

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 626 571**

51 Int. Cl.:

**C12N 1/20** (2006.01)

**C12N 1/14** (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: **05.11.2004 E 04026279 (2)**

97 Fecha y número de publicación de la concesión europea: **15.03.2017 EP 1529837**

54 Título: **Proceso para producir dipéptidos**

30 Prioridad:

**05.11.2003 JP 2003375823**

**25.06.2004 JP 2004189010**

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

**25.07.2017**

73 Titular/es:

**KYOWA HAKKO BIO CO., LTD. (100.0%)**

**1-6-1, OHTEMACHI, CHIYODA-KU**

**TOKYO, 100-8185, JP**

72 Inventor/es:

**HASHIMOTO, SHIN-ICHI y**

**KAZUHIKO, TABATA**

74 Agente/Representante:

**ARIAS SANZ, Juan**

ES 2 626 571 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

Proceso para producir dipéptidos

- 5 La presente invención se refiere a un microorganismo que produce un dipéptido y un proceso para producir un dipéptido usando el microorganismo.

Los dipéptidos son compuestos que son importantes como alimentos, fármacos, cosméticos y similares.

- 10 Los métodos conocidos para producir dipéptidos incluyen extracción de productos naturales, síntesis química y métodos enzimáticos. La extracción de productos naturales se puede usar solo para producir tipos limitados de dipéptidos, y la productividad es baja porque los contenidos de dipéptidos deseados en productos naturales son bajos. En la síntesis de dipéptidos por los métodos de síntesis química, son necesarias operaciones tal como la introducción y eliminación de grupos protectores para grupos funcionales, y también se forman racematos. Los métodos de síntesis química, por tanto, se consideran que son desventajosos con respecto a coste y eficacia. También son desfavorables desde el punto de vista de higiene medioambiental debido al uso de grandes cantidades de solventes orgánicos y similares.

- 20 Respecto a la síntesis de dipéptidos por los métodos enzimáticos, se conocen los siguientes métodos: un método que utiliza una reacción inversa de proteasa (J. Biol. Chem., 119, 707-720 (1937)); métodos que utilizan aminoacil ARNt sintetasa termoestable (Solicitud de patente japonesa no examinada publicada No. 146539/83, Solicitud de patente japonesa no examinada publicada No. 209991/83, Solicitud de patente japonesa no examinada publicada No. 209992/83 y Solicitud de patente japonesa no examinada publicada No. 106298/84); métodos que utilizan péptido sintetasa no ribosómica (de aquí en adelante denominado NRPS) (Chem. Biol., 7, 373-384 (2000), FEBS Lett., 498, 42-45 (2001), patente en EE UU No. 5.795.738 y patente en EE UU No. 5.652.116); un método que utiliza D-Ala-D-Ala ligasa (Biochemistry, 35, 10464-10471 (1996)); métodos que utilizan bacilisina sintetasa (J. Ind. Microbiol., 2, 201-208 (1987) y Enzyme. Microbiol. Technol., 29, 400-406 (2001)); y un método que utiliza amida de L-aminoácido hidrolasa o prolina iminopeptidasa (panfleto WO03/010307).

- 30 Sin embargo, el método que utiliza la reacción inversa de proteasas requiere la introducción y eliminación de grupos protectores para grupos funcionales de aminoácidos usados como sustratos, que produce dificultades en aumentar la eficacia de la reacción de formación de péptido y prevenir la reacción de degradación de péptido. Los métodos que utilizan aminoacil ARNt sintetasa termoestable tienen los defectos de que la expresión de la enzima y la prevención de reacciones que forman subproductos son difíciles.

- 35 Por otra parte, los métodos que utilizan NRPS, D-Ala-D-Ala ligasa y bacilisina sintetasa no tienen los problemas descritos anteriormente y son capaces de producir dipéptidos que tienen secuencias específicas. Sin embargo, no son métodos eficaces porque implican reacciones que requieren donantes de energía tal como ATP.

- 40 Con respecto al método que utiliza amida de L-aminoácido hidrolasa o prolina iminopeptidasa, se divulga que se pueden producir dipéptidos usando un cultivo de un microorganismo que produce la enzima, células de microorganismo aisladas del cultivo o un material tratado de las células del microorganismo. Sin embargo, las cantidades de dipéptidos producidas por este método no son suficientes.

- 45 Las células vivas tienen sistemas metabólicos en los que las proteínas innecesarias se descomponen a aminoácidos constituyentes y los aminoácidos formados se usan para la síntesis de proteínas necesarias. Como esta función es esencial para la supervivencia y proliferación de células, se sabe que muchos tipos de proteasas y peptidasas existen en organismos vivos.

- 50 Por ejemplo, se sabe que en Escherichia coli, existen al menos siete tipos de dipeptidasas, así como otras proteasas y peptidasas capaces de descomponer dipéptidos (FEMS Microbiol. Rev., 63, 265-276 (1989)). En los ADN cromosómicos de Corynebacterium glutamicum, Bacillus subtilis y Saccharomyces cerevisiae, existen 25 o más, 14 o más, y 20 o más, respectivamente, genes identificados como gen de peptidasa.

- 55 Sin embargo, no se sabe que la introducción de deficiencia en peptidasa en un microorganismo que produce un dipéptido aumente la producción del dipéptido.

- 60 También se sabe que las células vivas tienen sistemas de incorporación de péptidos plurales. Por ejemplo, se ha descrito que Escherichia coli tiene tres sistemas para incorporar dipéptidos (Chem. Biol., 5, 489-504 (1998) y Microbiol., 145, 2891-2901 (1999)). Además, se ha confirmado de información de ADN genómico que los microorganismos que pertenecen a los géneros Escherichia, Bacillus, Corynebacterium y Saccharomyces tienen todos tres o más tipos de sistemas de incorporación de péptidos.

- 65 Sin embargo, no se sabe si los dipéptidos sintetizados en células se descargan fuera de las células o no. Aún se sabe menos que la producción de un dipéptido por un microorganismo productor de dipéptidos aumenta al introducir deficiencia de una proteína implicada en la incorporación del péptido en el microorganismo.

Un objeto de la presente invención es proporcionar un microorganismo que produce un dipéptido y un proceso para producir un dipéptido usando el microorganismo.

5 La presente invención se refiere a los siguientes (1) a (3).

(1) Un proceso para producir un dipéptido, que comprende:  
permitir que una fuente de enzima y uno o más tipos de aminoácidos estén presentes en un medio acuoso, dicha fuente de enzima es

10 (a) un cultivo de un microorganismo; o  
(b) un material tratado del cultivo de (a);  
permitir que el dipéptido se forme y acumule en el medio acuoso; y  
recuperar el dipéptido del medio,

15 en donde dicho microorganismo es un microorganismo en el que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominadas también proteínas transportadoras de péptidos) se reducen o pierden comparado con la cepa parental, dicha cepa parental pertenece al género, Escherichia, Bacillus, Corynebacterium o Saccharomyces, y en donde dicho microorganismo tiene la capacidad para producir un dipéptido, en donde dicha capacidad para producir un dipéptido es la capacidad para producir una proteína según cualquiera de [1] a [4] a continuación:

[1] una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68;  
[2] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se delecionan, sustituyen o añaden en la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido;

25 [3] una proteína que consiste en una secuencia de aminoácidos que tiene el 65% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido; y

30 [4] una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología a la secuencia de aminoácidos mostrada en SEQ ID NO: 33 y que tiene la capacidad de sintetizar un dipéptido,

y en donde el material tratado del cultivo es cultivo concentrado, cultivo seco, células obtenidas centrifugando el cultivo, o un producto obtenido sometiendo las células a secado, liofilización o inmovilización.

35 (2) Un proceso para producir un dipéptido, que comprende:  
permitir que una fuente de enzima, un éster de L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de un microorganismo o un material tratado del cultivo;  
permitir que el dipéptido se forme y acumule en el medio acuoso; y  
recuperar el dipéptido del medio;

40 en donde dicho microorganismo es un microorganismo en el que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominadas también proteínas transportadoras de péptidos) se reducen o pierden comparado con la cepa parental, dicha cepa parental pertenece al género, Escherichia, Bacillus, Corynebacterium o Saccharomyces, y en donde dicho microorganismo tiene la capacidad para producir un dipéptido, en donde la peptidasa es una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4, o una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4 y que tiene actividad peptidasa,  
45 y en donde dicha capacidad para producir un dipéptido es la capacidad para formar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido produciendo prolina iminopeptidasa,  
50 y en donde el material tratado del cultivo es cultivo concentrado, cultivo seco, células obtenidas centrifugando el cultivo, o un producto obtenido sometiendo las células a secado, liofilización o inmovilización.

(3) Un proceso para producir un dipéptido, que comprende:  
permitir que una fuente de enzima, una amida de L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de un microorganismo o un material tratado del cultivo;  
55 permitir que el dipéptido se forme y acumule en el medio acuoso; y  
recuperar el dipéptido del medio,

60 en donde dicho microorganismo es un microorganismo en el que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominadas también proteínas transportadoras de péptidos) se reducen o pierden comparado con la cepa parental, dicha cepa parental pertenece al género, Escherichia, Bacillus, Corynebacterium o Saccharomyces, y en donde dicho microorganismo tiene la capacidad para producir un dipéptido, en donde dicha capacidad para producir un dipéptido es la capacidad para formar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido produciendo una proteína que tiene actividad amida hidrolasa,  
65 y en donde el material tratado del cultivo es cultivo concentrado, cultivo seco, células obtenidas centrifugando el cultivo, o un producto obtenido sometiendo las células a secado, liofilización o inmovilización.

Se describe los siguientes elementos:

- 5 (1) Un microorganismo en el que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominadas también proteínas transportadoras de péptidos) se reducen o pierden y que tiene la capacidad para producir un dipéptido.
- 10 (2) Un microorganismo en el que las actividades de tres o más tipos de peptidasas se reducen o pierden y que tiene la capacidad para producir un dipéptido.
- 15 (3) El microorganismo según los anteriores (1) o (2), en donde la peptidasa es una proteína que tiene una secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4, o una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4 y que tiene actividad peptidasa. Según esto, se prefiere que la peptidasa sea una proteína que tiene al menos el 90%, más preferiblemente al menos el 95%, incluso más preferiblemente al menos el 98% y lo más preferiblemente al menos el 99% de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4 y que tiene actividad peptidasa.
- 20 (4) El microorganismo según los anteriores (1) o (3), en donde la proteína transportadora de péptidos es una proteína que tiene una secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 5 a 9, o una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 5 a 9 y que tiene actividad transportadora de péptidos.
- 25 (5) El microorganismo según cualquiera de los anteriores (1) a (4), en donde el microorganismo que tiene la capacidad para producir un dipéptido es un microorganismo que tiene la capacidad para producir una proteína según cualquiera de [1] a [4] a continuación:
- 30 [1] una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68;  
[2] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se delecionan, sustituyen o añaden en la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido;
- 35 [3] una proteína que consiste en una secuencia de aminoácidos que tiene al menos el 65% o más de homología, preferiblemente al menos el 80%, más preferiblemente al menos el 90%, incluso más preferiblemente al menos el 95%, más preferiblemente al menos el 98% y lo más preferiblemente al menos el 99% de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido; y
- 40 [4] una proteína que tiene una secuencia de aminoácidos que tiene al menos el 80% o más de homología, preferiblemente al menos el 90%, más preferiblemente al menos el 95%, incluso más preferiblemente al menos el 98% y lo más preferiblemente al menos el 99% de homología respecto a la secuencia de aminoácidos mostrada en SEQ ID NO: 33 y que tiene la capacidad de sintetizar un dipéptido.
- 45 (6) El microorganismo según cualquiera de los anteriores (1) a (4), en donde el microorganismo que tiene la capacidad para producir un dipéptido es un microorganismo que porta un ADN según cualquiera de [1] a [4] a continuación:
- [1] ADN que codifica la proteína según cualquiera de [1] a [4] del anterior (5);
- 50 [2] ADN que tiene la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65;
- [3] ADN que hibrida con ADN que tiene la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65 en condiciones rigurosas y que codifica una proteína que tiene la actividad para sintetizar un dipéptido; y
- 55 [4] ADN que tiene una secuencia de nucleótido que tiene al menos el 80% o más de homología, preferiblemente al menos el 90%, más preferiblemente al menos el 95%, incluso más preferiblemente al menos el 98% y lo más preferiblemente al menos el 99% de homología respecto a la secuencia de nucleótidos mostrada en SEQ ID NO: 34 y que codifica una proteína que tiene la actividad para sintetizar un dipéptido
- 60 (7) El microorganismo según cualquiera de los anteriores (1) a (6), en donde el microorganismo que tiene la capacidad para producir un dipéptido es un microorganismo que porta un ADN recombinante en el que el ADN según cualquiera de [1] a [4] del anterior (6) está ligado a un ADN vector.
- (8) El microorganismo según cualquiera de los anteriores (1) a (7), en donde el microorganismo es un microorganismo que pertenece al género, Escherichia, Bacillus, Corynebacterium o Saccharomyces.
- 65

- (9) El microorganismo según cualquiera de los anteriores (1) a (4), en donde el microorganismo que tiene la capacidad para producir un dipéptido es un microorganismo que tiene la capacidad de formar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido.
- 5 (10) El microorganismo según el anterior (9), en donde el microorganismo que tiene la capacidad para formar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido es un microorganismo que produce prolina iminopeptidasa (EC 3.4.11.5).
- 10 (11) El microorganismo según los anteriores (9) o (10), en donde el microorganismo que tiene la capacidad para formar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido es un microorganismo que porta un ADN recombinante en el que el ADN que codifica una proteína que tiene actividad prolina iminopeptidasa está ligado a un ADN vector.
- 15 (12) El microorganismo según cualquiera de los anteriores (9) a (11), en donde el microorganismo es un microorganismo que pertenece al género, Escherichia, Bacillus, Corynebacterium o Saccharomyces.
- 20 (13) El microorganismo según cualquiera de los anteriores (1) a (4), en donde el microorganismo que tiene la capacidad para producir un dipéptido es un microorganismo que tiene la capacidad para formar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido.
- (14) El microorganismo según el anterior (13), en donde el microorganismo que tiene la capacidad para formar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido es un microorganismo que produce una proteína que tiene actividad amida de L-aminoácido hidrolasa.
- 25 (15) El microorganismo según el anterior (13), en donde el microorganismo que tiene la capacidad para formar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido es un microorganismo que porta un ADN recombinante en el que el ADN que codifica una proteína que tiene actividad amida de L-aminoácido hidrolasa está ligado a un ADN vector.
- 30 (16) El microorganismo según cualquiera de los anteriores (13) a (15), en donde el microorganismo es un microorganismo que pertenece al género, Escherichia, Bacillus, Corynebacterium o Saccharomyces.
- (17) Un proceso para producir un dipéptido que comprende:  
 35 permitir que una fuente de enzima y uno o más tipos de aminoácidos estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de un microorganismo según cualquiera de los anteriores (1) a (16) o un material tratado del cultivo;  
 permitir que el dipéptido se forme y acumule en el medio acuoso; y  
 recuperar el dipéptido del medio.
- 40 (18) El proceso según el anterior (17), en donde el dipéptido es un dipéptido representado por la siguiente fórmula (I):
- $$R^1-R^2 \quad (I)$$
- 45 (en donde  $R^1$  y  $R^2$ , que pueden ser iguales o diferentes, representan cada uno un aminoácido).
- (19) Un proceso para producir un dipéptido, que comprende:  
 50 permitir que una fuente de enzima, un éster de L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de un microorganismo según cualquiera de los anteriores (9) a (12) o un material tratado del cultivo;  
 permitir que el dipéptido se forme y acumule en el medio acuoso; y  
 recuperar el dipéptido del medio.
- 55 (20) Un proceso para producir un dipéptido, que comprende:  
 permitir que una fuente de enzima, una amida de L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de un microorganismo según cualquiera de los anteriores (13) a (16) o un material tratado del cultivo;  
 60 permitir que el dipéptido se forme y acumule en el medio acuoso; y  
 recuperar el dipéptido del medio.
- (21) El proceso según cualquiera de los anteriores (17) a (20), en donde el material tratado del cultivo es cultivo concentrado, cultivo seco, células obtenidas centrifugando el cultivo, o un producto obtenido sometiendo las células a secado, liofilización, tratamiento con un tensioactivo, ultrasonificación, fricción mecánica, tratamiento con un solvente, tratamiento enzimático o inmovilización.
- 65

(22) El proceso según cualquiera de los anteriores (17) a (21), que comprende además la formulación del dipéptido producido en un aditivo alimenticio, una composición farmacéutica o un cosmético.

5 Según la presente invención, se proporcionan un microorganismo que produce un dipéptido y un proceso para producir un dipéptido usando el microorganismo.

La figura 1 muestra los pasos para construir el vector de expresión del gen ywfE etiquetado con His, pQE60ywfE.

10 La figura 2 muestra los pasos para construir el vector de expresión aumentada del gen ywfE pPE56.

Explicación de los símbolos

PT5: promotor T5

P<sub>trp</sub>: promotor de triptófano.

15 La presente invención se define por las reivindicaciones.

### 1. Microorganismos

20 Los microorganismos descritos en el presente documento son microorganismo en los que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominadas proteínas transportadoras de péptidos) se reducen o pierden y que tienen la capacidad para producir un dipéptido, o microorganismos en los que las actividades de tres o más tipos de peptidasas se reducen o pierden y que tienen la capacidad para producir un dipéptido.

25 Los microorganismos en los que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas transportadoras de péptidos se reducen o pierden incluyen microorganismos en los que las actividades de uno más tipos arbitrarios de peptidasas y uno o más tipos arbitrarios de proteínas transportadoras de péptidos se reducen o pierden siempre que los microorganismos puedan crecer normalmente, específicamente, microorganismos en los que las actividades de preferiblemente de uno a nueve tipos, más preferiblemente de uno a siete tipos, más preferiblemente de uno a cuatro tipos de peptidasas y preferiblemente de uno a cinco tipos, más preferiblemente de uno o tres tipos, más preferiblemente uno o dos tipos, en particular preferiblemente un tipo de proteína transportadora de péptidos se reducen o pierden.

35 Los ejemplos de tales microorganismos son microorganismos en los que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas transportadoras de péptidos se reducen o pierden porque las secuencias de nucleótidos de uno o más tipos de genes que codifican peptidasas (de aquí en adelante denominados genes de peptidasas) y uno o más tipos de genes que codifican proteínas transportadoras de péptidos (de aquí en adelante denominados genes de proteínas transportadoras de péptidos) entre los genes de peptidasas y los genes de proteínas transportadoras de péptidos que existen en el ADN genómico de los microorganismo están deletados del todo o parcialmente o dichas secuencias de nucleótidos contienen sustituciones o adiciones de nucleótidos.

40 La expresión "la actividad de peptidasa se reduce" significa que la actividad degradante de péptidos/actividad peptidolítica está reducida comparada con la peptidasa que no tiene ninguna de las anteriores delecciones, sustituciones y adiciones de nucleótidos.

45 Preferiblemente, la reducción de la actividad de la peptidasa es al menos el 20%, preferiblemente al menos el 50%, más preferiblemente al menos el 80%, más preferiblemente al menos el 90%, comparada con la peptidasa que no tiene ninguna de las anteriores delecciones, sustituciones y adiciones de nucleótidos, tal como la peptidasa codificada por un gen de tipo salvaje.

50 La actividad degradante de péptidos/actividad peptidolítica de un microorganismo se puede medir permitiendo que un péptido como sustrato y células de microorganismo estén presentes en un medio acuoso, realizándose de esta manera la reacción de degradación de péptidos/reacción peptidolítica, y después determinando la cantidad del péptido restante por un método conocido, por ejemplo, análisis por HPLC.

55 Las peptidasas anteriores pueden ser cualquier proteína que tenga actividad degradante de péptido. Se prefieren proteínas que tienen alta actividad hidrolizante de dipéptido. Son más preferidas las dipeptidasas.

60 Los ejemplos de peptidasas incluyen: las que existen en Escherichia coli tal como PepA que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 1, PepB que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 2, PepD que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 3, PepN que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 4, PepP [No. de acceso en GenBank (de aquí en adelante abreviado como GenBank AAC75946], PepQ (GenBank AAC76850), PepE (GenBank AAC76991), PepT (GenBank AAC74211), Dcp (GenBank AAC74611) e ladA (GenBank AAC77284); las que existen en Bacillus subtilis tal como AmpS (GenBank AF012285), PepT (GenBank X99339), YbaC (GenBank Z99104), YcdD (GenBank Z99105), YjbG (GenBank Z99110), YkvY (GenBank Z99111), YqjE (GenBank Z99116) y YwaD (GenBank Z99123); las que existen

en *Corynebacterium glutamicum* tal como las proteínas que tienen las secuencias de aminoácidos representadas por BAB97732, BAB97858, BAB98080, BAB98880, BAB98892, BAB99013, BAB99598 and BAB99819 (Nos. de registro del banco de datos de ADN de Japón); y las que existen en *Saccharomyces cerevisiae* tal como OCT1 (GenBank NC\_001143), SPC2 (GenBank NC\_003143), SPY2 [base de datos del genoma de *Saccharomyces*] (http://www.yeastgenome.org/) No. de acceso L0002875] y YIM1 (GenBank NC\_001145). Los ejemplos de dipeptidasas incluyen PepA, PepB, PepD y PepN que tienen las secuencias de aminoácidos mostradas en SEQ ID NO: 1 a 4, PepQ, PepE e ladA. Las proteínas que tienen secuencias de aminoácidos que tienen el 80% o más, preferiblemente el 90% o más, más preferiblemente el 95% o más de homología a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4 y que tienen actividad peptidasa también están incluidas en las proteínas que tienen alta actividad degradante de dipéptidos.

La homología entre secuencias de aminoácidos y secuencias de nucleótidos se puede determinar usando el algoritmo BLAST por Karlin y Altschul [Proc. 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 por BLASTN en base a BLAST, los parámetros, por ejemplo, son como sigue: puntuación=100 y longitud de palabra=12. Cuando se analiza una secuencia de aminoácidos por BLASTX en base a BLAST, los parámetros, por ejemplo, son como sigue: puntuación=50 y longitud de palabra=3. Cuando se usan los programas BLAST y Gapped BLAST, se usan los parámetros por defecto de cada programa. Las técnicas específicas para estos análisis son conocidas (http://www.ncbi.nlm.nih.gov).

La expresión "la actividad de una proteína transportadora de péptidos/proteína que incorpora péptidos se reduce" significa que la actividad transportadora de péptidos/actividad de incorporación de péptidos se reduce comparada con una proteína transportadora de péptidos/proteína que incorpora péptidos que no tiene ninguna de las anteriores delecciones, sustituciones y adiciones de nucleótidos.

Preferiblemente, la reducción de la actividad de la peptidasa es al menos el 20%, preferiblemente al menos el 50%, más preferiblemente al menos el 80%, más preferiblemente al menos el 90%, comparada con la peptidasa que no tiene ninguna de las anteriores delecciones, sustituciones y adiciones de nucleótidos, tal como la peptidasa codificada por un gen de tipo salvaje.

La actividad transportadora de péptidos/actividad de incorporación de péptidos de un microorganismo se puede medir permitiendo que un péptido como sustrato y células de microorganismo estén presentes en un medio acuoso, realizándose de esta manera la reacción transportadora de péptidos/reacción de incorporación de péptidos, y después determinando la cantidad de péptido restante por un método conocido, por ejemplo, análisis por HPLC.

Las proteínas transportadoras de péptidos anteriores pueden ser cualquier proteína implicada en el transporte de péptidos de microorganismos, por ejemplo, proteínas codificadas por genes que forman un operón en ADN cromosómico que forma un complejo en la membrana celular para expresar actividad transportadora de dipéptidos y esas que tienen actividad transportadora de péptidos como proteínas individuales. Se prefieren proteínas que tienen alta actividad transportadora de péptidos.

Los ejemplos de proteínas transportadoras de péptidos incluyen: las existentes en *Escherichia coli* tal como DppA que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 5, DppB que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 6, DppC que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 7, DppD que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 8, DppF que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 9, OppA (GenBank AAC76569), OppB (GenBank AAC76568), OppC (GenBank AAC76567), OppD (GenBank AAC76566), OppF (GenBank AAC76565), YddO (GenBank AAC74556), YddP (GenBank AAC74557), YddQ (GenBank AAC74558), YddR (GenBank AAC74559), YddS (GenBank AAC74560), YbiK (GenBank AAC73915), MppA (GenBank AAC74411), SapA (GenBank AAC74376), SapB (GenBank AAC74375), SapC (GenBank AAC74374), SapD (GenBank AAC74373) y SapF (GenBank AAC74372); las existentes en *Bacillus subtilis* tal como DppA (GenBank CAA40002), DppB (GenBank CAA40003), DppC (GenBank CAA40004), DppD (GenBank CAA40005), DppE (GenBank CAA40006), OppA (GenBank CAA39787), OppB (GenBank CAA39788), OppC (GenBank CAA39789), OppD (GenBank CAA39790), OppF (GenBank CAA39791), AppA (GenBank CAA62358), AppB (GenBank CAA62359), AppC (GenBank CAA62360), AppD (GenBank CAA62356), AppF (GenBank CAA62357), Yc1F (GenBank CAB12175) y Yk1D (GenBank CAB13157); las existentes en *Corynebacterium glutamicum* tal como las proteínas que tienen las secuencias de aminoácidos representadas por BAB99048, BAB99383, BAB99384, BAB99385, BAB99713, BAB99714, BAB99715, BAB99830, BAB99831 y BAB99832 (Nos. de registro del banco de datos de ADN de Japón); y las existentes en *Saccharomyces cerevisiae* tal como OPT1 (GenBank NP\_012323), OPT2 (GenBank NP\_015520) y PTR2 (GenBank CAA82172). Los ejemplos de las proteínas que tienen alta actividad transportadora de péptidos incluyen DppA, DppB, DppC, DppD y DppF que tienen las secuencias de aminoácidos mostradas en SEQ ID NO: 5 a 9, y proteínas que tienen secuencias de aminoácidos que tienen el 80% o más, preferiblemente el 90% o más, más preferiblemente el 95% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 5 a 9.

La homología entre secuencias de aminoácidos se puede determinar usando programas tal como BLAST y FASTA descritos anteriormente.

5 El término "hibridar" como se usa en el presente documento se refiere a un proceso en el que polinucleótidos hibridan con la secuencia de ácido nucleico enumerada o partes de la misma. Por tanto, dicha secuencia de ácido nucleico puede ser útil como sondas en análisis por transferencia Northern o Southern de preparaciones de ARN o ADN, respectivamente, o se pueden usar como cebadores oligonucleotídicos en análisis de PCR dependiente de su tamaño respectivo. Preferiblemente, dichos polinucleótidos que hibridan comprenden al menos 10, más preferiblemente al menos 15 nucleótidos mientras que un polinucleótido que hibrida de la presente invención que se va a usar como sonda preferiblemente comprende al menos 200, o lo más preferiblemente al menos 500 nucleótidos.

15 Se sabe bien en la técnica como realizar experimentos de hibridación con moléculas de ácido nucleico, es decir, el experto en la materia sabe que condiciones de hibridación tiene que usar según la presente invención. Se hace referencia a tales condiciones de hibridación en libros de texto estándar tal como Sambrook et al.; Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press, 2ª edición 1989 y 3ª edición 2001; Gerhardt et al.; Methods for General and Molecular Bacteriology; ASM Press, 1994; Lefkovits; Immunology Methods Manual: The Comprehensive Sourcebook of Techniques; Academic Press, 1997; Golemis; Protein-Protein Interactions: A Molecular Cloning Manual; Cold Spring Harbor Laboratory Press, 2002) y otros manuales de laboratorio estándar que conocen los expertos en la materia o como se ha recitado anteriormente. Según la presente invención se prefieren condiciones de hibridación rigurosas.

25 "Condiciones de hibridación rigurosas" se refiere, por ejemplo, a una incubación durante la noche a 42°C en una solución que comprende formamida al 50%, SSC 5x (NaCl 750 mM, citrato de sodio 75 mM), fosfato de sodio 50 mM (pH 7,6), solución de Denhardt 5x, sulfato de dextrano al 10% y ADN de esperma de salmón desnaturalizado, cortado 20 µg/ml, seguido, por ejemplo, por lavar los filtros en SSC 0,2x a aproximadamente 65°C. También se contemplan moléculas de ácido nucleico que hibridan en condiciones menos rigurosas. Los cambios en la rigurosidad de la hibridación y la detección de la señal se logran principalmente mediante la manipulación de la concentración de formamida (menores porcentajes de formamida producen rigurosidad disminuida); condiciones de sal; o temperatura. Por ejemplo, las condiciones de menor rigurosidad incluyen una incubación durante la noche a 37°C en una solución que comprende SSPE 6X (SSPE 20X = NaCl 3 M; NaH<sub>2</sub>PO<sub>4</sub> 0,2 M; EDTA 0,02 M, pH 7,4), SDS al 0,5%, formamida al 30%, ADN bloqueante de esperma de salmón 100 µg/ml; seguido por lavados a 50°C con SSPE 1X, SDS al 0,1%. Además, para alcanzar rigurosidad incluso menor, los lavados realizados después de hibridación rigurosa se pueden hacer a mayor concentración de sal (por ejemplo, SSC 5X). Se debe advertir que se pueden lograr variaciones en las condiciones anteriores mediante la inclusión y/o sustitución de reactivos bloqueantes alternativos usados para suprimir el fondo en experimentos de hibridación. Los reactivos bloqueantes típicos incluyen reactivo de Denhardt, BLOTTO, heparina, ADN de esperma de salmón desnaturalizado, y formulaciones registradas comercialmente disponibles. La inclusión de reactivos bloqueantes específicos puede requerir modificación de las condiciones de hibridación descritas anteriormente, debido a problemas con compatibilidad.

45 Los microorganismos en los que las actividades de tres o más tipos de peptidasas se reducen o pierden incluyen microorganismos en los que las actividades de tres o más tipos arbitrarios de peptidasas se reducen o pierden siempre que el microorganismo puede crecer normalmente, específicamente, microorganismos en los que las actividades de preferiblemente de tres o nueve tipos, más preferiblemente de tres a seis tipos, más preferiblemente tres o cuatro tipos de peptidasas se reducen o pierden.

50 Los ejemplos de peptidasas incluyen las peptidasas y dipeptidasas anteriormente descritas existentes en Escherichia coli, Bacillus subtilis, Corynebacterium glutamicum y Saccharomyces cerevisiae. Las proteínas que consisten en secuencias de aminoácidos que tienen el 80% o más, preferiblemente el 90% o más, más preferiblemente el 95% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4 y que tienen actividad peptidasa también están incluidas en las proteínas que tienen alta actividad degradante de dipéptidos.

55 La homología entre secuencias de aminoácidos se puede determinar usando programas tal como BLAST y FASTA descritos anteriormente.

60 No hay ninguna restricción específica respecto al microorganismo que tiene la capacidad para producir un dipéptido siempre que tenga la capacidad para producir un dipéptido. Los microorganismos adecuados incluyen microorganismos que producen proteínas que tienen la actividad para sintetizar un dipéptido por condensación y ligación de uno o más tipos de aminoácidos, microorganismos que producen proteínas que tienen la actividad para sintetizar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido, y microorganismos que tienen la actividad para sintetizar un dipéptido a partir de una amida de un L-aminoácido y un L-aminoácido.

Los microorganismos que producen proteínas que tienen la actividad para sintetizar un dipéptido por condensación y ligación de uno o más tipos de aminoácidos incluyen microorganismo que producen una proteína seleccionada del grupo que consiste en NRPS, D-Ala-D-Ala ligasa y bacilisina sintetasa.

- 5 Los ejemplos de microorganismo que producen NRPS incluyen procariotas tal como microorganismos del género Bacillus, eucariotas tal como microorganismos del género Penicillium, microorganismos que producen BacA, BacB y BacC (GenBank AF007865), microorganismos que producen TycA, TycB y TycC (GenBank AF004835), microorganismos que producen PcbAB (GenBank M57425), y microorganismos que producen una proteína que tiene una secuencia de aminoácidos que tienen el 80% o más, preferiblemente el 90% o más, más preferiblemente el 95% o más de homología respecto a cualquier proteína seleccionada de BacA, BacB, BacC, TycA, TycB, TycC y PcbAB y que tenga actividad NRPS.

15 Los ejemplos del microorganismo que produce D-Ala-D-Ala ligasa incluyen procariotas que forman peptidoglucanos, microorganismos que producen DdlA (no. de acceso de GenBank M58467), microorganismos que producen DdlB (No. de acceso de GenBank AE000118), microorganismos que producen DdlC (No. de acceso de GenBank D88151), y microorganismos que producen una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se han deletado, sustituido o añadido en la secuencia de aminoácidos de cualquier proteína seleccionada de DdlA, DdlB y DdlC y que tengan actividad D-Ala-D-Ala ligasa.

20 La homología entre secuencias de aminoácidos se puede determinar usando programas tal como BLAST y FASTA descritos anteriormente.

25 Los ejemplos de los microorganismos que producen bacilisina sintetasa incluyen microorganismos que pertenecen al género Bacillus, preferiblemente Bacillus subtilis, Bacillus amyloliquefaciens, Bacillus coagulans, Bacillus licheniformis, Bacillus megaterium y Bacillus pumilus, y microorganismos que producen una proteína seleccionada de los siguientes [1] a [4]:

- [1] una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68;  
 [2] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se deletan, sustituyen o añaden en la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido;  
 [3] una proteína que consiste en una secuencia de aminoácidos que tiene el 65% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido; y  
 [4] una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos mostrada en SEQ ID NO: 33 y que tiene la capacidad de sintetizar un dipéptido.

