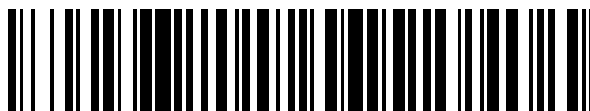


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 526 490**

51 Int. Cl.:

C12N 1/21 (2006.01)

C12N 15/09 (2006.01)

C12P 21/02 (2006.01)

C07K 5/06 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.04.2008 E 08739864 (0)**

97 Fecha y número de publicación de la concesión europea: **10.12.2014 EP 2143786**

54 Título: **Método para producir un dipéptido**

30 Prioridad:

06.04.2007 JP 2007099956

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

12.01.2015

73 Titular/es:

**KYOWA HAKKO BIO CO., LTD. (100.0%)
1-6-1, Ohtemachi, Chiyoda-ku
Tokyo 100-8185, JP**

72 Inventor/es:

**HAYASHI, MIKIRO;
TABATA, KAZUHIKO;
YAGASAKI, MAKOTO y
YONETANI, YOSHIYUKI**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 526 490 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 un dipéptido

Campo Técnico

5 La presente invención se refiere a un microorganismo que tiene la capacidad de producir una proteína que presenta actividad de síntesis de un dipéptido, y en el cual la actividad de una proteína para transportar un dipéptido de la célula microbiana al exterior de la célula microbiana es mayor que la de la cepa original, y un proceso para producir un dipéptido usando el microorganismo.

Antecedentes de la invención

10 Como proceso para producir un dipéptido donde dos aminoácidos L están unidos mediante un enlace α , se conoce un proceso que utiliza el producto génico *ywfE* que es una de las sintasas de la bacilisina, un dipéptido antibiótico derivado de un microorganismo que pertenece al género *Bacillus* (documentos de patentes 1 – 8).

15 Como método para mejorar la capacidad productora de una cepa en fermentación de aminoácidos L, se conoce un método que incluye alterar un sistema de eflujo para aminoácidos L en una célula microbiana (documentos de patentes 9 – 14). Adicionalmente, también se conoce la presencia de un sistema de eflujo para Microcina J25, un antibiótico peptídico (documento no patente 1).

20 También en la producción de un dipéptido, se considera que la productividad de dipéptido mejora al potenciar la actividad de un sistema de exportación de dipéptidos, a saber, una proteína que presente actividad para transportar un dipéptido de la célula microbiana hacia el exterior de las células microbianas (desde este punto será referida como actividad de transporte de dipéptidos). Sin embargo, hasta el momento no se ha descubierto ninguna proteína que tenga actividad de transporte de dipéptidos.

25 Es sabido que el gen *bcr* de *Escherichia coli* es un gen que codifica proteína de membrana que está relacionado con la resistencia a biciclomicina (documento no patente 2), y el gen *norE* es un gen que codifica una bomba de eflujo que está relacionada con la resistencia a quinolona (documento no patente 3). Se ha publicado que el gen *emrD* es un gen que codifica una proteína transportadora de SDS (documento no patente 4). Se predice que el gen *ydeE* es un gen que codifica una proteína transportadora de fármaco, pero su actividad no ha sido confirmada (documento no patente 4). La función del gen *yeeO* permanece completamente desconocida.

30 Además, previamente se han descrito métodos para producir L-cisteína en microorganismos en los que el gen *bcr* podría regularse al alza (Patente de EE.UU. 2005/221453). Adicionalmente, se ha descrito la sobreexpresión de proteínas transportadoras de fármaco (que incluyen el *bcr*) en *E. coli* (*Applied and Environmental Microbiology*, 72(7), 4735-4742 (2006)). Finalmente, se han descrito métodos para producir dipéptidos en microorganismos, en particular *E. coli*, donde la *E. coli* ha sido transformada con el gen *ywfE* de *Bacillus subtilis* (EP 1 616 963).

Documento de Patente 1: WO 2004/058960

Documento de Patente 2: WO 2005/045006

Documento de Patente 3: WO 2005/052153

35 Documento de Patente 4: WO 2005/052177

Documento de Patente 5: WO 2005/103260

Documento de Patente 6: WO 2006/001379

Documento de Patente 7: WO 2006/001381

Documento de Patente 8: WO 2006/001382

40 Documento de Patente 9: WO 97/23597 A2

Documento de Patente 10: US 2003/0113899 A1

Documento de Patente 11: JP-A-2000-116390

Documento de Patente 12: JP-A-2000-189177

Documento de Patente 13: JP-A-2005-287333

45 Documento de Patente 14: JP-A-2005-237379

Documento no patente 1: J. Bacteriol., 187, 3465-3470 (2005)

Documento no patente 2: Gene, 127, 117-120 (1993)

Documento no patente 3: J. Antimicrob. Chemother., 51, 545-56 (2003)

Documento no patente 4: J. Bacteriol., 183, 5803-5812 (2001)

Descripción de la invención

5 Problemas a resolver con la invención

La presente invención tiene el objetivo de proporcionar un microorganismo que produzca de manera eficiente un dipéptido, y un proceso para producir un dipéptido usando el microorganismo.

Medios para resolver los problemas

La presente invención se refiere a los siguientes puntos (1) – (6).

- 10 (1) Un microorganismo procariótico que sobreexpresa cualquiera entre ligasa de aminoácido L, prolina imino peptidasa y amida hidrolasa de aminoácido L, donde dicha ligasa de aminoácido L, dicha prolina imino peptidasa y dicha amida hidrolasa de aminoácido L presentan actividad de síntesis de dipéptidos y
- 15 en el que la actividad de una proteína para transportar un dipéptido de una célula microbiana al exterior de la célula microbiana (denominada a partir de este punto actividad de transporte de dipéptidos) es mayor que la de la cepa original, mediante la transformación de la cepa original con un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos,
- donde dicha proteína que tiene actividad de transporte de dipéptidos se describe en cualquiera de los siguientes puntos [1] a [3]:
- 20 [1] una proteína que tiene una secuencia de aminoácidos como la mostrada por cualquiera de las SEQ ID Nos: 6 a 10,
- [2] una proteína que consiste en una secuencia de aminoácidos en la que se han eliminado, sustituido o añadido de 1 a 20 aminoácidos a la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID Nos: 6 a 10, y que tiene actividad de transporte de dipéptidos,
- 25 [3] una proteína que tiene una secuencia de aminoácidos que tiene un 80% o más de homología con la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID Nos: 6 a 10, y que tiene actividad de transporte de aminoácidos.
- (2) El microorganismo de (1), que se obtiene transformando la cepa original con un ADN descrito en cualquiera de los siguientes puntos [1] a [3]:
- [1] un ADN que codifica una proteína de cualquiera de los puntos [1] a [3] de la reivindicación 1,
- 30 [2] un ADN que tiene la secuencia de nucleótidos de la región codificadora tal como se muestra en cualquiera de las SEQ ID Nos: 1 a 5,
- [3] un ADN que se hibrida con un ADN que consiste en una secuencia de nucleótidos que es complementaria a una secuencia de nucleótidos de una región codificadora de la secuencia de nucleótidos mostrada en cualquiera de las SEQ ID Nos: 1 a 5 en condiciones severas, y que codifica la proteína que
- 35 tiene actividad de transporte.
- (3) El microorganismo de (1) ó (2), que pertenece al género Escherichia, Corynebacterium, Bacillus, Serratia, Pseudomonas o Streptomyces.
- (4) Un proceso para producir un dipéptido que comprende:
- 40 permitir que un cultivo o un cultivo tratado del microorganismo de uno cualquiera de los puntos (1) a (3) y un aminoácido, éster de aminoácido o amida de aminoácido, esté presente en un medio acuoso, permitiendo que el dipéptido se forme y se acumule en el medio acuoso, y recuperar el dipéptido del medio acuoso.
- (5) Un proceso para producir un dipéptido que comprende:
- 45 cultivar el microorganismo de (1) ó (2) y que tiene la capacidad de producir al menos un tipo de aminoácido diferente a los aminoácidos que constituyen el dipéptido en un medio, permitir que se forme el dipéptido y que se acumule en el medio, y recuperar el dipéptido del cultivo.

- (6) El proceso de (4) ó (5), donde el microorganismo pertenece al género *Escherichia*, *Corynebacterium*, *Bacillus*, *Serratia*, *Pseudomonas* o *Streptomyces*.

Efecto de la Invención

- 5 Según la presente invención, potenciando la actividad de una proteína que tiene una actividad de transporte de dipéptidos en una célula microbiana de un microorganismo hacia el exterior de la célula microbiana, el dipéptido puede producirse de forma eficiente usando dicho microorganismo.

Mejor método para llevar a cabo la invención

1. Microorganismo de la presente invención

(1) Microorganismo procariótico en el que la actividad de transporte de dipéptidos es superior a la de la cepa original.

- 10 Los ejemplos del microorganismo en el que la actividad de una proteína que tiene actividad de transporte de dipéptidos es superior a la de la cepa original incluyen (a) en el contexto de la presente invención se describe i) un microorganismo en el que la actividad específica de la proteína se ve mejorada en comparación con la cepa original, e ii) un microorganismo que muestra un mayor producción de la proteína que tiene actividad de transporte de dipéptidos en comparación con la cepa original, que se obtiene modificando un gen del ADN cromosomal de la cepa original, y que codifica la proteína que tiene actividad de transporte de dipéptidos, e incluye (b) un microorganismo obtenido transformando la cepa original con un ADN que codifica una proteína que tienen actividad de transporte de dipéptidos. La cepa original de la presente especificación es la cepa original que es una diana para modificación o transformación, y puede ser una cepa natural o un mutante. Los ejemplos de cepa original incluyen, cuando el microorganismo es *Escherichia coli*, una cepa natural tal como la cepa *E. coli* K-12, la cepa B y la cepa B/r, o un mutante de las mismas. Los ejemplos de mutantes incluyen *E. coli* XL1-Blue, *E. coli* XL2-Blue, *E. coli* DH1, *E. coli* MC1000, *E. coli* ATCC 12435, *E. coli* W1485, *E. coli* JM109, *E. coli* HB101, *E. coli* N° 49, *E. coli* W3110, *E. coli* NY49, *E. coli* MP347, *E. coli* NM522, *E. coli* BL21, *E. coli* ME8415 y otros similares.

Los ejemplos de proteína que tiene actividad transportadora de dipéptidos incluyen una proteína de cualquiera de los siguientes [1] – [3],

- 25 [1] una proteína que tiene una secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10,
[2] una proteína que consiste en una secuencia de aminoácidos en la que se han eliminado, sustituido o añadido entre uno y 20 aminoácidos a la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10, y que tiene actividad de transporte de dipéptidos, y
30 [3] una proteína que tiene una secuencia de aminoácidos que tiene un 80% o más de homología con respecto a la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10 y que tiene actividad de transporte de dipéptidos.

- La proteína anterior que consiste en una secuencia de aminoácidos en la que se han eliminado, sustituido o añadido entre uno y 20 aminoácidos y que tiene actividad de transporte de dipéptidos, se puede obtener, por ejemplo, introduciendo una mutación sito-dirigida en el ADN que codifica una proteína que consiste en la secuencia de aminoácidos de cualquiera de las SEQ ID NO: 6 a 10 mediante mutagénesis sito-dirigida como se describe en
35 “Molecular Cloning, A Laboratory Manual”, Tercera Edición, Cold Spring Harbor Laboratory Press (1989) (referido a partir de este punto como “Molecular Cloning”, tercera edición); en “Current Protocols in Molecular Biology”, John Wiley & Sons (1987-1997) (referido a partir de este punto como “Current Protocols in Molecular Biology”); en *Nucleic Acids Research*, 10, 6487 (1982); *Proc. Natl. Acad. Sci. USA*, 79, 6409 (1982); *Gene*, 34, 315 (1985); *Nucleic Acids Research*, 13, 4431 (1985); *Proc. Natl. Acad. Sci. USA*, 82, 488 (1985), etc.

El número de residuos de aminoácido que se eliminan, sustituyen o añaden se encuentra dentro del rango en el que la eliminación, sustitución o adición es posible empleando métodos conocidos tales como la mencionada mutagénesis sito-dirigida. El número preferiblemente está entre 1 y 20, más preferiblemente entre 1 y 10, aún más preferiblemente entre 1 y 5.

- 45 La expresión “uno o más residuos de aminoácido son eliminados, sustituidos o añadidos en la secuencia de aminoácidos mostrada en las SEQ ID NO: 6 a 10” significa que la secuencia de aminoácidos puede contener la eliminación, sustitución o adición de un único residuo de aminoácido o de varios residuos de aminoácido en una posición arbitraria de las mismas.

- 50 La eliminación o adición de residuos de aminoácido puede darse, por ejemplo, a entre uno y 10 residuos de aminoácido del extremo N o el extremo C de la secuencia de aminoácidos de cualquiera de las SEQ ID NO: 6 a 10.

La eliminación, sustitución y adición pueden estar contenidas simultáneamente en una secuencia, y los aminoácidos que van a sustituirse o añadirse pueden ser naturales o no. Los ejemplos de aminoácidos naturales son L-alanina, L-

asparagina, ácido L-aspártico, L-arginina, L-glutamina, ácido L-glutámico, glicina, L-histidina, L-isoleucina, L-leucina, L-lisina, L-metionina, L-fenilalanina, L-prolina, L-serina, L-treonina, L-triptófano, L-tirosina, L-valina y L-cisteína.

A continuación se presentan ejemplos de los aminoácidos capaces de sustitución mutua. Los aminoácidos del mismo grupo pueden sustituirse mutuamente.

- 5 Grupo A: leucina, isoleucina, norleucina, valina, norvalina, alanina, ácido 2-aminobutanoico, metionina, O-metilserina, t-butilglicina, t-butilalanina, ciclohexilalanina.

Grupo B: ácido aspártico, ácido glutámico, ácido isoaspártico, ácido isoglutámico, ácido 2-aminoadípico, ácido 2-aminosubérico.

Grupo C: asparagina, glutamina.

- 10 Grupo D: lisina, arginina, ornitina, ácido 2,4-diaminobutanoico, ácido 2,3-diaminopropiónico.

Grupo E: prolina, 3-hidroxiprolina, 4-hidroxiprolina.

Grupo F: serina, treonina, homoserina.

Grupo G: fenilalanina, tirosina.

- 15 Adicionalmente, los ejemplos de proteínas que tienen actividad de transporte de dipéptidos incluyen una proteína que consiste en la secuencia de aminoácidos que tiene un 80% o más, preferiblemente un 90% o más, más preferiblemente un 95% o más, aún más preferiblemente un 97% o más, de forma particularmente preferible un 98% o más, lo más preferiblemente un 99% o más de homología con respecto a una secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10, y una proteína que tiene actividad de transporte de dipéptidos.

- 20 La homología entre secuencias de aminoácidos y secuencias de nucleótidos se puede determinar usando el algoritmo BLAST de Karlin y Altschul [Pro. Natl. Acad. Sci. USA, 90, 5873 (1993)] y FASTA [Methods Enzymol., 183, 63 (1990)]. En base al algoritmo BLAST, se han desarrollado programas tales como BLASTN y BLASTX [J. Mol. Biol., 215, 403 (1990)]. Cuando se analiza una secuencia de nucleótidos mediante BLASTN en base a BLAST, los parámetros, por ejemplo, son los siguientes: puntuación = 100 y longitud de palabra = 12. Cuando se analiza una secuencia de aminoácidos mediante BLASTX en base a BLAST, los parámetros, por ejemplo, son los siguientes:
25 puntuación = 50 y longitud de palabra = 3. Cuando se usan programas BLAST y BLAST con huecos ("gapped"), se emplean los parámetros por defecto de cada programa. Las técnicas específicas para estos análisis son bien conocidas.

- 30 El hecho de que la proteína que consiste en una secuencia de aminoácidos, donde se eliminan, sustituyen o añaden entre uno y 20 aminoácidos de la secuencia de aminoácidos mostrada por las SEQ ID NO: 6 a 10, tenga actividad de transporte de dipéptidos se puede confirmar, por ejemplo, preparando un transformante que tenga una actividad de proteína superior a la cepa original mediante la transformación de la cepa original con un ADN que codifica la proteína cuya actividad se quiere confirmar, añadiendo el cultivo de la cepa original o de la cepa transformada y los aminoácidos que constituyen el dipéptido objetivo a una disolución tampón, y comparando las cantidades de dipéptido formado y acumulado en el medio acuoso.

- 35 Los ejemplos de microorganismos de los mencionados anteriormente en (a) i), que tienen una actividad específica mejorada de la proteína que tiene actividad de transporte de dipéptidos en comparación con la cepa original, incluyen un microorganismo que contenga una proteína mutante con actividad mejorada con respecto a la proteína con actividad de transporte de dipéptidos de la cepa original, ya que la proteína tiene una secuencia de aminoácidos igual en la que entre uno y 20 aminoácidos, preferiblemente 1 – 10 aminoácidos, más preferiblemente 1 – 5
40 aminoácidos, aún más preferiblemente 1 – 3 aminoácidos, han sido sustituidos en la secuencia de aminoácidos de la proteína de la cepa original.

- Se describen ejemplos de microorganismos con una producción mejorada de la proteína que tiene actividad de transporte de dipéptidos, con respecto a la cepa original, de los mencionados anteriormente en (a) ii), que incluyen un microorganismo que muestra una mejora en la cantidad producida de la proteína con respecto a la proteína que tiene actividad transportadora de dipéptidos de la cepa original, ya que tiene una región promotora en la que se ha sustituido no menos de una base, preferiblemente entre 1 y 10 bases, más preferiblemente entre 1 y 5 bases, aún más preferiblemente entre 1 y 3 bases, de la secuencia de nucleótidos de la región reguladora de la transcripción o de la región promotora del gen que codifica la proteína, que está presente en el ADN cromosomal de la cepa original.

- 50 Los ejemplos de microorganismos obtenidos al transformar la cepa original de los mencionados anteriormente en (b) con un ADN que codifica la proteína que tiene actividad de transporte de dipéptidos incluyen un microorganismo obtenido mediante la transformación de la cepa original con

[4] un ADN que codifica una proteína de cualquiera de los mencionados anteriormente en [1] a [3],

[5] un ADN que tiene una secuencia de nucleótidos de la región codificadora en la secuencia de nucleótidos mostrada por cualquiera de las SEQ ID NO: 1 a 5, o

5 [6] un ADN que se hibrida con un ADN que consiste en una secuencia de nucleótidos complementaria a una secuencia de nucleótidos de una región codificadora de la secuencia de nucleótidos mostrada por cualquiera de las SEQ ID NO: 1 a 5 en condiciones severas, y que codifica la proteína que tiene actividad de transporte de dipéptidos.

Los ejemplos de microorganismos incluyen i) un microorganismo que tiene un ADN exógeno que codifica una proteína que tiene actividad de transporte de dipéptidos en el ADN cromosomal y ii) un microorganismo que tiene un ADN exógeno que codifica una proteína que tiene actividad de transporte de dipéptidos fuera del cromosoma. En otras palabras, el microorganismo de i) significa que la cepa original no tiene un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos y uno o más ADNs recién introducidos están presentes en el ADN cromosomal, o la cepa original tiene intrínsecamente un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos, lo que incluye el ADN recién introducido, están presentes en el ADN cromosomal. El microorganismo de ii) es un microorganismo que tiene un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos el ADN plásmido.

15 "Hibridar" como se ha mencionado anteriormente se refiere a una etapa de hibridación de ADN con ADN que tiene una secuencia de nucleótidos específica o una parte del ADN puede usarse como sonda para análisis Northern o Southern blot como un cebador de oligonucleótido para análisis de PCR. Los ADNs usados como sonda incluyen ADNs que comprenden al menos 100 o más bases, preferiblemente 200 o más bases, más preferiblemente 500 o más bases, y los ADNs usados como cebador incluyen ADNs que comprenden al menos 10 o más bases, preferiblemente 15 o más bases.

Los métodos de los experimentos de hibridación de ADN son bien conocidos y, por ejemplo, los especialistas en la técnica pueden determinar las condiciones de hibridación según la descripción de la especificación de la presente solicitud. Se pueden emplear condiciones de hibridación acordes a muchos libros de texto, que incluyen la descripción de "Molecular Cloning", Segunda Edición, Tercera Edición (2001), "Methods for General and Molecular Bacteriology", ASM Press (1994) e "Immunology Methods Manual", Academic Press (Molecular).

Por ejemplo, la hibridación en condiciones severas mencionada anteriormente se lleva a cabo como se indica a continuación. Se incuban un filtro y un ADN sonda en una disolución que comprende un 50% de formamida, 5xSSC (750 mmol/l de cloruro sódico y 75 mmol/l de citrato sódico), 50 mmol/l de fosfato sódico (pH 7,6), 5xdisolución de Denhardt, un 10% de sulfato de dextrano y 20 µg/L de ADN de esperma de salmón desnaturalizado a 42°C durante una noche, y tras la incubación, por ejemplo, el filtro se lava preferiblemente en disolución 0,2xSSC a aproximadamente 65°C. También se pueden emplear condiciones menos severas. La modificación de las condiciones severas se puede realizar ajustando la concentración de formamida (las condiciones se hacen menos severas al reducir la concentración de formamida) y cambiando las condiciones de concentración salina y temperatura. La hibridación en condiciones menos severas se lleva a cabo, por ejemplo, incubando un filtro con ADN inmovilizado sobre él y una sonda de ADN en una disolución que comprende 6xSSCE (20xSSCE: 3 mol/L de cloruro sódico, 0,2 mol/L de dihidrogenofosfato sódico y 0,02 mol/L de EDTA, pH 7,4), 0,5% de SDS, 30% de formamida y 100 µg/L de ADN de esperma de salmón desnaturalizado a 37°C durante una noche, y lavando el filtro con disolución 1xSSC que contiene 0,1% de SDS a 50°C. La hibridación en condiciones aún menos severas se lleva a cabo mediante hibridación en las condiciones menos severas anteriores usando una disolución que tenga una elevada concentración salina (por ejemplo, 5xSSC), y lavando el filtro.

También se pueden establecer varias condiciones descritas anteriormente añadiendo un reactivo de bloqueo usado para reducir el ruido de fondo de la hibridación o para cambiar el reactivo. La adición de dicho reactivo bloqueante puede venir acompañada de cambios en las condiciones para la hibridación para hacer que las condiciones sean más adecuadas para el propósito.

45 El anterior ADN capaz de hibridación en condiciones severas incluye ADN que tiene al menos un 90% o más, preferiblemente un 95% o más, más preferiblemente un 97% o más, aún más preferiblemente un 98% o más, de forma particularmente preferida un 99% o más, de homología con respecto al ADN que tiene la secuencia de nucleótidos de la región codificadora de la secuencia de nucleótidos de cualquiera de las SEQ ID NO: 1 a 5, calculada mediante el uso de programas tales como BLAST y FASTA descritos antes en base a los parámetros anteriores.

(2) Microorganismos que tienen la capacidad de producir una proteína que tenga actividad de síntesis de dipéptidos

El microorganismo que tiene capacidad para producir una proteína que tiene actividad de síntesis de dipéptidos no está limitado particularmente, siempre que presente dicha capacidad. Los ejemplos del mismo incluyen un microorganismo que produce una proteína que tiene actividad para sintetizar dipéptidos condensando y ligando uno o más tipos de aminoácidos [referidos a partir de este punto como ligasa de aminoácido L (EC 6.3.2.28)], un microorganismo que produce una proteína que tiene actividad de síntesis de dipéptidos a partir de un éster de aminoácido L y un aminoácido L, un microorganismo que produce una proteína que tiene actividad de síntesis de dipéptidos a partir de amida de aminoácido L y aminoácido L, y similares.

Los ejemplos del microorganismo que produce una proteína que tiene actividad de síntesis de dipéptidos condensando y ligando uno o más tipos de aminoácidos incluyen un microorganismo que produce una proteína seleccionada del grupo que consiste en NRPS, D-Ala-D-Ala ligasa y ligasa de aminoácido L.

5 Los ejemplos del microorganismo que produce NRPS incluyen procariontes que incluyen *Bacillus*, eucariontes que incluyen *Penicillium*, microorganismos que producen TycA, TycB y TycC (GenBank AF004835), un microorganismo que produce PcbAB (GenBank M57425), y un microorganismo que produce una proteína que tiene un 80% o más, preferiblemente un 90% o más, más preferiblemente un 95% o más, de homología con respecto a la secuencia de aminoácidos de una proteína seleccionada entre BacA, BacB, BacC, TycA, TycB, TycC y PcbAB y que tienen actividad de NRPS, y otros similares.

10 Los ejemplos de microorganismo que produce D-Ala-D-Ala ligasa incluyen un procarionte que forma un peptidoglicano, un microorganismo que produce DdlA (n° de acceso de GenBank M58467), un microorganismo que produce DdlC (n° de acceso de GenBank D88151), y un microorganismo que produce una proteína que tiene una secuencia de aminoácidos donde uno o más aminoácidos han sido eliminados, sustituidos o añadidos en la secuencia de aminoácidos de uno cualquiera seleccionado entre DdlA, DdlB y DdlC, y que tiene actividad de D-Ala-D-Ala ligasa, y otros similares.

15 La homología de la secuencia de aminoácidos se puede determinar usando un programa tal como los anteriormente mencionados BLAST, FASTA y similares.

20 Los ejemplos de microorganismos que producen ligasa de aminoácido L incluyen microorganismos que pertenecen al género *Bacillus*, preferiblemente *Bacillus subtilis*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, y el microorganismo que produce la proteína descrita en el documento WO 2004/058960.

Los ejemplos de ligasa de aminoácido L incluyen microorganismos capaces de producir una proteína seleccionada de entre los siguientes puntos [7] a [10]:

[7] Una proteína que consiste en una secuencia de aminoácidos mostrada en cualquiera de las SEQ ID N: 11 a 18,

25 [8] Una proteína que consiste en una secuencia de aminoácidos, donde se han eliminado, sustituido o añadido uno o más aminoácidos a la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 11 a 18, y que tiene actividad de ligasa de aminoácido L,

30 [9] Una proteína que consiste en una secuencia de aminoácidos que tiene una homología del 80% o más, preferiblemente del 90% o más, más preferiblemente del 95% o más, aún más preferiblemente del 97% o más, de forma particularmente preferida del 98% o más, lo más preferiblemente del 99% o más, con respecto a la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 11 a 18, y que tiene actividad de ligasa de aminoácido L, y

35 [10] Una proteína que tiene una secuencia de aminoácidos que tiene una homología del 80% o más, preferiblemente del 90% o más, más preferiblemente del 95% o más, aún más preferiblemente del 97% o más, de forma particularmente preferida del 98% o más, lo más preferiblemente del 99% o más, con respecto a la secuencia de aminoácidos mostrada en la SEQ ID NO: 19, y que tiene actividad de ligasa de aminoácido L.

40 Los ejemplos de microorganismos que producen una proteína que tiene una actividad de síntesis de dipéptidos a partir de un éster de aminoácido L y aminoácido L incluyen un microorganismo que produce prolina imino peptidasa; específicamente, microorganismos que pertenecen al género *Bacillus*, *Corynebacterium*, *Pseudomonas*, y similares. Específicamente, los ejemplos de microorganismo incluyen *Bacillus subtilis* ATCC6633, *Bacillus coagulans* EK01 [J. Bacteriol., 174, 7919 (1992)], *Corynebacterium glutamicum* ATCC13286, *Pseudomonas putida* AJ-2402 (FERM BP-8101), *Pseudomonas putida* ATCC12633, *Pseudomonas putida* AJ2048 (FERM BP-8123) (los microorganismos anteriores se describen en el documento WO 03/010307) y otros similares. Los ejemplos de microorganismos productores de prolina imino peptidasa incluyen *Arthrobacter nicotianae* [FEMS Microbiol. Lett., 78, 191 (1999)], *Escherichia coli* (JP-A-2-113887), *Flavobacterium meningosepticum* [Arch. Biochem. Biophys., 336, 35 (1996)], *Hafnia alvei* [J. Biochem., 119, 468 (1996)], *Lactobacillus delbrueckii* [Microbiology, 140, 527 (1994)], *Bacillus coagulans* [J. Bacteriol., 174, 7919 (1994)], *Aeromonas sobria* [J. Biochem., 116, 818 (1994)], *Xanthomonas campestris* (JP-A-9-121860), *Neisseria gonorrhoeae* [Mol. Microbiol., 9, 1203 (1993)], *Propionibacterium freudenreichii* [Appl. Environ. Microbiol., 64, 4736 (1998)], *Serratia marcescens* [J. Biochem., 122, 601 (1997)], *Corynebacterium variabilis* [J. Appl. Microbiol., 90, 449 (2001)], *Thermoplasma acidophilum* [FEBS Lett., 398, 101 (1996)], *Pseudomonas aeruginosa* [Nature, 406, 959 (2000)] y otros similares.

Además, los ejemplos de microorganismos que producen prolina imino peptidasa incluyen un microorganismo que tenga la capacidad de producir una proteína seleccionada de entre los siguientes puntos [11] – [13]:

55 [11] prolina imino peptidasa descrita en cualquiera de los siguientes documentos, WO 03/010307, FEMS Microbiol. Lett., 78, 191 (1999), JP-A-2-113887, Arch. Biochem. Biophys., 336, 35 (1996), J. Biochem., 119, 468 (1996), Microbiology, 140, 527 (1994), J. Bacteriol., 174, 7919 (1994), J. Biochem., 116, 818 (1994), JP-A-9-121860, Mol.

