

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 380 105**

51 Int. Cl.:
C12N 9/52 (2006.01)
C12N 15/57 (2006.01)
C12N 5/10 (2006.01)
C12N 15/82 (2006.01)
A23K 1/165 (2006.01)
A23L 1/305 (2006.01)
C11D 3/386 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **05748647 .4**
96 Fecha de presentación: **17.06.2005**
97 Número de publicación de la solicitud: **1766001**
97 Fecha de publicación de la solicitud: **28.03.2007**

54 Título: **Proteasas de Nocardiosis**

30 Prioridad:
21.06.2004 DK 200400969

45 Fecha de publicación de la mención BOPI:
08.05.2012

45 Fecha de la publicación del folleto de la patente:
08.05.2012

73 Titular/es:
**NOVOZYMES A/S
KROGSHØJVEJ 36
2880 BAGSVAERD, DK**

72 Inventor/es:
**LASSEN, Søren, Flensted;
SJØHOLM, Carsten;
ØSTERGAARD, Peter, Rahbek y
FISCHER, Morten**

74 Agente/Representante:
Tomas Gil, Tesifonte Enrique

ES 2 380 105 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

Proteasas de *Nocardioopsis*

Campo de la invención

5 [0001] La presente invención se refiere a polipéptidos aislados con actividad de proteasa y secuencias de ácidos nucleicos aisladas que codifican los polipéptidos. La invención también se refiere a construcciones de ácidos nucleicos, vectores y células huésped, incluyendo células de plantas y animales, comprendiendo las secuencias de ácidos nucleicos, al igual que métodos para producción y uso de los polipéptidos, en particular, el uso de los polipéptidos en alimento para animales y detergentes.

Antecedentes de la invención

10 [0002] Proteasas derivadas de *Nocardioopsis* sp. NRRL 18262 y *Nocardioopsis dassonvillei* NRRL 18133 se describen en la WO 88/03947. El ADN y secuencias de aminoácidos de la proteasa derivados de *Nocardioopsis* sp. NRRL 18262 se muestran en DK aplicación nº 1996 00013. La WO 01/58276 divulga el uso en alimento para animales de proteasas de ácido estables relacionadas con la proteasa derivada de *Nocardioopsis* sp. NRRL 18262, al igual que una proteasa derivada de *Nocardioopsis alba* DSM 14010.

15 [0003] La JP 2-255081-A divulga una proteasa derivada de *Nocardioopsis* sp. cepa OPC-210 (FERM P-10508), sin embargo, sin información de secuencia. La cepa ya no está disponible, ya que el depósito se retiró.

[0004] La DD 200432|8 divulga una preparación proteolítica derivada de *Nocardioopsis dassonvillei* cepa ZIMET 43647, sin embargo, sin información de secuencia. La cepa parece no estar ya disponible.

20 [0005] La JP 2003284571-A divulga la secuencia de aminoácidos y la secuencia de ADN correspondiente de una proteasa derivada de *Nocardioopsis* sp. TOA-1 (FERM P-18676). La secuencia se ha introducido en GENESEQP con nº ADF43564.

[0006] Es un objeto de la presente invención proporcionar proteasas alternativas, en particular, para uso en alimento para animales y/o detergentes.

Resumen de la invención

25 [0007] Varias proteasas se clonaron, purificaron y caracterizaron. Estas proteasas se designan de la siguiente manera: proteasa L1a derivadas de *Nocardioopsis dassonvillei* subesp. *dassonvillei* DSM 43235 (ver SEC ID nº 1 y 2); proteasa L1b derivada de *Nocardioopsis prasina* DSM 15649 (ver SEC ID nº: 3 y 4); proteasa L1c derivada de *Nocardioopsis prasina* (previamente *alba*) DSM 14010 (ver SEC ID nº: 5 y 6); proteasa L2a derivada de *Nocardioopsis* sp. DSM 16424 (ver SEC ID nº: 7 y 8); proteasa L2b derivada de *Nocardioopsis alkaliphila* DSM 44657 (ver SEC ID nº: 9 y 10); y proteasa L2c derivada de *Nocardioopsis lucentensis* DSM 44048 (ver SEC ID nº: 11 y 12).

[0008] Un polipéptido aislado con actividad de proteasa, seleccionado del grupo que consiste en: (a) un polipéptido con una secuencia de aminoácidos que tiene un grado de identidad para aminoácidos 1-192 de la SEC ID nº: 2 de al menos 80%; (b) un fragmento de (a) que tiene actividad de proteasa.

35 [0009] La invención también se refiere a secuencias de ácidos nucleicos aisladas que codifican los polipéptidos anteriores y a construcciones de ácidos nucleicos, vectores y células huésped comprendiendo las secuencias de ácidos nucleicos al igual que métodos para producción y uso de los polipéptidos, en particular, en alimento para animales y detergentes.

Descripción detallada de la invención

Polipéptidos con actividad de proteasa

40 [0010] Polipéptidos con actividad de proteasa, o proteasas, se designan también a veces también peptidasas, proteinasas, hidrolasas peptídicas o enzimas proteolíticas. Las proteasas pueden ser del tipo exo que hidrolizan péptidos empezando en cualquier extremo de estos, o del tipo endo que actúan internamente en las cadenas polipeptídicas (endopeptidasas). Endopeptidasas muestran actividad en sustratos de péptido bloqueados N- y C-terminalmente que son pertinentes para la especificidad de la proteasa en cuestión.

45 [0011] El término "proteasa" se define aquí como una enzima que hidroliza enlaces peptídicos. Incluye cualquier enzima que pertenece al grupo enzimático EC 3.4 (incluyendo cada una de las trece subclases de esta). El número EC se refiere a Enzyme Nomenclature 1992 de NC-IUBMB, Academic Press, San Diego, California, incluyendo suplementos 1-5 publicados en *Eur. J. Biochem.* 1994, 223, 1-5; *Eur. J. Biochem.* 1995, 232, 1-6; *Eur. J. Biochem.* 1996, 237, 1-5; *Eur. J. Biochem.* 1997, 250, 1-6; y *Eur. J. Biochem.* 1999, 264, 610-650 respectivamente. La nomenclatura es regularmente complementada y actualizada; ver, por ejemplo, la World Wide Web (WWW) en <http://www.chem.qmw.ac.uk/iubmb/enzyme/index.html>.

[0012] Las proteasas se clasifican basándose en sus mecanismos catalíticos en los siguientes grupos: serina proteasas (S), cisteína proteasas (C), proteasas aspárticas (A), metalo proteasas (M), y desconocido, o aún sin clasificar, proteasas (U), ver *Handbook of Proteolytic Enzymes*, A.J.Barrett, N.D.Rawlings, J.F.Woessner (eds), Academic Press (1998), en particular la parte de introducción general.

5 [0013] En formas de realización particulares, las proteasas de la invención y para uso según la invención son seleccionadas del grupo que consisten en:

- (a) proteasas del grupo enzimático EC 3.4.-.-;
- (b) serina proteasas del grupo S del manual anterior;
- (c) serina proteasas de la familia de peptidasa S2A; y/o

10 (d) serina proteasas de la familia de peptidasa S1 E como se describe en *Biochem.J.* 290:205-218 (1993) y en *MEROPS protease database, release 6.20, March 24, 2003*, (www.merops.ac.uk). La base de datos se describe en Rawlings, N.D., O'Brien, E. A. & Barrett, A.J. (2002) *MEROPS: the protease database.Nucleic Acids Res.* 30, 343-346.

15 [0014] Para determinar si una proteasa dada es una serina proteasa, y una familia de proteasa S2A, se hace referencia al manual anterior y a los principios indicados en este. Tal determinación puede llevarse a cabo para todo tipo de proteasas, sean proteasas de origen natural o de tipo salvaje; o creadas genéticamente o proteasas sintéticas.

20 [0015] Actividad de proteasa puede medirse usando cualquier ensayo, en el que un sustrato se emplea, que incluye enlaces peptídicos pertinentes para la especificidad de la proteasa en cuestión. PH de ensayo y temperatura de ensayo se adaptan asimismo a la proteasa en cuestión. Ejemplos de valores de pH de ensayo son pH 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, o 12. Ejemplos de temperaturas de ensayo son 30, 35, 37, 40, 45, 50, 55, 60, 65, 70, 80, 90, o 95°C.

25 [0016] Ejemplos de sustratos de proteasa son caseína, como caseína entrecruzada con azurina (caseína AZCL). Tres ensayos de proteasa se describen en los ejemplos 4-5 aquí, cualquiera de los cuales se puede usar para determinar actividad de proteasa. Para los fines de esta invención, el ensayo denominado pNA es un ensayo preferido.

30 [0017] No hay limitaciones en el origen de la proteasa de la invención y/o para uso según la invención. Así, el término proteasa incluye no solo proteasas naturales o de tipo salvaje obtenidas de microorganismos de cualquier género, sino también cualquier mutaciones, variantes, fragmentos etc. de estas que muestran actividad de proteasa, al igual que proteasas sintéticas, como proteasas transpuestas y proteasas de consenso. Tales proteasas creadas genéticamente se pueden preparar como se conoce generalmente en la técnica, por ejemplo, por mutagénesis dirigida, por PCR (usando un fragmento de PCR con la mutación deseada como uno de los cebadores en las reacciones de la PCR), o por mutagénesis aleatoria. La preparación de proteínas de consenso se describe en, por ejemplo, el documento EP 897985. Transposición de genes se describe generalmente en, por ejemplo, los documentos WO 95/22625 y WO 96/00343. Recombinación de genes de proteasa puede hacerse independientemente de la secuencia específica de los progenitores por transposición sintética como se describe en Ness, J.E. et al, en *Nature Biotechnology*, Vol. 20 (12), pp. 1251-1255, 2002. Oligonucleótidos sintéticos degenerados en su secuencia de ADN para proporcionar la posibilidad de que todos los aminoácidos encontrados en el conjunto de proteasas progenitoras se diseñan y los genes se ensamblan según la referencia. La transposición puede llevarse a cabo para la secuencia de longitud completa o para solo parte de la secuencia y luego más tarde combinada con el resto del gen para dar una secuencia de longitud completa. Las proteasas de las SEC ID n°: 2, 4, 6, 8, 10, y 12, al igual que las proteasas de *Nocardiosis* descritas en los documentos previos catalogados anteriormente, son ejemplos particulares de tales proteasas progenitoras que se pueden someter a transposición como se ha descrito anteriormente, para proporcionar proteasas adicionales de la invención. El término "obtenido de" como se utiliza en este caso en relación con una fuente dada debe significar que el polipéptido codificado por la secuencia de ácidos nucleicos se produce por la fuente o por una célula en la que está presente la secuencia de ácidos nucleicos de la fuente. En una forma de realización preferida, el polipéptido es segregado extracelularmente.

45 [0018] En una forma de realización específica, la proteasa es una variante hipoalergénica, diseñada para invocar una respuesta inmunológica reducida cuando se expone a animales, incluyendo al hombre. El término respuesta inmunológica debe entenderse como cualquier reacción del sistema inmunológico de un animal expuesto a la proteasa. Un tipo de respuesta inmunológica es una respuesta alérgica llevando a niveles aumentados de IgE en el animal expuesto. Variantes hipoalergénicas pueden prepararse usando técnicas conocidas en el técnica. Por ejemplo, la proteasa se puede conjugar con partes de protección de fracciones poliméricas o epítopos de la proteasa implicada en una respuesta inmunológica. Conjugación con polímeros puede implicar acoplamiento químico *in vitro* de polímero a la proteasa, por ejemplo, como se describe en los documentos WO 96/17929, WO 98/30682, WO 98/35026, y/o WO 99/00489. Conjugación puede en adición o alternativamente a esto implicar acoplamiento *in vivo* de polímeros a la proteasa. Tal conjugación se puede conseguir por ingeniería genética de la secuencia de nucleótidos que codifica la proteasa, por inserción de sitios de glicosilación de codificación de secuencias de consenso adicionales en la proteasa y expresión la proteasa en un huésped capaz de glicosilación de la proteasa,

ver, por ejemplo, WO 00/26354. Otra manera de provisión de variantes hipoalergénicas es ingeniería genética de la secuencia de nucleótidos que codifica la proteasa para provocar que las proteasas se auto-oligomericen, provocando que monómeros de proteasa puedan proteger los epítomos de otros monómeros de proteasa y reduciendo así la antigenicidad de los oligómeros. Tales productos y su preparación se describe, por ejemplo, en
 5 WO 96/16177. Epítomos implicados en una respuesta inmunológica se pueden identificar por varios métodos, como el método de visualización de fago descrito en WO 00/26230 y WO 01/83559, o el método aleatorio descrito en EP 561907. Una vez un epítomo ha sido identificado, su secuencia de aminoácidos se puede alterar para producir propiedades inmunológicas alteradas de la proteasa por técnicas de manipulación de gen conocidas, como mutagénesis dirigida al sitio (ver, por ejemplo, WO 00/26230, WO 00/26354 y/o WO 00/22103) y/o conjugación de un
 10 polímero puede realizarse en proximidad suficiente al epítomo para que el polímero proteja el epítomo.

[0019] Varios aspectos de la presente invención se refieren a polipéptidos aislados con actividad de proteasa (para abreviar "proteasas"), al igual que las secuencias de ácidos nucleicos aisladas correspondientes, dichos polipéptidos, o ácidos nucleicos, respectivamente, comprendiendo una secuencia de aminoácidos, o una secuencia de ácidos nucleicos, respectivamente, con un cierto grado de identidad a un fragmento específico de una secuencia
 15 de aminoácidos, o una secuencia de ácidos nucleicos, respectivamente, con una SEC ID n° específica. Los fragmentos específicos corresponden a los polipéptidos maduros, o las partes de codificación de polipéptidos maduros de las secuencias de ácidos nucleicos, respectivamente.

[0020] Para fines de la presente invención, el grado de identidad entre dos secuencias de aminoácidos, al igual que el grado de identidad entre dos secuencias de nucleótidos, se determina por el programa "align" que es un alineamiento de Needleman- Wunsch (es decir, un alineamiento global). El programa se usa para alineamiento de polipéptido, al igual que secuencias de nucleótidos. La matriz de puntuación por defecto BLOSUM50 se usa para alineamientos de polipéptido y la matriz de identidad de por defecto se usa para alineamientos de nucleótido. La penalización para el primer residuo de un espacio es -12 para polipéptidos y -16 para nucleótidos. Las penalizaciones para otros residuos de un espacio son -2 para polipéptidos y -4 para nucleótidos.
 20

[0021] "Align" es parte de la versión v20u6 del paquete FASTA (ver W. R. Pearson y D. J. Lipman (1988), "Improved Tools for Biological Sequence Analysis", PNAS 85:2444-2448, y W. R. Pearson (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA," Methods in Enzymology 183:63- 98) Alineamientos de proteína FASTA usan el algoritmo de Smith-Waterman con ninguna limitación en tamaño del espacio (ver "Smith-Waterman algorithm", T. F. Smith and M. S. Waterman (1981) J. Mol. Biol. 147:195-197).
 25

[0022] En formas de realización particulares, el polipéptido de la invención tiene un grado de identidad a las partes maduras de SEC ID n°: 2, de al menos 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97°/6, 98%, o al menos 99%.
 30

[0023] En otras formas de realización particulares, la secuencia de ácidos nucleicos de la invención tienen un grado de identidad a la parte de codificación de péptidos madura de la SEC ID n°: 1 de al menos 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o al menos 99%.
 35

[0024] En todavía otra forma de realización particular, la proteasa de la invención comprende la secuencia de aminoácidos de las partes maduras de cualquiera de las SEC ID n°: 2, 4, 6, 8, 10, o 12 o es una variante alélica de esta; o un fragmento de esta que tiene actividad de proteasa.

[0025] En otra forma de realización preferida, los polipéptidos de la presente invención consisten en la parte de péptido maduro de la SEC ID n°: 2, ; o variantes alélicas de esta; o fragmentos de esta que tienen actividad de proteasa.
 40

[0026] Un fragmento de la SEC ID n°: 2, es un polipéptido con uno o más aminoácidos eliminado del amino y/o carboxilo terminal de estas secuencias de aminoácidos. En una forma de realización, un fragmento contiene al menos 75 residuos de aminoácidos, o al menos 100 residuos de aminoácidos, o al menos 125 residuos de aminoácidos, o al menos 150 residuos de aminoácidos, o al menos 160 residuos de aminoácidos, o al menos 165 residuos de aminoácidos, o al menos 170 residuos de aminoácidos, o al menos 175 residuos de aminoácidos.
 45

[0027] Una variante alélica denota cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. Variación alélica surge naturalmente a través de mutación y puede suponer polimorfismo dentro de poblaciones. Mutaciones de gen pueden ser silenciosas (ningún cambio en el polipéptido codificado) o puede codificar polipéptidos con secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.
 50

[0028] Las secuencias de ácidos nucleicos de la SEC ID n°: 1, o las partes de codificación de péptidos maduros de estas, o una subsecuencia de estas, al igual que las secuencias de aminoácidos de la SEC ID n°: 2, o un fragmento de esta, se pueden utilizar para diseño de una sonda de ácido nucleico para identificar y clonar polipéptidos que codifican ADN con actividad de proteasa de cepas de este o géneros diferentes o especies según métodos bien conocidos en la técnica. En particular, tales sondas se pueden usar para hibridación con el genómico o ADNc del género o especies de interés, siguiendo procedimientos de transferencia de Southern estándar, para identificar y aislar el gen correspondiente en este. Tales sondas pueden ser considerablemente más cortas que la secuencia
 55

entera, pero debería ser al menos 15, preferiblemente al menos 25, y más preferiblemente al menos 35 nucleótidos de longitud. Sondas más largas también se puede usar. Ambas sondas de ADN y ARN pueden usarse. Las sondas se marcan típicamente para detectar el gen correspondiente (por ejemplo, con ^{32}P , ^3H , ^{35}S , biotina, o avidina).