40 Los microorganismos que producen proteínas que tienen la actividad para sintetizar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido incluyen microorganismos que producen prolina iminopeptidasa, específicamente los que pertenecen a los géneros Bacillus, Corynebacterium y Pseudomonas. Los ejemplos de tales microorganismos son Bacillus subtilis ATCC 6633, Bacillus coagulans EK01 [J. Bacteriol., 174, 7919 (1992)], Corynebacterium glutamicum ATCC 13286, Pseudomonas putida AJ-2402 (FERM BP-8101), Pseudomonas putida ATCC 12633 y Pseudomonas putida AJ-2048 (FERM BP-8123) (microorganismos descritos en el documento WO03/010307). Los microorganismos que producen prolina iminopeptidasa también incluyen Arthrobacter nicotianae [FEMS Microbiol. Lett., 78, 191 (1999)], Escherichia coli (Solicitud de patente japonesa no examinada publicada No. 113887/90), 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 (Solicitud de patente japonesa no examinada publicada No. 121860/97), 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)] y Pseudomonas aeruginosa [Nature, 406, 959 (2000)].

55 Además, los microorganismos que producen prolina iminopeptidasa incluyen microorganismos que tienen la capacidad para producir una proteína seleccionada de los siguientes [1] a [3]:

- [1] prolina iminopeptidasa descrita en cualquiera de documento WO03/010307; FEMS Microbiol. Lett., 78, 191 (1999); Solicitud de patente japonesa no examinada publicada No. 113887/90; 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); Solicitud de patente japonesa no examinada publicada No. 121860/97; 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);  
 [2] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se deletan, sustituyen o añaden en la secuencia de aminoácidos de cualquier prolina iminopeptidasa del anterior [1] y que tiene actividad iminopeptidasa; y

[3] una proteína que consiste en una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos de cualquier prolina iminopeptidasa del anterior [1] y que tiene actividad prolina iminopeptidasa.

5 Los microorganismos que producen proteínas que tienen la actividad para sintetizar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido incluyen microorganismos que producen amida de L-aminoácido hidrolasa, específicamente los que pertenecen a los géneros, Bacillus, Corynebacterium, Erwinia, Rhodococcus, Chryseobacterium, Micrococcus, Pseudomonas, Cryptococcus, Trichosporon, Rhodospiridium, Sporobolomyces, Tremella, Torulaspora, Sterigmatomyces y Rhodotolura. Se prefieren microorganismos que pertenecen a los géneros Bacillus, Corynebacterium y Pseudomonas. Ejemplos más preferidos son Bacillus megaterium AJ3284 (FERM BP-8090), Corynebacterium glutamicum ATCC 13286, Micrococcus luteus ATCC 9341 y Pseudomonas saccharophila ATCC 15946 (microorganismos descritos en el documento WO03/010187).

15 Los microorganismos que producen proteínas que tienen actividad amida de L-aminoácido hidrolasa incluyen microorganismos que tienen la capacidad para producir una proteína de los siguientes [1] o [2]:

[1] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos están deletados, sustituidos o añadido en la secuencia de aminoácidos de la amida de L-aminoácido hidrolasa descrita en el documento WO03/010187 y que tiene actividad amida de L-aminoácido hidrolasa;

[2] una proteína que consiste en una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos de la amida de L-aminoácido hidrolasa descrita en el documento WO03/010187 y que tiene actividad amida de L-aminoácido hidrolasa.

25 La proteína anterior que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos están deletados, sustituidos o añadidos y que tiene la actividad para sintetizar un dipéptido se puede obtener, por ejemplo, introduciendo un mutación dirigida en el ADN que codifica una proteína seleccionada de una proteína que consiste en la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68, una proteína que tiene actividad prolina iminopeptidasa y una proteína que tiene actividad amida de L-aminoácidos hidrolasa por mutagénesis dirigida descrita en Molecular Cloning, A Laboratory Manual, Segunda Edición, Cold Spring Harbor Laboratory Press (1989) (de aquí en adelante denominado Molecular Cloning, Segunda Edición); Current Protocols in Molecular Biology, John Wiley & Sons (1987-1997) (de aquí en adelante denominado Current Protocols in Molecular Biology); 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 deletan, sustituyen o añaden no está específicamente limitado, pero está en el intervalo donde la delección, sustitución o adición es posible por métodos conocidos tal como la anterior mutagénesis dirigida. El número adecuado es de 1 a docenas, preferiblemente de 1 a 20, más preferiblemente de 1 a 10, más preferiblemente de 1 a 5.

45 La expresión "uno o más residuos de aminoácido se deletan, sustituyen o añaden en cualquiera de las secuencias de aminoácidos mostradas en SEQ ID NO: 19 a 25 y 68 y las secuencias de aminoácidos de una proteína que tiene actividad prolina iminopeptidasa y una proteína que tiene actividad amida de L-aminoácido hidrolasa" significa que la secuencia de aminoácidos puede contener delección, sustitución o adición de un único aminoácido o residuos de aminoácidos plurales en una posición arbitraria en las mismas.

50 La delección, sustitución y adición pueden estar simultáneamente contenidas en una secuencia, y los aminoácidos que se van a sustituir o añadir pueden ser naturales o no. Los ejemplos de los aminoácidos naturales con L-alanina, L-asparagina, ácido L-aspartico, L-arginina, L-glutamina, ácido L-glutámico, glicina, L-histidina, L-iso-leucina, L-leucina, L-lisina, L-metionina, L-fenilalanina, L-prolina, L-serina, L-treonina, L-triptófano, L-tirosina, L-valina y L-cisteína.

55 Los siguientes son ejemplos de los aminoácidos capaces de sustitución mutua. Los aminoácidos en el mismo grupo se pueden sustituir mutuamente.

- 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: asparragina, glutamina
- Grupo D: lisina, arginina, ornitina, ácido 2,4-diaminobutanoico, ácido 2,3-aminopropiónico
- Grupo E: prolina, 3-hidroxi prolina, 4-hidroxi prolina
- Grupo F: serina, treonina, homoserina
- 65 Grupo G: fenilalanina, tirosina,

No hay ninguna restricción específica respecto a la posición donde la anterior delección, sustitución o adición de uno o más residuos de aminoácidos se introduce, siempre que una proteína que tiene una secuencia de aminoácidos que tiene la mutación introducida tenga la actividad de sintetizar dipéptidos. Los ejemplos adecuados incluyen residuos de aminoácidos que no están conservados en una o más secuencias de aminoácidos entre las secuencias de aminoácidos mostradas en SEQ ID NO: 19 a 25 y 68 cuando estas secuencias se comparan.

Las proteínas anteriores que consisten en secuencias de aminoácidos en donde uno o más residuos de aminoácidos se delecionan, sustituyen o añaden y que tienen la actividad para sintetizar un dipéptido incluyen una proteína que consiste en una secuencia de aminoácidos que habitualmente tiene el 65% o más, preferiblemente el 80% o más, más preferiblemente el 90% o más, en particular preferiblemente el 95% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68, y una proteína que consiste en una secuencia de aminoácidos que habitualmente tiene el 80% o más, preferiblemente el 90% o más, más preferiblemente el 95% o más de homología respecto a la secuencia de aminoácidos de prolina iminopeptidasa o amida de L-aminoácido hidrolasa.

La homología entre secuencias de aminoácidos y secuencias de nucleótidos se pueden determinar usando programas tales como BLAST y FASTA descritos anteriormente.

La secuencia de aminoácidos mostrada en SEQ ID NO: 33 es una región conservada entre las proteínas que tienen las secuencias de aminoácidos mostradas en SEQ ID NO: 19 a 25 y también es una región correspondiente a la secuencia consenso de proteínas que tienen actividad Ala-Ala ligasa derivadas de varios microorganismos.

Los microorganismos que producen una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más, preferiblemente el 90% o más, más preferiblemente el 95% o más de homología respecto a la secuencia de aminoácidos mostrada en SEQ ID NO: 33 y que tienen la actividad para sintetizar un dipéptido también están incluidas en los microorganismos productores de dipéptidos.

Para que la proteína que tiene una secuencia de aminoácidos que tiene el 80% o más, preferiblemente el 90% o más, más preferiblemente el 95% o más de homología respecto a la secuencia de aminoácidos mostrada en SEQ ID NO: 33 pueda tener la actividad para sintetizar un dipéptido, es deseable que la homología de su secuencia de aminoácidos respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 sea al menos el 80% o más, habitualmente el 90% o más, y particularmente el 95% o más.

La homología entre secuencias de aminoácidos se puede determinar usando programas tales como BLAST y FASTA descritos anteriormente.

Los microorganismos descritos en el presente documento también incluyen microorganismos que tiene un ADN recombinante obtenido por ligación, a un ADN vector, de ADN que codifica una proteína que tiene la actividad para sintetizar un dipéptido por condensación y ligación de uno o más tipos de aminoácidos, el ADN que codifica una proteína que tiene la actividad para sintetizar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido, o el DNA que codifica una proteína que tiene la actividad para sintetizar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido.

Los ejemplos de los microorganismos incluyen los que pertenecen a los géneros Escherichia, Bacillus, Corynebacterium y Saccharomyces.

Los ADN que codifican una proteína que tiene la actividad para sintetizar un dipéptido por condensación y ligación de uno o más tipos de aminoácidos incluyen los ADN que codifican NRPS, D-Ala-D-Ala ligasa o bacilisina sintetasa.

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

Los ejemplos de los ADN que codifican D-Ala-D-Ala ligasa incluyen los ADN que codifican una proteína seleccionada del grupo que consiste en DdlA, DdlB y DdlC.

Los ejemplos de los ADN que codifican bacilisina sintetasa incluyen ADN que codifican proteínas de los siguientes [1] a [4]:

- [1] una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68;
- [2] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se delecionan, sustituyen o añaden en la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido;
- [3] una proteína que consiste en una secuencia de aminoácidos que tiene el 65% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido; y

[4] una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos mostrada en SEQ ID NO: 33 y que tiene la capacidad de sintetizar un dipéptido;

y ADN de los siguientes [5] a [7]:

[5] ADN que tiene la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65;

[6] ADN que hibrida con ADN que tiene la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65 en condiciones rigurosas y que codifica una proteína que tiene la actividad para sintetizar un dipéptido; y

[7] ADN que tiene una secuencia de nucleótido que tiene el 80% o más de homología con la secuencia de nucleótidos mostrada en SEQ ID NO: 34 y que codifica una proteína que tiene la actividad para sintetizar un dipéptido.

Los ejemplos de los ADN que codifican una proteína que tiene la actividad para sintetizar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido incluyen ADN que codifican proteínas de los siguientes [1] a [3]:

[1] prolina iminopeptidasa descrita en cualquiera de documento WO03/010307; FEMS Microbiol. Lett., 78, 191 (1999); Solicitud de patente japonesa no examinada publicada No. 113887/90; 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); Solicitud de patente japonesa no examinada publicada No. 121860/97; 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);

[2] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se deletorean, sustituyen o añaden en la secuencia de aminoácidos de cualquier prolina iminopeptidasa del anterior [1] y que tiene actividad iminopeptidasa; y

[3] una proteína que consiste en una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos de cualquier prolina iminopeptidasa del anterior [1] y que tiene actividad prolina iminopeptidasa

y ADN de los siguientes [4] y [5]:

[4] ADN que codifica prolina iminopeptidasa y que tiene la secuencia de nucleótidos descrita en cualquiera del documento WO03/010307; FEMS Microbiol. Lett., 78, 191 (1999); Solicitud de patente japonesa no examinada publicada No. 113887/90; 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); Solicitud de patente japonesa no examinada publicada No. 121860/97; 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

[5] ADN que hibrida con cualquier ADN que codifica prolina iminopeptidasa del anterior [4] en condiciones rigurosas y que codifica una proteína que tiene actividad prolina iminopeptidasa.

Los ejemplos de los ADN que codifican una proteína que tiene la actividad para sintetizar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido incluyen ADN que codifican proteínas de los siguientes [1] y [2]:

[1] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos están deletreados, sustituidos o añadido en la secuencia de aminoácidos de la amida de L-aminoácido hidrolasa descrita en el documento WO03/010187 y que tiene actividad amida de L-aminoácido hidrolasa;

[2] una proteína que consiste en una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos de la amida de L-aminoácido hidrolasa descrita en el documento WO03/010187 y que tiene actividad amida de L-aminoácido hidrolasa;

y ADN de los siguientes [3] y [4]:

[3] ADN que tiene la secuencia de nucleótidos descrita en el documento WO03/010187 y que codifica amida de L-aminoácido hidrolasa; y

[4] ADN que hibrida con ADN que consiste en la secuencia de nucleótidos descrita en el documento WO03/010187 y que codifica amida de L-aminoácido hidrolasa en condiciones rigurosas y que codifica una proteína que tiene actividad amida de L-aminoácido hidrolasa.

Los ADN anteriores capaces de hibridación en condiciones rigurosas se refieren a ADN que se obtiene por hibridación en colonia, hibridación en placa, hibridación por transferencia Southern, o similares usando una parte o todo de cualquiera de los ADN anteriores como sonda. Un ejemplo específico de tal ADN es ADN que se puede identificar realizando la hibridación a 65°C en presencia de cloruro de sodio de 0,7 a 1,0 mol/l, preferiblemente cloruro de sodio 0,9 mol/l usando un filtro con ADN derivado de colonia o placa inmovilizado en el mismo, y después lavando el filtro a 65°C con solución SSC de 0,1 a 2 veces concentrada, preferiblemente 0,1 veces concentrada (solución SSC 1 vez concentrada: cloruro de sodio 150 mmol/l y citrato de sodio 15 mmol/l). La hibridación se puede llevar a cabo según los métodos descritos en Molecular Cloning, Segunda Edición; Current Protocols in Molecular Biology; DNA Cloning 1: Core Techniques, A Practical Approach, Segunda Edición, Oxford University (1995), etc.

Específicamente, el ADN hibridable incluye ADN que tiene al menos el 80% o más de homología, preferiblemente el 90% o más de homología, más preferiblemente el 95% o más de homología respecto a la secuencia de nucleótidos de cualquiera de los ADN anteriores calculado mediante el uso de un programa tal como BLAST o FASTA descritos anteriormente basado en los parámetros anteriores.

5 Es posible confirmar que el ADN que hibrida con el ADN anterior en condiciones rigurosas es ADN que codifica una proteína que tiene la actividad para sintetizar un dipéptido de la siguiente manera. Es decir, se prepara un ADN recombinante que expresa el ADN y el ADN recombinante se introduce en una célula huésped para obtener un microorganismo que se va a usar como una fuente de enzima. A continuación, 1) se permite que la fuente de enzima y uno o más tipos de aminoácidos estén presentes en un medio acuoso, seguido por análisis por HPLC o similar para saber si un dipéptido se forma y acumula en el medio acuoso, 2) se permite que la fuente de enzima, un éster de un L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, seguido por análisis por HPLC o similar para saber si un dipéptido se forma y acumula en el medio acuoso, o 3) se permite que la fuente de enzima, una amida de un L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, seguido por análisis por HPLC o similar para saber si un dipéptido se forma y acumula en el medio acuoso.

La homología entre secuencias de nucleótidos se puede determinar usando programas tales como BLAST y FASTA descritos anteriormente.

## 20 2. Métodos para preparar los microorganismos

Los microorganismos descritos en el presente documento se pueden obtener por cualquiera de los siguientes métodos: 1) métodos de impartir la capacidad de producir dipéptidos a microorganismos en los que las funciones de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos se reducen o pierden, o microorganismos en los que las funciones de tres o más tipos de peptidasas se reducen o pierden; y 2) métodos de reducir o producir la pérdida de las funciones de a) uno o más tipos de peptidasas y uno o más tipos de proteínas transportadoras de péptidos o b) tres o más tipos de peptidasas de microorganismos que tiene la capacidad para producir un dipéptido.

30 (1) Métodos de impartir la capacidad de producir dipéptidos a microorganismos en los que las funciones de uno o más tipos de peptidasas y uno o más tipos de proteínas transportadoras de péptidos se reducen o pierden, o microorganismos en los que las funciones de tres o más tipos de peptidasas se reducen o pierden

35 (i) Preparación de microorganismos en los que las actividades de peptidasas y proteínas transportadoras de péptidos se reducen o pierden

Los microorganismos en los que las actividades de peptidasas y proteínas transportadoras de péptidos se reducen o pierden se pueden obtener por cualquier método capaz de preparar tales microorganismos. Por ejemplo, se pueden obtener introduciendo una delección, sustitución o adición de un nucleótido en genes de peptidasas y genes que codifican proteínas transportadoras de péptidos en ADN cromosómicos de microorganismos como se describe a continuación.

Los métodos para introducir una delección, sustitución o adición de un nucleótido en un gen en el ADN cromosómico de un microorganismo incluyen métodos que utilizan recombinación homóloga. Un ejemplo de los métodos que utilizan recombinación homóloga general es un método que usa un plásmido para recombinación homóloga preparado ligando un gen mutante que tiene una delección, sustitución o adición de nucleótidos introducida con un ADN de plásmido incapaz de replicación autónoma en una célula huésped en la que la delección de nucleótidos o similar se va a introducir y que lleva un gen de resistencia a fármaco.

50 El plásmido para la recombinación homóloga se introduce en la célula huésped por un método habitual, seguido por selección de un transformante en el que el plásmido para la recombinación homóloga se ha integrado en el ADN cromosómico por recombinación homóloga usando la resistencia al fármaco como un marcador. El transformante obtenido se cultiva usando un medio que no contiene el fármaco durante varias horas a un día, y después se extiende en un medio con agar que contiene el fármaco y en un medio con agar sin el fármaco. Seleccionando una cepa que no crece en el primer medio, pero puede crecer el último, se puede obtener la cepa en la que la segunda recombinación homóloga se ha producido. La introducción de una delección, sustitución o adición de nucleótidos en un gen deseado en el ADN cromosómico se puede confirmar determinando la secuencia de nucleótidos de una región del ADN cromosómico que contiene el gen en el que se ha introducido la delección o similar.

60 Mediante el uso del método anterior, se puede introducir una delección, sustitución o adición de nucleótidos en genes deseados en ADN cromosómicos de microorganismos tal como los que pertenecen a los géneros Escherichia, Bacillus, Corynebacterium y Saccharomyces.

Además, una delección, sustitución o adición de nucleótidos se puede introducir eficazmente en genes plurales utilizando recombinación homóloga según un método que usa ADN lineal.

Específicamente, un ADN lineal que contiene un gen en que una delección, sustitución o adición de nucleótidos se va a introducir se incorpora a una célula para producir recombinación homóloga entre el ADN cromosómico y el ADN lineal introducido. Este método es aplicable a cualquier microorganismo capaz de incorporar eficazmente un ADN lineal. Los microorganismos preferidos son los que pertenecen a los géneros *Escherichia* y *Bacillus*. *Escherichia coli* es más preferido, y *Escherichia coli* que expresa un grupo de proteínas recombinantes derivadas del fago  $\lambda$  (sistema de recombinación Red) es más preferido.

Un ejemplo de *Escherichia coli* que expresa el sistema de recombinación Red de  $\lambda$  es *Escherichia coli* JM101 que porta pKD46, que es un ADN plásmido que comprende un gen del sistema de recombinación Red de  $\lambda$  (disponible del *Escherichia coli* Genetic Stock Center, Universidad de Yale, EE UU).

Los ejemplos de los ADN útiles para recombinación homóloga son como sigue:

(a) ADN lineal en el que los ADN presentes en el exterior de una región de ADN cromosómico que se va a someter a introducción de una delección, sustitución o adición de nucleótidos o ADN que tienen homología a los ADN están presentes en ambos extremos del gen de resistencia a fármaco;

(b) ADN lineal en el que los ADN presentes en el exterior de una región de ADN cromosómico que se va a someter a introducción de una delección, sustitución o adición de nucleótidos o ADN que tienen homología a los ADN están directamente ligados entre sí;

(c) ADN lineal en el que los ADN presentes en el exterior de una región de ADN cromosómico que se va a someter a introducción de una delección, sustitución o adición de nucleótidos o ADN que tienen homología a los ADN están presentes en ambos extremos del gen de resistencia a fármaco y un gen que se puede usar para selección negativa; y

(d) ADN lineal del anterior (a) en el que una secuencia de nucleótidos reconocida por la recombinasa Fip derivada de levadura [Proc. Natl. Acad. Sci. USA., 82, 5875 (1985)] está presente además entre el gen de resistencia al fármaco y los ADN presentes en el exterior de una región de ADN cromosómico o ADN que tienen homología a los ADN.

Como el gen de resistencia a fármaco, cualquier gen de resistencia a un fármaco al que el microorganismo huésped muestra sensibilidad se puede usar. Cuando se usa *Escherichia coli* como el microorganismo huésped, los ejemplos de los genes de resistencia a fármacos son gen de resistencia a kanamicina, gen de resistencia cloranfenicol, gen de resistencia a gentamicina, gen de resistencia a espectinomicina, gen de resistencia a tetraciclina y gen de resistencia a ampicilina.

El "gen que se puede usar para selección negativa" se refiere a un gen que es letal para un microorganismo huésped en ciertas condiciones de cultivo cuando el gen se expresa en el microorganismo huésped. Los ejemplos de los genes son el gen *sacB* derivado de un microorganismo que pertenece al género *Bacillus* [Appl. Environ. Microbiol., 59, 1361-1366 (1993)] y el gen *rpsL* derivado de un microorganismo que pertenece al género *Escherichia* [Genomics, 72, 99-104 (2001)].

Los ADN presentes en el exterior de una región de ADN cromosómico que se va a someter a la introducción de una sustitución o delección o los ADN que tienen homología respecto a los ADN en los ADN lineales anteriores se localizan en la misma dirección que ese en el ADN cromosómico, y su longitud es preferiblemente aproximadamente de 10 pb a 100 pb, más preferiblemente aproximadamente de 20 pb a 50 pb, y más preferiblemente aproximadamente de 30 pb a 40 pb.

La secuencia de nucleótidos reconocida por la recombinasa Fip derivada de levadura no está específicamente limitada siempre que sea una secuencia de nucleótidos reconocida por la dicha proteína y que catalice recombinación homóloga. Los ejemplos preferidos son ADN que tienen la secuencia de nucleótidos mostrada en SEQ ID NO: 39, y ADN que tiene una secuencia de nucleótidos en donde de uno o varios nucleótidos se delecionan, sustituyen o añaden en dicho ADN y que tiene una secuencia de nucleótido reconocida por la recombinasa Fip derivada de levadura y que cataliza recombinación homóloga.

El "ADN que tiene homología" se refiere a ADN que tiene tal grado de identidad que permite la existencia de recombinación homóloga entre la región objeto del ADN cromosómico y el anterior ADN lineal, específicamente, ADN que tiene el 80% o más de homología, preferiblemente el 90% o más de homología, más preferiblemente el 95% o más de homología, más preferiblemente el 100% de homología. En la medida en que los términos "homología" e "identidad" se usan como sinónimos en el contexto de la presente invención.

La homología entre secuencias de nucleótidos se puede determinar usando programas tal como BLAST y FASTA descritos anteriormente.

Los fragmentos de ADN lineales anteriores se pueden preparar por PCR. El ADN lineal deseado también se puede obtener construyendo ADN que contiene el ADN lineal anterior en plásmido y después llevando a cabo tratamiento con enzimas de restricción.

Los ejemplos de los métodos para introducir una delección, sustitución o adición de nucleótidos en el ADN cromosómico de un microorganismo incluyen los siguientes métodos 1 a 4.

Método 1:

5 Un método que comprende introducir el ADN lineal de los anteriores (a) o (d) en un microorganismo huésped y seleccionar un transformante que porta el ADN lineal insertado en su ADN cromosómico por recombinación homóloga usando la resistencia a fármaco como marcador.

10 Método 2:

15 Un método que comprende introducir el ADN, en el que los ADN presentes en el exterior de una región del ADN cromosómico que se va a someter a introducción de una delección, sustitución o adición de nucleótidos o ADN que tienen homología a los ADN están directamente ligados entre sí, en el transformante obtenido según el anterior método 1 y eliminar el gen de resistencia a fármaco insertado en su ADN cromosómico según el método para sustituir o deleccionar una región del ADN cromosómico del microorganismo.

Método 3:

20 Un método que comprende:

[1] introducir el ADN lineal del anterior (c) en un microorganismo huésped y seleccionar un transformante que porta el ADN lineal insertado en su ADN cromosómico por recombinación homóloga usando la resistencia al fármaco como marcador;

25 [2] sintetizar ADN ligando ADN que tienen homología a los ADN presentes en el exterior de una región de ADN cromosómico que se va a someter a la introducción de una sustitución o delección en la misma dirección que esa en el ADN cromosómico, e introducir el ADN sintetizado en el transformante obtenido en el anterior [1]; y

30 [3] cultivar el transformante sometido a la operación del anterior [2] en condiciones tales que el gen que se puede usar para selección negativa se exprese, y seleccionar una cepa capaz de crecer por el cultivo como una cepa en la que el gen de resistencia al fármaco y el gen que se puede usar para selección negativa se eliminan del ADN cromosómico.

35 Método 4:

Un método que comprende:

40 [1] introducir el ADN lineal del anterior (d) en un microorganismo huésped y seleccionar un transformante que porta el ADN lineal insertado en su ADN cromosómico por recombinación homóloga usando la resistencia al fármaco como marcador; y

45 [2] introducir un plásmido de expresión del gen de la recombinasa Flp en el transformante obtenido en el anterior [1], y después de la expresión del gen, obtener una cepa sensible al fármaco usado en el anterior [1].

50 En los métodos anteriores, la introducción del ADN lineal en un microorganismo huésped se puede llevar a cabo por cualquiera de los métodos para introducir ADN en el microorganismo, por ejemplo, el método que usa ion calcio [Proc. Natl. Acad. Sci. USA, 69, 2110 (1972)], el método de protoplastos (Solicitud de patente japonesa no examinada publicada No. 248394/88) y electroporación [Nucleic Acids Res., 16, 6127 (1988)].

Usando un ADN en el que un gen arbitrario que se va a insertar al ADN cromosómico se incorpora en la parte central del ADN usado en el método 2 o método 3 [2], es posible eliminar el gen de resistencia al fármaco y al mismo tiempo insertar un gen arbitrario en el ADN cromosómico.

55 Los métodos 2 a 4 anteriores son métodos que no dejan genes exógenos tal como un gen de resistencia a fármaco y un gen utilizable para selección negativa en el ADN cromosómico del transformante que se va a obtener, por último. Por tanto, es posible producir fácilmente un microorganismo que tiene delecciones, sustituciones o adiciones de nucleótidos en dos o más regiones diferentes del ADN cromosómico repitiendo las operaciones del método 2, método 3 [1] a [3] y método 4 [1] y [2] usando el mismo gen de resistencia a fármaco y el mismo gen utilizable para selección negativa.

60 (ii) Métodos de impartir la capacidad de producir dipéptidos

Los métodos de impartir la capacidad de producir dipéptidos al microorganismo huésped usado en el anterior (i) incluyen los siguientes métodos.

65

(a) Preparación de ADN que codifica una proteína que tiene actividad para sintetizar un dipéptido.

El ADN anterior que codifica una proteína que tiene la actividad para sintetizar un dipéptido se puede obtener por los métodos descritos a continuación que utilizan información de la secuencia de nucleótidos en el ADN.

Por ejemplo, el ADN que codifica bacilisina sintetasa se puede obtener por hibridación Southern de un genoteca de ADN cromosómico de un microorganismo que pertenece al género Bacillus usando una sonda diseñada basada en la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65, o por PCR [PCR Protocols, Academic Press (1990)] usando ADN cebadores diseñados basados en la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65 y, como molde, el ADN cromosómico de un microorganismo que pertenece al género Bacillus.

El ADN que codifica una proteína que tiene la actividad para sintetizar un dipéptido también se puede obtener realizando una búsqueda en varias bases de datos de secuencias de genes para una secuencia que tiene el 65% o más de homología, preferiblemente el 80% o más de homología, más preferiblemente el 90% o más de homología, más preferiblemente el 95% o más de homología respecto a la secuencia de nucleótidos del ADN que codifica la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25, 33 y 68, y obteniendo el ADN deseado, basado en la secuencia de nucleótidos obtenida por la búsqueda, de una genoteca de ADN cromosómico o ADNc de un organismo que tiene la secuencia de nucleótidos según el método descrito anteriormente.

El ADN obtenido, como tal o después de corte con enzimas de restricción apropiadas, se inserta en un vector por un método convencional para obtener un ADN recombinante. Un ADN de plásmido se extrae de un transformante obtenido introduciendo el ADN recombinante en Escherichia coli. A continuación, la secuencia de nucleótidos del ADN se puede determinar por un método de secuenciación convencional tal como el método del dideoxi [Proc. Natl. Acad. Sci., USA, 74, 5463 (1977)] o usando un secuenciador de nucleótidos tal como 373A DNA Sequencer (Perkin-Elmer Corp.).

En casos donde el ADN obtenido se encuentra que es un ADN parcial mediante el análisis de la secuencia de nucleótidos, el ADN de longitud completa se puede obtener por hibridación Southern de una genoteca de ADN cromosómico usando el ADN parcial como sonda.

También es posible preparar el ADN deseado por síntesis química usando un sintetizador de ADN (por ejemplo, modelo 8905, PerSeptive Biosystems) basado en la secuencia de nucleótidos determinada del ADN.

Los ejemplos de los ADN que se pueden obtener por el método descrito anteriormente son ADN que tienen las secuencias de nucleótidos mostradas en SEQ ID NO: 26 a 32, 64 y 65.

Los ejemplos de los vectores para insertar el ADN incluyen pBluescriptII KS(+) (Stratagene), pDIRECT [Nucleic Acids Res., 18, 6069 (1990)], pCR-Script Amp SK(+) (Stratagene), pT7 Blue (Novagen, Inc.), pCR II (Invitrogen Corp.) y pCR-TRAP (Genhunter Corp.).

Los ejemplos de Escherichia coli incluyen Escherichia coli XL1-Blue, Escherichia coli XL2-Blue, Escherichia coli DH1, Escherichia coli MC1000, Escherichia coli KY3276, Escherichia coli W1485, Escherichia coli JM101, Escherichia coli JM109, Escherichia coli HB101, Escherichia coli No. 49, Escherichia coli W3110, Escherichia coli NY49, Escherichia coli MP347, Escherichia coli NM522 y Escherichia coli ME8415.

La introducción del ADN recombinante se puede llevar a cabo por cualquiera de los métodos para introducir ADN en las células huésped anteriores, por ejemplo, el método que usa ion calcio [Proc. Natl. Acad. Sci. USA, 69, 2110 (1972)], el método de protoplastos (Solicitud de patente japonesa no examinada publicada No. 248394/88) y electroporación [Nucleic Acids Res., 16, 6127 (1988)].

Un ejemplo del microorganismo que porta el ADN que codifica una proteína que tiene la actividad sintetizadora de dipéptidos obtenido por el método anterior es Escherichia coli NM522/pPE43, que es un microorganismo que porta un ADN recombinante que comprende ADN que tiene la secuencia mostrada en SEQ ID NO: 19 descrito posteriormente.

(b) Producción de una proteína que tiene la actividad sintetizadora de dipéptidos

La proteína que tiene la actividad sintetizadora de dipéptidos se puede producir expresando el ADN obtenido por los métodos en el anterior (a) en células huésped usando los métodos descritos en Molecular Cloning, Segunda Edición, Current Protocols in Molecular Biology, etc., por ejemplo, de la siguiente manera.

En base al ADN obtenido por los métodos descritos en el anterior (a), se prepara un fragmento de ADN de una longitud apropiada que comprende una región que codifica una proteína que tiene la actividad sintetizadora de dipéptidos según la necesidad. La productividad de la proteína se puede aumentar sustituyendo un nucleótido en la

secuencia de nucleótidos de la región que codifica la proteína de modo que se haga el codón más adecuado para la expresión en una célula huésped.

5 El fragmento de ADN se inserta después de un promotor en un vector de expresión adecuado para preparar un ADN recombinante.

Se puede obtener un transformante que produce la proteína que tiene la actividad sintetizadora de dipéptidos introduciendo el ADN recombinante en una célula huésped adecuada para el vector de expresión.

10 Como la célula huésped, se puede usar cualquier microorganismo tal como células bacterianas y células de levadura que sean capaces de expresar el gen deseado.

15 Los vectores de expresión que se pueden emplear son los capaces de replicación autónoma o integración en el cromosoma en las células huéspedes anteriores y que comprenden un promotor en una posición apropiada para la transcripción del ADN que codifica la proteína que tiene la actividad sintetizadora de dipéptidos.

20 Cuando se usa un procarionta tal como una bacteria como la célula huésped, se prefiere que el ADN recombinante que comprende el ADN que codifica una proteína que tenga la actividad sintetizadora de dipéptidos sea un ADN recombinante que es capaz de replicación autónoma en el procarionta y que comprende un promotor, una secuencia de unión a ribosomas, el ADN que codifica una proteína que tiene actividad sintetizadora de dipéptidos, y una secuencia de terminación de la transcripción. El ADN recombinante puede comprender además un gen que regula el promotor.

25 Los ejemplos de vectores de expresión adecuados son pBTrp2, pBTac1 y pBTac2 (productos de Boehringer Mannheim GmbH), pHelix1 (Roche Diagnostics Corp.), pKK233-2 (Amersham Pharmacia Biotech), pSE280 (Invitrogen Corp.), pGEMEX-1 (Promega Corp.), pQE-8 (Qiagen, Inc.), pET-3 (Novagen, Inc.), pKYP10 (Solicitud de patente japonesa no examinada publicada No. 110600/83), pKYP200 [Agric. Biol. Chem., 48, 669 (1984)], pLSA1 [Agric. Biol. Chem., 53, 277 (1989)], pGEL1 [Proc. Natl. Acad. Sci. USA, 82, 4306 (1985)], pBluescript II SK(+), pBluescript II KS(-) (Stratagene), pTrS30 [preparado de *Escherichia coli* JM109/pTrS30 (FERM BP-5407)], pTrS32 [preparado de *Escherichia coli* JM109/pTrS32 (FERM BP-5408)], pPAC31 (documento WO98/12343), pUC19 [Gene, 33, 103 (1985)], pSTV28 (Takara Shuzo Co., Ltd.), pUC118 (Takara Shuzo Co., Ltd.) y pPA1 (Solicitud de patente japonesa no examinada publicada No. 233798/88).