Microbiol., 9, 1203 (1993), Appl. Environ. Microbiol., 64, 4736 (1998), J. Biochem., 122, 601 (1997), FEBS Lett., 398, 101 (1996), y Nature, 406, 959 (2000),

5 [12] una proteína que consiste en una secuencia de aminoácidos, donde se han eliminado, sustituido o añadido uno o más aminoácidos de secuencia de aminoácidos mencionada anteriormente de prolina imino péptidos de [11], y que tiene actividad de prolina imino peptidasa, y

10 [13] una proteína que consiste en una secuencia de aminoácidos que tiene una homología del 80% o más, preferiblemente del 90% o más, más preferiblemente del 95% o más, aún más preferiblemente del 97% o más, de forma particularmente preferida del 98% o más, lo más preferiblemente del 99% o más, con respecto a la secuencia de aminoácidos de cualquiera de las prolina imino peptidasas mencionadas anteriormente en [11], y que tiene actividad de prolina imino peptidasa.

15 Los ejemplos de microorganismos que producen una proteína que tiene actividad de síntesis de dipéptidos a partir de amida de aminoácido L y de aminoácido L incluyen un microorganismo que produce hidrolasa de amida de aminoácido L; específicamente, microorganismos que pertenecen al género Bacillus, Corynebacterium, Erwinia, Rhodococcus, Chryseobacterium, Micrococcus, Pseudomonas, Cryptococcus, Trichosporon, Rhodosporidium, Sporobolomyces, Tremella, Torulaspora, Sterigmatomyces y Rhodotorula, preferiblemente un microorganismo que pertenece al género Bacillus, Corynebacterium y Pseudomonas, más preferiblemente Bacillus megaterium AJ3284 (FERM BP-8090), Corynebacterium glutamicum ATCC13286, Micrococcus luteus ATCC9341, Pseudomonas saccharophila ATCC15946 (los microorganismos anteriores se describen en el documento WO 03/010187) y similares.

20 Adicionalmente, los ejemplos de microorganismos que producen una proteína que tiene actividad de hidrolasa de amida de aminoácido L incluyen un microorganismo que tienen la capacidad de producir una proteína seleccionada de los siguientes puntos [14] – [16]:

[14] la hidrolasa de amida de aminoácido L descrita en WO 03/010187,

25 [15] una proteína que consiste en una secuencia de aminoácido donde se han eliminado, sustituido o añadido uno o más aminoácidos en la secuencia de aminoácidos de la hidrolasa de amida de aminoácido L descrita en el documento WO 03/010187, y que tiene actividad de hidrolasa de amida de aminoácido L, y

30 [16] una proteína que consiste en una secuencia de aminoácidos que tiene una homología del 80% o más, preferiblemente del 90% o más, más preferiblemente del 95% o más, aún más preferiblemente del 97% o más, de forma particularmente preferida del 98% o más, lo más preferiblemente del 99% o más, con respecto a la secuencia de aminoácidos de la hidrolasa de amida de aminoácido L descrita en el documento WO 03/010187, y que tiene actividad de hidrolasa de amida de aminoácido L.

35 En lo anterior, la proteína que tiene una secuencia de aminoácidos en la que se han eliminado, sustituido o añadido uno o más aminoácidos, y que tiene actividad de síntesis de dipéptidos, puede adquirirse siguiendo un método similar al mencionado anteriormente en (1). El número, tipo y similar de aminoácidos a eliminar, sustituir o añadir es igual a lo mencionado anteriormente en (1).

La homología de secuencias de aminoácido se puede determinar usando un programa tal como los mencionados BLAST, FASTA y similares.

40 Adicionalmente, un microorganismo que tiene un ADN recombinante obtenido ligando un ADN que codifica una proteína que tiene actividad de síntesis de dipéptidos condensando y ligando uno o más tipos de aminoácidos, un ADN que codifica una proteína que tiene actividad de síntesis de dipéptidos a partir de éster de aminoácido L y de aminoácido L, o un ADN que codifica una proteína que tiene actividad de síntesis de dipéptidos a partir de amida de aminoácido L y aminoácido L y un ADN vector también es un microorganismo que tiene capacidad para producir una proteína que tiene actividad de síntesis de dipéptidos.

45 Los ejemplos de microorganismos incluyen microorganismos que pertenecen al género Escherichia, Bacillus, Corynebacterium o Saccharomyces.

Los ejemplos de ADN que codifica una proteína que tiene actividad de síntesis de dipéptidos condensando y ligando uno o más tipos de aminoácidos incluyen ADNs que codifican NRPS, D-Ala-D-Ala ligasa, ligasa de aminoácido L, y otros similares.

50 Los ejemplos de ADN que codifica NRPS incluyen ADN que codifica una proteína seleccionada del grupo que consiste en BacA, BacB, BacC, TycA, TycB, TycC y PcbAB.

Los ejemplos de ADN que codifica D-Ala-D-Ala ligasa incluyen un ADN que codifica una proteína seleccionada del grupo que consiste en DdIA, DdIB y DdIC.

Los ejemplos de ADN que codifican ligasa de aminoácido L incluyen ADN seleccionado de los siguientes puntos [17] – [20]:

[17] un ADN que codifica la ligasa de aminoácido L descrita en cualquiera de los puntos mencionados anteriormente [7] a [10],

5 [18] un ADN que tiene una secuencia de nucleótidos mostrada por cualquiera de las SEQ ID NO: 20 a 28,

[19] un ADN que se hibrida con un ADN que tiene una secuencia de nucleótidos complementaria a una secuencia de nucleótidos mostrada por cualquiera de las SEQ ID NO: 20 a 28 en condiciones severas, y que codifica una proteína que tiene actividad de ligasa de aminoácido L, y

10 [20] un ADN que tiene una secuencia de nucleótidos que tiene una homología del 80% o más, preferiblemente del 90% o más, más preferiblemente del 95% o más, aún más preferiblemente del 97% o más, de forma particularmente preferida del 98% o más, lo más preferiblemente del 99% o más, con respecto a la secuencia de nucleótidos mostrada por la SEQ ID NO: 29, y que codifica una proteína que tiene actividad de ligasa de aminoácido L.

Los ejemplos de ADN que codifican una proteína que tiene actividad de síntesis de dipéptidos a partir de éster de aminoácido L y aminoácido L incluyen un ADN seleccionado de los siguientes apartados [21] – [23]:

15 [21] un ADN que codifica la prolina imino peptidasa descrita en cualquiera de los puntos mencionados anteriormente [11] a [13],

[22] un ADN que codifica prolina imino peptidasa que tiene la secuencia de nucleótidos descrita en cualquiera de WO 03/010307, FEMS Microbiol. Lett., 78, 191 (1999), JP-A-2-113887, Arch. Biochem. Biophys., 336, 35 (1996), J. Biochem., 119, 468 (1996), Microbiology, 140, 527 (1994), J. Bacteriol., 174, 7919 (1994), J. Biochem., 116, 818 (1994), JP-A-9-121860, Mol. Microbiol., 9, 1203 (1993), Appl. Environ. Microbiol., 64, 4736 (1998), J. Biochem., 122, 601 (1997), FEBS Lett., 398, 101 (1996), y Nature, 406, 959 (2000), y

20 [23] un ADN que se hibrida con un ADN de cadena complementaria del ADN que codifica prolina imino peptidasa de cualquiera de las mencionadas anteriormente en [21] en condiciones severas, y que codifica una proteína que tiene actividad de prolina imino peptidasa.

25 Los ejemplos de ADN que codifican una proteína que tiene actividad de síntesis de dipéptidos a partir de amida de aminoácido L y de aminoácido L incluyen ADN seleccionado de los siguientes apartados [24] – [26]:

[24] un ADN que codifica la hidrolasa de amida de aminoácido L de cualquiera de los apartados mencionados anteriormente [14] – [16],

30 [25] un ADN que codifica una hidrolasa de amida de aminoácido L que tiene la secuencia de nucleótidos descrita en el documento WO 03/010187, y

[26] un ADN que se hibrida con una cadena complementaria de un ADN que codifica una hidrolasa de amida de aminoácido L que tiene la secuencia de nucleótidos descrita en el documento WO 03/010187 en condiciones severas, y que codifica una proteína que tiene actividad de hidrolasa de amida de aminoácido L.

35 El ADN que puede hibridarse en condiciones severas es el mismo de la definición del anteriormente mencionado (1). El hecho de que el ADN que se hibrida con el ADN mencionado anteriormente en condiciones severas sea un ADN que codifica una proteína que tiene actividad de síntesis de dipéptidos se puede confirmar empleando un método que incluye la preparación de un ADN recombinante que expresa el ADN, usando un microorganismo obtenido introduciendo el ADN recombinante en una célula hospedante como fuente enzimática, 1) añadiendo la fuente de enzima y uno o más tipos de aminoácidos a un medio acuoso, y analizando si se forma y acumula un dipéptido o no en el medio acuoso mediante HPLC o similar, 2) añadiendo la fuente de enzimas, éster de aminoácido L y aminoácido L en un medio acuoso, y analizando si se forma y se acumula un dipéptido o no en el medio acuoso mediante HPLC o similar, o 3) añadiendo la fuente de enzimas, amida de aminoácido L y aminoácido L en un medio acuoso, y analizando si se forma y se acumula un dipéptido en el medio acuoso mediante HPLC o similar.

45 La homología de la secuencia de nucleótidos se puede determinar usando un programa como los mencionados anteriormente BLAST, FASTA y similares.

2. Preparación de un microorganismo procariótico usado en la presente invención

(1) Preparación de un microorganismo en el que una proteína que tiene actividad de transporte de dipéptidos presenta una actividad mayor que la de la cepa original

50 De los microorganismos en los que una proteína que tiene actividad de transporte de dipéptidos presenta mayor actividad que la de la cepa original, en la presente memoria se describe un microorganismo en el que la actividad específica es mayor que la de la proteína que tiene actividad de transporte de dipéptidos de la cepa original, que puede obtenerse sometiendo un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos a un

tratamiento de mutación in vitro con un mutágeno, o a PCR con tendencia a error y otros similares para introducir una mutación en el ADN, sustituyendo un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos antes de la introducción de la mutación, que está presente en el ADN cromosomal de la cepa original, con el ADN mutante por un método conocido [Proc. Natl. Acad. Sci. U.S.A., 97, 6640 (2000)] para proporcionar una cepa modificada que exprese el ADN mutante, y comparar la actividad de transporte de dipéptidos de la cepa original y de la cepa modificada de acuerdo a los métodos mencionados anteriormente.

Adicionalmente, de los microorganismos en los que la proteína que tiene actividad de transporte de dipéptidos muestra una actividad superior a la de la cepa original, en la presente memoria se describe un microorganismo en el que la cantidad de proteína producida es superior a la de la cepa original, lo que puede confirmarse mediante un método que incluye someter un ADN que tiene una región reguladora de la transcripción y una región promotora del gen que codifica la proteína que tiene actividad de transporte de dipéptidos de la cepa original, por ejemplo, la secuencia de nucleótidos de 200 pb, preferiblemente 100 pb, por encima del codón de inicio de la proteína, a un tratamiento de mutación in vitro, o a PCR con tendencia a error o similar, para introducir una mutación en el ADN, sustituyendo la región reguladora de la transcripción y la región promotora del gen que codifica la proteína que tiene actividad de transporte de dipéptidos antes de la introducción de la mutación, que está presente en el ADN cromosomal de la cepa original, con el ADN mutante mediante un método conocido [Proc. Natl. Acad. Sci. U.S.A., 97, 6640 (2000)] para proporcionar una cepa modificada que tenga una región reguladora de la transcripción mutante o una región promotora mutante, y comparar las cantidades de transcrito de los genes que codifican la proteína que tiene actividad de transporte de dipéptidos de la cepa original y de la cepa modificada mediante RT-PCR, hibridación Northern y similares, o comparando las cantidades producidas de proteína con actividad de transporte de dipéptidos de la cepa original y de la cepa modificada mediante SDS-PAGE y similares.

Adicionalmente, en la presente memoria se describe que un microorganismo en el que se ha promocionado el nivel de producción de la proteína que tiene actividad de transporte de dipéptidos en comparación con la cepa original también se puede obtener sustituyendo la región promotora de un gen que codifica la proteína que tiene actividad de transporte de dipéptidos de la cepa original con una secuencia promotora fuerte conocida.

Los ejemplos de dicho promotor incluyen promotores derivados de *Escherichia coli*, fagos y similares, que son funcionales en *E. coli*, tal como el promotor *trp* (P_{trp}), el promotor *lac* (P_{lac}), el promotor P_L , el promotor P_R , el promotor P_{SE} y otros similares, el promotor SPO1, el promotor SPO2, el promotor *penP* y otros similares. Adicionalmente, también se pueden mencionar promotores construidos artificialmente tales como un promotor que tenga dos P_{trp} conectados en tándem, el promotor *tac*, el promotor *lacT7*, el promotor *let I* y otros similares.

A continuación se explica en detalle el método de preparación de un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos, y un método de preparación de un microorganismo obtenido mediante la transformación de la cepa original con el ADN.

(a) Preparación de un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos

Un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos se puede obtener, por ejemplo, mediante hibridación Southern de una biblioteca de ADN cromosomal de un microorganismo tal como *E. coli* y similar, usando un ADN sonda que puede diseñarse en base a la secuencia de nucleótidos mostrada por cualquiera de las SEQ ID NO: 6 a 10, o mediante PCR [PCR Protocols, Academic Press (1990)] usando un ADN cebador que puede diseñarse en base a la secuencia de nucleótidos y el ADN cromosomal de un microorganismo, preferiblemente *E. coli*, como plantilla.

Adicionalmente, es posible buscar en varias bases de datos de secuencias génicas una secuencia que tenga una homología del 80% o más, preferiblemente del 90% o más, más preferiblemente del 95% o más, aún más preferiblemente del 97% o más, lo más preferiblemente del 99% o más, con respecto a la secuencia de nucleótidos de un ADN que codifica una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10, y, en base a la secuencia de nucleótidos obtenida mediante dicha búsqueda, obtener un ADN que codifique una proteína que tenga actividad de transporte de dipéptidos de acuerdo a los métodos mencionados anteriormente a partir de ADN cromosomal, una biblioteca de ADNc, y otros similares, del microorganismo que tiene la secuencia de nucleótidos.

La secuencia de nucleótidos del ADN puede determinarse incorporando el ADN obtenido directamente o tras realizar la ruptura con una enzima de restricción adecuada, y similar, en un vector de acuerdo a un método convencional, e introduciendo el ADN recombinante obtenido en una célula hospedante, y analizando la secuencia mediante un método de análisis de secuencia de nucleótidos usado de manera general, por ejemplo, el método dideoxi [Proc. Natl. Acad. Sci., USA, 74, 5463 (1977)], o usando un aparato de análisis de secuencia de nucleótidos tal como el analizador 3700 DNA (fabricado por Applied Biosystems) y otros similares.

Como vector se puede mencionar pBluescriptII KS (+) fabricado por Stratagene), pDIRECT [Nucleic Acids Res., 18, 6069 (1990)], pCR-Script Amp SK (+) (fabricado por Stratagene), pT7Blue (fabricado por Novagen), pCR II (fabricado por Invitrogen) y pCR-TRAP (fabricado por GenHunter Corporation) y similares.

Como célula hospedante, se pueden mencionar los microorganismos que pertenecen al género *Escherichia* y similar. Los ejemplos de microorganismos que pertenecen al género *Escherichia* incluyen *E. coli* XL1-Blue, *E. coli* XL2-Blue, *E. coli* DH1, *E. coli* MC1000, *E. coli* ATCC 12435, *E. coli* W1485, *E. coli* JM109, *E. coli* HB101, *E. coli* No.49, *E. coli* W3110, *E. coli* NY49, *E. coli* MP347, *E. coli* NM522, *E. coli* BL21, *E. coli* ME8415 y similares.

- 5 Para la introducción de ADN recombinante, se puede usar cualquier método siempre que el ADN sea introducido en la célula hospedante mencionada anteriormente. Por ejemplo, se puede mencionar un método que usa ion calcio [Proc. Natl. Acad. Sci., USA, 69, 2110 (1972)], el método de protoplasto (JP-A-63-248394), el método de electroporación [Nucleic Acids Res., 16, 6127 (1988)] y similares.

- 10 Como resultado de la determinación de la secuencia de nucleótidos, cuando es ADN obtenido es un ADN parcial, se puede obtener el ADN de longitud completa sometiendo a una biblioteca de ADN cromosomal a un método de hibridación Southern usando el ADN parcial como sonda, y similares.

Además, el ADN objeto se puede preparar en base a la secuencia de nucleótidos de ADN determinada mediante síntesis química usando un sintetizador de ADN de tipo 8905 fabricado por Perceptive Biosystems y similar.

- 15 Los ejemplos de ADN obtenido como se ha mencionado anteriormente incluyen un ADN que codifica una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10 y un ADN que tiene una región codificadora en la secuencia de nucleótidos mostrada en cualquiera de las SEQ ID NO: 1 a 5.

(b) Preparación del microorganismo transformado con vector de plásmido que expresa la proteína que tiene actividad de transporte de dipéptidos

- 20 En base al ADN que codifica una proteína que tiene actividad de transporte de dipéptidos, que se obtiene mediante el método del aparato (a) anterior, se prepara si es necesario un fragmento de ADN con una longitud adecuada y que contiene una parte que codifica la proteína que tiene actividad de transporte de dipéptidos. Adicionalmente, se puede obtener un transformante que muestra un aumento de la cantidad de proteína sustituyendo la base de la secuencia de nucleótidos de la parte que codifica la proteína que tiene actividad de transporte de dipéptidos, de tal modo que se pueda obtener un codón óptimo para la expresión en la célula hospedante.

- 25 El ADN recombinante se prepara insertando el fragmento de ADN por debajo de un promotor de un vector de expresión adecuado.

Al introducir el ADN recombinante en una célula hospedante adecuada para el vector de expresión, se puede obtener un transformante en el que la actividad de una proteína que tenga actividad de transporte de dipéptidos mejora con respecto a la de la célula hospedante, a saber, la cepa original.

- 30 Como célula hospedante, se pueden usar microorganismos procarióticos, más preferiblemente bacterias, más preferiblemente microorganismos que pertenezcan al género *Escherichia*, lo más preferiblemente *E. coli*.

Los vectores de expresión que pueden emplearse son aquellos capaces de replicación o integración autónomas en el cromosoma de las células hospedantes anteriores y que contienen un promotor en una posición apropiada para la transcripción del ADN que codifica la proteína que tiene actividad de transporte de dipéptidos.

- 35 Cuando se usa un procarionte como célula hospedante, se prefiere que el ADN recombinante que comprende el ADN que sea capaz de replicación autónoma en el procarionte y que comprenda un promotor, una secuencia de unión a ribosoma, el ADN que codifica la proteína que tiene actividad de transporte de dipéptidos y una secuencia de terminación de la transcripción. El ADN recombinante puede comprender además un gen que regule el promotor.

- 40 Los ejemplos de vectores de expresión son pColdI (fabricado por Takara Bio Inc.), pCDF-1b, pRSF-1b (ambos fabricados por Novagen, Inc.), pMAL-c2x (fabricado por New England Biolabs), pGEX-4T-1 (fabricado por GE Healthcare Bioscience), pTrcHis (fabricado por Invitrogen), pSE280 (fabricado por Invitrogen), pGEMEX-1 (fabricado por Promega Corp.), pQE-30 (fabricado por Qiagen, Inc.), pET-3 (fabricado por Novagen, Inc.), pKYP10 (JP-A-58-110600), pKYP200 [Agric. Biol. Chem., 48, 669 (1984)], pLSA1 [Agric. Biol. Chem., 53, 277 (1989)], pGEL1 [Proc. Natl. Acad. Sci., USA, 82, 4306 (1985)], pBluescriptII SK(+), pBluescript II KS(-) (fabricado por Stratagene), pTrs30 [preparado a partir de *Escherichia coli* JM109/pTrs30 (FERM BP-5407)], pTrs32 [preparado a partir de *Escherichia coli* JM109/pTrs32 (FERM BP-5408)], pPAC31 (WO 98/12343), pUC19 [Gene, 33, 103 (1985)], pSTV28 (fabricado por Takara Bio Inc.), pUC118 (fabricado por Takara Bio Inc.), pPA1 (JP-A-63-233798) y otros similares.

- 50 Como promotor, se puede usar cualquier promotor capaz de funcionar en células hospedantes tales como *E. coli*. Por ejemplo, también se pueden usar promotores derivados de *E. coli*, fagos y otros similares, tales como el promotor *trp* (P_{trp}), el promotor *lac* (P_{lac}), el promotor P_L , el promotor P_R , el promotor P_{SE} , el promotor SPO1, el promotor SPO2, el promotor *penP*. También se pueden usar promotores modificados y diseñados artificialmente tales como un promotor que tenga dos P_{trp} combinados en tándem, el promotor *tac*, el promotor *lacT7* y el promotor *let I*, etc.

También se pueden usar promotores tales como el promotor *xylA* para la expresión en microorganismos que pertenezcan al género *Bacillus* [Appl. Microbiol. Biotechnol., 35, 594-599 (1991)] y el promotor P54-6 para la expresión en microorganismos que pertenezcan al género *Corynebacterium* [Appl. Microbiol. Biotechnol., 53, 674-679 (2000)].

- 5 Se prefiere usar un plásmido en el que la distancia entre la secuencia de Shine-Dalgarno (secuencia de unión de ribosomas) y el codón de inicio se ajuste a una longitud apropiada (p.ej., de 6 a 18 bases).

En un ADN recombinante en el que un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos se une a un vector de expresión, no siempre es necesaria una secuencia terminadora de la transcripción. Sin embargo, es preferible disponer la secuencia terminadora de la transcripción inmediatamente por debajo del gen estructural.

10

Los ejemplos de dicho ADN recombinante incluyen pSbcr, pSnorE, pSydeE, pSemrD y pSyeeO, mostrados más adelante.

Como hospedante de ADN recombinante, se pueden mencionar los procariontes, más preferiblemente bacterias.

Los ejemplos de procariontes incluyen microorganismos que pertenecen a los géneros *Escherichia*, *Serratia*, *Bacillus*, *Brevibacterium*, *Corynebacterium*, *Microbacterium*, *Pseudomonas*, *Agrobacterium*, *Alicyclobacillus*, *Anabaena*, *Anacystis*, *Arthrobacter*, *Azotobacter*, *Chromatium*, *Erwinia*, *Methylobacterium*, *Phormidium*, *Rhodobacter*, *Rhodopseudomonas*, *Rhodospirillum*, *Scenedesmus*, *Streptomyces*, *Synechococcus*, *Zymomonas* y otros similares. Por ejemplo, se pueden mencionar los siguientes, *Escherichia coli*, *Bacillus subtilis*, *Bacillus megaterium*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus pumilus*, *Brevibacterium ammoniagenes*, *Brevibacterium immariophilum*, *Brevibacterium saccharolyticum*, *Brevibacterium flavum*, *Brevibacterium lactofermentum*, *Corynebacterium glutamicum*, *Corynebacterium acetoacidophilum*, *Microbacterium ammoniaphilum*, *Serratia ficaria*, *Serratia fonticola*, *Serratia liquefaciens*, *Serratia marcescens*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Anabaena cylindrica*, *Anabaena doliolum*, *Anabaena flos-aquae*, *Arthrobacter aureus*, *Arthrobacter citreus*, *Arthrobacter globiformis*, *Arthrobacter hydrocarboglutamicus*, *Arthrobacter mysorens*, *Arthrobacter nicotianae*, *Arthrobacter paraffineus*, *Arthrobacter protophormiae*, *Arthrobacter roseoparaffinus*, *Arthrobacter sulfureus*, *Arthrobacter ureafaciens*, *Chromatium buderi*, *Chromatium tepidum*, *Chromatium vinosum*, *Chromatium warmingii*, *Chromatium fluviatile*, *Erwinia uredovora*, *Erwinia carotovora*, *Erwinia ananas*, *Erwinia herbicola*, *Erwinia punctata*, *Erwinia terreus*, *Methylobacterium rhodesianum*, *Methylobacterium extorquens*, *Phormidium sp. ATCC29409*, *Rhodobacter capsulatus*, *Rhodobacter sphaeroides*, *Rhodopseudomonas blastica*, *Rhodopseudomonas marina*, *Rhodopseudomonas palustris*, *Rhodospirillum rubrum*, *Rhodospirillum salalexigens*, *Rhodospirillum salinarum*, *Streptomyces ambofaciens*, *Streptomyces aureofaciens*, *Streptomyces aureus*, *Streptomyces fungicidicus*, *Streptomyces griseochromogenes*, *Streptomyces griseus*, *Streptomyces lividans*, *Streptomyces olivogriseus*, *Streptomyces rameus*, *Streptomyces tanashiensis*, *Streptomyces vinaceus*, *Zymomonas mobilis* y otros similares.

15

20

25

30

Los ejemplos de procariontes preferidos incluyen bacterias pertenecientes a los géneros *Escherichia*, *Serratia*, *Bacillus*, *Brevibacterium*, *Corynebacterium*, *Pseudomonas*, *Streptomyces* y otros similares. Por ejemplo, se pueden mencionar las especies anteriores que pertenecen al género *Escherichia*, *Serratia*, *Bacillus*, *Brevibacterium*, *Corynebacterium*, *Pseudomonas*, *Streptomyces* y otros similares. Los ejemplos de bacterias más preferidas incluyen *Escherichia coli*, *Corynebacterium glutamicum*, *Corynebacterium ammoniagenes*, *Corynebacterium lactofermentum*, *Corynebacterium flavum*, *Corynebacterium efficiens*, *Bacillus subtilis*, *Bacillus megaterium*, *Serratia marcescens*, *Pseudomonas putida*, *Pseudomonas aeruginosa*, *Streptomyces coelicolor* y *Streptomyces lividans*. De forma particularmente preferida, se puede mencionar la *Escherichia coli*.

35

40

(c) Preparación de un microorganismo en el que se incorpora un ADN que codifica una proteína que tiene actividad de transporte de dipéptidos al ADN cromosomal

Mediante la incorporación de un ADN que codifica una proteína que tenga actividad de transporte de dipéptidos obtenida por el método mencionado anteriormente en (a) en una posición del ADN cromosomal, también se puede obtener un microorganismo en el que la actividad de la proteína que tiene actividad de transporte de dipéptidos sea superior a la de la cepa original.

45

Como método para la incorporación de un ADN que codifique una proteína que tenga actividad de transporte de dipéptidos en cualquier posición del ADN cromosomal de un microorganismo, se puede mencionar un método que utiliza recombinación homóloga. Si se usa *E. coli* como hospedante, es decir, una cepa original, se puede mencionar el método descrito en Proc. Natl. Acad. Sci. U.S.A., 97, 6640 (2000).

50

(2) Preparación de un microorganismo en el que la actividad de la proteína con actividad de síntesis de dipéptidos es superior a la de la cepa original

De los microorganismos en los que la actividad de una proteína con actividad de síntesis de dipéptidos es superior a la de la cepa original, se puede obtener un microorganismo que muestre una actividad específica superior a la de la

55

proteína de la cepa original, y un microorganismo en el que la cantidad producida de la proteína se incremente con respecto a la de la cepa original, como se han indicado en el anterior apartado (1) sustituyendo el gen de la cepa original por un gen de enzima mutante obtenido al someter a un ADN que codifica una proteína con actividad de síntesis de dipéptidos a un tratamiento de mutación in vitro, PCR con tendencia a error y otros métodos similares.

- 5 Cultivando respectivamente el microorganismo obtenido que tiene un gen de enzima mutante y la cepa original en un medio de cultivo líquido, y midiendo y comparando la cantidad de dipéptido contenida en el cultivo mediante un método conocido, se puede confirmar la mayor actividad de la proteína con actividad de síntesis de dipéptidos con respecto a la de la cepa original.

10 A continuación se explica un método para preparar ADN que codifica la proteína que tiene actividad de síntesis de dipéptidos y un método para preparar el microorganismo obtenido mediante la transformación de la cepa original con el ADN.

(a) Preparación de un ADN que codifica una proteína con actividad de síntesis de dipéptidos

15 Se puede obtener un ADN que codifica una proteína con actividad de síntesis de dipéptidos según la secuencia de nucleótidos de un ADN que codifica una proteína que tiene actividad de síntesis de dipéptidos del apartado anterior 1(2) del mismo modo que el apartado anterior 2(1) (a).

Los ejemplos de ADN que puede obtenerse mediante el método mencionado anteriormente incluyen un ADN que tiene una región codificadora en la secuencia de nucleótidos mostrada como cualquiera de las SEQ ID NO: 20 a 28, que codifica ligasa de aminoácido L que tiene las secuencias de aminoácido mostradas por las SEQ ID NO: 11 – 18.

20 (b) Preparación de un microorganismo transformado con un vector plásmido que expresa una proteína que tiene actividad de síntesis de dipéptidos

Se puede obtener un vector plásmido que exprese una proteína con actividad de síntesis de dipéptidos usando el ADN que codifica una proteína que tiene actividad de síntesis de dipéptidos, que se obtiene en el apartado anterior 2(2) (a), y del mismo modo que en el apartado anterior 2(1) (b).

25 Los ejemplos de vector plásmido que expresan una proteína con actividad de síntesis de dipéptidos, que puede obtenerse mediante el método mencionado anteriormente, incluyen el pPE86usp, que se menciona más adelante.

