Met Arg Thr Asn Asp Arg Phe Gly Arg Ala Glu Pro Asp Ile Ala	
145	150
Ala Leu Ala Glu Leu Ala Ala Leu Gln Leu Leu Pro Arg Leu Asn Glu	
	165
	170
	175

ES 2 403 199 T3

Gly Leu Val Ile Thr Gln Gly Phe Ile Gly Ser Glu Asn Lys Gly Arg  
 180 185 190  
 Thr Thr Thr Leu Gly Arg Gly Gly Ser Asp Tyr Thr Ala Ala Leu Leu  
 195 200 205  
 Ala Glu Ala Leu His Ala Ser Arg Val Asp Ile Trp Thr Asp Val Pro  
 210 215 220  
 Gly Ile Tyr Thr Thr Asp Pro Arg Val Val Ser Ala Ala Lys Arg Ile  
 225 230 235 240  
 Asp Glu Ile Ala Phe Ala Glu Ala Ala Glu Met Ala Thr Phe Gly Ala  
 245 250 255  
 Lys Val Leu His Pro Ala Thr Leu Leu Pro Ala Val Arg Ser Asp Ile  
 260 265 270  
 Pro Val Phe Val Gly Ser Ser Lys Asp Pro Arg Ala Gly Gly Thr Leu  
 275 280 285  
 Val Cys Asn Lys Thr Glu Asn Pro Pro Leu Phe Arg Ala Leu Ala Leu  
 290 295 300  
 Arg Arg Asn Gln Thr Leu Leu Thr Leu His Ser Leu Asn Met Leu His  
 305 310 315 320  
 Ser Arg Gly Phe Leu Ala Glu Val Phe Gly Ile Leu Ala Arg His Asn  
 325 330 335  
 Ile Ser Val Asp Leu Ile Thr Thr Ser Glu Val Ser Val Ala Leu Thr  
 340 345 350  
 Leu Asp Thr Thr Gly Ser Thr Ser Thr Gly Asp Thr Leu Leu Thr Gln  
 355 360 365  
 Ser Leu Leu Met Glu Leu Ser Ala Leu Cys Arg Val Glu Val Glu Glu  
 370 375 380  
 Gly Leu Ala Leu Val Ala Leu Ile Gly Asn Asp Leu Ser Lys Ala Cys  
 385 390 395 400  
 Gly Val Gly Lys Glu Val Phe Gly Val Leu Glu Pro Phe Asn Ile Arg  
 405 410 415  
 Met Ile Cys Tyr Gly Ala Ser Ser His Asn Leu Cys Phe Leu Val Pro  
 420 425 430  
 Gly Glu Asp Ala Glu Gln Val Val Gln Lys Leu His Ser Asn Leu Phe  
 435 440 445  
 Glu

<210> 23  
 <211> 879  
 5 <212> ADN  
 <213> *Escherichia coli*

<220>  
 <221> CDS  
 10 <222> (1)..(879)

<400> 23

atg ttc acg gga agt att gtc gcg att gtt act ccg atg gat gaa aaa 48  
 Met Phe Thr Gly Ser Ile Val Ala Ile Val Thr Pro Met Asp Glu Lys  
 1 5 10 15  
 ggt aat gtc tgt cgg gct agc ttg aaa aaa ctg att gat tat cat gtc 96  
 Gly Asn Val Cys Arg Ala Ser Leu Lys Lys Leu Ile Asp Tyr His Val  
 20 25 30  
 gcc agc ggt act tcg gcg atc gtt tct gtt ggc acc act ggc gag tcc 144  
 Ala Ser Gly Thr Ser Ala Ile Val Ser Val Gly Thr Thr Gly Glu Ser  
 35 40 45  
 gct acc tta aat cat gac gaa cat gct gat gtg gtg atg atg acg ctg 192  
 Ala Thr Leu Asn His Asp Glu His Ala Asp Val Val Met Met Thr Leu  
 50 55 60  
 gat ctg gct gat ggg cgc att ccg gta att gcc ggg acc ggc gct aac 240