35 Como el promotor, cualquier promotor capaz de funcionar en células huésped tal como *Escherichia coli* se puede usar. Por ejemplo, promotores derivados de *Escherichia coli* o fagos, tal como el promotor *trp* ( $P_{trp}$ ), promotor *lac* ( $P_{lac}$ ), promotor  $P_L$ , promotor  $P_R$  y promotor  $P_{SE}$ , promotor SPO1, promotor SPO2 y promotor penP se pueden usar. Promotores artificialmente diseñados y modificados tal como un promotor en que dos  $P_{trp}$  se combinan en tándem, promotor *lac*, promotor *lacT7* y promotor *letI*, etc. también se pueden usar.

40 También son útiles promotores tal como el promotor *xylA* para la expresión en bacterias que pertenecen al género *Bacillus* [Appl. Microbiol. Biotechnol., 35, 594-599 (1991)] y el promotor P54-6 para la expresión en bacterias que pertenecen al género *Corynebacterium* [Appl. Microbiol. Biotechnol., 53, 674-679 (2000)].

45 Se prefiere usar un plásmido en el que la distancia entre la secuencia Shine-Dalgarno (secuencia de unión al ribosoma) y el codón de iniciación se ajuste a una longitud apropiada (por ejemplo, de 6 a 18 nucleótidos).

50 En el ADN recombinante en donde el ADN que codifica la proteína que tiene la actividad sintetizadora de dipéptidos está ligado a un vector de expresión, la secuencia de terminación de la transcripción no es esencial, pero se prefiere colocar la secuencia de terminación de la transcripción después del gen estructural.

Un ejemplo de tal ADN recombinante es pPE43 descrito posteriormente.

55 Los ejemplos de procariontas adecuados para uso como células huésped incluyen microorganismos que pertenecen a los géneros *Escherichia*, *Bacillus* y *Corynebacterium*. Los ejemplos específicos son *Escherichia coli* XL1-Blue, *Escherichia coli* XL2-Blue, *Escherichia coli* DH1, *Escherichia coli* DH5 $\alpha$ , *Escherichia coli* MC1000, *Escherichia coli* KY3276, *Escherichia coli* W1485, *Escherichia coli* JM101, *Escherichia coli* JM109, *Escherichia coli* HB101, *Escherichia coli* No. 49, *Escherichia coli* W3110, *Escherichia coli* NY49, *Escherichia coli* MP347, *Escherichia coli* NM522, *Bacillus subtilis* ATCC 33712, *Bacillus megaterium*, *Bacillus* sp. FERM BP-6030, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus pumilus*, *Corynebacterium glutamicum* ATCC 13032 y *Corynebacterium glutamicum* ATCC 14297.

60 La introducción del ADN recombinante se puede llevar a cabo por cualquiera de los métodos para introducir ADN en las células huéspedes anteriores, por ejemplo, el método que usa ion calcio [Proc. Natl. Acad. Sci. USA, 69, 2110 (1972)], el método del protoplasto (Solicitud de patente japonesa no examinada publicada No. 248394/88) y electroporación [Nucleic Acids Res., 16, 6127 (1988)].

Cuando se usa una cepa que pertenece al género Saccharomyces como la célula huésped, YEp13 (ATCC 37115), YEp24 (ATCC 37051), YCp50 (ATCC 37419), pHSI9, pHS15, etc. se puede usar como el vector de expresión.

5 Como el promotor, se puede usar cualquier promotor capaz de funcionar en las cepas que pertenecen al género Saccharomyces. Los promotores adecuados incluyen, promotor PHO5, promotor PGK, promotor GAP, promotor ADH, promotor gal 1, promotor gal 10, promotor del polipéptido de choque térmico, promotor MFa1 y promotor CUP 1.

10 Los ejemplos de células huéspedes adecuadas son cepas que pertenecen al género Saccharomyces, específicamente, Saccharomyces cerevisiae.

15 La introducción del ADN recombinante se puede llevar a cabo por cualquiera de los métodos para introducir ADN en levaduras, por ejemplo, electroporación [Methods Enzymol., 194, 182 (1990)], el método de esferoplasto [Proc. Natl. Acad. Sci. USA, 81, 4889 (1984)] y el método del acetato de litio [J. Bacteriol., 153, 163 (1983)].

(2) Métodos de reducir o producir la pérdida de las funciones de a) uno o más tipos de peptidasas y uno o más tipos de proteínas transportadoras de péptidos o b) tres o más tipos de peptidasas de microorganismos que tienen la capacidad de producir un dipéptido

20 Los microorganismos que tienen la capacidad de producir un dipéptido se pueden preparar llevando a cabo los métodos descritos en el anterior (1) (ii) usando microorganismos arbitrarios como células huésped. Al llevar a cabo los métodos descritos en el anterior (1) (i) usando los microorganismos así preparados, se pueden preparar microorganismos en los que las funciones de uno o más tipos de peptidasas y uno o más tipos de proteínas transportadoras de péptidos o las funciones de tres o más tipos de peptidasas se reducen o pierden y que tienen la capacidad de producir un dipéptido.

25 Los microorganismos descritos en el presente documento también se pueden obtener llevando a cabo los métodos descritos en el anterior (1) (i) usando microorganismos que tienen inherentemente la capacidad de producir un dipéptido.

30 Los ejemplos de los microorganismos anteriores incluyen los que pertenecen a los géneros Escherichia, Bacillus, Corynebacterium y Saccharomyces. Se prefieren Escherichia coli, Bacillus subtilis, Corynebacterium glutamicum y Saccharomyces cerevisiae.

### 35 3. Proceso para producir un dipéptido

Los procesos de producción descritos en el presente documento incluyen:

40 (i) un proceso para producir un dipéptido, que comprende: dejar que una fuente de enzima y uno o más tipo de aminoácidos estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo del microorganismo de la presente invención o un material tratado del cultivo; dejar que el dipéptido se forme y acumule en el medio acuoso; y recuperar el dipéptido del medio;

45 (ii) un proceso para producir un dipéptido, que comprende: dejar que una fuente de enzima, un éster de L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo del microorganismo de la presente invención o un material tratado del cultivo; dejar que el dipéptido se forme y acumule en el medio acuoso; y recuperar el dipéptido del medio; y

50 (iii) un proceso para producir un dipéptido, que comprende: dejar que una fuente de enzima, una amida de L-aminoácido y un L-aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo del microorganismo de la presente invención o un material tratado del cultivo; dejar que el dipéptido se forme y acumule en el medio acuoso; y recuperar el dipéptido del medio.

55 Uno o más tipos, preferiblemente uno o dos tipos de aminoácidos usados como como sustratos en el anterior proceso de producción (i) son aminoácidos, preferiblemente aminoácidos seleccionados del grupo que consiste en L-aminoácidos, glicina (Gly) y β-alanina (β-Ala), y se pueden usar en cualquier combinación. Los ejemplos de L-aminoácidos son 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-asparragina (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.

65 Los aminoácidos que se usan más preferiblemente en el anterior proceso (i) incluyen los siguientes: una combinación de un tipo de aminoácidos seleccionado del grupo que consiste en L-Ala, Gly, L-Met, L-Ser, L-Thr 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- $\alpha$ -AB,  $\beta$ -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; y una combinación de L- $\alpha$ -AB y L-Gln, L-Arg o L- $\alpha$ -AB. Aminoácidos más 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- $\alpha$ -AB, L-azaserina, L-Cit y L-teanina; una combinación de Gly y 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- $\alpha$ -AB y L-Cit; una combinación de L-Met y un tipo de aminoácido seleccionado del grupo que consiste 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- $\alpha$ -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- $\alpha$ -AB; una combinación de L-Gln y L-Phe; una combinación de  $\beta$ -Ala y un tipo de aminoácido seleccionado del grupo que consiste en L-Phe, L-Met, L-His y L-Cit; y una combinación de L- $\alpha$ -AB y L-Gln, L-Arg o L- $\alpha$ -AB.

En el anterior proceso de producción (i), el L-aminoácido usado como un sustrato se añade al medio acuoso al inicio o en el curso de la reacción para dar una concentración de 0,1 a 500 g/l, preferiblemente de 0,2 a 200 g/l.

Los dipéptidos producidos por el anterior proceso (i) incluyen los dipéptidos representados por la siguiente fórmula (I):



(en donde  $R^1$  y  $R^2$ , que pueden ser iguales o diferentes, representan cada uno un aminoácido). Los dipéptidos preferidos son los representados por la anterior fórmula (I) en donde  $R^1$  y  $R^2$ , que pueden ser iguales o diferentes, representan cada uno un 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- $\alpha$ -AB,  $\beta$ -Ala, L-azaserina, L-teanina, L-4-HYP, L-3-HYP, L-Orn y L-6-diazo-5-oxo-norleucina. Más preferidos son dipéptidos donde  $R^1$  es L-Ala, Gly, L-Met, L-Ser, L-Thr o  $\beta$ -Ala, y  $R^2$  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-Asp, L- $\alpha$ -AB,  $\beta$ -Ala, L-azaserina; L-teanina, L-4-HYP, L-3-HYP, L-Orn o L-6-diazo-5-oxo-norleucina. Dipéptidos más preferidos son: dipéptidos en donde  $R^1$  es L-Ala y  $R^2$  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- $\alpha$ -AB, L-azaserina o L-teanina; dipéptidos en donde  $R^1$  es Gly y  $R^2$  es L-Gln, Gly, L-Trp, L-Met, L-Ser, L-Thr, L-Cys, L-Tyr, L-Lys, L-Arg o L- $\alpha$ -AB; dipéptidos en donde  $R^1$  es L-Met y  $R^2$  es L-Phe, L-Met, L-Cys, L-Tyr, L-Lys o L-His; dipéptidos en donde  $R^1$  es L-Ser y  $R^2$  es L-Gln, Gly, L-Phe, L-Met, L-Ser, L-Thr, L-Tyr, L-His o L- $\alpha$ -AB; dipéptidos en donde  $R^1$  es L-Thr y  $R^2$  es L-Gln, L-Gly, L-Phe, L-Met, L-Ser, L-Thr o L- $\alpha$ -AB; dipéptidos en donde  $R^1$  es L-Gln y  $R^2$  es L-Phe o L- $\alpha$ -AB; un dipéptido en donde  $R^1$  es L-Phe y  $R^2$  es L-Gln; un dipéptido en donde  $R^1$  es L-Trp y  $R^2$  es Gly; dipéptidos en donde  $R^1$  es L-Cys y  $R^2$  es L-Ala, L-Gln, Gly o L-Met; dipéptidos en donde  $R^1$  es L-Lys y  $R^2$  es L-Ala, Gly o L-Met; un dipéptido en donde  $R^1$  es L-Arg y  $R^2$  es L- $\alpha$ -AB; un dipéptido en donde  $R^1$  es L-His y  $R^2$  es L-Met; y dipéptidos en donde  $R^1$  es L- $\alpha$ -AB y  $R^2$  es L-Ala, L-Gln, Gly, L-Ser, L-Thr, L-Arg o L- $\alpha$ -AB.

Además, en el proceso anterior, se pueden añadir compuestos que pueden ser metabolizados por el microorganismo de la presente invención para producir ATP, por ejemplo, azúcares tal como glucosa, alcoholes tal como etanol, y ácidos orgánicos tal como ácido acético, como fuente de ATP, al medio acuoso según la necesidad.

El éster de L-aminoácido y L-aminoácido usados como sustratos en el anterior proceso de producción (ii) pueden ser cualquier éster de L-aminoácido y L-aminoácido que se pueden usar como sustratos por el microorganismo de la presente invención para formar un dipéptido, y se pueden usar en cualquier combinación. Preferiblemente, el éster de L-aminoácido se selecciona 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-treonina, éster de L-glutamina, éster de L-tirosina, éster de L-arginina,  $\alpha$ -éster de ácido L-aspártico,  $\beta$ -éster de ácido L-aspártico, éster de L-leucina, éster de L-asparragina, éster de L-lisina, éster  $\alpha,\beta$ -dimetilico de ácido L-aspártico y  $\gamma$ -éster de L-glutamina, y el L-aminoácido se selecciona 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 anterior proceso (ii), el éster de L-aminoácido y L-aminoácido usados como sustratos se añaden al medio acuoso al inicio o en el curso de la reacción para dar una concentración de 0,1 a 500 g/l, preferiblemente de 0,2 a 200 g/l.

La amida de L-aminoácido y L-aminoácido usados como sustratos en el anterior proceso de producción (iii) pueden ser cualquier amida de L-aminoácido y L-aminoácido que puedan ser usados como sustratos por el microorganismo de la presente invención para formar un dipéptido, y se pueden usar en cualquier combinación. Preferiblemente, la amida de L-aminoácido se selecciona del grupo que consiste en amida de L-alanina, amida de glicina y amida de ácido L-aspártico, y el L-aminoácido se selecciona 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.

En el anterior proceso (iii), la amida de L-aminoácido y L-aminoácido usados como sustratos se añaden al medio acuoso al inicio o en el curso de la reacción para dar una concentración de 0,1 a 500 g/l, preferiblemente de 0,2 a 200 g/l.

5 El medio acuoso usado en los procesos de producción puede comprender cualquier componente y puede tener cualquier composición siempre que la reacción de formación del dipéptido no esté inhibida. Los medios acuosos adecuados incluyen agua y tampones tal como tampón fosfato, tampón carbonato, tampón acetato, tampón borato, tampón citrato y tampón Tris. El medio acuoso puede comprender alcoholes tal como metanol y etanol, ésteres tal como acetato de etilo, cetonas tal como acetona, y amidas tal como acetamida.

10 La reacción de formación de dipéptido se lleva a cabo en el medio acuoso a pH de 5 a 11, preferiblemente pH de 6 a 10, de 20 a 60°C, preferiblemente de 25 a 45°C, durante de 2 a 150 horas, preferiblemente de 6 a 120 horas.

15 Si es necesario, se puede añadir además un tensioactivo o un solvente orgánico al medio acuoso.

Se puede usar cualquier tensioactivo que fomente la formación de un dipéptido. Los tensioactivos adecuados incluyen tensioactivos no iónicos tal como polioxietileno octadecilamina (por ejemplo, Nymeen S-215, NOF Corporation), tensioactivos catiónicos tal como bromuro de cetiltrimetilamonio y cloruro de alquildimetilbencilamonio (por ejemplo, Cation F2-40E, NOF Corporation), tensioactivos aniónicos tal como sarcosinato de lauroilo, y aminas terciarias tal como alquildimetilamina (por ejemplo, Tertiary Amine FB, NOF Corporation), que se pueden usar solos o en combinación. El tensioactivo habitualmente se usa a una concentración de 0,1 a 50 g/l. Como el solvente orgánico, se pueden usar xileno, tolueno, alcoholes alifáticos, acetona, acetato de etilo, etc., habitualmente a una concentración de 0,1 a 50 ml/l.

25 Los materiales tratados del cultivo incluyen cultivo concentrado, cultivo seco, células obtenidas por centrifugación del cultivo, y productos obtenidos tratando las células por varios medios tal como secado, liofilización, tratamiento con un tensioactivo, ultrasonificación, ficción mecánica, tratamiento con un solvente, tratamiento enzimático e inmovilización. Los materiales tratados del cultivo descritos en el presente documento también incluyen extractos crudos de proteína obtenidos eliminando materiales insolubles y similares de los materiales tratados obtenidos tratando las células anteriores por medios tales como tratamiento con un tensioactivo, ultrasonificación, ficción mecánica, tratamiento con un solvente, tratamiento enzimático.

35 Cuando el cultivo o un material tratado del cultivo se usa como la fuente de enzima, la cantidad de la fuente de enzima que se va a añadir varía según su actividad específica, etc., pero es, por ejemplo, de 5 a 1000 mg, preferiblemente de 10 a 400 mg por mg de aminoácido, éster de L-aminoácido o amida de L-aminoácido usados como sustrato.

40 La recuperación del dipéptido formado y acumulado en el medio acuoso se puede llevar a cabo por métodos habituales usando carbono activo, resinas de intercambio iónico, etc., o por medios tales como extracción con un solvente orgánico, cristalización, cromatografía de capa fina y cromatografía líquida de alto rendimiento.

Además, los anteriores procesos de producción (ii) y (iii) se pueden llevar a cabo según las descripciones en los documentos WO03/010189 o WO03/010187.

45 Ejemplos experimentales de la presente invención se muestran a continuación.

#### Ejemplo experimental 1

50 Construcción de un plásmido que expresa el gen ywfE derivado de Bacillus subtilis

Se obtuvo un fragmento del gen ywfE de Bacillus subtilis de la siguiente manera.

Usando un sintetizador de ADN (modelo 8905, PerSeptive Biosystems, Inc.), se sintetizaron ADN que tenían las secuencias de nucleótidos mostradas en SEQ ID NO: 35 y 36 (de aquí en adelante denominadas cebador A y cebador B, respectivamente). El cebador A tiene una secuencia de nucleótidos que contiene una región en donde el codón de iniciación del gen ywfE (atg) se sustituye por la secuencia de reconocimiento de NcoI (ccatgg). El cebador B tiene una secuencia de nucleótidos que contiene una región en donde el codón de terminación del gen ywfE se sustituye por la secuencia de reconocimiento de BamHI (ggatcc).

60 Se llevó a cabo PCR usando el ADN cromosómico de Bacillus subtilis como molde y los anteriores cebador A y cebador B como un conjunto de cebadores. Es decir, se llevó a cabo PCR mediante 30 ciclos, un ciclo consiste en reacción a 94°C durante un minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos, usando 40 µl de una mezcla de reacción que comprendía 0,1 µg del ADN cromosómico, 0,5 µmol/l de cada uno de los cebadores, 2,5 unidades de ADN polimerasa Pfu, 4 µl de tampón para ADN polimerasa Pfu (10x) y 200 µmol/l de cada uno de los dNTP.

Un décimo de la mezcla de reacción resultante se sometió a electroforesis en gel de agarosa para confirmar que un fragmento de aprox. 1,4 kb correspondiente al gen ywfE se amplificó. A continuación, la mezcla de reacción restante se mezcló con una cantidad igual de fenol/cloroformo saturado con TE. La mezcla resultante se centrifugó, y la fase superior obtenida se mezcló con un volumen de dos veces de etanol frío y se dejó reposar a -80°C durante 30 minutos. La solución resultante se centrifugó, y el precipitado de ADN obtenido se disolvió en 20 µl de TE.

La solución así obtenida (5 µl) se sometió a reacción para cortar el ADN amplificado con las enzimas de restricción NcoI y BamHI. Se separaron los fragmentos de ADN por electroforesis en gel de agarosa, y un fragmento de ADN de 1,4 kb que contenía el gen ywfE se recuperó usando el kit GENE CLEAN II (Bio 101).

Se cortó el vector de expresión recombinante con etiqueta His C-terminal pQE60 (Qiagen, Inc.) (0,2 µg) con las enzimas de restricción NcoI y BamHI. Los fragmentos de ADN se separaron por electroforesis en gel de agarosa, y un fragmento de ADN de 3,4 kb se recuperó de la misma manera que anteriormente.

El fragmento de ADN de 1,4 kb que contenía el gen ywfE y el fragmento de ADN de 3,4 kb obtenidos anteriormente se sometieron a reacción de ligación usando un kit de ligación (Takara Shuzo Co., Ltd.) a 16°C durante 16 horas.

Se transformó Escherichia coli NM52 (Stratagene) usando la mezcla de reacción de ligación según el método que usa ion calcio [Proc. Natl. Acad. Sci. USA, 69, 2110 (1972)], se extendió en medio LB agar que contenía ampicilina 50 µg/ml, y se cultivó durante la noche a 30°C.

Se extrajo un plásmido de una colonia del transformante que creció en el medio según un método conocido, por lo cual se obtuvo pQE60ywfE, que es un vector de expresión del gen ywfE con etiqueta de His C-terminal. La estructura del vector se confirmó por digestión con enzimas de restricción (Fig. 1).

#### Ejemplo experimental 2

##### Adquisición de un producto del gen ywfE

Se inoculó Escherichia coli NM52/pQE60ywfE que lleva pQE60ywfE en 8 ml de medio LB que contenía ampicilina 50 µg/ml en un tubo de ensayo, y se cultivó a 28°C durante 17 horas. El cultivo resultante se inoculó en 50 ml de medio LB que contenía ampicilina 50 µg/ml en matraz Erlenmeyer de 250 ml, y se cultivó a 30°C durante 3 horas. A continuación, se añadió isopropil-β-D-tiogalactopiranosido (IPTG) para dar una concentración final de 1 mmol/l, seguido por cultivo adicional a 30°C durante 4 horas. El cultivo resultante se centrifugó para obtener células húmedas, y una enzima recombinante con etiqueta His se purificó de las células húmedas usando HisTrap (kit de purificación de proteínas con etiqueta His, Amersham Pharmacia Biotech) según las instrucciones adjuntas al mismo.

#### Ejemplo experimental 3

##### Producción de dipéptidos usando la enzima recombinante con etiqueta His (1)

(i) Se preparó una mezcla de reacción (0,1 ml) que comprendía 0,04 mg de la enzima recombinante con etiqueta de His purificada obtenida en el ejemplo experimental 2, Tris-HCl 100 mmol/l (pH 8,0), cloruro de magnesio 60 mmol/l, ATP 60 mmol/l, L-Ala 30 mmol/l y L-Gln 30 mmol/l, y la reacción se llevó a cabo a 37°C durante 16 horas.

Después de completar la reacción, el producto de reacción se derivó por el método del dinitrofenol y después se analizó por HPLC. El análisis por HPLC se llevó a cabo usando, como columna de separación, columna Lichrosorb-RP-18 (Kanto Kagaku) y, como solución eluyente, ácido fosfórico al 1% (v/v) y acetonitrilo al 25% (v/v) a una velocidad de flujo de 0,7 ml/min. Como resultado, se confirmó que L-Ala-L-Gln 3,7 g/l y L-alanil-L-alanina (L-Ala-L-Ala) 0,3 g/l se formaron y acumularon en la mezcla de reacción.

(ii) Se llevaron a cabo reacciones en las mismas condiciones que el anterior (i) usando mezclas de reacción que tenían la misma composición que la de la mezcla de reacción del anterior (i) excepto que se usaron 0,01 mg de la enzima y L-Phe, L-Met, L-Leu y L-Val, respectivamente, se usaron en lugar de L-Gln.

Después de completar las reacciones, los productos de reacción se analizaron de la misma manera que en anterior (i), por lo que se confirmó que los siguientes dipéptidos se formaron y acumularon en las respectivas mezclas de reacción: L-alanil-L-fenilalanina (L-Ala-L-Phe) solo 7,0 g/l; L-alanil-L-metionina (L-Ala-L-Met) 7,0 g/l y L-Ala-L-Ala 0,03 g/l; L-alanil-L-leucina (L-Ala-L-Leu) 5,0 g/l y L-Ala-L-Ala 0,2 g/l; y L-alanil-L-valina (L-Ala-L-Val) 1,6 g/l y L-Ala-L-Ala 0,3 g/l.

(iii) Se llevaron a cabo reacciones en las mismas condiciones que el anterior (i) usando mezclas de reacciones que tenían la misma composición que la de la mezcla de reacción del anterior (i) excepto que se usaron 0,01 mg de la enzima, Gly se usó en lugar de L-Ala y L-Phe y L-Met, respectivamente se usaron en lugar de L-Glu.

Después de completar las reacciones, los productos de reacción se analizaron de la misma manera que en anterior (i), por lo que se confirmó que glicil-L-fenilalanina (Gly-L-Phe) 5,2 g/l y glicil-L-metionina (Gly-L-Met) 1,1 g/l se formaron y acumularon en las respectivas mezclas de reacción.

5 Cuando se excluyó ATP de las composiciones de las mezclas de reacción anteriores, no se formaron dipéptidos.

Los resultados anteriores revelaron que el producto del gen *ywfE* tiene la actividad para producir, en presencia de ATP, los siguientes dipéptidos: L-Ala-L-Gln más L-Ala-L-Ala, L-Ala-L-Phe, L-Ala-L-Met más L-Ala-L-Ala, L-Ala-L-Leu más L-Ala-L-Ala, o L-Ala-L-Val más L-Ala-L-Ala de L-Ala más L-Gln, L-Phe, L-Met, L-Leu o L-Val; y Gly-L-Phe o Gly-L-Met de Gly más L-Phe o L-Met.

Ejemplo experimental 4

15 Producción de dipéptidos usando la enzima recombinante con etiqueta His (2)

Se preparó una mezcla de reacción (0,1 ml) que comprendía 0,04 mg de la enzima recombinante con etiqueta de His purificada obtenida en el ejemplo experimental 2, Tris-HCl 100 mmol/l (pH 8,0), cloruro de magnesio 60 mmol/l y ATP 60 mmol/l. A esta mezcla se añadieron respectivamente combinaciones de varios L-aminoácidos, Gly y β-Ala seleccionados de los aminoácidos mostrados en la primera fila de la tabla 1 y en la columna más a la izquierda de la tabla 1 para dar una concentración de 30 mmol/l cada uno, y las mezclas resultantes se sometieron a reacción a 37°C durante 16 horas. Después de completar las reacciones, los productos de reacción se analizaron por HPLC, por lo cual se confirmó que se formaron los dipéptidos mostrados en la tabla 1.

Tabla 1-1

	Ala	Gln	Glu	Gly	Val	Leu	Ile	Pro
Ala	AlaAla	AlaGln AlaAla	AlaAla	AlaGly AlaAla	AlaVal AlaAla	AlaLeu AlaAla	AlaIle AlaAla	AlaAla
Gln		×	×	GlyGln GlyGly	×	×	×	×
Glu				GlyGly				
Gly				GlyGly				GlyGly
Val								
Leu								
Ile								
Pro								
Phe								
Trp								
Met								
Ser								
Thr								
Cys								
Asn								
Tyr								
Lys								
Arg								
His								
Asp								
α AB								
β -Ala								
Cit								

25

Tabla 1-2

	Phe	Trp	Met	Ser	Thr	Cys	Asn	Tyr
Ala	AlaPhe AlaAla	AlaTrp AlaAla	AlaMet	AlaSer AlaAla	AlaThr AlaAla	AlaAla ○	AlaAsn AlaAla	AlaTyr AlaAla
Gln	○	×	MetMet	SerGln SerSer	ThrGln ThrThr	○	×	×
Glu								
Gly	GlyPhe	GlyGly ○	GlyMet GlyGly	GlySer GlyGly SerGly SerSer	GlyThr GlyGly ThrGly ThrThr	GlyGly ○	GlyGly	GlyTyr GlyGly
Val			×					
Leu			MetMet		ThrLeu			
Ile			MetMet					
Pro			MetMet	SerSer	ThrThr			
Phe			MetPhe MetMet	SerPhe	ThrPhe ThrThr			
Trp								
Met			MetMet	SerMet	ThrMet ThrThr	MetMet ○		MetTyr MetMet
Ser				SerSer	SerThr SerSer ThrSer ThrThr			SerTyr SerSer
Thr					ThrThr			
Cys								
Asn								
Tyr								
Lys								
Arg								
His								
Asp								
α-AB								
β-Ala								
Cit								

Tabla 1-3

	Lys	Arg	His	Asp	$\alpha$ -AB	$\beta$ -Ala	Cit	Aza-serina	Tea-nina
Ala	AlaAla ○	AlaArg AlaAla	AlaHis AlaAla	AlaAla	AlaAla ○		AlaAla ○	AlaAla ○	AlaAla ○
Gln	×	×	×	×	○				
Glu									
Gly	GlyGly ○	GlyArg GlyGly	GlyGly	GlyGly	GlyGly ○		○		
Val									
Leu									
Ile									
Pro									
Phe				×		○			
Trp									
Met	MetMet ○		MetMet ○			○			
Ser			SerHis		SerSer ○				
Thr					ThrThr ○				
Cys									
Asn									
Tyr									
Lys									
Arg					○				
His						$\beta$ -AlaHis			
Asp									
$\alpha$ -AB					○				
$\beta$ -Ala									
Cit						○			

5 Los dipéptidos formados por la reacción que usa, como sustratos, dos (o uno) tipos de L-aminoácidos, Gly y  $\beta$ -Ala mostrados en la primera fila y la columna más a la izquierda de la tabla 1 se muestran en las respectivas células de la tabla. En la tabla, O significa que se formó un dipéptido, aunque su secuencia no se identificó; X significa que la formación de un dipéptido no se confirmó; y un blanco significa que la reacción no llevó a cabo.

10 Ejemplo experimental 5

Producción de un dipéptido usando la cepa que expresa la enzima recombinante con etiqueta de His

15 Se inoculó *Escherichia coli* NM522/pQE60ywfE obtenida en el ejemplo experimental 1 en 8 ml de medio LB que contenía ampicilina 50  $\mu$ g/ml en un tubo de ensayo, y se cultivó a 28°C durante 17 horas. El cultivo resultante se inoculó en 50 ml de medio LB que contenía ampicilina 50  $\mu$ g/ml en matraz Erlenmeyer de 250 ml, y se cultivó a 30°C durante 3 horas. A continuación, se añadió IPTG para dar una concentración final de 1 mmol/l, seguido por cultivo adicional a 30°C durante 4 horas. El cultivo resultante se centrifugó para obtener células húmedas.

20 Una mezcla de reacción (20 ml, pH 7,2) que comprendía células húmedas 200 g/l, glucosa 50 g/l, ácido fítico (diluido a neutralidad con solución de hidróxido de sodio concentrado al 33%) 5 g/l, dihidrogenofosfato de potasio 15 g/l, sulfato de magnesio heptahidrato 5 g/l, Nymeen S-125 4 g/l, xileno 10 ml/l, L-Ala 200 mmol/l y L-Gln 200 mmol/l se puso en un vaso de precipitado de 50 ml, y la reacción se llevó a cabo a 32°C a 900 rpm durante 2 horas. Durante la reacción, el pH de la mezcla de reacción se mantuvo a 7,2 usando hidróxido de potasio 2 mol/l.

25 El producto de reacción se analizó por el mismo método que en el ejemplo experimental 3, por lo cual se confirmó que se acumuló L-Ala-L-Gln 25 mg/l.

30 Ejemplo experimental 6

Clonación de genes correspondientes al gen ywfE de varios microorganismos del género Bacillus y análisis de los mismos

5 En base a la secuencia de nucleótidos mostrada en SEQ ID NO: 26, se obtuvieron genes correspondientes al gen ywfE que existe en Bacillus subtilis ATCC 15245, ATCC 6633, IAM 1213, IAM 1107, IAM 1214, ATCC 9466, IAM 1033 y ATCC 21555, Bacillus amyloliquefaciens IFO 3022 y Bacillus pumilus NRRL B-12025 de la siguiente manera.

10 Es decir, Bacillus subtilis ATCC 15245, ATCC 6633, IAM 1213, IAM 1107, IAM 1214, ATCC 9466, IAM 1033 y ATCC 21555, Bacillus amyloliquefaciens IFO 3022 y Bacillus pumilus NRRL B-12025 se inocularon respectivamente en medio LB y se sometieron a cultivo estático durante la noche a 30°C. Después del cultivo, los ADN cromosómicos de los respectivos microorganismos se aislaron y purificaron según el método que usa fenol saturado descrito en Current Protocols in Molecular Biology.

15 Usando un sintetizador de ADN (Modelo 8905, PerSpective Biosystems, Inc.), se sintetizaron ADN que tenían las secuencias de nucleótidos mostradas en SEQ ID NO: 37 y 38 (de aquí en adelante denominadas cebador C y cebador D, respectivamente). El cebador C tiene una secuencia que contiene la región anterior del codón de iniciación del gen ywfE del ADN cromosómico de Bacillus subtilis 168, y el cebador D tiene una secuencia complementaria a una secuencia que contiene una región posterior del codón de terminación del gen ywfE.

20 Se llevó a cabo PCR usando cada uno de los ADN cromosómicos de Bacillus subtilis ATCC 15245, ATCC 6633, IAM 1213, IAM 1107, IAM 1214, ATCC 9466, IAM 1033 y ATCC 21555, y Bacillus amyloliquefaciens IFO 3022 como molde y los anteriores cebador C y cebador D como un conjunto de cebadores. Es decir, se llevó a cabo PCR mediante 30 ciclos, un ciclo consiste en reacción a 94°C durante un minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos, usando 40 µl de una mezcla de reacción que comprendía 0,1 µg del ADN cromosómico, 0,5 µmol/l de cada uno de los cebadores, 2,5 unidades de ADN polimerasa Pfu, 4 µl de tampón para ADN polimerasa Pfu (10x) y 200 µmol/l de cada uno de los dNTP.

30 Un décimo de cada una de las mezclas de reacción resultantes se sometió a electroforesis en gel de agarosa para confirmar que un fragmento de aprox. 1,4 kb correspondiente al gen ywfE se amplificó. A continuación, la mezcla de reacción restante se mezcló con una cantidad igual de fenol/cloroforno saturado con TE. La solución resultante se centrifugó, y la fase superior obtenida se mezcló con un volumen de dos veces de etanol frío y se dejó reposar a -80°C durante 30 minutos. La solución resultante se centrifugó, y el precipitado de ADN obtenido se disolvió en 20 µl de TE.

35 Cada uno de los fragmentos de ADN de 1,4 kb así obtenidos derivados de los ADN cromosómicos de las respectivas cepas y pCR-blunt (Invitrogen Corp.) se sometieron a reacción de ligación usando un kit de ligación a 16°C durante 16 horas.