(c) Preparación de un microorganismo en el que se incorpora al ADN cromosomal un ADN que codifica la proteína con actividad de síntesis de dipéptidos

30 Un microorganismo en el que la actividad de una proteína con actividad de síntesis de dipéptidos sea superior a la de la cepa original también puede obtenerse incorporando el ADN que codifica la proteína con actividad de síntesis de dipéptidos, que se obtiene mediante el método del apartado anterior 2(2) (a), en cualquier posición de un ADN cromosomal.

El ADN se puede incorporar en cualquier posición de un ADN cromosomal de un microorganismo del mismo modo que en el apartado anterior 2(1) (c).

35 (3) Preparación de un microorganismo que tenga la capacidad de producir al menos un tipo de aminoácido de entre los aminoácidos que constituyen el dipéptido

40 El microorganismo que tiene la capacidad de producir al menos un tipo de aminoácidos de entre los aminoácidos que constituyen el dipéptido, que se usa en el proceso de producción de un dipéptido de la presente invención, puede ser cualquiera siempre que tenga dicha capacidad. Cuando una cepa aislada del mundo natural presenta dicha capacidad por sí misma, se puede usar la cepa *per se*. Alternativamente, se puede usar un microorganismo al que se le ha conferido la capacidad de producir al menos un tipo de aminoácido de entre los aminoácidos que constituyen el dipéptido mediante un método conocido, y similar.

Los ejemplos del método conocido incluyen:

- (a) un método para moderar o liberar al menos uno de los mecanismos que regulan la biosíntesis de un aminoácido,
 (b) un método para potenciar la expresión de al menos una enzima implicada en la biosíntesis de un aminoácido,
 45 (c) un método para aumentar el número de copias de al menos uno de los genes de enzima implicados en la biosíntesis de un aminoácido,
 (d) un método para debilitar o bloquear al menos una ruta metabólica paralela a la ruta de biosíntesis de un aminoácido que genere otro metabolito diferente al aminoácido, y
 50 (e) un método para seleccionar la línea celular que tenga una elevada resistencia a un análogo de aminoácido en comparación con la cepa natural,

y otros similares. Los métodos conocidos mencionados pueden usarse solos o en combinación.

El método (a) descrito antes se describe, por ejemplo, en *Agric. Biol. Chem.*, 43, 105-111 (1979), *J. Bacteriol.*, 110, 761-763 (1972) y en *Appl. Microbiol. Biotechnol.*, 39, 318-323 (1993) y otros similares; el método (b) descrito anteriormente se describe, por ejemplo, en *Agric. Biol. Chem.*, 43, 105-111 (1979) y en *J. Bacteriol.*, 110, 761-763 (1972) y otros similares; el método (c) descrito anteriormente se describe, por ejemplo, en *Appl. Microbiol. Biotechnol.*, 39, 318-323 (1993) y en *Agric. Biol. Chem.*, 39, 371-377 (1987) y otros similares; el método (d) descrito anteriormente se describe, por ejemplo, en *Appl. Environ. Microbiol.*, 38, 181-190 (1979) y en *Agric. Biol. Chem.*, 42, 1773-1778 (1978) y otros similares; y el método (e) descrito anteriormente se describe, por ejemplo, en *Agric. Biol. Chem.*, 36, 1675-1684 (1972), *Agric. Biol. Chem.*, 41, 109-116 (1977), *Agric. Biol. Chem.*, 37, 2013-2023 (1973) y en *Agric. Biol. Chem.*, 51, 2089-2094 (1987) y otros similares. En referencia a las publicaciones mencionadas, y otras similares, se puede preparar un microorganismo que tenga la capacidad de producir diversos aminoácidos.

Adicionalmente, muchos ejemplos de métodos de preparación de un microorganismo que tenga la capacidad de producir un aminoácido a partir de cualquiera de las vías (a) – (e) mencionadas anteriormente, o de una combinación de las mismas, se describe en *Biotechnology 2nd ed.*, Vol. 6, *Products of Primary Metabolism* (VCH Verlagsgesellschaft mbH, Weinheim, 1996), sección 14a, 14b, y en *Advances in Biochemical Engineering/Biotechnology*, 79, 1-35 (2003), y en *Amino Acid Fermentation*, Japan Scientific Societies Press, Hiroshi Aida et al. (1986). Además de las anteriores, se presentan muchos informes relativos a un método de preparación de un microorganismo con capacidad para producir un aminoácido específico en los siguientes documentos, JP-A-2003-164297, *Agric. Biol. Chem.*, 39, 153-160 (1975), *Agric. Biol. Chem.*, 39, 1149-1153 (1975), JP-A-58-13599, *J. Gen. Appl. Microbiol.*, 4, 272-283 (1958), JP-A-63-94985, *Agric. Biol. Chem.*, 37, 2013-2023 (1973), WO 97/15673, JP-A-56-18596, JP-A-56-144092, Publicación Nacional de Solicitud de Patente Internacional N° 2003-511086 y otros similares. Un microorganismo con capacidad para producir uno o más tipos de aminoácidos se puede preparar usando como referencia las publicaciones mencionadas, y otras similares.

Los ejemplos de microorganismos que tienen capacidad de producir un aminoácido, que pueden ser preparados a través de los métodos anteriores, incluyen un microorganismo como la bacteria productora de L-glutamina, que tiene defecto de gen *glnE* y/o gen *glnB*, un microorganismo como la bacteria productora de L-alanina, que muestra un aumento en la expresión del gen de alanina deshidrogenasa (gen *ald*), un microorganismo como el microorganismo productor de L-prolina, que expresa el gen *pheA* insensible a fenilalanina y/o el gen *aroF* insensible a tirosina, y otros similares.

El microorganismo mencionado anteriormente que forma y acumula un aminoácido puede ser cualquiera siempre que se le puedan aplicar los métodos de los apartados (a) – (e) anteriores o que tenga las propiedades genéticas mencionadas anteriormente. Se prefieren los procariontes y son más preferidas las bacterias. Los procariontes y las bacterias son los mismos indicados en el apartado anterior 2(1).

Los ejemplos de microorganismo productor de un aminoácido incluyen la *Escherichia coli* JGL1, la *Escherichia coli* JGLBE1 y otros similares, como la cepa productora de L-glutamina, la *Escherichia coli* JM101 que porta el plásmido de expresión de gen *ald* y otros similares, como cepa productora de L-alanina, la *Escherichia coli* JM101 que porta el plásmido de expresión del gen *pPHEA2* y/o *aroF* y otros similares, como cepa productora de L-fenilalanina, la *Escherichia coli* JGLE1 y la *Escherichia coli* JGLBE1 que porta el plásmido de expresión de gen *ald* y otros similares, como cepa productora de L-glutamina y L-alanina, la *Escherichia coli* JM101 que porta el plásmido de expresión de gen *ald* y el plásmido de expresión de gen *pPHEA2* y/o *aroF* y otros similares, como cepa productora de L-alanina y L-fenilalanina, la cepa ATCC21277 que porta el plásmido de expresión de gen *pPHEA2* y/o *aroF* y otros similares, como cepa productora de L-treonina y L-fenilalanina, se describen en el documento WO 2006/001379.

Específicamente, los ejemplos de microorganismos que tienen la capacidad de producir un aminoácido incluyen el FERM BP-5807, el ATCC13032 y otros similares, como cepa productora de ácido L-glutámico, el FERM P-4806, el ATCC14751 y otros similares como cepa productora de L-glutamina, ATCC21148, ATCC21277, ATCC21650 y otros similares como cepa productora de L-treonina, FERM P-5084, ATCC13286 y otros similares como cepa productora de L-lisina, FERM P-5479, VKPM B-2175, ATCC21608 y otros similares como cepa productora de L-metionina, FERM BP-3757, ATCC14310 y otros similares como cepa productora de L-isoleucina, ATCC13005, ATCC19561 y otras similares como cepa productora de L-valina, FERM BP-4704, ATCC21302 y otros similares como cepa productora de L-leucina, FERM BP-4121, ATCC15108 y otros similares como cepa productora de L-alanina, ATCC21523, FERM BP-6576 y otros similares como cepa productora de L-serina, FERM BP-2807, ATCC19224 y otros similares como cepa productora de L-prolina, FERM P-5616, ATCC21831 y otros similares como cepa productora de L-arginina, ATCC13232 y otros similares como cepa productora de L-ornitina, FERM BP-6674, ATCC21607 y otros similares como cepa productora de L-histidina, DSM10118, DSM10121, DSM10123, FERM BP-1777 y otros similares como la cepa productora de L-triptófano, ATCC13281, ATCC21669 y otros similares como cepa productora de L-fenilalanina, ATCC21652 y otros similares como cepa productora de L-tirosina, W3110/pHC34 (descrito en la Publicación Nacional de Solicitud de Patente Internacional N° 2003-511086) y otros similares como cepa productora de L-cisteína, *Escherichia coli* SOLR/pHR71 (descrita en el documento WO 96/27669) y otros similares como cepa productora de L-4-hidroxiprolina, FERM BP-5026, FERM BP-5409 y otros similares como cepa productora de L-3-hidroxiprolina, FERM P-5643, FERM P-1645 y otros similares como cepa productora de L-citrulina.

Las cepas mostradas con la identificación FERM mencionada anteriormente se encuentran disponibles en el "International Patent Organism Depository", National Institute of Advanced Industrial Science and Technology (Japón), las cepas mostradas con la identificación ATCC pueden obtenerse en el "American Type Culture Collection" (EE.UU.), las cepas identificadas como VKPM pueden obtenerse en el "Russian National Collection of Industrial Microorganisms" (Rusia), y las cepas identificadas como DSM pueden obtenerse en el "Deutsche Sammlung von Mikroorganismen und Zellkulturen" (Alemania).

(4) Preparación de microorganismos en los que la actividad de una proteína que tiene actividad de peptidasa o de captación de péptidos es inferior a la de la cepa original o se ha perdido.

El microorganismo de la presente invención y un microorganismo que va a ser usado en el método de producción de dipéptidos de la presente invención tienen la capacidad de producir una proteína que tiene actividad de síntesis de dipéptidos, donde la actividad de la proteína que tiene actividad de transporte de dipéptidos es superior a la de la cepa original, y la actividad de la proteína que tiene actividad de peptidasa o de captación de péptidos puede ser inferior a la de la cepa original o haberse perdido.

El microorganismo en el que la actividad de una proteína que tiene actividad de peptidasa o de captación de péptidos puede ser inferior a la de la cepa original, o se ha perdido, se puede preparar según el método descrito en el documento WO 2005/045006.

3. Proceso para la producción del dipéptido de la presente invención.

(1) Proceso para producir el dipéptido usando un cultivo de microorganismos o un cultivo tratado del mismo como fuente de enzimas

Se puede obtener un cultivo de los microorganismos de la presente invención cultivando los microorganismos en un medio natural o en un medio sintético que contenga fuentes de carbono, fuentes de nitrógeno, sales inorgánicas y otros similares, utilizables por los microorganismos, y que permita un cultivo eficiente de un transformante.

Como fuentes de carbono, se puede usar cualquier fuente de carbono que sea asimilada por el microorganismo. Los ejemplos de fuentes de carbono incluyen carbohidratos tales como glucosa, fructosa, sacarosa, melazas que los contengan, almidón e hidrolisato de almidón; ácidos orgánicos tales como ácido acético y ácido propiónico; y alcoholes tales como etanol y propanol.

Como fuentes de nitrógeno, se puede usar amoníaco, sales de amonio de ácidos inorgánicos u orgánicos tales como cloruro de amonio, sulfato de amonio, acetato de amonio y fosfato de amonio, y otros compuestos que contengan nitrógeno, así como peptona, extracto de carne, extracto de levadura, licor de maíz, hidrolisato de caseína, torta de soja, hidrolisato de torta de soja, y diversas células microbianas fermentadas, y productos digeridos de los mismos.

Los ejemplos de sales inorgánicas incluyen dihidrogenofosfato potásico, hidrogenofosfato dipotásico, fosfato de magnesio, sulfato de magnesio, cloruro sódico, sulfato ferroso, sulfato de manganeso, sulfato de cobre y carbonato cálcico.

El cultivo se lleva a cabo habitualmente en condiciones aerobias, por ejemplo, agitando el cultivo o con un agitador sumergido con aeración al cultivo. La temperatura de cultivo preferiblemente va de 15 a 40°C, y el tiempo de cultivo normalmente es de 5 horas a 7 días. El pH se mantiene entre 3,0 y 9,0 durante el cultivo. El ajuste de pH se lleva a cabo usando un ácido inorgánico u orgánico, una disolución alcalina, urea, carbonato cálcico, amoníaco, etc.

Si es necesario, se pueden añadir antibióticos como la ampicilina y la tetraciclina al medio de cultivo durante el cultivo.

Cuando se cultiva un microorganismo transformado con un vector de expresión que comprende un promotor inducible, se puede añadir un inductor al medio de cultivo, si es necesario. Por ejemplo, en el caso de un microorganismo transformado con un vector de expresión que comprende el promotor lac, se puede añadir isopropil- β -D-tiogalactopiranosido, u otro similar, al medio de cultivo; y en el caso de un microorganismo transformado con un vector de expresión que comprende el promotor trp, se puede añadir ácido indolacrílico, u otro similar.

Los ejemplos de cultivos tratados incluyen cultivos concentrados obtenidos mediante el método mencionado anteriormente, cultivos sometidos a secado, células microbianas centrifugadas y obtenidas a partir del cultivo, células microbianas sometidas a secado, células microbianas secadas por congelación, células microbianas tratadas con tensioactivos, células microbianas sometidas a ultrasonidos, células microbianas molidas mecánicamente, células microbianas tratadas con disolvente, células microbianas tratadas con enzimas y cultivos tratados que contengan células microbianas viables tales como células microbianas inmovilizadas.

Mediante el uso de un cultivo o un cultivo tratado de los microorganismos de la presente invención, que pueden obtenerse a través del método mencionado anteriormente, como fuente de enzimas, se puede producir un dipéptido que se forme y se acumule en un medio acuoso que contenga la fuente de enzimas y uno o más tipos de sustratos

seleccionados entre un aminoácido, un éster de aminoácido y una amida de aminoácido, y recuperando el dipéptido del medio.

El proceso anterior se explica en detalle a continuación, dividido en los apartados (i) a (iii):

5 (i) un proceso que usa un cultivo o un cultivo tratado del microorganismo como fuente de enzimas, donde el dipéptido se forma y se acumula en un medio acuoso que contiene la fuente de enzimas y uno o más tipos, preferiblemente uno o dos tipos, de aminoácidos, y se recupera del medio,

10 (ii) un proceso que usa un cultivo o un cultivo tratado del microorganismo como fuente de enzimas, donde se forma y se acumula un dipéptido en un medio acuoso que contiene la fuente de enzimas, uno o más tipos, preferiblemente un tipo, de éster de aminoácido y uno o más tipos, preferiblemente un tipo, de aminoácido, y se recupera del medio, y

(iii) un proceso que usa un cultivo o un cultivo tratado del microorganismo como fuente de enzimas, donde se forma y se acumula un dipéptido en un medio acuoso que contiene la fuente de enzimas, uno o más tipos, preferiblemente un tipo, de amida de aminoácido y uno o más tipos, preferiblemente un tipo, de aminoácido, y se recupera del medio.

15 En el proceso de producción del apartado anterior (i), uno o más tipos, preferiblemente uno o dos tipos de aminoácidos usados como sustratos pueden ser cualquier aminoácido, preferiblemente aminoácidos seleccionados del grupo que consiste en aminoácidos L, glicina (Gly) y β -alanina (β -Ala), que pueden usarse en cualquier combinación. Los ejemplos de aminoácido L incluyen L-alanina (L-Ala), L-glutamina (L-Gln), ácido L-glutámico (L-Glu), L-valina (L-Val), L-leucina (L-Leu), L-isoleucina (L-Ile), L-prolina (L-Pro), L-fenilalanina (L-Phe), L-triptófano (L-Trp), L-metionina (L-Met), L-serina (L-Ser), L-treonina (L-Thr), L-cisteína (L-Cys), L-asparagina (L-Asn), L-tirosina (L-Tyr), L-lisina (L-Lys), L-arginina (L-Arg), L-histidina (L-His), ácido L-aspártico (L-Asp), ácido L- α -aminobutírico (L- α -AB), L-azaserina, L-teanina, L-4-hidroxiprolina (L-4-HYP), L-3-hidroxiprolina (L-3-HYP), L-ornitina (L-Orn), L-citrulina (L-Cit) y L-6-diazo-5-oxo norleucina y otros similares.

25 Los aminoácidos que se usan más preferiblemente en el proceso de producción del apartado anterior (i) son una combinación de un tipo de aminoácido seleccionado del grupo que consiste en L-Ala, Gly, L-Met, L-Ser, L-Thr, L-Cys, L- α -AB y β -Ala, y un tipo de aminoácido seleccionado del grupo que consiste en L-Ala, L-Gln, L-Glu, Gly, L-Val, L-Leu, L-Ile, L-Pro, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L-Asp, L- α -AB, β -Ala, L-azaserina, L-teanina, L-4-HYP, L-3-HYP, L-Orn, L-Cit y L-6-diazo-5-oxo-norleucina; una combinación de L-Gln y L-Phe. Otros aminoácidos preferidos son: una combinación de L-Ala y un tipo de aminoácido seleccionado del grupo que consiste en L-Ala, L-Gln, Gly, L-Val, L-Leu, L-Ile, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L- α -AB, L-Azaserina, L-Cit y L-teanina; una combinación de Gly un tipo de aminoácido seleccionado del grupo que consiste en L-Gln, Gly, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Tyr, L-Lys, L-Arg, L- α -AB y K-Cit; una combinación de L-Met y un tipo de aminoácido seleccionado del grupo que consiste en L-Phe, L-Met, L-Ser, L-Thr, L-Cys, L-Tyr, L-Lys y L-His; una combinación de L-Ser y un tipo de aminoácido seleccionado del grupo que consiste en L-Gln, L-Phe, L-Ser, L-Thr, L-Tyr, L-His y L- α -AB; una combinación de L-Thr y un tipo de aminoácido seleccionado del grupo que consiste en L-Gln, L-Phe, L-Leu, L-Thr y L- α -AB; una combinación de L-Gln y L-Phe; una combinación de β -Ala y un tipo de aminoácido seleccionado del grupo que consiste en L-Phe, L-Met, L-His y L-Cit; una combinación de L- α -AB y L-Gln, L-Arg o L- α -AB.

40 En el proceso de producción del apartado anterior (i), el aminoácido que va a usarse como sustrato se añade a un medio acuoso hasta una concentración de 0,1 – 500 g/L, preferiblemente de 0,2 – 200 g/L, desde el inicio o durante la reacción.

Los ejemplos de dipéptido producido mediante el proceso de producción del apartado anterior (i) incluyen un dipéptido representado por la siguiente fórmula (I)



45 donde R^1 y R^2 son iguales o diferentes y cada uno es un aminoácido. Los ejemplos preferibles incluyen un dipéptido representado por la fórmula (I) anterior donde R^1 y R^2 son iguales o diferentes y cada uno es un aminoácido seleccionado de L-Ala, L-Gln, L-Glu, Gly, L-Val, L-Leu, L-Ile, L-Pro, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L-Asp, L- α -AB, β -Ala, L-azaserina, L-teanina, L-4-HYP, L-3-HYP, L-Orn y L-6-diazo-5-oxo-norleucina. Más preferido es un dipéptido en el que R^1 es L-Ala, Gly, L-Met, L-Ser, L-Thr, L-Cys, L- α -AB ó β -Ala, entonces R^2 es L-Ala, L-Gln, L-Glu, Gly, L-Val, L-Leu, L-Ile, L-Pro, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L-Asp, L- α -AB, β -Ala, L-azaserina, L-teanina, L-4-HYP, L-3-HYP, L-Orn ó L-6-diazo-5-oxo-norleucina, y aún más preferido es un dipéptido en el que cuando R^1 es L-Ala, entonces R^2 es L-Ala, L-Gln, Gly, L-Val, L-Leu, L-Ile, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L- α -AB, L-azaserina o L-teanina, cuando R^1 es Gly, entonces R^2 es L-Gln, Gly, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Tyr, L-Lys, L-Arg o L- α -AB, cuando R^1 es L-Met, entonces R^2 es L-Phe, L-Met, L-Cys, L-Tyr, L-Lys o L-His, cuando R^1 es L-Ser, entonces R^2 es L-Gln, Gly, L-Phe, L-Met, L-Ser, L-Thr, L-Tyr, L-His ó L- α -AB, cuando R^1 es L-Thr, entonces R^2 es L-Gln, L-Gly, L-Phe, L-Met, L-Ser, L-Thr ó L- α -AB, cuando R^1 es L-Gln, entonces R^2 es L-Phe ó L- α -AB, cuando R^1 es L-Phe, entonces R^2 es L-Gln, cuando R^1 es L-Trp, entonces R^2 es Gly, cuando R^1 es L-Cys, entonces R^2 es L-Ala, L-Gln,

Gly ó L-Met, cuando R¹ es L-Lys, entonces R² es L-Ala, Gly ó L-Met, cuando R¹ es L-Arg, entonces R² es L- α -AB, cuando R¹ es L-His, entonces R² es L-Met, o cuando R¹ es L- α -AB, entonces R² es L-Ala, L-Gln, Gly, L-Ser, L-Thr, L-Arg ó L- α -AB.

- 5 En el proceso de producción mencionado anteriormente, además, se puede añadir un compuesto que pueda ser metabolizado por el microorganismo de la presente invención para producir ATP, por ejemplo, sacáridos tales como glucosa, alcoholes tales como etanol, ácidos orgánicos tales como ácido acético y otros similares, según sea necesario como fuente de suministro de ATP a un medio acuoso.

- 10 En el proceso de producción del apartado anterior (ii), uno o más tipos de ésteres de aminoácido y uno o más tipos de aminoácidos que van a ser usados como sustratos y uno o más tipos de aminoácidos que van a ser usados como sustratos pueden ser cualquier combinación de cualquier éster de aminoácido y cualquier aminoácido, siempre que el microorganismo que va a ser usado como fuente de enzimas en el proceso de producción de la presente invención pueda formar un dipéptido usándolos como sustratos. Se prefiere una combinación de un tipo de éster de aminoácido y un tipo de aminoácido, en la que el aminoácido preferiblemente es un aminoácido L o glicina. Los ejemplos de una combinación más preferible de un tipo de éster de aminoácido y un tipo de aminoácido incluyen una combinación de éster de aminoácido, que sea de un tipo seleccionado del grupo que consiste en éster de L-alanina, éster de glicina, éster de L-valina, éster de L-isoleucina, éster de L-metionina, éster de L-fenilalanina, éster de L-serina, éster de L-metionina, éster de L-fenilalanina, éster de L-serina, éster de L-treonina, éster de L-glutamina, éster de L-tirosina, éster de L-arginina, α -éster de ácido L-aspártico, β -éster de ácido L-aspártico, éster de L-leucina, éster de L-asparagina, éster de L-lisina, α,β -dimetiléster de ácido L-aspártico y γ -éster de L-glutamina, y aminoácido, que sea de un tipo seleccionado del grupo que consiste en L-Gln, L-Asn, Gly, L-Ala, L-Leu, L-Met, L-Pro, L-Phe, L-Trp, L-Ser, L-Thr, L-Tyr, L-Lys, L-Arg, L-His y L-Glu.

En el proceso de producción del apartado anterior (ii), el éster de aminoácido y el aminoácido que van a ser usados como sustratos se añaden a un medio acuoso hasta una concentración de 0,1 – 500 g/L, preferiblemente de 0,2 – 200 g/L, desde el inicio o durante la reacción.

- 25 En el proceso de producción del apartado anterior (iii), uno o más tipos de amidas de aminoácido y uno o más tipos de aminoácidos que van a ser usados como sustratos pueden ser cualquier combinación de cualquier amida de aminoácido y cualquier aminoácido, siempre que el microorganismo que va a ser usado como fuente de enzimas en el método de producción de la presente invención pueda formar un dipéptido usándolos como sustratos. Se prefiere una combinación de un tipo de amida de aminoácido y un tipo de aminoácido. Como aminoácido, son preferibles los aminoácidos L y la glicina. Los ejemplos de combinaciones de un tipo de amida de aminoácido y un tipo de aminoácido incluyen una combinación de un tipo de amida de aminoácido seleccionada del grupo que consiste en amida de L-alanina, amida de glicina y amida de ácido L-aspártico, y un tipo de aminoácido seleccionado del grupo que consiste en L-Gln, L-Asn, Gly, L-Ala, L-Val, L-Leu, L-Ile, L-Met, L-Pro, L-Phe, L-Trp, L-Ser, L-Thr, L-Tyr, L-Lys, L-Arg, L-His y L-Glu.
- 30
- 35 En el proceso de producción del apartado anterior (iii), la amida de aminoácido y el aminoácido que van a ser usados como sustratos se añaden a un medio acuoso hasta una concentración de 0,1 – 500 g/L, preferiblemente de 0,2 – 200 g/L, desde el inicio o durante la reacción.

- 40 El medio acuoso que va a ser usado en el proceso de producción de la presente invención puede tener cualesquier componentes y cualquier composición, siempre que no inhiba la reacción de producción de un dipéptido. Por ejemplo, se pueden mencionar agua, tampones tales como fosfato, carbonato, acetato, borato, citrato, tris, etc., y otros similares. Adicionalmente, puede contener alcoholes tales como metanol, etanol y otros similares, ésteres tales como acetato de etilo y otros similares, cetonas tales como acetona y otros similares, y amidas tales como acetamida y otros similares.

- 45 La reacción de producción de un dipéptido se lleva a cabo en un medio acuoso en las condiciones de pH 5 – 11, preferiblemente pH 6 – 10, a 20 – 60°C, preferiblemente 25 – 45°C, durante 2 – 150 h, preferiblemente 6 – 120 h.

Cuando sea necesario, también se puede añadir un tensioactivo o un disolvente orgánico al medio acuoso.

- 50 El tensioactivo puede ser cualquiera siempre que promueva la formación de un dipéptido, tal como los tensioactivos no iónicos tales como polioxietilen octadecilamina (p.ej., NYMEEN S-215, fabricado por NOF Corporation), tensioactivos catiónicos tales como bromuro de cetil trimetilamonio y cloruro de alquildimetil bencilamonio (p.ej., cation F2-40E fabricado por NOF Corporation), tensioactivos aniónicos tales como sarcosinato de lauroilo, aminas terciarias tales como alquildimetilamina (p.ej., amina terciaria FB, fabricada por NOF Corporation) y otros similares. Uno tipo o varios tipos de los mismos se pueden usar mezclados. La concentración de tensioactivo generalmente es de 0,1 – 50 g/L. Los ejemplos de disolvente orgánico incluyen xileno, tolueno, alcohol alifático, acetona, acetato de etilo y otros similares, y generalmente se usa en una concentración de 0,1 – 50 mL/L.

- 55 Aunque la cantidad de cultivo o de producto tratado de cultivo que va a ser usado como fuente de enzimas varía dependiendo de la actividad específica y el tipo de fuente de enzima, por ejemplo es de 5 – 1000 mg,

preferiblemente de 10 – 400 mg, por cada 1 mg de aminoácido, éster metílico de aminoácido o amida de aminoácido como sustrato.

5 El dipéptido formado y acumulado en el medio acuoso puede recolectarse empleando un método general que use carbón activado, resina de intercambio iónico y otros similares, o mediante extracción con un disolvente orgánico, cristalización, cromatografía de capa fina, cromatografía líquida de alta resolución, y otros similares.

Además, los procesos de producción de los apartados anteriores (ii) y (iii) se pueden llevar a cabo según la descripción de los documentos de patente WO 03/010189 y WO 03/010187.

(2) Producción mediante un método de fermentación

10 Se puede producir un dipéptido cultivando el microorganismo de la presente invención que tiene capacidad para producir al menos un tipo de aminoácido de entre los aminoácidos que constituyen el dipéptido en un medio de cultivo para permitir la formación y acumulación del dipéptido en el medio de cultivo, y recuperando el dipéptido del cultivo.

15 El método para cultivar el microorganismo en un medio de cultivo es similar al método de cultivo en el apartado anterior (1). Cuando se desee, el medio de cultivo puede contener al menos un tipo de aminoácido que constituya el dipéptido deseado.

Los ejemplos de dipéptido producido mediante el proceso mencionado incluyen un dipéptido en el que uno o dos tipos de aminoácido están unidos en configuración α , dándole preferencia al dipéptido en el que el aminoácido es un aminoácido L o glicina. Más preferidos son los dipéptidos representados por la siguiente fórmula (II)

R^1-R^2 (II)

20 donde R^1 y R^2 son iguales o diferentes y cada uno es un aminoácido seleccionado de L-Ala, L-Gln, L-Glu, Gly, L-Val, L-Leu, L-Ile, L-Pro, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L-Asp, L- α -AB, L-4-HYP, L-3-HYP, L-ornitina (L-Orn) y L-citrulina (L-Cit). Más preferido es un dipéptido en el que cuando R^1 es L-Ala, Gly, L-Met, L-Ser, L-Cys, L- α -AB ó L-Thr, entonces R^2 es L-Gln, L-Glu, Gly, L-Val, L-Leu, L-Ile, L-Pro, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L- α -AB, L-4-HYP, L-3-HYP, L-Orn ó L-Cit.