ES 2 403 199 T3

Asp Leu Ala Asp Gly Arg Ile Pro Val Ile Ala Gly Thr Gly Ala Asn  
65 70 75 80  
gct act gcg gaa gcc att agc ctg acg cag cgc ttc aat gac agt ggt 288  
Ala Thr Ala Glu Ala Ile Ser Leu Thr Gln Arg Phe Asn Asp Ser Gly  
85 90 95  
atc gtc ggc tgc ctg acg gta acc cct tac tac aat cgt ccg tgc caa 336  
Ile Val Gly Cys Leu Thr Val Thr Pro Tyr Tyr Asn Arg Pro Ser Gln  
100 105 110  
gaa ggt ttg tat cag cat ttc aaa gcc atc gct gag cat act gac ctg 384  
Glu Gly Leu Tyr Gln His Phe Lys Ala Ile Ala Glu His Thr Asp Leu  
115 120 125  
ccg caa att ctg tat aat gtg ccg tcc cgt act ggc tgc gat ctg ctc 432  
Pro Gln Ile Leu Tyr Asn Val Pro Ser Arg Thr Gly Cys Asp Leu Leu  
130 135 140  
ccg gaa acg gtg ggc cgt ctg gcg aaa gta aaa aat att atc gga atc 480  
Pro Glu Thr Val Gly Arg Leu Ala Lys Val Lys Asn Ile Ile Gly Ile  
145 150 155 160  
aaa gag gca aca ggg aac tta acg cgt gta aac cag atc aaa gag ctg 528  
Lys Glu Ala Thr Gly Asn Leu Thr Arg Val Asn Gln Ile Lys Glu Leu  
165 170 175  
gtt tca gat gat ttt gtt ctg ctg agc ggc gat gat gcg agc gcg ctg 576  
Val Ser Asp Asp Phe Val Leu Leu Ser Gly Asp Asp Ala Ser Ala Leu  
180 185 190  
gac ttc atg caa ttg ggc ggt cat ggg gtt att tcc gtt acg gct aac 624  
Asp Phe Met Gln Leu Gly Gly His Gly Val Ile Ser Val Thr Ala Asn  
195 200 205  
gtc gca gcg cgt gat atg gcc cag atg tgc aaa ctg gca gca gaa ggg 672  
Val Ala Ala Arg Asp Met Ala Gln Met Cys Lys Leu Ala Ala Glu Gly  
210 215 220  
cat ttt gcc gag gca cgc gtt att aat cag cgt ctg atg cca tta cac 720  
His Phe Ala Glu Ala Arg Val Ile Asn Gln Arg Leu Met Pro Leu His  
225 230 235 240  
aac aaa cta ttt gtc gaa ccc aat cca atc ccg gtg aaa tgg gca tgt 768  
Asn Lys Leu Phe Val Glu Pro Asn Pro Ile Pro Val Lys Trp Ala Cys  
245 250 255  
aag gaa ctg ggt ctt gtg gcg acc gat acg ctg cgc ctg cca atg aca 816  
Lys Glu Leu Gly Leu Val Ala Thr Asp Thr Leu Arg Leu Pro Met Thr  
260 265 270  
cca atc acc gac agt ggt cgt gag acg gtc aga gcg gcg ctt aag cat 864  
Pro Ile Thr Asp Ser Gly Arg Glu Thr Val Arg Ala Ala Leu Lys His  
275 280 285  
gcc ggt ttg ctg taa 879  
Ala Gly Leu Leu  
290

<210> 24  
<211> 292  
5 <212> PRT  
<213> *Escherichia coli*

<400> 24

Met Phe Thr Gly Ser Ile Val Ala Ile Val Thr Pro Met Asp Glu Lys  
1 5 10 15  
Gly Asn Val Cys Arg Ala Ser Leu Lys Lys Leu Ile Asp Tyr His Val  
20 25 30  
Ala Ser Gly Thr Ser Ala Ile Val Ser Val Gly Thr Thr Gly Glu Ser  
35 40 45  
Ala Thr Leu Asn His Asp Glu His Ala Asp Val Val Met Met Thr Leu  
50 55 60  
10 Asp Leu Ala Asp Gly Arg Ile Pro Val Ile Ala Gly Thr Gly Ala Asn