40 Se transformó Escherichia coli NM522 usando cada mezcla de reacción de ligación según el método que usa ion calcio, se extendió en medio LB agar que contenía ampicilina 50 µg/ml, y se cultivó durante la noche a 30°C.

45 Se extrajo un plásmido de una colonia de cada transformante que creció en el medio según un método conocido y la estructura de cada plásmido se analizó usando enzimas de restricción. Como resultado, se confirmó que se obtuvieron los siguientes plásmidos contenían un gen correspondiente al gen ywfE: pYWFE1 (derivado de ATCC 15245, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 65), pYWFE2 (derivado de ATCC 6633, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 27), pYWFE3 (derivado de IAM 1213, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 28), pYWFE4 (derivado de IAM 1107, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 29), pYWFE5 (derivado de IAM 1214, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 30), pYWFE6 (derivado de ATCC 9466, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 26), pYWFE7 (derivado de IAM 1033, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 65), pYWFE8 (derivado de ATCC 21555, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 31) y pYWFE9 (derivado de IFO 3022, ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 32).

55 Por otra parte, se obtuvo un gen correspondiente al gen ywfE de Bacillus pumilus NRRL B-12025 (ADN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 64) de la siguiente manera.

60 Se llevó a cabo PCR usando el ADN cromosómico de la cepa NRRL B-12025 preparado anteriormente como molde y ADN que consistían, respectivamente, en las secuencias de nucleótidos mostradas en SEQ ID NO: 66 y 67 como un conjunto de cebadores. Es decir, se llevó a cabo PCR mediante 30 ciclos, un ciclo consistía en reacción a 98°C durante 5 segundos, reacción a 55°C durante 30 segundos y reacción a 72°C durante un minuto, usando 50 µl de una mezcla de reacción que comprendía 0,1 µg del ADN cromosómico, 0,5 µmol/l de cada uno de los cebadores, 2,5 unidades de polimerasa Z-taq (Takara Bio Inc.), 5 µl de tampón para polimerasa Z-taq (10x) (Takara Bio, Inc.) y 200 µmol/l de cada uno de los dNTP.

65

## ES 2 626 571 T3

- Un décimo de la mezcla de reacción resultante se sometió a electroforesis en gel de agarosa para confirmar que un fragmento de aprox. 0,8 kb se amplificó. A continuación, la mezcla de reacción restante se mezcló con una cantidad igual de fenol/cloroformo saturado con TE. La solución resultante se centrifugó, y la fase superior obtenida se mezcló con un volumen de dos veces de etanol frío y se dejó reposar a -80°C durante 30 minutos. La solución resultante se centrifugó, y el precipitado de ADN obtenido se disolvió en 20 µl de TE.
- El fragmento de 0,8 kb así obtenido derivado del ADN cromosómico y pGEM T-easy (Promega Corp.) se sometieron a reacción de ligación usando un kit de ligación a 16°C durante 16 horas.
- Se transformó *Escherichia coli* DH5α usando la mezcla de reacción de ligación según el método que usa ion calcio, se extendió en medio LB agar que contenía ampicilina 50 µg/ml, y se cultivó durante la noche a 30°C.
- Se extrajo un plásmido del transformante obtenido anteriormente y la secuencia de nucleótidos del inserto de ADN de aprox. 0,8 kb se determinó, por lo cual se confirmó una secuencia de nucleótidos de 358 a 1160 en la secuencia de nucleótidos mostrada en SEQ ID NO: 64.
- El plásmido anterior se cortó con *EcoRI* y después se sometió a electroforesis en gel de agarosa para separar un fragmento de ADN. El fragmento de ADN se purificó usando el kit GENE CLEAN II, y aproximadamente 0,5 µg del fragmento de ADN purificado se marcaron con DIG usando el kit DIG-High Prime DNA Labeling & Detection Starter Kit I (Roche Diagnostics Corp.) según las instrucciones adjuntas en el mismo.
- Se llevó a cabo análisis por Southern del ADN cromosómico de la cepa NRRL B-12025 usando el ADN marcado con DIG obtenido anteriormente.
- El ADN cromosómico de la cepa NRRL B-12025 se digirió completamente con *BamHI*, *EcoRI*, *HindIII*, *KpnI*, *PstI*, *SacI*, *Sall* y *SphI*, respectivamente, y se sometió a electroforesis en gel de agarosa para separar fragmentos de ADN, seguido por transferencia a membrana de nailon con carga positiva (Roche Diagnostics Corp.) según un método habitual.
- Después de fijar los fragmentos de ADN a la membrana de nailon por irradiación UV, se llevó a cabo hibridación Southern usando el ADN sonda anterior y la membrana de nailon.
- La hibridación se llevó a cabo poniendo en contacto la membrana de nailon con el ADN sonda a 65°C durante 16 horas, lavando la membrana de nailon dos veces con una solución que consistía en SDS al 0,1% y SSC 2x a temperatura ambiente durante 5 minutos, y lavando además la membrana dos veces con una solución que consistía en SDS al 0,1% y SSC 0,5x a 65°C durante 15 minutos. Las otras operaciones y condiciones y detección del ADN hibridado se llevaron a cabo según las instrucciones adjuntas al DIG-High Prime DNA Labeling & Detection Starter Kit I.
- Como resultado, se observó desarrollo de color a aproximadamente 3,5 kpb de los fragmentos completamente digeridos con *HindIII* y *PstI*.
- Posteriormente, el ADN cromosómico de la cepa NRRL B-12025 se digirió completamente con *HindIII* y *PstI*, respectivamente, y se sometió a electroforesis en gel de agarosa para separar fragmentos de ADN. De los respectivos ADN digeridos con enzimas de restricción, se purificaron fragmentos de 3-4 kpb usando el kit GENE CLEAN II, seguido por autociclación usando un kit de ligación.
- En base a la secuencia de nucleótidos del fragmento de ADN de 0,8 kb determinada anteriormente, se diseñaron y sintetizaron las secuencias de nucleótidos mostradas en SEQ ID NO: 71 y 72, y se usaron en PCR usando el ADN ciclado obtenido anteriormente como molde. Se llevó a cabo PCR mediante 30 ciclos, un ciclo consistía en reacción a 98°C durante 5 segundos, reacción a 55°C durante 30 segundos y reacción a 72°C durante un minuto, usando 50 µl de una mezcla de reacción que comprendía 10 ng del ADN ciclado, 0,5 µmol/l de cada uno de los cebadores, 2,5 unidades de polimerasa pyrobest (Takara Bio Inc.), 5 µl de tampón para polimerasa pyrobest (10x) (Takara Bio, Inc.) y 200 µmol/l de cada uno de los dNTP.
- Un décimo de la mezcla de reacción resultante se sometió a electroforesis en gel de agarosa para confirmar que un fragmento de aprox. 3,0 kb se amplificó. A continuación, la mezcla de reacción restante se mezcló con una cantidad igual de fenol/cloroformo saturado con TE. La solución resultante se centrifugó, y la fase superior obtenida se mezcló con un volumen de dos veces de etanol frío y se dejó reposar a -80°C durante 30 minutos. La solución resultante se centrifugó, y el precipitado de ADN obtenido se disolvió en 20 µl de TE.
- El fragmento de ADN así obtenido y Zero Blunt PCR Cloning Kit (Invitrogen Corp.) se sometieron a reacción de ligación usando un kit de ligación.
- Se transformó *Escherichia coli* NM522 usando la mezcla de reacción según el método que usa ion calcio, se extendió en medio LB agar que contenía ampicilina 50 µg/ml, y se cultivó durante la noche a 30°C.

Se extrajo un plásmido de una colonia del transformante que creció en el medio según un método conocido y la estructura del plásmido se analizó usando enzimas de restricción. Como resultado, se confirmó que se obtuvo el plásmido pYWFE10 (derivado de NRRL B-12025, ADN que tiene la secuencia de nucleótidos mostrada en SEQ UD NO: 64) que contenía un gen correspondiente al gen ywfE.

Las secuencias de nucleótidos de los genes correspondientes al gen ywfE que están respectivamente contenidas en los plásmidos pYWFE1 a pYWFE10 obtenidos anteriormente se determinaron usando el secuenciador de ADN 373A.

Las secuencias de aminoácidos de las proteínas codificadas por los genes respectivamente contenidos en pYWFE1, pYWFE6 y pYWFE7 eran idénticas a la secuencia de aminoácidos de la proteína codificada por el gen ywfE, mientras que las de las proteínas codificadas por los genes respectivamente contenidos en pYWFE2, pYWFE3, pYWFE4, pYWFE5, pYWFE8, pYWFE9 y pYWFE10 eran diferentes de la secuencia de aminoácidos de la proteína codificada por el gen ywfE.

Las secuencias de aminoácidos de las proteínas por los genes correspondientes al gen ywfE que están contenidos en pYWFE2, pYWFE3, pYWFE4, pYWFE5, pYWFE8, pYWFE9 y pYWFE10, y pYWFE1, y pYWFE7 se muestran en SEQ ID NO: 20 a 25, 68 y 19, respectivamente, y las secuencias de nucleótidos de estos genes se muestran en SEQ ID NO: 27 a 32, 65 y 26, respectivamente.

#### Ejemplo experimental 7

Purificación de dipéptido sintetasa recombinante con etiqueta de His C-terminal

Se llevó a cabo PCR usando cada uno de los ADN cromosómicos de Bacillus subtilis ATCC 15245, ATCC 6633, IAM 1213, IAM 1107, IAM 1214, ATCC 9466, IAM 1033 y ATCC 21555, y Bacillus amyloliquefaciens IFO 3022 como molde y el cebador A y cebador B descritos en el ejemplo experimental 1, como un conjunto de cebadores. Es decir, se llevó a cabo PCR mediante 30 ciclos, un ciclo consiste en reacción a 94°C durante un minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos, usando 40 µl de una mezcla de reacción que comprendía 0,1 µg del ADN cromosómico, 0,5 µmol/l de cada uno de los cebadores, 2,5 unidades de ADN polimerasa Pfu, 4 µl de tampón para ADN polimerasa Pfu (10x) y 200 µmol/l de cada uno de los dNTP.

Cuando se usó como molde el ADN cromosómico de Bacillus pumilus NRRL B-12025, se llevó a cabo PCR usando ADN que tienen respectivamente las secuencias de nucleótidos mostradas en SEQ ID NO: 69 y 70 como un conjunto de cebadores en las mismas condiciones que anteriormente.

Un décimo de cada una de las mezclas de reacción resultantes se sometió a electroforesis en gel de agarosa para confirmar que un fragmento de aprox. 1,4 kb correspondiente al fragmento ywfE se amplificó. A continuación, la mezcla de reacción restante se mezcló con una cantidad igual de fenol/cloroformo saturado con TE. La mezcla resultante se centrifugó, y la fase superior obtenida se mezcló con un volumen de dos veces de etanol frío y se dejó reposar a -80°C durante 30 minutos. La solución resultante se centrifugó, y el precipitado de ADN obtenido se disolvió en 20 µl de TE.

Cada una de las soluciones así obtenidas (5 µl) se sometió a reacción para cortar el ADN amplificado con las enzimas de restricción NcoI y BamHI. Los fragmentos de ADN se separaron por electroforesis en gel de agarosa, y un fragmento de ADN de 1,4 kb que contenía un gen correspondiente al gen ywfE se recuperó usando el kit GENECLAN II.

Posteriormente, se cortaron 0,2 µg del vector de expresión recombinante con etiqueta de His C-terminal pQE60 con las enzimas de restricción NcoI y BamHI. Se separaron fragmentos de ADN por electroforesis en gel de agarosa, y un fragmento de ADN de 3,4 kb se recuperó de la misma manera que anteriormente.

Cada uno de los fragmentos de ADN de 1,4 kb que contenía un gen correspondiente al gen ywfE de Bacillus subtilis 168 y el fragmento de ADN de 3,4 kb obtenidos anteriormente se sometieron a reacción de ligación usando un kit de ligación a 16°C durante 16 horas. Se transformó Escherichia coli NM52 usando cada mezcla de reacción de ligación según el método que usa ion calcio, se extendió en medio LB agar que contenía ampicilina 50 µg/ml, y se cultivó durante la noche a 30°C.

Se extrajo un plásmido de cada colonia de cada transformante que creció en el medio según un método conocido y la estructura de cada plásmido se analizó usando enzimas de restricción. Como resultado, se confirmó que se obtuvieron los siguientes vectores de expresión de genes con etiqueta de His C-terminal: pQE60ywfE1 (un vector que contiene el gen derivado de ATCC 15245), pQE60ywfE2 (un vector que contiene el gen derivado de ATCC 6633), pQE60ywfE3 (un vector que contiene el gen derivado de IAM 1213), pQE60ywfE4 (un vector que contiene el gen derivado de IAM 1107), pQE60ywfE5 (un vector que contiene el gen derivado de IAM 1214), pQE60ywfE6 (un vector que contiene el gen derivado de ATCC 9466), pQE60ywfE7 (un vector que contiene el gen derivado de IAM

1033), pQE60ywfE8 (un vector que contiene el gen derivado de ATCC 21555), pQE60ywfE9 (un vector que contiene el gen derivado de IFO 3022) y pQE60ywfE10 (un vector que contiene el gen derivado de NRRL B-12025).

5 Las cepas *Escherichia coli* NM522/pQE60ywfE1 a NM522/pQE60ywfE10 obtenidas anteriormente se inocularon respectivamente en 8 ml de medio LB, que contenía ampicilina 50 µg/ml en un tubo de ensayo, y se cultivó a 28°C durante 17 horas. Cada uno de los cultivos resultantes se inoculó en 50 ml de medio LB que contenía ampicilina 50 µg/ml en matraz Erlenmeyer de 250 ml, y se cultivó a 30°C durante 3 horas. A continuación, se añadió IPTG para dar una concentración final de 1 mmol/l, seguido por cultivo adicional a 30°C durante 4 horas. El cultivo resultante se centrifugó para obtener células húmedas, y las enzimas recombinantes con etiqueta His se purificaron de las  
10 respectivas células húmedas usando HisTrap según las instrucciones adjuntas a las mismas.

#### Ejemplo experimental 8

Producción de dipéptidos usando enzimas purificadas

15 Se prepararon mezclas de reacción (0,1 ml cada una) que comprendían 0,04 mg de las respectivas enzimas recombinantes obtenidas en el ejemplo experimental 7, Tris-HCl 100 mmol/l (pH 8,0), cloruro de magnesio 60 mmol/l, ATP 60 mmol/l, L-Ala 30 mmol/l y L-Gln 30 mmol/l, y las reacciones se llevaron a cabo a 37°C durante 16 horas.

20 Después de completar las reacciones, las mezclas de reacción se analizaron por el método descrito en el ejemplo experimental 3, por lo cual se confirmó que L-Ala-L-Gln de 3,0 a 3,5 g/l y L-Ala-L-Ala de 0,25 a 0,30 g/l se formaron y acumularon.

25 Cuando se excluyó el ATP de las composiciones de las mezclas de reacción anteriores, no se formaron en absoluto L-Ala-L-Gln o L-Ala-L-Ala.

Los resultados anteriores revelaron que todos los productos de los genes obtenidos en el ejemplo experimental 7 tienen la actividad para producir L-Ala-L-Gln y L-Ala-L-Ala a partir de L-Ala y L-Gln en presencia de ATP.

30

#### Ejemplo experimental 9

Construcción de *Escherichia coli* para expresión aumentada del gen ywfE

35 Usando un sintetizador de ADN (modelo 8905, PerSpective, Inc.), se sintetizaron ADN que tenían las secuencias mostradas en SEQ ID NO: 60 a 63 (de aquí en adelante denominados cebador E, cebador F, cebador G y cebador H, respectivamente). La secuencia de SEQ ID NO: 60 es una secuencia en donde se añade una secuencia que contiene la secuencia de reconocimiento de XhoI en el extremo 5' de una región que contiene la secuencia Shine-Dalgarno (secuencia de unión a ribosomas) del gen ywfE en el plásmido pQE60ywfE. La secuencia de SEQ ID NO:  
40 61 es una secuencia en donde se añade una secuencia que contiene la secuencia de reconocimiento de BamHI al extremo 5' de una secuencia complementaria a una secuencia que contiene el codón de terminación del gen ywfE. La secuencia de SEQ ID NO: 62 es una secuencia en donde se añade una secuencia que contiene la secuencia de reconocimiento de EcoRI al extremo 5' de la secuencia de la región promotora trp del vector de expresión pTrS30 que contiene el promotor trp. La secuencia de SEQ ID NO: 63 es una secuencia en donde se añade una secuencia  
45 que contiene la secuencia de reconocimiento de XhoI en el extremo 5' de una secuencia complementaria a la secuencia de la región promotora trp del vector de expresión pTrS30 que contiene el promotor trp.

50 Se amplificaron un fragmento del gen ywfE y un fragmento de la región promotora trp por PCR usando los anteriores cebadores E y F y cebadores G y H como un conjunto de cebadores, respectivamente, y el plásmido pQE60ywfE como molde. Se llevó a cabo PCR mediante 30 ciclos, un ciclo consiste en reacción a 94°C durante un minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos, usando 40 µl de una mezcla de reacción que comprendía 10 ng de pQE60ywfE, 0,5 µmol/l de cada uno de los cebadores, 2,5 unidades de ADN polimerasa Pfu, 4 µl de tampón para ADN polimerasa Pfu (10x) y 200 µmol/l de cada uno de los dNTP.

55 Un décimo de cada una de las mezclas de reacción resultantes se sometió a electroforesis en gel de agarosa para confirmar que un fragmento de aprox. 1,4 kb correspondiente al fragmento del gen ywfE y un fragmento de aprox. 0,3 kb correspondiente al fragmento de la región promotora trp se amplificaron respectivamente en la PCR usando el cebador E y el cebador F y la PCR usando el cebador G y el cebador H. A continuación, la mezcla de reacción restante se mezcló con una cantidad igual de fenol/cloroformo saturado con TE. La mezcla resultante se centrifugó,  
60 y la fase superior obtenida se mezcló con un volumen de dos veces de etanol frío y se dejó reposar a -80°C durante 30 minutos. La solución resultante se centrifugó, y el precipitado de ADN obtenido se disolvió en 20 µl de TE.

Las soluciones de ADN así obtenidas (5 µl de cada una) se sometieron respectivamente a reacción para cortar el ADN amplificado usando el cebador E y cebador F con las enzimas de restricción XhoI y BamHI y a una reacción  
65 para cortar el ADN amplificado usando el cebador G y cebador H con las enzimas de restricción EcoRI y XhoI. Los fragmentos de ADN se separaron por electroforesis en gel de agarosa, y un fragmento de 1,4 kb que contenía el gen

ywfE y un fragmento de 0,3 kb que contenía la región promotora trp se recuperaron respectivamente usando el kit GENECLAN II.

5 pTrs30 (un vector de expresión que contiene el promotor trp, 0,2 µg) se cortó con las enzimas de restricción EcoRI y BamHI. Se separaron fragmentos de ADN por electroforesis en gel de agarosa y un fragmento de ADN de 4,5 kb se recuperó de la misma manera que anteriormente.

10 El fragmento de 1,4 kb que contenía el gen ywfE, el fragmento de 0,3 kb que contenía la región promotora trp y el fragmento de ADN de 4,5 kb obtenido anteriormente se sometieron a reacción de ligación usando un kit de ligación a 16°C durante 16 horas.

Se transformó Escherichia coli NM52 usando la mezcla de reacción según el método que usa ion calcio, se extendió en medio LB agar que contenía ampicilina 50 µg/ml, y se cultivó durante la noche a 30°C.

15 Se extrajo un plásmido de cada colonia del transformante que creció en el medio según un método conocido, por lo cual se obtuvo el vector de expresión pPE56 que contiene el gen ywfE en una posición posterior del promotor trp. La estructura del vector se confirmó por digestión con enzimas de restricción (Fig. 2).

20 Ciertas formas de realización de la presente invención se ilustran en los siguientes ejemplos. Estos ejemplos no se deben interpretar como limitantes del ámbito de la invención.

#### Ejemplo 1

25 Preparación de cepas que tienen deleciones del gen pepD, gen pepN, gen pepB, gen pepA y operón dpp

Se prepararon cepas en que genes específicos en el ADN cromosómico de Escherichia coli se delecionan según el método que utiliza el sistema de recombinación homóloga del fago lambda [Proc. Natl. Acad. Sci. USA, 97, 6641-6645 (2000)].

30 Se prepararon los plásmidos pKD46, pKD3 y pCP20 usados posteriormente por extracción de cepas de Escherichia coli que los llevan que se obtuvieron del Escherichia coli Genetic Stock Center, Universidad de Yale, EE UU.

#### (1) Clonación de fragmentos de ADN para la delección de genes

35 Para el fin de deleccionar los siguientes genes que existen en el ADN cromosómico de Escherichia coli K12, se sintetizaron ADN que tienen secuencias de nucleótidos homólogas a secuencias de nucleótidos de 36 pb que están antes y después de los genes respectivos que se van a deleccionar en el ADN cromosómico de Escherichia coli K12 y la secuencia de nucleótidos mostrada en SEQ ID NO: 39 que está reconocida por la recombinasa derivada de levadura Flp, usando un sintetizador de ADN (modelo 8905, PerSpective, Inc.). Los genes que se van a deleccionar son el gen pepD que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 10, gen pepN que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 11, gen pepB que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 12, gen pepA que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 13, gen dppA que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 14, gen dppB que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 15, gen dppC que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 16, gen dppD que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 17 y gen dppF que tiene la secuencia de nucleótidos mostrada en SEQ ID NO: 18. En el caso del gen dppA, gen dppB, gen dppC, gen dppD y gen dppF, que forman un operón, se sintetizaron ADN que tienen secuencias de nucleótidos homólogas a las secuencias de nucleótidos que están antes y después del operón.

50 Es decir, se sintetizaron ADN que consistían en las siguientes secuencias de nucleótidos como conjuntos respectivos de cebadores para amplificación de fragmentos de ADN para delección de genes: SEQ ID NO: 40 y 41 para delección del gen pepD; SEQ ID NO: 42 y 43 para delección del gen pepN; SEQ ID NO: 44 y 45 para delección del gen pepA; SEQ ID NO: 46 y 47 para delección del gen pepB; y SEQ ID NO: 48 y 49 para delección del operón dpp.

55 Posteriormente, se llevó a cabo PCR usando cada conjunto de los ADN sintéticos anteriores como un conjunto de cebadores y ADN de pKD3 como molde. Es decir, se llevó a cabo PCR mediante 30 ciclos, un ciclo consiste en reacción a 94°C durante un minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos, usando 40 µl de una mezcla de reacción que comprendía 10 ng de ADN del plásmido, 0,5 µmol/l de cada uno de los cebadores, 2,5 unidades de ADN polimerasa Pfu (Stratagene), 4 µl de tampón para ADN polimerasa Pfu (10x) (Stratagene) y 200 µmol/l de cada uno de los desoxiNTP (dATP, dGTP, dCTP y dTTP).

60 Un décimo de cada una de las mezclas de reacción resultantes se sometió a electroforesis en gel de agarosa para confirmar que el fragmento deseado se amplificó. A continuación, la mezcla de reacción restante se mezcló con una cantidad igual de fenol/cloroformo (1 vol/1 vol) saturado con TE [Tris-HCl 10 mmol/l (pH 8,0), EDTA 1 mmol/l].

65

La mezcla resultante se centrifugó, y la fase superior obtenida se mezcló con un volumen de dos veces de etanol frío y se dejó reposar a -80°C durante 30 minutos, seguido por centrifugación. Mediante este procedimiento, se obtuvieron fragmentos de ADN que contienen el gen de resistencia a cloranfenicol para la delección del gen pepD, gen pepN, gen pepB, gen pepA y operón dpp.

(2) Preparación de Escherichia coli JM101 que tiene delección del gen pepD

Se transformó Escherichia coli JM101 con pKD46, se extendió en medio LB agar que contenía ampicilina 100 mg/l y se cultivó a 30°C para seleccionar un transformante.

El plásmido pKD46 lleva insertado el gen de la recombinasa Red de  $\lambda$  y se diseña de modo que la expresión del gen se induzca por L-arabinosa. Según esto, cuando la Escherichia coli hecha crecer en presencia de L-arabinosa se transforma usando un ADN lineal, se produce recombinación homóloga con alta frecuencia. Además, como pKD46 tiene un origen de replicación termosensible, el curado del plásmido se puede producir fácilmente cultivando la cepa a 42°C.

El fragmento de ADN que contiene el gen de resistencia a cloranfenicol para la delección del gen pepD obtenido anteriormente se introdujo en Escherichia coli JM101/pKD46 obtenida cultivando en presencia de L-arabinosa 10 mmol/l y ampicilina 50  $\mu$ g/ml por electroporación. Las células resultantes se extendieron en medio LB agar (bactotripton 10 g/l, extracto de bacto-levadura 5 g/l, cloruro de sodio 5 g/l y agar 15 g/l) que contenía cloranfenicol 25 mg/l y se cultivaron a 30°C para seleccionar un transformante en el que el fragmento de ADN que contiene el gen de resistencia a cloranfenicol para la delección del gen pepD se integró en el ADN cromosómico de Escherichia coli JM101 por recombinación homóloga.

La cepa resistente a cloranfenicol seleccionada se inoculó en medio LB agar que contenía cloranfenicol 25 mg/l y se cultivó a 42°C durante 14 horas, seguido por aislamiento de colonias individuales. Se hicieron réplicas de las colonias obtenidas en medio LB agar que contenía cloranfenicol 25 mg/l y medio LB agar que contenía ampicilina 100 mg/l, seguido por cultivo a 37°C. Al seleccionar una colonia que muestra resistencia a cloranfenicol y sensibilidad a ampicilina, se obtuvo una cepa pKD46-curada.

La cepa pKD46-curada así obtenida se transformó usando pCP20, seguido por selección en medio LB agar que contenía ampicilina 100 mg/l para obtener una cepa pKD46-curada que tiene pCP20.

El plásmido pCP20 lleva insertado el gen de la recombinasa derivada de levadura F1p y está diseñado de modo que la expresión del gen se induce a una temperatura de 42°C.

Los fragmentos de ADN que contienen el gen de resistencia a cloranfenicol para delección del gen pepD, gen pepN, gen pepB, gen pepA y operón dpp preparados anteriormente contienen secuencias de nucleótidos reconocidas por la recombinasa F1p en ambos extremos del gen de resistencia a cloranfenicol. Por tanto, el gen de resistencia se puede deleccionar fácilmente por recombinación homóloga catalizada por la recombinasa F1p.

Además, como pCP20 tiene un origen de replicación termosensible, la expresión de la recombinasa F1p y curado de pCP20 se puede inducir simultáneamente cultivando la cepa que tiene pCP20 a 42°C.

La cepa pKD46-curada que tiene pCP20 obtenida anteriormente se inoculó en medio LB agar sin fármacos y se cultivó a 42°C durante 14 horas, seguido por aislamiento de colonias individuales. Se hicieron réplicas de las colonias obtenidas en medio LB agar sin fármacos, medio LB agar que contenía cloranfenicol 25 mg/l y medio LB agar que contenía ampicilina 100 mg/l, seguido por cultivo a 30°C. A continuación, las colonias que mostraban sensibilidad a cloranfenicol y sensibilidad a ampicilina se seleccionaron.

Se prepararon ADN cromosómicos de las respectivas cepas seleccionadas anteriormente según un método habitual [Seibutsukogaku Jikkensho (Experiments in Biotechnology), editado por The Society for Biotechnology, Japón, p. 97-98, Baifukan (1992)]. Se llevó a cabo PCR usando, como un conjunto de cebadores, los ADN que tienen las secuencias de nucleótidos mostradas en SEQ ID NO: 50 y 51 que se diseñaron basadas en una secuencia de nucleótidos interna del gen pepD que se va a deleccionar, y usando cada uno de los ADN cromosómicos como molde. Es decir, se llevó a cabo PCR mediante 30 ciclos, un ciclo consiste en reacción a 94°C durante un minuto, reacción a 55°C durante 2 minutos y reacción a 72°C durante 3 minutos, usando 40  $\mu$ l de una mezcla de reacción que comprendía 0,1  $\mu$ g del ADN cromosómico, 0,5  $\mu$ mol/l de cada uno de los cebadores, 2,5 unidades de ADN polimerasa Pfu, 4  $\mu$ l de tampón para ADN polimerasa Pfu (10x) y 200  $\mu$ mol/l de cada uno de los desoxiNTP.

Una cepa en la que no se detectó fragmento de ADN amplificado en la PCR anterior se identificó como una cepa que tenía delección del gen pepD y se designó Escherichia coli JPD1.

(3) Preparación de una cepa en la que el gen pepD y el gen pepN en el ADN cromosómico de Escherichia coli JM101 se deleccionan

Se transformó *Escherichia coli* JPD1 obtenida en el anterior (2) con pKD46, se extendió en medio LB agar que contenía ampicilina 100 mg/l, y se cultivó a 30°C para seleccionar un transformante. El fragmento de ADN que tiene el gen de resistencia a cloranfenicol para la delección del gen pepN se introdujo en el transformante obtenido (*Escherichia coli* JPD1/pKD46) por electroporación para obtener un transformante en el que el fragmento de ADN que contiene el gen de resistencia a cloranfenicol para la delección del gen pepN se integró en el ADN cromosómico de *Escherichia coli* JPD1/pKD46 por recombinación homóloga.

Posteriormente, se llevó a cabo el mismo procedimiento que el anterior (2) para obtener una cepa en la que el gen de resistencia a cloranfenicol se delecionó del ADN cromosómico, que se designó *Escherichia coli* JPDN2.

(4) Preparación de cepas en las que el gen pepN, gen pepA, gen pepB u operón ddp en el ADN cromosómico de *Escherichia coli* JM101 se deleciona y cepas que tiene delección múltiple de genes

Se prepararon las cepas que tienen delección del gen pepN, gen pepA, gen pepB u operón ddp según el mismo procedimiento que en el anterior (2) usando los respectivos fragmentos de ADN que contienen del gen de resistencia a cloranfenicol para delección de genes u operón preparados en el anterior (1).

La adquisición de las cepas que tienen delecciones de genes por el método anterior se confirmó llevando a cabo PCR de la misma manera que en el anterior (2) usando, como conjuntos de cebadores, ADN que tiene las secuencias de nucleótidos mostradas en SEQ ID NO: 52 a 59 que se diseñaron y sintetizaron basadas en secuencias de nucleótidos internas de los respectivos genes que se van a delecionar. Es decir, se usaron ADN que tienen las siguientes secuencias de nucleótidos como los respectivos conjuntos de cebadores para la confirmación de delección de genes: SEQ ID NO: 52 y 53 para delección del gen pepN; SEQ ID NO: 54 y 55 para delección del gen pepA; SEQ ID NO: 56 y 57 para delección del gen pepB; y SEQ ID NO: 58 y 59 para delección del operón ddp.

La cepa con operón ddp deleccionado, cepa con gen pepN deleccionado, cepa con gen pepA deleccionado y cepa con gen pepB deleccionado así obtenidas se designaron como *Escherichia coli* JDPP1, *Escherichia coli* JPN1, *Escherichia coli* JPA1 y *Escherichia coli* JPB7, respectivamente.

Además, cepas que tenían múltiples delecciones de genes, es decir, delecciones de dos o más genes u operón seleccionados del grupo que consiste en el gen pepD, gen pepN, gen pepB, gen pepA y operón ddp se prepararon según el método del anterior (3). La adquisición de las cepas que tienen múltiples delecciones de genes se confirmó por PCR similar a la del anterior (2). La cepa con doble gen deleccionado así obtenida que tiene delecciones del gen pepD y operón ddp así obtenida se designó *Escherichia coli* JPDP49, la cepa con triple delección de genes que tiene delecciones del gen pepB, gen pepD y gen pepN como *Escherichia coli* JPDNB43, la cepa con triple delección de genes que tiene delecciones del gen pepD, gen pepN y operón ddp como *Escherichia coli* JPNDDP36, la cepa con cuádruple delección de genes que tiene delecciones del gen pepA, gen pepD, gen pepN y operón ddp como *Escherichia coli* JPNDAP5, y la cepa con cuádruple delección de genes que tiene delecciones del gen pepB, gen pepD, gen pepN y operón ddp como *Escherichia coli* JPNDBP7. Los genes deleccionados en las cepas con genes deleccionados se muestran en la tabla 2.

Tabla 2

Cepa	Gen deleccionado
JM101	ninguno
JDPP1	operón ddp
JPN1	pepN
JPA1	pepA
JPB7	pepB
JPD1	pepD
JPDN2	pepD, pepN
JPNDB43	pepB, pepD, pepN
JPDP49	pepD, operón ddp
JPNDDP36	pepD, pepN, operón ddp
JPNDAP5	pepA, pepD, pepN, operón ddp
JPNDBP7	pepB, pepD, pepN, operón ddp

Ejemplo 2

Evaluación de la productividad de L-alanil-L-glutamina (de aquí en adelante denominado AlaGln) y L-alanil-L-alanina (de aquí en adelante denominada AlaAla) por cepas de *Escherichia coli* en que las actividades peptidasa y transportadora de péptidos se pierden

Las cepas que tienen delecciones de los genes que codifican varias peptidasas y un operón que codifica una proteína transportadora de péptidos que se obtuvieron en el ejemplo 1 se transformaron usando el plásmido pPE56 construido en el ejemplo experimental 8 para obtener transformantes resistentes a ampicilina.