5 [0029] Así, un ADN genómico o genoteca de ADNc obtenida a partir de otros organismos se pueden seleccionar para ADN que hibridiza con las sondas anteriormente descritas y que codifica un polipéptido con actividad de proteasa. Genómico u otro ADN de otros organismos se pueden separar por agarosa o electroforesis en gel de poliacrilamida, u otras técnicas de separación. ADN de las bibliotecas o el ADN separado se puede transferir a e
10 inmovilizar en nitrocelulosa u otro material portador adecuado. Para identificar un clon o ADN que es homólogo a la SEC ID n°: 1, o una subsecuencia de esta, el material portador se usa en una transferencia de Southern. Hibridación indica que la secuencia de ácidos nucleicos hibridiza a una sonda de ácido nucleico marcada correspondiente a la
15 secuencia de ácidos nucleicos mostrada en cualquiera de estas SEC ID n°, su cadena complementaria, o una subsecuencia de esta, bajo condiciones de astringencia de muy bajas a muy altas. Moléculas a la que la sonda de ácido nucleico hibridiza bajo estas condiciones son detectadas usando película radiográfica.

15 [0030] En particular, la sonda de ácido nucleico es una secuencia de ácidos nucleicos que codifica las partes de péptido maduro de SEC ID n°: 2, o subsecuencias de esta. También, la sonda de ácido nucleico es aquellos nucleótidos de la SEC ID n°: 1, que corresponden a las regiones de codificación de polipéptidos maduros.

20 [0031] Para sondas largas de al menos 100 nucleótidos de longitud, condiciones de astringencia de muy bajas a muy altas se definen como prehibridación e hibridación a 42°C en 5X SSPE, 0,3% SDS, 200 µg/ml de ADN de esperma de salmón desnaturalizado y cortado, y bien 25% de formamida para astringencias bajas y muy bajas, 35%
25 de formamida para medio y astringencias medias altas, o 50% de formamida para astringencias altas y muy altas, seguido de procedimientos de transferencia de Southern estándar.

25 [0032] Para sondas largas de al menos 100 nucleótidos de longitud, el material portador se lava finalmente tres veces cada una durante 15 minutos usando 2 x SSC, 0,2% SDS preferiblemente al menos a 45°C (astringencia muy baja), más preferiblemente al menos a 50°C (astringencia baja), más preferiblemente al menos a 55°C (astringencia
30 media), más preferiblemente al menos a 60°C (astringencia media alta), incluso más preferiblemente al menos a 65°C (astringencia alta), y más preferiblemente al menos a 70°C (astringencia muy alta). Preferiblemente, el lavado se conduce usando bien 0,2 X SSC, 0,1 X SSC o 0,02 X SSC, las otras condiciones de lavado que no tienen modificaciones (es decir, lavar tres veces, cada una durante 15 minutos; incluyen 0,2% SDS, lavado preferiblemente
35 al menos a 45°C (astringencia muy baja), más preferiblemente al menos a 50°C (astringencia baja), más preferiblemente al menos a 55°C (astringencia media), más preferiblemente al menos a 60°C (astringencia media alta), incluso más preferiblemente al menos a 65°C (astringencia alta), y más preferiblemente al menos a 70°C (astringencia muy alta).

35 [0033] Para ondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, condiciones de astringencia se definen como prehibridación, hibridación y post-hibridación de lavado de 5°C a 10°C por debajo de la T_m calculada usando el cálculo según Bolton y McCarthy (1962, *Proceedings of the National Academy of Sciences EEUU* 48:1390) en 0,9 M de NaCl, 0,09 M de Tris-HCl, pH 7,6, 6 mM de EDTA, 0,5% de NP-40, solución de Denhardt 1X, 1 mM de pirofosfato de sodio, 1 mM de fosfato monobásico de sodio, 0,1 mM de ATP y 0,2 mg de levadura ARN por ml siguiendo procedimientos de transferencia de Southern estándar.

40 [0034] Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, el material portador se lava una vez en 6X SSC más 0,1% de SDS durante 15 minutos y dos veces cada durante 15 minutos usando 6X SSC de 5°C a 10°C por debajo de la T_m calculada.

45 [0035] Las secuencias de aminoácidos de los polipéptidos variantes pueden diferir de la secuencia de aminoácidos de las partes maduras de la SEC ID n°: 2, por una inserción o delección de uno o más residuos de aminoácidos y/o la sustitución de uno o más residuos de aminoácidos por residuos de aminoácidos diferentes. Preferiblemente,
50 cambios de aminoácidos son de una naturaleza menor, que es sustituciones de aminoácido conservadoras que significativamente no afectan al pliegue y/o actividad de la proteína; delecciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; extensiones carboxilo o amino terminal pequeñas, como un residuo de metionina amino terminal; un pequeño péptido de enlace de hasta aproximadamente 20-25 residuos; o una pequeña extensión que facilita purificación cambiando la carga de red u otra función, como un tracto de polihistidina, un epítipo antigénico o un dominio de unión.

55 [0036] Ejemplos de sustituciones conservadoras están dentro del grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina) y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Por consiguiente, por ejemplo, un polipéptido con, o comprendiendo, una secuencia como se expone en la SEC ID n°: 2, donde sustituciones de aminoácido conservadoras comprenden sustituciones, una a otra, entre los aminoácidos básicos (arginina, lisina y histidina), entre los aminoácidos ácidos (ácido glutámico y ácido aspártico), entre los aminoácidos polares (glutamina y asparagina), entre los aminoácidos hidrofóbicos (alanina, leucina, isoleucina y valina), entre los aminoácidos aromáticos (fenilalanina, triptófano y tirosina), y entre los aminoácidos pequeños (glicina, alanina, serina, treonina y

metionina), o cualquier combinación de estos, o fragmentos activos de los estos. Sustituciones de aminoácido que generalmente no alteran la actividad específica se conocen en la técnica y se describen, por ejemplo, por H. Neurath and R.L. Hill, 1979, en *The Proteins*, Academic Press, Nueva York. Los intercambios que ocurren más frecuentemente son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, y Asp/Gly al igual que estos a la inversa.

[0037] Un polipéptido de la presente invención puede ser un polipéptido fúngico o bacteriano. El polipéptido fúngico se puede derivar de un hongo filamentoso o de una levadura.

[0038] En formas de realización particulares, el polipéptido de la invención es i) una proteasa bacteriana; ii) una proteasa del filo Actinobacteria; iii) de la clase Actinobacteria; iv) del orden actinomycetales; v) de la familia Nocardioaceae; vi) del género *Nocardopsis*; y/o una proteasa derivado de vii) especie *Nocardopsis*, como *Nocardopsis alba*, *Nocardopsis alkathiphila*, *Nocardopsis antarctica*, *Nocardopsis prasina*, *Nocardopsis composta*, *Nocardopsis exhalans*, *Nocardopsis halophile*, *Nocardopsis halotolerans*, *Nocardopsis kunsanensis*, *Nocardopsis listeri*, *Nocardopsis lucentensis*, *Nocardopsis metallicus*, *Nocardopsis synnemataformans*, *Nocardopsis trehalosi*, *Nocardopsis tropica*, *Nocardopsis umidischolae*, *Nocardopsis xinjiangensis*, o *Nocardopsis dassonvillei*, por ejemplo, de bien *Nocardopsis dassonvillei* subesp. *dassonvillei* DSM 43235, *Nocardopsis prasina* DSM 15649, *Nocardopsis prasina* (previamente *alba*) DSM 14010, *Nocardopsis* sp. DSM 16424, *Nocardopsis alkaliphila* DSM 44657, o fde *Nocardopsis lucentensis* DSM 44048.

[0039] En una forma de realización particular, la proteasa deriva de *Nocardopsis alba*, *Nocardopsis alkaliphila*, *Nocardopsis dassonvillei*, *Nocardopsis lucentensis*, *Nocardopsis prasina*, o *Nocardopsis* sp.

[0040] La taxonomía anterior está según el capítulo: *The road map to the Manual by G.M. Garrity & J. G. Holt en Bergey's Manual of Systematic Bacteriology, 2001, second edition, volume 1, David R. Boone, Richard W. Castenholz.*

[0041] Se entenderá que para las especies mencionadas, la invención incluye los estados imperfectos y perfectos y otros equivalentes taxonómicos, por ejemplo, anamorfos, independientemente del nombre de especies por las que se conocen. Expertos en la técnica fácilmente reconocen la identidad de equivalentes apropiados.

[0042] Cepas de estas especies son fácilmente accesibles al público en un número de colecciones de cultivo, como la American Type Culture Collection (ATCC), Deutsche Sammlung von.

[0043] Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS) y Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

[0044] Además, tales polipéptidos se pueden identificar y obtener de otras fuentes incluyendo microorganismos aislados de naturaleza (p. ej., suelo, abonos, agua, etc.) usando las sondas mencionadas anteriormente. Técnicas para aislamiento de microorganismos de hábitats naturales se conocen en la técnica. La secuencia de ácidos nucleicos puede luego derivar por selección de forma similar de un genómico o genoteca de ADN de otro microorganismo. Una vez que una secuencia de ácidos nucleicos que codifica un polipéptido ha sido detectada con la sonda(s), la secuencia puede aislarse o clonarse por utilización de técnicas que se conocen por aquellos de habilidad común en la técnica (ver, por ejemplo, Sambrook et al., 1989, supra).

[0045] Tal y como se define aquí, un polipéptido "aislado" es un polipéptido que está esencialmente libre de otros polipéptidos de no proteasa, por ejemplo, al menos aproximadamente 20% puro, preferiblemente al menos aproximadamente 40% puro, más preferiblemente aproximadamente 60% puro, incluso más preferiblemente aproximadamente 80% puro, de forma más preferida aproximadamente 90% puro, e incluso de forma más preferida aproximadamente 95% puro, como se determina por SDS-PAGE.

[0046] Polipéptidos codificados por secuencias de ácidos nucleicos de la presente invención también incluyen polipéptidos fundidos o polipéptidos de fusión divisibles en el que otro polipéptido se funde al N-terminal o C-terminal del polipéptido o fragmento de este. Un polipéptido fusionado se produce por fusión de una secuencia de ácido nucleico (o una parte de esta) codificando otro polipéptido para una secuencia de ácido nucleico (o una parte de esta) de la presente invención. Técnicas para producir polipéptidos de fusión se conocen en la técnica, e incluyen ligar las secuencias de codificación que codifican los polipéptidos, de modo que estas están en marco y que la expresión del polipéptido fusionado está bajo control del mismo promotor(es) y terminador.

[0047] En particular, los polipéptidos son ácido-estables. Para los presentes fines, el término ácido-estable significa que la actividad residual después de 2 horas de incubación a pH 2,0, pH 2,5 o pH 3,0 y 37°C, es al menos 50%, en comparación con la actividad residual de una muestra correspondiente incubada durante 2 horas a pH 9,0 y 5°C. En una forma de realización particular, la actividad residual es al menos 55%, 60%, 65%, 70%, 75%, 80%, 85%, o al menos 90%. Un ensayo adecuado para determinar estabilidad de ácido es el ensayo de estabilidad de pH del ejemplo 2.

[0048] También, los polipéptidos son alcali-estables. Para los presentes fines, el término alcali-estable significa que la actividad residual después de 2 horas de incubación a pH 12,0 y 37°C, es al menos 85%, en comparación con la actividad residual de una muestra correspondiente incubada durante 2 horas a pH 9,0 y 5°C. En una forma de

realización particular, la actividad residual es al menos 86%, 87%, 88%, 89%, 90%, 91%, o al menos 92%. Un ensayo adecuado para determinar estabilidad de álcali es el ensayo de estabilidad de pH del ejemplo 4.

5 [0049] Además, los polipéptidos tienen i) una actividad relativa a 15°C y pH 9 de al menos 0,02, 0,04, 0,06, 0,08, 0,10, o al menos 0,11 ii) una actividad relativa a 25°C y pH 9 de al menos 0,05, 0,10, 0,15, o al menos 0,17 y/o iii) una actividad relativa a 37°C y pH 9 de al menos 0,05, 0,10, 0,15, 0,20, 0,25, o al menos 0,30. La prueba de perfil de temperatura del ejemplo 4 se usa para estas determinaciones.

[0050] Además, los polipéptidos tienen una T_m , como determinada por DSC, de al menos 76,6°C, o de al menos 77, 78, o al menos 78,2°C. T_m se determina a pH 7,0 como se describe en el ejemplo 7.

10 [0051] Adicionalmente, la proteasa puede mostrar una actividad específica en la hemoglobina a pH 7,5 y 25°C de al menos 38,4 AU/g. La actividad específica se puede determinar como se describe en el ejemplo 5. La proteasa puede mostrar una actividad específica de al menos 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 49,8, o al menos 50 AU/g.

15 [0052] Además, la proteasa, es capaz de mejorar, por lo menos 13% en comparación con el blanco, el nivel de proteína digerible de una dieta de comida de maíz/semilla de soja (SBM:Maíz = 2:3 (p/p)) en un modelo de digestión monogástrico *in vitro*. El modelo incluye un paso de digestión gástrica (1,0 hora a pH 3,0 y 40°C), y un paso de digestión intestinal posterior (4,5 horas a pH 6,8 y 40°C). El modelo también incluye adición de pepsina (3000 U/g, en el paso de digestión gástrica) y de pancreatina (8 mg/g, en el paso de digestión intestinal). La dosificación de proteasa es 100 mg de proteína de enzima proteasa (EP) por kg de dieta. Un modelo adecuado se describe en el ejemplo 8. El nivel de mejora puede ser al menos 14%, 15%, o al menos 16%.

20 [0053] La invención excluye la proteasa derivada de (i) *Nocardiosis dassonvillei* NRRL 18133 que se describe en WO 88/03947; (ii) *Nocardiosis sp.* cepa OPC-210 (FERM P-10508) que se describe en JP 2-255081-A; y/o (iii) la proteasa derivada de cepa ZIMET 43647 de la especie *Nocardiosis dassonvillei* que se describe en DD 200432|8.

Secuencias de ácidos nucleicos

25 [0054] La presente invención también se refiere a secuencias de ácidos nucleicos aisladas que codifican un polipéptido de la presente invención. Secuencias de ácidos nucleicos particulares de la invención son (i) nucleótidos 1-1143, 1-87, 88-567 y 568-1143 de la SEC ID n°: 1.

[0055] Nucleótidos particularmente preferidos son nucleótidos 568-1143 de la SEC ID n°: 1 correspondientes a las partes de codificación de polipéptidos maduros o regiones.

30 [0056] La presente invención también incluye secuencias de ácidos nucleicos que codifican un polipéptido con las partes maduras de la SEC ID n°: 2, que difieren de las partes correspondientes de la SEC ID n°: 1, en virtud de la degeneración del código genético. La presente invención también se refiere a subsecuencias de la SEC ID n°: 1, que codifica n fragmentos de la SEC ID n°: 2, y que tiene actividad de proteasa.

35 [0057] Una subsecuencia de SEC ID n°: 1, es una secuencia de ácidos nucleicos incluida por la SEC ID n°: 1, excepto que uno o más nucleótidos del extremo 5' y/o 3' han sido eliminados. Preferiblemente, una subsecuencia contiene al menos 100, 125, 150, 175, 200, o al menos 225 nucleótidos, más preferiblemente al menos 300 nucleótidos, incluso más preferiblemente al menos 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, o al menos 560 nucleótidos.

[0058] La presente invención también se refiere a secuencias de nucleótidos que tienen un grado de identidad a las partes de codificación de péptidos maduros de la SEC ID n°: 1, de al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o al menos 99%.

40 [0059] Para determinar el grado de identidad de nucleótido, el programa "align" se usa el que es referido anteriormente.

45 [0060] La presente invención también se refiere a secuencias de ácidos nucleicos mutantes comprendiendo al menos una mutación en cualquiera de los nucleótidos catalogados arriba, preferiblemente las partes de codificación de péptidos maduras de estos, en las que la secuencia de ácidos nucleicos mutante codifica un polipéptido que (i) consiste en las secuencias de aminoácidos de la SEC ID n°: 2, preferiblemente las partes de péptido maduras de estas, o (ii) es una variante alélica de cualquiera de las secuencias de (i), o (ii) es un fragmento de cualquiera de las secuencias de (i).

[0061] Las técnicas usadas para aislar o clonar una secuencia de ácidos nucleicos que codifica un polipéptido se conocen en la técnica e incluyen aislamiento de ADN genómico, preparación de ADNc, o una combinación de estos.

50 La clonación de las secuencias de ácidos nucleicos de la presente invención de tal ADN genómico puede ser efectuada, por ejemplo, usando la bien conocida reacción en cadena de la polimerasa (PCR) o selección de anticuerpos de genotecas para detectar fragmentos de ADN clonados con características estructurales compartidas.

Ver, por ejemplo, Innis et al., 1990, *PCR: A Guide to Methods and Application*, Academic Press, Nueva York. Otros procedimientos de amplificación de ácido nucleico como reacción en cadena de la ligasa (LCR), transcripción activada ligada (LAT) y amplificación a base de secuencia de ácido nucleico (NASBA) pueden utilizarse. La secuencia de ácidos nucleicos se puede clonar de una cepa de *Nocardiosis*, u otra u organismo relacionado y así, por ejemplo, puede ser una variante alélica o de especie de la región de codificación de polipéptido de la secuencia de ácidos nucleicos.

[0062] El término "secuencia de ácidos nucleicos aislada" como se utiliza en este caso se refiere a una secuencia de ácidos nucleicos que está esencialmente libre de otras secuencias de ácidos nucleicos, por ejemplo, al menos aproximadamente 20% puro, preferiblemente al menos aproximadamente 40% puro, más preferiblemente al menos aproximadamente 60% puro, incluso más preferiblemente al menos aproximadamente 80% puro, y de forma más preferida al menos aproximadamente 90% puro, como se determina por electroforesis de agarosa. Por ejemplo, una secuencia de ácidos nucleicos aislada se puede obtener por procedimientos de clonación estándar usados en ingeniería genética para recolocar la secuencia de ácidos nucleicos de su ubicación natural a un sitio diferente donde este será reproducido. Los procedimientos de clonación pueden implicar escisión y aislamiento de un fragmento de ácido nucleico deseado comprendiendo la secuencia de ácidos nucleicos que codifica el polipéptido, inserción del fragmento en una molécula de vector, e incorporación del vector recombinante en una célula huésped donde copias múltiples o clones de la secuencia de ácidos nucleicos se replicarán. La secuencia de ácidos nucleicos puede ser de origen genómico, ADNc, ARN, semisintético, sintético, o cualquier combinación de estos.