65					70					75					80
Ala	Thr	Ala	Glu	Ala	Ile	Ser	Leu	Thr	Gln	Arg	Phe	Asn	Asp	Ser	Gly
				85					90					95	
Ile	Val	Gly	Cys	Leu	Thr	Val	Thr	Pro	Tyr	Tyr	Asn	Arg	Pro	Ser	Gln
		100					105						110		
Glu	Gly	Leu	Tyr	Gln	His	Phe	Lys	Ala	Ile	Ala	Glu	His	Thr	Asp	Leu
		115					120						125		
Pro	Gln	Ile	Leu	Tyr	Asn	Val	Pro	Ser	Arg	Thr	Gly	Cys	Asp	Leu	Leu
	130					135						140			
Pro	Glu	Thr	Val	Gly	Arg	Leu	Ala	Lys	Val	Lys	Asn	Ile	Ile	Gly	Ile
145					150						155				160
Lys	Glu	Ala	Thr	Gly	Asn	Leu	Thr	Arg	Val	Asn	Gln	Ile	Lys	Glu	Leu
				165					170					175	
Val	Ser	Asp	Asp	Phe	Val	Leu	Leu	Ser	Gly	Asp	Asp	Ala	Ser	Ala	Leu
		180						185						190	
Asp	Phe	Met	Gln	Leu	Gly	Gly	His	Gly	Val	Ile	Ser	Val	Thr	Ala	Asn
	195						200						205		
Val	Ala	Ala	Arg	Asp	Met	Ala	Gln	Met	Cys	Lys	Leu	Ala	Ala	Glu	Gly
	210					215						220			
His	Phe	Ala	Glu	Ala	Arg	Val	Ile	Asn	Gln	Arg	Leu	Met	Pro	Leu	His
225					230						235				240
Asn	Lys	Leu	Phe	Val	Glu	Pro	Asn	Pro	Ile	Pro	Val	Lys	Trp	Ala	Cys
			245						250					255	
Lys	Glu	Leu	Gly	Leu	Val	Ala	Thr	Asp	Thr	Leu	Arg	Leu	Pro	Met	Thr
			260					265						270	
Pro	Ile	Thr	Asp	Ser	Gly	Arg	Glu	Thr	Val	Arg	Ala	Ala	Leu	Lys	His
		275					280						285		
Ala	Gly	Leu	Leu												
	290														

<210> 25  
 <211> 822  
 5 <212> ADN  
 <213> *Escherichia coli*

<220>  
 <221> CDS  
 10 <222> (1)..(822)

<400> 25

atg	cat	gat	gca	aac	atc	cgc	gtt	gcc	atc	gcg	gga	gcc	ggg	ggg	cgt	48
Met	His	Asp	Ala	Asn	Ile	Arg	Val	Ala	Ile	Ala	Gly	Ala	Gly	Gly	Arg	
1			5					10					15			
atg	ggc	cgc	cag	ttg	att	cag	gcg	gcg	ctg	gca	tta	gag	ggc	gtg	cag	96
Met	Gly	Arg	Gln	Leu	Ile	Gln	Ala	Ala	Leu	Ala	Leu	Glu	Gly	Val	Gln	
			20					25					30			
ttg	ggc	gct	gcg	ctg	gag	cgt	gaa	gga	tct	tct	tta	ctg	ggc	agc	gac	144
Leu	Gly	Ala	Ala	Leu	Glu	Arg	Glu	Gly	Ser	Ser	Leu	Leu	Gly	Ser	Asp	
			35				40						45			
gcc	ggt	gag	ctg	gcc	gga	gcc	ggg	aaa	aca	ggc	ggt	acc	gtg	caa	agc	192
Ala	Gly	Glu	Leu	Ala	Gly	Ala	Gly	Lys	Thr	Gly	Val	Thr	Val	Gln	Ser	
	50					55					60					
agc	ctc	gat	gcg	gta	aaa	gat	gat	ttt	gat	gtg	ttt	atc	gat	ttt	acc	240
Ser	Leu	Asp	Ala	Val	Lys	Asp	Asp	Phe	Asp	Val	Phe	Ile	Asp	Phe	Thr	
65					70					75					80	
cgt	ccg	gaa	ggt	acg	ctg	aac	cat	ctc	gct	ttt	tgt	cgc	cag	cat	ggc	288
Arg	Pro	Glu	Gly	Thr	Leu	Asn	His	Leu	Ala	Phe	Cys	Arg	Gln	His	Gly	
				85					90					95		
aaa	ggg	atg	gtg	atc	ggc	act	acg	ggg	ttt	gac	gaa	gcc	ggt	aaa	caa	336
Lys	Gly	Met	Val	Ile	Gly	Thr	Thr	Gly	Phe	Asp	Glu	Ala	Gly	Lys	Gln	