5 Cada uno de los transformantes obtenidos se inoculó en 8 ml de medio LB que contenía ampicilina 50 µg/ml en un tubo de ensayo y se cultivaron a 28°C durante 17 horas. El cultivo resultante se inoculó en 8 ml de un medio de producción [hidrogenofosfato dipotásico 16 g/l, dihidrogenofosfato potásico 14 g/l, sulfato de amonio 5 g/l, ácido cítrico (anhidro) 1 g/l, casaminoácidos (Difco) 0,5 g/l, L-Pro 1 g/l, L-Ala 2,5 g/l, L-Gln 2,5 g/l, glucosa 10 g/l, vitamina B<sub>1</sub> 10 mg/l, sulfato de magnesio heptahidrato 25 mg/l y sulfato ferroso heptahidrato 50 mg/l; pH ajustado a 7,2 con solución de hidróxido de sodio 10 mol/l; L-Gln se añadió después de la esterilización por filtración de una solución concentrada 10 veces; glucosa, vitamina B<sub>1</sub>, sulfato de magnesio heptahidrato y sulfato ferroso heptahidrato se añadieron después de autoclavado separado] que contenía ampicilina 100µg/ml en un tubo de ensayo en una cantidad del 1% y se cultivó a 30°C durante 24 horas. El cultivo resultante se centrifugó para obtener un sobrenadante de cultivo.

15 El producto en el sobrenadante de cultivo se derivó por el método de F-moc y después se analizó por HPLC. El análisis por HPLC se llevó a cabo usando ODS-HG5 (Nomura Kagaku, Co., Ltd.) como columna de separación y una solución A (ácido acético 6 ml/l y acetonitrilo al 20% (v/v), pH ajustado a 4,8 con trietilamina) y solución B (ácido acético 6 ml/l y acetonitrilo al 70% (v/v), pH ajustado a 4,8 con trietilamina) como soluciones de elución. La proporción de solución A respecto a la solución B fue 8:2 durante los primeros 5 minutos de elución y después de ello cambió con un gradiente lineal de modo que la proporción se volvió 1:1 a los 20 minutos después del inicio de la elución. Los resultados del análisis se muestran en la tabla 3.

20 Tabla 3

Cepa	Gen deletado	AlaGln (g/l)	AlaAla (g/l)
JM101	ninguno	0	0
JDPP1	operón ddp	0,02	0,01
JPN1	pepN	0,01	0,01
JPA1	pepA	0,01	0,01
JPB7	pepB	0,01	0,01
JPD1	pepD	0,01	0,01
JPDN2	pepD, pepN	0,02	0,03
JPNDB43	pepB, pepD, pepN	0,05	0,12
JPDP49	pepD, operón ddp	0,11	0,08
JPNDDP36	pepD, pepN, operón ddp	0,16	0,21
JPNDAP5	pepA, pepD, pepN, operón ddp	0,28	0,26
JPNDBP7	pepB, pepD, pepN, operón ddp	0,43	0,22

25 Como se puede ver de la tabla 3, pequeñas cantidades de péptidos se formaron y acumularon mediante el uso de los microorganismos que tienen deleciones de dos o menos tipos de genes de peptidasas o un tipo de gen de proteína transportadora de péptidos, mientras que las cantidades de dipéptidos formados y acumulados estaban muy aumentadas mediante el uso de los microorganismo que tienen deleciones de uno o más tipos de genes de peptidasas y un tipo de gen de proteína transportadora de péptidos o microorganismos que tienen deleciones de tres o más tipos de genes de peptidasas.

### 30 Ejemplo 3

Evaluación de la productividad de L-alanil-L-valina (de aquí en adelante denominado AlaVal) por cepas de Escherichia coli en que las actividades peptidasa y transportadora de péptidos se pierden

35 De forma similar al ejemplo 2, las cepas de Escherichia coli que tienen deleciones de varios genes de peptidasas y un operón que codifica una proteína transportadora de péptidos se transformaron usando pPE56. Cada uno de los transformantes obtenidos se inoculó en 8 ml de medio LB que contenía ampicilina 50 µg/ml en un tubo de ensayo y se cultivaron a 28°C durante 17 horas. El cultivo resultante se inoculó en 8 ml de un medio de producción [hidrogenofosfato dipotásico 16 g/l, dihidrogenofosfato potásico 14 g/l, sulfato de amonio 5 g/l, ácido cítrico (anhidro) 1 g/l, casaminoácidos (Difco) 0,5 g/l, L-Pro 1 g/l, L-Ala 2,5 g/l, L-Val 2,5 g/l, glucosa 10 g/l, vitamina B<sub>1</sub> 10 mg/l, sulfato de magnesio heptahidrato 25 mg/l y sulfato ferroso heptahidrato 50 mg/l; pH ajustado a 7,2 con solución de hidróxido de sodio 10 mol/l; glucosa, vitamina B<sub>1</sub>, sulfato de magnesio heptahidrato y sulfato ferroso heptahidrato se añadieron después de autoclavado separado] que contenía ampicilina 100µg/ml en un tubo de ensayo en una cantidad del 1% y se cultivó a 30°C durante 24 horas. El cultivo resultante se centrifugó para obtener un sobrenadante de cultivo.

45 El producto en el sobrenadante de cultivo se analizó por el método descrito en el ejemplo 2. Los resultados se muestran en la tabla 4.

Tabla 4

Cepa	Gen deletado	AlaVal (g/l)
JM101	ninguno	0
JDPP1	operón ddp	0
JPN1	pepN	0
JPA1	pepA	0
JPB7	pepB	0
JPD1	pepD	0
JPDN2	pepD, pepN	0
JPNDB43	pepB, pepD, pepN	0,04
JPDP49	pepD, operón ddp	0,11
JPNDDP36	pepD, pepN, operón ddp	0,22
JPNDP7	pepB, pepD, pepN, operón ddp	0,20

5 Como se puede ver de la tabla 4, el dipéptido no se produjo mediante el uso de los microorganismos que tienen deleciones de dos o menos tipos de genes de peptidasa o un tipo de gen de proteína transportadora de péptidos, mientras que el dipéptido se produjo mediante el uso de microorganismo que tienen deleciones de tres o más tipos de genes de peptidasa o microorganismos que tienen deleciones de uno o más tipos de peptidasas y un tipo de gen de proteína transportadora de péptidos.

#### 10 Ejemplo 4

Evaluación de la productividad de glicil-L-glutamina (de aquí en adelante denominada GlyGln) por cepas de *Escherichia coli* en que las actividades peptidasa y transportadora de péptidos se pierden

15 De forma similar al ejemplo 2, las cepas de *Escherichia coli* que tienen deleciones de varios genes de peptidasas y un operón que codifica una proteína transportadora de péptidos se transformaron usando pPE56. Cada uno de los transformantes obtenidos se inoculó en 8 ml de medio LB que contenía ampicilina 50 µg/ml en un tubo de ensayo y se cultivaron a 28°C durante 17 horas.

20 El cultivo resultante se inoculó en 8 ml de un medio de producción [hidrogenofosfato dipotásico 16 g/l, dihidrogenofosfato potásico 14 g/l, sulfato de amonio 5 g/l, ácido cítrico (anhidro) 1 g/l, casaminoácidos (Difco) 0,5 g/l, L-Pro 1 g/l, Gly 2,5 g/l, L-Gln 2,5 g/l, glucosa 10 g/l, vitamina B<sub>1</sub> 10 mg/l, sulfato de magnesio heptahidrato 25 mg/l y sulfato ferroso heptahidrato 50 mg/l; pH ajustado a 7,2 con solución de hidróxido de sodio 10 mol/l; L-Gln se añadió después de la esterilización por filtración de una solución concentrada 10 veces; glucosa, vitamina B<sub>1</sub>, sulfato de magnesio heptahidrato y sulfato ferroso heptahidrato se añadieron después de autoclavado separado] que contenía ampicilina 100µg/ml en un tubo de ensayo en una cantidad del 1% y se cultivó a 30°C durante 24 horas. El cultivo resultante se centrifugó para obtener un sobrenadante de cultivo.

25 El producto en el sobrenadante de cultivo se analizó por el método descrito en el ejemplo 2. los resultados se muestran en la tabla 5.

30

Tabla 5

Cepa	Gen deletado	GlyGln(g/l)
JM101	ninguno	0
JDPP1	operón ddp	0
JPDN2	pepD, pepN	0
JPNDB43	pepB, pepD, pepN	0,01
JPNDDP36	pepD, pepN, operón ddp	0,02
JPNDP7	pepB, pepD, pepN, operón ddp	0,03

35 Como se puede ver de la tabla 5, el dipéptido no se produjo mediante el uso de los microorganismos que tienen deleciones de dos o menos tipos de genes de peptidasa o un tipo de gen de proteína transportadora de péptidos, mientras que el dipéptido se produjo mediante el uso de microorganismo que tienen deleciones de tres o más tipos de genes de peptidasa o microorganismos que tienen deleciones de uno o más tipos de peptidasas y un tipo de gen de proteína transportadora de péptidos.

#### 40 Lista de secuencias texto libre

40

SEQ ID NO: 35 – Descripción de secuencia artificial: ADN sintético

SEQ ID NO: 36 – Descripción de secuencia artificial: ADN sintético

SEQ ID NO: 37 – Descripción de secuencia artificial: ADN sintético

SEQ ID NO: 38 – Descripción de secuencia artificial: ADN sintético

45

SEQ ID NO: 39 – Descripción de secuencia artificial: ADN sintético

SEQ ID NO: 40 – Descripción de secuencia artificial: ADN sintético

- SEQ ID NO: 41 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 42 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 43 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 44 – Descripción de secuencia artificial: ADN sintético  
 5 SEQ ID NO: 45 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 46 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 47 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 48 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 49 – Descripción de secuencia artificial: ADN sintético  
 10 SEQ ID NO: 50 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 51 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 52 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 53 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 54 – Descripción de secuencia artificial: ADN sintético  
 15 SEQ ID NO: 55 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 56 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 57 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 58 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 59 – Descripción de secuencia artificial: ADN sintético  
 20 SEQ ID NO: 60 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 61 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 62 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 63 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 66 – Descripción de secuencia artificial: ADN sintético  
 25 SEQ ID NO: 67 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 69 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 70 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 71 – Descripción de secuencia artificial: ADN sintético  
 SEQ ID NO: 72 – Descripción de secuencia artificial: ADN sintético  
 30

**Lista de secuencias**

- <110> KYOWA HAKKO KOGYO CO., LTD.  
 35 <120> Microorganismos que producen dipéptidos y procesos para producir dipéptidos usando los microorganismos  
 <130> K 2817 EP  
 <150> JP 2003-375823  
 40 <151> 5-11-2003  
 <150> JP 2004-189010  
 <151> 25-06-2004  
 45 <160> 72  
 <170> PatentIn ver. 2.1  
 <210> 1  
 50 <211> 503  
 <212> PRT  
 <213> Escherichia coli  
 <400> 1

ES 2 626 571 T3

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

ES 2 626 571 T3

Gln Gly Ser Gln Asn Glu Ser Leu Met Ser Val Ile Glu Tyr Lys Gly  
 245 250 255  
 Asn Ala Ser Glu Asp Ala Arg Pro Ile Val Leu Val Gly Lys Gly Leu  
 260 265 270  
 Thr Phe Asp Ser Gly Gly Ile Ser Ile Lys Pro Ser Glu Gly Met Asp  
 275 280 285  
 Glu Met Lys Tyr Asp Met Cys Gly Ala Ala Ala Val Tyr Gly Val Met  
 290 295 300  
 Arg Met Val Ala Glu Leu Gln Leu Pro Ile Asn Val Ile Gly Val Leu  
 305 310 315 320  
 Ala Gly Cys Glu Asn Met Pro Gly Gly Arg Ala Tyr Arg Pro Gly Asp  
 325 330 335  
 Val Leu Thr Thr Met Ser Gly Gln Thr Val Glu Val Leu Asn Thr Asp  
 340 345 350  
 Ala Glu Gly Arg Leu Val Leu Cys Asp Val Leu Thr Tyr Val Glu Arg  
 355 360 365  
 Phe Glu Pro Glu Ala Val Ile Asp Val Ala Thr Leu Thr Gly Ala Cys  
 370 375 380  
 Val Ile Ala Leu Gly His His Ile Thr Gly Leu Met Ala Asn His Asn  
 385 390 395 400  
 Pro Leu Ala His Glu Leu Ile Ala Ala Ser Glu Gln Ser Gly Asp Arg  
 405 410 415  
 Ala Trp Arg Leu Pro Leu Gly Asp Glu Tyr Gln Glu Gln Leu Glu Ser  
 420 425 430  
 Asn Phe Ala Asp Met Ala Asn Ile Gly Gly Arg Pro Gly Gly Ala Ile  
 435 440 445  
 Thr Ala Gly Cys Phe Leu Ser Arg Phe Thr Arg Lys Tyr Asn Trp Ala  
 450 455 460  
 His Leu Asp Ile Ala Gly Thr Ala Trp Arg Ser Gly Lys Ala Lys Gly  
 465 470 475 480  
 Ala Thr Gly Arg Pro Val Ala Leu Leu Ala Gln Phe Leu Leu Asn Arg  
 485 490 495  
 Ala Gly Phe Asn Gly Glu Glu  
 500

<210> 2  
 <211> 427  
 <212> PRT  
 <213> Escherichia coli

5

<400> 2  
 Met Thr Glu Ala Met Lys Ile Thr Leu Ser Thr Gln Pro Ala Asp Ala  
 1 5 10 15  
 Arg Trp Gly Glu Lys Ala Thr Tyr Ser Ile Asn Asn Asp Gly Ile Thr  
 20 25 30  
 Leu His Leu Asn Gly Ala Asp Asp Leu Gly Leu Ile Gln Arg Ala Ala  
 35 40 45

ES 2 626 571 T3

Arg Lys Ile Asp Gly Leu Gly Ile Lys His Val Gln Leu Ser Gly Glu  
 50 55 60  
 Gly Trp Asp Ala Asp Arg Cys Trp Ala Phe Trp Gln Gly Tyr Lys Ala  
 65 70 75 80  
 Pro Lys Gly Thr Arg Lys Val Val Trp Pro Asp Leu Asp Asp Ala Gln  
 85 90 95  
 Arg Gln Glu Leu Asp Asn Arg Leu Met Ile Ile Asp Trp Val Arg Asp  
 100 105 110  
 Thr Ile Asn Ala Pro Ala Glu Glu Leu Gly Pro Ser Gln Leu Ala Gln  
 115 120 125  
 Arg Ala Val Asp Leu Ile Ser Asn Val Ala Gly Asp Arg Val Thr Tyr  
 130 135 140  
 Arg Ile Thr Lys Gly Glu Asp Leu Arg Glu Gln Gly Tyr Met Gly Leu  
 145 150 155 160  
 His Thr Val Gly Arg Gly Ser Glu Arg Ser Pro Val Leu Leu Ala Leu  
 165 170 175  
 Asp Tyr Asn Pro Thr Gly Asp Lys Glu Ala Pro Val Tyr Ala Cys Leu  
 180 185 190  
 Val Gly Lys Gly Ile Thr Phe Asp Ser Gly Gly Tyr Ser Ile Lys Gln  
 195 200 205  
 Thr Ala Phe Met Asp Ser Met Lys Ser Asp Met Gly Gly Ala Ala Thr  
 210 215 220  
 Val Thr Gly Ala Leu Ala Phe Ala Ile Thr Arg Gly Leu Asn Lys Arg  
 225 230 235 240  
 Val Lys Leu Phe Leu Cys Cys Ala Asp Asn Leu Ile Ser Gly Asn Ala  
 245 250 255  
 Phe Lys Leu Gly Asp Ile Ile Thr Tyr Arg Asn Gly Lys Lys Val Glu  
 260 265 270  
 Val Met Asn Thr Asp Ala Glu Gly Arg Leu Val Leu Ala Asp Gly Leu  
 275 280 285  
 Ile Asp Ala Ser Ala Gln Lys Pro Glu Met Ile Ile Asp Ala Ala Thr  
 290 295 300  
 Leu Thr Gly Ala Ala Lys Thr Ala Leu Gly Asn Asp Tyr His Ala Leu  
 305 310 315 320  
 Phe Ser Phe Asp Asp Ala Leu Ala Gly Arg Leu Leu Ala Ser Ala Ala  
 325 330 335  
 Gln Glu Asn Glu Pro Phe Trp Arg Leu Pro Leu Ala Glu Phe His Arg  
 340 345 350  
 Ser Gln Leu Pro Ser Asn Phe Ala Glu Leu Asn Asn Thr Gly Ser Ala  
 355 360 365  
 Ala Tyr Pro Ala Gly Ala Ser Thr Ala Ala Gly Phe Leu Ser His Phe  
 370 375 380  
 Val Glu Asn Tyr Gln Gln Gly Trp Leu His Ile Asp Cys Ser Ala Thr  
 385 390 395 400  
 Gly Val Arg Thr Ile Ala Asn Leu Leu Thr Ala  
 420 425

- <210> 3
- 5 <211> 485
- <212> PRT
- <213> Escherichia coli
- <400> 3

ES 2 626 571 T3

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

ES 2 626 571 T3

Val Ala Asn Asp Lys Ala Ala Leu Ile Ala Lys Ser Arg Asp Thr Phe  
 305 310 315 320  
 Ile Arg Leu Leu Asn Ala Thr Pro Asn Gly Val Ile Arg Asn Ser Asp  
 325 330 335  
 Val Ala Lys Gly Val Val Glu Thr Ser Leu Asn Val Gly Val Val Thr  
 340 345 350  
 Met Thr Asp Asn Asn Val Glu Ile His Cys Leu Ile Arg Ser Leu Ile  
 355 360 365  
 Asp Ser Gly Lys Asp Tyr Val Val Ser Met Leu Asp Ser Leu Gly Lys  
 370 375 380  
 Leu Ala Gly Ala Lys Thr Glu Ala Lys Gly Ala Tyr Pro Gly Trp Gln  
 385 390 395 400  
 Pro Asp Ala Asn Ser Pro Val Met His Leu Val Arg Glu Thr Tyr Gln  
 405 410 415  
 Arg Leu Phe Asn Lys Thr Pro Asn Ile Gln Ile Ile His Ala Gly Leu  
 420 425 430  
 Glu Cys Gly Leu Phe Lys Lys Pro Tyr Pro Glu Met Asp Met Val Ser  
 435 440 445  
 Ile Gly Pro Thr Ile Thr Gly Pro His Ser Pro Asp Glu Gln Val His  
 450 455 460  
 Ile Glu Ser Val Gly His Tyr Trp Thr Leu Leu Thr Glu Leu Leu Lys  
 465 470 475 480  
 Glu Ile Pro Ala Lys  
 485

<210> 4  
 <211> 870  
 <212> PRT  
 <213> Escherichia coli

5

<400> 4  
 Met Thr Gln Gln Pro Gln Ala Lys Tyr Arg His Asp Tyr Arg Ala Pro  
 1 5 10 15  
 Asp Tyr Gln Ile Thr Asp Ile Asp Leu Thr Phe Asp Leu Asp Ala Gln  
 20 25 30  
 Lys Thr Val Val Thr Ala Val Ser Gln Ala Val Arg His Gly Ala Ser  
 35 40 45  
 Asp Ala Pro Leu Arg Leu Asn Gly Glu Asp Leu Lys Leu Val Ser Val  
 50 55 60  
 His Ile Asn Asp Glu Pro Trp Thr Ala Trp Lys Glu Glu Glu Gly Ala  
 65 70 75 80  
 Leu Val Ile Ser Asn Leu Pro Glu Arg Phe Thr Leu Lys Ile Ile Asn  
 85 90 95  
 Glu Ile Ser Pro Ala Ala Asn Thr Ala Leu Glu Gly Leu Tyr Gln Ser  
 100 105 110  
 Gly Asp Ala Leu Cys Thr Gln Cys Glu Ala Glu Gly Phe Arg His Ile  
 115 120 125  
 Thr Tyr Tyr Leu Asp Arg Pro Asp Val Leu Ala Arg Phe Thr Thr Lys  
 130 135 140

ES 2 626 571 T3

Ile Ile Ala Asp Lys Ile Lys Tyr Pro Phe Leu Leu Ser Asn Gly Asn  
145 150 155 160  
Arg Val Ala Gln Gly Glu Leu Glu Asn Gly Arg His Trp Val Gln Trp  
165 170 175  
Gln Asp Pro Phe Pro Lys Pro Cys Tyr Leu Phe Ala Leu Val Ala Gly  
180 185 190  
Asp Phe Asp Val Leu Arg Asp Thr Phe Thr Thr Arg Ser Gly Arg Glu  
195 200 205  
Val Ala Leu Glu Leu Tyr Val Asp Arg Gly Asn Leu Asp Arg Ala Pro  
210 215 220  
Trp Ala Met Thr Ser Leu Lys Asn Ser Met Lys Trp Asp Glu Glu Arg  
225 230 235 240  
Phe Gly Leu Glu Tyr Asp Leu Asp Ile Tyr Met Ile Val Ala Val Asp  
245 250 255  
Phe Phe Asn Met Gly Ala Met Glu Asn Lys Gly Leu Asn Ile Phe Asn  
260 265 270  
Ser Lys Tyr Val Leu Ala Arg Thr Asp Thr Ala Thr Asp Lys Asp Tyr  
275 280 285  
Leu Asp Ile Glu Arg Val Ile Gly His Glu Tyr Phe His Asn Trp Thr  
290 295 300  
Gly Asn Arg Val Thr Cys Arg Asp Trp Phe Gln Leu Ser Leu Lys Glu  
305 310 315 320  
Gly Leu Thr Val Phe Arg Asp Gln Glu Phe Ser Ser Asp Leu Gly Ser  
325 330 335  
Arg Ala Val Asn Arg Ile Asn Asn Val Arg Thr Met Arg Gly Leu Gln  
340 345 350  
Phe Ala Glu Asp Ala Ser Pro Met Ala His Pro Ile Arg Pro Asp Met  
355 360 365  
Val Ile Glu Met Asn Asn Phe Tyr Thr Leu Thr Val Tyr Glu Lys Gly  
370 375 380  
Ala Glu Val Ile Arg Met Ile His Thr Leu Leu Gly Glu Glu Asn Phe  
385 390 395 400  
Gln Lys Gly Met Gln Leu Tyr Phe Glu Arg His Asp Gly Ser Ala Ala  
405 410 415  
Thr Cys Asp Asp Phe Val Gln Ala Met Glu Asp Ala Ser Asn Val Asp  
420 425 430  
Leu Ser His Phe Arg Arg Trp Tyr Ser Gln Ser Gly Thr Pro Ile Val  
435 440 445  
Thr Val Lys Asp Asp Tyr Asn Pro Glu Thr Glu Gln Tyr Thr Leu Thr  
450 455 460  
Ile Ser Gln Arg Thr Pro Ala Thr Pro Asp Gln Ala Glu Lys Gln Pro  
465 470 475 480  
Leu His Ile Pro Phe Ala Ile Glu Leu Tyr Asp Asn Glu Gly Lys Val  
485 490 495  
Ile Pro Leu Gln Lys Gly Gly His Pro Val Asn Ser Val Leu Asn Val

ES 2 626 571 T3

500					505					510					
Thr	Gln	Ala	Glu	Gln	Thr	Phe	Val	Phe	Asp	Asn	Val	Tyr	Phe	Gln	Pro
		515					520					525			
Val	Pro	Ala	Leu	Leu	Cys	Glu	Phe	Ser	Ala	Pro	Val	Lys	Leu	Glu	Tyr
	530					535					540				
Lys	Trp	Ser	Asp	Gln	Gln	Leu	Thr	Phe	Leu	Met	Arg	His	Ala	Arg	Asn
545				550						555					560
Asp	Phe	Ser	Arg	Trp	Asp	Ala	Ala	Gln	Ser	Leu	Leu	Ala	Thr	Tyr	Ile
			565						570					575	
Lys	Leu	Asn	Val	Ala	Arg	His	Gln	Gln	Gly	Gln	Pro	Leu	Ser	Leu	Pro
			580					585					590		
Val	His	Val	Ala	Asp	Ala	Phe	Arg	Ala	Val	Leu	Leu	Asp	Glu	Lys	Ile
		595					600					605			
Asp	Pro	Ala	Leu	Ala	Ala	Glu	Ile	Leu	Thr	Leu	Pro	Ser	Val	Asn	Glu
	610					615					620				
Met	Ala	Glu	Leu	Phe	Asp	Ile	Ile	Asp	Pro	Ile	Ala	Ile	Ala	Glu	Val
625					630					635					640
Arg	Glu	Ala	Leu	Thr	Arg	Thr	Leu	Ala	Thr	Glu	Leu	Ala	Asp	Glu	Leu
				645					650					655	
Leu	Ala	Ile	Tyr	Asn	Ala	Asn	Tyr	Gln	Ser	Glu	Tyr	Arg	Val	Glu	His
			660					665					670		
Glu	Asp	Ile	Ala	Lys	Arg	Thr	Leu	Arg	Asn	Ala	Cys	Leu	Arg	Phe	Leu
		675					680					685			
Ala	Phe	Gly	Glu	Thr	His	Leu	Ala	Asp	Val	Leu	Val	Ser	Lys	Gln	Phe
	690					695					700				
His	Glu	Ala	Asn	Asn	Met	Thr	Asp	Ala	Leu	Ala	Ala	Leu	Ser	Ala	Ala
705					710					715					720
Val	Ala	Ala	Gln	Leu	Pro	Cys	Arg	Asp	Ala	Leu	Met	Gln	Glu	Tyr	Asp
			725						730					735	
Asp	Lys	Trp	His	Gln	Asn	Gly	Leu	Val	Met	Asp	Lys	Trp	Phe	Ile	Leu
			740					745					750		
Gln	Ala	Thr	Ser	Pro	Ala	Ala	Asn	Val	Leu	Glu	Thr	Val	Arg	Gly	Leu
		755					760					765			
Leu	Gln	His	Arg	Ser	Phe	Thr	Met	Ser	Asn	Pro	Asn	Arg	Ile	Arg	Ser
	770					775					780				
Leu	Ile	Gly	Ala	Phe	Ala	Gly	Ser	Asn	Pro	Ala	Ala	Phe	His	Ala	Glu
785					790					795					800
Asp	Gly	Ser	Gly	Tyr	Leu	Phe	Leu	Val	Glu	Met	Leu	Thr	Asp	Leu	Asn
				805					810					815	
Ser	Arg	Asn	Pro	Gln	Val	Ala	Ser	Arg	Leu	Ile	Glu	Pro	Leu	Ile	Arg
			820					825					830		
Leu	Lys	Arg	Tyr	Asp	Ala	Lys	Arg	Gln	Glu	Lys	Met	Arg	Ala	Ala	Leu
		835					840					845			
Glu	Gln	Leu	Lys	Gly	Leu	Glu	Asn	Leu	Ser	Gly	Asp	Leu	Tyr	Glu	Lys
	850						855				860				
Ile	Thr	Lys	Ala	Leu	Ala										
865				870											

<210> 5  
 <211> 535  
 <212> PRT

5

ES 2 626 571 T3

<213> Escherichia coli

<400> 5

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



ES 2 626 571 T3

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

His Lys Lys

<210> 7

<211> 300

5 <212> PRT

<213> Escherichia coli

<400> 7

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



ES 2 626 571 T3

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

<210> 9  
 <211> 334  
 <212> PRT  
 <213> Escherichia coli

5

<400> 9  
 Met Ser Thr Gln Glu Ala Thr Leu Gln Gln Pro Leu Leu Gln Ala Ile  
 1 5 10 15  
 Asp Leu Lys Lys His Tyr Pro Val Lys Lys Gly Met Phe Ala Pro Glu  
 20 25 30  
 Arg Leu Val Lys Ala Leu Asp Gly Val Ser Phe Asn Leu Glu Arg Gly  
 35 40 45  
 Lys Thr Leu Ala Val Val Gly Glu Ser Gly Cys Gly Lys Ser Thr Leu  
 50 55 60  
 Gly Arg Leu Leu Thr Met Ile Glu Met Pro Thr Gly Gly Glu Leu Tyr  
 65 70 75 80

ES 2 626 571 T3

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

<210> 10  
 <211> 1509  
 <212> ADN  
 <213> Escherichia coli

5

<400> 10  
 atg gag ttt agt gta aaa agc ggt agc ccg gag aaa cag cgg agt gcc 48  
 Met Glu Phe Ser Val Lys Ser Gly Ser Pro Glu Lys Gln Arg Ser Ala  
 1 5 10 15  
 tgc atc gtc gtg ggc gtc ttc gaa cca cgt cgc ctt tct ccg att gca 96  
 Cys Ile Val Val Gly Val Phe Glu Pro Arg Arg Leu Ser Pro Ile Ala  
 20 25 30  
 gaa cag ctc gat aaa atc agc gat ggg tac atc agc gcc ctg cta cgt 144  
 Glu Gln Leu Asp Lys Ile Ser Asp Gly Tyr Ile Ser Ala Leu Leu Arg  
 35 40 45  
 cgg ggc gaa ctg gaa gga aaa ccg ggg cag aca ttg ttg ctg cac cat 192

ES 2 626 571 T3

Arg	Gly	Glu	Leu	Glu	Gly	Lys	Pro	Gly	Gln	Thr	Leu	Leu	Leu	His	His		
	50					55					60						
ggt	ccg	aat	gta	ctt	tcc	gag	cga	att	ctc	ctt	att	ggt	tgc	ggc	aaa	240	
Val	Pro	Asn	Val	Leu	Ser	Glu	Arg	Ile	Leu	Leu	Ile	Gly	Cys	Gly	Lys	80	
	65				70					75							
gaa	cgt	gag	ctg	gat	gag	cgt	cag	tac	aag	cag	ggt	att	cag	aaa	acc	288	
Glu	Arg	Glu	Leu	Asp	Glu	Arg	Gln	Tyr	Lys	Gln	Val	Ile	Gln	Lys	Thr	95	
				85					90								
att	aat	acg	ctg	aat	gat	act	ggc	tca	atg	gaa	gcg	gtc	tgc	ttt	ctg	336	
Ile	Asn	Thr	Leu	Asn	Asp	Thr	Gly	Ser	Met	Glu	Ala	Val	Cys	Phe	Leu	110	
			100					105					110				
act	gag	ctg	cac	ggt	aaa	ggc	cgt	aac	aac	tac	tgg	aaa	gtg	cgt	cag	384	
Thr	Glu	Leu	His	Val	Lys	Gly	Arg	Asn	Asn	Tyr	Trp	Lys	Val	Arg	Gln	125	
		115					120					125					
gct	gtc	gag	acg	gca	aaa	gag	acg	ctc	tac	agt	ttc	gat	cag	ctg	aaa	432	
Ala	Val	Glu	Thr	Ala	Lys	Glu	Thr	Leu	Tyr	Ser	Phe	Asp	Gln	Leu	Lys	140	
	130					135					140						
acg	aac	aag	agc	gaa	ccg	cgt	cgt	ccg	ctg	cgt	aag	atg	gtg	ttc	aac	480	
Thr	Asn	Lys	Ser	Glu	Pro	Arg	Arg	Pro	Leu	Arg	Lys	Met	Val	Phe	Asn	160	
					150					155							
gtg	ccg	acc	cgc	cgt	gaa	ctg	acc	agc	ggt	gag	cgc	gcg	atc	cag	cac	528	
Val	Pro	Thr	Arg	Arg	Glu	Leu	Thr	Ser	Gly	Glu	Arg	Ala	Ile	Gln	His	175	
					165				170					175			
ggt	ctg	gcg	att	gcc	gcc	ggg	att	aaa	gca	gca	aaa	gat	ctc	ggc	aat	576	
Gly	Leu	Ala	Ile	Ala	Ala	Gly	Ile	Lys	Ala	Ala	Lys	Asp	Leu	Gly	Asn	190	
			180					185					190				
atg	ccg	ccg	aat	atc	tgt	aac	gcc	gct	tac	ctc	gct	tca	caa	gcg	cgc	624	
Met	Pro	Pro	Asn	Ile	Cys	Asn	Ala	Ala	Tyr	Leu	Ala	Ser	Gln	Ala	Arg	205	
		195					200					205					
cag	ctg	gct	gac	agc	tac	agc	aag	aat	gtc	atc	acc	cgc	ggt	atc	ggc	672	
Gln	Leu	Ala	Asp	Ser	Tyr	Ser	Lys	Asn	Val	Ile	Thr	Arg	Val	Ile	Gly	220	
	210					215					220						
gaa	cag	cag	atg	aaa	gag	ctg	ggg	atg	cat	tcc	tat	ctg	gcg	gtc	ggt	720	
Glu	Gln	Gln	Met	Lys	Glu	Leu	Gly	Met	His	Ser	Tyr	Leu	Ala	Val	Gly	240	
					230					235							
cag	ggt	tcg	caa	aac	gaa	tcg	ctg	atg	tcg	gtg	att	gag	tac	aaa	ggc	768	
Gln	Gly	Ser	Gln	Asn	Glu	Ser	Leu	Met	Ser	Val	Ile	Glu	Tyr	Lys	Gly	255	
				245					250					255			
aac	gcg	tcg	gaa	gat	gca	cgc	cca	atc	gtg	ctg	gtg	ggt	aaa	ggt	tta	816	
Asn	Ala	Ser	Glu	Asp	Ala	Arg	Pro	Ile	Val	Leu	Val	Gly	Lys	Gly	Leu	270	
			260					265					270				
acc	ttc	gac	tcc	ggc	ggt	atc	tcg	atc	aag	cct	tca	gaa	ggc	atg	gat	864	
Thr	Phe	Asp	Ser	Gly	Gly	Ile	Ser	Ile	Lys	Pro	Ser	Glu	Gly	Met	Asp	285	
		275					280					285					
gag	atg	aag	tac	gat	atg	tgc	ggt	gcg	gca	gcg	ggt	tac	ggc	gtg	atg	912	
Glu	Met	Lys	Tyr	Asp	Met	Cys	Gly	Ala	Ala	Ala	Val	Tyr	Gly	Val	Met	300	
	290					295											
cgg	atg	gtc	gcg	gag	cta	caa	ctg	ccg	att	aac	ggt	atc	ggc	gtg	ttg	960	
Arg	Met	Val	Ala	Glu	Leu	Gln	Leu	Pro	Ile	Asn	Val	Ile	Gly	Val	Leu	320	
					310					315							
gca	ggc	tgc	gaa	aac	atg	cct	ggc	gga	cga	gcc	tat	cgt	ccg	ggc	gat	1008	