[0063] Modificación de una secuencia de ácidos nucleicos que codifica un polipéptido de la presente invención puede ser necesaria para la síntesis de polipéptidos sustancialmente similares al polipéptido. El término "sustancialmente similar" al polipéptido se refiere a de forma de origen no natural o formas del polipéptido. Estos polipéptidos puede diferir en alguna manera creada genéticamente del polipéptido aislado de su fuente nativa, por ejemplo, variantes que difieren en la actividad específica, termoestabilidad, pH óptimo, alergenicidad, o similares. La secuencia variante puede construirse basándose en la secuencia de ácidos nucleicos presentada como la parte de codificación de polipéptido, o parte de codificación de péptido madura, de la SEC ID n°: 1, por ejemplo, una subsecuencia de esta, y/o por introducción de sustituciones de nucleótido que no dan lugar a otra secuencia de aminoácidos del polipéptido codificadas por la secuencia de ácidos nucleicos, pero que corresponden al uso de codón del organismo huésped destinado para la producción de la proteasa, o por introducción de sustituciones de nucleótido que puede dar lugar a una secuencia de aminoácidos diferente. Para una descripción general de sustitución de nucleótido, ver, por ejemplo, Ford et al., 1991, *Protein Expression and Purification 2: 95-107*. Polipéptidos hipoalergénicos pueden, por ejemplo, prepararse como se ha descrito anteriormente.

[0064] Será aparente a los expertos en la técnica que tales sustituciones pueden hacerse fuera de las regiones importantes a la función de la molécula y todavía suponer un polipéptido activo. Residuos de aminoácidos esenciales para la actividad del polipéptido codificados por la secuencia de ácidos nucleicos aislada de la invención y, por lo tanto, preferiblemente no sujetos a sustitución, se pueden identificar según procedimientos conocidos en la técnica, como mutagénesis dirigida o mutagénesis de escaneado de alanina (ver, por ejemplo, Cunningham and Wells, 1989, *Science 244: 1081-1085*). En esta técnica, mutaciones se introducen en cada residuo positivamente cargado en la molécula, y las moléculas mutantes resultantes se evalúan para actividad de proteasa para identificar residuos de aminoácidos que son importantes para la actividad de la molécula. Sitios de interacción de sustrato de proteasa pueden también determinarse por análisis de la estructura tridimensional como determinado por tales técnicas como análisis de resonancia magnética nuclear, cristalografía o etiquetado de fotoafinidad (ver, por ejemplo, de Vos et al., 1992, *Science 255: 306-312* ; Smith et al., 1992, *Journal of Molecular Biology 224: 899-904* ; Wlodaver et al., 1992, *FEBS Letters 309: 59-64*).

Métodos para producir secuencias de ácidos nucleicos mutantes

[0065] La presente invención además se refiere a métodos para producir una secuencia de ácidos nucleicos mutante, comprendiendo introducción de al menos una mutación en las partes de codificación de polipéptidos maduros de la SEC ID n°: 1, o una subsecuencia de la misma, donde la secuencia de ácidos nucleicos mutante codifica un polipéptido que consiste de los péptidos maduros de la SEC ID n° 2 o un fragmento de estos que tiene actividad de proteasa.

[0066] La introducción de una mutación en la secuencia de ácidos nucleicos para intercambiar un nucleótido por otro nucleótido se puede realizar por mutagénesis dirigida que usa cualquiera de los métodos conocidos en la técnica. Particularmente útil es el procedimiento que utiliza un vector de ADN bicatenario superenrollado con un inserto de interés y dos cebadores sintéticos con la mutación deseada. Los cebadores oligonucleótidos, cada complementario a hilos opuestos del vector, se extienden durante variación cíclica de la temperatura mediante Pfu ADN polimerasa. En la incorporación de los cebadores, un plásmido mutado con cortes en bisel se genera. Después del ciclo de temperatura, el producto se trata con DpnI que es específico para ADN hemimetilado y metilado para digerir el molde de ADN parental y para seleccionar para ADN sintetizado con mutación. Otros procedimientos conocidos en la técnica puede usarse también.

Construcciones de ácidos nucleicos

[0067] La presente invención también se refiere a construcciones de ácidos nucleicos comprendiendo una secuencia de ácidos nucleicos de la presente invención unidos operativamente a una o más secuencias de control que dirigen la expresión de la secuencia codificante en una célula huésped adecuada bajo condiciones compatibles con las secuencias de control. Expresión se entenderá que incluye cualquier paso implicado en la producción del polipéptido incluyendo, pero no limitado a, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

[0068] "Construcción de ácido nucleico" se define aquí como una molécula de ácido nucleico, o bien bicatenaria o monocatenaria, que se aísla de un gen de origen natural o que ha sido modificado para contener segmentos de ácido nucleico combinados y superpuesto en cierto modo que de otra manera no existiría en la naturaleza. El término construcción de ácidos nucleicos es sinónimo con la expresión de término *cassette* cuando la construcción de ácidos nucleicos contiene todas las secuencias de control requeridas para la expresión de una secuencia codificante de la presente invención. El término "secuencia codificante" se define aquí como una secuencia de ácidos nucleicos que especifica directamente la secuencia de aminoácidos de su producto proteínico. Los límites de la secuencia codificante se determinan generalmente por un sitio de unión del ribosoma (procariotas) o por codón de inicio ATG (eucariotas) localizados justo arriba del marco de lectura abierto en el extremo 5' del ARNm y una secuencia del terminador de transcripción localizada justo abajo del marco de lectura abierto en el extremo 3' del ARNm. Una secuencia codificante puede incluir, pero no se limita a, ADN, ADNc y secuencias de ácidos nucleicos recombinantes.

[0069] Una secuencia de ácidos nucleicos que codifica aislada un polipéptido de la presente invención se puede manipular en una variedad de vías para proveer a expresión del polipéptido. Manipulación de la secuencia de ácidos nucleicos antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar secuencias de ácidos nucleicos utilizando métodos de ADN recombinante se conocen bien en la técnica.

[0070] El término "secuencias de control" se define aquí para incluir todos componentes que son necesarios o ventajosos para la expresión de un polipéptido de la presente invención. Cada secuencia de control puede ser nativa o extranjera a la secuencia de ácidos nucleicos que codifica el polipéptido. Tales secuencias de control incluyen, pero de forma no limitativa, una líder, secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal y terminador de transcripción. A un mínimo, las secuencias de control incluyen un promotor y señales de parada traduccionales y transcripcionales. Las secuencias de control se pueden proporcionar con conectores con motivo de introducir sitios de restricción específicos facilitando ligamiento de las secuencias de control con la región de codificación de la secuencia de ácidos nucleicos que codifica un polipéptido. El término "unido operativamente" se define aquí como una configuración en la que una secuencia de control se coloca apropiadamente en una posición relativa a la secuencia codificante de la secuencia de ADN, de manera que la secuencia de control dirige la expresión de un polipéptido.

[0071] La secuencia de control puede ser una secuencia del promotor apropiada, una secuencia de ácidos nucleicos que es reconocida por una célula huésped para la expresión de la secuencia de ácidos nucleicos. La secuencia del promotor contiene secuencias de control transcripcionales que median la expresión del polipéptido. El promotor puede ser cualquier secuencia de ácidos nucleicos que muestra actividad transcripcional en la célula huésped de elección, incluyendo mutante, truncado y promotores híbridos, y se pueden obtener de genes que codifican polipéptidos intracelulares o extracelulares bien homólogos o heterólogos a la célula huésped.

[0072] Ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácidos nucleicos de la presente invención, especialmente en una célula huésped bacteriana, son los promotores obtenido de la *E. coli lac operon*, gen de agarasa de *Streptomyces coelicolor* (dagA), gen de levansucrasa de *Bacillus subtilis* (sacB), gen de alfa-amilasa de *Bacillus licheniformis* (amyL), gen de amilasa maltogénica de *Bacillus stearothermophilus* (amyM), gen de alfa-amilasa de *Bacillus amyloliquefaciens* (amyQ), gen de penicilinas de *Bacillus licheniformis* (penP), genes xylA y xylB de *Bacillus subtilis* y gen de beta lactamasa procariótica (cVilla-Kamaroff et al., 1978, *Proceedings of the National Academy of Sciences USA 75: 3727-3731*), al igual que el promotor tac (DeBoer et al., 1983, *Proceedings of the National Academy of Sciences USA 80: 21-25*). Además promotores se describen en "Useful proteins from recombinant bacteria" en Scientific American, 1980, 242: 74-94 ; y en Sambrook et al., 1989, supra.

[0073] Ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácidos nucleicos de la presente invención en una célula huésped filamentosa fúngica son promotores obtenidos de los genes para TAKA amilasa de *Aspergillus oryzae*, proteinasa aspártica de *Rhizomucor miehei*, alfa-amilasa neutra *Aspergillus niger*, ácido estable de alfa amilasa de *Aspergillus niger*, glucoamilasa de *Aspergillus niger* o de *Aspergillus awamori* (glaA), lipasa de *Rhizomucor miehei*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, acetamidasa de *Aspergillus nidulans* y proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), al igual que el promotor NA2-tpi (un híbrido de los promotores de los genes para alfa-amilasa neutra de *Aspergillus niger* y triosa fosfato isomerasa de *Aspergillus oryzae*) y promotores mutantes, truncados e híbridos de estos.

- 5 [0074] En un huésped de levadura, promotores útiles se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO 1), galactocinasa de *Saccharomyces cerevisiae* (GAL1), alcohol dehidrogenasa/gliceraldehído-3-fosfato dehidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP) y 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para células huésped de levadura son descritos por Romanos et al., 1992, *Yeast* 8: 423-488.
- [0075] La secuencia de control puede también ser una secuencia del terminador de transcripción adecuada, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia del terminador se une operativamente al 3' terminal de la secuencia de ácidos nucleicos que codifica el polipéptido. Cualquier terminador que es funcional en la célula huésped de elección se puede utilizar en la presente invención.
- 10 [0076] Terminadores preferidos para células huésped filamentosas fúngicas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, alfa-glucosidasa de *Aspergillus niger* y proteasa de tipo tripsina de *Fusarium oxysporum*.
- [0077] Terminadores preferidos para células huésped de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo de *Saccharomyces cerevisiae* C (CYC1) y gliceraldehído-3-fosfato-deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles para células huésped de levadura se describen por Romanos et al., 1992, supra.
- 15 [0078] Terminadores preferidos para células huésped bacterianas, como una célula huésped de *Bacillus*, son los terminadores de gen de alfa-amilasa de *Bacillus licheniformis* (amyL), el gen de amilasa maltogénica de *Bacillus stearothermophilus* (amyM), o el gen de alfa-amilasa de *Bacillus amyloliquefaciens* (amyQ).
- 20 [0079] La secuencia de control puede también ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para traducción por la célula huésped. La secuencia líder se une operativamente al 5' terminal de la secuencia de ácidos nucleicos que codifica el polipéptido. Cualquier secuencia líder que es funcional en la célula huésped de elección se puede utilizar en la presente invención.
- [0080] Líderes preferidos para células huésped filamentosas fúngicas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.
- 25 [0081] Líderes adecuados para células huésped de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO 1), 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*, alfa-factor de *Saccharomyces cerevisiae*, y alcohol dehidrogenasa/gliceraldehído-3-fosfato dehidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP).
- 30 [0082] La secuencia de control puede también ser una secuencia de poliadenilación, una secuencia operativamente unida al 3' terminal de la secuencia de ácidos nucleicos y que, cuando se transcribe, es reconocida por la célula huésped como una señal para añadir residuos de poliadenosina para ARNm transcrito. Cualquier secuencia de poliadenilación que es funcional en la célula huésped de elección se puede utilizar en la presente invención.
- [0083] Secuencias de poliadenilación preferidas para células huésped filamentosas fúngicas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, proteasa de tipo tripsina de *Fusarium oxysporum* y alfa-glucosidasa de *Aspergillus niger*.
- 35 [0084] Secuencias de poliadenilación útiles para células huésped de levadura se describen por Guo and Sherman, 1995, *Molecular Cellular Biology* 15: 5983-5990.
- [0085] La secuencia de control puede también ser una región de codificación de péptido señal que codifica para una secuencia de aminoácidos única al amino terminal de un polipéptido y dirige el polipéptido codificado en la vía secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de ácidos nucleicos puede intrínsecamente contener una región de codificación de péptido de señal naturalmente única en el marco de lectura de traducción con el segmento de la región de codificación que codifica el polipéptido secretado. Alternativamente, el extremo 5' de la secuencia codificante puede contener una región de codificación de péptido de señal que es extranjero a la secuencia codificante. La región de codificación del péptido de señal foráneo puede requerirse donde la secuencia codificante naturalmente no contiene una región de codificación de péptido de señal. Alternativamente, la región de codificación de péptido de señal foráneo puede simplemente reemplazar la región de codificación del péptido señal natural para mejorar secreción del polipéptido. No obstante, cualquier región de codificación del péptido de señal que dirige el polipéptido expresado en la vía secretora de una célula huésped de elección se puede utilizar en la presente invención.
- 40 [0086] Regiones de codificación de péptido de señal eficaces para células huésped bacterianas son las regiones de codificación de péptido de señal obtenidas de los genes para amilasa maltogénica de *Bacillus* NCIB 11837, alfa-amilasa de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, alfa-amilasa de *Bacillus licheniformis*, proteasas neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM) y *Bacillus subtilis* prsA. Además péptidos de señal se describen por Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137.
- 45 [0087] Secuencias de control que son reconocidas por una célula huésped de elección para iniciar la transcripción de un ARNm se describen en el presente documento.
- 50 [0088] Secuencias de control que son reconocidas por una célula huésped de elección para iniciar la transcripción de un ARNm se describen en el presente documento.
- 55 [0089] Secuencias de control que son reconocidas por una célula huésped de elección para iniciar la transcripción de un ARNm se describen en el presente documento.

[0087] Regiones de codificación de péptido de señal eficaces para células huésped filamentosas fúngicas son las regiones de codificación de péptido de señal obtenidas de los genes para TAKA amilasa de *Aspergillus oryzae*, amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, proteinasa aspártica de *Rhizomucor miehei*, celulasa de *Humicola insolens* y lipasa *Humicola lanuginosa*.

- 5 [0088] Péptidos de señal útiles para células huésped de levadura se obtienen de los genes para factor alfa de *Saccharomyces cerevisiae* invertasa de *Saccharomyces cerevisiae*. Otras regiones de codificación de péptido de señal útiles se describen por Romanos et al., 1992, supra.

[0089] La secuencia de control puede también ser una región de codificación de propéptido que codifica para una secuencia de aminoácidos situada en el amino terminal de un polipéptido. El polipéptido resultante es conocido como una proenzima o propolipéptido (o un zimógeno en algunos casos). Un propolipéptido es generalmente inactivo y se puede convertir a un polipéptido maduro activo por escisión autocatalítica o catalítica del propéptido del propolipéptido. La región de codificación de propéptido se puede obtener de los genes para proteasa alcalina de *Bacillus subtilis* (aprE), proteasa neutra de *Bacillus subtilis* (nprT), alfa-factor de *Saccharomyces cerevisiae*, proteinasa aspártica de *Rhizomucor miehei* y lacasa de *Myceliophthora thermophila* (WO 95/33836).

- 15 [0090] Preferiblemente, la región de codificación de péptido de señal es la región de codificación del péptido de señal de la SEC ID nº: 1

[0091] También preferiblemente, la región de codificación de propéptido es las regiones de codificación de propéptido de la SEC ID nº: 1.

- 20 [0092] Donde ambos péptido de señal y regiones de propéptido están presentes en el amino terminal de un polipéptido, la región de propéptido está situada junto al amino terminal de un polipéptido y la región de péptido señal está situada junto al amino terminal de la región de propéptido.

[0093] Puede también ser deseable para añadir secuencias reguladoras que permiten la regulación de la expresión del polipéptido relativas al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son los que causan que la expresión del gen se gire ON u OFF en respuesta a un estímulo físico o químico, incluyendo la presencia de un compuesto regulador. Sistemas reguladores en sistemas procariotas incluyen los sistemas operadores lac, tac y trp. En la levadura, el sistema ADH2 o sistema GAL1 puede utilizarse. En hongos filamentosos, el promotor de TAKA alfa-amilasa, promotor de glucoamilasa de *Aspergillus niger* y promotor de glucoamilasa de *Aspergillus oryzae* se pueden utilizar como secuencias reguladoras. Otros ejemplos de secuencias reguladoras son las que permiten amplificación génica. En sistemas eucariotas, estas incluyen el gen de dihidrofolato-reductasa que se amplifica en presencia de metotrexato y los genes de metalotioneína que se amplifican con metales pesados. En estos casos, la secuencia de ácidos nucleicos que codifica el polipéptido se uniría operativamente con la secuencia reguladora.

Vectores de expresión

[0094] La presente invención también se refiere a vectores de expresión recombinantes comprendiendo una secuencia de ácidos nucleicos de la presente invención, un promotor y señales de parada traduccionales y transcripcionales. Las varias secuencias de ácidos nucleicos y de control descritas anteriormente pueden juntarse para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir inserción o sustitución de la secuencia de ácidos nucleicos que codifica el polipéptido en tales sitios. Alternativamente, la secuencia de ácidos nucleicos de la presente invención se puede expresar por inserción de la secuencia de ácidos nucleicos o una construcción de ácidos nucleicos comprendiendo la secuencia en un vector apropiado para la expresión. En la creación del vector de expresión, la secuencia codificante se localiza en el vector de modo que la secuencia codificante se une operativamente con las secuencias de control apropiadas para la expresión.