ES 2 403 199 T3

			100					105				110										
gca	att	cgt	gac	gcc	gct	gcc	gat	att	gcg	att	gtc	ttt	gct	gcc	aat							384
Ala	Ile	Arg	Asp	Ala	Ala	Ala	Asp	Ile	Ala	Ile	Val	Phe	Ala	Ala	Asn							
			115					120				125										
ttt	agc	ggt	ggc	ggt	aac	gtc	atg	ctt	aag	ctg	ctg	gag	aaa	gca	gcc							432
Phe	Ser	Val	Gly	Val	Asn	Val	Met	Leu	Lys	Leu	Leu	Glu	Lys	Ala	Ala							
			130					135				140										
aaa	gtg	atg	ggt	gac	tac	acc	gat	atc	gaa	att	att	gaa	gca	cat	cat							480
Lys	Val	Met	Gly	Asp	Tyr	Thr	Asp	Ile	Glu	Ile	Ile	Glu	Ala	His	His							
			145				150					155			160							
aga	cat	aaa	ggt	gat	gcg	ccg	tca	ggc	acc	gca	ctg	gca	atg	gga	gag							528
Arg	His	Lys	Val	Asp	Ala	Pro	Ser	Gly	Thr	Ala	Leu	Ala	Met	Gly	Glu							
				165								170			175							
gcg	atc	gcc	cac	gcc	ctt	gat	aaa	gat	ctg	aaa	gat	tgc	gcg	gtc	tac							576
Ala	Ile	Ala	His	Ala	Leu	Asp	Lys	Asp	Leu	Lys	Asp	Cys	Ala	Val	Tyr							
			180					185				190										
agt	cgt	gaa	ggc	cac	acc	ggg	gaa	cgt	gtg	cct	ggc	acc	att	ggg	ttt							624
Ser	Arg	Glu	Gly	His	Thr	Gly	Glu	Arg	Val	Pro	Gly	Thr	Ile	Gly	Phe							
			195					200				205										
gcc	acc	gtg	cgt	gca	ggg	gac	atc	ggt	ggg	gaa	cat	acc	gcg	atg	ttt							672
Ala	Thr	Val	Arg	Ala	Gly	Asp	Ile	Val	Gly	Glu	His	Thr	Ala	Met	Phe							
			210				215					220										
gcc	gat	att	ggc	gag	cgt	ctg	gag	atc	acc	cat	aag	gcg	tcc	agc	cgt							720
Ala	Asp	Ile	Gly	Glu	Arg	Leu	Glu	Ile	Thr	His	Lys	Ala	Ser	Ser	Arg							
				225			230				235			240								
atg	aca	ttt	gct	aac	ggc	gcg	gta	aga	tcg	gct	ttg	tgg	ttg	agt	ggg							768
Met	Thr	Phe	Ala	Asn	Gly	Ala	Val	Arg	Ser	Ala	Leu	Trp	Leu	Ser	Gly							
				245					250					255								
aag	gaa	agc	ggg	ctt	ttt	gat	atg	cga	gat	gta	ctt	gat	ctc	aat	aat							816
Lys	Glu	Ser	Gly	Leu	Phe	Asp	Met	Arg	Asp	Val	Leu	Asp	Leu	Asn	Asn							
			260					265					270									
ttg	taa																					822
Leu																						