ES 2 626 571 T3

Ala Gly Cys Glu Asn Met Pro Gly Gly Arg Ala Tyr Arg Pro Gly Asp  
 325 330 335  
 gtg tta acc acc atg tcc ggt caa acc gtt gaa gtg ctg aac acc gac 1056  
 Val Leu Thr Thr Met Ser Gly Gln Thr Val Glu Val Leu Asn Thr Asp  
 340 345 350  
 gct gaa ggc cgc ctg gta ctg tgc gac gtg tta act tac gtt gag cgt 1104  
 Ala Glu Gly Arg Leu Val Leu Cys Asp Val Leu Thr Tyr Val Glu Arg  
 355 360 365  
 ttt gag ccg gaa gcg gtg att gac gtg gcg acg ctg acc ggt gcc tgc 1152  
 Phe Glu Pro Glu Ala Val Ile Asp Val Ala Thr Leu Thr Gly Ala Cys  
 370 375  
 gtg atc gcg ctg ggt cat cat att act ggt ctg atg gcg aac cat aat 1200  
 Val Ile Ala Leu Gly His His Ile Thr Gly Leu Met Ala Asn His Asn  
 385 390 400  
 ccg ctg gcc cat gaa ctg att gcc gcg tct gaa caa tcc ggt gac cgc 1248  
 Pro Leu Ala His Glu Leu Ile Ala Ala Ser Glu Gln Ser Gly Asp Arg  
 405 410 415  
 gca tgg cgc tta ccg ctg ggt gac gag tat cag gaa caa ctg gag tcc 1296  
 Ala Trp Arg Leu Pro Leu Gly Asp Glu Tyr Gln Glu Gln Leu Glu Ser  
 420 425 430  
 aat ttt gcc gat atg gcg aac att ggc ggt cgt cct ggt ggg gcg att 1344  
 Asn Phe Ala Asp Met Ala Asn Ile Gly Gly Arg Pro Gly Gly Ala Ile  
 435 440 445  
 acc gca ggt tgc ttc ctg tca cgc ttt acc cgt aag tac aac tgg gcg 1392  
 Thr Ala Gly Cys Phe Leu Ser Arg Phe Thr Arg Lys Tyr Asn Trp Ala  
 450 455 460  
 cac ctg gat atc gcc ggt acc gcc tgg cgt tct ggt aaa gca aaa ggc 1440  
 His Leu Asp Ile Ala Gly Thr Ala Trp Arg Ser Gly Lys Ala Lys Gly  
 465 470 475  
 gcc acc ggt cgt ccg gta gcg ttg ctg gca cag ttc ctg tta aac cgc 1488  
 Ala Thr Gly Arg Pro Val Ala Leu Leu Ala Gln Phe Leu Leu Asn Arg  
 485 490 495  
 gct ggg ttt aac ggc gaa gag  
 Ala Gly Phe Asn Gly Glu Glu 1509  
 500

<210> 11  
 <211> 1281  
 <212> ADN  
 <213> Escherichia coli

<400> 11  
 atg aca gaa gcg atg aag att acc ctc tct acc caa cct gcc gac gcg 48  
 Met Thr Glu Ala Met Lys Ile Thr Leu Ser Thr Gln Pro Ala Asp Ala  
 1 5 10 15  
 cgc tgg gga gaa aaa gca act tac agc att aat aat gac ggc att acc 96  
 Arg Trp Gly Glu Lys Ala Thr Tyr Ser Ile Asn Asn Asp Gly Ile Thr  
 20 25 30  
 ctg cat ttg aac ggg gca gac gat ctg ggg ctg atc cag cgt gcg gcg 144  
 Leu His Leu Asn Gly Ala Asp Asp Leu Gly Leu Ile Gln Arg Ala Ala  
 35 40 45  
 cgc aag att gac ggt ctg ggc atc aag cat gtt cag tta agc ggt gaa 192  
 Arg Lys Ile Asp Gly Leu Gly Ile Lys His Val Gln Leu Ser Gly Glu

ES 2 626 571 T3

50			55			60										
ggc Gly 65	tgg Trp	gat Asp	gcg Ala	gat Asp	cgc Arg 70	tgc Cys	tgg Trp	gca Ala	ttc Phe	tgg Trp 75	caa Gln	ggt Gly	tac Tyr	aaa Lys	gcc Ala 80	240
ccg Pro	aaa Lys	ggc Gly	acg Thr	cgt Arg 85	aaa Lys	gtg Val	gtg Val	tgg Trp	ccg Pro 90	gat Asp	ctg Leu	gac Asp	gat Asp	gcc Ala 95	cag Gln	288
cgc Arg	cag Gln	gaa Glu	ctg Leu 100	gat Asp	aac Asn	cgc Arg	ctg Leu	atg Met 105	atc Ile	atc Ile	gac Asp	tgg Trp	gtg Val 110	cgt Arg	gac Asp	336
acc Thr	atc Ile	aac Asn 115	gca Ala	ccg Pro	gca Ala	gaa Glu	gaa Glu 120	ttg Leu	gga Gly	cca Pro	tcg Ser	caa Gln 125	ctg Leu	gca Ala	cag Gln	384
cgt Arg	gct Ala 130	gtt Val	gat Asp	ctg Leu	atc Ile	agc Ser 135	aac Asn	gtc Val	gcg Ala	ggc Gly	gat Asp 140	cgt Arg	gtg Val	act Thr	tat Tyr	432
cgg Arg 145	atc Ile	acc Thr	aaa Lys	ggc Gly	gaa Glu 150	gat Asp	ctg Leu	cgt Arg	gag Glu	caa Gln 155	ggc Gly	tat Tyr	atg Met	ggg Gly	ctg Leu 160	480
cac His	aca Thr	gtc Val	gga Gly	cgc Arg 165	ggc Gly	tca Ser	gaa Glu	cgt Arg	tct Ser 170	ccg Pro	gta Val	ttg Leu	ctg Leu	gcg Ala 175	ctg Leu	528
gat Asp	tac Tyr	aac Asn 180	cca Pro	act Thr	ggc Gly	gat Asp	aaa Lys	gaa Glu 185	gcg Ala	cca Pro	gtg Val	tac Tyr	gcg Ala 190	tgc Cys	ctg Leu	576
gta Val	ggc Gly	aaa Lys 195	ggc Gly	atc Ile	act Thr	ttt Phe	gac Asp 200	tcc Ser	ggc Gly	ggc Gly	tac Tyr	agc Ser 205	atc Ile	aaa Lys	cag Gln	624
act Thr	gcg Ala 210	ttt Phe	atg Met	gac Asp	tcg Ser	atg Met 215	aag Lys	tcg Ser	gac Asp	atg Met	ggc Gly 220	ggc Gly	gcg Ala	gca Ala	acg Thr	672
gtt Val 225	acc Thr	ggg Gly	gcg Ala	ctg Leu	gca Ala 230	ttt Phe	gcc Ala	att Ile	acg Thr	cgc Arg 235	gga Gly	ctg Leu	aac Asn	aag Lys	cgc Arg 240	720
gtg Val	aag Lys	ctg Leu	ttc Phe	ctc Leu 245	tgc Cys	tgt Cys	gcg Ala	gat Asp	aac Asn 250	ctg Leu	att Ile	agc Ser	ggc Gly	aat Asn 255	gcg Ala	768
ttc Phe	aag Lys	ctg Leu	ggc Gly 260	gat Asp	atc Ile	atc Ile	acc Thr	tat Tyr 265	cgc Arg	aac Asn	ggc Gly	aaa Lys	aaa Lys 270	ggt Val	gaa Glu	816
gtg Val	atg Met	aac Asn 275	act Thr	gat Asp	gcc Ala	gaa Glu	ggg Gly 280	cgt Arg	ctg Leu	gtg Val	ctt Leu	gcc Ala 285	gat Asp	ggc Gly	ctg Leu	864
att Ile	gat Asp 290	gcc Ala	agt Ser	gcg Ala	cag Gln	aaa Lys 295	ccg Pro	gaa Glu	atg Met	atc Ile	att Ile 300	gat Asp	gcg Ala	gcg Ala	acc Thr	912
ctc Leu 305	acc Thr	ggg Gly	gcg Ala	gcg Ala	aaa Lys 310	act Thr	gcg Ala	ctg Leu	ggc Gly	aat Asn 315	gat Asp	tat Tyr	cac His	gcg Ala	ctg Leu 320	960
ttc Phe	agt Ser	ttt Phe	gac Asp	gat Asp	gcg Ala	ctg Leu	gcc Ala	ggc Gly	cgc Arg	ttg Leu	ctg Leu	gcg Ala	agt Ser	gcc Ala	gcg Ala	1008

ES 2 626 571 T3

				325					330					335			
cag	gag	aac	gaa	ccg	ttc	tgg	cg	ctg	ccg	ctg	gcg	gag	ttc	cac	cg		1056
Gln	Glu	Asn	Glu	Pro	Phe	Trp	Arg	Leu	Pro	Leu	Ala	Glu	Phe	His	Arg		
			340					345					350				
agc	cag	ctg	ccg	tct	aac	ttt	gcc	gaa	ctg	aac	aat	acc	gga	agc	gcg		1104
Ser	Gln	Leu	Pro	Ser	Asn	Phe	Ala	Glu	Leu	Asn	Asn	Thr	Gly	Ser	Ala		
			355				360						365				
gcg	tat	ccg	gca	ggc	gcg	agc	acg	gcg	gcg	ggc	ttc	ctg	tcg	cac	ttt		1152
Ala	Tyr	Pro	Ala	Gly	Ala	Ser	Thr	Ala	Ala	Gly	Phe	Leu	Ser	His	Phe		
	370					375					380						
ggt	gag	aac	tat	cag	caa	ggc	tgg	ctg	cat	atc	gac	tgc	tcg	gcg	act		1200
Val	Glu	Asn	Tyr	Gln	Gln	Gly	Trp	Leu	His	Ile	Asp	Cys	Ser	Ala	Thr		
	385				390					395					400		
tac	cg	aaa	gcg	ccg	ggt	gaa	cag	tgg	tct	gcg	ggc	gct	acg	gga	ctt		1248
Tyr	Arg	Lys	Ala	Pro	Val	Glu	Gln	Trp	Ser	Ala	Gly	Ala	Thr	Gly	Leu		
				405					410					415			
ggt	gtg	cg	acg	ata	gct	aat	ctg	tta	acg	gcg							1281
Gly	Val	Arg	Thr	Ile	Ala	Asn	Leu	Leu	Thr	Ala							
			420					425									

<210> 12  
 <211> 1455  
 <212> ADN  
 <213> Escherichia coli

5

gtg	tct	gaa	ctg	tct	caa	tta	tct	cca	cag	ccg	ctg	tgg	gat	att	ttt		48
Val	Ser	Glu	Leu	Ser	Gln	Leu	Ser	Pro	Gln	Pro	Leu	Trp	Asp	Ile	Phe		
	1			5					10					15			
gcc	aaa	atc	tgt	tct	att	cct	cac	ccg	tcc	tat	cat	gaa	gag	caa	ctc		96
Ala	Lys	Ile	Cys	Ser	Ile	Pro	His	Pro	Ser	Tyr	His	Glu	Glu	Gln	Leu		
			20					25					30				
gct	gaa	tac	att	ggt	ggt	tgg	gca	aaa	gag	aaa	ggt	ttc	cat	gtc	gaa		144
Ala	Glu	Tyr	Ile	Val	Gly	Trp	Ala	Lys	Glu	Lys	Gly	Phe	His	Val	Glu		
		35				40						45					
cg	gat	cag	gta	ggt	aat	atc	ctg	att	cg	aaa	cct	gct	acc	gca	ggt		192
Arg	Asp	Gln	Val	Gly	Asn	Ile	Leu	Ile	Arg	Lys	Pro	Ala	Thr	Ala	Gly		
	50					55					60						
atg	gaa	aat	cg	aaa	ccg	gtc	gtc	tta	cag	gcc	cac	ctc	gat	atg	gtg		240
Met	Glu	Asn	Arg	Lys	Pro	Val	Val	Leu	Gln	Ala	His	Leu	Asp	Met	Val		
	65				70					75					80		
ccg	cag	aaa	aat	aac	gac	acc	gtg	cat	gac	ttc	acg	aaa	gat	cct	atc		288
Pro	Gln	Lys	Asn	Asn	Asp	Thr	Val	His	Asp	Phe	Thr	Lys	Asp	Pro	Ile		
				85					90					95			
cag	cct	tat	att	gat	ggc	gaa	tgg	ggt	aaa	gcg	cg	ggc	acc	acg	ctg		336
Gln	Pro	Tyr	Ile	Asp	Gly	Glu	Trp	Val	Lys	Ala	Arg	Gly	Thr	Thr	Leu		
			100					105					110				
ggt	gcg	gat	aac	ggc	att	ggt	atg	gcc	tct	gcg	ctg	gcg	ggt	ctg	gct		384
Gly	Ala	Asp	Asn	Gly	Ile	Gly	Met	Ala	Ser	Ala	Leu	Ala	Val	Leu	Ala		
		115				120						125					
gac	gaa	aac	gtg	ggt	cac	ggc	ccg	ctg	gaa	gtg	ctg	ctg	acc	atg	acc		432
Asp	Glu	Asn	Val	Val	His	Gly	Pro	Leu	Glu	Val	Leu	Leu	Thr	Met	Thr		
	130					135					140						

ES 2 626 571 T3

gaa Glu 145	gaa Glu	gcc Ala	ggt Gly	atg Met	gac Asp 150	ggt Gly	gcg Ala	ttc Phe	ggc Gly	tta Leu 155	cag Gln	ggc Gly	aac Asn	tgg Trp	ttg Leu 160	480
cag Gln	gct Ala	gat Asp	att Ile	ctg Leu 165	att Ile	aac Asn	acc Thr	gac Asp	tcc Ser 170	gaa Glu	gaa Glu	gaa Glu	ggt Gly	gaa Glu 175	atc Ile	528
tac Tyr	atg Met	ggt Gly	tgt Cys 180	gcg Ala	ggg Gly	ggt Gly	atc Ile	gac Asp 185	ttc Phe	acc Thr	tcc Ser	aac Asn	ctg Leu 190	cat His	tta Leu	576
gat Asp	cgt Arg	gaa Glu 195	gcg Ala	ggt Val	cca Pro	gct Ala	ggg Gly 200	ttt Phe	gaa Glu	acc Thr	ttc Phe	aag Lys 205	tta Leu	acc Thr	tta Leu	624
aaa Lys	ggt Gly 210	ctg Leu	aaa Lys	ggc Gly	ggt Gly	cac His 215	tcc Ser	ggc Gly	ggg Gly	gaa Glu	atc Ile 220	cac His	ggt Val	ggg Gly	ctg Leu	672
ggt Gly 225	aat Asn	gcc Ala	aac Asn	aaa Lys	ctg Leu 230	ctg Leu	gtg Val	cgc Arg	ttc Phe	ctg Leu 235	gcg Ala	ggt Gly	cat His	gcg Ala	gaa Glu 240	720
gaa Glu	ctg Leu	gat Asp	ctg Leu	cgc Arg 245	ctt Leu	atc Ile	gat Asp	ttc Phe	aac Asn 250	ggc Gly	ggc Gly	aca Thr	ctg Leu	cgf Arg 255	aac Asn	768
gcc Ala	atc Ile	ccg Pro	cgf Arg 260	gaa Glu	gcc Ala	ttt Phe	gcg Ala	acc Thr 265	att Ile	gct Ala	gtc Val	gca Ala	gct Ala 270	gat Asp	aaa Lys	816
gtc Val	gac Asp	gtc Val 275	ctg Leu	aaa Lys	tct Ser	ctg Leu	gtg Val 280	aat Asn	acc Thr	tat Tyr	cag Gln	gag Glu 285	atc Ile	ctg Leu	aaa Lys	864
aac Asn	gag Glu 290	ctg Leu	gca Ala	gaa Glu	aaa Lys	gag Glu 295	aaa Lys	aat Asn	ctg Leu	gcc Ala	ttg Leu 300	ttg Leu	ctg Leu	gac Asp	tct Ser	912
gta Val 305	gcg Ala	aac Asn	gat Asp	aaa Lys	gct Ala 310	gcc Ala	ctg Leu	att Ile	gcg Ala	aaa Lys 315	tct Ser	cgf Arg	gat Asp	acc Thr	ttt Phe 320	960
att Ile	cgf Arg	ctg Leu	ctg Leu	aac Asn 325	gcc Ala	acc Thr	ccg Pro	aac Asn	ggt Gly 330	gtg Val	att Ile	cgf Arg	aac Asn	tcc Ser 335	gat Asp	1008
gta Val	gcc Ala	aaa Lys	ggt Gly 340	gtg Val	ggt Val	gaa Glu	acc Thr 345	tcc Ser	ctg Leu	aac Asn	gtc Val	ggt Gly	gtg Val 350	gtg Val	acc Thr	1056
atg Met	act Thr	gac Asp 355	aat Asn	aac Asn	gta Val	gaa Glu	att Ile 360	cac His	tgc Cys	ctg Leu	atc Ile	cgf Arg 365	tca Ser	ctg Leu	atc Ile	1104
gac Asp	agc Ser	ggt Gly 370	aaa Lys	gac Asp	tac Tyr	gtg Val 375	gtg Val	agc Ser	atg Met	ctg Leu	gat Asp 380	tcg Ser	ctg Leu	ggt Gly	aaa Lys	1152
ctg Leu 385	gct Ala	ggc Gly	gcg Ala	aaa Lys	acc Thr 390	gaa Glu	gcg Ala	aaa Lys	ggc Gly	gca Ala 395	tat Tyr	cct Pro	ggc Gly	tgg Trp	cag Gln 400	1200
ccg Pro	gac Asp	gct Ala	aat Asn	tct Ser 405	ccg Pro	gtg Val	atg Met	cat His	ctg Leu 410	gta Val	cgf Arg	gaa Glu	acc Thr	tat Tyr 415	cag Gln	1248

ES 2 626 571 T3

cgc ctg ttc aac aag acg ccg aac atc cag att atc cac gcg ggc ctg 1296  
 Arg Leu Phe Asn Lys Thr Pro Asn Ile Gln Ile Ile His Ala Gly Leu  
 420 425 430  
 gaa tgt ggt ctg ttc aaa aaa ccg tat ccg gaa atg gac atg gtt tct 1344  
 Glu Cys Gly Leu Phe Lys Lys Pro Tyr Pro Glu Met Asp Met Val Ser  
 435 440 445  
 atc ggg cca act atc acc ggt cca cac tct ccg gat gag caa gtt cac 1392  
 Ile Gly Pro Thr Ile Thr Gly Pro His Ser Pro Asp Glu Gln Val His  
 450 455 460  
 atc gaa agc gta ggt cat tac tgg aca ctg ctg act gaa ctg ctg aaa 1440  
 Ile Glu Ser Val Gly His Tyr Trp Thr Leu Leu Thr Glu Leu Leu Lys  
 465 470 475 480  
 gaa att ccg gcg aag 1455  
 Glu Ile Pro Ala Lys  
 485  
 <210> 13  
 <211> 2610  
 <212> ADN  
 <213> Escherichia coli  
 <400> 13  
 atg act caa cag cca caa gcc aaa tac cgt cac gat tat cgt gcg ccg 48  
 Met Thr Gln Gln Pro Gln Ala Lys Tyr Arg His Asp Tyr Arg Ala Pro  
 1 5 10 15  
 gat tac cag att act gat att gac ttg acc ttt gac ctc gac gcg caa 96  
 Asp Tyr Gln Ile Thr Asp Ile Asp Leu Thr Phe Asp Leu Asp Ala Gln  
 20 25 30  
 aag acg gtc gtt acc gcg gtc agc cag gct gtc cgt cat ggt gca tca 144  
 Lys Thr Val Val Thr Ala Val Ser Gln Ala Val Arg His Gly Ala Ser  
 35 40 45  
 gat gct ccc ctt cgt ctc aac ggc gaa gac ctc aaa ctg gtt tct gtt 192  
 Asp Ala Pro Leu Arg Leu Asn Gly Glu Asp Leu Lys Leu Val Ser Val  
 50 55 60  
 cat att aat gat gag ccg tgg acc gcc tgg aaa gaa gaa gag ggc gca 240  
 His Ile Asn Asp Glu Pro Trp Thr Ala Trp Lys Glu Glu Glu Gly Ala  
 65 70 75 80  
 ctg gtt atc agt aat ttg ccg gag cgt ttt acg ctt aag atc att aat 288  
 Leu Val Ile Ser Asn Leu Pro Glu Arg Phe Thr Leu Lys Ile Ile Asn  
 85 90 95  
 gaa ata agc ccg gcg gcg aat acc gcg ctg gaa ggg ctt tat cag tca 336  
 Glu Ile Ser Pro Ala Ala Asn Thr Ala Leu Glu Gly Leu Tyr Gln Ser  
 100 105 110  
 ggc gat gcg ctt tgc acc cag tgt gaa gcc gaa ggt ttc cgc cat att 384  
 Gly Asp Ala Leu Cys Thr Gln Cys Glu Ala Glu Gly Phe Arg His Ile  
 115 120 125  
 acg tat tat ctc gac cgc ccg gac gtg ctg gcg cgt ttt acc acc aaa 432  
 Thr Tyr Tyr Leu Asp Arg Pro Asp Val Leu Ala Arg Phe Thr Thr Lys  
 130 135 140  
 att att gcc gat aaa atc aaa tat ccc ttc ctg ctt tcc aat ggt aac 480  
 Ile Ile Ala Asp Lys Ile Lys Tyr Pro Phe Leu Leu Ser Asn Gly Asn  
 145 150 155 160

5

ES 2 626 571 T3

cgc gtt gcg caa ggc gaa ctg gaa aac gga cgc cat tgg gta cag tgg	528
Arg Val Ala Gln Gly Glu Leu Glu Asn Gly Arg His Trp Val Gln Trp	
	165
	170
	175
cag gac ccg ttc ccg aaa ccg tgc tac ctg ttt gcg ctg gtg gca ggc	576
Gln Asp Pro Phe Pro Lys Pro Cys Tyr Leu Phe Ala Leu Val Ala Gly	
	180
	185
	190
gac ttt gat gta ctg cgc gat acc ttt acc acg cgt tct ggt cgc gaa	624
Asp Phe Asp Val Leu Arg Asp Thr Phe Thr Thr Arg Ser Gly Arg Glu	
	195
	200
	205
gta gca ctg gag ctg tac gtc gat cgc ggc aac ctt gat cgc gcg ccg	672
Val Ala Leu Glu Leu Tyr Val Asp Arg Gly Asn Leu Asp Arg Ala Pro	
	210
	215
	220
tgg gcg atg acc tcg ctg aaa aac tcc atg aaa tgg gat gaa gaa cgc	720
Trp Ala Met Thr Ser Leu Lys Asn Ser Met Lys Trp Asp Glu Glu Arg	
	225
	230
	235
	240
ttt ggc ctg gag tat gac ctc gac atc tat atg atc gtc gcg gtg gat	768
Phe Gly Leu Glu Tyr Asp Leu Asp Ile Tyr Met Ile Val Ala Val Asp	
	245
	250
	255
ttc ttc aat atg ggc gca atg gag aat aag ggg ctg aat atc ttt aac	816
Phe Phe Asn Met Gly Ala Met Glu Asn Lys Gly Leu Asn Ile Phe Asn	
	260
	265
	270
tcc aaa tat gtg ctg gcc cgc acc gac acc gcc acc gac aaa gat tac	864
Ser Lys Tyr Val Leu Ala Arg Thr Asp Thr Ala Thr Asp Lys Asp Tyr	
	275
	280
	285
ctc gat att gaa cgc gtt atc ggc cat gaa tat ttc cat aac tgg acc	912
Leu Asp Ile Glu Arg Val Ile Gly His Glu Tyr Phe His Asn Trp Thr	
	290
	295
	300
ggt aac cga gtg acc tgt cgc gac tgg ttc cag ctc agc ctg aaa gaa	960
Gly Asn Arg Val Thr Cys Arg Asp Trp Phe Gln Leu Ser Leu Lys Glu	
	305
	310
	315
	320
ggt tta acc gtc ttc cgc gat cag gag ttc agc tct gac ctt ggt tcc	1008
Gly Leu Thr Val Phe Arg Asp Gln Glu Phe Ser Ser Asp Leu Gly Ser	
	325
	330
	335
cgc gca gtt aac cgc atc aat aat gta cgc acc atg cgc gga ttg cag	1056
Arg Ala Val Asn Arg Ile Asn Asn Val Arg Thr Met Arg Gly Leu Gln	
	340
	345
	350
ttt gca gaa gac gcc agc ccg atg gcg cac ccg atc cgc ccg gat atg	1104
Phe Ala Glu Asp Ala Ser Pro Met Ala His Pro Ile Arg Pro Asp Met	
	355
	360
	365
gtc att gag atg aac aac ttc tac acc ctg acc gtt tac gag aag ggc	1152
Val Ile Glu Met Asn Asn Phe Tyr Thr Leu Thr Val Tyr Glu Lys Gly	
	370
	375
	380
gcg gaa gtg att cgc atg atc cac acc ctg ctt ggc gaa gaa aac ttc	1200
Ala Glu Val Ile Arg Met Ile His Thr Leu Leu Gly Glu Glu Asn Phe	
	385
	390
	395
400	
cag aaa ggg atg cag ctt tat ttc gag cgt cat gat ggt agt gca gcg	1248
Gln Lys Gly Met Gln Leu Tyr Phe Glu Arg His Asp Gly Ser Ala Ala	
	405
	410
	415
acc tgt gac gac ttt gtg cag gcg atg gaa gat gcg tcg aat gtc gat	1296
Thr Cys Asp Asp Phe Val Gln Ala Met Glu Asp Ala Ser Asn Val Asp	
	420
	425
	430

ES 2 626 571 T3

ctc tcc cat ttc cgc cgt tgg tac agc cag tcc ggt aca ccg att gtg Leu Ser His Phe Arg Arg Trp Tyr Ser Gln Ser Gly Thr Pro Ile Val 435 440 445	1344
acc gtc aaa gac gac tac aat ccg gaa acc gag cag tac acc ctg acc Thr Val Lys Asp Asp Tyr Asn Pro Glu Thr Glu Gln Tyr Thr Leu Thr 450 455 460	1392
atc agc cag cgc acg cca gcc acg ccg gat cag gca gaa aaa cag ccg Ile Ser Gln Arg Thr Pro Ala Thr Pro Asp Gln Ala Glu Lys Gln Pro 465 470 475 480	1440
ctg cat att ccg ttt gcc atc gaa ctg tat gat aac gaa ggc aaa gtg Leu His Ile Pro Phe Ala Ile Glu Leu Tyr Asp Asn Glu Gly Lys Val 485 490 495	1488
atc ccg ttg cag aaa ggc ggt cat ccg gtg aat tcc gtg ctg aac gtc Ile Pro Leu Gln Lys Gly Gly His Pro Val Asn Ser Val Leu Asn Val 500 505 510	1536
act cag gcg gaa cag acc ttt gtc ttt gat aat gtc tac ttc cag ccg Thr Gln Ala Glu Gln Thr Phe Val Phe Asp Asn Val Tyr Phe Gln Pro 515 520 525	1584
gtg cct gcg ctg ctg tgc gaa ttc tct gcg cca gtg aaa ctg gaa tat Val Pro Ala Leu Leu Cys Glu Phe Ser Ala Pro Val Lys Leu Glu Tyr 530 535 540	1632
aag tgg agc gat cag caa ctg acc ttc ctg atg cgt cat gcg cgt aat Lys Trp Ser Asp Gln Gln Leu Thr Phe Leu Met Arg His Ala Arg Asn 545 550 555 560	1680
gat ttc tcc cgc tgg gat gcg gcg caa agt ttg ctg gca acc tac atc Asp Phe Ser Arg Trp Asp Ala Ala Gln Ser Leu Leu Ala Thr Tyr Ile 565 570 575	1728
aag ctg aac gtc gcg cgt cat cag caa ggt cag ccg ctg tct ctg ccg Lys Leu Asn Val Ala Arg His Gln Gln Gly Gln Pro Leu Ser Leu Pro 580 585 590	1776
gtg cat gtg gct gat gct ttc cgc gcg gta ctg ctt gat gag aag att Val His Val Ala Asp Ala Phe Arg Ala Val Leu Leu Asp Glu Lys Ile 595 600 605	1824
gat cca gcg ctg gcg gca gaa atc ctg acg ctg cct tct gtc aat gaa Asp Pro Ala Leu Ala Ala Glu Ile Leu Thr Leu Pro Ser Val Asn Glu 610 615 620	1872
atg gct gaa ttg ttc gat atc atc gac ccg att gct att gcc gaa gta Met Ala Glu Leu Phe Asp Ile Ile Asp Pro Ile Ala Ile Ala Glu Val 625 630 635 640	1920
cgc gaa gca ctc act cgt act ctg gcg act gaa ctg gcg gat gag cta Arg Glu Ala Leu Thr Arg Thr Leu Ala Thr Glu Leu Ala Asp Glu Leu 645 650 655	1968
ctg gct att tac aac gcg aat tac cag agc gag tac cgt gtt gag cat Leu Ala Ile Tyr Asn Ala Asn Tyr Gln Ser Glu Tyr Arg Val Glu His 660 665 670	2016
gaa gat att gca aaa cgc act ctg cgt aat gcc tgc ctg cgt ttc ctc Glu Asp Ile Ala Lys Arg Thr Leu Arg Asn Ala Cys Leu Arg Phe Leu 675 680 685	2064
gct ttt ggt gaa acg cat ctg gct gat gtg ctg gtg agc aag cag ttc Ala Phe Gly Glu Thr His Leu Ala Asp Val Leu Val Ser Lys Gln Phe 690 695 700	2112

ES 2 626 571 T3

cac gaa gca aac aat atg act gat gcg ctg gcg gcg ctt tct gcg gcg 2160  
 His Glu Ala Asn Asn Met Thr Asp Ala Leu Ala Ala Leu Ser Ala Ala  
 705 710 715 720  
 gtt gcc gca cag ctg cct tgc cgt gac gcg ctg atg cag gag tac gac 2208  
 Val Ala Ala Gln Leu Pro Cys Arg Asp Ala Leu Met Gln Glu Tyr Asp  
 725 730 735  
 gac aag tgg cat cag aac ggt ctg gtg atg gat aaa tgg ttt atc ctg 2256  
 Asp Lys Trp His Gln Asn Gly Leu Val Met Asp Lys Trp Phe Ile Leu  
 740 745 750  
 caa gcc acc agc ccg gcg gcg aat gtg ctg gag acg gtg cgc ggc ctg 2304  
 Gln Ala Thr Ser Pro Ala Ala Asn Val Leu Glu Thr Val Arg Gly Leu  
 755 760 765  
 ttg cag cat cgc tca ttt acc atg agc aac ccg aac cgt att cgt tcg 2352  
 Leu Gln His Arg Ser Phe Thr Met Ser Asn Pro Asn Arg Ile Arg Ser  
 770 775 780  
 ttg att ggc gcg ttt gcg ggc agc aat ccg gca gcg ttc cat gcc gaa 2400  
 Leu Ile Gly Ala Phe Ala Gly Ser Asn Pro Ala Ala Phe His Ala Glu  
 785 790 800  
 gat ggc agc ggt tac ctg ttc ctg gtg gaa atg ctt acc gac ctc aac 2448  
 Asp Gly Ser Gly Tyr 805 Leu Phe Leu Val Glu Met Leu Thr Asp Leu Asn  
 810 815  
 agc cgt aac ccg cag gtg gct tca cgt ctg att gaa ccg ctg att cgc 2496  
 Ser Arg Asn Pro Gln Val Ala Ser Arg 825 Leu Ile Glu Pro Leu Ile Arg  
 830  
 ctg aaa cgt tac gat gcc aaa cgt cag gag aaa atg cgc gcg gcg ctg 2544  
 Leu Lys Arg Tyr Asp Ala Lys Arg Gln Glu Lys Met Arg Ala Ala Leu  
 835 840 845  
 gaa cag ttg aaa ggg ctg gaa aat ctc tct ggc gat ctg tac gag aag 2592  
 Glu Gln Leu Lys Gly Leu Glu Asn Leu Ser Gly Asp Leu Tyr Glu Lys  
 850 855 860  
 ata act aaa gca ctg gct 2610  
 Ile Thr Lys Ala Leu Ala  
 865 870

<210> 14  
 <211> 1605  
 <212> ADN  
 <213> Escherichia coli

5

<400> 14  
 atg cgt att tcc ttg aaa aag tca ggg atg ctg aag ctt ggt ctc agc 48  
 Met Arg Ile Ser Leu Lys Lys Ser Gly Met Leu Lys Leu Gly Leu Ser  
 1 5 10 15  
 ctg gtg gct atg acc gtc gca gca agt gtt cag gct aaa act ctg gtt 96  
 Leu Val Ala Met Thr Val Ala Ala Ser Val Gln Ala Lys Thr Leu Val  
 20 25 30  
 tat tgc tca gaa gga tct ccg gaa ggg ttt aac ccg cag ctg ttt acc 144  
 Tyr Cys Ser Glu Gly Ser Pro Glu Gly Phe Asn Pro Gln Leu Phe Thr  
 35 40 45  
 tcc ggc acc acc tat gac gcc tct tcc gtc ccg ctt tat aac cgt ctg 192  
 Ser Gly Thr Thr Tyr Asp Ala Ser Ser Val Pro Leu Tyr Asn Arg Leu  
 50 55 60  
 gtt gaa ttt aaa atc ggc acc acc gaa gtg atc ccg ggc ctc gct gaa 240