25 Particularmente preferido es un dipéptido en el que cuando R^1 es L-Ala, entonces R^2 es L-Gln, Gly, L-Val, L-Leu, L-Ile, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Asn, L-Tyr, L-Lys, L-Arg, L-His, L- α -AB ó L-Cit, cuando R^1 es Gly, entonces R^2 es L-Gln, Gly, L-Phe, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Tyr, L-Lys, L-Arg, L- α -AB ó L-Cit, cuando R^1 es L-Met, entonces R^2 es L-Phe, L-Met, L-Cys, L-Tyr, L-Lys ó L-His, cuando R^1 es L-Ser, entonces R^2 es L-Gln, Gly, L-Phe, L-Met, L-Ser, L-Thr, Tyr, L-His ó L- α -AB, cuando R^1 es L-Thr, entonces R^2 es L-Gln, L-Leu, L-Phe, L-Met, L-Ser, L-Thr ó L- α -AB, cuando R^1 es L-Gln, entonces R^2 es L-Phe ó L- α -AB, cuando R^1 es L-Phe, entonces R^2 es L-Gln, cuando R^1 es L-Trp, entonces R^2 es Gly, cuando R^1 es L-Cys, entonces R^2 es L-Ala, L-Gln, Gly ó L-Met, cuando R^1 es L-Lys, entonces R^2 es L-Ala, Gly ó L-Met, cuando R^1 es L-Arg, entonces R^2 es L- α -AB, cuando R^1 es L-His, entonces R^2 es L-Met, o cuando R^1 es L- α -AB, entonces R^2 es L-Ala, L-Gln, Gly, L-Ser, L-Thr, L-Arg ó L- α -AB, y los más preferidos son L-alanil-L-alanina (L-Ala-L-Ala), L-alanil-L-glutamina (L-Ala-L-Gln), L-alanil-L-fenilalanina (L-Ala-L-Phe) y L-treonil-L-fenilalanina (L-Thr-L-Phe).

35

La recuperación del dipéptido formado y acumulado en el medio o cultivo acuoso puede llevarse a cabo empleando métodos ordinarios mediante el uso de carbón activado, resinas de intercambio iónico, etc., o por medio de extracción con un disolvente orgánico, cromatografía de capa fina, cromatografía de líquidos de alta resolución y otros similares.

40 La presente invención se explica detalladamente a continuación en referencia a los Ejemplos, que no pretenden ser limitativos.

[Ejemplo 1]

Construcción de un plásmido de expresión de gen bcr

Se construye un plásmido de expresión de gen bcr mediante el siguiente método.

45 Se inoculó la cepa JM101 de Escherichia coli en medio LB [10 g/L de Bacto Tripton (fabricado por Difco), 5 g/L de extracto de levadura (fabricado por Difco), 5 g/L de cloruro sódico] y se cultivó en reposo a 30°C durante una noche. Tras el cultivo, se aisló el ADN cromosomal del microorganismo y se purificó mediante un método que usa fenol saturado descrito en "Current Protocols in Molecular Biology".

50 En base a la secuencia de nucleótidos mostrada en la SEQ ID NO: 1, se sintetizaron ADNs que consistían en la secuencia de nucleótidos mostrada en la SEQ ID N: 30 ó 31 como ADNs cebadores para la amplificación de gen bcr, y se llevó a cabo una PCR usando los ADNs sintéticos como conjunto de cebadores.

ES 2 526 490 T3

Se llevó a cabo una PCR preparando 50 µL de una mezcla de reacción que comprende 0,1 µg de ADN cromosomal como plantilla, 0,5 µmol/L de cada uno de los cebadores, 2,5 unidades de ADN polimerasa Pyrobest (10x) (fabricado por Takara Bio Inc.) y 200 µmol/L de dNTPs (dATP, dGTP, dCTP y dTTP) y repitiendo 30 veces un ciclo que consiste en reacción a 96°C durante 15 s, reacción a 55°C durante 30 s y reacción a 72°C durante 1 minuto.

- 5 Tras confirmación de la amplificación de aproximadamente de 1,2 kb de fragmento de ADN, el fragmento de ADN fue purificado mediante un método convencional.

El vector de expresión pTrs30 [preparado a partir de Escherichia coli JM109/pTrs30 (FERM BP-5407)] que contenía el fragmento de ADN y promotor de trp fue digerido con HindIII, SacI, respectivamente, los fragmentos de ADN fueron separados mediante electroforesis en gel de agarosa y los fragmentos de ADN digeridos con enzima de restricción fueron recuperados usando un kit GENE CLEAN II (fabricado por BIO 101).

- 10

El fragmento de 1,2 kb que contenía el gen bcr obtenido antes, y el fragmento pTrs30 digerido con enzima de restricción fueron ligados uno al otro usando un Kit de Ligadura (fabricado por Takara Bio Inc.).

La cepa DH5α de Escherichia coli (fabricada por Takara Bio Inc.) fue transformada usando el ADN ligado obtenido, y se seleccionó un transformante con resistencia a ampicilina como índice.

- 15 Se extrajo un plásmido de la colonia del transformante seleccionado empleando un método conocido, y se analizó su estructura usando una enzima de restricción, mediante lo cual se confirmó que el vector de expresión pTbcr en el que se había ligado el gen bcr al promotor de trp se había adquirido.

A continuación, se digirieron el pTbcr y el pSTV28 con EcoRI y SacI, respectivamente, y el fragmento de 1,6 kb en el que se había ligado el gen bcr al promotor trp y el fragmento de pSTV28 digerido con enzima de restricción fueron recuperados del mismo modo que antes.

- 20

Fueron ligados del mismo modo que antes, la cepa DH5α de Escherichia coli fue transformada con el ADN ligado obtenido, y se seleccionó un transformante con resistencia a cloranfenicol como índice.

Se extrajo un plásmido de la colonia del transformante seleccionado empleando un método conocido, y se analizó su estructura usando una enzima de restricción, mediante lo cual se confirmó que el vector de expresión en el que el gen bcr había sido ligado al promotor trp se había obtenido. El vector de expresión se designó como pSbcr.

- 25

[Ejemplo 2]

Construcción de plásmido de expresión de gen norE

En base a la secuencia de nucleótidos mostrada en la SEQ ID NO: 2, se sintetizaron los ADNs cebadores que consistían en una secuencia de nucleótidos mostrada en las SEQ ID NO: 32 ó 33 del mismo modo que en el Ejemplo 1 como ADNs cebadores para la amplificación de gen norE. Se llevó a cabo una PCR usando dichos ADNs sintéticos como conjunto de cebadores.

- 30

La PCR se llevó a cabo en condiciones similares al Ejemplo 1, excepto que se empleó el conjunto de cebadores mencionado antes como ADN cebador.

El fragmento de ADN amplificado obtenido mediante PCR y pTrs30 fueron digeridos con HindIII y BamHI, respectivamente, ambos ADNs fueron ligados uno al otro del mismo modo que en el Ejemplo 1, y la cepa DH5α de Escherichia coli fue transformada con el ADN ligado.

- 35

Se construyó un vector de expresión pTnorE en el que el gen norE estaba ligado al promotor trp mediante extracción del transformante obtenido. Análogamente al Ejemplo 1, pTnorE y pSTV28 fueron digeridos con EcoRI y BamHI, respectivamente. Ambos ADNs fueron ligados uno al otro y se construyó el vector de expresión en el que el gen norE está ligado al promotor trp del mismo modo que en el Ejemplo Experimental 1 y se designó como pSnor.

- 40

[Ejemplo 3]

Construcción de plásmido de expresión de gen ydeE

En base a la secuencia de nucleótidos mostrada en la SEQ ID NO: 3, se sintetizaron los ADNs cebadores que consistían en una secuencia de nucleótidos mostrada en las SEQ ID NO: 34 ó 35 del mismo modo que en el Ejemplo 1 como ADNs cebadores para la amplificación del gen ydeE. Se llevó a cabo una PCR usando los ADNs sintéticos como conjunto de cebadores.

- 45

La PCR se llevó a cabo en condiciones similares al Ejemplo 1 excepto que se usó el conjunto de cebadores mencionado antes como ADN cebador.

El fragmento de ADN amplificado obtenido mediante PCR y pSTV28 fueron digeridos con EcoRI y BamHI, respectivamente. Ambos ADNs fueron ligados uno al otro del mismo modo que en el Ejemplo 1, y la cepa DH5α de

- 50

Escherichia coli fue transformada con el ADN ligado. Se construyó el vector de expresión en el que el gen ydeE fue ligado al promotor lac mediante el método mencionado anteriormente y se designó como pSydeE.

[Ejemplo 4]

Construcción de plásmido de expresión de gen emrD

5 En base a la secuencia de nucleótidos mostrada en la SEQ ID NO: 4, se sintetizaron los ADNs cebadores que consistían en una secuencia de nucleótidos mostrada en las SEQ ID NO: 36 ó 37 del mismo modo que en el Ejemplo 1 como ADNs cebadores para la amplificación del gen emrD. Se llevó a cabo una PCR usando los ADNs sintéticos como conjunto de cebadores.

10 La PCR se llevó a cabo en condiciones similares al Ejemplo 1 excepto que se usó el conjunto de cebadores mencionado antes como ADN cebador.

El fragmento de ADN amplificado obtenido mediante PCR y pSTV28 fueron digeridos con EcoRI y PstI, respectivamente. Ambos ADNs fueron ligados uno al otro del mismo modo que en el Ejemplo 1, y la cepa DH5α de Escherichia coli fue transformada con el ADN ligado. Se construyó el vector de expresión en el que el gen emrD fue ligado al promotor lac mediante el método mencionado anteriormente y se designó como pSemrD.

15 **[Ejemplo 5]**

Construcción de plásmido de expresión de gen yeeO

20 En base a la secuencia de nucleótidos mostrada en la SEQ ID NO: 5, se sintetizaron los ADNs cebadores que consistían en una secuencia de nucleótidos mostrada en las SEQ ID NO: 38 ó 39 del mismo modo que en el Ejemplo 1 como ADNs cebadores para la amplificación del gen yeeO. Se llevó a cabo una PCR usando los ADNs sintéticos como conjunto de cebadores.

La PCR se llevó a cabo en condiciones similares al Ejemplo 1 excepto que se usó el conjunto de cebadores mencionado antes como ADN cebador.

25 El fragmento de ADN amplificado obtenido mediante PCR y pSTV28 fueron digeridos con EcoRI y PstI, respectivamente. Ambos ADNs fueron ligados uno al otro del mismo modo que en el Ejemplo 1, y la cepa DH5α de Escherichia coli fue transformada con el ADN ligado. Se construyó el vector de expresión en el que el gen yeeO fue ligado al promotor lac mediante el método mencionado anteriormente y se designó como pSyeeO.

[Ejemplo 6]

Construcción de plásmido de expresión de gen ywfE y de gen ald

(1) Construcción del plásmido de expresión pTrSQE30

30 Se llevó a cabo una PCR usando ADNs que consistían en la secuencia de nucleótidos mostrada por las SEQ ID NO: 40 ó 41 como conjunto de cebadores y el vector de expresión pQE60 (fabricado por QIAGEN K.K.) como plantilla.

35 La PCR se llevó a cabo preparando 40 µL de una mezcla de reacción que comprende 10 ng de ADN plásmido, 0,5 µmol/L de cada uno de los cebadores, 2,5 unidades de Pfu ADN polimerasa, 4 µL de tampón para Pfu ADN polimerasa (10x) y 200 µmol/L de cada uno de los dNTPs, y repitiendo 30 veces un ciclo que consistía en reacción a 94°C durante 1 minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos.

40 El fragmento de ADN amplificado obtenido mediante PCR y pTrS30 fueron digeridos con ClaI y SphI, respectivamente. Ambos ADNs fueron ligados uno al otro del mismo modo que en el Ejemplo 1 y se transformó la cepa NM522 de Escherichia coli con el ADN ligado. Se construyó un vector de expresión de proteína marcada con His en el C terminal que tenía el promotor trp mediante el método mencionado anteriormente y se designó pTrSQE30.

(2) Construcción del plásmido de expresión pUATQE30

Se realizó una PCR usando ADNs que consistían en las secuencias de nucleótidos mostradas en las SEQ ID NO: 42 ó 43 como conjunto de cebadores y ADN cromosomal de Escherichia coli W3110 como plantilla.

45 La PCR se llevó a cabo preparando 40 µL de una mezcla de reacción que comprende 0,1 µg de ADN cromosomal, 0,5 µmol/L de cada uno de los cebadores, 2,5 unidades de Pfu ADN polimerasa, 4 µL de tampón para Pfu ADN polimerasa (10x) y 200 µmol/L de cada uno de los dNTPs, y repitiendo 30 veces un ciclo que consistía en reacción a 94°C durante 1 minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos.

El fragmento de ADN amplificado obtenido mediante PCR y pTrSQE30 fueron digeridos con EcoRI y ClaI, respectivamente. Ambos ADNs fueron ligados uno al otro del mismo modo que en el Ejemplo 1 y se transformó la

cepa NM522 de *Escherichia coli* con el ADN ligado. Se construyó un vector de expresión de proteína marcada con His en el C terminal que tenía un promotor de proteína de estrés (promotor *uspA*) mediante el método mencionado anteriormente y se designó pUATQE30.

(3) Construcción de plásmido de expresión de gen *ywfE* y de gen *ald*

- 5 Se llevó a cabo una PCR usando ADNs que consisten en las secuencias de nucleótidos mostradas en las SEQ ID NO: 41 ó 42 como conjunto de cebadores y pUATQE30 obtenido antes como plantilla.

La PCR se llevó a cabo preparando 40 µL de una mezcla de reacción que comprende 10 ng de plásmido, 0,5 µmol/L de cada uno de los cebadores, 2,5 unidades de Pfu ADN polimerasa, 4 µL de tampón para Pfu ADN polimerasa (10x) y 200 µmol/L de cada uno de los dNTPs, y repitiendo 30 veces un ciclo que consistía en reacción a 94°C durante 1 minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos.

10 El fragmento de ADN amplificado obtenido mediante PCR y pPE86 que deriva de plásmido de expresión de gen *ywfE* de *Bacillus subtilis* y de gen *ald* de *Bacillus subtilis* (descritos en el documento 2006/001379) fue digerido con EcoRI y NcoI, respectivamente. Ambos ADNs fueron ligados uno al otro del mismo modo que en el Ejemplo 1 y se transformó la cepa NM522 de *Escherichia coli* con el ADN ligado. El plásmido de expresión en el que se ligan el gen *ywfE* y el gen *ald* con el promotor *uspA* fue construido mediante el método mencionado anteriormente y se designó como pPE86*usp*.

[Ejemplo 7]

Preparación de una cepa con eliminación del gen *pepD*, gen *pepN*, gen *pepB*, gen *pepA*, operón *dpp*, gen *glnE* y gen *glnB*

- 20 Usando como cepa original *Escherichia coli* JPNDBP7 (WO 2005/045006) donde se habían eliminado el gen *pepD*, el gen *pepN*, el gen *pepB* y el operón *dpp*, se preparó una cepa con el gen *pepA*, el gen *glnE* y el gen *glnB* en el ADN cromosomal de acuerdo a un método que utiliza el sistema de recombinación homóloga de fago lambda [Proc. Natl. Acad. Sci. USA, 97, 6641-6645 (2000)].

25 Se prepararon los plásmidos pKD46, pKD3 y pCP20 descritos más adelante por extracción de cepas de *Escherichia coli* que porta los plásmidos, que fue obtenida del "Escherichia coli Genetic Stock Center" (EE.UU., Universidad de Yale), de acuerdo a un método conocido.

(1) Clonación de fragmento de ADN para eliminación de gen

30 Las secuencias de nucleótidos del gen *pepA* que codifica peptidasa de la cepa K12 de *Escherichia coli*, el gen *putA* implicado en la degradación de L-prolina, cada uno de los genes *glnE* y *glnB* implicados en el control de la biosíntesis de L-glutamina, ya han sido esclarecidas [Science, 5331, 1453-1474 (1977)].

Para eliminar los respectivos genes de *pepA*, *putA*, *glnE* y *glnB*, se sintetizaron los ADNs que tienen una secuencia de nucleótidos homóloga a la secuencia de nucleótidos de 36 pb por encima y por debajo de los respectivos genes que van a ser eliminados del ADN cromosomal de la cepa K12 de *Escherichia coli* y la secuencia de nucleótidos que va a ser reconocida por Flp recombinasa derivada de levadura en base a las secuencias de nucleótidos reportadas.

35 Es decir, se sintetizaron respectivamente los ADNs que tiene las secuencias de nucleótidos mostradas en las SEQ ID NO: 40 a 41 como conjunto de cebadores para amplificación del fragmento de ADN para eliminación del gen *pepA*, los ADNs que consisten en las secuencias de nucleótidos mostradas en las SEQ ID NO: 42 y 43 como conjunto de cebadores para amplificación del fragmento de ADN para eliminación del gen *putA*, los ADNs que tiene las secuencias de nucleótidos mostradas en las SEQ ID NO: 44 a 45 como conjunto de cebadores para amplificación del fragmento de ADN para eliminación del gen *glnE*, los ADNs que consisten en las secuencias de nucleótidos mostradas en las SEQ ID NO: 46 y 47 como conjunto de cebadores para amplificación del fragmento de ADN para eliminación del gen *glnB*.

45 A continuación, se llevó a cabo una PCR usando los ADNs sintéticos mencionados anteriormente como conjunto de cebadores y pKD3DNA como plantilla. La PCR se llevó a cabo repitiendo 30 veces un ciclo que consistía en reacción a 94°C durante 1 minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos, usando 40 µL de la mezcla de reacción que comprende 10 ng de ADN plásmido, 0,5 µmol/L de cada uno de los cebadores, 2,5 unidades de Pfu ADN polimerasa, 4 µL de tampón para Pfu ADN polimerasa (10x) y 200 µmol/L de cada uno de los dNTPs.

50 Los fragmentos de ADN para la eliminación del gen *pepA*, el gen *putA*, el gen *glnE* y el gen *glnB*, que contenían el gen de resistencia a cloranfenicol, fueron obtenidos mediante la PCR mencionada antes.

(2) Preparación de *Escherichia coli* JPNDBP7 con defecto de gen *pepA*

La cepa JPNDBP7 de *Escherichia coli* fue transformada con pKD46, el transformante se propagó sobre un medio de agar LB que contenía 100 mg/L de ampicilina y se cultivó a 30°C para seleccionar la cepa JPNDBP7 de *Escherichia coli* que porta pKD46 (denominado a partir de este punto como *Escherichia coli* JPNDBP7/pKD46).

5 El plásmido pKD46 tiene el gen de recombinasa λ Red y la expresión de dicho gen puede ser inducida por L-arabinosa. Por tanto, se produce una recombinación homóloga con una alta frecuencia cuando *Escherichia coli* que porta pKD46 en presencia de L-arabinosa es transformada con ADN lineal. Además, puesto que el pKD46 tiene un origen de replicación sensible a la temperatura, el plásmido puede ser curado fácilmente mediante cultivo a 42°C.

10 Un fragmento para la eliminación del gen pepD obtenido anteriormente, que contiene un gen de resistencia a cloranfenicol, fue introducido en *Escherichia coli* JPNDBP7/pKD46 obtenida mediante cultivo en presencia de 10 mmol/L de L-arabinosa y 50 μ g/mL de ampicilina con el método de pulso eléctrico. Se incorporó un transformante en el que el fragmento de ADN para la eliminación de gen pepA que contiene un gen de resistencia a cloranfenicol en el ADN cromosomal de *Escherichia coli* JPNDBP7, y se fue propagado en un medio de agar LB (Bacto Tripton 10 g/L, Extracto de Levadura Bacto 5 g/L, cloruro sódico 5 g/L, agar 15 g/L) que contenía 25 mg/L de cloranfenicol y se cultivó a 30°C para la selección.

15 La cepa de resistencia de cloranfenicol seleccionada fue inoculada en un medio de agar LB que contenía 25 mg/L de cloranfenicol, se cultivó durante 14 h a 42°C y se separaron colonias sencillas. Cada una de las colonias obtenidas fue replicada en un medio agar LB que contenía 25 mg/L de cloranfenicol y un medio de agar LB que contenía 100 mg/L de ampicilina cultivada 37°C, y las colonias que mostraron resistencia a cloranfenicol y sensibilidad a ampicilina fueron seleccionadas para producir la cepa curada de pKD46.

20 A continuación, la cepa curada de pKD46 obtenida antes fue transformada con pCP20 y seleccionada sobre un medio de agar LB que contenía 100 mg/L de ampicilina para producir la cepa curada de pKD46 que porta pCP20.

El plásmido pCP20 tiene el gen de recombinasa F₁p derivado de levadura y la expresión de este gen puede ser inducida a 42°C.

25 Asimismo, ambos extremos de los fragmentos de ADN para la eliminación del gen pepA, el gen putA, el gen glnE o el gen glnB que contienen un gen de resistencia a cloranfenicol, que fueron preparados antes, tienen una secuencia de nucleótidos reconocible por recombinasa F₁p. Por lo tanto, el gen de resistencia puede ser curado fácilmente mediante recombinación homóloga catalizada por recombinasa F₁p.

Además, la expresión de recombinasa F₁p y el curado de pCP20 se pueden introducir simultáneamente cultivando una cepa que porta pCP20 a 42°C, ya que el pCP20 tiene un origen de replicación sensible a la temperatura.

30 La cepa curada de pKD46 que porta pCP20 obtenida antes fue inoculada en un medio de agar LB libre de fármaco, cultivada durante 14 h a 42°C y se separaron colonias individuales. Cada una de las colonias obtenidas fue replicada en un medio de agar LB libre de fármaco, un medio de agar LB que contiene 25 mg/L de cloranfenicol y un medio de agar LB que contiene 100 mg/L de ampicilina, cultivada a 30°C y se seleccionaron las colonias que mostraron sensibilidad a cloranfenicol y sensibilidad a ampicilina.

35 Se preparó el ADN cromosomal de cada cepa seleccionada antes. La cepa confirmada mediante PCR como defectuosa en gen pepA sobre el ADN cromosomal fue tomada como cepa defectuosa en gen pepA y se designó como cepa *Escherichia coli* JPNDABP.

40 A continuación, usando la cepa *Escherichia coli* JPNDABP como cepa original y de acuerdo a un método similar al anterior, se introdujo la eliminación de gen putA, la eliminación de gen glnE y la eliminación de gen glnB en este orden para producir la cepa que tiene eliminados los genes pepD, pepN, pepA, pepB, putA, glnE y glnB, así como el multigen operón dpp. La cepa obtenida fue designada como cepa *Escherichia coli* JPNDABPUTGEB.

[Ejemplo 8]

Producción L-alanil-L-glutamina (L-Ala-L-Gln)

45 El JPNDABPUTGEB obtenido en el Ejemplo 7 fue transformado con el pPE86usp obtenido en el Ejemplo 6 para producir *Escherichia coli* JPNDABPUTGEB/pPE86usp que tiene la capacidad de producir una proteína que tiene actividad de dipéptido sintasa.

50 A continuación, la *Escherichia coli* JPNDABPUTGEB/pPE86usp fue transformada con pSbcr, pSnorE, pSydeE, pSemrD o pSyeeO, que fueron obtenidos en los Ejemplos 1 – 5. Los transformantes obtenidos fueron designados *Escherichia coli* JPNDABPUTGEB/pPE86usp/pSbcr, JPNDABPUTGEB/pPE86usp/pSnorE, JPNDABPUTGEB/pPE86usp/pSydeE, JPNDABPUTGEB/pPE86usp/pSemrD, JPNDABPUTGEB/pPE86usp/pSyeeO, respectivamente. Del mismo modo, también se obtuvo un transformante (JPNDABPUTGEB/pPE86usp/pSTV28) que portaba pSTV28.

El transformante obtenido antes fue inoculado en un tubo de ensayo de mayor tamaño que contenía 8 mL de medio LB que contiene 50 µg/mL de ampicilina y 25 µg/mL de cloranfenicol y la mezcla se cultivó durante 17 h a 30°C. El medio cultivado se añadió a un tubo de ensayo que contenía 8 mL de un medio de cultivo (16 g/L de hidrogenofosfato dipotásico, 14 g/L de dihidrogenofosfato potásico, 2 g/L de sulfato amónico, 1 g/L de ácido cítrico (anhidro), 1 g/L de ácido casamino (fabricado por Difco), 10 g/L de glucosa, 10 mg/L de vitamina B₁, 2 g/L de sulfato magnésico heptahidratado, 10 mg/L de sulfato de manganeso pentahidratado, 50 mg/L de sulfato férrico heptahidratado, 0,1 g/L de L-Pro; ajustado a pH 7,2 con 10 mol/L de disolución de hidróxido sódico; se añadió glucosa, vitamina B₁, sulfato magnésico heptahidratado, sulfato férrico heptahidratado y L-Pro tras autoclavado por separado) al 1%, y la mezcla se cultivó durante 24 h a 30°C. El medio de cultivo se centrifugó para obtener el sobrenadante. El producto cultivado resultante del sobrenadante del cultivo fue derivatizado con el método F-moc, y el producto resultante se analizó con HPLC. Los resultados se muestran en la Tabla 1.

Tabla 1

cepa	L-Ala-L-Gln (g/L)
JPNDABPUTGEB/pPE86usp/pSTV28	0,34
JPNDABPUTGEB/pPE86usp/pSbcr	0,55
JPNDABPUTGEB/pPE86usp/pSnorE	0,62
JPNDABPUTGEB/pPE86usp/pSydeE	0,69
JPNDABPUTGEB/pPE86usp/pSemrD	0,52
JPNDABPUTGEB/pPE86usp/pSyeeO	0,55

Tal como se muestra en la Tabla 1, la cantidad acumulada de L-Ala-L-Gln en el medio de cultivo aumentó al potenciar la expresión del gen bcr, norE, ydeE, emrD ó yeeO. Estos resultados demuestran que los productos génicos anteriores son proteínas que tienen actividad de transporte de dipéptidos, mediante el cual el dipéptido de la célula es exportado extracelularmente.

[Ejemplo 9]

Producción de L-alanil-L-leucina (L-Ala-L-Leu), L-alanil-L-valina (L-Ala-L-Val, L-alanil-L-isoleucina (L-Ala-L-Ile), L-alanil-L-tirosina (L-Ala-L-Tyr).

Se transformó *Escherichia coli* JPNDDP36 (WO 05/45006) con pPE86usp obtenido en el Ejemplo 6 y el transformante obtenido se designó JPNDDP36/pPE86usp. La *Escherichia coli* JPNDDP36/pPE86usp se transformó con pSbcr, pSnorE, pSydeE, pSemrD y pSyeeO, que fueron obtenidos en los Ejemplos 1 – 5, respectivamente, y los transformantes obtenidos fueron designados como *Escherichia coli* JPNDDP36/pPE86usp/pSbcr, JPNDDP36/pPE86usp/pSnorE, JPNDDP36/pPE86usp/pSydeE, JPNDDP36/pPE86usp/pSemrD y JPNDDP36/pPE86usp/pSyeeO, respectivamente. Del mismo modo, también se obtuvo un transformante (JPNDDP36/pPE86usp/pSTV28) que porta pSTV28.

El transformante obtenido fue inoculado en un tubo de ensayo de mayor tamaño que contenía 8 mL de medio LB que contiene 50 µg/mL de ampicilina y 25 µg/mL de cloranfenicol y la mezcla se cultivó durante 17 h a 30°C. El medio cultivado se añadió a un tubo de ensayo que contenía 8 mL de un medio acuoso que contenía 100 µg/mL de ampicilina, 25 µg/mL de cloranfenicol y aminoácido (L-Leu ó L-Val ó L-Ile ó L-Tyr) (16 g/L de hidrogenofosfato dipotásico, 14 g/L de dihidrogenofosfato potásico, 2 g/L de sulfato amónico, 1 g/L de ácido cítrico (anhidro), 1 g/L de ácido casamino (fabricado por Difco), 0,1 g/L de L-Pro, 2 g/L de L-Leu (ó L-Val, L-Ile, L-Tyr), 10 g/L de glucosa, 10 mg/L de vitamina B₁, 2 g/L de sulfato magnésico heptahidratado, 50 mg/L de sulfato férrico heptahidratado, 10 mg/L de sulfato de manganeso pentahidratado; ajustado a pH 7,2 con 10 mol/L de disolución de hidróxido sódico; se añadió glucosa, vitamina B₁, sulfato magnésico heptahidratado, sulfato férrico heptahidratado y L-Pro tras autoclavado por separado) al 1%, y la mezcla se cultivó durante 24 h a 30°C. El medio acuoso se centrifugó para obtener el sobrenadante.

El producto resultante del sobrenadante del cultivo fue derivatizado con el método F-moc, y el producto resultante se analizó con HPLC. Los resultados se muestran en la Tabla 2.

Tabla 2

cepa	L-Ala-L-Leu (g/L)
JPNDDP36/pPE86usp/pSTV28	0,12
JPNDDP36/pPE86usp/pSbcr	0,29
JPNDDP36/pPE86usp/pSnorE	0,39

JPNDDP36/pPE86usp/pSydeE	0,39
JPNDDP36/pPE86usp/pSemrD	0,56
JPNDDP36/pPE86usp/pSyeeO	0,26
	L-Ala-L-Val (g/L)
JPNDDP36/pPE86usp/pSTV28	0,68
JPNDDP36/pPE86usp/pSbcr	1,65
JPNDDP36/pPE86usp/pSnorE	1,40
JPNDDP36/pPE86usp/pSydeE	2,21
JPNDDP36/pPE86usp/pSemrD	1,25
JPNDDP36/pPE86usp/pSyeeO	1,46
	L-Ala-L-Ile (g/L)
JPNDDP36/pPE86usp/pSTV28	0,26
JPNDDP36/pPE86usp/pSbcr	1,13
JPNDDP36/pPE86usp/pSnorE	0,84
JPNDDP36/pPE86usp/pSydeE	1,15
JPNDDP36/pPE86usp/pSemrD	0,43
JPNDDP36/pPE86usp/pSyeeO	0,51
	L-Ala-L-Tyr (g/L)
JPNDDP36/pPE86usp/pSTV28	0,25
JPNDDP36/pPE86usp/pSbcr	0,33
JPNDDP36/pPE86usp/pSnorE	0,81
JPNDDP36/pPE86usp/pSydeE	0,54
JPNDDP36/pPE86usp/pSemrD	0,32
JPNDDP36/pPE86usp/pSyeeO	0,22

Tal como se muestra en la Tabla 2, las cantidades acumuladas de dipéptidos diferentes a L-Ala-L-Gln en el medio de cultivo también aumentó al potenciar la expresión del gen bcr, norE, ydeE, emrD ó yeeO. Estos resultados demuestran que los productos génicos anteriores son proteínas que tienen actividad de transporte de dipéptidos en general.