[0095] El vector de expresión recombinante puede ser cualquier vector (p. ej., un plásmido o virus) que puede ser convenientemente sometido a procedimientos de ADN recombinante y puede provocar la expresión de la secuencia de ácidos nucleicos. La elección del vector típicamente dependerá de la compatibilidad del vector con la célula huésped en la que el vector debe ser introducido. Los vectores puede ser plásmidos lineales o circulares cerrados.

[0096] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, la replicación del cual es independiente de replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma, o un cromosoma artificial. El vector puede contener cualquier medio para asegurar autorreplicación. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica con la cromosoma(es) en los que ha sido integrado. Además, un único vector o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total para ser introducido en el genoma de la célula huésped, o un transposón pueden utilizarse.

[0097] Los vectores de la presente invención preferiblemente contienen uno o más marcadores seleccionables que permiten selección fácil de células transformadas. Una etiqueta seleccionable es un gen, el producto del cual proporciona resistencia vírica o biocida, resistencia a metales pesados, prototrofia para auxótrofos y similares. Ejemplos de marcadores seleccionables bacterianos son los genes de *Bacillus subtilis* o *Bacillus licheniformis*, o

5 marcadores que confieren resistencia antibiótica, como ampicilina, canamicina, cloranfenicol o resistencia a tetraciclina. Marcadores adecuados para células huésped de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3. Marcadores seleccionables para uso en una célula huésped filamentosa fúngica incluyen, pero de forma no limitativa, amdS (acetamidasa), argB (ornitina carbamoiltransferasa), bar (fosfonitrilina acetiltransferasa), hygB (higromicina fosfotransferasa), niaD (nitrato-reductasa), pyrG (orotidina-5'-fosfato-descarboxilasa), sC (adeniltransferasa de sulfato), trpC (sintasa de antranilato), al igual que equivalentes de estos. Preferido para uso en una célula de *Aspergillus* son los genes amdS y pyrG de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen bar de *Streptomyces higroscopicus*.

10 [0098] Los vectores de la presente invención preferiblemente contienen un elemento(s) que permite integración estable del vector en el genoma huésped de la célula o replicación autónoma del vector en la célula independiente del genoma.

15 [0099] Para integración en el genoma de la célula huésped, el vector puede depender de la secuencia de ácidos nucleicos que codifica el polipéptido o cualquier otro elemento del vector para integración estable del vector en el genoma por recombinación homóloga o no homóloga. Alternativamente, el vector puede contener secuencias de ácidos nucleicos adicionales para dirigir integración por recombinación homóloga en el genoma de la célula huésped. Las secuencias de ácidos nucleicos adicionales permiten al vector ser integrado en el genoma de la célula huésped en una ubicación(es) precisa en el cromosoma(s). Para aumentar la probabilidad de integración en una ubicación precisa, los elementos integracionales deberían preferiblemente contener un número suficiente de ácidos nucleicos, como de 100 a 1500 pares de bases, preferiblemente de 400 a 1500 pares de bases, y de forma más preferida de 800 a 1500 pares de bases, que son altamente homólogos a la secuencia objetivo correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que es homóloga a la secuencia objetivo en el genoma de la célula huésped. Además, los elementos integracionales pueden ser secuencias de ácidos nucleicos codificantes o no codificantes. Por otro lado, el vector se puede integrar en el genoma de la célula huésped por recombinación no homóloga.

25 [0100] Para replicación autónoma, el vector puede comprender además un origen de replicación permitiendo al vector replicar de manera autónoma en la célula huésped en cuestión. Ejemplos de orígenes bacterianos de replicación son los orígenes de replicación de plásmidos pBR322, pUC19, pACYC177 y pACYC184 permitiendo replicación en *E. coli*, y pUB110, pE194; pTA1060 y pAMR1 permitiendo replicación en *Bacillus*. Ejemplos de orígenes de replicación para uso en una célula huésped de levadura son el origen de micra de replicación 2, ARS1, ARS4, la combinación de ARS1 y CEN3 y la combinación de ARS4 y CEN6. El origen de replicación puede ser uno con una mutación que hace su funcionamiento termosensible en la célula huésped (ver, por ejemplo, Ehrlich, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1433).

35 [0101] Más de una copia de una secuencia de ácidos nucleicos de la presente invención se puede insertar en la célula huésped para aumentar producción del producto genético. Un aumento en el número de copia de la secuencia de ácidos nucleicos se puede obtener por integración de al menos una copia adicional de la secuencia en el genoma de la célula huésped o incluyendo un gen marcador seleccionable amplificable con la secuencia de ácidos nucleicos donde células con copias amplificadas del gen marcador seleccionable, y así copias adicionales de la secuencia de ácidos nucleicos, se pueden seleccionar para cultivo de las células en presencia del agente apropiado seleccionable.

40 [0102] Los procedimientos usados para unir los elementos anteriormente descritos para construcción de los vectores de expresión recombinantes de la presente invención se conocen por un experto en la técnica (ver, por ejemplo, Sambrook et al., 1989, supra).

[0103] La proteasa puede también ser coexpresada junto con al menos otra enzima de interés para alimento para animales, como una amilasa, fitasa, xilanasas, galactanasas, alfa-galactosidasas, proteasa, fosfolipasas, y/o una beta-galactanasas.

45 [0104] Las enzimas se pueden coexpresar de distintos vectores, de un vector, o usando una mezcla de ambas técnicas. Cuando se usan vectores diferentes, los vectores pueden tener marcadores seleccionables diferentes y orígenes de replicación diferentes. Cuando se usa solo un vector, los genes se pueden expresar de uno o más promotores. Si se clona bajo la regulación de un promotor (di- o multicistónico), la orden en la que los genes se clonan puede afectar a los niveles de expresión de las proteínas. La proteasa puede también expresarse como una proteína de fusión, es decir, que la proteasa de gen que codifica se ha fundido en marco al gen que codifica otra proteína. Esta proteína puede ser otra enzima o un dominio funcional de otra enzima.

Células huésped

55 [0105] La presente invención también se refiere a células huésped recombinantes, comprendiendo una secuencia de ácidos nucleicos de la invención, que se usan ventajosamente en la producción recombinante de los polipéptidos. Un vector comprendiendo una secuencia de ácidos nucleicos de la presente invención se introduce en una célula huésped de modo que el vector se mantiene como un integrante cromosómico o como un vector extracromosómico autorreplicante como se describe anteriormente. El término "célula huésped" encierra cualquier progenie de una

célula madre que no es idéntica a la célula madre debido a mutaciones que ocurren durante replicación. La elección de una célula huésped en gran parte depende del gen que codifica el polipéptido y su fuente.

[0106] La célula huésped puede ser un microorganismo unicelular, por ejemplo, un procarionta, o un microorganismo no unicelular, por ejemplo, un eucariota.

5 [0107] Células unicelulares útiles son células bacterianas como bacterias gram positivas incluyendo, pero no limitado a, una célula de *Bacillus*, o una célula de *Streptomyces*, o células de bacterias del ácido láctico; o bacterias gram negativas, como *E. coli* y *Pseudomonas* sp. Bacterias del ácido láctico incluyen, pero de forma no limitativa, especies de los géneros *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Streptococcus*, *Pediococcus* y *Enterococcus*.

10 [0108] La introducción de un vector en una célula huésped bacteriana puede, por ejemplo, efectuarse por transformación de protoplasto (ver, por ejemplo, Chang y Cohen, 1979, *Molecular General Genetics* 168: 111-115), usando células competentes (ver, por ejemplo, Young y Spizizin, 1961, *Journal of Bacteriology* 81: 823-829, o Dubnau y Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221), electroporación (ver, por ejemplo, Shigekawa y Dower, 1988, *Biotechniques* 6: 742-751), o conjugación (ver, por ejemplo, Shigekawa y Dower, 1988, *Biotechniques* 6: 742-751).

15 [0109] La célula huésped puede ser una eucariota, como una célula animal no humana, una célula de insecto, una célula vegetal o una célula fúngica.

20 [0110] En una forma de realización particular, la célula huésped es una célula fúngica. "Hongos" como se utiliza en este caso incluye el filo *Ascomycota*, *Basidiomycota*, *Chytridiomycota* y *Zygomycota* (como definido por Hawksworth et al., en *Ainsworth and Bisby's Dictionary of The Fungi, 8th edition*, 1995, CAB International, University Press, Cambridge, UK) al igual que la *Oomycota* (como citado en Hawksworth et al., 1995, supra, page 171) y todos los hongos mitospóricos (Hawksworth et al., 1995, supra).

25 [0111] En otra forma de realización particular, la célula huésped fúngica es una célula de levadura. "Levadura", como se usa aquí, incluye levadura de ascosporogenous (Endomycetales), levadura de basidiosporogenous y levadura de *Fungi Imperfecti* (Blastomycetes). Ya que la clasificación de levadura puede cambiar en el futuro, para los fines de esta invención, levadura debe definirse como descrita en *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., y Davenport, R.R., eds, *Soc. App. Bacteriol. Symposium Series No. 9, 1980*).

[0112] La célula huésped de levadura puede ser una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*.

30 [0113] La célula huésped fúngica puede ser una célula micótica filamentosa. "Hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión *Eumycota* y *Oomycota* (como definido por Hawksworth et al., 1995, supra). Los hongos filamentosos se caracterizan por el hecho de que tienen una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. Crecimiento vegetativo es por alargamiento hifal y catabolismo de carbono es estrictamente aeróbico. En cambio, crecimiento vegetativo por levaduras como *Saccharomyces cerevisiae* es por injerto de un talo unicelular y catabolismo de carbono puede ser fermentativo.

35 Ejemplos de células huésped filamentosas fúngicas son células de especies de, pero no limitado a, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolyopcladium*, o *Trichoderma*.

40 [0114] Células fúngicas se pueden transformar por un proceso implicando formación de protoplasto, transformación de los protoplastos y regeneración de la pared celular en una manera conocida *per se*. Procedimientos adecuados para transformación de células de *Aspergillus* huésped se describen en EP 238 023 y Yelton et al., 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470-1474. Métodos adecuados para transformación de especies de *Fusarium* se describen por Malardier et al., 1989, *Gene* 78: 147-156 y WO 96/00787. Levadura puede transformarse usando los procedimientos descritos por Becker y Guarente, en *Abelson, J.N. y Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187*, Academic Press, Inc., New York; Ito et al., 1983, *Journal of Bacteriology* 153: 163; y Hinnen et al., 1978, *Proceedings of the National Academy of Sciences USA* 75: 1920.

Métodos de producción

50 [0115] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención comprendiendo (a) cultivo de una cepa, que en su forma de tipo salvaje es capaz de producir el polipéptido, para producir un sobrenadante comprendiendo el polipéptido; y (b) recuperación del polipéptido. En una forma de realización preferida, la cepa es de filo *Actinobacteria*, preferiblemente de la clase *Actinobacteria*, más preferiblemente del orden *Actinomycetales*, incluso más preferiblemente de la familia *Nocardioseae*, y de forma más preferida del género *Nocardioseae*, por ejemplo, cualquiera de la especie *Nocardioseae*, como las cepas específicas catalogadas precedentes.

[0116] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención comprendiendo (a) cultivo de una célula huésped bajo condiciones propicias para la producción del polipéptido; y (b) recuperación el polipéptido.

5 [0117] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención comprendiendo (a) cultivo de una célula huésped bajo condiciones propicias para la producción del polipéptido, donde la célula huésped comprende una secuencia de ácidos nucleicos mutante comprendiendo al menos una mutación en las partes de codificación de péptidos maduros de la SEC ID n°: 1, en la que la secuencia de ácidos nucleicos mutante codifica un polipéptido que (i) consiste en los péptidos maduros de la SEC ID n°: 2, o (ii) es una variante alélica de cualquiera de las secuencias de (i), o (ii) es un fragmento de cualquiera de las secuencias de (i).

10 [0118] En los métodos de producción de la presente invención, las células se cultivan en un medio nutritivo adecuado para la producción de un polipéptido usando los métodos conocidos en la técnica. Por ejemplo, la célula se puede cultivar por cultivo en matraz de agitación, fermentación a gran escala o a escala pequeña (incluyendo fermentaciones de estado sólido continua, por lotes o por alimentación por lotes) en el laboratorio o fermentadores industriales realizados en un medio adecuado y bajo condiciones que permiten al polipéptido expresarse y/o aislarse. El cultivo se desarrolla en un medio nutritivo adecuado comprendiendo fuentes de nitrógeno y carbono y sales inorgánicas, usando procedimientos conocidos en la técnica. Medios adecuados están disponibles de proveedores comerciales o se pueden preparar según composiciones publicadas (p. ej., en catálogos de la American Type Culture Collection). Si el polipéptido se secreta en el medio nutritivo, el polipéptido se puede recuperar directamente del medio. Si el polipéptido no se secreta, este se puede recuperar de lisados celulares.

20 [0119] Los polipéptidos pueden detectarse usando métodos conocidos en la técnica que son específicos para los polipéptidos. Estos métodos de detección pueden incluir uso de anticuerpos específicos, formación de un producto de proteasa, o desaparición de un sustrato de proteasa. Por ejemplo, un ensayo de proteasa se puede utilizar para determinar la actividad del polipéptido como se describe en este caso.

25 [0120] El polipéptido resultante se puede recuperar por métodos conocidos en la técnica. Por ejemplo, el polipéptido puede recuperarse del medio nutritivo por procedimientos convencionales incluyendo, pero no limitado a, centrifugado, filtración, extracción, secado por pulverización, evaporación, o precipitación.

30 [0121] Los polipéptidos de la presente invención se pueden purificar por una variedad de procedimientos conocidos en la técnica incluyendo, pero no limitado a, cromatografía (p. ej., intercambio iónico, afinidad, hidrofóbica, cromatofoco y exclusión de tamaño), procedimientos electroforéticos (p. ej., isoelectrofoque preparativo), solubilidad diferencial (p. ej., precipitación con sulfato amónico), SDS-PAGE, o extracción (ver, por ejemplo, *Protein Purification, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989*).

Plantas

35 [0122] La presente invención también se refiere a una planta transgénica, parte de planta, o célula vegetal que ha sido transformada con una secuencia de ácidos nucleicos que codifica un polipéptido con actividad de proteasa de la presente invención para expresar y producir el polipéptido en cantidades recuperables. El polipéptido se puede recuperar de la planta o parte de planta. Alternativamente, la planta o parte de planta con el polipéptido recombinante se puede utilizar como tal para mejorar la calidad de un alimento o alimento, por ejemplo, mejora de valor nutritivo, palatabilidad y propiedades reológicas, o para destruir un factor antinutritivo.

40 [0123] En una forma de realización particular, el polipéptido se provee a las vacuolas de almacenamiento de endospermo en semillas. Este se puede obtener por sintetización de este como un precursor con un péptido señal adecuado, ver Horvath et al en *PNAS*, Feb. 15, 2000, vol. 97, no. 4, p. 1914-1919.

45 [0124] La planta transgénica puede ser *dicotyledonous* (una dicotiledónea) o *monocotyledonous* (una monocotiledónea) o variantes creadas genéticamente de estas. Ejemplos de plantas monocotiledóneas son hierbas, como césped de pradera (pasto azul, *Poa*), gramínea forrajera como *Festuca*, *Lolium*, césped templado, como *Agrostis*, y cereales, por ejemplo, trigo, avena, centeno, cebada, arroz, sorgo, triticale (híbrido de estabilizado trigo (*Triticum*) y centeno (*Secale*) y maíz (com). Ejemplos de planta dicotiledónea son tabaco, leguminosas, como girasol (*Helianthus*), algodón (*Gossypium*), altramuces, patata, remolacha azucarera, guisante, judía y semilla de soja y plantas crucíferas (familia *Brassicaceae*), como coliflor, semilla de colza y el organismo de modelo estrechamente relacionado *Arabidopsis thaliana*. Plantas de bajo fitato como se describe, por ejemplo, en la patente US n°. 5,689,054 y patente US n°. 6,111,168 son ejemplos de plantas creadas genéticamente.

50 [0125] Ejemplos de partes de planta son vástago, callo, hojas, raíz, frutas, semillas y tubérculos, al igual que los tejidos individuales comprendiendo estas partes, por ejemplo, epidermis, mesófilo, parénquima, tejidos vasculares, meristemas. También compartimentos de célula vegetal específicos, como cloroplasto, apoplasto, mitocondria, vacuola, peroxisomas y citoplasma se consideran que son una parte de planta. Además, cualquier célula vegetal, cualquiera que sea el origen del tejido, se considera que es una parte de planta. Asimismo, partes de planta como tejidos específicos y células aisladas para facilitar la utilización de la invención son también consideradas partes de planta, por ejemplo, embriones, endospermas, aleurona y revestimientos de semilla.

[0126] También incluido dentro del campo de la presente invención están la progenie de tales plantas, partes de planta y células vegetales.

5 [0127] La planta transgénica o célula vegetal que expresa un polipéptido de la presente invención se pueden construir conforme a métodos conocidos en la técnica. Brevemente, la planta o célula vegetal se construye por incorporación de una o más construcciones de expresión que codifican un polipéptido de la presente invención en el genoma huésped de la planta y propagando la planta modificada resultante o célula vegetal en una planta transgénica o célula vegetal.

10 [0128] Convenientemente, la construcción de expresión es una construcción de ácidos nucleicos que comprende una secuencia de ácidos nucleicos que codifica un polipéptido de la presente invención operativamente unido con secuencias reguladoras apropiadas requeridas para expresión de la secuencia de ácidos nucleicos en la planta o parte de planta de elección. Además, la construcción de expresión puede comprender una etiqueta seleccionable útil para identificación de células huésped en la que la construcción de expresión ha sido integrada y secuencias de ADN necesarias para introducción de la construcción en la planta en cuestión (esto depende del método de introducción de ADN que se usa).