<210> 26  
 <211> 273  
 <212> PRT  
 <213> *Escherichia coli*

5

<400> 26

Met	His	Asp	Ala	Asn	Ile	Arg	Val	Ala	Ile	Ala	Gly	Ala	Gly	Gly	Arg							
1				5					10					15								
Met	Gly	Arg	Gln	Leu	Ile	Gln	Ala	Ala	Leu	Ala	Leu	Glu	Gly	Val	Gln							
			20					25					30									
Leu	Gly	Ala	Ala	Leu	Glu	Arg	Glu	Gly	Ser	Ser	Leu	Leu	Gly	Ser	Asp							
			35				40					45										
Ala	Gly	Glu	Leu	Ala	Gly	Ala	Gly	Lys	Thr	Gly	Val	Thr	Val	Gln	Ser							
			50			55					60											
Ser	Leu	Asp	Ala	Val	Lys	Asp	Asp	Phe	Asp	Val	Phe	Ile	Asp	Phe	Thr							
65					70					75				80								
Arg	Pro	Glu	Gly	Thr	Leu	Asn	His	Leu	Ala	Phe	Cys	Arg	Gln	His	Gly							
				85					90					95								
Lys	Gly	Met	Val	Ile	Gly	Thr	Thr	Gly	Phe	Asp	Glu	Ala	Gly	Lys	Gln							
			100					105					110									
Ala	Ile	Arg	Asp	Ala	Ala	Ala	Asp	Ile	Ala	Ile	Val	Phe	Ala	Ala	Asn							
			115					120				125										
Phe	Ser	Val	Gly	Val	Asn	Val	Met	Leu	Lys	Leu	Leu	Glu	Lys	Ala	Ala							
			130				135					140										
Lys	Val	Met	Gly	Asp	Tyr	Thr	Asp	Ile	Glu	Ile	Ile	Glu	Ala	His	His							
145					150						155				160							

ES 2 403 199 T3

Arg	His	Lys	Val	Asp	Ala	Pro	Ser	Gly	Thr	Ala	Leu	Ala	Met	Gly	Glu
				165					170					175	
Ala	Ile	Ala	His	Ala	Leu	Asp	Lys	Asp	Leu	Lys	Asp	Cys	Ala	Val	Tyr
			180					185				190			
Ser	Arg	Glu	Gly	His	Thr	Gly	Glu	Arg	Val	Pro	Gly	Thr	Ile	Gly	Phe
		195					200					205			
Ala	Thr	Val	Arg	Ala	Gly	Asp	Ile	Val	Gly	Glu	His	Thr	Ala	Met	Phe
	210					215					220				
Ala	Asp	Ile	Gly	Glu	Arg	Leu	Glu	Ile	Thr	His	Lys	Ala	Ser	Ser	Arg
225					230					235					240
Met	Thr	Phe	Ala	Asn	Gly	Ala	Val	Arg	Ser	Ala	Leu	Trp	Leu	Ser	Gly
				245					250					255	
Lys	Glu	Ser	Gly	Leu	Phe	Asp	Met	Arg	Asp	Val	Leu	Asp	Leu	Asn	Asn
			260					265					270		
Leu															