5

“Texto Libre de Listado de Secuencias”

SEQ ID NO: 30 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 31 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 32 – explicación de secuencia artificial: ADN sintético.

10 SEQ ID NO: 33 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 34 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 35 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 36– explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 37 – explicación de secuencia artificial: ADN sintético.

15 SEQ ID NO: 38 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 39 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 40 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 41 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 42 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 43 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 44 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 45 – explicación de secuencia artificial: ADN sintético.

SEQ ID NO: 46 – explicación de secuencia artificial: ADN sintético.

5 SEQ ID NO: 47 – explicación de secuencia artificial: ADN sintético.

Aplicabilidad Industrial

Según la presente invención, potenciando la actividad de una proteína para transportar un dipéptido de la célula microbiana al exterior de las células microbianas de un microorganismo, se puede producir de manera eficiente un dipéptido usando el microorganismo.

10

ES 2 526 490 T3

Listado de secuencias

<110> KYOWA HAKKO KOGYO CO., LTD.

5 <120> Ppara producir dipéptidos

<130> 1944

<150> JP2007-99956

10 <151> 2007-04-06

<160> 47

<170> PatentIn Ver. 3.3

15 <210> 1

<211> 1491

<212> DNA

<213> Escherichia coli

20 <220>

<221> CDS

<222> (201)..(1388)

25 <400> 1

```

gtctgaccat cagcgaaggg cgttatcatc aggtgaaacg catgttcgcc gccgtgggta      60
accacgtggg tgagctgcat cgtgaacgta ttggcgggat tacgctggat gctgatttag      120
cccccggtga atatcgtccg ttaactgaag aagaaattgc cagcgtcgtc taacttttca      180
atgactttca ggagcccgtt  gtg acc acc cga cag cat tcg tcg ttt gct att      233
                    Met Thr Thr Arg Gln His Ser Ser Phe Ala Ile
                    1             5             10

ggt ttt atc ctt ggc ctg ctg gcc atg ttg atg ccg ctg tcg att gat      281
Val Phe Ile Leu Gly Leu Leu Ala Met Leu Met Pro Leu Ser Ile Asp
                    15             20             25

atg tat ctg ccc gcg cta ccg gta att tca gcg cag ttt ggc gta ccg      329
Met Tyr Leu Pro Ala Leu Pro Val Ile Ser Ala Gln Phe Gly Val Pro
                    30             35             40

gcg ggc agt acg cag atg acc ctc agt act tat att ctg ggc ttt gcg      377
Ala Gly Ser Thr Gln Met Thr Leu Ser Thr Tyr Ile Leu Gly Phe Ala
                    45             50             55

ttg ggg cag tta atc tac ggg ccg atg gca gac agc ttc ggg cgt aag      425
Leu Gly Gln Leu Ile Tyr Gly Pro Met Ala Asp Ser Phe Gly Arg Lys
                    60             65             70             75

ccg gtg gtg ctc ggc ggt acg ctg gtg ttt gcc gcc gcc gcg gtg gcg      473
Pro Val Val Leu Gly Gly Thr Leu Val Phe Ala Ala Ala Ala Val Ala
                    80             85             90

tgt gcg ttg gca aac acc atc gat cag ctg att gtg atg cgt ttc ttc      521
Cys Ala Leu Ala Asn Thr Ile Asp Gln Leu Ile Val Met Arg Phe Phe
                    95             100             105

cac ggg ctg gct gcg gct gcg gcc agc gtg gtc att aac gcc ctg atg      569

```

ES 2 526 490 T3

His	Gly	Leu	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Val	Val	Ile	Asn	Ala	Leu	Met	
		110						115					120				
cgc	gat	att	tac	ccg	aaa	gaa	gag	ttc	tcg	cgg	atg	atg	tcg	ttt	gtc		617
Arg	Asp	Ile	Tyr	Pro	Lys	Glu	Glu	Phe	Ser	Arg	Met	Met	Ser	Phe	Val		
	125					130					135						
atg	ctg	gtg	aca	acc	att	gca	ccg	ctg	atg	gca	ccg	ata	gtt	ggc	ggc		665
Met	Leu	Val	Thr	Thr	Ile	Ala	Pro	Leu	Met	Ala	Pro	Ile	Val	Gly	Gly		
	140				145					150				155			
tgg	gtg	ctg	gtg	tgg	ctg	agc	tgg	cat	tac	atc	ttc	tgg	atc	ctg	gca		713
Trp	Val	Leu	Val	Trp	Leu	Ser	Trp	His	Tyr	Ile	Phe	Trp	Ile	Leu	Ala		
				160				165						170			
tta	gcg	gcg	att	ctg	gct	tcg	gca	atg	att	ttc	ttc	ctg	att	aaa	gaa		761
Leu	Ala	Ala	Ile	Leu	Ala	Ser	Ala	Met	Ile	Phe	Phe	Leu	Ile	Lys	Glu		
			175					180					185				
acc	tta	cca	ccg	gag	cgt	cgt	cag	cca	ttt	cac	att	cgt	acc	act	att		809
Thr	Leu	Pro	Pro	Glu	Arg	Arg	Gln	Pro	Phe	His	Ile	Arg	Thr	Thr	Ile		
		190					195					200					
ggt	aac	ttt	gcg	gcg	ctg	ttc	cgc	cat	aaa	cgt	gtc	ctg	agc	tac	atg		857
Gly	Asn	Phe	Ala	Ala	Leu	Phe	Arg	His	Lys	Arg	Val	Leu	Ser	Tyr	Met		
	205					210					215						
ctt	gcc	agt	ggt	ttc	agc	ttt	gcc	ggg	atg	ttc	tca	ttc	tta	agc	gcc		905
Leu	Ala	Ser	Gly	Phe	Ser	Phe	Ala	Gly	Met	Phe	Ser	Phe	Leu	Ser	Ala		
	220				225					230					235		
gga	ccg	ttt	ggt	tat	att	gaa	att	aac	cac	gtc	gcg	ccg	gaa	aac	ttt		953
Gly	Pro	Phe	Val	Tyr	Ile	Glu	Ile	Asn	His	Val	Ala	Pro	Glu	Asn	Phe		
				240					245					250			
ggt	tat	tac	ttt	gcg	cta	aac	att	ggt	ttt	ctg	ttc	gtg	atg	acc	atc		1001
Gly	Tyr	Tyr	Phe	Ala	Leu	Asn	Ile	Val	Phe	Leu	Phe	Val	Met	Thr	Ile		
			255					260					265				
ttt	aac	agc	cgc	ttc	gtc	cgc	cgc	att	ggc	gcg	tta	aat	atg	ttc	cgc		1049
Phe	Asn	Ser	Arg	Phe	Val	Arg	Arg	Ile	Gly	Ala	Leu	Asn	Met	Phe	Arg		
		270					275					280					
tcg	ggg	ttg	tgg	ata	caa	ttt	att	atg	gca	gcg	tgg	atg	gtc	atc	agt		1097
Ser	Gly	Leu	Trp	Ile	Gln	Phe	Ile	Met	Ala	Ala	Trp	Met	Val	Ile	Ser		
	285					290					295						
gcg	ctg	ctg	ggg	ctg	gga	ttt	tgg	tcg	ctg	gtg	ggt	ggc	ggt	gcg	gcg		1145
Ala	Leu	Leu	Gly	Leu	Gly	Phe	Trp	Ser	Leu	Val	Val	Gly	Val	Ala	Ala		
	300				305					310				315			
ttt	gtg	ggc	tgc	gtg	tcg	atg	gtg	tca	tcc	aat	gcg	atg	gcg	gtc	att		1193
Phe	Val	Gly	Cys	Val	Ser	Met	Val	Ser	Ser	Asn	Ala	Met	Ala	Val	Ile		
				320				325						330			
ctt	gat	gag	ttt	ccc	cat	atg	gcg	gga	acg	gca	tct	tcg	ctg	gca	gga		1241
Leu	Asp	Glu	Phe	Pro	His	Met	Ala	Gly	Thr	Ala	Ser	Ser	Leu	Ala	Gly		
			335					340					345				
acc	ttc	cgt	ttt	ggc	ata	ggg	gca	att	ggt	ggc	gca	ttg	ctt	tct	ctt		1289
Thr	Phe	Arg	Phe	Gly	Ile	Gly	Ala	Ile	Val	Gly	Ala	Leu	Leu	Ser	Leu		

ES 2 526 490 T3

	350		355		360		
	gcg acc ttt aac tct gca tgg ccg atg att tgg tca att gca ttc tgc					1337	
	Ala Thr Phe Asn Ser Ala Trp Pro Met Ile Trp Ser Ile Ala Phe Cys						
	365		370		375		
	gca acc agc tcc att ctc ttc tgt ctg tac gcc agt cgg ccg aaa aaa					1385	
	Ala Thr Ser Ser Ile Leu Phe Cys Leu Tyr Ala Ser Arg Pro Lys Lys						
	380		385		390		395
	cgg tgatctattg cacaacgagg aagctaaaag gcttcctttg ttgatgcatg					1438	
	Arg						
	tcaaccacaa atctatcatt cccccgatat atgtttattt tatgtaaaat caa					1491	
	<210> 2						
	<211> 1674						
5	<212> DNA						
	<213> Escherichia coli						
	<220>						
	<221> CDS						
10	<222> (201)..(1571)						
	<400> 2						
	aaatcagtta agacattctg ttcagcacaa tagcaggtgg aaaacgcctt taccagtgaa					60	
	ggggtaagaa tggctatttt ttcactggag aattaataaa tcctcgctac aatagactga					120	
	atttcccctg cttcttcttt ttgctgcca ttcaggcggc tttttagtct ctcataaac					180	
	tacaaataaa aggtgttcac gtg cag aag tat atc agt gaa gcg cgt ctg tta					233	
	Met Gln Lys Tyr Ile Ser Glu Ala Arg Leu Leu						
	1		5		10		
	tta gca tta gca atc ccg gtg att ctc gcg caa atc gcc caa act gcg					281	
	Leu Ala Leu Ala Ile Pro Val Ile Leu Ala Gln Ile Ala Gln Thr Ala						
	15		20		25		
	atg ggt ttt gtc gat acc gtg atg gcg ggc ggc tat agt gcc acc gac					329	
	Met Gly Phe Val Asp Thr Val Met Ala Gly Gly Tyr Ser Ala Thr Asp						
	30		35		40		
	atg gcg gcg gtc gct atc ggt act tct atc tgg ctt ccg gcg atc ctc					377	
	Met Ala Ala Val Ala Ile Gly Thr Ser Ile Trp Leu Pro Ala Ile Leu						
	45		50		55		
	ttt ggt cac gga ctg ctg ctg gca tta acg ccg gtt atc gcg caa tta					425	
	Phe Gly His Gly Leu Leu Leu Ala Leu Thr Pro Val Ile Ala Gln Leu						
	60		65		70		75
	aat ggt tcc ggt cga cgt gag cgc att gcg cat cag gtg cga caa ggt					473	
	Asn Gly Ser Gly Arg Arg Glu Arg Ile Ala His Gln Val Arg Gln Gly						
	80		85		90		
	ttc tgg ctg gca ggt ttt gtt tcc gtt ctc att atg ctg gtg ctg tgg					521	
	Phe Trp Leu Ala Gly Phe Val Ser Val Leu Ile Met Leu Val Leu Trp						
	95		100		105		

15

ES 2 526 490 T3

aat gca ggt tac att atc cgc tcc atg gaa aac atc gat ccg gct ctg 569
 Asn Ala Gly Tyr Ile Ile Arg Ser Met Glu Asn Ile Asp Pro Ala Leu
 110 115 120

gcg gac aaa gcc gtg ggt tat ctg cgt gcg ttg ttg tgg gcc gcg ccg 617
 Ala Asp Lys Ala Val Gly Tyr Leu Arg Ala Leu Leu Trp Gly Ala Pro
 125 130 135

gga tat ctg ttc ttc cag gtt gcc cgt aac cag tgt gaa ggt ctg gca 665
 Gly Tyr Leu Phe Phe Gln Val Ala Arg Asn Gln Cys Glu Gly Leu Ala
 140 145 150 155

aaa acc aag ccg ggt atg gta atg ggc ttt atc ggc ctg ctg gtg aac 713
 Lys Thr Lys Pro Gly Met Val Met Gly Phe Ile Gly Leu Leu Val Asn
 160 165 170

atc ccg gtg aac tat atc ttt att tat ggt cat ttc ggt atg cct gag 761
 Ile Pro Val Asn Tyr Ile Phe Ile Tyr Gly His Phe Gly Met Pro Glu
 175 180 185

ctc ggt ggc gtt ggt tgt ggc gtg gct act gcg gcg gtg tat tgg gtc 809
 Leu Gly Gly Val Gly Cys Gly Val Ala Thr Ala Ala Val Tyr Trp Val
 190 195 200

atg ttc ctt gcc atg gtt tct tac att aaa cgc gcc cgc tcc atg cgc 857
 Met Phe Leu Ala Met Val Ser Tyr Ile Lys Arg Ala Arg Ser Met Arg
 205 210 215

gat att cgt aac gaa aaa ggc acc gca aaa ccc gat cct gcg gtt atg 905
 Asp Ile Arg Asn Glu Lys Gly Thr Ala Lys Pro Asp Pro Ala Val Met
 220 225 230 235

aaa cga ctg att caa ctc ggt ttg ccg att gcg ctg gca ctg ttc ttt 953
 Lys Arg Leu Ile Gln Leu Gly Leu Pro Ile Ala Leu Ala Leu Phe Phe
 240 245 250

gaa gtg aca ctg ttt gcc gtc gtg gct ctg tta gtg tct ccg ctc ggt 1001
 Glu Val Thr Leu Phe Ala Val Val Ala Leu Leu Val Ser Pro Leu Gly
 255 260 265

att gtt gat gtc gca gga cac cag att gcc ctg aac ttt agt tca cta 1049
 Ile Val Asp Val Ala Gly His Gln Ile Ala Leu Asn Phe Ser Ser Leu
 270 275 280

atg ttc gtg ctt cca atg tcc ctg gcg gca gcg gta act atc cgc gta 1097
 Met Phe Val Leu Pro Met Ser Leu Ala Ala Ala Val Thr Ile Arg Val
 285 290 295

ggt tat cgt ctg ggt cag gcc tca acg ctg gat gcg caa acc gct gcg 1145
 Gly Tyr Arg Leu Gly Gln Gly Ser Thr Leu Asp Ala Gln Thr Ala Ala
 300 305 310 315

cgg acc ggg ctt atg gtg ggt gtc tgt atg gca acc ctg acg gcc att 1193
 Arg Thr Gly Leu Met Val Gly Val Cys Met Ala Thr Leu Thr Ala Ile
 320 325 330

ttc acg gtt tca ctg cgg gag caa atc gcc ctg ttg tac aac gac aat 1241
 Phe Thr Val Ser Leu Arg Glu Gln Ile Ala Leu Leu Tyr Asn Asp Asn
 335 340 345

ES 2 526 490 T3

```

ccc gag gtt gta acg ctg gct gcg cat ttg atg ttg ctg gcg gcg gta      1289
Pro Glu Val Val Thr Leu Ala Ala His Leu Met Leu Leu Ala Ala Val
      350                      355                      360

tat cag att tct gac tca atc cag gtg att ggc agt ggg att ttg cgt      1337
Tyr Gln Ile Ser Asp Ser Ile Gln Val Ile Gly Ser Gly Ile Leu Arg
      365                      370                      375

ggg tat aaa gat acg cgt tcc att ttc tat att acc ttt acg gct tac      1385
Gly Tyr Lys Asp Thr Arg Ser Ile Phe Tyr Ile Thr Phe Thr Ala Tyr
      380                      385                      390                      395

tgg gtg ctg ggc ttg cca agc ggc tat att ctg gca ctg acc gat ctg      1433
Trp Val Leu Gly Leu Pro Ser Gly Tyr Ile Leu Ala Leu Thr Asp Leu
      400                      405                      410

gtc gtt gaa cct atg ggg cca gca ggc ttc tgg ata ggc ttt att att      1481
Val Val Glu Pro Met Gly Pro Ala Gly Phe Trp Ile Gly Phe Ile Ile
      415                      420                      425

ggc ctg acg tcg gca gcc att atg atg atg ttg cgt atg cgg ttc ctg      1529
Gly Leu Thr Ser Ala Ala Ile Met Met Met Leu Arg Met Arg Phe Leu
      430                      435                      440

caa cgt ctg ccg tca gcc atc att ctg caa cga gca tcc cgc      1571
Gln Arg Leu Pro Ser Ala Ile Ile Leu Gln Arg Ala Ser Arg
      445                      450                      455

```

taataaaagac aaggcgcaac cttcacgggt tgcgcctgta tttttacgca ggctggagcg 1631

ttgcgccaat cccgtcttcg tctggctgta atttcagagc gtt 1674

<210> 3
 <211> 1488
 5 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 10 <222> (201)..(1385)

<400> 3

```

cgacattcta cgcctctga atttcatctt ttgtaagcaa tcaacttagc tgaatttact      60

tttctttaac agttgattcg ttagtcgccg gttacgacgg cattaatgcg caataagtc      120

gctataacttc ggatttttgc catgctatctt ctttacatct ctaaaacaaa acataacgaa      180

acgcactgcc ggacagacaa atg aac tta tcc cta cga cgc tct acc agc gcc      233
      Met Asn Leu Ser Leu Arg Arg Ser Thr Ser Ala
      1                      5                      10

ctt ctt gcc tcg tcg ttg tta tta acc atc gga cgc ggc gct acg ctg      281
Leu Leu Ala Ser Ser Leu Leu Leu Thr Ile Gly Arg Gly Ala Thr Leu
      15                      20                      25

cca ttt atg acc att tac ttg agt cgc cag tac agc ctg agt gtc gat      329
Pro Phe Met Thr Ile Tyr Leu Ser Arg Gln Tyr Ser Leu Ser Val Asp
      30                      35                      40

```

15

ES 2 526 490 T3

cta atc ggt tat gcg atg aca att gcg ctc act att ggc gtc gtt ttt Leu Ile Gly Tyr Ala Met Thr Ile Ala Leu Thr Ile Gly Val Val Phe 45 50 55	377
agc ctc ggt ttt ggt atc ctg gcg gat aag ttc gac aag aaa cgc tat Ser Leu Gly Phe Gly Ile Leu Ala Asp Lys Phe Asp Lys Lys Arg Tyr 60 65 70 75	425
atg tta ctg gca att acc gcc ttc gcc agc ggt ttt att gcc att act Met Leu Leu Ala Ile Thr Ala Phe Ala Ser Gly Phe Ile Ala Ile Thr 80 85 90	473
tta gtg aat aac gtg acg ctg gtt gtg ctc ttt ttt gcc ctc att aac Leu Val Asn Asn Val Thr Leu Val Val Leu Phe Phe Ala Leu Ile Asn 95 100 105	521
tgc gcc tat tct gtt ttt gct acc gtg ctg aaa gcc tgg ttt gcc gac Cys Ala Tyr Ser Val Phe Ala Thr Val Leu Lys Ala Trp Phe Ala Asp 110 115 120	569
aat ctt tcg tcc acc agc aaa acg aaa atc ttc tca atc aac tac acc Asn Leu Ser Ser Thr Ser Lys Thr Lys Ile Phe Ser Ile Asn Tyr Thr 125 130 135	617
atg cta aac att ggc tgg acc atc ggt ccg ccg ctc ggc acg ctg ttg Met Leu Asn Ile Gly Trp Thr Ile Gly Pro Pro Leu Gly Thr Leu Leu 140 145 150 155	665
gta atg cag agc atc aat ctg ccc ttc tgg ctg gca gct atc tgt tcc Val Met Gln Ser Ile Asn Leu Pro Phe Thr Leu Ala Ala Ile Cys Ser 160 165 170	713
gcg ttt ccc atg ctt ttc att caa att tgg gta aag cgc agc gag aaa Ala Phe Pro Met Leu Phe Ile Gln Ile Trp Val Lys Arg Ser Glu Lys 175 180 185	761
atc atc gcc acg gaa aca ggc agt gtc tgg tcg ccg aaa gtt tta tta Ile Ile Ala Thr Glu Thr Gly Ser Val Trp Ser Pro Lys Val Leu Leu 190 195 200	809
caa gat aaa gca ctg ttg tgg ttt acc tgc tct ggt ttt ctg gct tct Gln Asp Lys Ala Leu Leu Trp Phe Thr Cys Ser Gly Phe Leu Ala Ser 205 210 215	857
ttt gta agc ggc gca ttt gct tca tgc att tca caa tat gtg atg gtg Phe Val Ser Gly Ala Phe Ala Ser Cys Ile Ser Gln Tyr Val Met Val 220 225 230 235	905
att gct gat ggg gat ttt gcc gaa aag gtg gtc gcg gtt gtt ctt ccg Ile Ala Asp Gly Asp Phe Ala Glu Lys Val Val Ala Val Val Leu Pro 240 245 250	953
gtg aat gct gcc atg gtg gtt acg ttg caa tat tcc gtg ggc cgc cga Val Asn Ala Ala Met Val Val Thr Leu Gln Tyr Ser Val Gly Arg Arg 255 260 265	1001
ctt aac ccg gct aac atc cgc gcg ctg atg aca gca ggc acc ctc tgt Leu Asn Pro Ala Asn Ile Arg Ala Leu Met Thr Ala Gly Thr Leu Cys 270 275 280	1049

ES 2 526 490 T3

ttc gtc atc ggt ctg gtc ggt ttt att ttt tcc ggc aac agc ctg cta 1097
Phe Val Ile Gly Leu Val Gly Phe Ile Phe Ser Gly Asn Ser Leu Leu
285 290 295

ttg tgg ggt atg tca gct gcg gta ttt act gtc ggt gaa atc att tat 1145
Leu Trp Gly Met Ser Ala Ala Val Phe Thr Val Gly Glu Ile Ile Tyr
300 305 310 315

gcg ccg ggc gag tat atg ttg att gac cat att gcg ccg cca gaa atg 1193
Ala Pro Gly Glu Tyr Met Leu Ile Asp His Ile Ala Pro Pro Glu Met
320 325 330

aaa gcc agc tat ttt tcc gcc cag tct tta ggc tgg ctt ggt gcc gcg 1241
Lys Ala Ser Tyr Phe Ser Ala Gln Ser Leu Gly Trp Leu Gly Ala Ala
335 340 345

att aac cca tta gtg agt ggc gta gtg cta acc agc ctg ccg cct tcc 1289
Ile Asn Pro Leu Val Ser Gly Val Val Leu Thr Ser Leu Pro Pro Ser
350 355 360

tcg ctg ttt gtc atc tta gcg ttg gtg atc att gct gcg tgg gtg ctg 1337
Ser Leu Phe Val Ile Leu Ala Leu Val Ile Ile Ala Ala Trp Val Leu
365 370 375

atg tta aaa ggg att cga gca aga ccg tgg ggg cag ccc gcg ctt tgt 1385
Met Leu Lys Gly Ile Arg Ala Arg Pro Trp Gly Gln Pro Ala Leu Cys
380 385 390 395

tgatttaagt cgaacacaat aaagatntaa ttcagccttc gtttaggtta cctctgctaa 1445

tatctttctc attgagatga aaattaaggt aagcgaggaa aca 1488

<210> 4
<211> 1485
5 <212> DNA
<213> Escherichia coli

<220>
10 <221> CDS
<222> (201)..(1382)

<400> 4

atgctgacgc atcttatccg cctaccatc tctcccggca acatttattg ccgcttttgt 60

ttacatattc tgccgctaaa caattcccca ttctggcgt atatctggct aacattcatc 120

aatgtgatag attcctctcc cgcatttatg ggaatgcgta gtgacttatt ctaattattt 180

ttataaaagc atccgtgata atg aaa agg caa aga aac gtc aat ttg tta ttg 233
Met Lys Arg Gln Arg Asn Val Asn Leu Leu Leu
1 5 10

atg ttg gta tta ctc gtg gcc gtc ggt cag atg gcg caa acc att tat 281
Met Leu Val Leu Leu Val Ala Val Gly Gln Met Ala Gln Thr Ile Tyr
15 20 25

att cca gct att gcc gat atg gcg cgc gat ctc aac gtc cgt gaa ggg 329
Ile Pro Ala Ile Ala Asp Met Ala Arg Asp Leu Asn Val Arg Glu Gly
30 35 40

15

ES 2 526 490 T3

gcg gtg cag agc gta atg ggc gct tat ctg ctg act tac ggt gtc tca 377
Ala Val Gln Ser Val Met Gly Ala Tyr Leu Leu Thr Tyr Gly Val Ser
45 50 55

cag ctg ttt tat ggc ccg att tcc gac cgc gtg ggc cgc cga ccg gtg 425
Gln Leu Phe Tyr Gly Pro Ile Ser Asp Arg Val Gly Arg Arg Pro Val
60 65 70 75

atc ctc gtc gga atg tcc att ttt atg ctg gca acg ctg gtc gcg gtc 473
Ile Leu Val Gly Met Ser Ile Phe Met Leu Ala Thr Leu Val Ala Val
80 85 90

acg acc tcc agt ttg acg gtg ttg att gcc gcc agc gcg atg cag ggg 521
Thr Thr Ser Ser Leu Thr Val Leu Ile Ala Ala Ser Ala Met Gln Gly
95 100 105

atg ggc acc ggc gtt ggc ggc gta atg gcg cgt act tta ccg cga gat 569
Met Gly Thr Gly Val Gly Gly Val Met Ala Arg Thr Leu Pro Arg Asp
110 115 120

tta tat gaa cgg aca cag ttg cgc cat gct aac agc ctg tta aac atg 617
Leu Tyr Glu Arg Thr Gln Leu Arg His Ala Asn Ser Leu Leu Asn Met
125 130 135

ggg att ctc gtc agt ccg ttg ctc gca ccg cta atc ggc ggt ctg ctg 665
Gly Ile Leu Val Ser Pro Leu Leu Ala Pro Leu Ile Gly Gly Leu Leu
140 145 150 155

gat acg atg tgg aac tgg cgc gcc tgt tat ctc ttt ttg ttg gtt ctt 713
Asp Thr Met Trp Asn Trp Arg Ala Cys Tyr Leu Phe Leu Leu Val Leu
160 165 170

tgt gct ggt gtg acc ttc agt atg gcc cgc tgg atg ccg gaa acg cgt 761
Cys Ala Gly Val Thr Phe Ser Met Ala Arg Trp Met Pro Glu Thr Arg
175 180 185

ccg gtc gat gca ccg cgc acg cgc ctg ctt acc agt tat aaa acg ctt 809
Pro Val Asp Ala Pro Arg Thr Arg Leu Leu Thr Ser Tyr Lys Thr Leu
190 195 200

ttc ggt aac agc ggt ttt aac tgt tat ttg ctg atg ctg att ggc ggt 857
Phe Gly Asn Ser Gly Phe Asn Cys Tyr Leu Leu Met Leu Ile Gly Gly
205 210 215

ctg gcc ggg att gcc gcc ttt gaa gcc tgc tcc ggc gtg ctg atg ggc 905
Leu Ala Gly Ile Ala Ala Phe Glu Ala Cys Ser Gly Val Leu Met Gly
220 225 230 235

gcg gtg tta ggg ctg agc agt atg acg gtc agt att ttg ttt att ctg 953
Ala Val Leu Gly Leu Ser Ser Met Thr Val Ser Ile Leu Phe Ile Leu
240 245 250

ccg att ccg gca gcg ttt ttt ggc gca tgg ttt gcc gga cgt ccc aat 1001
Pro Ile Pro Ala Ala Phe Phe Gly Ala Trp Phe Ala Gly Arg Pro Asn
255 260 265

aaa cgc ttc tcc acg tta atg tgg cag tgc gtt atc tgc tgc ctg ctg 1049
Lys Arg Phe Ser Thr Leu Met Trp Gln Ser Val Ile Cys Cys Leu Leu
270 275 280