15 [0129] La elección de secuencias reguladoras, como secuencias promotoras y terminadoras y opcionalmente secuencias de señal o de tránsito se determinan, por ejemplo, basándose en cuándo, dónde y cómo el polipéptido se desea que se exprese. Por ejemplo, la expresión del gen que codifica un polipéptido de la presente invención puede ser constitutiva o inducible, o puede ser desarrollable, fase o tejido específico y el producto genético puede estar dirigido a un compartimiento de célula específico, tejido o parte de planta como semillas u hojas. Secuencias reguladoras se describen, por ejemplo, por Tague et al., 1988, *Plant Physiology* 86: 506.

20 [0130] Para expresión constitutiva, los siguientes promotores pueden ser utilizados: el promotor 35S-CaMV (Franck et al., 1980, *Cell* 21: 285-294), la ubiquitina de maíz 1 (Christensen AH, Sharrock RA and Quail 1992. *GMaize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation*), o el promotor de actina de arroz 1 (*Plant Mo. Biol.* 18, 675-689; Zhang W, McElroy D. and Wu R 1991, *Analysis of rice Act1 5' region activity in transgenic rice plants. Plant Cell* 3, 1155-1165). Promotores específicos a un órgano pueden ser, por ejemplo, un promotor de tejidos sumidero de almacenamiento como semillas, tubérculos de patata y frutas (Edwards & Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), o de tejidos sumidero metabólicos como meristemas (ito et al., 1994, *Plant Mol. Biol.* 24: 863-878), un promotor específico de semilla como la glutelina, prolamina, globulina, o promotor de albúmina de arroz (Wu et al., 1998, *Plant and Cell Physiology* 39: 885-889), un promotor de *Vicia faba* de la legúmina B4 y el gen de proteína de semilla desconocida de *Vicia faba* (Conrad et al., 1998, *Journal of Plant Physiology* 152: 708-711), un promotor de una proteína de cuerpo de aceite de semilla (Chen et al., 1998, *Plant and Cell Physiology* 39: 935-941), el promotor de la proteína de almacenamiento napA de *Brassica napus*, o cualquier otro promotor específico de semilla conocido en la técnica, por ejemplo, como se describe en WO 91/14772. Además, el promotor puede ser un promotor específico de hoja como el promotor de rbcS de arroz o tomate (Kyoizuka et al., 1993, *Plant Physiology* 102: 991-1000), el promotor de gen de metiltransferasa de adenina de virus de *chlorella* (Mittra and Higgins, 1994, *Plant Molecular Biology* 26: 85-93), o el promotor aldP de gen de arroz (Kagaya et al., 1995, *Molecular and General Genetics* 248: 668-674), o un promotor inducible dañado como el promotor pin2 de patata (Xu et al., 1993, *Plant Molecular Biology* 22: 573-588). Asimismo, el promotor puede ser inducible por tratamientos abióticos como temperatura, sequía o alteraciones en la salinidad o inducibles por sustancias aplicadas exógenamente que activan el promotor, por ejemplo, etanol, estrógenos, hormonas de planta como etileno, ácido abscísico, ácido giberélico y/o metales pesados.

35 [0131] Un elemento promotor intensificador puede también usarse para conseguir expresión más alta de la proteasa en la planta. Por ejemplo, el elemento promotor intensificador puede ser un intrón que se coloca entre el promotor y la secuencia de nucleótidos que codifica un polipéptido de la presente invención. Por ejemplo, Xu et al., 1993, supra revela el uso del primer intrón del gen de actina de arroz 1 para mejorar expresión.

40 [0132] Aún más, el uso de codón se puede optimizar para las especies de planta en cuestión para mejorar expresión (ver Horvath et al referido anteriormente).

[0133] El gen marcador seleccionable y cualquier otra parte de la construcción de expresión se pueden elegir de aquellas disponibles en la técnica.

50 [0134] La construcción de ácidos nucleicos se incorpora en el genoma de la planta según técnicas convencionales conocidas en la técnica, incluyendo transformación mediada por *Agrobacterium*, transformación mediada por virus, microinyección, bombardeo de partículas, transformación biolística y electroporación (Gasser et al., 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto et al., 1989, *Nature* 338: 274).

55 [0135] Actualmente, transferencia de gen mediada por *Agrobacterium tumefaciens* es el método de elección para generar dicotiledóneas transgénicas (para una revisión, ver Hooykas and Schilperoort, 1992, *Plant Molecular Biology* 19: 15-38), y esto puede también usarse para transformación de monocotiledóneas, aunque otros métodos de transformación son más frecuentemente usados para estas plantas. Actualmente, el método de elección para generar monocotiledóneas transgénicas, complementando el método de *Agrobacterium*, es bombardeo de partícula

(oro microscópico o partículas de tungsteno revestido con la transformación ADN) de callos embrionarios o embriones de revelando (Christou, 1992, *Plant Journal* 2: 275-281; Shimamoto, 1994, *Current Opinion Biotechnology* 5: 158-162; Vasil et al., 1992, *Bio/Technology* 10: 667-674). Un método alternativo para transformación de monocotiledóneas se basa en transformación de protoplasto como se describe por Omirulleh et al., 1993, *Plant Molecular Biology* 21: 415-428.

[0136] Después de la transformación, los transformantes que se han incorporado aquí, la construcción de expresión se selecciona y se regenera en plantas enteras según métodos bien conocidos en la técnica. Frecuentemente, el procedimiento de transformación se diseña para la eliminación selectiva de genes de selección bien durante regeneración o en las siguientes generaciones usando, por ejemplo, co-transformación con dos construcciones de T-ADN separadas o escisión específica de sitio del gen de selección por una recombinasa específica.

[0137] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención comprendiendo (a) cultivo de una planta transgénica o una célula vegetal comprendiendo una secuencia de ácidos nucleicos que codifica un polipéptido con actividad de proteasa de la presente invención bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

15 Animales

[0138] La presente invención también se refiere a un animal transgénico no humano y productos o elementos de este, ejemplos de este son líquidos biológicos como leche y sangre, órganos, carne y células animales. Técnicas para expresión de proteínas, por ejemplo, en células mamíferas, se conocen en la técnica, ver, por ejemplo, *Handbook Protein Expression: A Practical Approach*, Higgins and Hames (eds), Oxford University Press (1999), y los otros tres manuales en esta serie acerca transcripción de gen, maduración del RNA y procesamiento postraduccional. En términos generales, para preparar un animal transgénico, células seleccionadas de un animal seleccionado se transforman con una secuencia de ácidos nucleicos que codifica un polipéptido con actividad de proteasa de la presente invención para expresar y producir el polipéptido. El polipéptido puede recuperarse del animal, por ejemplo, de la leche de animales hembra, o el polipéptido se puede expresar al beneficio del animal mismo, por ejemplo, para ayudar la digestión del animal. Ejemplos de animales se mencionan más adelante en la sección titulada Alimento para animales.

[0139] Para producir un animal transgénico con el propósito de recuperar proteasa de la leche del animal, un gen que codifica la proteasa se puede insertar en los huevos fertilizados de un animal en cuestión, por ejemplo, usando un vector de expresión de transgen que comprende un promotor de proteína de la leche adecuado y la proteasa de gen de codificación. El vector de expresión de transgen se microinyecta en huevos fertilizados y preferiblemente se integra permanentemente en el cromosoma. Un vez el huevo empieza a crecer y dividirse, el embrión potencial se implanta en una madre portadora y animales que llevan el transgen se identifican. El animal resultante puede después ser multiplicado por cultivo convencional. El polipéptido se puede purificar de la leche del animal, ver, por ejemplo, Meade, H.M. et al (1999): *Expression of recombinant proteins in the milk of transgenic animals*, *Gene expression systems: Using nature for the art of expression*. J. M. Fernandez and J. P. Hoeffler (eds.), Academic Press.

[0140] En la alternativa, para producir un animal transgénico no humano que lleva en el genoma de sus células somáticas y/o germinales una secuencia de ácidos nucleicos que incluye una construcción de transgen heterólogo que incluye una proteasa de codificación de transgen, el transgen puede unirse operativamente a una primera secuencia reguladora para expresión específica de glándula salival de proteasa, como descrito en WO 2000064247.

Composiciones

[0141] En todavía otro aspecto, la presente invención se refiere a composiciones comprendiendo un polipéptido de la presente invención.

[0142] Las composiciones de polipéptido se pueden preparar conforme a métodos conocidos en la técnica y pueden estar en forma de líquido o una composición seca. Por ejemplo, la composición de polipéptido puede estar en forma de un granulado o un microgranulado. El polipéptido para ser incluido en la composición se puede estabilizar conforme a métodos conocidos en la técnica.

[0143] Ejemplos se dan más adelante de usos preferidos de los polipéptidos o composiciones de polipéptido de la invención.

50 Alimento para animales

[0144] La presente invención se dirige también a métodos para uso de los polipéptidos con actividad de proteasa en el alimento para animales, al igual que para composiciones alimenticias y aditivos alimenticios comprendiendo los polipéptidos de la invención.

[0145] El término animal incluye todos los animales, incluyendo seres humanos. Ejemplos de animales son no rumiantes y rumiantes. Animales rumiantes incluyen, por ejemplo, animales como oveja, cabras, caballos y ganado

- 5 bovino, por ejemplo, ganado bovino para carne, vacas y terneros jóvenes. En una forma de realización particular, el animal es un animal no rumiante. Animales no rumiantes incluyen animales monogástricos, por ejemplo, cerdos o puercos (incluyendo, pero no limitado a, lechones, cerdos en crecimiento y cerdas); aves como pavos, patos y pollos (incluyendo, pero no limitado a, pollos para asar, estratos); terneros jóvenes; y peces (incluyendo, pero no limitado a, salmón, trucha, tilapia, bagre y carpas; y crustáceos (incluyendo, pero no limitado, a gambas y cigalas).
- [0146] El término alimento o composición alimenticia significa cualquier compuesto, preparaciones, mezcla, o composición adecuada para, o destinado para, toma por un animal.
- [0147] En el uso según la invención la proteasa se puede suministrar al animal antes, después, o simultáneamente con la dieta.
- 10 Lo último se prefiere.
- [0148] En una forma de realización particular, la proteasa, en la forma en la que se añade al alimento, o cuando se incluye en un aditivo de alimento, está bien definido. Bien definido significa que la preparación de proteasa es al menos 50% puro como determinado por cromatografía de exclusión de tamaño (véase ejemplo 12 de la WO 01/58275). En otras formas de realización particulares, la preparación de proteasa es al menos 60, 70, 80, 85, 88, 15 90, 92, 94, o al menos 95% puro como determinado por este método.
- [0149] Una preparación bien definida de proteasa es ventajosa. Por ejemplo, es mucho más fácil a dosis correctamente al alimentar una proteasa que es esencialmente libre de interferir o contaminar otras proteasas. El término dosis correctamente se refiere en particular al objetivo de obtener resultados constantes y consistentes, y a la capacidad de optimización de dosificación basada en el efecto deseado.
- 20 [0150] Para uso en el alimento para animales, no obstante, la proteasa no necesita ser tan pura; esta puede, por ejemplo, incluir otras enzimas, en cuyo caso esto podría denominarse una preparación de proteasa.
- [0151] La preparación de proteasa puede ser (a) añadida directamente al alimento (o usada directamente en un proceso de tratamiento de proteína), o (b) esta se puede usar en la producción de una o más composiciones intermedias como aditivos alimenticios o premezclas que se añaden posteriormente al alimento (o se usan en un proceso de tratamiento). El grado de pureza anteriormente descrito se refiere a la pureza de la preparación de 25 proteasa original, si se usa según (a) o (b) de arriba.
- [0152] Preparaciones de proteasa con purezas de este orden de magnitud son en particular obtenibles usando métodos recombinantes de producción, mientras que no son así fácilmente obtenidas y también están sujetas a una variación entre lotes mucho más alta cuando la proteasa se produce por métodos de fermentación tradicionales.
- 30 [0153] Tal preparación de proteasa puede, por supuesto ser mezclada con otras enzimas.
- [0154] En otra forma de realización particular, la proteasa para uso según la invención es capaz de solubilización de proteínas según el modelo *in vitro* del ejemplo 8 de este documento.
- [0155] La proteína puede ser una proteína animal, como harina de carne y huesos, y/o harina de pescado; o esta puede ser una proteína vegetal.
- 35 [0156] El término proteínas vegetales como se utiliza en este caso se refiere a cualquier compuesto, composición, preparación o mezcla que incluye al menos una proteína derivada de u originada de un vegetal, incluyendo proteínas modificadas y derivados de proteína. En formas de realización particulares, el contenido de proteína de las proteínas vegetales es al menos 10, 20, 30, 40, 50, o 60% (p/p).
- 40 [0157] Proteínas vegetales pueden derivar de fuentes de proteína vegetal, como legumbres y cereales, por ejemplo, materiales de plantas de las familias *Fabaceae* (*Leguminosae*), *Cruciferaeae*, *Chenopodiaceae* y *Poaceae*, como harina de soja, harina de lupino y comida de semilla de colza.
- [0158] En una forma de realización particular, la fuente de proteína vegetal es material de una o más plantas de la familia *Fabaceae*, por ejemplo, semilla de soja, altramuz, guisante, o judía.
- 45 [0159] En otra forma de realización particular, la fuente de proteína vegetal es material de una o más plantas de la familia *Chenopodiaceae*, por ejemplo, remolacha, remolacha azucarera, espinaca o quinua.
- [0160] Otros ejemplos de fuentes de proteína vegetal son semilla de colza, semilla de girasol, semilla de algodón y repollo.
- [0161] Semilla de soja es una fuente de proteína vegetal preferida.
- 50 [0162] Otros ejemplos de fuentes de proteína vegetal son cereales como cebada, trigo, centeno, avena, maíz (grano), arroz, triticale y sorgo.

- 5 [0163] El tratamiento según la invención de proteínas con al menos una proteasa de la invención produce una solubilización aumentada de proteínas, en comparación con la blanco. Al menos 101%, o 102%, 103%, 104%, 105%, 106%, o al menos 107% de proteína solubilizada puede ser obtenible usando las proteasas de la invención, haciendo referencia al modelo *in vitro* del ejemplo 8 de este documento. El término solubilización de proteínas básicamente significa llevar proteína(s) en la solución. Tal solubilización se puede deber a liberación mediada por proteasa de proteína de otros componentes de las composiciones naturales normalmente complejas, como alimento.
- Solubilización puede ser medida como un aumento en la cantidad de proteínas solubles, por referencia a una muestra con ningún tratamiento de proteasa (véase ejemplo 8).
- 10 [0164] El tratamiento según la invención de proteínas con al menos una proteasa de la invención produce una digestibilidad aumentada de proteínas, en comparación con la blanco. Al menos 101%, o 102%, 103%, 104%, 105%, 106%, 107%, 108%, 109%, 110%, 111%, 112%, 113%, 114%, 115%, o al menos 116% de proteína digerible puede ser obtenible usando las proteasas de la invención, haciendo referencia al modelo *in vitro* del ejemplo 8 de este documento.
- 15 [0165] En una forma de realización particular de un proceso de tratamiento, la proteasa(s) en cuestión afecta (o actúa en, o ejerce su influencia de solubilización en) las proteínas, como proteínas vegetales o fuentes de proteína. Para conseguir esto, la proteína o fuente de proteína se suspende típicamente en un solvente, por ejemplo, un solvente acuoso tal como agua, y el pH y valores de temperatura son ajustados prestando la debida atención a las características de la enzima en cuestión. Por ejemplo, el tratamiento puede ocurrir a un valor de pH en el que la actividad de la proteasa real es al menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, o al menos 90%.
- 20 Asimismo, por ejemplo, el tratamiento puede ocurrir a una temperatura en la que la actividad de la proteasa real es al menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, o al menos 90%. Las indicaciones de porcentajes de actividad anteriores son en relación a las actividades máximas. La reacción enzimática es continua hasta que el resultado deseado se consigue, seguido de que esto puede o no puede detenerse inactivando la enzima, por ejemplo, por un paso de tratamiento térmico.
- 25 [0166] En otra forma de realización particular de un proceso de la invención de tratamiento, la acción de la proteasa se sostiene, significando, por ejemplo, que la proteasa se añade a las proteínas, pero su influencia de solubilización es por así decirlo no accionada hasta más tarde cuando se desea, una vez que condiciones de solubilización se establecen, o una vez que cualquier inhibidor enzimático se inactiva, o cualquier otros medios podrían haber sido aplicados para posponer la acción de la enzima.
- 30 [0167] En una forma de realización el tratamiento es un pretratamiento de alimento para animales o proteínas para uso en el alimento para animales, es decir, las proteínas se solubilizan antes de la toma.
- [0168] El término mejora del valor nutritivo de un alimento para animales significa mejorar la disponibilidad de las proteínas, conduciendo así a extracción de proteína aumentada, rendimiento de proteína mayor, y/o utilización de proteína mejorada. El valor nutritivo del alimento se aumenta por lo tanto y el índice de crecimiento y/o ganancia de peso y/o conversión de alimento (es decir, el peso de alimento ingerido en relación a aumento de peso) del animal es/son mejorado.
- 35 [0169] La proteasa se puede añadir al alimento en cualquier forma, ser este como una proteasa relativamente pura, o en la mezcla con otros componentes destinados para adición para alimento para animales, es decir, en forma de aditivos de alimento, como las denominadas pre-mezclas para alimento para animales.
- 40 [0170] En otro aspecto, la presente invención se refiere a composiciones para uso en el alimento para animales, como alimento para animales y aditivos de alimento, por ejemplo, premezclas.
- [0171] Además de la proteasa de la invención, los aditivos de alimento de la invención contienen al menos una vitamina liposoluble y/o al menos una vitamina hidrosoluble y/o al menos un oligoelemento y/o al menos un macromineral.
- 45 [0172] Además, ingredientes con aditivos de alimentos opcionales son agentes colorantes, por ejemplo, carotenoides como beta-caroteno, astaxantina y luteína; compuestos aromáticos; estabilizadores; péptidos antimicrobianos; ácidos grasos poliinsaturados; especies generadores de oxígeno reactivo; y/o al menos otra enzima seleccionada de entre fitasa (EC 3.1.3.8 o 3.1.3.26); xilanasas (EC 3.2.1.8); galactanasas (EC 3.2.1.89); alfa-galactosidasas (EC 3.2.1.22); proteasa (EC 3.4.-.-), fosfolipasa A1 (EC 3.1.1.32); fosfolipasa A2 (EC 3.1.1.4);
- 50 lisofosfolipasa (EC 3.1.1.5); fosfolipasa C (3.1.4.3); fosfolipasa D (EC 3.1.4.4); amilasa como, por ejemplo, alfa-amilasa (EC 3.2.1.1); y/o beta-glucanasas (EC 3.2.1.4 o EC3.2.1.6).
- [0173] En una forma de realización particular, estas otras enzimas son bien definidas (como definidas arriba para preparaciones de proteasa).
- 55 [0174] Ejemplos de péptidos antimicrobianos (AMP) son CAP18, Leucocin A, Trirpticin, Protegrin-1, Thanatin, Defensin, Lactoferrin, Lactoferricin y Ovispirin como Novispirin (Robert Lehrer, 2000), Plectasins y Statins,

incluyendo los compuestos y polipéptidos descritos en las WO 03/044049 y WO 03/048148, al igual que variantes o fragmentos de lo anterior que retienen actividad antimicrobiana.