ES 2 626 571 T3

Val 65	Glu	Phe	Lys	Ile	Gly 70	Thr	Thr	Glu	Val	Ile 75	Pro	Gly	Leu	Ala	Glu 80		
aag	tgg	gaa	gtc	agc	gaa	gac	ggt	aaa	acc	tat	acc	ttc	cat	ctg	cgt	288	
Lys	Trp	Glu	Val	Ser 85	Glu	Asp	Gly	Lys	Thr 90	Tyr	Thr	Phe	His	Leu 95	Arg		
aaa	ggt	gtg	aag	tgg	cac	gac	aat	aaa	gaa	ttc	aaa	ccg	acg	cgt	gaa	336	
Lys	Gly	Val	Lys 100	Trp	His	Asp	Asn	Lys 105	Glu	Phe	Lys	Pro	Thr 110	Arg	Glu		
ctg	aac	gcc	gat	gat	gtg	gtg	ttc	tcg	ttc	gat	cgt	cag	aaa	aac	gcg	384	
Leu	Asn	Ala 115	Asp	Asp	Val	Val	Phe 120	Ser	Phe	Asp	Arg	Gln 125	Lys	Asn	Ala		
caa	aac	ccg	tac	cat	aaa	gtt	tct	ggc	ggc	agc	tac	gaa	tac	ttc	gaa	432	
Gln	Asn 130	Pro	Tyr	His	Lys	Val 135	Ser	Gly	Gly	Ser	Tyr 140	Glu	Tyr	Phe	Glu		
ggc	atg	ggc	ttg	cca	gag	ctg	atc	agt	gaa	gtg	aaa	aag	gtg	gac	gac	480	
Gly	Met	Gly	Leu	Pro	Glu 150	Leu	Ile	Ser	Glu	Val 155	Lys	Lys	Val	Asp	Asp 160		
aac	acc	gtt	cag	ttt	gtg	ctg	act	cgc	ccg	gaa	gcg	ccg	ttc	ctc	gct	528	
Asn	Thr	Val	Gln 165	Phe	Val	Leu	Thr	Arg	Pro 170	Glu	Ala	Pro	Phe	Leu 175	Ala		
gac	ctg	gca	atg	gac	ttc	gcc	tct	att	ctg	tca	aaa	gaa	tat	gct	gat	576	
Asp	Leu	Ala 180	Met	Asp	Phe	Ala	Ser	Ile 185	Leu	Ser	Lys	Glu	Tyr 190	Ala	Asp		
gcg	atg	atg	aaa	gcc	ggt	aca	ccg	gaa	aaa	ctg	gac	ctc	aac	cca	atc	624	
Ala	Met	Met 195	Lys	Ala	Gly	Thr	Pro 200	Glu	Lys	Leu	Asp	Leu 205	Asn	Pro	Ile		
gga	acc	ggt	ccg	ttc	cag	tta	cag	cag	tat	caa	aaa	gat	tcc	cgt	atc	672	
Gly	Thr 210	Gly	Pro	Phe	Gln	Leu 215	Gln	Gln	Tyr	Gln	Lys 220	Asp	Ser	Arg	Ile		
cgc	tac	aaa	gcg	ttt	gat	ggc	tac	tgg	ggc	acc	aaa	ccg	cag	atc	gat	720	
Arg	Tyr	Lys	Ala	Phe	Asp 230	Gly	Tyr	Trp	Gly	Thr 235	Lys	Pro	Gln	Ile	Asp 240		
acg	ctg	gtt	ttc	tct	att	acc	cct	gac	gct	tcc	gtg	cgt	tac	gcg	aaa	768	
Thr	Leu	Val	Phe 245	Ser	Ile	Thr	Pro	Asp	Ala 250	Ser	Val	Arg	Tyr	Ala 255	Lys		
ttg	cag	aag	aat	gaa	tgc	cag	gtg	atg	ccg	tac	ccg	aac	ccg	gca	gat	816	
Leu	Gln	Lys	Asn 260	Glu	Cys	Gln	Val	Met 265	Pro	Tyr	Pro	Asn	Pro 270	Ala	Asp		
atc	gct	cgc	atg	aag	cag	gat	aaa	tcc	atc	aat	ctg	atg	gaa	atg	ccg	864	
Ile	Ala	Arg 275	Met	Lys	Gln	Asp	Lys 280	Ser	Ile	Asn	Leu	Met 285	Glu	Met	Pro		
ggg	ctg	aac	gtc	ggt	tat	ctc	tcg	tat	aac	gtg	cag	aaa	aaa	cca	ctc	912	
Gly	Leu 290	Asn	Val	Gly	Tyr	Leu 295	Ser	Tyr	Asn	Val 300	Gln	Lys	Lys	Pro	Leu		
gat	gac	gtg	aaa	ggt	cgc	cag	gct	ctg	acc	tac	gcg	gtg	aac	aaa	gac	960	
Asp	Asp	Val	Lys	Val	Arg 310	Gln	Ala	Leu	Thr	Tyr 315	Ala	Val	Asn	Lys	Asp 320		
gcg	atc	atc	aaa	gcg	ggt	tat	cag	ggc	gcg	ggc	gta	tca	gcg	aaa	aac	1008	
Ala	Ile	Ile	Lys 325	Ala	Val	Tyr	Gln	Gly	Ala 330	Gly	Val	Ser	Ala	Lys 335	Asn		
ctg	atc	ccg	cca	acc	atg	tgg	ggc	tat	aac	gac	gac	ggt	cag	gac	tac	1056	

ES 2 626 571 T3

Leu	Ile	Pro	Pro 340	Thr	Met	Trp	Gly	Tyr 345	Asn	Asp	Asp	Val	Gln 350	Asp	Tyr		
acc	tac	gat	cct	gaa	aaa	gcg	aaa	gcc	ttg	ctg	aaa	gaa	gcg	ggt	ctg	1104	
Thr	Tyr	Asp 355	Pro	Glu	Lys	Ala	Lys 360	Ala	Leu	Leu	Lys	Glu 365	Ala	Gly	Leu		
gaa	aaa	ggt	ttc	tcc	atc	gac	ctg	tgg	gcg	atg	ccg	gta	caa	cgt	ccg	1152	
Glu	Lys 370	Gly	Phe	Ser	Ile	Asp 375	Leu	Trp	Ala	Met	Pro 380	Val	Gln	Arg	Pro		
tat	aac	ccg	aac	gct	cgc	cgc	atg	gcg	gag	atg	att	cag	gca	gac	tgg	1200	
Tyr	Asn	Pro	Asn	Ala	Arg 390	Arg	Met	Ala	Glu	Met 395	Ile	Gln	Ala	Asp	Trp 400		
gcg	aaa	gtc	ggc	gtg	cag	gcc	aaa	att	gtc	acc	tac	gaa	tgg	ggt	gag	1248	
Ala	Lys	Val	Gly	Val 405	Gln	Ala	Lys	Ile	Val 410	Thr	Tyr	Glu	Trp	Gly 415	Glu		
tac	ctc	aag	cgt	gcg	aaa	gat	ggc	gag	cac	cag	acg	gta	atg	atg	ggc	1296	
Tyr	Leu	Lys	Arg 420	Ala	Lys	Asp	Gly	Glu 425	His	Gln	Thr	Val	Met 430	Met	Gly		
tgg	act	ggc	gat	aac	ggg	gat	ccg	gat	aac	ttc	ttc	gcc	acc	ctg	ttc	1344	
Trp	Thr	Gly 435	Asp	Asn	Gly	Asp	Pro 440	Asp	Asn	Phe	Phe	Ala 445	Thr	Leu	Phe		
agc	tgc	gcc	gcc	tct	gaa	caa	ggc	tcc	aac	tac	tca	aaa	tgg	tgc	tac	1392	
Ser	Cys 450	Ala	Ala	Ser	Glu	Gln 455	Gly	Ser	Asn	Tyr	Ser 460	Lys	Trp	Cys	Tyr		
aaa	ccg	ttt	gaa	gat	ctg	att	caa	ccg	gcg	cgt	gct	acc	gac	gac	cac	1440	
Lys	Pro	Phe	Glu	Asp	Leu 470	Ile	Gln	Pro	Ala	Arg 475	Ala	Thr	Asp	Asp	His 480		
aat	aaa	cgc	ggt	gaa	ctg	tac	aaa	caa	gcg	cag	gtg	gtg	atg	cac	gat	1488	
Asn	Lys	Arg	Val 485	Glu	Leu	Tyr	Lys	Gln	Ala 490	Gln	Val	Val	Met	His 495	Asp		
cag	gct	ccg	gca	ctg	atc	atc	gct	cac	tcc	acc	gtg	ttt	gaa	ccg	gta	1536	
Gln	Ala	Pro	Ala 500	Leu	Ile	Ile	Ala	His 505	Ser	Thr	Val	Phe	Glu 510	Pro	Val		
cgt	aaa	gaa	ggt	aaa	ggc	tat	gtg	ggt	gat	cca	tta	ggc	aaa	cat	cac	1584	
Arg	Lys	Glu 515	Val	Lys	Gly	Tyr	Val 520	Val	Asp	Pro	Leu	Gly 525	Lys	His	His		
ttc	gaa	aac	gtc	tct	atc	gaa										1605	
Phe	Glu 530	Asn	Val	Ser	Ile	Glu 535											

<210> 15  
 <211> 1017  
 <212> ADN  
 <213> Escherichia coli

5

<400> 15																	
atg	ttg	cag	ttt	att	ctc	cga	cgt	ttg	gga	ctc	gtc	atc	ccc	acg	ttt	48	
Met	Leu	Gln	Phe	Ile 5	Leu	Arg	Arg	Leu	Gly 10	Leu	Val	Ile	Pro	Thr 15	Phe		
1																	
atc	ggt	att	acc	ctt	ctc	aca	ttt	gcc	ttt	gtc	cac	atg	atc	ccg	ggc	96	
Ile	Gly	Ile	Thr 20	Leu	Leu	Thr	Phe	Ala 25	Phe	Val	His	Met	Ile 30	Pro	Gly		
gat	ccg	gtg	atg	atc	atg	gcg	ggc	gaa	cgt	ggg	atc	tcc	cca	gag	cgt	144	
Asp	Pro	Val	Met	Ile	Met	Ala	Gly	Glu	Arg	Gly	Ile	Ser	Pro	Glu	Arg		

ES 2 626 571 T3

35			40			45										
cac His	gcg Ala 50	cag Gln	ctg Leu	ctg Leu	gct Ala	gaa Glu 55	ctc Leu	ggc Gly	tta Leu	gat Asp	aaa Lys 60	ccg Pro	atg Met	tgg Trp	cag Gln	192
cag Gln 65	tat Tyr	ctc Leu	cat His	tac Tyr	att Ile 70	tgg Trp	ggc Gly	gtt Val	atg Met	cat His 75	ggc Gly	gat Asp	cta Leu	ggc Gly	att Ile 80	240
tca Ser	atg Met	aaa Lys	agc Ser	cgc Arg 85	atc Ile	ccg Pro	gtt Val	tgg Trp	gaa Glu 90	gag Glu	ttc Phe	gtg Val	ccg Pro	cgc Arg 95	ttc Phe	288
cag Gln	gcc Ala	acg Thr 100	ctg Leu	gaa Glu	ctt Leu	ggc Gly	gtc Val	tgc Cys 105	gcg Ala	atg Met	att Ile	ttt Phe	gct Ala 110	acg Thr	gca Ala	336
gtc Val	ggg Gly	att Ile 115	ccg Pro	gtc Val	ggc Gly	gtg Val	ctg Leu 120	gct Ala	gcg Ala	gtt Val	aaa Lys	cgc Arg 125	ggg Gly	tcc Ser	att Ile	384
ttc Phe 130	gat Asp	cac His	aca Thr	gcg Ala	ggt Val	ggc Gly 135	ctg Leu	gcg Ala	ctg Leu	aca Thr	ggg Gly 140	tat Tyr	tca Ser	atg Met	cct Pro	432
atc Ile 145	ttc Phe	tgg Trp	tgg Trp	ggc Gly	atg Met 150	atg Met	ctg Leu	atc Ile	atg Met	ctg Leu 155	ggt Val	tcg Ser	gtg Val	cac His	tgg Trp 160	480
aac Asn	ctg Leu	acg Thr	ccc Pro	gtc Val 165	tcc Ser	ggg Gly	cgc Arg	gtg Val	agc Ser 170	gat Asp	atg Met	gtg Val	ttc Phe	ctc Leu 175	gat Asp	528
gac Asp	tcc Ser	aat Asn	ccg Pro	tta Leu	acc Thr	ggg Gly	ttt Phe	atg Met 185	cta Leu	atc Ile	gac Asp	acc Thr	gcc Ala 190	atc Ile	tgg Trp	576
ggg Gly	gaa Glu	gac Asp 195	ggc Gly	aac Asn	ttt Phe	atc Ile	gat Asp 200	gcc Ala	gtc Val	gcc Ala	cat His	atg Met 205	atc Ile	ttg Leu	cct Pro	624
gcc Ala	att Ile 210	gtg Val	ctg Leu	ggg Gly	act Thr	att Ile 215	ccg Pro	ctg Leu	gcg Ala	gtc Val	att Ile 220	gtg Val	cgt Arg	atg Met	aca Thr	672
cgc Arg 225	tcc Ser	tcg Ser	atg Met	ctg Leu	gaa Glu 230	gtg Val	ctg Leu	ggc Gly	gag Glu	gat Asp 235	tac Tyr	atc Ile	cgc Arg	acc Thr	gcg Ala 240	720
cgc Arg	gcc Ala	aaa Lys	ggg Gly	cta Leu 245	acc Thr	cgc Arg	atg Met	cgg Arg	gtg Val 250	att Ile	atc Ile	gtc Val	cat His	gcg Ala 255	ctg Leu	768
cgt Arg	aac Asn	gcg Ala	atg Met	ctg Leu	ccg Pro	gtg Val	gtg Val	acc Thr 265	gtt Val	atc Ile	ggc Gly	ctg Leu	cag Gln 270	gtg Val	gga Gly	816
aca Thr	ttg Leu	ctg Leu 275	gcg Ala	ggg Gly	gcg Ala	att Ile 280	ctg Leu	acc Thr	gaa Glu	acc Thr	atc Ile	ttc Phe 285	tcg Ser	tgg Trp	ccc Pro	864
ggg Gly	ctg Leu 290	gga Gly	cgc Arg	tgg Trp	ttg Leu	att Ile 295	gac Asp	gca Ala	ctg Leu	caa Gln	cgc Arg 300	cgc Arg	gac Asp	tat Tyr	ccg Pro	912
gta Val 305	gtg Val	cag Gln	ggc Gly	ggc Gly	gta Val 310	ttg Leu	ctg Leu	gtg Val	gcg Ala	acg Thr	atg Met	att Ile	atc Ile	ctc Leu	gtc Val 320	960
aac Asn	ttg Leu	ctg Leu	gtc Val	gat Asp 325	ctg Leu	ctg Leu	tac Tyr	ggc Gly	gtg Val 330	gtg Val	aac Asn	ccg Pro	cgt Arg	att Ile 335	cgt Arg	1008
cat His	aag Lys	aag Lys														1017

ES 2 626 571 T3

<210> 16  
 <211> 900  
 <212> ADN  
 <213> Escherichia coli

5

<400> 16  
 atg tca cag gtt act gaa aat aaa gtg att agc gca ccg gtg ccg atg 48  
 Met Ser Gln Val Thr Glu Asn Lys Val Ile Ser Ala Pro Val Pro Met  
 1 5 10 15  
 acc ccg tta cag gag ttc tgg cac tat ttt aaa cgc aac aaa ggc gcg 96  
 Thr Pro Leu Gln Glu Phe Trp His Tyr Phe Lys Arg Asn Lys Gly Ala  
 20 25 30  
 gtc gtc ggg ctg gtt tac gtc gtc atc gtg ctg ttc atc gcg atc ttt 144  
 Val Val Gly Leu Val Tyr Val Val Ile Val Leu Phe Ile Ala Ile Phe  
 35 40 45  
 gcc aac tgg att gca ccc tat aac ccg gcg gaa cag ttc cgc gat gca 192  
 Ala Asn Trp Ile Ala Pro Tyr Asn Pro Ala Glu Gln Phe Arg Asp Ala  
 50 55 60  
 ctg ctc gcc ccg cca gcc tgg cag gaa ggc ggc agc atg gcg cac ttg 240  
 Leu Leu Ala Pro Pro Ala Trp Gln Glu Gly Gly Ser Met Ala His Leu  
 65 70 75 80  
 ctg ggc acc gat gac gta ggc cgt gat gtg ctg tcg cgc ctg atg tac 288  
 Leu Gly Thr Asp Asp Val Gly Arg Asp Val Leu Ser Arg Leu Met Tyr  
 85 90 95  
 ggt gcg cgc ctg tcg ctg ctg gtt ggc tgt ctg gta gtt gtg tta tcg 336  
 Gly Ala Arg Leu Ser Leu Leu Val Gly Cys Leu Val Val Val Leu Ser  
 100 105 110  
 ctg att atg ggc gtt att ctc ggc ctg atc gcc ggt tac ttt ggc ggc 384  
 Leu Ile Met Gly Val Ile Leu Gly Leu Ile Ala Gly Tyr Phe Gly Gly  
 115 120 125  
 ctg gtc gat aac atc att atg cgc gtg gtc gat atc atg ctg gcg ctg 432  
 Leu Val Asp Asn Ile Ile Met Arg Val Val Asp Ile Met Leu Ala Leu  
 130 135 140  
 cca agt ctg ctg ctg gcg ctg gtg ctg gtg gca att ttc ggc ccg tcg 480  
 Pro Ser Leu Leu Leu Ala Leu Val Leu Val Ala Ile Phe Gly Pro Ser  
 145 150 155 160  
 att ggt aac gcc gcg ctg gca ctg acc ttc gtt gcc ttg ccg cac tat 528  
 Ile Gly Asn Ala Ala Leu Ala Leu Thr Phe Val Ala Leu Pro His Tyr  
 165 170 175  
 gtg gcg tta acc ccg gcc gcc gtg ctg gtg gaa gtt aac cgc gat tac 576  
 Val Arg Leu Thr Arg Ala Ala Val Leu Val Glu Val Asn Arg Asp Tyr  
 180 185 190  
 gtc acc gcg tct ccg gtg gcg ggt gcc ggc ggc atg cgt cag atg ttt 624  
 Val Thr Ala Ser Arg Val Ala Gly Ala Gly Ala Met Arg Gln Met Phe  
 195 200 205

ES 2 626 571 T3

att aac atc ttc ccg aac tgc ctt gcg ccg ctg att gtt cag gcg tcg Ile Asn Ile Phe Pro Asn Cys Leu Ala Pro Leu Ile Val Gln Ala Ser	672
ctc ggt ttc tct aac gcc att ctc gat atg gct gct ctt ggt ttc ctc Leu Gly Phe Ser Asn Ala Ile Leu Asp Met Ala Ala Leu Gly Phe Leu	720
ggc atg ggg gca cag ccg cca acg cct gag tgg ggc acc atg ctc tcc Gly Met Gly Ala Gln Pro Pro Thr Pro Glu Trp Gly Thr Met Leu Ser	768
gac gtg ttg cag ttc gcg caa agc gcc tgg tgg gtc gtg acc ttc ccg Asp Val Leu Gln Phe Ala Gln Ser Ala Trp Trp Val Val Thr Phe Pro	816
ggt ctg gcg atc ctg ctg acg gtg ctg gca ttt aac ctg atg ggt gac Gly Leu Ala Ile Leu Leu Thr Val Leu Ala Phe Asn Leu Met Gly Asp	864
ggt ctg cgt gac gcg ctc gat ccc aaa ctg aag cag Gly Leu Arg Asp Ala Leu Asp Pro Lys Leu Lys Gln	900
<210> 17	
<211> 981	
<212> ADN	
<213> Escherichia coli	
<400> 17	
atg gcg tta tta aat gta gat aaa tta tcg gtg cat ttc ggc gac gaa Met Ala Leu Leu Asn Val Asp Lys Leu Ser Val His Phe Gly Asp Glu	48
agc gcg ccg ttc cgc gcc gta gac cgc atc agc tac agc gta aaa cag Ser Ala Pro Phe Arg Ala Val Asp Arg Ile Ser Tyr Ser Val Lys Gln	96
ggc gaa gtg gtc ggg att gtg ggt gag tcc ggc tcc ggt aag tcg gtc Gly Glu Val Val Gly Ile Val Gly Glu Ser Gly Ser Gly Lys Ser Val	144
agt tca ctg gcg att atg ggg ctg att gat tat ccg ggc cgc gta atg Ser Ser Leu Ala Ile Met Gly Leu Ile Asp Tyr Pro Gly Arg Val Met	192
gca gaa aaa ctg gag ttt aac gcc cag gat ttg cag cgt atc tca gaa Ala Glu Lys Leu Glu Phe Asn Gly Gln Asp Leu Gln Arg Ile Ser Glu	240
aaa gag cgc cgc aac ctg gtg ggt gcc gaa gtg gcg atg atc ttc cag Lys Glu Arg Arg Asn Leu Val Gly Ala Glu Val Ala Met Ile Phe Gln	288
gac ccg atg acc agc ctt aac ccg tgc tac acc gtg ggt ttc cag att Asp Pro Met Thr Ser Leu Asn Pro Cys Tyr Thr Val Gly Phe Gln Ile	336
atg gaa gcg att aag gtg cat cag gcc gcc aac aaa agt acc cgc cgt Met Glu Ala Ile Lys Val His Gln Gly Gly Asn Lys Ser Thr Arg Arg	384
cag cga gcg att gac ctg ctg aat cag gtc ggt att ccc gat ccg gca Gln Arg Ala Ile Asp Leu Leu Asn Gln Val Gly Ile Pro Asp Pro Ala	432
tcg cgt ctg gat gtt tac ccg cat cag ctt tcc ggc ggc atg agc cag	480

5

ES 2 626 571 T3

Ser 145	Arg	Leu	Asp	Val	Tyr 150	Pro	His	Gln	Leu	Ser 155	Gly	Gly	Met	Ser	Gln 160	
cgc	gtg	atg	atc	gcc	atg	gcg	att	gcc	tgt	cgg	cca	aaa	ctg	ctg	att	528
Arg	Val	Met	Ile	Ala 165	Met	Ala	Ile	Ala	Cys 170	Arg	Pro	Lys	Leu	Leu 175	Ile	
gcc	gat	gaa	ccg	acc	acc	gcg	ctg	gac	gtg	acc	att	cag	gcg	caa	atc	576
Ala	Asp	Glu	Pro 180	Thr	Thr	Ala	Leu	Asp 185	Val	Thr	Ile	Gln	Ala 190	Gln	Ile	
atc	gaa	cta	ctg	ctg	gag	cta	cag	cag	aaa	gag	aac	atg	gcg	ctg	gtg	624
Ile	Glu	Leu 195	Leu	Leu	Glu	Leu	Gln 200	Gln	Lys	Glu	Asn	Met 205	Ala	Leu	Val	
tta	att	acc	cat	gac	ctg	gcg	ctg	gtg	gcg	gaa	gcg	gca	cat	aaa	atc	672
Leu	Ile 210	Thr	His	Asp	Leu	Ala 215	Leu	Val	Ala	Glu	Ala 220	Ala	His	Lys	Ile	
atc	gtg	atg	tat	gca	ggc	cag	gtg	gtg	gaa	acc	ggg	gat	gcg	cac	gcc	720
Ile	Val	Met	Tyr	Ala	Gly 230	Gln	Val	Val	Glu	Thr 235	Gly	Asp	Ala	His	Ala 240	
atc	ttc	cat	gcg	ccg	cgt	cac	ccg	tat	act	cag	gca	ttg	ctg	cgt	gcg	768
Ile	Phe	His	Ala	Pro 245	Arg	His	Pro	Tyr	Thr 250	Gln	Ala	Leu	Leu	Arg 255	Ala	
ctg	cca	gaa	ttt	gct	cag	gac	aaa	gaa	cgt	ctg	gcg	tcg	ttg	cca	ggt	816
Leu	Pro	Glu	Phe 260	Ala	Gln	Asp	Lys	Glu 265	Arg	Leu	Ala	Ser	Leu 270	Pro	Gly	
gtc	gtt	ccc	ggc	aag	tac	gac	cgc	ccg	aac	ggc	tgc	ctg	ctt	aac	ccg	864
Val	Val	Pro 275	Gly	Lys	Tyr	Asp	Arg 280	Pro	Asn	Gly	Cys	Leu 285	Leu	Asn	Pro	
cgc	tgc	ccc	tat	gcc	act	gac	aga	tgt	cgc	gct	gaa	gaa	ccg	gcg	ctg	912
Arg	Cys 290	Pro	Tyr	Ala	Thr	Asp 295	Arg	Cys	Arg	Ala	Glu 300	Glu	Pro	Ala	Leu	
aat	atg	ctc	gct	gac	ggg	cgt	cag	tcc	aaa	tgc	cat	tac	cca	ctt	gat	960
Asn	Met	Leu	Ala	Asp	Gly 310	Arg	Gln	Ser	Lys	Cys 315	His	Tyr	Pro	Leu	Asp 320	
gat	gcc	ggg	agg	ccg	aca	cta										981
Asp	Ala	Gly	Arg	Pro 325	Thr	Leu										

<210> 18  
 <211> 1002  
 <212> ADN  
 <213> Escherichia coli

5

<400> 18	atg	agt	acg	caa	gag	gcc	acc	ctg	caa	caa	ccg	ctg	ttg	cag	gct	atc	48
	Met	Ser	Thr	Gln	Glu	Ala	Thr	Leu	Gln	Gln	Pro	Leu	Leu	Gln	Ala	Ile	
	1				5					10					15		
	gac	ctg	aaa	aaa	cat	tat	ccg	gtg	aag	aaa	ggc	atg	ttc	gcg	ccg	gaa	96
	Asp	Leu	Lys	Lys	His	Tyr	Pro	Val	Lys	Lys	Gly	Met	Phe	Ala	Pro	Glu	
			20						25					30			
	cgt	ctg	gtt	aaa	gcg	ctg	gat	ggc	gtt	tcg	ttt	aac	ctt	gaa	cgt	ggc	144
	Arg	Leu	Val	Lys	Ala	Leu	Asp	Gly	Val	Ser	Phe	Asn	Leu	Glu	Arg	Gly	
			35					40					45				
	aaa	acg	ctg	gca	gta	gtg	ggc	gaa	tct	ggc	tgc	ggg	aaa	tcg	acc	ctc	192
	Lys	Thr	Leu	Ala	Val	Val	Gly	Glu	Ser	Gly	Cys	Gly	Lys	Ser	Thr	Leu	

ES 2 626 571 T3

50	55	60														
ggt Gly 65	cgg Arg	ttg Leu	ctg Leu	acg Thr	atg Met 70	att Ile	gaa Glu	atg Met	ccc Pro	acc Thr 75	ggt Gly	ggc Gly	gag Glu	ctg Leu	tat Tyr 80	240
tac Tyr	cag Gln	ggg Gly	cag Gln	gat Asp 85	ctg Leu	ctt Leu	aag Lys	cac His	gat Asp 90	ccg Pro	cag Gln	gcg Ala	cag Gln	aag Lys 95	ctg Leu	288
cgt Arg	cgg Arg	cag Gln	aaa Lys 100	atc Ile	cag Gln	atc Ile	gtc Val	ttc Phe 105	cag Gln	aac Asn	cct Pro	tac Tyr	ggt Gly 110	tcg Ser	ctg Leu	336
aat Asn	ccg Pro	cgt Arg 115	aaa Lys	aaa Lys	gtc Val	ggg Gly	caa Gln 120	att Ile	ctt Leu	gaa Glu	gag Glu	ccg Pro 125	ctg Leu	ctg Leu	atc Ile	384
aac Asn	acc Thr 130	agc Ser	tta Leu	agc Ser	aaa Lys	gaa Glu 135	cag Gln	cgt Arg	cgg Arg	gaa Glu	aaa Lys 140	gcc Ala	ctg Leu	tcg Ser	atg Met	432
atg Met 145	gcg Ala	aaa Lys	gtc Val	ggc Gly	ctg Leu 150	aaa Lys	acc Thr	gag Glu	cac His	tat Tyr 155	gac Asp	cgc Arg	tat Tyr	ccg Pro	cat His 160	480
atg Met	ttc Phe	tcc Ser	ggc Gly	ggt Gly 165	cag Gln	cgt Arg	cag Gln	cgt Arg	atc Ile 170	gcc Ala	atc Ile	gcc Ala	cgt Arg	ggt Gly 175	ctg Leu	528
atg Met	ctc Leu	gac Asp	ccg Pro 180	gat Asp	gtg Val	gtg Val	att Ile	gcc Ala 185	gat Asp	gaa Glu	ccg Pro	gtt Val	tcc Ser 190	gcg Ala	ctg Leu	576
gat Asp	gtt Val	tca Ser 195	gtg Val	cgc Arg	gcg Ala	cag Gln	gtg Val 200	ctg Leu	aat Asn	ctg Leu	atg Met	atg Met	gat Asp	ttg Leu	cag Gln	624
cag Gln	gag Glu 210	ttg Leu	ggg Gly	ctg Leu	tct Ser	tat Tyr 215	gtc Val	ttt Phe	atc Ile	tcc Ser	cac His 220	gac Asp	ctg Leu	tcg Ser	gtg Val	672
gtg Val 225	gag Glu	cac His	att Ile	gct Ala	gat Asp 230	gaa Glu	gtg Val	atg Met	gtg Val	atg Met 235	tac Tyr	ctg Leu	ggc Gly	cgc Arg	tgc Cys 240	720
gtg Val	gag Glu	aag Lys	gga Gly	acg Thr 245	aaa Lys	gac Asp	caa Gln	atc Ile	ttc Phe 250	aat Asn	aac Asn	ccg Pro	cgc Arg	cat His 255	ccg Pro	768
tac Tyr	act Thr	cag Gln	gcg Ala 260	cta Leu	ctt Leu	tcc Ser	gcg Ala	acg Thr 265	ccg Pro	cgc Arg	ctg Leu	aac Asn	ccg Pro 270	gac Asp	gat Asp	816
cgc Arg	cgc Arg	gag Glu 275	cgc Arg	atc Ile	aag Lys	ctc Leu	agc Ser 280	ggt Gly	gaa Glu	cta Leu	cca Pro	agc Ser 285	cca Pro	ctg Leu	aat Asn	864
cca Pro	ccg Pro 290	ccg Pro	ggt Gly	tgc Cys	gcc Ala	ttc Phe 295	aac Asn	gcc Ala	cgc Arg	tgt Cys	cgt Arg 300	cgg Arg	cgc Arg	ttc Phe	ggc Gly	912
ccc Pro 305	tgc Cys	acc Thr	cag Gln	ttg Leu	cag Gln 310	ccg Pro	cag Gln	cta Leu	aaa Lys	gac Asp 315	tac Tyr	ggc Gly	ggt Gly	caa Gln	ctg Leu 320	960
gta Val	gct Ala	tgt Cys	ttt Phe	gct Ala	ggt Val	gat Asp	cag Gln	gat Asp	gaa Glu	aat Asn	ccg Pro	cag Gln	cgt Arg			1002

<210> 19

<211> 472

<212> PRT

<213> Bacillus subtilis 168

5

ES 2 626 571 T3

<400> 19

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

ES 2 626 571 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 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> 20

<211> 472

5 <212> PRT

<213> Bacillus subtilis ATCC6633

<400> 20

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

ES 2 626 571 T3

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

<210> 21  
 <211> 472  
 <212> PRT  
 <213> Bacillus subtilis IAM1213

5

<400> 21  
 Met Glu Arg Lys Thr Val Leu Val Ile Ala Asp Leu Gly Gly Cys Pro

ES 2 626 571 T3

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	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	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	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	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

ES 2 626 571 T3

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> 22

<211> 472

<212> PRT

<213> Bacillus subtilis IAM1107

5

<400> 22

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  
 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



ES 2 626 571 T3

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  
 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

ES 2 626 571 T3

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> 24  
 <211> 472  
 <212> PRT  
 <213> Bacillus subtilis ATCC21555

5

<400> 24  
 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  
 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

ES 2 626 571 T3

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  
 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> 25  
 <211> 472  
 <212> PRT  
 <213> Bacillus amyloliquefaciens IFO3022

5

<400> 25  
 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

ES 2 626 571 T3

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 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  
 165 170 175  
 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

- <210> 26
- <211> 1416
- <212> ADN
- <213> Bacillus subtilis 168

ES 2 626 571 T3

<400> 26  
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 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 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 gcc 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 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

ES 2 626 571 T3

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 tcg 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 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> 27  
 <211> 1416  
 <212> ADN  
 <213> Bacillus subtilis ATCC6633

5

<400> 27  
 atg gag aga aaa aca gta ttg gtc atc gct gat ctt gga ggc tgc ccg 48

ES 2 626 571 T3

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

ES 2 626 571 T3

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 Val Phe Gly Leu Asp Met Ala Gln Leu Leu Leu Asp  
 340 345 350

ggt 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> 28  
 <211> 1416  
 <212> ADN  
 <213> Bacillus subtilis IAM1213