ES 2 526 490 T3

gct ggc ttg ctg atg tgg atc ccc gac tgg ttt ggc gtg atg aat gtc 1097
 Ala Gly Leu Leu Met Trp Ile Pro Asp Trp Phe Gly Val Met Asn Val
 285 290 295

tgg acg ctg ctc gtt ccc gcc gcg ctg ttc ttt ttc ggt gcc ggg atg 1145
 Trp Thr Leu Leu Val Pro Ala Ala Leu Phe Phe Phe Gly Ala Gly Met
 300 305 310 315

ctg ttt ccg ctg gcg acc agc ggc gcg atg gag ccg ttc ccc ttc ctg 1193
 Leu Phe Pro Leu Ala Thr Ser Gly Ala Met Glu Pro Phe Pro Phe Leu
 320 325 330

gcg ggc acg gct ggc gcg ctg gtc ggc ggt ctg caa aac att ggt tcc 1241
 Ala Gly Thr Ala Gly Ala Leu Val Gly Gly Leu Gln Asn Ile Gly Ser
 335 340 345

ggc gtg ctg gcg tcg ctc tct gcg atg ttg ccg caa acc ggt cag ggc 1289
 Gly Val Leu Ala Ser Leu Ser Ala Met Leu Pro Gln Thr Gly Gln Gly
 350 355 360

agc ctg ggg ttg ttg atg acc tta atg gga ttg ttg atc gtg ctg tgc 1337
 Ser Leu Gly Leu Leu Met Thr Leu Met Gly Leu Leu Ile Val Leu Cys
 365 370 375

tgg ctg ccg ctg gcg acg cgg atg tcg cat cag ggg cag ccc gtt 1382
 Trp Leu Pro Leu Ala Thr Arg Met Ser His Gln Gly Gln Pro Val
 380 385 390

taagcgcacg tcaccgcagc atcgtcatca gctccatggg agaacgatgc tgctttatca 1442

gatcacgcat caccgcgata tgcggtgcgg agtaagaata aaa 1485

<210> 5
 <211> 1944
 5 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 10 <222> (201)..(1841)

<400> 5

tttgtaatgc aggactgtca gttgaggata tcaagcggca ggtgatggaa agcagtgaag 60

aggaaattga ttatcagata tagcaaaaat cccgctataa aagcgggctt ttcaggaatt 120

tggctcctct gactggactc gaaccagtga catacggatt aacagtcgcg cgttctaccg 180

actgaactac agaggaatcg ttg ttg agg cac atc tta acg gcg aaa aat ctt 233
 Met Leu Arg His Ile Leu Thr Ala Lys Asn Leu
 1 5 10

ttg tca aac ccg att ttt aaa ttc ccc aac tgt ttg ccg ttt cta tca 281
 Leu Ser Asn Pro Ile Phe Lys Phe Pro Asn Cys Leu Pro Phe Leu Ser
 15 20 25

aca gtt tgt tgc att tgc aga caa ttt gtt ggc gaa aat ctt tgc agc 329
 Thr Val Cys Cys Ile Cys Arg Gln Phe Val Gly Glu Asn Leu Cys Ser
 30 35 40

15

ES 2 526 490 T3

ttt gct gat tct ccc tca tta ttt gaa atg tgg ttt cac ttt ctg caa	377
Phe Ala Asp Ser Pro Ser Leu Phe Glu Met Trp Phe His Phe Leu Gln	
45 50 55	
tta agg tcg gct ttg aat atc tcc tct gct tta cgc cag gtt gtt cac	425
Leu Arg Ser Ala Leu Asn Ile Ser Ser Ala Leu Arg Gln Val Val His	
60 65 70 75	
ggc act cgc tgg cac gct aaa cgc aag agc tac aaa gtg ttg ttc tgg	473
Gly Thr Arg Trp His Ala Lys Arg Lys Ser Tyr Lys Val Leu Phe Trp	
80 85 90	
cgc gag ata acc ccg ctt gct gtt cct atc ttc atg gag aat gcc tgt	521
Arg Glu Ile Thr Pro Leu Ala Val Pro Ile Phe Met Glu Asn Ala Cys	
95 100 105	
gtc ctg ttg atg ggg gtt ctg agc act ttt ctg gtc agc tgg ctg gga	569
Val Leu Leu Met Gly Val Leu Ser Thr Phe Leu Val Ser Trp Leu Gly	
110 115 120	
aaa gat gcg atg gcc ggc gtg gga ttg gcg gac agc ttc aat atg gtc	617
Lys Asp Ala Met Ala Gly Val Gly Leu Ala Asp Ser Phe Asn Met Val	
125 130 135	
att atg gct ttt ttt gct gct atc gat ctt ggt act act gtc gtt gtg	665
Ile Met Ala Phe Phe Ala Ala Ile Asp Leu Gly Thr Thr Val Val Val	
140 145 150 155	
gca ttt agt ctc ggt aag cgg gat cga cga cga gcg agg gtg gcg acg	713
Ala Phe Ser Leu Gly Lys Arg Asp Arg Arg Arg Ala Arg Val Ala Thr	
160 165 170	
cgg cag tca ttg gtg atc atg acg ttg ttt gcc gta ctg ttg gca acg	761
Arg Gln Ser Leu Val Ile Met Thr Leu Phe Ala Val Leu Leu Ala Thr	
175 180 185	
ctt att cat cat ttt ggc gaa caa att att gat ttc gtc gcg ggt gat	809
Leu Ile His His Phe Gly Glu Gln Ile Ile Asp Phe Val Ala Gly Asp	
190 195 200	
gcc acg aca gaa gtt aaa gca ctg gcg ttg act tat ctg gag ctg acg	857
Ala Thr Thr Glu Val Lys Ala Leu Ala Leu Thr Tyr Leu Glu Leu Thr	
205 210 215	
gta ctc agt tat cca gca gct gcc atc act ctt att ggt agc ggg gca	905
Val Leu Ser Tyr Pro Ala Ala Ala Ile Thr Leu Ile Gly Ser Gly Ala	
220 225 230 235	
ctt cgt ggt gca ggg aat acg aaa ata ccg cta ttg att aac ggt agc	953
Leu Arg Gly Ala Gly Asn Thr Lys Ile Pro Leu Leu Ile Asn Gly Ser	
240 245 250	
ctg aat att ctt aat att att att agc ggc ata ttg att tac ggc ctt	1001
Leu Asn Ile Leu Asn Ile Ile Ile Ser Gly Ile Leu Ile Tyr Gly Leu	
255 260 265	
ttc tcc tgg ccg gga ctg gga ttt gtc ggg gca ggg ctg ggt tta acc	1049
Phe Ser Trp Pro Gly Leu Gly Phe Val Gly Ala Gly Leu Gly Leu Thr	
270 275 280	

ES 2 526 490 T3

att tct cgt tat att ggc gca gtt gca att ttg tgg gtg ctg gcg att Ile Ser Arg Tyr Ile Gly Ala Val Ala Ile Leu Trp Val Leu Ala Ile 285 290 295	1097
ggc ttt aat cct gcg cta agg att tcg tta aag agc tat ttt aaa ccg Gly Phe Asn Pro Ala Leu Arg Ile Ser Leu Lys Ser Tyr Phe Lys Pro 300 305 310 315	1145
ctg aat ttt agc att atc tgg gaa gtc atg ggg att ggt att ccc gcg Leu Asn Phe Ser Ile Ile Trp Glu Val Met Gly Ile Gly Ile Pro Ala 320 325 330	1193
agt gtc gaa tca gtg tta ttt acc agt ggt cgg tta tta acc caa atg Ser Val Glu Ser Val Leu Phe Thr Ser Gly Arg Leu Leu Thr Gln Met 335 340 345	1241
ttc gtt gcc ggg atg ggg acc agt gtt att gcc gga aat ttt atc gcg Phe Val Ala Gly Met Gly Thr Ser Val Ile Ala Gly Asn Phe Ile Ala 350 355 360	1289
ttt tca att gcg gct ctt atc aac tta ccc gga agt gcg ctc ggc tct Phe Ser Ile Ala Ala Leu Ile Asn Leu Pro Gly Ser Ala Leu Gly Ser 365 370 375	1337
gct tct acg atc att aca ggc cga agg ttg ggg gta ggg cag ata gcg Ala Ser Thr Ile Ile Thr Gly Arg Arg Leu Gly Val Gly Gln Ile Ala 380 385 390 395	1385
caa gca gag att cag ttg cgg cat gtg ttc tgg ctt tcc act ctt gga Gln Ala Glu Ile Gln Leu Arg His Val Phe Trp Leu Ser Thr Leu Gly 400 405 410	1433
tta acg gcc atc gcc tgg cta acg gct ccc ttt gcc ggg gtt atg gca Leu Thr Ala Ile Ala Trp Leu Thr Ala Pro Phe Ala Gly Val Met Ala 415 420 425	1481
tcg ttt tac acc cag gat cca cag gtt aaa cat gtc gtt gtg att ctg Ser Phe Tyr Thr Gln Asp Pro Gln Val Lys His Val Val Val Ile Leu 430 435 440	1529
att tgg cta aat gct tta ttt atg cct att tgg tcc gcc tca tgg gtg Ile Trp Leu Asn Ala Leu Phe Met Pro Ile Trp Ser Ala Ser Trp Val 445 450 455	1577
cta ccc gct gga ttt aaa ggt gct cgt gat gcc cgt tac gcc atg tgg Leu Pro Ala Gly Phe Lys Gly Ala Arg Asp Ala Arg Tyr Ala Met Trp 460 465 470 475	1625
gtt tcg atg ttg agc atg tgg ggt tgt cgg gtt gta gtc ggt tat gtg Val Ser Met Leu Ser Met Trp Gly Cys Arg Val Val Val Gly Tyr Val 480 485 490	1673
ctg gga atc atg ctt ggc tgg ggt gtg gtt ggt gtc tgg atg gga atg Leu Gly Ile Met Leu Gly Trp Gly Val Val Gly Val Trp Met Gly Met 495 500 505	1721
ttt gcc gac tgg gct gtg cgg gcc gtg ctg ttt tac tgg cga atg gtt Phe Ala Asp Trp Ala Val Arg Ala Val Leu Phe Tyr Trp Arg Met Val 510 515 520	1769
act gga cgt tgg cta tgg aaa tac cct cga ccc gag ccg caa aag tgt	1817

ES 2 526 490 T3

Thr Gly Arg Trp Leu Trp Lys Tyr Pro Arg Pro Glu Pro Gln Lys Cys
 525 530 535
 gaa aaa aag cca gtt gtg tcg gaa taaacgacaa aatgcagatt atttcagcaa 1871
 Glu Lys Lys Pro Val Val Ser Glu
 540 545
 acgatttcaa atttaaaaaa caggctttga cattgtgggt gggcatcgct aatattcgcc 1931
 tcgttctcac gat 1944

<210> 6
 <211> 396
 <212> PRT
 <213> Escherichia coli

5

<400> 6

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

10

ES 2 526 490 T3

Ala Ser Ala Met Ile Phe Phe Leu Ile Lys Glu Thr Leu Pro Pro Glu
 180 185 190

Arg Arg Gln Pro Phe His Ile Arg Thr Thr Ile Gly Asn Phe Ala Ala
 195 200 205

Leu Phe Arg His Lys Arg Val Leu Ser Tyr Met Leu Ala Ser Gly Phe
 210 215 220

Ser Phe Ala Gly Met Phe Ser Phe Leu Ser Ala Gly Pro Phe Val Tyr
 225 230 235 240

Ile Glu Ile Asn His Val Ala Pro Glu Asn Phe Gly Tyr Tyr Phe Ala
 245 250 255

Leu Asn Ile Val Phe Leu Phe Val Met Thr Ile Phe Asn Ser Arg Phe
 260 265 270

Val Arg Arg Ile Gly Ala Leu Asn Met Phe Arg Ser Gly Leu Trp Ile
 275 280 285

Gln Phe Ile Met Ala Ala Trp Met Val Ile Ser Ala Leu Leu Gly Leu
 290 295 300

Gly Phe Trp Ser Leu Val Val Gly Val Ala Ala Phe Val Gly Cys Val
 305 310 315 320

Ser Met Val Ser Ser Asn Ala Met Ala Val Ile Leu Asp Glu Phe Pro
 325 330 335

His Met Ala Gly Thr Ala Ser Ser Leu Ala Gly Thr Phe Arg Phe Gly
 340 345 350

Ile Gly Ala Ile Val Gly Ala Leu Leu Ser Leu Ala Thr Phe Asn Ser
 355 360 365

Ala Trp Pro Met Ile Trp Ser Ile Ala Phe Cys Ala Thr Ser Ser Ile
 370 375 380

Leu Phe Cys Leu Tyr Ala Ser Arg Pro Lys Lys Arg
 385 390 395

<210> 7
 <211> 457
 <212> PRT
 <213> Escherichia coli

5

<400> 7

ES 2 526 490 T3

Met Gln Lys Tyr Ile Ser Glu Ala Arg Leu Leu Leu Ala Leu Ala Ile
 1 5 10 15

Pro Val Ile Leu Ala Gln Ile Ala Gln Thr Ala Met Gly Phe Val Asp
 20 25 30

Thr Val Met Ala Gly Gly Tyr Ser Ala Thr Asp Met Ala Ala Val Ala
 35 40 45

Ile Gly Thr Ser Ile Trp Leu Pro Ala Ile Leu Phe Gly His Gly Leu
 50 55 60

Leu Leu Ala Leu Thr Pro Val Ile Ala Gln Leu Asn Gly Ser Gly Arg
 65 70 75 80

Arg Glu Arg Ile Ala His Gln Val Arg Gln Gly Phe Trp Leu Ala Gly
 85 90 95

Phe Val Ser Val Leu Ile Met Leu Val Leu Trp Asn Ala Gly Tyr Ile
 100 105 110

Ile Arg Ser Met Glu Asn Ile Asp Pro Ala Leu Ala Asp Lys Ala Val
 115 120 125

Gly Tyr Leu Arg Ala Leu Leu Trp Gly Ala Pro Gly Tyr Leu Phe Phe
 130 135 140

Gln Val Ala Arg Asn Gln Cys Glu Gly Leu Ala Lys Thr Lys Pro Gly
 145 150 155 160

Met Val Met Gly Phe Ile Gly Leu Leu Val Asn Ile Pro Val Asn Tyr
 165 170 175

Ile Phe Ile Tyr Gly His Phe Gly Met Pro Glu Leu Gly Gly Val Gly
 180 185 190

Cys Gly Val Ala Thr Ala Ala Val Tyr Trp Val Met Phe Leu Ala Met
 195 200 205

Val Ser Tyr Ile Lys Arg Ala Arg Ser Met Arg Asp Ile Arg Asn Glu
 210 215 220

Lys Gly Thr Ala Lys Pro Asp Pro Ala Val Met Lys Arg Leu Ile Gln

ES 2 526 490 T3

225	230	235	240
Leu Gly Leu Pro	Ile Ala Leu Ala Leu Phe Phe Glu Val Thr Leu Phe		
	245	250	255
Ala Val Val Ala Leu Leu Val Ser Pro Leu Gly Ile Val Asp Val Ala			
	260	265	270
Gly His Gln Ile Ala Leu Asn Phe Ser Ser Leu Met Phe Val Leu Pro			
	275	280	285
Met Ser Leu Ala Ala Ala Val Thr Ile Arg Val Gly Tyr Arg Leu Gly			
	290	295	300
Gln Gly Ser Thr Leu Asp Ala Gln Thr Ala Ala Arg Thr Gly Leu Met			
	305	310	315
Val Gly Val Cys Met Ala Thr Leu Thr Ala Ile Phe Thr Val Ser Leu			
	325	330	335
Arg Glu Gln Ile Ala Leu Leu Tyr Asn Asp Asn Pro Glu Val Val Thr			
	340	345	350
Leu Ala Ala His Leu Met Leu Leu Ala Ala Val Tyr Gln Ile Ser Asp			
	355	360	365
Ser Ile Gln Val Ile Gly Ser Gly Ile Leu Arg Gly Tyr Lys Asp Thr			
	370	375	380
Arg Ser Ile Phe Tyr Ile Thr Phe Thr Ala Tyr Trp Val Leu Gly Leu			
	385	390	395
Pro Ser Gly Tyr Ile Leu Ala Leu Thr Asp Leu Val Val Glu Pro Met			
	405	410	415
Gly Pro Ala Gly Phe Trp Ile Gly Phe Ile Ile Gly Leu Thr Ser Ala			
	420	425	430
Ala Ile Met Met Met Leu Arg Met Arg Phe Leu Gln Arg Leu Pro Ser			
	435	440	445
Ala Ile Ile Leu Gln Arg Ala Ser Arg			
	450	455	

5 <210> 8
 <211> 395
 <212> PRT
 <213> Escherichia coli
 <400> 8

ES 2 526 490 T3

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

ES 2 526 490 T3

Phe Ala Ser Cys Ile Ser Gln Tyr Val Met Val Ile Ala Asp Gly Asp
225 230 235 240

Phe Ala Glu Lys Val Val Ala Val Val Leu Pro Val Asn Ala Ala Met
245 250 255

Val Val Thr Leu Gln Tyr Ser Val Gly Arg Arg Leu Asn Pro Ala Asn
260 265 270

Ile Arg Ala Leu Met Thr Ala Gly Thr Leu Cys Phe Val Ile Gly Leu
275 280 285

Val Gly Phe Ile Phe Ser Gly Asn Ser Leu Leu Leu Trp Gly Met Ser
290 295 300

Ala Ala Val Phe Thr Val Gly Glu Ile Ile Tyr Ala Pro Gly Glu Tyr
305 310 315 320

Met Leu Ile Asp His Ile Ala Pro Pro Glu Met Lys Ala Ser Tyr Phe
325 330 335

Ser Ala Gln Ser Leu Gly Trp Leu Gly Ala Ala Ile Asn Pro Leu Val
340 345 350

Ser Gly Val Val Leu Thr Ser Leu Pro Pro Ser Ser Leu Phe Val Ile
355 360 365

Leu Ala Leu Val Ile Ile Ala Ala Trp Val Leu Met Leu Lys Gly Ile
370 375 380

Arg Ala Arg Pro Trp Gly Gln Pro Ala Leu Cys
385 390 395

<210> 9

<211> 394

5 <212> PRT

<213> Escherichia coli

<400> 9

Met Lys Arg Gln Arg Asn Val Asn Leu Leu Leu Met Leu Val Leu Leu
1 5 10 15

Val Ala Val Gly Gln Met Ala Gln Thr Ile Tyr Ile Pro Ala Ile Ala
20 25 30

10 Asp Met Ala Arg Asp Leu Asn Val Arg Glu Gly Ala Val Gln Ser Val
35 40 45

ES 2 526 490 T3

Met Gly Ala Tyr Leu Leu Thr Tyr Gly Val Ser Gln Leu Phe Tyr Gly
50 55 60

Pro Ile Ser Asp Arg Val Gly Arg Arg Pro Val Ile Leu Val Gly Met
65 70 75 80

Ser Ile Phe Met Leu Ala Thr Leu Val Ala Val Thr Thr Ser Ser Leu
85 90 95

Thr Val Leu Ile Ala Ala Ser Ala Met Gln Gly Met Gly Thr Gly Val
100 105 110

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

Gln Leu Arg His Ala Asn Ser Leu Leu Asn Met Gly Ile Leu Val Ser
130 135 140

Pro Leu Leu Ala Pro Leu Ile Gly Gly Leu Leu Asp Thr Met Trp Asn
145 150 155 160

Trp Arg Ala Cys Tyr Leu Phe Leu Leu Val Leu Cys Ala Gly Val Thr
165 170 175

Phe Ser Met Ala Arg Trp Met Pro Glu Thr Arg Pro Val Asp Ala Pro
180 185 190

Arg Thr Arg Leu Leu Thr Ser Tyr Lys Thr Leu Phe Gly Asn Ser Gly
195 200 205

Phe Asn Cys Tyr Leu Leu Met Leu Ile Gly Gly Leu Ala Gly Ile Ala
210 215 220

Ala Phe Glu Ala Cys Ser Gly Val Leu Met Gly Ala Val Leu Gly Leu
225 230 235 240

Ser Ser Met Thr Val Ser Ile Leu Phe Ile Leu Pro Ile Pro Ala Ala
245 250 255

Phe Phe Gly Ala Trp Phe Ala Gly Arg Pro Asn Lys Arg Phe Ser Thr
260 265 270

Leu Met Trp Gln Ser Val Ile Cys Cys Leu Leu Ala Gly Leu Leu Met
275 280 285

ES 2 526 490 T3

Trp Ile Pro Asp Trp Phe Gly Val Met Asn Val Trp Thr Leu Leu Val
 290 295 300

Pro Ala Ala Leu Phe Phe Phe Gly Ala Gly Met Leu Phe Pro Leu Ala
 305 310 315 320

Thr Ser Gly Ala Met Glu Pro Phe Pro Phe Leu Ala Gly Thr Ala Gly
 325 330 335

Ala Leu Val Gly Gly Leu Gln Asn Ile Gly Ser Gly Val Leu Ala Ser
 340 345 350

Leu Ser Ala Met Leu Pro Gln Thr Gly Gln Gly Ser Leu Gly Leu Leu
 355 360 365

Met Thr Leu Met Gly Leu Leu Ile Val Leu Cys Trp Leu Pro Leu Ala
 370 375 380

Thr Arg Met Ser His Gln Gly Gln Pro Val
 385 390

<210> 10
 <211> 547
 <212> PRT
 <213> Escherichia coli

5

<400> 10

Met Leu Arg His Ile Leu Thr Ala Lys Asn Leu Leu Ser Asn Pro Ile
 1 5 10 15

Phe Lys Phe Pro Asn Cys Leu Pro Phe Leu Ser Thr Val Cys Cys Ile
 20 25 30

Cys Arg Gln Phe Val Gly Glu Asn Leu Cys Ser Phe Ala Asp Ser Pro
 35 40 45

Ser Leu Phe Glu Met Trp Phe His Phe Leu Gln Leu Arg Ser Ala Leu
 50 55 60

Asn Ile Ser Ser Ala Leu Arg Gln Val Val His Gly Thr Arg Trp His
 65 70 75 80

Ala Lys Arg Lys Ser Tyr Lys Val Leu Phe Trp Arg Glu Ile Thr Pro
 85 90 95

10 Leu Ala Val Pro Ile Phe Met Glu Asn Ala Cys Val Leu Leu Met Gly

ES 2 526 490 T3

	100						105							110	
Val	Leu	Ser	Thr	Phe	Leu	Val	Ser	Trp	Leu	Gly	Lys	Asp	Ala	Met	Ala
		115					120					125			
Gly	Val	Gly	Leu	Ala	Asp	Ser	Phe	Asn	Met	Val	Ile	Met	Ala	Phe	Phe
	130					135					140				
Ala	Ala	Ile	Asp	Leu	Gly	Thr	Thr	Val	Val	Val	Ala	Phe	Ser	Leu	Gly
145					150					155					160
Lys	Arg	Asp	Arg	Arg	Arg	Ala	Arg	Val	Ala	Thr	Arg	Gln	Ser	Leu	Val
				165					170					175	
Ile	Met	Thr	Leu	Phe	Ala	Val	Leu	Leu	Ala	Thr	Leu	Ile	His	His	Phe
			180					185					190		
Gly	Glu	Gln	Ile	Ile	Asp	Phe	Val	Ala	Gly	Asp	Ala	Thr	Thr	Glu	Val
		195					200					205			
Lys	Ala	Leu	Ala	Leu	Thr	Tyr	Leu	Glu	Leu	Thr	Val	Leu	Ser	Tyr	Pro
	210					215					220				
Ala	Ala	Ala	Ile	Thr	Leu	Ile	Gly	Ser	Gly	Ala	Leu	Arg	Gly	Ala	Gly
225					230					235					240
Asn	Thr	Lys	Ile	Pro	Leu	Leu	Ile	Asn	Gly	Ser	Leu	Asn	Ile	Leu	Asn
				245					250					255	
Ile	Ile	Ile	Ser	Gly	Ile	Leu	Ile	Tyr	Gly	Leu	Phe	Ser	Trp	Pro	Gly
			260					265					270		
Leu	Gly	Phe	Val	Gly	Ala	Gly	Leu	Gly	Leu	Thr	Ile	Ser	Arg	Tyr	Ile
		275					280					285			
Gly	Ala	Val	Ala	Ile	Leu	Trp	Val	Leu	Ala	Ile	Gly	Phe	Asn	Pro	Ala
	290					295					300				
Leu	Arg	Ile	Ser	Leu	Lys	Ser	Tyr	Phe	Lys	Pro	Leu	Asn	Phe	Ser	Ile
305					310					315					320
Ile	Trp	Glu	Val	Met	Gly	Ile	Gly	Ile	Pro	Ala	Ser	Val	Glu	Ser	Val
				325					330					335	
Leu	Phe	Thr	Ser	Gly	Arg	Leu	Leu	Thr	Gln	Met	Phe	Val	Ala	Gly	Met
			340					345					350		

ES 2 526 490 T3

Gly Thr Ser Val Ile Ala Gly Asn Phe Ile Ala Phe Ser Ile Ala Ala
 355 360 365

Leu Ile Asn Leu Pro Gly Ser Ala Leu Gly Ser Ala Ser Thr Ile Ile
 370 375 380

Thr Gly Arg Arg Leu Gly Val Gly Gln Ile Ala Gln Ala Glu Ile Gln
 385 390 395 400

Leu Arg His Val Phe Trp Leu Ser Thr Leu Gly Leu Thr Ala Ile Ala
 405 410 415

Trp Leu Thr Ala Pro Phe Ala Gly Val Met Ala Ser Phe Tyr Thr Gln
 420 425 430

Asp Pro Gln Val Lys His Val Val Ile Leu Ile Trp Leu Asn Ala
 435 440 445

Leu Phe Met Pro Ile Trp Ser Ala Ser Trp Val Leu Pro Ala Gly Phe
 450 455 460

Lys Gly Ala Arg Asp Ala Arg Tyr Ala Met Trp Val Ser Met Leu Ser
 465 470 475 480

Met Trp Gly Cys Arg Val Val Val Gly Tyr Val Leu Gly Ile Met Leu
 485 490 495

Gly Trp Gly Val Val Gly Val Trp Met Gly Met Phe Ala Asp Trp Ala
 500 505 510

Val Arg Ala Val Leu Phe Tyr Trp Arg Met Val Thr Gly Arg Trp Leu
 515 520 525

Trp Lys Tyr Pro Arg Pro Glu Pro Gln Lys Cys Glu Lys Lys Pro Val
 530 535 540

Val Ser Glu
 545

<210> 11

<211> 472

5 <212> PRT

<213> Bacillus subtilis 168

<400> 11

ES 2 526 490 T3

Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45
 Glu Lys Tyr Ser Val Ala Val Ile Lys Asp Lys Asp Tyr Phe Lys Ser
 50 55 60
 Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80
 His Asn Lys Pro Glu Glu Glu Val Val Glu Gln Ile Val Lys Val Ala
 85 90 95
 Glu Met Phe Gly Ala Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
 100 105 110
 Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly
 115 120 125
 Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala
 130 135 140
 Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr
 145 150 155 160
 Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr Pro Leu Ile
 165 170 175
 Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Thr
 180 185 190
 Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn Asp Tyr Leu
 195 200 205
 Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile
 210 215 220
 Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu
 225 230 235 240
 Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu
 245 250 255
 Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr
 260 265 270
 Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys
 275 280 285
 Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln
 290 295 300
 Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Pro
 305 310 315 320
 Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro

ES 2 526 490 T3

325 330 335
 Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350
 Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp
 355 360 365
 Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380
 Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu
 385 390 395 400
 Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Val
 405 410 415
 Ser Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445
 Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460
 Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 12
 <211> 472
 <212> PRT
 <213> Bacillus subtilis ATCC6633

5

<400> 12

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

10

ES 2 526 490 T3

Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala
130 135 140

Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr
145 150 155 160

Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr Pro Leu Ile
165 170 175

Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Thr
180 185 190

Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn Asp Tyr Leu
195 200 205

Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile
210 215 220

Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu
225 230 235 240

Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu
245 250 255

Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr
260 265 270

Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys
275 280 285

Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln
290 295 300

Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Pro
305 310 315 320

Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
325 330 335

Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp
340 345 350

Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp
355 360 365

Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
370 375 380

Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu
385 390 395 400

Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Val
405 410 415

Thr Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
420 425 430

Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
435 440 445

Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
450 455 460

Thr Ala Lys Tyr Val Leu Pro Val
465 470

- 5 <210> 13
- <211> 472
- <212> PRT
- <213> Bacillus subtilis IAM1213

ES 2 526 490 T3

<400> 13

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

ES 2 526 490 T3

Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr
 260 265 270

Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys
 275 280 285

Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln
 290 295 300

Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Pro
 305 310 315 320

Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
 325 330 335

Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350

Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp
 355 360 365

Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380

Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu
 385 390 395 400

Ala Ile Asp Leu Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Val
 405 410 415

Ser Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430

Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445

Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460

Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 14

<211> 472

5 <212> PRT

<213> Bacillus subtilis IAM1107

<400> 14

Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15

Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30

Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45

Glu Lys Tyr Ser Val Ala Val Val Lys Asp Lys Asp Tyr Phe Lys Ser
 50 55 60

10

ES 2 526 490 T3

Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80
 His Asn Lys Pro Glu Glu Glu Val Val Glu Gln Ile Val Lys Val Ala
 85 90 95
 Glu Met Phe Gly Ala Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
 100 105 110
 Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly
 115 120 125
 Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala
 130 135 140
 Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr
 145 150 155 160
 Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr Pro Leu Ile
 165 170 175
 Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Thr
 180 185 190
 Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn Asp Tyr Leu
 195 200 205
 Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile
 210 215 220
 Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu
 225 230 235 240
 Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu
 245 250 255
 Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr
 260 265 270
 Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys
 275 280 285
 Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln
 290 295 300
 Asn Cys Ala Thr His Thr Glu Val Lys Leu Met Lys Asn Arg Glu Pro
 305 310 315 320
 Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
 325 330 335
 Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350
 Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp
 355 360 365
 Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380

ES 2 526 490 T3

Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu
 385 390 395 400
 Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Phe
 405 410 415
 Ser Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445
 Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460
 Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 15