5 [0175] Ejemplos de polipéptidos antifúngicos (AFP) son el *Aspergillus giganteus*, péptidos de *Aspergillus niger*, al igual que variantes y fragmentos de estos que retienen actividad antifúngica, como descritos en las WO 94/01459 y WO 02/090384.

[0176] Ejemplos de ácidos grasos poliinsaturados son ácidos grasos poliinsaturados C18, C20 y C22, como ácido araquidónico, ácido docosahexaenoico, ácido eicosapentanoico y ácido gamma-linolénico.

[0177] Ejemplos de especies generadoras de oxígeno reactivo son productos químicos como perborato, persulfato, o percarbonato; y enzimas como una oxidasa, una oxigenasa o una sintetasa.

10 [0178] Normalmente, grasa y vitaminas hidrosolubles, al igual que oligoelementos forman parte de una denominada premezcla destinada para adición al alimento, mientras que macrominerales se añaden normalmente por separado al alimento. Cualquiera de estos tipos de composición, cuando enriquecido con una proteasa de la invención, es un aditivo de alimento de la invención.

15 [0179] En una forma de realización particular, el aditivo de alimento de la invención se destina para ser incluido (o prescrito como que tiene que ser incluido) en dietas para animales o alimento a niveles de 0,01 a 10,0%; más particularmente 0,05 a 5,0%; o 0,2 a 1,0% (% significa g aditivo por 100 g de alimento).

Esto es así en particular para premezclas.

[0180] Lo siguiente son listas no exclusivas de ejemplos de estos componentes:

20 Ejemplos de vitaminas liposolubles son vitamina A, vitamina D3, vitamina E y vitamina K, por ejemplo, vitamina K3.

Ejemplos de vitaminas hidrosolubles son vitamina B12, biotina y colina, vitamina B1, vitamina B2, vitamina B6, niacina, ácido fólico y pantotenato, por ejemplo, Ca-D-pantotenato.

Ejemplos de oligoelementos son manganeso, zinc, hierro, cobre, yodo, selenio y cobalto.

Ejemplos de macrominerales son calcio, fósforo y sodio.

25 [0181] Los requisitos nutritivos de estos componentes (ejemplificados con ave y lechones/cerdos) se catalogan en la tabla A de WO 01/58275. Requisito nutritivo significa que estos componentes deberían ser proporcionados en la dieta en las concentraciones indicadas.

30 [0182] En la alternativa, el aditivo de alimento de la invención comprende al menos uno de los componentes individuales específicos en la tabla A de WO 01/58275. Al menos uno significa cualquiera de, uno o más de, uno, o dos, o tres, o cuatro, etcétera todos hasta trece, o todos hasta quince componentes individuales. Más específicamente, este al menos un componente individual se incluye en el aditivo de la invención en tal cantidad en cuanto a proporcionar una concentración en alimento en el intervalo indicado en la columna cuatro, o columna cinco, o columna seis de la tabla de A.

35 [0183] En todavía otra forma de realización, el aditivo de alimento de la invención comprende al menos una de las vitaminas de debajo, preferiblemente para proporcionar una concentración en peso en los intervalos específicos en la tabla I de debajo (para dietas de lechón y dietas de engorde, respectivamente).

Tabla I: Recomendaciones de vitamina típicas

Vitamina	Dieta de lechón	Dieta de engorde
Vitamina A	10,000-15,000 IU/kg de alimento	8-12,500 IU/kg de alimento
Vitamina D3	1800-2000 IU/kg de alimento	3000-5000 IU/kg de alimento
Vitamina E	60-100 mg/kg de alimento	150-240 mg/kg de alimento
Vitamina K3	2-4 mg/kg de alimento	2-4 mg/kg de alimento
Vitamina B1	2-4 mg/kg de alimento	2-3 mg/kg de alimento
Vitamina B2	6-10 mg/kg de alimento	7-9 mg/kg de alimento
Vitamina B6	4-8 mg/kg de alimento	3-6 mg/kg de alimento
Vitamin B12	0,03-0,05 mg/kg de alimento	0,015-0,04 mg/kg de alimento
Niacina (Vitamina B3)	30-50 mg/kg de alimento	50-80 mg/kg de alimento
Ácido Pantoténico	20-40 mg/kg de alimento	10-18 mg/kg de alimento
Ácido fólico	1-2 mg/kg de alimento	1-2 mg/kg de alimento
Biotina	0,15-0,4 mg/kg de alimento	0,15-0,3 mg/kg de alimento
Cloruro de colina	200-400 mg/kg de alimento	300-600 mg/kg de alimento

5 [0184] La presente invención también se refiere a composiciones de alimento para animales. Composiciones de alimento para animales o dietas tienen un contenido relativamente alto de proteína. Dietas de ave y de cerdo se pueden caracterizar como indicadas en la tabla B de la WO 01/58275, columnas 2-3. Dietas de pescado se pueden caracterizar como indicadas en la columna 4 de esta tabla B. Además tales dietas de pescado normalmente tienen un contenido de grasa cruda de 200-310 g/kg.

10 [0185] La WO 01/58275 corresponde a la US 09/779334.

[0186] Una composición de alimento para animales según la invención tiene un contenido bruto de proteína de 50-800 g/kg y comprende además al menos una proteasa como se reivindica aquí.

15 [0187] Además, o en la alternativa (al contenido bruto de proteína indicado anteriormente), la composición de alimento para animales de la invención tiene un contenido de energía metabolizable de 10-30 MJ/kg; y/o un contenido de calcio de 0,1-200 g/kg; y/o un contenido de fósforo disponible de 0,1-200 g/kg; y/o un contenido de metionina de 0,1-100 g/kg; y/o un contenido de metionina más cisteína de 0,1-150 g/kg; y/o un contenido de lisina de 0,5-50 g/kg.

20 [0188] En formas de realización particulares, el contenido de energía metabolizable, proteína cruda, calcio, fósforo, metionina, metionina más cisteína y/o lisina está dentro de todos los intervalos 2, 3, 4 o 5 en la tabla B de la WO 01/58275 (R. 2-5).

[0189] Proteína cruda se calcula como nitrógeno (N) multiplicado por un factor 6,25, es decir, proteína cruda (g/kg)= N (g/kg) x 6,25. El contenido de nitrógeno se determina por el método Kjeldahl (A.O.A.C., 1984, *Official Methods of Analysis 14th ed.*, Association of Official Analytical Chemists, Washington DC).

5 [0190] Energía metabolizable puede calcularse basándose en la *NRC publication Nutrient requirements in swine, ninth revised edition 1988, subcommittee on swine nutrition, committee on animal nutrition, board of agriculture, national research council. National Academy Press, Washington, D.C., pp. 2-6*, y *European Table of Energy Values for Poultry Feed-stuffs, Spelderholt centre for poultry research and extension, 7361 DA Beekbergen, The Netherlands. Grafisch bedrijf Ponsen & looijen bv, Wageningen. ISBN 90-71463-12-5*.

10 [0191] El contenido dietético de calcio, fósforo disponible y aminoácidos en dietas para animales completas se calcula basándose en tablas de aliento como *Veevoedertabel 1997, gegevens over chemische samenstelling, verteerbaarheid en voederwaarde van voedermiddelen, Central Veevoederbureau, Runderweg 6, 8219 pk Lelystad.*

ISBN 90-72839-13-7.

15 [0192] En una forma de realización particular, la composición de alimento para animales de la invención contiene al menos una proteína vegetal como se ha definido anteriormente. Puede también contener proteína animal, como harina de carne y huesos y/o harina de pescado, típicamente en una cantidad de 0-25%.

[0193] En todavía otras formas de realización particulares, la composición de alimento para animales de la invención contiene 0-80% de maíz; y/o 0-80% de sorgo; y/o 0-70% de trigo; y/o 0-70% de cebada; y/o 0-30% de avena; y/o 0-40% de harina de soja; y/o 0-25% de harina de pescado; y/o 0-25% de harina de carne y huesos; y/o 0-20% de lactosuero.

20 [0194] Dietas para animales pueden, por ejemplo, fabricarse como alimento triturado (no granulado) o alimento granulado. Típicamente, los materiales de alimento molidos se mezclan y cantidades suficientes de vitaminas esenciales y minerales se agregan según las especificaciones para las especies en cuestión. Enzimas se pueden añadir como formulaciones enzimáticas sólidas o líquidas. Por ejemplo, una formulación de enzima sólida se añade típicamente antes o durante el paso de mezcla; y una preparación enzimática líquida se añade típicamente después del paso de granulación. La enzima puede también ser incorporada en un aditivo de alimento o premezcla.

25 [0195] La concentración de enzima final en la dieta está en el intervalo de 0,01-200 mg de proteína enzimática por kg de dieta, por ejemplo, en el intervalo de 0,5-25 mg de proteína enzimática por kg de dieta animal.

30 [0196] La proteasa debería, por supuesto, ser aplicada en una cantidad eficaz, es decir, en una cantidad adecuada para mejora de solubilización, digestibilidad, y/o mejora de valor nutritivo de alimento. Actualmente, se contempla que la enzima se administra en una o más de las cantidades siguientes (intervalos de dosificación): 0,01-200, 0,01-100, 0,5-100, 1-50, 5-100, 10-100, 0,05-50 o 0,10-10, todos estos intervalos son en mg de proteína de proteasa por kg de alimento (ppm).

35 [0197] Para determinar mg de proteína de proteasa por kg de alimento, la proteasa se purifica de la composición alimentaria y la actividad específica de la proteasa purificada se determina usando un ensayo pertinente (ver debajo actividad de proteasa, sustratos y ensayos). La actividad de proteasa de la composición alimentaria como tal se determina también usando el mismo ensayo y basándose en estas dos determinaciones, la dosificación en mg de proteína de proteasa por kg de alimento se calcula.

40 [0198] Los mismos principios pueden determinar mg de proteína de proteasa en aditivos alimenticios. Por supuesto, si una muestra está disponible de la proteasa usada para preparar el aditivo de alimento o el alimento, la actividad específica se determina de esta muestra (sin necesidad de purificar la proteasa de la composición alimentaria o el aditivo).

Composiciones detergentes

[0199] La proteasa de la invención se puede añadir a y así hacerse un componente de una composición detergente.

45 [0200] La composición detergente de la invención puede, por ejemplo, formularse como composición de detergente para ropa a mano o a máquina que incluye una composición de aditivo de lavandería adecuada para pretratamiento de tejidos manchados y una composición de suavizante de enjuague añadida, o ser formulada como una composición detergente para uso en operaciones de limpieza de superficies duras del hogar en general, o ser formulada operaciones de lavado de la vajilla a mano o a máquina.

50 [0201] En un aspecto específico, la invención proporciona un aditivo detergente comprendiendo la proteasa de la invención. El aditivo detergente al igual que la composición detergente puede comprender una o más otras enzimas como otra proteasa, como proteasas alcalinas de *Bacillus*, una lipasa, una cutinasa, una amilasa, una carbohidrasa, una celulasa, una pectinasa, una mananasa, una arabinasa, una galactanasa, una xilanasa, una oxidasa, por ejemplo, una lacasa y/o una peroxidasa.

[0202] En general, las propiedades de la enzima(s) elegida deberían ser compatibles con el detergente seleccionado, (es decir, pH óptimo, compatibilidad con otros ingredientes no enzimáticos y enzimáticos, etc.), y la enzima(s) debería estar presente en cantidades eficaces.

5 [0203] Lipasas adecuadas incluyen aquellas de origen fúngico o bacteriano. Proteínas químicamente modificadas o mutantes creadas genéticamente se incluyen. Ejemplos de lipasas útiles incluyen lipasas de *Humicola* (sinónimo de *Thermomyces*), por ejemplo, de *H. lanuginosa* (*T. Lanuginosus*) como se describe en las EP 258068 y EP 305216 o de *H. insolens* como se describe en la WO 96/13580, una lipasa de *Pseudomonas*, por ejemplo, *P. alcaligenes* o *P. pseudoalcaligenes* (EP 218272), *P. cepacia* (EP 331376), *P. stutzeri* (GB 1,372,034), *P. fluorescens*, *Pseudomonas* sp. cepa SD 705 (WO 95/06720 y WO 96/27002), *P. wisconsinensis* (WO 96/12012), una lipasa de *Bacillus*, por ejemplo, de *B. subtilis* (Dartois et al. (1993), *Biochemica et Biophysica Acta*, 1131, 253-360), *B. stearothermophilus* (JP 64/744992) o *B. pumilus* (WO 91/16422). Otros ejemplos son variantes de lipasa tales como las descritas en las WO 92/05249, WO 94/01541, EP 407225, EP 260105, WO 95/35381, WO 96/00292, WO 95/30744, WO 94/25578, WO 95/14783, WO 95/22615, WO 97/04079 y WO 97/07202. Enzimas de lipasa preferidas disponibles comercialmente incluyen Lipolase™ y Lipolase Ultra™ (Novozymes A/S). Amilasas adecuadas (beta y/o alfa)

10 incluyen aquellas de origen fúngico o bacteriano. Proteínas químicamente modificadas o mutantes creadas genéticamente se incluyen. Amilasas incluyen, por ejemplo, alfa-amilasas obtenidas de *Bacillus*, por ejemplo, de una cepa especial de *B. licheniformis*, descrito con más detalle en la GB 1,296,839. Ejemplos de amilasas útiles son las variantes descritas en las WO 94/02597, WO 94/18314, WO 95/26397, WO 96/23873, WO 97/43424, WO 00/60060 y WO 01/66712, especialmente las variantes con sustituciones en una o más de las siguientes posiciones: 15, 23,

15 105, 106, 124, 128, 133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408, y 444. Amilasas disponibles comercialmente son Natalase™, Supramyl™, Stainzyme™, Duramyl™, Termamyl™, Fungamyl™ y BAN™ (Novozymes A/S), Rapidase™ y Purastar™ (de Genencor International Inc.).

20

[0204] Celulasas adecuadas incluyen aquellas de origen fúngico o bacteriano. Proteínas químicamente modificadas o mutantes creadas genéticamente se incluyen. Celulasas adecuadas incluyen celulasas del género *Bacillus*,

25 *Pseudomonas*, *Humicola*, *Fusarium*, *Thielavia*, *Acremonium*, por ejemplo, las celulasas fúngicas producidas de *Humicola insolens*, *Myceliophthora thermophila* y *Fusarium oxysporum* descritas en las US 4,435,307, US 5,648,263, US 5,691,178, US 5,776,757 y WO 89/09259. Celulasas especialmente adecuadas son las celulasas neutras o alcalinas con beneficios de cuidado de color. Ejemplos de tales celulasas son celulasas descritas en las EP 0 495257, EP 531372, WO 96/11262, WO 96/29397, WO 98/08940. Otros ejemplos son variantes de celulasas tales como las descritas en las WO 94/07998, EP 0 531 315, US 5,457,046, US 5,686,593, US 5,763,254, WO 95/24471, WO 98/12307 y WO 99/01544. Celulasas disponibles comercialmente incluyen Celluzyme™, y Carezyme™ (Novozymes A/S), Clazinase™ y Puradax HA™ (Genencor internacional Inc.), y KAC- 500(B)™ (Kao Corporation).

30

[0205] Peroxidasas/oxididasas adecuadas incluyen aquellas origen de plantas, fúngico o bacteriano. Proteínas químicamente modificadas o mutantes creadas genéticamente se incluyen. Ejemplos de peroxidasas útiles incluyen peroxidasas de *Coprinus*, por ejemplo, de *C. cinereus* y variantes de esta como aquellas descritas en las WO 93/24618, WO 95/10602 y WO 98/15257. Peroxidasas disponibles comercialmente incluyen Guardzyme™ (Novozymes).

35

[0206] La enzima(s) detergente se puede incluir en una composición detergente añadiendo aditivos separados con una o más enzimas, o añadiendo un aditivo combinado comprendiendo todas estas enzimas. Un aditivo detergente de la invención, es decir, un aditivo separado o un aditivo combinado, se puede formular por ejemplo, como un granulado, un líquido, un lodo, etc. Formulaciones de aditivo detergente preferidas son granulados, en particular, granulados no polvorientos, líquidos, en particular, líquidos estabilizados, o compuestos acuosos.