## REIVINDICACIONES

- 5 1. Método para producir L-lisina, que comprende hacer crecer en cultivo en un medio una *Escherichia coli* que tiene capacidad productora de L-lisina y recoger L-lisina del medio,
- en el que la *Escherichia coli* se ha modificado para disminuir la actividad o actividades de una o más clases de enzimas de la ruta de síntesis del ácido meso- $\alpha,\epsilon$ -diaminopimélico, y
- 10 en el que se ha introducido un gen que codifica para diaminopimelato deshidrogenasa en la *Escherichia coli*.
2. Método según la reivindicación 1, en el que las enzimas de la ruta de síntesis del ácido meso- $\alpha,\epsilon$ -diaminopimélico se seleccionan de 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa, succinildiaminopimelato transaminasa, succinildiaminopimelato desuccinilasa y diaminopimelato epimerasa.
- 15 3. Método según la reivindicación 2, en el que la 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa, succinildiaminopimelato transaminasa, succinildiaminopimelato desuccinilasa y diaminopimelato epimerasa se codifican por el gen *dapD*, gen *dapC*, gen *dapE* y gen *dapF*, respectivamente.
- 20 4. Método según la reivindicación 3, en el que las actividades de las enzimas de la ruta de síntesis del ácido meso- $\alpha,\epsilon$ -diaminopimélico se disminuyen mediante la disminución de la expresión de los genes o mediante la alteración de los genes.
- 25 5. Método según una cualquiera de las reivindicaciones 2 a 4, en el que la *Escherichia coli* se ha modificado para disminuir al menos la actividad de 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa.
- 30 6. Método según una cualquiera de las reivindicaciones 1 a 5, en el que el gen que codifica para diaminopimelato deshidrogenasa es el gen *ddh* de una bacteria corineforme.
7. Método según una cualquiera de las reivindicaciones 2 a 6, en el que la 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa es una proteína definida en (A) o (B) siguientes:
- (A) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 2,
- 35 (B) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 2 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa.
- 40 8. Método según una cualquiera de las reivindicaciones 2 a 7, en el que la succinildiaminopimelato transaminasa es una proteína definida en (C) o (D) siguientes:
- (C) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 4,
- 45 (D) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 4 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de succinildiaminopimelato transaminasa.
- 50 9. Método según una cualquiera de las reivindicaciones 2 a 8, en el que la succinildiaminopimelato desuccinilasa es una proteína definida en (E) o (F) siguientes:
- (E) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 6,
- 55 (F) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 6 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de succinildiaminopimelato desuccinilasa.
- 60 10. Método según una cualquiera de las reivindicaciones 2 a 9, en el que la diaminopimelato epimerasa es una proteína definida en (G) o (H) siguientes:
- (G) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 8,
- 65 (H) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 8 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de diaminopimelato epimerasa.
11. Método según una cualquiera de las reivindicaciones 3 a 10, en el que el gen *dapD* es un ADN definido en (a) o (b) siguientes:

- (a) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 1, o
- 5 (b) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 1 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de 2,3,4,5-tetrahidropiridina-2,6-dicarboxilato N-succiniltransferasa.
12. Método según una cualquiera de las reivindicaciones 3 a 11, en el que el gen *dapC* es un ADN definido en (c) o (d) siguientes:
- 10 (c) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 3, o
- (d) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 3 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de succinildiaminopimelato transaminasa.
- 15 13. Método según una cualquiera de las reivindicaciones 3 a 12, en el que el gen *dapE* es un ADN definido en (e) o (f) siguientes:
- 20 (e) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 5, o
- (f) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 5 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de succinildiaminopimelato desuccinilasa.
- 25 14. Método según una cualquiera de las reivindicaciones 3 a 13, en el que el gen *dapF* es un ADN definido en (g) o (h) siguientes:
- 30 (g) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 7, o
- (h) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 7 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de diaminopimelato epimerasa.
- 35 15. Método según una cualquiera de las reivindicaciones 1 a 14, en el que la diaminopimelato deshidrogenasa es una proteína definida en (I) o (J) siguientes:
- (I) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 10, 12, 14, 16 ó 18,
- 40 (J) una proteína que tiene la secuencia de aminoácidos de SEQ ID NO: 10, 12, 14, 16 ó 18 que incluye sustituciones, deleciones, inserciones o adiciones de uno o varios residuos de aminoácidos y que tiene la actividad de diaminopimelato deshidrogenasa.
- 45 16. Método según una cualquiera de las reivindicaciones 6 a 13, en el que el gen *ddh* es un ADN definido en (i) o (j) siguientes:
- (i) un ADN que comprende la secuencia de nucleótidos de SEQ ID NO: 9, 11, 13, 15 ó 17, o
- 50 (j) un ADN que puede hibridarse con la secuencia de nucleótidos de SEQ ID NO: 9, 11, 13, 15 ó 17 o una sonda que puede prepararse a partir de la secuencia de nucleótidos en condiciones rigurosas y que codifica para una proteína que tiene la actividad de diaminopimelato deshidrogenasa.
- 55 17. Método según una cualquiera de las reivindicaciones 1 a 16, en el que la *Escherichia coli* tiene además una dihidrodipicolinato sintasa que está desensibilizada para la retroinhibición por L-lisina y una aspartocinasa que está desensibilizada para la retroinhibición por L-lisina y tiene actividad potenciada de dihidrodipicolinato reductasa.