<211> 472

5 <212> PRT

<213> Bacillus subtilis IAM1214

<400> 15

Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45
 Glu Lys Tyr Ser Val Ala Val Ile Lys Asp Lys Asp Tyr Phe Gln Ser
 50 55 60
 Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80
 His Asp Lys Pro Glu Glu Glu Val Val Glu Gln Ile Val Lys Val Ala
 85 90 95
 Gln Met Phe Glu Ala Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
 100 105 110
 Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly
 115 120 125
 Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala
 130 135 140
 Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr
 145 150 155 160
 Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr Pro Leu Ile
 165 170 175
 Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Thr
 180 185 190

10

ES 2 526 490 T3

Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn Asp Tyr Leu
 195 200 205

Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile
 210 215 220

Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu
 225 230 235 240

Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu
 245 250 255

Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr
 260 265 270

Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys
 275 280 285

Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln
 290 295 300

Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Pro
 305 310 315 320

Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
 325 330 335

Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350

Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp
 355 360 365

Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380

Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu
 385 390 395 400

Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Val
 405 410 415

Thr Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430

Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445

Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460

Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 16

<211> 472

5 <212> PRT

<213> Bacillus subtilis ATCC21555

<400> 16

ES 2 526 490 T3

Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45
 Glu Lys Tyr Ser Ile Ala Val Ile Lys Asp Lys Asp Tyr Phe Lys Ser
 50 55 60
 Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80
 His Asp Lys Pro Glu Glu Glu Val Val Glu Glu Ile Val Lys Val Ala
 85 90 95
 Asp Met Phe Gly Val Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
 100 105 110
 Ala Pro Met Ala Lys Ala Cys Lys Arg Leu Gly Leu Arg Gly Ala Gly
 115 120 125
 Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Ala Ala
 130 135 140
 Phe Asn Arg Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr
 145 150 155 160
 Leu Glu Asp Phe Arg Ala Ala Leu Gln Glu Ile Gly Thr Pro Leu Ile
 165 170 175
 Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Lys
 180 185 190
 Glu Met Glu Thr Ala Glu Ala Glu Phe Asn Arg Val Asn Glu Tyr Leu
 195 200 205
 Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile
 210 215 220
 Ala Glu Glu Phe Leu Gln Gly Glu Tyr Asp Asp Trp Tyr Glu Thr Ser
 225 230 235 240
 Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu
 245 250 255
 Tyr Phe Pro Val Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr
 260 265 270
 Glu Thr Ala His Ile Thr Pro Ser Ile Leu Asp Asp Asp Ala Lys Arg
 275 280 285
 Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Glu
 290 295 300
 Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Ala
 305 310 315 320

ES 2 526 490 T3

Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
 325 330 335
 Asn Ile Lys Lys Val Phe Gly Val Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350
 Val Leu Cys Tyr Gly Lys Glu Ala Asp Leu Pro Lys Gly Leu Leu Glu
 355 360 365
 Gln Glu Pro Cys Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380
 Lys Glu Asn Gly Gln Leu Pro Glu Thr Val Val Asp Phe Val Ile Glu
 385 390 395 400
 Ser Ile Glu Ile Pro Asp Gly Val Leu Lys Gly Asp Thr Glu Leu Val
 405 410 415
 Ser Phe Ser Ala Ala Glu Ala Gly Thr Ser Val Asp Leu Arg Leu Phe
 420 425 430
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445
 Asn Asp Val Ala Glu Ser Ile Lys Gln Ile Gln Gln Gln Ala Lys Leu
 450 455 460
 Thr Ala Lys Tyr Ala Leu Ser Val
 465 470

<210> 17
 <211> 472
 5 <212> PRT
 <213> Bacillus amyloliquefaciens IF03022

<400> 17

Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45
 Glu Lys Tyr Ser Val Ala Val Ile Lys Asp Lys Asp Tyr Phe Lys Ser
 50 55 60
 Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80
 His Asp Lys Pro Glu Glu Glu Val Val Glu Glu Ile Val Lys Val Ala
 85 90 95
 Gly Met Phe Ala Val Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
 100 105 110
 10 Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly

ES 2 526 490 T3

115 120 125

Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Ala Ala
130 135 140

Phe Asn Arg Ala Gly Val Lys Ser Ile Lys Asn Arg Arg Val Thr Thr
145 150 155 160

Leu Glu Asp Phe Arg Ala Ala Leu Gln Glu Ile Gly Thr Pro Leu Ile
165 170

Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Lys
180 185 190

Glu Arg Glu Thr Ala Glu Ala Glu Phe Asn Arg Val Asn Glu Tyr Leu
195 200 205

Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile
210 215 220

Ala Glu Glu Phe Leu Gln Gly Glu Tyr Asp Asp Trp Tyr Glu Thr Ser
225 230 235 240

Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu
245 250 255

Tyr Phe Pro Val Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr
260 265 270

Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Asp Asp Ala Lys Arg
275 280 285

Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Glu
290 295 300

Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Ala
305 310 315 320

Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
325 330 335

Asn Ile Lys Lys Val Phe Gly Val Asp Met Ala Gln Leu Leu Leu Asp
340 345 350

Val Leu Cys Phe Gly Lys Glu Ala Asp Leu Pro Lys Gly Leu Leu Glu
355 360 365

Gln Glu Pro Cys Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
370 375 380

Lys Glu Asn Gly Gln Leu Pro Glu Thr Ala Val Asp Phe Val Ile Glu
385 390 395 400

Ser Ile Asp Ile Pro Asp Gly Val Leu Lys Gly Asp Thr Glu Ile Val
405 410 415

Ser Phe Ser Ala Ala Glu Ala Gly Thr Ser Val Asp Leu Arg Leu Phe
420 425 430

Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
435 440 445

Gly Asp Val Ala Glu Ser Ile Lys Gln Ile Gln Gln Gln Ala Lys Leu
450 455 460

Thr Ala Lys Tyr Ala Leu Pro Val
465 470

- 5 <210> 18
- <211> 476
- <212> PRT
- <213> Bacillus pumilus NRRL B-12025

ES 2 526 490 T3

<400> 18

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

5

ES 2 526 490 T3

245 250 255
 Glu Gly Glu Tyr Leu Pro Phe Val Ile His Asp Lys Thr Pro Gln Ile
 260 265 270
 Gly Phe Thr Glu Thr Ala His Ile Thr Pro Thr Ile Leu Asp Asn Glu
 275 280 285
 Ala Lys Gln Ile Ile Ile Glu Ala Ala Arg Lys Ala Asn Glu Gly Leu
 290 295 300
 Gly Leu Glu His Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn
 305 310 315 320
 Arg Glu Thr Gly Leu Ile Glu Ala Ala Ala Arg Phe Ala Gly Trp Asn
 325 330 335
 Met Ile Pro Asn Ile Lys Lys Val Phe Gly Val Asp Met Ala Lys Leu
 340 345 350
 Leu Ile Asp Val Leu Val Asp Gly Lys Lys Ala Val Leu Pro Lys Gln
 355 360 365
 Leu Leu Ser Gly His Thr Phe Tyr Val Ala Asp Cys His Leu Tyr Pro
 370 375 380
 Gln His Phe Lys Glu Ser Gly Leu Ile Pro Pro Glu Ala Thr His Ile
 385 390 395 400
 Thr Ile Asp His Val Ser Ile Pro Gln Glu Ala Phe Val Gly Asp Thr
 405 410 415
 Ala Ile Val Ser Gln Ser Phe Pro Ala Lys Gly Thr Ile Val Asp Leu
 420 425 430
 Glu Leu Phe Glu Ala Phe Asn Gly Ile Val Ser Leu Glu Leu Lys Gly
 435 440 445
 Ser Ser Ser Gln Asp Val Ala Ala Ser Ile Arg Asn Ile Gln Lys Gln
 450 455 460
 Ala Thr Ile Gln Leu Met Asp Glu Leu Val Lys Gly
 465 470 475

<210> 19

<211> 93

5 <212> PRT

<213> Bacillus subtilis 168

<400> 19

Gly Ala Gly Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met
 1 5 10 15
 Arg Asp Ala Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg
 20 25 30
 Val Thr Thr Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr
 35 40 45
 Pro Leu Ile Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr
 50 55 60
 Leu Ile Thr Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn
 65 70 75 80
 Asp Tyr Leu Lys Ser Ile Asn Val Pro Lys Ala Val Thr
 85 90

10

ES 2 526 490 T3

<210> 20
 <211> 1416
 <212> DNA
 <213> Bacillus subtilis 168

5

<400> 20

```

atg gag aga aaa aca gta ttg gtc atc gct gat ctt gga ggc tgc ccg 48
Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
  1                               10                          15

ccg cac atg ttt tat aaa agc gct gct gaa aaa tat aac ctg gtc agc 96
Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
                20                          25                          30

ttt att cca aga cct ttt gca att aca gcc tcc cat gca gca ttg att 144
Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
                35                          40                          45

gaa aaa tac tcg gtc gcg gtc ata aaa gat aaa gac tat ttt aag agt 192
Glu Lys Tyr Ser Val Ala Val Ile Lys Asp Lys Asp Tyr Phe Lys Ser
                50                          55                          60

tta gct gat ttt gaa cac cct gat tcc att tat tgg gcg cat gaa gat 240
Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
  65                          70                          75                          80

cat aac aag cct gag gaa gag gtc gtc gag caa atc gtc aag gtt gcc 288
His Asn Lys Pro Glu Glu Val Val Glu Gln Ile Val Lys Val Ala
                85                          90                          95

gaa atg ttt ggg gcg gat gcc atc aca aca aac aat gaa tta ttc att 336
Glu Met Phe Gly Ala Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
                100                          105                          110

gct ccg atg gcg aaa gcc tgt gaa cgt ctg ggc ttg aga ggt gcc ggc 384
Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly
                115                          120                          125

gtg cag gca gcc gaa aat gcc aga gat aaa aat aaa atg agg gac gct 432
Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala
                130                          135                          140

ttt aat aag gcc gga gtc aaa tcg atc aaa aac aaa cga gtc aca act 480
Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr
                145                          150                          155                          160

ctt gaa gat ttc cgt gct gct ctt gaa gag atc ggc aca cct ctt atc 528
Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr Pro Leu Ile
                165                          170                          175
    
```

ES 2 526 490 T3

tta aag cct aca tac tta gcg agt tct atc ggt gta acg ctg att acg	576
Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Thr	
180 185 190	
gac act gag acg gca gaa gat gaa ttt aac aga gtc aat gac tat ctg	624
Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn Asp Tyr Leu	
195 200 205	
aaa tca att aac gtg cca aag gcg gtt acg ttt gaa gcg ccg ttt atc	672
Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile	
210 215 220	
gct gaa gaa ttt tta cag ggt gag tac gga gac tgg tat caa aca gaa	720
Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu	
225 230 235 240	
ggg tac tcc gac tat atc agt ata gaa ggc atc atg gct gac ggt gag	768
Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu	
245 250 255	
tat ttc ccg atc gcc att cat gat aaa acg ccg caa atc ggg ttt aca	816
Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr	
260 265 270	
gag aca tcc cac att acg ccg tcc att ctg gat gaa gag gca aaa aag	864
Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys	
275 280 285	
aaa att gtc gaa gct gcc aaa aag gca aat gaa ggg ctt gga ctg caa	912
Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln	
290 295 300	
aat tgc gca aca cat aca gag atc aag cta atg aaa aac aga gaa ccg	960
Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Pro	
305 310 315 320	
ggt tta ata gag tcc gca gcc aga ttt gcc ggc tgg aat atg atc ccc	1008
Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro	
325 330 335	
aat att aaa aag gtc ttt ggc ctt gat atg gcg caa tta tta tta gat	1056
Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp	
340 345 350	
gtc ctc tgt ttc gga aaa gac gcc gat ctg ccg gac gga tta ttg gat	1104
Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp	
355 360 365	
caa gag cct tat tat gtt gcc gac tgc cat ttg tac ccg cag cat ttc	1152
Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe	
370 375 380	
aaa caa aat ggc caa att cct gaa act gct gag gat ttg gtc att gaa	1200
Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu	
385 390 395 400	
gcg atc gat att ccg gac ggg ctt tta aaa ggg gat act gaa atc gtt	1248
Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Val	
405 410 415	
tct ttt tcc gcc gca gca cca ggc act tca gtt gat ttg aca ttg ttt	1296

ES 2 526 490 T3

Ser Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430
 gaa gct ttc aat tcc att gct gca ttt gaa ctg aaa ggc agt aat tca 1344
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445
 cag gat gtg gct gaa tca atc aga caa att cag cag cat gcg aag ctg 1392
 Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460
 acg gca aag tat gtg ctg cca gta 1416
 Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 21

<211> 1416

5 <212> DNA

<213> Bacillus subtilis ATCC6633

<400> 21

atg gag aga aaa aca gta ttg gtc atc gct gat ctt gga ggc tgc ccg 48
 Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15
 ccg cac atg ttt tat aaa agc gct gct gaa aaa tat aac ctg gtt agc 96
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30
 ttt att ccg aga cct ttt gca ata aca gcc tcc cat gca gca ctg att 144
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45
 gaa aaa tac tcg gtc gcg gtc ata aaa gat aaa gac tat ttt cag agc 192
 Glu Lys Tyr Ser Val Ala Val Ile Lys Asp Lys Asp Tyr Phe Gln Ser
 50 55 60
 tta gct gat ttt gag cat ccc gat tca att tat tgg gcg cat gag gat 240
 Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80
 cat gac aag cct gaa gaa gag gtt gtc gag caa atc gtc aag gtt gcc 288
 His Asp Lys Pro Glu Glu Glu Val Val Glu Gln Ile Val Lys Val Ala
 85 90 95
 caa atg ttt gag gcg gac gcc atc aca aca aac aat gaa tta ttc att 336
 Gln Met Phe Glu Ala Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
 100 105 110
 gcc ccg atg gcg aaa gcc tgt gaa cgc ctt ggc ctg agg ggc gcc gga 384
 Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly
 115 120 125
 gtg cag gca gcg gaa aat gcc aga gat aaa aat aaa atg agg gac gct 432
 Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala
 130 135 140
 ttt aat aag gcg gga gtc aaa tcg atc aaa aac aaa cga gtc aca act 480
 Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr

10

ES 2 526 490 T3

145	150	155	160	
ctt gag gat ttt cgt gct gca ctt gaa gag atc ggc aca cct cta atc				528
Leu Glu Asp Phe	Arg Ala Ala	Leu Glu Ile Gly Thr	Pro Leu Ile	
	165	170	175	
tta aag cct aca tac tta gcg agt tca atc ggc gta acg ctg att acc				576
Leu Lys Pro Thr Tyr	Leu Ala Ser Ser	Ile Gly Val Thr	Leu Ile Thr	
	180	185	190	
gac acg gag acg gca gaa gat gaa ttt aac aga gtc aat gac tac ctg				624
Asp Thr Glu Thr Ala Glu Asp	Glu Phe Asn Arg	Val Asn Asp Tyr	Leu	
	195	200	205	
aaa tcg att aac gtg ccg aag gcg gtc aca ttt gaa gca ccg ttt att				672
Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe	Glu Ala Pro Phe	Ile		
	210	215	220	
gct gag gaa ttt tta cag ggt gag tac gga gac tgg tat caa aca gaa				720
Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu				
	225	230	235	240
ggg tac tcc gac tat atc agc ata gaa ggc att atg gca gat ggt gag				768
Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu				
	245	250	255	
tat ttt ccg atc gcc att cat gac aaa acg ccg caa att gga ttt aca				816
Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr				
	260	265	270	
gag aca tca cat att acg cca tcc att ctg gat gaa gag gcg aaa aag				864
Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys				
	275	280	285	
aaa att gtc gaa gcg gct aaa aag gca aat gaa ggg ctt gga ctg caa				912
Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln				
	290	295	300	
aat tgc gca aca cat aca gaa atc aag cta atg aaa aac aga gaa ccg				960
Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Pro				
	305	310	315	320
ggt tta ata gag tcg gct gcc aga ttc gca ggc tgg aat atg att cct				1008
Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro				
	325	330	335	
aac att aaa aag gtt ttc ggc ctt gat atg gcg caa tta tta tta gat				1056
Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp				
	340	345	350	
gtt ctc tgt ttc gga aaa gat gct gat ctg ccg gac ggg tta ttg gat				1104
Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp				
	355	360	365	
caa gag cct tac tat gtt gct gac tgc cat ctg tac cct cag cat ttc				1152
Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe				
	370	375	380	
aaa caa aat ggc cag atc cct gaa act gcc gag gat ttg gta atc gaa				1200
Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu				
	385	390	395	400

ES 2 526 490 T3

gcg atc gat att ccg gat ggg ctt ttg aag ggt gat aca gaa atc gtt 1248
 Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Val
 405 410 415

act ttt tcg gct gcg gca cca gga aca tca gtt gat ttg aca ctg ttt 1296
 Thr Phe Ser Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430

gaa gcc ttc aac tcc att gct gca ttt gaa ctg aaa ggc agc aat tca 1344
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445

cag gat gtg gct gaa tca atc aga caa att cag cag cat gcg aag ctg 1392
 Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460

acg gca aag tat gtg ctg cca gta 1416
 Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 22
 <211> 1416
 <212> DNA
 <213> Bacillus subtilis IAM1213

5

<400> 22

atg gag aga aaa aca gta ttg gtc atc gct gat ctt gga ggc tgc ccg 48
 Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15

ccg cac atg ttt tat aaa agc gct gct gaa aaa tat aac ctg gtc agc 96
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30

ttt att cca aga cct ttt gca att aca gcc tcc cat gca gca ttg att 144
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45

gaa aaa tac tcg gtc gcg gtc ata aaa gat aaa gac tat ttt aag agt 192
 Glu Lys Tyr Ser Val Ala Val Ile Lys Asp Lys Asp Tyr Phe Lys Ser
 50 55 60

tta gct gat ttt gag cat cct gac tcc att tat tgg gcg cat gag gat 240
 Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80

cat aac aag cct gag gaa gag gtc gtc gag caa atc gtc aag gtt gcc 288
 His Asn Lys Pro Glu Glu Glu Val Val Glu Gln Ile Val Lys Val Ala
 85 90 95

gaa atg ttt ggg gcg gat gcc atc aca aca aac aat gaa tta ttc att 336
 Glu Met Phe Gly Ala Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile
 100 105 110

gct ccg atg gcg aaa gcc tgt gaa cgt ctg ggc ctg aga ggt gcc ggc 384
 Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly
 115 120 125

10

ES 2 526 490 T3

gtg	cag	gca	gcc	gaa	aat	gcc	aga	gat	aaa	aat	aaa	atg	agg	gac	gct	432
Val	Gln	Ala	Ala	Glu	Asn	Ala	Arg	Asp	Lys	Asn	Lys	Met	Arg	Asp	Ala	
	130					135					140					
ttt	aat	aag	gcc	gga	gtc	aaa	tcg	atc	aaa	aac	aaa	cga	gtc	aca	act	480
Phe	Asn	Lys	Ala	Gly	Val	Lys	Ser	Ile	Lys	Asn	Lys	Arg	Val	Thr	Thr	
	145				150					155					160	
ctc	gaa	gat	ttc	cgt	gct	gct	ctt	gaa	gag	atc	ggc	aca	cct	ctt	atc	528
Leu	Glu	Asp	Phe	Arg	Ala	Ala	Leu	Glu	Glu	Ile	Gly	Thr	Pro	Leu	Ile	
				165					170					175		
tta	aag	cct	aca	tac	tta	gcg	agt	tca	atc	ggg	gta	acg	ctg	att	acg	576
Leu	Lys	Pro	Thr	Tyr	Leu	Ala	Ser	Ser	Ile	Gly	Val	Thr	Leu	Ile	Thr	
			180					185						190		
gac	act	gag	acg	gca	gaa	gat	gaa	ttt	aac	aga	gtc	aat	gac	tat	ctg	624
Asp	Thr	Glu	Thr	Ala	Glu	Asp	Glu	Phe	Asn	Arg	Val	Asn	Asp	Tyr	Leu	
		195					200					205				
aaa	tca	att	aac	gtg	cca	aag	gcg	ggt	acg	ttt	gaa	gcg	ccg	ttt	atc	672
Lys	Ser	Ile	Asn	Val	Pro	Lys	Ala	Val	Thr	Phe	Glu	Ala	Pro	Phe	Ile	
	210					215					220					
gct	gaa	gaa	ttt	tta	cag	ggt	gag	tac	gga	gac	tgg	tat	caa	aca	gaa	720
Ala	Glu	Glu	Phe	Leu	Gln	Gly	Glu	Tyr	Gly	Asp	Trp	Tyr	Gln	Thr	Glu	
	225				230					235					240	
ggg	tac	tcc	gac	tat	atc	agt	ata	gaa	ggc	atc	atg	gct	gac	ggg	gag	768
Gly	Tyr	Ser	Asp	Tyr	Ile	Ser	Ile	Glu	Gly	Ile	Met	Ala	Asp	Gly	Glu	
				245					250					255		
tat	ttc	ccg	atc	gcc	att	cat	gat	aaa	acg	ccg	caa	atc	ggg	ttt	aca	816
Tyr	Phe	Pro	Ile	Ala	Ile	His	Asp	Lys	Thr	Pro	Gln	Ile	Gly	Phe	Thr	
			260					265						270		
gag	aca	tcc	cac	att	acg	ccg	tcc	att	ctg	gat	gaa	gag	gca	aaa	aag	864
Glu	Thr	Ser	His	Ile	Thr	Pro	Ser	Ile	Leu	Asp	Glu	Glu	Ala	Lys	Lys	
		275					280					285				
aaa	att	gtc	gaa	gct	gcc	aaa	aag	gca	aat	gaa	ggt	ctt	ggc	ctg	caa	912
Lys	Ile	Val	Glu	Ala	Ala	Lys	Lys	Ala	Asn	Glu	Gly	Leu	Gly	Leu	Gln	
	290					295					300					
aat	tgc	gca	aca	cat	aca	gag	atc	aag	cta	atg	aaa	aat	aga	gaa	ccg	960
Asn	Cys	Ala	Thr	His	Thr	Glu	Ile	Lys	Leu	Met	Lys	Asn	Arg	Glu	Pro	
	305				310					315					320	
ggt	tta	att	gaa	tcg	gca	gcc	aga	ttc	gcc	ggc	tgg	aat	atg	atc	ccc	1008
Gly	Leu	Ile	Glu	Ser	Ala	Ala	Arg	Phe	Ala	Gly	Trp	Asn	Met	Ile	Pro	
				325					330					335		
aat	att	aaa	aag	gtc	ttt	ggc	ctt	gat	atg	gcg	caa	tta	tta	tta	gat	1056
Asn	Ile	Lys	Lys	Val	Phe	Gly	Leu	Asp	Met	Ala	Gln	Leu	Leu	Leu	Asp	
			340					345					350			
gtc	ctt	tgt	ttc	gga	aaa	gac	gcc	gat	ctg	ccg	gac	gga	tta	ttg	gat	1104
Val	Leu	Cys	Phe	Gly	Lys	Asp	Ala	Asp	Leu	Pro	Asp	Gly	Leu	Leu	Asp	
		355					360					365				
caa	gag	cct	tat	tat	gtt	gcc	gac	tgc	cat	ttg	tac	ccg	caa	cat	ttc	1152

ES 2 526 490 T3

Gln	Glu	Pro	Tyr	Tyr	Val	Ala	Asp	Cys	His	Leu	Tyr	Pro	Gln	His	Phe		
	370					375					380						
aaa	caa	aat	ggc	cag	att	cca	gaa	act	gct	gag	gat	ttg	gtc	att	gaa	1200	
Lys	Gln	Asn	Gly	Gln	Ile	Pro	Glu	Thr	Ala	Glu	Asp	Leu	Val	Ile	Glu		
385				390					395					400			
gcg	atc	gat	ctg	cct	gac	ggg	ctt	tta	aaa	ggg	gat	act	gag	atc	gtt	1248	
Ala	Ile	Asp	Leu	Pro	Asp	Gly	Leu	Leu	Lys	Gly	Asp	Thr	Glu	Ile	Val		
			405					410					415				
tct	ttt	tcg	gcc	gca	gca	cca	gga	act	tca	gtt	gat	ttg	aca	ttg	ttt	1296	
Ser	Phe	Ser	Ala	Ala	Ala	Pro	Gly	Thr	Ser	Val	Asp	Leu	Thr	Leu	Phe		
			420					425					430				
gaa	gct	ttc	aat	tcc	att	gct	gca	ttt	gaa	ctg	aaa	ggc	agt	aat	tca	1344	
Glu	Ala	Phe	Asn	Ser	Ile	Ala	Ala	Phe	Glu	Leu	Lys	Gly	Ser	Asn	Ser		
		435					440					445					
cag	gat	gtg	gct	gaa	tca	atc	aga	caa	att	cag	cag	cat	gcg	aag	ctg	1392	
Gln	Asp	Val	Ala	Glu	Ser	Ile	Arg	Gln	Ile	Gln	Gln	His	Ala	Lys	Leu		
	450					455					460						
acg	gca	aag	tat	gtg	ctg	cca	gta									1416	
Thr	Ala	Lys	Tyr	Val	Leu	Pro	Val										
465					470												
<210>	23																
<211>	1416																
<212>	DNA																
<213>	Bacillus subtilis IAM1107																
<400>	23																
atg	gag	aga	aaa	aca	gta	ttg	gtc	atc	gct	gat	ctt	gga	ggc	tgc	ccg	48	
Met	Glu	Arg	Lys	Thr	Val	Leu	Val	Ile	Ala	Asp	Leu	Gly	Gly	Cys	Pro		
1				5				10						15			
ccg	cac	atg	ttt	tat	aaa	agc	gct	gct	gaa	aaa	tat	aac	ctg	gtc	agc	96	
Pro	His	Met	Phe	Tyr	Lys	Ser	Ala	Ala	Glu	Lys	Tyr	Asn	Leu	Val	Ser		
			20					25					30				
ttt	att	cca	aga	cct	ttt	gca	att	aca	gcc	tcc	cat	gca	gca	ttg	att	144	
Phe	Ile	Pro	Arg	Pro	Phe	Ala	Ile	Thr	Ala	Ser	His	Ala	Ala	Leu	Ile		
		35				40						45					
gaa	aaa	tac	tcg	gtc	gcg	gtc	gta	aaa	gat	aaa	gac	tat	ttt	aag	agt	192	
Glu	Lys	Tyr	Ser	Val	Ala	Val	Val	Lys	Asp	Lys	Asp	Tyr	Phe	Lys	Ser		
	50					55					60						
tta	gct	gat	ttt	gag	cat	cct	gac	tcc	att	tat	tgg	gcg	cat	gag	gat	240	
Leu	Ala	Asp	Phe	Glu	His	Pro	Asp	Ser	Ile	Tyr	Trp	Ala	His	Glu	Asp		
	65				70					75					80		
cat	aac	aag	cct	gag	gaa	gag	gtc	gtc	gag	caa	atc	gtc	aag	gtt	gcc	288	
His	Asn	Lys	Pro	Glu	Glu	Glu	Val	Val	Glu	Gln	Ile	Val	Lys	Val	Ala		
			85					90						95			
gaa	atg	ttc	ggg	gcg	gat	gcc	atc	aca	aca	aac	aat	gaa	tta	ttc	att	336	
Glu	Met	Phe	Gly	Ala	Asp	Ala	Ile	Thr	Thr	Asn	Asn	Glu	Leu	Phe	Ile		

5

10

ES 2 526 490 T3

	100	105	110	
	gct ccg atg gcg aaa gcc tgt gaa cgt ctg ggc ttg aga ggt gcc ggc Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly 115	120	125	384
	gtg cag gca gcc gaa aat gcc aga gat aaa aat aaa atg agg gac gct Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala 130	135	140	432
	ttt aat aag gcc gga gtc aaa tcg atc aaa aac aaa cga gtc aca act Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr 145	150	155	480
	ctt gaa gat ttc cgt gct gct ctt gaa gag atc ggc aca cct ctt atc Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr Pro Leu Ile 165	170	175	528
	tta aag cct aca tac tta gcg agt tct atc ggt gta acg ctg att acg Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Thr 180	185	190	576
	gac act gag acg gca gaa gat gaa ttt aac aga gtc aat gac tat ctg Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn Asp Tyr Leu 195	200	205	624
	aaa tca att aac gtg cca aag gcg gtt acg ttt gaa gcg ccg ttt atc Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile 210	215	220	672
	gct gaa gaa ttt tta cag ggt gag tac gga gac tgg tat caa aca gaa Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu 225	230	235	720
	ggg tac tcc gac tat atc agt ata gaa ggc atc atg gct gac ggt gag Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu 245	250	255	768
	tat ttc ccg atc gcc att cat gat aaa acg ccg caa atc ggg ttt aca Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr 260	265	270	816
	gag aca tcc cac att acg ccg tcc att ctg gat gaa gag gca aaa aag Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys 275	280	285	864
	aaa att gtc gaa gct gcc aaa aag gca aat gaa ggg ctt ggc ctg caa Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln 290	295	300	912
	aat tgc gca aca cat aca gag gtc aag cta atg aaa aac aga gaa ccg Asn Cys Ala Thr His Thr Glu Val Lys Leu Met Lys Asn Arg Glu Pro 305	310	315	960
	ggt tta att gaa tcg gca gcc aga ttt gcc ggc tgg aat atg atc cct Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro 325	330	335	1008
	aac att aaa aag gtt ttc ggc ctt gat atg gcg caa tta tta tta gat Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp 340	345	350	1056