40

[0207] Granulados no polvorientos pueden producirse, por ejemplo, como se describe en las US 4,106,991 y 4,661,452 y pueden opcionalmente revestirse por métodos conocidos en la técnica. Ejemplos de materiales de recubrimiento cerosos son productos de poli(óxido de etileno) (polietilenoglicol; PEG) con pesos molares medios de 1000 a 20000; nonilfenoles etoxilados con de 16 a 50 unidades de óxido de etileno; alcoholes etoxilados grasos en los que el alcohol contiene de 12 a 20 átomos de carbono y en los cuales hay de 15 a 80 unidades de óxido de etileno; alcoholes grasos; ácidos grasos; y mono y di- y triglicéridos de ácidos grasos. Ejemplos de materiales de recubrimiento que forman películas adecuadas para aplicación por técnicas de lecho fluidificado se dan en la GB 1483591. Preparaciones enzimáticas líquidas pueden, por ejemplo, estabilizarse añadiendo un poliol tal como propilenglicol, un azúcar o alcohol de azúcar, ácido láctico o ácido bórico según métodos establecidos. Enzimas protegidas se pueden preparar según el método descrito en la EP 238216.

45

50

[0208] La composición detergente de la invención puede estar en cualquier forma conveniente, por ejemplo, una barra, una pastilla, un polvo, un gránulo, una pasta o un líquido. Un detergente líquido puede ser acuoso, típicamente con hasta 70 % de agua y 0-30 % disolvente orgánico, o no acuoso.

55

[0209] La composición detergente comprende uno o más tensioactivos, que pueden ser no iónico incluyendo semipolar y/o aniónico y/o zwitteriónico y/o catiónico. Los tensioactivos están presentes típicamente a un nivel de 0,1 % a 60% en peso.

[0210] Cuando se incluye aquí el detergente normalmente contiene de aproximadamente 1% a aproximadamente 40% de un surfactante aniónico tal como alquilbencenosulfonato lineal, alfa-olefinsulfonato, sulfato de alquilo (sulfato de alcohol graso), etoxisulfato alcohólico, alcanosulfonato secundario, alfa-sulfo ácido graso de metil éster, ácido alquil- o alqueni succínico o jabón.

5 [0211] Cuando se incluye aquí el detergente normalmente contiene de aproximadamente 0,2% a aproximadamente 40% de un surfactante no iónico tal como alcohol etoxilato, nonilfenol, etoxilato alquilpoliglicósido, óxido de alquildimetilamina, monoetanolamida de ácido graso etoxilado, monoetanolamida de ácido graso, amida de polidroxil alquil ácido graso, o derivados de N-acil N-alquil glucosamina ("glucamidas").

10 [0212] El detergente puede contener 0-65 % de un constructor de detergente o agente complejo tal como zeolita, difosfato, trifosfato, fosfonato, carbonato, citrato, ácido nitrilotriacético, ácido etilendiaminatetraacético, ácido dietileno-triaminopentaacético, ácido alquil- o alqueni succínico, silicatos solubles o silicatos estratificados (p. ej. SKS-6 de Hoechst).

15 [0213] El detergente puede comprender uno o más polímeros. Ejemplos son carboximetilcelulosa, poli(vinilpirrolidona), poli(etilenglicol), poli(vinil alcohol), poli(vinilpiridina-N-óxido), poli(vinilimidazole), policarboxilatos tales como poliacrilatos, copolímeros de ácido maléico/acrílicos y copolímeros de ácido de lauril metacrilato/acrílico.

[0214] El detergente puede contener un sistema blanqueante que puede comprender una fuente de H₂O₂ tal como perborato o percarbonato que se puede combinar con un activador blanqueante de formación de perácido tal como tetraacetil-etilendiamina o nonanoiloxibencenosulfonato. Alternativamente, el sistema blanqueante puede comprender peroxiácidos del tipo, por ejemplo, amida, imida, o sulfona.

20 [0215] La enzima(s) de la composición detergente de la invención puede estabilizarse usando agentes convencionales estabilizantes, por ejemplo, un poliol tal como propilenglicol o glicerol, un azúcar o alcohol de azúcar, ácido láctico, ácido bórico, o un derivado de ácido bórico, por ejemplo, un éster de borato aromático, o un derivado de ácido fenil borónico tal como 4-formilfenil ácido borónico, y la composición se puede formular como se describe en, por ejemplo, las WO 92/19709 y WO 92/19708.

25 [0216] El detergente puede también contener otros ingredientes de detergentes convencionales como, por ejemplo, acondicionadores de tejidos incluyendo arcillas, reforzadores de espuma, supresores de espuma, agentes anticorrosivos, agentes de suspensión de suciedad, agentes de reposición antisuciedad, tintes, bactericidas, blanqueadores ópticos, hidrotropos, inhibidores de decoloración, o perfumes.

30 [0217] Actualmente, se contempla que en las composiciones detergentes cualquier enzima, en particular, la enzima de la invención, se puede añadir en una cantidad correspondiente a 0,01-100 mg de proteína enzimática por litro de líquido de lavado, preferiblemente 0,05-5 mg de proteína enzimática por litro de líquido de lavado, en particular, 0,1-1 mg de proteína enzimática por litro de líquido de lavado.

[0218] La enzima de la invención puede adicionalmente incorporarse en las formulaciones detergentes descritas en la WO 97/07202.

35 **Formas de realización particulares**

[0219] La invención también se refiere a las siguientes formas de realización particulares, en lo que sigue numerado 42-58;

40 42. Un polipéptido aislado con actividad de proteasa, seleccionado del grupo que consiste en: (a) un polipéptido con una secuencia de aminoácidos que tiene un grado de identidad para aminoácidos 1-192 de la SEC ID n°: 2 de al menos 20

(b) una variante alélica de (a), y (c) un fragmento de un, (a), o (b) que tiene actividad de proteasa.

45 43. Una secuencia de ácidos nucleicos aislada comprendiendo una secuencia de ácidos nucleicos que codifica un polipéptido con actividad de proteasa, y que (a) codifica el polipéptido de cualquiera de las formas de realización 42, y/o (c) tiene un grado de identidad para nucleótidos 568-1143 de la SEC ID n°: 1 de al menos 81,2%.

44. Una construcción de ácidos nucleicos comprendiendo la secuencia de ácidos nucleicos de la forma de realización 43 unida operativamente a una o más secuencias de control que dirigen la producción del polipéptido en un huésped de expresión adecuada.

50 45. Un vector de expresión recombinante comprendiendo la construcción de ácidos nucleicos de la forma de realización 44.

46. Una célula huésped recombinante comprendiendo la construcción de ácidos nucleicos de la forma de realización 44 o el vector de la forma de realización 45.

47. Una planta transgénica, o parte de planta, capaz de expresión del polipéptido de la forma de realización 42.

48. Un animal transgénico no humano, o productos o elementos de este, que es capaz de expresión del polipéptido de cualquiera de la forma de realización 42.

5 49. Un método para la producción de un polipéptido de las formas de realización 42-43, el método comprendiendo (a) cultivo de una célula huésped recombinante de la forma de realización 49 para producir un sobrenadante comprendiendo el polipéptido; y (b) recuperación del polipéptido.

10 50. Un método para la producción de un polipéptido de la forma de realización 42, el método comprendiendo (a) cultivo de la siguiente cepa: *Nocardioopsis dassonvillei* subesp. *dassonvillei* DSM 43235, y (b) recuperación del polipéptido.

15 51. Uso de al menos una proteasa de la forma de realización 42 (i) en el alimento para animales; (ii) en aditivos de alimento; (iii) en la preparación de una composición para uso en el alimento para animales; (iv) para mejorar el valor nutritivo de un alimento para animales; (v) para aumentar la proteína soluble y/o digerible en el alimento para animales; (vi) para aumentar el grado de hidrólisis de proteínas en dietas para animales; y/o (vii) para el tratamiento de proteínas.

52. Un método para mejorar el valor nutritivo de un alimento para animales, donde al menos una proteasa de la forma de realización 42 se añade al alimento.

20 53. Un aditivo de alimento comprendiendo (a) al menos una proteasa de la forma de realización 42; y (b) al menos una vitamina grasa soluble, y/o (c) al menos una vitamina hidrosoluble, y/o (d) al menos un oligoelemento.

54. El aditivo de alimento de la forma de realización 53, que comprende además amilasa; fitasa; xilanasas; galactanasas; alfa-galactosidasas; proteasa, fosfolipasa; y/o beta-glucanasa.

55. Un alimento para animales con un contenido bruto de proteína de 50 a 800 g/kg y comprendiendo al menos una proteasa de la forma de realización 42.

25 56. Un método para el tratamiento de proteínas, comprendiendo el paso de adición de al menos una proteasa de la forma de realización 42 a al menos una proteína o fuente de proteína.

57. El método de la forma de realización 56, donde semilla de soja se incluye entre al menos una fuente de proteína.

58. Uso de al menos una proteasa de la forma de realización 42 en detergentes.

30 Depósito de material biológico

[0220] Los siguientes materiales biológicos han sido depositados según las condiciones del tratado de Budapest con el DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig; Alemania), y ha dado los siguientes números de registro:

Depósito	Fecha de acceso	Fecha de depósito
35 <i>Nocardioopsis</i> sp.	DSM 16424	24 de mayo de 2004
<i>Nocardioopsis prasina</i>	DSM 15649	30 de mayo de 2003
Prasina de <i>Nocardioopsis</i> (previamente <i>alba</i>)	DSM 14010	20 de enero de 2001

40 [0221] Estas cepas han sido depositadas bajo condiciones que aseguran que acceso al cultivo estará disponible durante la pendencia de esta solicitud de patente para uno determinado por el comisario de patentes y marcas registradas para ser titulado aquí bajo 37 C.F.R. §1.14 and 35 U.S.C. §122. El depósito representa un cultivo substancialmente puro de la cepa depositada. El depósito está disponible según sea necesario por leyes de patentes extranjeras en países donde contrapartes de la presente solicitud, o su progenie son solicitados. No obstante, se debe entender que la disponibilidad de un depósito no constituye una licencia para practicar la presente invención en derogación de los derechos de patente concedida por acción gubernamental.

45 [0222] Cepa DSM 15649 fue aislada en 2001 de una muestra de suelo de Dinamarca.

[0223] Las siguientes cepas están disponibles públicamente de DSM

Nocardiosis dassonvillei subesp. *dassonvillei* DSM 43235

Nocardiosis alkaliphila DSM 44657

Nocardiosis lucentensis DSM 44048

- 5 [0224] *Nocardiosis dassonvillei* subesp. cepa de *dassonvillei* DSM 43235 fue también depositada en otras instituciones de depósito de la siguiente manera: ATCC 23219, IMRU 1250, NCTC 10489.

Ejemplos

Ejemplo 1: Clonación y expresión de tres proteasas (L1a, L1b y L1c)

[0225]

Reactivos y medios

LB agar	Descrito en Ausubel, F. M. et al. (eds.) "Current protocols in Molecular Biology" John Wiley and Sons, 1995;
LB-PG agar	LB agar complementado con 0,5% de glucosa y 0,05 M de fosfato potásico, pH 7,0
PS-1	10% de sacarosa, 4% de harina de soja, 1% de Na ₃ PO ₄ · 12H ₂ O, 0,5% de CaCO ₃ y 0,01% ácido plurónico
TE	10 mM de Tris-HCl, pH 7,4 1 mM de EDTA, pH 8,0
TEL	50 mg/ml de Lisozym en el tampón TE
Thiocyanate	5M de tiocianato de guanidinio 100 mM de EDTA 0,6 % p/v de N-laurilsarcosina, sal de sodio 60 g thiocyanate, 20 ml 0,5 M de EDTA, pH 8,0, 20 ml de H ₂ O se disuelve a 65°C. Refrigeración a temperatura ambiente (RT) y adición de 0,6 g de N-laurilsarcosina. Adición de H ₂ O a 100 ml y filtrado a través de un filtro estéril 0,2 µ.
NH ₄ Ac	7,5 M de CH ₃ COONH ₄
TER	1 µg/ml de ribonucleasa pancreática en el tampón TE
CIA	cloroformo/isoamil alcohol 24:1

10

Fermentación de cepas de *Nocardiosis*

[0226] Cada una de las cepas de *Nocardiosis dassonvillei* subesp. *dassonvillei* DSM 43235, *Nocardiosis prasina* DSM 15649 y *Nocardiosis prasina* (previamente *alba*) DSM 14010 se cultivaron durante 3 días antes de recogida, en el siguiente medio a 30°C:

15

	Tripticasa	20 g
	Extracto de levadura	5 g
	Ferrocianuro	6 mg
5	Sulfato de magnesio	15 mg
	Agua destilada	hasta 1000 ml

pH ajustado a 9 por adición de carbonato de sodio.

Preparación de ADN genómico

[0227] ADN genómico se aisló según el siguiente procedimiento:

- 10 1. Recogida de 1,5 ml de cultivo y resuspender en 100 µl de TEL.
Incubar a 37°C durante 30 min.
2. Adición de 500 µl de tampón de tiocianato y dejar a temperatura ambiente durante 10 min.
3. Adición 250 µl de NH₄Ca y dejar en hielo durante 10 min.
4. Adición de 500 µl de CIA y mezclar.
- 15 5. Transferencia a un micro-centrifugador y giro durante 10 min. a toda velocidad.
6. Transferencia de sobrenadante a un nuevo tubo de Eppendorf y añadir 0,54 de volumen de isopropanol frío.
Mezclar completamente.
7. Girado y lavado del granulado de ADN con 70% de EtOH.
- 20 8. Resuspensión el ADN genómico en 100 µl de TER.

Construcción de cepas de expresión de *Bacillus subtilis* Sav-L1a, Sav-L1b y Sav-L1c

[0228] La región de codificación para la proteasa pro-madura L1a (nucleótidos 88-1143 de la SEC ID nº: 1) se amplificaron con los siguientes cebadores 1424 y 1485 en el ADN genómico aislado de *Nocardiopsis dassonvillei* subesp. *dassonvillei* DSM 43235:

- 25 Cebador 1485 (SEC ID nº: 14): 5'-gcttttagttcatcgatcgcatcggtgacgaccgtaccggccgagccag-3'
Cebador 1424 (SEC ID nº: 15): 5'-ggagcggattgaacatgcatgactaaccgggtaccagggacagcc-3'

[0229] La región de codificación para la proteasa pro-madura L1b (nucleótidos 88-1149 de la SEC ID nº: 3) se amplificó con los siguientes cebadores 1751 y 1753 en el ADN genómico aislado de *Nocardiopsis prasina* DSM15649:

- 30 1751 (SEC ID nº: 16): 5'-gttcatcgatcgatcggtgtcaccgacccaccgagcc-3'
1753 (SEC ID nº: 17): 5'-ggagcggattgaacatgcatgactgctggtgacgaggctgaggttc-3'

[0230] La región de codificación para la proteasa pro-madura L1c (nucleótidos 88-1149 de la SEC ID nº: 5) se amplificó con los siguientes cebadores 1755 y 1756 en el ADN genómico aislado de *Nocardiopsis prasina* DSM14010:

- 35 1755 (SEC ID nº: 18): 5'-gttcatcgatcgatcggtgtgaccgccccgagcc-3'
1756 (SEC ID nº: 19): 5'-ggagcggattgaacatgcatgactgctggtgacgaggctgaggttc-3'

[0231] Cada uno de estos polinucleótidos L1a, L1b, y L1c se fundieron, por PCR, en el marco a un fragmento ADN heterólogo que codifica un péptido de señal Sav (SEC ID nº: 13).

- 40 [0232] Cepas de *Bacillus subtilis* designada Sav-L1 a, Sav-L1b, y Sav-L1c, respectivamente, se construyeron por incorporación de estos genes (incluyendo la parte de codificación de péptido de señal) por recombinación homóloga en el genoma de *Bacillus subtilis* MB1053 de célula huésped (WO03/95658). Los genes se expresaron bajo el

5 control de un sistema promotor triple (como descrito en la WO 99/43835), que consiste en los promotores de gen de alfa-amilasa de *Bacillus licheniformis* (amyL), gen de alfa-amilasa de *Bacillus amyloliquefaciens* (amyQ), y el promotor cryIIIA de *Bacillus thuringiensis* incluyendo secuencia estabilizadora. La codificación de gen para acetiltransferasa de cloranfenicol fue usado como marcador (descrito en, por ejemplo, Diderichsen, B.; Poulsen, G.B.; Joergensen, S.T.; *A useful cloning vector for Bacillus subtilis. Plasmid 30:312* (1993)).

[0233] Transformantes resistentes al cloranfenicol se controlaron para actividad de proteasa en 1% de placas de agar LB-PG de leche desnatada (complementado con 6 µg/ml de cloranfenicol). Algunas colonias positivas de proteasa fueron además analizadas por secuenciación del ADN del inserto para confirmar la secuencia de ADN correcta, y una cepa para cada construcción se seleccionó.

10 **Fermentación de cepas de huésped de *Bacillus***

[0234] Cada una de las cepas huésped de *Bacillus subtilis* transformadas se fermentaron en una tabla de agitación giratoria (250 r.p.m.) en matraces de Erlenmeyer 500 ml disipados con 100 ml de medio PS-1 complementado con 6 µg/ml de cloranfenicol, a 37°C durante 16 horas, y a 26°C para 4 días extra.

Ejemplo 2: Clonación y expresión de proteasa L2a

15 [0235] La pro-forma de un gen de codificación de proteasa (nucleótidos 88-1152 de la SEC ID n°: 7) se aisló de *Nocardioopsis* sp. DSM 16424 por el procedimiento descrito en el ejemplo 1, salvo el uso de los siguientes cebadores:

1718 (SEC ID n°: 20): 5'-gttcatcgatcgatcggtgcccggccccgccccag-3'

1720 (SEC ID n°: 21): 5'-ggagcggattgaacatgcatgcatgctggtgcccgatgcaac-3'.

[0236] La proteasa correspondiente (SEC ID n°: 8) se designó L2a.

20 [0237] Una cepa huésped de *Bacillus subtilis* designada Sav-L2a se construyó, como también generalmente se describe en el ejemplo 1, y una colonia positiva de proteasa resistente al cloranfenicol se seleccionó y se analizó por secuenciación del ADN del inserto.