5

<400> 28

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

ES 2 626 571 T3

35					40					45						
gaa Glu	aaa Lys 50	tac Tyr	tcg Ser	gtc Val	gcg Ala	gtc Val 55	ata Ile	aaa Lys	gat Asp	aaa Lys	gac Asp 60	tat Tyr	ttt Phe	aag Lys	agt Ser	192
tta Leu 65	gct Ala	gat Asp	ttt Phe	gag Glu	cat His 70	cct Pro	gac Asp	tcc Ser	att Ile	tat Tyr 75	tgg Trp	gcg Ala	cat His	gag Glu	gat Asp 80	240
cat His	aac Asn	aag Lys	cct Pro	gag Glu 85	gaa Glu	gag Glu	gtc Val	gtc Val	gag Glu 90	caa Gln	atc Ile	gtc Val	aag Lys	gtt Val 95	gcc Ala	288
gaa Glu	atg Met	ttt Phe	ggg Gly 100	gcg Ala	gat Asp	gcc Ala	atc Ile	aca Thr 105	aca Thr	aac Asn	aat Asn	gaa Glu	tta Leu 110	ttc Phe	att Ile	336
gct Ala	ccg Pro	atg Met 115	gcg Ala	aaa Lys	gcc Ala	tgt Cys	gaa Glu 120	cgt Arg	ctg Leu	ggc Gly	ctg Leu	aga Arg 125	ggt Gly	gcc Ala	ggc Gly	384
gtg Val 130	cag Gln	gca Ala	gcc Ala	gaa Glu	aat Asn	gcc Ala 135	aga Arg	gat Asp	aaa Lys	aat Asn	aaa Lys 140	atg Met	agg Arg	gac Asp	gct Ala	432
ttt Phe 145	aat Asn	aag Lys	gcc Ala	gga Gly	gtc Val 150	aaa Lys	tcg Ser	atc Ile	aaa Lys	aac Asn 155	aaa Lys	cga Arg	gtc Val	aca Thr	act Thr 160	480
ctc Leu	gaa Glu	gat Asp	ttc Phe	cgt Arg 165	gct Ala	gct Ala	ctt Leu	gaa Glu	gag Glu 170	atc Ile	ggc Gly	aca Thr	cct Pro	ctt Leu 175	atc Ile	528
tta Leu	aag Lys	cct Pro	aca Thr 180	tac Tyr	tta Leu	gcg Ala	agt Ser	tca Ser 185	atc Ile	ggt Gly	gta Val	acg Thr	ctg Leu 190	att Ile	acg Thr	576
gac Asp	act Thr	gag Glu 195	acg Thr	gca Ala	gaa Glu	gat Asp	gaa Glu 200	ttt Phe	aac Asn	aga Arg	gtc Val	aat Asn 205	gac Asp	tat Tyr	ctg Leu	624
aaa Lys	tca Ser 210	att Ile	aac Asn	gtg Val	cca Pro	aag Lys 215	gcg Ala	ggt Val	acg Thr	ttt Phe	gaa Glu 220	gcg Ala	ccg Pro	ttt Phe	atc Ile	672
gct Ala 225	gaa Glu	gaa Glu	ttt Phe	tta Leu	cag Gln 230	ggt Gly	gag Glu	tac Tyr	gga Gly	gac Asp 235	tgg Trp	tat Tyr	caa Gln	aca Thr	gaa Glu 240	720
ggg Gly	tac Tyr	tcc Ser	gac Asp 245	tat Tyr	atc Ile	agt Ser	ata Ile	gaa Glu	ggc Gly 250	atc Ile	atg Met	gct Ala	gac Asp	ggt Gly 255	gag Glu	768
tat Tyr	ttc Phe	ccg Pro	atc Ile 260	gcc Ala	att Ile	cat His	gat Asp	aaa Lys 265	acg Thr	ccg Pro	caa Gln	atc Ile	ggg Gly 270	ttt Phe	aca Thr	816
gag Glu	aca Thr	tcc Ser 275	cac His	att Ile	acg Thr	ccg Pro	tcc Ser 280	att Ile	ctg Leu	gat Asp	gaa Glu	gag Glu 285	gca Ala	aaa Lys	aag Lys	864
aaa Lys	att Ile 290	gtc Val	gaa Glu	gct Ala	gcc Ala	aaa Lys 295	aag Lys	gca Ala	aat Asn	gaa Glu	ggt Gly 300	ctt Leu	ggc Gly	ctg Leu	caa Gln	912
aat Asn	tgc Cys	gca Ala	aca Thr	cat His	aca Thr	gag Glu	atc Ile	aag Lys	cta Leu	atg Met	aaa Lys	aat Asn	aga Arg	gaa Glu	ccg Pro	960

ES 2 626 571 T3

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 325		Ala Ala Arg Phe Ala Gly Trp Asn Met Ile 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
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		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> 29  
 <211> 1416  
 5 <212> ADN  
 <213> Bacillus subtilis IAM1107

<400> 29																
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																
				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								

ES 2 626 571 T3

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	
			100					105					110			
gct	ccg	atg	gcg	aaa	gcc	tgt	gaa	cg	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	cg	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	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	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	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	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	gtc	aag	cta	atg	aaa	aac	aga	gaa	ccg	960
Asn	Cys	Ala	Thr	His	Thr	Glu	Val	Lys	Leu	Met	Lys	Asn	Arg	Glu	Pro	
305					310				315						320	
ggt	tta	att	gaa	tcg	gca	gcc	aga	ttt	gcc	ggc	tgg	aat	atg	atc	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	ggt	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			

ES 2 626 571 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 cct 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> 30

<211> 1416

5 <212> ADN

<213> Bacillus subtilis IAM1214

<400> 30

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

ES 2 626 571 T3

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	
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	
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			
ggt	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					

ES 2 626 571 T3

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> 31  
 <211> 1416  
 <212> ADN  
 <213> Bacillus subtilis ATCC21555  
 <400> 31  
 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  
 50 55 60  
 ctg gct gat ttt gaa cat ccc gat tcg att tat 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 gtc gtc gaa gaa atc gtg aaa gtg gcc 288  
 His Asp Lys Pro Glu Glu Glu Val Val Glu Glu Ile Val Lys Val Ala  
 85 90 95  
 gac atg ttt ggg gtt gac gcc att acg acc aac aat gaa ctg ttt atc 336  
 Asp Met Phe Gly Val Asp Ala Ile Thr Thr Asn Asn Glu Leu Phe Ile  
 100 105 110  
 gct ccg atg gca aaa gcg tgt aaa cgt ctc ggc ctg cgg gga gcg gcc 384  
 Ala Pro Met Ala Lys Ala Cys Lys Arg Leu Gly Leu Arg Gly Ala Gly  
 115 120 125  
 gta cag gcc gct gaa aac gcc aga gat aaa aat aaa atg aga gcc gcc 432  
 Val Gln Ala Ala Glu Asn Ala Arg Asp Lys Asn Lys Met Arg Ala Ala  
 130 135 140  
 ttc aac ccg gcc ggc gtc aaa tcc atc aaa aac aaa cgg gtg acg acc 480

5

ES 2 626 571 T3

Phe 145	Asn	Arg	Ala	Gly	Val 150	Lys	Ser	Ile	Lys	Asn 155	Lys	Arg	Val	Thr	Thr 160	
ctg Leu	gaa Glu	gat Asp	ttc Phe	cgc Arg 165	gcc Ala	gcg Ala	ctt Leu	cag Gln	gaa Glu 170	atc Ile	gga Gly	acg Thr	ccg Pro	ctt Leu 175	att Ile	528
ctg Leu	aag Lys	cct Pro	aca Thr 180	tat Tyr	ctg Leu	gca Ala	agc Ser	tcg Ser 185	atc Ile	ggc Gly	gtg Val	acg Thr	ctt Leu 190	att Ile	aaa Lys	576
gag Glu	atg Met	gaa Glu 195	acg Thr	gcc Ala	gaa Glu	gct Ala	gaa Glu 200	ttc Phe	aac Asn	aga Arg	gtc Val	aat Asn 205	gag Glu	tac Tyr	ttg Leu	624
aaa Lys	tcg Ser 210	att Ile	aat Asn	gta Val	ccg Pro	aaa Lys 215	gcg Ala	gtg Val	acg Thr	ttt Phe	gaa Glu 220	gcg Ala	ccg Pro	ttt Phe	atc Ile	672
gcg Ala 225	gaa Glu	gaa Glu	ttc Phe	ttg Leu	cag Gln 230	ggc Gly	gag Glu	tat Tyr	gat Asp	gac Asp 235	tgg Trp	tac Tyr	gaa Glu	aca Thr	agc Ser 240	720
ggt Gly	tat Tyr	tcc Ser	gac Asp 245	tat Tyr	atc Ile	agc Ser	atc Ile	gaa Glu	ggc Gly 250	atc Ile	atg Met	gcc Ala	gac Asp	gga Gly 255	gaa Glu	768
tac Tyr	ttc Phe	ccc Pro	gtt Val 260	gcg Ala	atc Ile	cat His	gat Asp 265	aaa Lys	aca Thr	ccg Pro	caa Gln	atc Ile	gga Gly 270	ttc Phe	acg Thr	816
gag Glu	aca Thr 275	gcg Ala	cat His	att Ile	acg Thr	ccg Pro	tcc Ser 280	atc Ile	ctg Leu	gat Asp	gat Asp 285	gac Asp	gcc Ala	aag Lys	cgg Arg	864
aaa Lys	atc Ile 290	gtc Val	gaa Glu	gct Ala	gcc Ala	aag Lys 295	aag Lys	gcg Ala	aat Asn	gaa Glu	gga Gly 300	ctc Leu	ggc Gly	ctc Leu	gaa Glu	912
aac Asn 305	tgt Cys	gca Ala	acg Thr	cat His 310	aca Thr	gaa Glu	ata Ile	aaa Lys	tta Leu	atg Met 315	aaa Lys	aac Asn	cgg Arg	gaa Glu	gcc Ala 320	960
gga Gly	ctg Leu	att Ile	gag Glu	tca Ser 325	gcg Ala	gcc Ala	aga Arg	ttc Phe	gcg Ala 330	gga Gly	tgg Trp	aat Asn	atg Met 335	att Ile	ccg Pro	1008
aat Asn	att Ile	aaa Lys 340	aag Lys	gtc Val	ttc Phe	ggc Gly	gtt Val	gat Asp 345	atg Met	gcg Ala	cag Gln	cta Leu	tta Leu 350	ttg Leu	gat Asp	1056
gtt Val	ctc Leu	tgt Cys 355	tac Tyr	gga Gly	aaa Lys	gaa Glu	gct Ala 360	gat Asp	ctg Leu	ccg Pro	aaa Lys 365	gga Gly 365	tta Leu	ttg Leu	gag Glu	1104
cag Gln 370	gag Glu	cca Pro	tgc Cys	tat Tyr	gtc Val	gca Ala 375	gac Asp	tgc Cys	cac His	ttg Leu	tat Tyr 380	cct Pro	cag Gln	cat His	ttc Phe	1152
aaa Lys 385	gag Glu	aac Asn	ggc Gly	cag Gln	ctg Leu 390	cct Pro	gag Glu	acg Thr	gtt Val	gtc Val 395	gat Asp	ttc Phe	gtc Val	att Ile	gaa Glu 400	1200
agc Ser	att Ile	gaa Glu	att Ile	cct Pro 405	gac Asp	ggc Gly	gtc Val	tta Leu	aag Lys 410	gga Gly	gac Asp	act Thr	gaa Glu	ctc Leu 415	gtt Val	1248
tct	ttc	tca	gcg	gct	gaq	gcq	qqt	acq	tca	ata	qat	cta	cqq	ctq	ttc	1296

ES 2 626 571 T3

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 Ala Lys Leu  
450 455 460

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

<210> 32  
<211> 1416  
<212> ADN  
<213> Bacillus amyloliquefaciens IFO3022

<400> 32

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

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 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 ccg 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

5

ES 2 626 571 T3

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 ccg 864  
 Glu Thr Ser His Ile Thr Pro Ser Ile Leu 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

ggt 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 ccg ctg ttc 1296  
 Ser Phe Ser Ala Ala Glu Ala Gly Thr 425 Ser Val Asp Leu Arg Leu Phe  
 420 430

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

<210> 33  
 <211> 93  
 <212> PRT

5

ES 2 626 571 T3

<213> Bacillus subtilis 168

<400> 33

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

5

<210> 34

<211> 279

<212> ADN

<213> Bacillus subtilis 168

10

<400> 34

ggt 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

15

<210> 35

<211> 27

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: ADN sintético

<400> 35

ttaaccatgg agagaaaaac agtattg 27

25

<210> 36

<211> 30

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Descripción de la secuencia artificial: ADN sintético

	<400> 36 <b>atatggatcc tactggcagc acatactttg</b>	30
5	<210> 37 <211> 21 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: ADN sintético	
	<400> 37 <b>caccgcagac ggaggataca c</b>	21
15	<210> 38 <211> 22 <212> ADN <213> Secuencia artificial	
20	<220> <223> Descripción de la secuencia artificial: ADN sintético	
25	<400> 38 <b>cggacgtcac ccaataatcg tg</b>	22
30	<210> 39 <211> 34 <212> ADN <213> Secuencia artificial	
35	<220> <223> Descripción de la secuencia artificial: ADN sintético	
	<400> 39 <b>gaagttccta tacttttctag agaataggaa cttc</b>	34
40	<210> 40 <211> 56 <212> ADN <213> Secuencia artificial	
45	<220> <223> Descripción de la secuencia artificial: ADN sintético	
	<400> 40 <b>ctaaccctgt gacctgcaat actgttttgc gggtgagtgt aggctggagc tgcttc</b>	56
50	<210> 41 <211> 56 <212> ADN <213> Secuencia artificial	
55	<220> <223> Descripción de la secuencia artificial: ADN sintético	
	<400> 41 <b>gaaactgccg gaaggcgatt aaacgccatc cggcagcata tgaatatacct ccttag</b>	56
60	<210> 42 <211> 56 <212> ADN <213> Secuencia artificial	
65	<220> <223> Descripción de la secuencia artificial: ADN sintético	

ES 2 626 571 T3

<400> 42  
**ttacgcaaca ggaatagact gaacaccaga ctctatgtgt aggctggagc tgcttc** 56

5 <210> 43  
 <211> 56  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

<400> 43  
**agaaaacagg ggtaaattcc ccgaatggcg gcgctacata tgaatattcct ccttag** 56

15 <210> 44  
 <211> 56  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

<400> 44  
**atggagtta gtgtaaaaag cggtagcccg gagaaagtgt aggctggagc tgcttc** 56

25 <210> 45  
 <211> 56  
 <212> ADN  
 <213> Secuencia artificial

30 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

35 <400> 45  
**ttactcttcg ccgttaaacc cagcgcggtt taacagcata tgaatattcct ccttag** 56

40 <210> 46  
 <211> 56  
 <212> ADN  
 <213> Secuencia artificial

45 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

<400> 46  
**atgacagaag cgatgaagat taccctctct acccaagtgt aggctggagc tgcttc** 56

50 <210> 47  
 <211> 56  
 <212> ADN  
 <213> Secuencia artificial

55 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

<400> 47  
**ttacgccgtt aacagattag ctatcgtgcy cacaccata tgaatattcct ccttag** 56

60 <210> 48  
 <211> 56  
 <212> ADN  
 <213> Secuencia artificial

65 <220>

ES 2 626 571 T3

	<223> Descripción de la secuencia artificial: ADN sintético	
	<400> 48 gcatcccccac ctcataacgt tgacccgacc gggcaagtgt aggctggagc tgcttc	56
5	<210> 49 <211> 56 <212> ADN <213> Secuencia artificial	
10	<220> <223> Descripción de la secuencia artificial: ADN sintético	
15	<400> 49 ctgtacggca ttttgctatg cttgtcgcca ctgttgcata tgaatattcct ccttag	56
20	<210> 50 <211> 21 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de la secuencia artificial: ADN sintético	
30	<400> 50 gtgtctgaac tgtctcaatt a	21
35	<210> 51 <211> 21 <212> ADN <213> Secuencia artificial	
40	<220> <223> Descripción de la secuencia artificial: ADN sintético	
45	<400> 51 cggaatttct ttcagcagtt c	21
50	<210> 52 <211> 21 <212> ADN <213> Secuencia artificial	
55	<220> <223> Descripción de la secuencia artificial: ADN sintético	
60	<400> 52 atgactcaac agccacaagc c	21
65	<210> 53 <211> 21 <212> ADN <213> Secuencia artificial	
70	<220> <223> Descripción de la secuencia artificial: ADN sintético	
75	<400> 53 tgcttttagtt atctttctcgt a	21
80	<210> 54 <211> 21 <212> ADN <213> Secuencia artificial	
85		

<220>  
 <223> Descripción de la secuencia artificial: ADN sintético

5 <400> 54  
 agtgcctgca tcgtcgtggg c 21

<210> 55  
 <211> 21  
 <212> ADN  
 10 <213> Secuencia artificial

<220>  
 <223> Descripción de la secuencia artificial: ADN sintético

15 <400> 55  
 gggcctttt gctttaccag a 21

<210> 56  
 <211> 21  
 20 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Descripción de la secuencia artificial: ADN sintético

25 <400> 56  
 gacgcgcgct ggggagaaaa a 21

<210> 57  
 <211> 21  
 30 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Descripción de la secuencia artificial: ADN sintético

35 <400> 57  
 cgtagcggcc gcagaccact g 21

<210> 58  
 <211> 21  
 40 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Descripción de la secuencia artificial: ADN sintético

45 <400> 58  
 atgcgtatTT ccttgaaaaa g 21

50 <210> 59  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial

55 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

60 <400> 59  
 ttattcgata gagacgtttt c 21

<210> 60  
 <211> 29  
 <212> ADN  
 65 <213> Secuencia artificial

	<220> <223> Descripción de la secuencia artificial: ADN sintético	
5	<400> 60 tacactcgag attaaagagg agaaattaa	29
10	<210> 61 <211> 30 <212> ADN <213> Secuencia artificial	
15	<220> <223> Descripción de la secuencia artificial: ADN sintético	
	<400> 61 ttaggatcct catactggca gcacatactt	30
20	<210> 62 <211> 24 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de la secuencia artificial: ADN sintético	
	<400> 62 caagaattct catgtttgac agct	24
30	<210> 63 <211> 28 <212> ADN <213> Secuencia artificial	
35	<220> <223> Descripción de la secuencia artificial: ADN sintético	
40	<400> 63 taactcgaga ttcccttttt acgtgaac	28
45	<210> 64 <211> 1428 <212> ADN <213> Bacillus pumullis NRRL B-12025	
	<400> 64	

ES 2 626 571 T3

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	
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	ggt	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	
65				70						75					80	
cat	gat	gat	tat	cca	aaa	tca	gaa	gaa	gaa	ggt	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				
ggt	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	

ES 2 626 571 T3

				245					250					255		
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	gca	gct	cga	ttc	gct	ggc	tgg	aat	1008
Arg	Glu	Thr	Gly	Leu	Ile	Glu	Ala	Ala	Ala	Arg	Phe	Ala	Gly	Trp	Asn	
				325					330					335		
atg	atc	ccg	aat	att	aaa	aaa	gtc	ttt	ggc	gtc	gat	atg	gca	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	gca	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		
gca	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	gca	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> 65

<211> 1416

5 <212> ADN

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

<400> 65

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

ES 2 626 571 T3

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	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				
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	
					150					155					160	
ctt	gaa	gat	ttc	cg	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	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	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				

ES 2 626 571 T3

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  
 ggt tta ata gag tcg gca 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  
 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 gtt 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

5 <210> 66  
 <211> 23  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

<400> 66  
 ccgatggcra aagcstgtra acg 23

15 <210> 67  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético

<400> 67  
 cggcagatcr gcdtcttttc c 21

25 <210> 68

ES 2 626 571 T3

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

5 <400> 68  
 Val<sub>1</sub> Leu Ser Leu Ser<sub>5</sub> Lys Lys Thr Val Leu<sub>10</sub> Val Ile Ala Asp Leu<sub>15</sub> Gly  
 Gly Cys Pro Pro<sub>20</sub> His Met Phe Tyr Glu<sub>25</sub> Ser Val Ala Ala Ser<sub>30</sub> Tyr His  
 Ile Val Ser<sub>35</sub> Tyr Ile Pro Arg Pro<sub>40</sub> Phe Ala Ile Thr Lys<sub>45</sub> Gly His Ala  
 Glu Leu<sub>50</sub> Ile Glu Lys Tyr Ser<sub>55</sub> Ile Ala Val Ile Lys<sub>60</sub> Asp Arg Asp Tyr  
 Phe<sub>65</sub> Glu Thr His Pro Ser<sub>70</sub> Phe Glu His Pro Asp<sub>75</sub> Ser Ile Tyr Trp Ala<sub>80</sub>  
 His Asp Asp Tyr Pro<sub>85</sub> Lys Ser Glu Glu Glu<sub>90</sub> Val Val Glu Asp Phe<sub>95</sub> Ile  
 Arg Val Ala Ser<sub>100</sub> Phe Phe Lys Ala Asp<sub>105</sub> Ala Ile Thr Thr Asn<sub>110</sub> Asn Glu  
 Leu Phe Ile<sub>115</sub> Ala Pro Met Ala Lys<sub>120</sub> Ala Ala Glu Arg Leu<sub>125</sub> Gly Leu Arg  
 Gly Ala Gly Val Lys Ala Ala<sub>135</sub> Glu Met Ala Arg Asp<sub>140</sub> Lys Ser Gln Met  
 Arg Ala Ala Phe Asn Ala<sub>150</sub> Ser Gly Val Lys Ala<sub>155</sub> Val Lys Thr Gln Pro<sub>160</sub>  
 Val Thr Thr Leu Ser<sub>165</sub> Asp Phe Gln Gln Ala Ile Glu Ser Ile Gly<sub>175</sub> Thr  
 Pro Leu Ile Leu<sub>180</sub> Lys Pro Thr Tyr Leu<sub>185</sub> Ala Ser Ser Ile Gly<sub>190</sub> Val Thr  
 Leu Phe His<sub>195</sub> Asp Lys Ala Gly Ser<sub>200</sub> Asp Asp Leu Phe Leu<sub>205</sub> Gln Val Gln  
 Ser Tyr<sub>210</sub> Leu Glu Thr Ile Pro<sub>215</sub> Val Pro Asp Ala Val<sub>220</sub> Thr Tyr Glu Ala  
 Pro Phe Val Ala Glu Thr Tyr Leu Glu Gly Ala<sub>235</sub> Tyr Glu Asp Trp Tyr<sub>240</sub>  
 Glu Asp Glu Gly Tyr<sub>245</sub> Ala Asp Tyr Val Ser<sub>250</sub> Val Glu Gly Leu Val<sub>255</sub> Val  
 Glu Gly Glu Tyr<sub>260</sub> Leu Pro Phe Val Ile<sub>265</sub> His Asp Lys Thr Pro Gln Ile  
 Gly Phe Thr<sub>275</sub> Glu Thr Ala His Ile<sub>280</sub> Thr Pro Thr Ile Leu<sub>285</sub> Asp Asn Glu  
 Ala Lys Gln Ile Ile Ile Glu Ala Ala Arg Lys Ala<sub>300</sub> Asn Glu Gly Leu

ES 2 626 571 T3

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

- 5 <210> 69  
 <211> 44  
 <212> ADN  
 <213> Secuencia artificial
- 10 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético  
 <400> 69  
 catgccatgg agaaaaaac tgtacttgtc attgctgact tagg □□□□@ 44
- 15 <210> 70  
 <211> 38  
 <212> ADN  
 <213> Secuencia artificial
- 20 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético  
 <400> 70  
 cgcggatccc ttcactaatt catccattaa ctgaatcg □□□□@38
- 25 <210> 71  
 <211> 26  
 <212> ADN  
 <213> Secuencia artificial
- 30 <220>  
 <223> Descripción de la secuencia artificial: ADN sintético  
 <400> 71  
 gctaggtcctt gaacattgtg caaccc 26
- 35 <210> 72  
 <211> 23

<212> ADN  
<213> Secuencia artificial

<220>

5 <223> Descripción de la secuencia artificial: ADN sintético

<400> 72

ggtgttccga tagactcaat ggc

23

## REIVINDICACIONES

- 5 1. Un proceso para producir un dipéptido, que comprende:
- dejar que una fuente de enzima y uno o más tipos de aminoácidos estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de  
 (a) un microorganismo; o  
 (b) un material tratado del cultivo de (a)
- 10 dejar que el dipéptido se forme y acumule en el medio acuoso; y recuperar el dipéptido del medio,  
 en donde dicho microorganismo es un microorganismo en el que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominados también proteínas transportadoras de péptidos) se reducen o pierden comparadas con la cepa parental, dicha cepa parental pertenece al género *Escherichia*, *Bacillus*, *Corynebacterium* o *Saccharomyces*, y en donde dicho microorganismo tiene la capacidad para producir un dipéptido, en donde dicha capacidad para producir un dipéptido es la capacidad para producir una proteína según cualquiera de [1] a [4] a continuación:
- 15 [1] una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68;  
 [2] una proteína que consiste en una secuencia de aminoácidos en donde uno o más residuos de aminoácidos se delecionan, sustituyen o añaden en la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido;  
 [3] una proteína que consiste en una secuencia de aminoácidos que tiene el 65% o más de homología con la  
 20 secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 19 a 25 y 68 y que tiene la actividad para sintetizar un dipéptido; y  
 [4] una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología con la secuencia de aminoácidos mostrada en SEQ ID NO: 33 y que tiene la capacidad de sintetizar un dipéptido,
- 30 y en donde el material tratado del cultivo es cultivo concentrado, cultivo seco, células obtenidas centrifugando el cultivo, o un producto obtenido sometiendo las células a secado, liofilización o inmovilización.
- 35 2. El proceso según la reivindicación 1, en donde el dipéptido es un dipéptido representado por la siguiente fórmula (I):
- R1-R2 (I)
- 40 en donde R1 y R2, que pueden ser iguales o diferentes, representan cada uno un aminoácido.
3. Un proceso para producir un dipéptido que comprende:
- dejar que una fuente de enzima, un éster de L-aminoácido y un L- aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de un microorganismo o un material tratado del cultivo;
- 45 dejar que el dipéptido se forme y acumule en el medio acuoso; y recuperar el dipéptido del medio,  
 en donde dicho microorganismo es un microorganismo en el que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominados también proteínas transportadoras de péptidos) se reducen o pierden comparadas con la cepa parental, dicha cepa parental pertenece al género *Escherichia*, *Bacillus*, *Corynebacterium* o *Saccharomyces*, y en donde dicho microorganismo tiene la capacidad para producir un dipéptido,  
 50 en donde la peptidasa es una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4, o una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4 y que tiene actividad peptidasa,  
 y en donde dicha capacidad para producir un dipéptido es la capacidad para formar un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido produciendo prolina iminopeptidasa;  
 y en donde el material tratado del cultivo es cultivo concentrado, cultivo seco, células obtenidas centrifugando el cultivo, o un producto obtenido sometiendo las células a secado, liofilización o inmovilización.
- 60 4. Un proceso para producir un dipéptido, que comprende:
- dejar que una fuente de enzima, una amida de L-aminoácido y un L- aminoácido estén presentes en un medio acuoso, dicha fuente de enzima es un cultivo de un microorganismo o un material tratado del cultivo;
- 65 dejar que el dipéptido se forme y acumule en el medio acuoso; y recuperar el dipéptido del medio,

- en donde dicho microorganismo es un microorganismo en el que las actividades de uno o más tipos de peptidasas y uno o más tipos de proteínas que tienen actividad transportadora de péptidos (de aquí en adelante denominados también proteínas transportadoras de péptidos) se reducen o pierden comparadas con la cepa parental, dicha cepa parental pertenece al género *Escherichia*, *Bacillus*, *Corynebacterium* o *Saccharomyces*, y en donde dicho microorganismo tiene la capacidad para producir un dipéptido, en donde dicha capacidad para producir un dipéptido es la capacidad para formar un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido produciendo una proteína que tiene actividad amida de L-aminoácido hidrolasa,
- y en donde el material tratado del cultivo es cultivo concentrado, cultivo seco, células obtenidas centrifugando el cultivo, o un producto obtenido sometiendo las células a secado, liofilización o inmovilización.
5. El proceso según cualquiera de las reivindicaciones 1 a 4, que comprende además la formulación del dipéptido producido en un aditivo alimenticio, una composición farmacéutica o un cosmético.
  6. El proceso según la reivindicación 1 o 4, en donde la peptidasa es una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4, o una proteína que tiene una secuencia de aminoácidos que tiene el 80% o más de homología respecto a la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 1 a 4 y que tiene actividad peptidasa.
  7. El proceso según la reivindicación 1, 3 o 4, en donde la proteína transportadora de péptidos es una proteína que tiene la secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO: 5 a 9, o una proteína que tiene mostrada en cualquiera de SEQ ID NO: 5 a 9 y que tiene actividad transportadora de péptidos.
  8. El proceso según la reivindicación 1, en donde el microorganismo que tiene la capacidad para producir un dipéptido es un microorganismo que tiene un ADN según cualquiera de [1] a [4] a continuación:
    - [1] ADN que codifica la proteína según cualquiera de [1] a [4] de la reivindicación 2(a);
    - [2] ADN que tiene la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65;
    - [3] ADN que hibrida con ADN que tiene la secuencia de nucleótidos mostrada en cualquiera de SEQ ID NO: 26 a 32, 64 y 65 en condiciones rigurosas y que codifica una proteína que tiene la actividad para sintetizar un dipéptido; y
    - [4] ADN que tiene una secuencia de nucleótido que tiene el 80% o más de homología con la secuencia de nucleótidos mostrada en SEQ ID NO: 34 y que codifica una proteína que tiene la actividad para sintetizar un dipéptido.
  9. El proceso según la reivindicación 1, en donde el microorganismo que tiene la capacidad para producir un dipéptido es un microorganismo que tiene un ADN recombinante en el que el ADN según cualquiera de [1] a [4] de la reivindicación 8 está ligado a un ADN vector.
  10. El proceso según la reivindicación 3, en donde el microorganismo que tiene la capacidad para producir un dipéptido a partir de un éster de L-aminoácido y un L-aminoácido es un microorganismo que tiene un ADN recombinante en el que el ADN que codifica una proteína que tiene actividad prolina iminopeptidasa está ligado a un ADN vector.
  11. El proceso según la reivindicación 4, en donde el microorganismo que tiene la capacidad para producir un dipéptido a partir de una amida de L-aminoácido y un L-aminoácido es un microorganismo que tiene un ADN recombinante en el que el ADN que codifica una proteína que tiene actividad amida de L-aminoácido hidrolasa está ligado a un ADN vector.
  12. El proceso según cualquiera de las reivindicaciones 1, 3 o 4, en donde el microorganismo es un microorganismo que pertenece al género *Escherichia*, *Bacillus*, *Corynebacterium* o *Saccharomyces*.

Fig.1

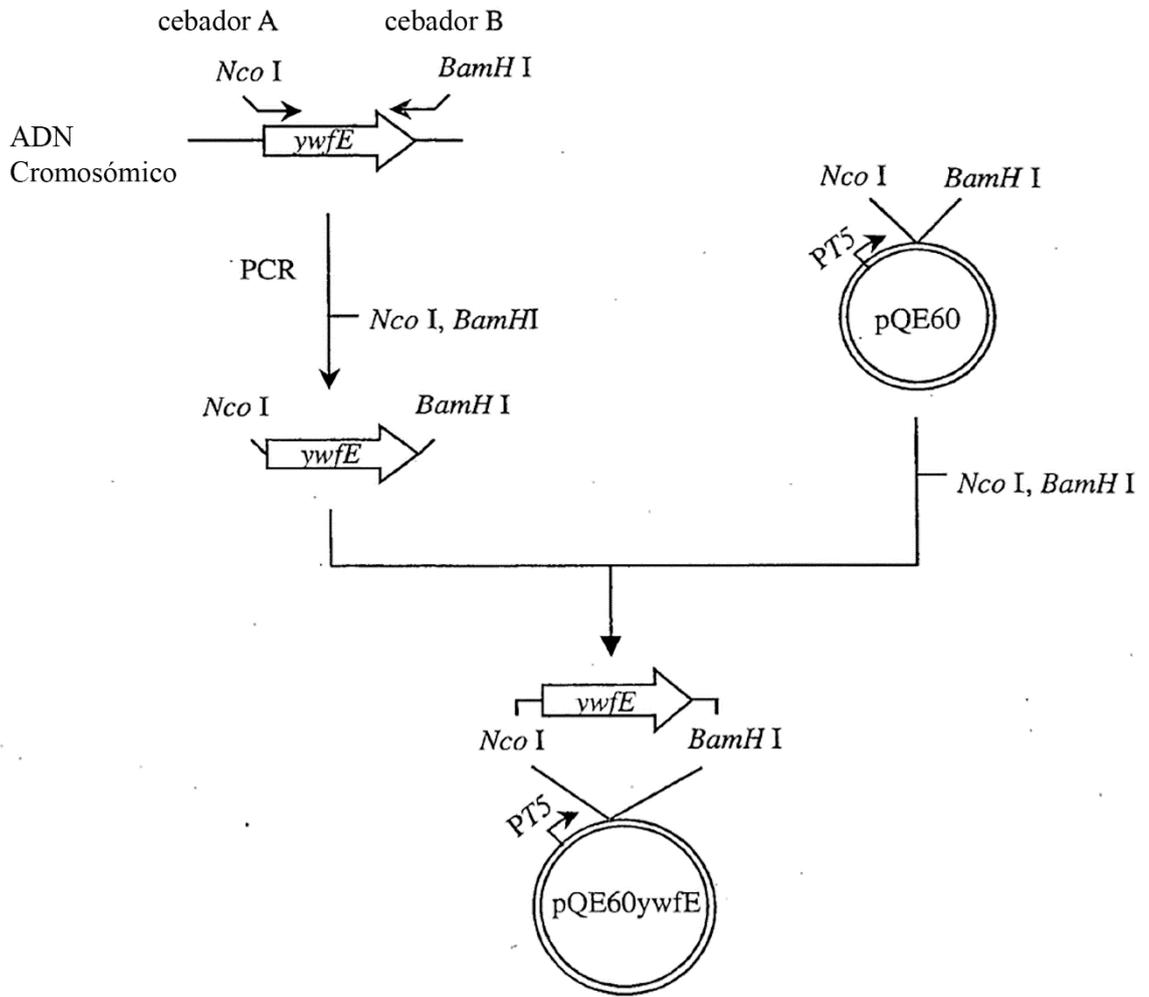


Fig. 2