ES 2 526 490 T3

gtc ctc tgt ttc gga aaa gat gcc gat ctg ccg gac gga tta ttg gat 1104
 Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp
 355 360 365

caa gag oct tac tat gtc gcc gac tgc cat ttg tac ccg cag cat ttc 1152
 Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380

aaa caa aat ggc cag att cca gaa acc gct gag gat ttg gtc att gaa 1200
 Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu
 385 390 395 400

gcg atc gat att ccg gac ggg ctt tta aaa ggg gat act gaa atc ttt 1248
 Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Phe
 405 410 415

tct ttt tcg gcc gca gca cca ggc act tca gtt gat ttg aca ttg ttt 1296
 Ser Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430

gaa gct ttc aat tcc att gct gca ttt gaa ctg aaa ggc agt aat tca 1344
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445

cag gat gtg gct gaa tca atc aga caa att cag cag cat gcg aag ctg 1392
 Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460

acg gca aag tat gtg ctg cca gta 1416
 Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 24
 <211> 1416
 5 <212> DNA
 <213> Bacillus subtilis IAM1214

<400> 24

atg gag aga aaa aca gta ttg gtc atc gct gat ctt gga ggc tgc ccg 48
 Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15

ccg cac atg ttt tat aaa agc gct gct gaa aaa tat aac ctg gtt agc 96
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30

ttt att ccg aga cct ttt gca ata aca gcc tcc cat gca gca ctg att 144
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45

gaa aaa tac tcg gtc gcg gtc ata aaa gat aaa gac tat ttt cag agc 192
 Glu Lys Tyr Ser Val Ala Val Ile Lys Asp Lys Asp Tyr Phe Gln Ser
 50 55 60

tta gct gat ttt gag cat ccc gat tca att tat tgg gcg cat gag gat 240
 Leu Ala Asp Phe Glu His Pro Asp Ser Ile Tyr Trp Ala His Glu Asp
 65 70 75 80

10

ES 2 526 490 T3

cat gac aag cct gaa gaa gag gtt gtc gag caa atc gtc aag gtt gcc	288
His Asp Lys Pro Glu Glu Glu Val Val Glu Gln Ile Val Lys Val Ala	
85 90 95	
caa atg ttt gag gcg gac gcc atc aca aca aac aat gaa tta ttc att	336
Gln Met Phe Glu Ala Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile	
100 105 110	
gcc ccg atg gcg aaa gcc tgt gaa cgc ctt ggc ctg agg gcc gcc gga	384
Ala Pro Met Ala Lys Ala Cys Glu Arg Leu Gly Leu Arg Gly Ala Gly	
115 120 125	
gtg cag gca gcg gaa aat gcc aga gat aaa aat aaa atg agg gac gct	432
Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Asp Ala	
130 135 140	
ttt aat aag gcg gga gtc aaa tcg atc aaa aac aaa cga gtc aca act	480
Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg Val Thr Thr	
145 150 155 160	
ctt gag gat ttt cgt gct gca ctt gaa gag atc ggc aca cct cta atc	528
Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr Pro Leu Ile	
165 170 175	
tta aag cct aca tac tta gcg agt tca atc ggc gta acg ctg att acc	576
Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr Leu Ile Thr	
180 185 190	
gac acg gag acg gca gaa gat gaa ttt aac aga gtc aat gac tac ctg	624
Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn Asp Tyr Leu	
195 200 205	
aaa tcg att aac gtg ccg aag gcg gtc aca ttt gaa gca ccg ttt att	672
Lys Ser Ile Asn Val Pro Lys Ala Val Thr Phe Glu Ala Pro Phe Ile	
210 215 220	
gct gag gaa ttt tta cag ggt gag tac gga gac tgg tat caa aca gaa	720
Ala Glu Glu Phe Leu Gln Gly Glu Tyr Gly Asp Trp Tyr Gln Thr Glu	
225 230 235 240	
ggg tac tcc gac tat atc agc ata gaa ggc att atg gca gat ggt gag	768
Gly Tyr Ser Asp Tyr Ile Ser Ile Glu Gly Ile Met Ala Asp Gly Glu	
245 250 255	
tat ttt ccg atc gcc att cat gac aaa acg ccg caa att gga ttt aca	816
Tyr Phe Pro Ile Ala Ile His Asp Lys Thr Pro Gln Ile Gly Phe Thr	
260 265 270	
gag aca tca cat att acg cca tcc att ctg gat gaa gag gcg aaa aag	864
Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Glu Glu Ala Lys Lys	
275 280 285	
aaa att gtc gaa gcg gct aaa aag gca aat gaa ggg ctt gga ctg caa	912
Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Gln	
290 295 300	
aat tgc gca aca cat aca gaa atc aag cta atg aaa aac aga gaa ccg	960
Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Pro	
305 310 315 320	
ggg tta ata gag tcg gct gcc aga ttc gca ggc tgg aat atg att cct	1008

ES 2 526 490 T3

Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
 325 330 335

aac att aaa aag gtt ttc ggc ctt gat atg gcg caa tta tta tta gat 1056
 Asn Ile Lys Lys Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350

gtt ctc tgt ttc gga aaa gat gct gat ctg ccg gac ggg tta ttg gat 1104
 Val Leu Cys Phe Gly Lys Asp Ala Asp Leu Pro Asp Gly Leu Leu Asp
 355 360 365

caa gag cct tac tat gtt gct gac tgc cat ctg tac cct cag cat ttc 1152
 Gln Glu Pro Tyr Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380

aaa caa aat ggc cag atc cct gaa act gcc gag gat ttg gta atc gaa 1200
 Lys Gln Asn Gly Gln Ile Pro Glu Thr Ala Glu Asp Leu Val Ile Glu
 385 390 395 400

gcg atc gat att ccg gat ggg ctt ttg aag ggt gat aca gaa atc gtt 1248
 Ala Ile Asp Ile Pro Asp Gly Leu Leu Lys Gly Asp Thr Glu Ile Val
 405 410 415

act ttt tcg gct gcg gca cca gga aca tca gtt gat ttg aca ctg ttt 1296
 Thr Phe Ser Ala Ala Ala Pro Gly Thr Ser Val Asp Leu Thr Leu Phe
 420 425 430

gaa gcc ttc aac tcc att gct gca ttt gaa ctg aaa ggc agc aat tca 1344
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445

cag gat gtg gct gaa tca atc aga caa att cag cag cat gcg aag ctg 1392
 Gln Asp Val Ala Glu Ser Ile Arg Gln Ile Gln Gln His Ala Lys Leu
 450 455 460

acg gca aag tat gtg ctg cca gta 1416
 Thr Ala Lys Tyr Val Leu Pro Val
 465 470

<210> 25

<211> 1416

5 <212> DNA

<213> Bacillus subtilis ATCC21555

<400> 25

atg gag aga aaa aca gta ttg gtt atc gct gat ctt ggg ggc tgc ccg 48
 Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15

ccg cat atg ttt tac aaa agc gca gcc gaa aaa tac aac ctc gtc agc 96
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30

ttt att ccg aga ccc ttt gca att aca gcc tct cat gcg gcc tta att 144
 Phe Ile Pro Arg Pro Phe Ala Ile Thr Ala Ser His Ala Ala Leu Ile
 35 40 45

gaa aaa tac tcg att gcg gtc att aaa gat aaa gac tat ttt aag agt 192
 Glu Lys Tyr Ser Ile Ala Val Ile Lys Asp Lys Asp Tyr Phe Lys Ser

10

ES 2 526 490 T3

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

ES 2 526 490 T3

aac tgt gca acg cat aca gaa ata aaa tta atg aaa aac cgg gaa gcc 960
 Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Ala
 305 310 315 320

gga ctg att gag tca gcg gcc aga ttc gcg gga tgg aat atg att ccg 1008
 Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
 325 330 335

aat att aaa aag gtc ttc ggc gtt gat atg gcg cag cta tta ttg gat 1056
 Asn Ile Lys Lys Val Phe Gly Val Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350

gtt ctc tgt tac gga aaa gaa gct gat ctg ccg aaa gga tta ttg gag 1104
 Val Leu Cys Tyr Gly Lys Glu Ala Asp Leu Pro Lys Gly Leu Leu Glu
 355 360 365

cag gag cca tgc tat gtc gca gac tgc cac ttg tat cct cag cat ttc 1152
 Gln Glu Pro Cys Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380

aaa gag aac ggc cag ctg cct gag acg gtt gtc gat ttc gtc att gaa 1200
 Lys Glu Asn Gly Gln Leu Pro Glu Thr Val Val Asp Phe Val Ile Glu
 385 390 395 400

agc att gaa att cct gac ggc gtc tta aag gga gac act gaa ctc gtt 1248
 Ser Ile Glu Ile Pro Asp Gly Val Leu Lys Gly Asp Thr Glu Leu Val
 405 410 415

tct ttc tca gcg gct gag gcg ggt acg tca gtg gat ctg ccg ctg ttc 1296
 Ser Phe Ser Ala Ala Glu Ala Gly Thr Ser Val Asp Leu Arg Leu Phe
 420 425 430

gaa gcg ttc aac agc att gcg gcg ttt gag ctg aaa gga agc aat tcg 1344
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445

aac gac gtg gcc gaa tca atc aaa caa att cag cag cag gcg aag ctg 1392
 Asn Asp Val Ala Glu Ser Ile Lys Gln Ile Gln Gln Gln Ala Lys Leu
 450 455 460

act gca aag tat gcg tta tcg gta 1416
 Thr Ala Lys Tyr Ala Leu Ser Val
 465 470

<210> 26

<211> 1416

5 <212> DNA

<213> Bacillus amyloliquefaciens IF03022

<400> 26

atg gag aga aaa aca gta ttg gtt atc gct gac ctt ggg gga tgc ccg 48
 Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro
 1 5 10 15

ccg cat atg ttt tac aaa agc gca gcc gaa aaa tac aac ctc gtc agc 96
 Pro His Met Phe Tyr Lys Ser Ala Ala Glu Lys Tyr Asn Leu Val Ser
 20 25 30

10

ES 2 526 490 T3

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

ES 2 526 490 T3

Glu Thr Ser His Ile Thr Pro Ser Ile Leu Asp Asp Asp Ala Lys Arg
 275 280 285

aaa atc gtc gaa gca gcc aaa aag gcg aat gaa gga ctc ggc ctc gaa 912
 Lys Ile Val Glu Ala Ala Lys Lys Ala Asn Glu Gly Leu Gly Leu Glu
 290 295 300

aac tgc gca acc cat aca gag att aaa tta atg aaa aac cgg gaa gcc 960
 Asn Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn Arg Glu Ala
 305 310 315 320

gga ctg att gaa tca gcg gca cga ttt gcg ggc tgg aac atg att ccg 1008
 Gly Leu Ile Glu Ser Ala Ala Arg Phe Ala Gly Trp Asn Met Ile Pro
 325 330 335

aat att aaa aag gtc ttc ggc gtc gat atg gcg cag ctg tta ttg gat 1056
 Asn Ile Lys Lys Val Phe Gly Val Asp Met Ala Gln Leu Leu Leu Asp
 340 345 350

gtt ctc tgt ttc gga aaa gaa gcc gat ctg ccg aaa gga tta ttg gag 1104
 Val Leu Cys Phe Gly Lys Glu Ala Asp Leu Pro Lys Gly Leu Leu Glu
 355 360 365

cag gag ccg tgc tat gtc gcc gac tgc cac ttg tat cct cag cat ttc 1152
 Gln Glu Pro Cys Tyr Val Ala Asp Cys His Leu Tyr Pro Gln His Phe
 370 375 380

aaa gag aac ggc cag ctg cct gag acg gct gtc gat ttc gtc att gaa 1200
 Lys Glu Asn Gly Gln Leu Pro Glu Thr Ala Val Asp Phe Val Ile Glu
 385 390 395 400

agc att gac att ccc gac ggc gtc tta aag gga gac acc gaa atc gtt 1248
 Ser Ile Asp Ile Pro Asp Gly Val Leu Lys Gly Asp Thr Glu Ile Val
 405 410 415

tct ttc tcg gcg gcc gag gcg ggt aca tcc gtg gat ctg cgg ctg ttc 1296
 Ser Phe Ser Ala Ala Glu Ala Gly Thr Ser Val Asp Leu Arg Leu Phe
 420 425 430

gaa gcg ttc aac agc att gcg gcg ttc gag ctg aaa gga agc aat tcg 1344
 Glu Ala Phe Asn Ser Ile Ala Ala Phe Glu Leu Lys Gly Ser Asn Ser
 435 440 445

ggt gac gtg gcc gaa tca atc aaa caa att cag cag cag gcg aag ctg 1392
 Gly Asp Val Ala Glu Ser Ile Lys Gln Ile Gln Gln Gln Ala Lys Leu
 450 455 460

act gca aag tat gcg tta ccg gta 1416
 Thr Ala Lys Tyr Ala Leu Pro Val
 465 470

<210> 27

<211> 1428

5 <212> DNA

<213> Bacillus pumilus NRRL B-12025

<400> 27

10 gtg ctt tca ttg agt aaa aaa act gta ctt gtc att gct gac tta gga 48
 Val Leu Ser Leu Ser Lys Lys Thr Val Leu Val Ile Ala Asp Leu Gly

ES 2 526 490 T3

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

ES 2 526 490 T3

gag ggc gaa tat ctc cct ttt gtc ata cat gat aaa acc cct caa atc 816
 Glu Gly Glu Tyr Leu Pro Phe Val Ile His Asp Lys Thr Pro Gln Ile
 260 265 270

ggc ttt aca gaa acg gct cat atc act ccg acg atc tta gac aat gaa 864
 Gly Phe Thr Glu Thr Ala His Ile Thr Pro Thr Ile Leu Asp Asn Glu
 275 280 285

gcc aag caa atc atc att gaa gca gca agg aag gca aat gaa ggg cta 912
 Ala Lys Gln Ile Ile Ile Glu Ala Ala Arg Lys Ala Asn Glu Gly Leu
 290 295 300

ggt ctt gaa cat tgt gca acc cat aca gaa atc aaa ctc atg aaa aat 960
 Gly Leu Glu His Cys Ala Thr His Thr Glu Ile Lys Leu Met Lys Asn
 305 310 315 320

cga gaa act gga ctg atc gag gca gcg gct cga ttc gct ggc tgg aat 1008
 Arg Glu Thr Gly Leu Ile Glu Ala Ala Arg Phe Ala Gly Trp Asn
 325 330 335

atg atc ccg aat att aaa aaa gtc ttt ggc gtc gat atg gcg aag cta 1056
 Met Ile Pro Asn Ile Lys Lys Val Phe Gly Val Asp Met Ala Lys Leu
 340 345 350

ttg att gat gta tta gtt gat ggt aaa aag gct gta ctg cca aaa cag 1104
 Leu Ile Asp Val Leu Val Asp Gly Lys Lys Ala Val Leu Pro Lys Gln
 355 360 365

ctg ctt tct gga cat aca ttt tat gta gcg gac tgc cac ctg tac cct 1152
 Leu Leu Ser Gly His Thr Phe Tyr Val Ala Asp Cys His Leu Tyr Pro
 370 375 380

cag cat ttt aaa gag agt ggg ctt atc ccg cct gaa gcc aca cat att 1200
 Gln His Phe Lys Glu Ser Gly Leu Ile Pro Pro Glu Ala Thr His Ile
 385 390 395 400

acc att gat cat gtg tct att ccg cag gaa gca ttc gtt gga gat act 1248
 Thr Ile Asp His Val Ser Ile Pro Gln Glu Ala Phe Val Gly Asp Thr
 405 410 415

gcg att gtc agt caa tca ttc cct gcc aaa ggg act att gtg gat ctt 1296
 Ala Ile Val Ser Gln Ser Phe Pro Ala Lys Gly Thr Ile Val Asp Leu
 420 425 430

gaa tta ttt gaa gct ttt aat gga atc gta tct ctt gaa tta aaa gga 1344
 Glu Leu Phe Glu Ala Phe Asn Gly Ile Val Ser Leu Glu Leu Lys Gly
 435 440 445

tca tcc tca caa gat gtt gcc gcg tcc atc cgc aac att cag aaa cag 1392
 Ser Ser Ser Gln Asp Val Ala Ala Ser Ile Arg Asn Ile Gln Lys Gln
 450 455 460

gca acg att cag tta atg gat gaa tta gtg aag gga 1428
 Ala Thr Ile Gln Leu Met Asp Glu Leu Val Lys Gly
 465 470 475

<210> 28

<211> 1416

5 <212> DNA

<213> Bacillus subtilis ATCC 15245 y Bacillus subtilis IAM 1033

<400> 28

ES 2 526 490 T3

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

ES 2 526 490 T3

Ala	Glu	Glu	Phe	Leu	Gln	Gly	Glu	Tyr	Gly	Asp	Trp	Tyr	Gln	Thr	Glu	
225					230					235					240	
ggg	tac	tcc	gac	tat	atc	agt	ata	gaa	ggc	atc	atg	gct	gac	ggg	gag	768
Gly	Tyr	Ser	Asp	Tyr	Ile	Ser	Ile	Glu	Gly	Ile	Met	Ala	Asp	Gly	Glu	
				245					250					255		
tat	ttc	ccg	atc	gcc	att	cat	gat	aaa	acg	ccg	caa	atc	ggg	ttt	aca	816
Tyr	Phe	Pro	Ile	Ala	Ile	His	Asp	Lys	Thr	Pro	Gln	Ile	Gly	Phe	Thr	
			260				265						270			
gag	aca	tcc	cac	att	acg	ccg	tcc	att	ctg	gat	gaa	gag	gca	aaa	aag	864
Glu	Thr	Ser	His	Ile	Thr	Pro	Ser	Ile	Leu	Asp	Glu	Glu	Ala	Lys	Lys	
		275					280					285				
aaa	att	gtc	gaa	gct	gcc	aaa	aag	gca	aat	gaa	ggg	ctt	ggc	ctg	caa	912
Lys	Ile	Val	Glu	Ala	Ala	Lys	Lys	Ala	Asn	Glu	Gly	Leu	Gly	Leu	Gln	
	290					295					300					
aat	tgc	gca	aca	cat	aca	gag	atc	aag	cta	atg	aaa	aac	aga	gaa	ccg	960
Asn	Cys	Ala	Thr	His	Thr	Glu	Ile	Lys	Leu	Met	Lys	Asn	Arg	Glu	Pro	
305					310					315					320	
ggg	ttt	ccc	atc	gcc	att	cat	gat	aaa	acg	ccg	caa	atc	ggg	ttt	aca	816
Gly	Leu	Ile	Glu	Ser	Ala	Ala	Arg	Phe	Ala	Gly	Trp	Asn	Met	Ile	Pro	
			325						330					335		
aat	att	aaa	aag	gtc	ttt	ggc	ctt	gat	atg	gcg	caa	tta	tta	tta	gat	1056
Asn	Ile	Lys	Lys	Val	Phe	Gly	Leu	Asp	Met	Ala	Gln	Leu	Leu	Leu	Asp	
			340					345						350		
gtc	ctc	tgt	ttc	gga	aaa	gac	gcc	gat	ctg	ccg	gac	gga	tta	ttg	gat	1104
Val	Leu	Cys	Phe	Gly	Lys	Asp	Ala	Asp	Leu	Pro	Asp	Gly	Leu	Leu	Asp	
		355					360					365				
caa	gag	cct	tat	tat	gtt	gcc	gac	tgc	cat	ttg	tac	ccg	cag	cat	ttc	1152
Gln	Glu	Pro	Tyr	Tyr	Val	Ala	Asp	Cys	His	Leu	Tyr	Pro	Gln	His	Phe	
	370				375						380					
aaa	caa	aat	ggc	cag	att	cca	gaa	acc	gct	gag	gat	ttg	gtc	att	gaa	1200
Lys	Gln	Asn	Gly	Gln	Ile	Pro	Glu	Thr	Ala	Glu	Asp	Leu	Val	Ile	Glu	
385					390					395					400	
gcg	atc	gat	att	ccg	gac	ggg	ctt	tta	aaa	ggg	gat	act	gaa	atc	ggt	1248
Ala	Ile	Asp	Ile	Pro	Asp	Gly	Leu	Leu	Lys	Gly	Asp	Thr	Glu	Ile	Val	
			405						410					415		
tca	ttt	tca	gcc	gca	gca	cca	ggc	act	tca	gtt	gat	ttg	aca	ttg	ttt	1296
Ser	Phe	Ser	Ala	Ala	Ala	Pro	Gly	Thr	Ser	Val	Asp	Leu	Thr	Leu	Phe	
			420					425					430			
gaa	gct	ttc	aat	tcc	att	gct	gca	ttt	gaa	ctg	aaa	ggc	agt	aat	tca	1344
Glu	Ala	Phe	Asn	Ser	Ile	Ala	Ala	Phe	Glu	Leu	Lys	Gly	Ser	Asn	Ser	
		435				440						445				
cag	gat	gtg	gct	gaa	tca	atc	aga	caa	att	cag	cag	cat	gca	aag	ctg	1392
Gln	Asp	Val	Ala	Glu	Ser	Ile	Arg	Gln	Ile	Gln	Gln	His	Ala	Lys	Leu	
	450					455					460					
acg	gca	aag	tat	gtg	ctg	cca	gta									1416
Thr	Ala	Lys	Tyr	Val	Leu	Pro	Val									
465					470											

<210> 29

<211> 279

5 <212> DNA

<213> Bacillus subtilis 168

<400> 29

ES 2 526 490 T3

```

ggg gcc ggc gtg cag gca gcc gaa aat gcc aga gat aaa aat aaa atg   48
Gly Ala Gly Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met
  1                    5                    10                    15

agg gac gct ttt aat aag gcc gga gtc aaa tcg atc aaa aac aaa cga   96
Arg Asp Ala Phe Asn Lys Ala Gly Val Lys Ser Ile Lys Asn Lys Arg
                20                    25                    30

gtc aca act ctt gaa gat ttc cgt gct gct ctt gaa gag atc ggc aca   144
Val Thr Thr Leu Glu Asp Phe Arg Ala Ala Leu Glu Glu Ile Gly Thr
                35                    40                    45

cct ctt atc tta aag cct aca tac tta gcg agt tct atc ggt gta acg   192
Pro Leu Ile Leu Lys Pro Thr Tyr Leu Ala Ser Ser Ile Gly Val Thr
                50                    55                    60

ctg att acg gac act gag acg gca gaa gat gaa ttt aac aga gtc aat   240
Leu Ile Thr Asp Thr Glu Thr Ala Glu Asp Glu Phe Asn Arg Val Asn
  65                    70                    75                    80

gac tat ctg aaa tca att aac gtg cca aag gcg gtt acg               279
Asp Tyr Leu Lys Ser Ile Asn Val Pro Lys Ala Val Thr
                85                    90

```

```

5  <210> 30
   <211> 29
   <212> DNA
   <213> Secuencia Artificial

10 <220>
   <223> Descripción de la Secuencia Artificial: DNA sintético

   <400> 30
ttaaagcttt gactttcagg agcccgttg                29

15 <210> 31
   <211> 29
   <212> DNA
   <213> Secuencia Artificial

20 <220>
   <223> Descripción de la Secuencia Artificial: DNA sintético

   <400> 31
tttgagctca tatatcgggg gaatgatag                29

30 <210> 32
   <211> 29
   <212> DNA
   <213> Secuencia Artificial

   <220>
   <223> Descripción de la Secuencia Artificial: DNA sintético

35 <400> 32
ttaaagctta aaaggtgttc acgtgcaga                29

40 <210> 33
   <211> 29
   <212> DNA
   <213> Secuencia Artificial

   <220>
45 <223> Descripción de la Secuencia Artificial: DNA sintético

```

ES 2 526 490 T3

	<400> 33	
	tttggatcct tagcgggatg ctcgttgca	29
5	<210> 34 <211> 30 <212> DNA <213> Secuencia Artificial	
10	<220> <223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 34	
15	atagaattct taacagttga ttcgtagtc	30
	<210> 35 <211> 30 <212> DNA	
20	<213> Secuencia Artificial	
	<220> <223> Descripción de la Secuencia Artificial: DNA sintético	
25	<400> 35	
	ataggatcct caacaaagcg cgggctgccc	30
	<210> 36 <211> 30 <212> DNA	
30	<213> Secuencia Artificial	
	<220> <223> Descripción de la Secuencia Artificial: DNA sintético	
35	<400> 36	
	tttgaattcg gtgataatga aaaggcaaag	30
40	<210> 37 <211> 29 <212> DNA	
	<213> Secuencia Artificial	
45	<220> <223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 37	
50	tttctgcagt taaacgggct gccctgat	29
	<210> 38 <211> 29 <212> DNA	
55	<213> Secuencia Artificial	
	<220> <223> Descripción de la Secuencia Artificial: DNA sintético	
60	<400> 38	
	tttgaattcg ttgttgaggc acatcttaa	29
65	<210> 39	

	<211> 29	
	<212> DNA	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 39	
10	tttctgcagt tattccgaca caactggct	29
	<210> 40	
	<211> 56	
	<212> DNA	
15	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: DNA sintético	
20	<400> 40	
	atggagttta gtgtaaaaag cggtagcccg gagaaagtg aggctggagc tgcttc	56
	<210> 41	
25	<211> 56	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
30	<223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 41	
	ttactcttcg ccgtaaacc cagcgcgggt taacagcata tgaatattct ccttag	56
35	<210> 42	
	<211> 56	
	<212> DNA	
	<213> Secuencia Artificial	
40	<220>	
	<223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 42	
45	catcatggat atttcacgat aacgtaagt tgcaccgtg aggctggagc tgcttc	56
	<210> 43	
	<211> 56	
50	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: DNA sintético	
55	<400> 43	
	agatgccgga ggaggttga acatctccg gctaccata tgaatattct ccttag	56
60	<210> 44	
	<211> 56	
	<212> DNA	
	<213> Secuencia Artificial	
65	<220>	

	<223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 44	
5	gttgagcggc tgccagagcc tttagccgag gaatcagtgt aggctggagc tgcttc	56
	<210> 45	
	<211> 56	
	<212> DNA	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: DNA sintético	
15	<400> 45	
	ctgccagctt gcccgaccca gttcacgctc tgcggtcata tgaatattct ccttag	56
	<210> 46	
20	<211> 56	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
25	<223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 46	
	ctggacgatg tccgcgaagc actggccgaa gtcggtgtgt aggctggagc tgcttc	56
30	<210> 47	
	<211> 56	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Descripción de la Secuencia Artificial: DNA sintético	
	<400> 47	
40	tgccgcgtcg tcctctcac cggtaggat gccaatcata tgaatattct ccttag	56

REIVINDICACIONES

1. Un microorganismo procariótico que sobreexpresa cualquiera de entre ligasa de aminoácido L, prolina imino peptidasa e hidrolasa de amida de aminoácido L, donde dichas ligasa de aminoácido L, prolina imino peptidasa e hidrolasa de amida de aminoácido L tienen actividad de síntesis de dipéptidos y
- 5 en las que la actividad de una proteína para transportar un dipéptido de una célula microbiana hacia el exterior de dicha célula microbiana (referida a partir de este punto como actividad de transporte de dipéptidos) es superior a la de la cepa original, mediante la transformación de la cepa original con un ADN que codifica una proteína con actividad de transporte de dipéptidos,
- 10 donde dicha proteína que tiene actividad de transporte de dipéptidos se describe en cualquiera de los siguientes apartados [1] a [3]:7
- [1] una proteína que tiene una secuencia de aminoácidos mostrada por cualquiera de las SEQ ID NO: 6 a 10,
- [2] una proteína que consiste en una secuencia de aminoácidos en la que se han eliminado, sustituido o añadido entre 1 y 20 aminoácidos de la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10, y que tiene actividad de transporte de dipéptidos,
- 15 [3] una proteína que tiene una secuencia de aminoácidos con un 80% o más de homología con respecto a la secuencia de aminoácidos mostrada en cualquiera de las SEQ ID NO: 6 a 10, y que tiene actividad de transporte de dipéptidos.
2. El microorganismo de la reivindicación 1, que se obtiene transformando la cepa original con un ADN descrito en cualquiera de los siguientes apartados [1] a [3]:
- 20 [1] un ADN que codifica una proteína de cualquiera de los puntos [1] a [3] de la reivindicación 1,
- [2] un ADN que tiene la secuencia de nucleótidos de la región codificadora tal como se muestra en cualquiera de las SEQ ID NO: 1 a 5,
- [3] un ADN que se hibrida con un ADN que consiste en una secuencia de nucleótidos que es complementaria a una secuencia de nucleótidos de una región codificadora de la secuencia de nucleótidos mostrada en cualquiera de las SEQ ID NO: 1 a 5 en condiciones severas, y que codifica la proteína que tiene actividad de transporte de dipéptidos.
- 25
3. El microorganismo de la reivindicación 1 ó 2, que pertenece al género *Escherichia*, *Corynebacterium*, *Bacillus*, *Serratia*, *Pseudomonas* o *Streptomyces*.
4. Un proceso para producir un dipéptido que comprende:
- 30 permitir que un cultivo o un cultivo tratado del microorganismo de una cualquiera de las reivindicaciones 1 a 3 y un aminoácido, éster de aminoácido o amida de aminoácido estén presentes en un medio acuoso, dejar que se forme y se acumule el dipéptido en el medio acuoso y recuperar el dipéptido del medio acuoso.
5. Un proceso para producir un dipéptido que comprende:
- 35 cultivar el microorganismo de la reivindicación 1 ó 2 y que tiene la capacidad de producir al menos un tipo de aminoácido de entre los aminoácidos que constituyen el dipéptido en un medio, permitir que se forme y se acumule el dipéptido en el medio, y recuperar el dipéptido del cultivo.
6. El proceso de la reivindicación 4 ó 5, donde el microorganismo pertenece al género *Escherichia*, *Corynebacterium*, *Bacillus*, *Serratia*, *Pseudomonas* o *Streptomyces*.