Ejemplo 3: Clonación de dos proteasas adicionales

25 [0238] Las pro-formas de dos genes de codificación de proteasas adicionales (nucleótidos 88-1149 de la SEC ID n°: 9, y nucleótidos 88-1152 de la SEC ID n°: 11, respectivamente) se aislaron de *Nocardioopsis alkaliphila* DSM 44657, y de *Nocardioopsis lucentensis* DSM 44048, respectivamente, por el procedimiento descrito en el ejemplo 1, salvo el uso de cebadores 1728 y 1763; y 1747 y 1749, respectivamente:

1728 (SEC ID n°: 22): 5'-gttcatcgatcgatcggtgccccgccccagtc-3'

1763 (SEC ID n°: 23): 5'-ggagcggattgaacatgcatgtaggtgcccagacgagccccca-3';

30 1747 (SEC ID n°: 24): 5'-gttcatcgatcgatcggtggaaccgtaccacccccag-3'

1749 (SEC ID n°: 25): 5'-ggagcggattgaacatgcatgtagctggtgcccagtcgac-3'

[0239] Las proteasas correspondientes (SEC ID n°: 10 y 12, respectivamente) se designaron L2b, y L2c, respectivamente.

35 [0240] Cepas huésped de *Bacillus subtilis* designadas Sav-L2b y Sav-L2c, respectivamente, se construyen, como también se describe generalmente en el ejemplo 1, y, colonias positivas de proteasa resistentes al cloranfenicol se seleccionan y se analizan por secuenciación del ADN de los insertos.

Ejemplo 4: Purificación y caracterización de la proteasa L1a

Ensayos de proteasa

[0241]

40 1) ensayo pNA:

sustrato de pNA: Suc-AAPF-pNA (Bachem L-1400).

Temperatura: temperatura ambiente (25°C)

Tampones de ensayo: 100mM de ácido succínico, 100mM de HEPES, 100mM de CHES, 100mM de CABS, 1 mM de CaCl₂, 150mM de KCl, 0,01% de Tritón X-100 ajustado a valores de pH 2,0, 2,5, 3,0, 3,5, 4,0, 5,0, 6,0, 7,0, 8,0, 9,0, 10,0, 11,0 y 12,0 con HCl o NaOH.

45

20µl de proteasa (diluida en 0,01% de Tritón X-100) se mezcla con 100µl de tampón de ensayo.

El ensayo se comienza por adición de 100µl de sustrato pNA (50mg disuelto en 1,0ml de DMSO y además diluido 45x con 0,01% de Tritón X-100).

El aumento en OD405 se monitorea como una medida de la actividad de proteasa.

5 2) Ensayo de protazyme AK:

Sustrato: comprimido de Protazyme AK (reticulado y caseína teñida; de Megazyme)

Temperatura: controlada (ensayo temperatura).

10 Tampones de ensayo: 100mM de ácido succínico, 100mM de HEPES, 100mM de CHES, 100mM de CABS, 1 mM de CaCl₂, 150mM de KCl, 0,01% de Tritón X-100 ajustado a valores de pH 2,0, 2,5, 3,0, 3,5, 4,0, 5,0, 6,0, 7,0, 8,0, 9,0, 10,0 y 11,0 con HCl o NaOH.

15 Un comprimido de Protazyme AK se suspende en 2,0ml de Tritón X-100 0,01% por agitación suave. 50OR I de esta suspensión y 500µl de tampón de ensayo se mezclan en un tubo de Eppendorf y se colocan en hielo. 20µl de muestra de proteasa (diluida en Tritón X-100 0,01%) se añade. El ensayo se inicia por transferencia del tubo de Eppendorf a un termomezclador Eppendorf, que se fija a la temperatura de ensayo. El tubo se incuba durante 15 minutos en el termomezclador Eppendorf a su índice de agitación máximo (1400 r.p.m.). La incubación se detiene por transferencia del tubo de nuevo al baño de hielo. Luego, el tubo se centrifuga en un centrifugador *icecold* durante unos minutos y 200µl de sobrenadante se transfirieron a una placa de microtitulación. OD650 se lee como una medida de actividad de proteasa. Un tampón ciego se incluye en el ensayo (en vez de enzima).

Purificación

20 [0242] El huésped de *Bacillus* transformado que expresa la L1 una proteasa descrita en el ejemplo 1 se fermentó como se describe también en el ejemplo 1, pero a 26°C durante 6 días. El caldo de cultivo se centrifugó (20000 x g, 20 min) y los sobrenadantes se decantaron cuidadosamente de los precipitados. Los sobrenadantes combinados se filtraron a través de una placa Seitz EKS para eliminar las células huésped de *Bacillus* restantes. El EKS filtrado se transfirió a 50mM de H₃BO₃, 5mM de ácido succínico, 1mM de CaCl₂, pH 7 en una columna Sefadex G25. Sulfato amónico sólido se añadió a la solución enzimática de la columna Sefadex G25 para dar una concentración final de 25 1,6M de (NH₄)₂SO₄ en la solución enzimática. La solución enzimática se mezcló suavemente con un agitador magnético durante la adición de (NH₄)₂SO₄ y la agitación se continuó durante 30 minutos después de la adición para llevar al sistema al equilibrio. Luego, la solución enzimática se aplicó a una columna de Butil Toyopearl equilibrada en 100mM de H₃BO₃, 10mM de ácido succínico, 2mM de CaCl₂, 1,6M de (NH₄)₂SO₄, pH 7. Después del lavado de la 30 columna extensivamente con el tampón de equilibrio, la proteasa se eluyó con un gradiente lineal (NH₄)₂SO₄ (1,6 a 0M) en el mismo tampón. Proteasa con fracciones se agruparon y se transfirieron a 20mM de HEPES, pH 8 en una columna Sefadex G25 y aplicada a una columna Q sefarosa FF equilibrada en el mismo tampón. Después del lavado de la columna extensivamente con el tampón de equilibrio, la proteasa se eluyó con un gradiente de NaCl lineal (0 a 0,5M) en el mismo tampón. Fracciones de la columna se analizaron para actividad de proteasa (usando el 35 ensayo Suc-AAPF-pNA a pH 9) y fracciones activas se analizaron además por SDS-PAGE. Fracciones con una única banda (como juzgada por un gel de SDS-PAGE manchado de coomassie) se agruparon para proporcionar la preparación purificada que se usó para otra caracterización.

40 [0243] La proteasa L1a se caracterizó como se describe abajo, en comparación con la otra proteasa derivada de *Nocardiaopsis dassonvillei* subesp. *dassonvillei* DSM 43235, preparada: como se describe en la WO 2004/111220 (en lo que sigue para abreviar designado "la proteasa L2").

Actividad de pH, estabilidad de pH y actividad de temperatura

[0244] El ensayo pNA se usó para el obtener el perfil de la actividad de pH al igual que el perfil de estabilidad de pH.

45 Para el perfil estabilidad de pH, la proteasa se diluyó 10x en los tampones de ensayo y se incubó durante 2 horas a 37°C. Tras la incubación, las muestras de proteasa se transfirieron al mismo pH (pH 9), antes de ensayo para actividad residual, por dilución en el tampón de ensayo pH. El ensayo Protazyme AK se usó para el obtener el perfil de actividad de temperatura a pH 9. Los resultados se muestran en tablas 1-3 de debajo.

Tabla 1: Perfil de actividad de pH

pH	Proteasa L1a	Proteasa L2
2	0,00	0,00
3	0,00	0,00
4	0,02	0,03
5	0,10	0,11
6	0,25	0,21
7	0,38	0,37
8	0,66	0,71
9	0,97	0,97
10	1,00	1,00
11	0,99	0,94
12	0,94	-

Tabla 2: Perfil de estabilidad de pH

pH	Proteasa L1a	Proteasa L2
2,0	0,50	1,00
2,5	0,81	0,95
3,0	0,93	0,97
3,5	0,94	1,01
4,0	0,97	0,98
5,0	0,96	0,97
6,0	0,95	0,98
7,0	0,99	0,96
8,0	0,97	0,99
9,0	0,93	0,99
10,0	0,94	0,96
11,0	0,94	0,94
12,0	0,92	0,84
9,0 y después 2 horas a 5 °C	1,00	1,00

Tabla 3: Perfil de actividad de temperatura

Temperatura (°C)	Proteasa L1a	Proteasa L2
15	0,11	0,01
25	0,17	0,01
37	0,30	0,03
50	0,58	0,09
60	0,90	0,19
70	1,00	0,63
80	0,34	1,00
90	-	0,35

Otras características

- 5 [0245] La proteasa L1a es una proteasa alfa-lítica como enzima (familia de peptidasa S1E, notación antigua S2A) que se descubre que se inhibe por fenil metil sulfonil fluoruro (PMSF) y por el inhibidor de subtilisina de Streptomyces (SSI). Su peso molecular relativo como determinado por SDS-PAGE es Mr = 22kDa, y la secuencia N-terminal: ADIVGGEAY (SEC ID n°: 26).

Ejemplo 5: Actividad específica de la proteasa L1a

- 10 [0246] La preparación de proteasa purificada descrita en el ejemplo 4 se usó para determinación de la actividad específica. La pureza de la preparación fue de sobre 95% cuando se analizó por SDS-PAGE (determinado como se describe en el ejemplo 2A en la WO 01/58275). La muestra de proteasa se dividió en dos. Una parte se analizó para el contenido de proteína (mg/ml) por análisis de aminoácido, la otra parte se analizó para actividad de proteasa.

Análisis de aminoácido (AAA)/(mg/ml)

- 15 [0247] Los enlaces peptídicos de la muestra de proteasa se sometieron a hidrólisis ácida, seguido de separación y cuantificación de los aminoácidos liberados en un Biochrom 20 Plus Amino Acid Analyser, disponible comercialmente de Bie & Bentsen A/S, Sandbaekvej 5-7, DK-2610 Roedovre, Dinamarca, según las instrucciones del fabricante. Para la hidrólisis ácida, la muestra de proteína se secó en un centrifugador de vacío, se resolvió en 18,5% (vol/vol) de HCl + 0,1% (vol/vol) de fenol y se incubó durante 16hr a 110°C. Tras la incubación, la muestra se secó otra vez en el centrifugador de vacío, se resolvió en el tampón de carga (0,2 M de Na-citrato, pH 2,2) y se cargó sobre el Biochrom 20 Plus Amino Acid Analyser.

- 20 [0248] Para la cuantificación, la muestra hidrolizada se cargó sobre una columna de la resina de intercambio de catión UltroPac n°. 8, forma de sodio, que está disponible comercialmente de Bie & Bentsen A/S, catalogue no. 80-2104-15. Tampones de pH variable (de pH 1 a pH 8) y fuerza iónica se bombearon a través de la columna según las instrucciones del fabricante referido anteriormente, para separar varios aminoácidos. La temperatura de columna se controló con precisión, también según las instrucciones del fabricante (de 53°C a 92°C y de nuevo a 53°C) para asegurar la separación requerida. El eluyente de columna se mezcló con reactivo de ninhidrina (Bie & Bentsen, catálogo n°. 80-2038-07) y la mezcla se pasó a través de la bobina de reacción de alta temperatura del analizador de aminoácido. En la bobina de reacción, ninhidrina reaccionada con los aminoácidos para formar compuestos coloreados, la cantidad de lo cual fue directamente proporcional a la cantidad de aminoácido presente.

Ensayo de actividad de proteasa (AU/ml)

- 25 [0249] Hemoglobina desnaturada (0,65% (p/p) en 6,7mM de tampón KH₂PO₄/NaOH, pH 7,50) se degradó a 25°C durante 10 minutos por la proteasa y hemoglobina no asimilada se precipitó con (TCA) de ácido tricloroacético y se quitó por filtración. Los productos ATC-solubles de degradación de hemoglobina en el filtrado se determinaron con reactivo fenol de Folin & Ciocalteu, que da un color azul con diferentes aminoácidos. La unidad de actividad (AU) se definió y se midió por referencia a un estándar de ALCALASE™. Una descripción detallada del ensayo, al igual que

una muestra del estándar de ALCALASE™, está disponible a petición de Novozymes A/S, Krogshoejvej 36, DK-2880 Bagsvaerd, Dinamarca (ensayo n°. EB-SM-0349.02/01).

[0250] La actividad específica se calculó como: actividad específica (AU/g) = (actividad (AU/ml) / AAA (mg/ml)) x 1000 (mg/g).

- 5 [0251] La actividad específica de proteasa L1a fue 49,8 AU/g, en comparación con la actividad específica de la proteasa derivada de *Nocardioopsis* sp. NRRL 18262 de 38,3 AU/g.

Ejemplo 6: Purificación y caracterización de la proteasa L2a

10 [0252] El huésped de *Bacillus* transformado que expresa la proteasa L2a descrita en el ejemplo 2 se fermentó como se describe en el ejemplo 1, pero a 30°C durante 5 días. El caldo de cultivo se centrifugó (20000 x g, 20 min) y los sobrenadantes se decantaron cuidadosamente de los precipitados. Los sobrenadantes combinados se filtraron a través de una placa Seitz EKS para eliminar las células huésped de *Bacillus* restantes. El filtrado EKS se transfirió a 50mM de H₃BO₃, 5mM de ácido succínico, 1 mM de CaCl₂, pH 7 en una columna Sefadex G25 y se aplicó a una columna de sílice de bacitracina equilibrada en el mismo tampón. Después del lavado de la columna de bacitracina extensivamente con el tampón de equilibrio, la proteasa se eluyó por fases con 100mM de H₃BO₃, 10mM de ácido succínico, 2mM de CaCl₂, 1 M de NaCl, 25% de isopropanol, pH 7. El eluato de bacitracina se transfirió a 50mM de H₃BO₃, 10mM de CH₃COOH, 1mM de CaCl₂, pH 4,5 en una columna Sefadex G25 y se aplicó a una columna S sefarosa HP equilibrada en el mismo tampón. Después del lavado de la columna extensivamente con el tampón de equilibrio, la proteasa se eluyó con un gradiente de NaCl lineal (0 a 0,5M) en el mismo tampón. Fracciones de la columna se analizaron para actividad de proteasa (usando el ensayo Protazyme AK a 37°C y pH 9) y fracciones activas se analizaron además por SDS-PAGE. Fracciones con una única banda se agruparon (como juzgada por un gel de SDS-PAGE manchado de coomassie) para proporcionar la preparación purificada que fue usada para otra caracterización.

25 [0253] La proteasa L2a se caracterizó como se describe en el ejemplo 4 de arriba, en comparación con una proteasa conocida derivada de *Nocardioopsis* sp. NRRL 18262 (para abreviar designada "Proteasa 10"). Los resultados se muestran en las tablas 4-6 de debajo.

Tabla 4: Perfil de actividad pH

pH	Proteasa L2a	Proteasa 10
2	0,00	-
3	0,00	0,00
4	0,02	0,02
5	0,10	0,07
6	0,22	0,21
7	0,41	0,44
8	0,75	0,67
9	0,97	0,88
10	0,99	1,00
11	1,00	0,93
12	0,85	-

Tabla 5: Perfil de estabilidad de pH

pH	Proteasa L2a	Proteasa 10
2,0	0,67	0,78
2,5	0,93	1,00
3,0	0,95	1,03
3,5	0,96	0,98
4,0	0,97	0,99
5,0	0,94	1,02
6,0	0,95	1,00
7,0	0,97	1,01
8,0	0,96	0,98
9,0	0,95	0,99
10,0	0,96	0,99
11,0	0,90	0,86
12,0	0,60	-
y después 2 horas a 5 °C	1,00	1,00

Tabla 6: Perfil de estabilidad de pH

Temperatura (°C)	Proteasa <u>L2a</u>	Proteasa 10
15	0,02	0,02
25	0,02	0,02
37	0,05	0,07
50	0,13	0,20
60	0,31	0,51
70	0,79	1,00
80	1,00	0,39
90	0,28	-

Otras características

[0254] La proteasa L2a es una proteasa alfa-lítica como enzima (familia de peptidasa S1E, notación antigua S2A) que se ha descubierto que se inhibe por metil fenil sulfonil fluoruro (PMSF). Su peso molecular relativo como determinado por SDS-PAGE es Mr = 20kDa y la secuencia N-terminal: ANIIGGLAYT (SEC ID n°: 27).

Ejemplo 7: Temperatura de fusión de la proteasa L2a

5 Calorimetría diferencial de barrido (DSC)

[0255] Calorimetría por análisis diferencial se usó para determinar estabilidad de temperatura a pH 7,0 de la proteasa L2a derivada de *Nocardioopsis* sp. DSM 16424. La proteasa se purificó como se describe en el ejemplo 6 y se dializó durante la noche a 4°C contra 10 mM de fosfato sódico, 50 mM de cloruro sódico, pH 7,0 y se condujo en un instrumento de calorimetría por análisis diferencial de vasopresina (Micro Cal) con una tasa de barrido constante de 1,5°C/min de 20 a 100°C. Manipulación de datos se realizó usando el software Microcal Origin.

[0256] La desnaturalización resultante o temperatura de fusión (T_m o T_d), fue 78,2°C. La T_m para proteasa 10 es 76,5°C.

Ejemplo 8: Rendimiento de la proteasa L2a en un modelo de digestión monogástrica *in vitro*

[0257] El rendimiento de la proteasa L2a purificada descrita en ejemplo el 6 se evaluó en un modelo *in vitro* que simula la digestión en animales monogástricos, en comparación con la proteasa conocida derivada de *Nocardioopsis* sp. NRRL 18262 ("Proteasa 10"). En particular, la proteasa se evaluó para su capacidad para mejorar solubilización y digestión de maíz-SBM (comida de maíz/semilla de soja) proteínas. El sistema *in vitro* consistía en 18 matraces en el que el sustrato maíz/- SBM se incubó inicialmente con HCl/pepsina, simulación de digestión gástrica, y posteriormente con pancreatina, simulación de digestión intestinal. Ocho de los matraces se dosificaron con la proteasa al principio de la fase gástrica mientras que los diez matraces restantes sirvieron como blancos. Al final de la fase de incubación intestinal, muestras de digesta *in vitro* se quitaron y se analizaron para proteína digerida y solubilizada.

Perfil del procedimiento de digestión *in vitro*

[0258]

Componentes añadidos	pH	Temperatura	Curso del tiempo	Fase de digestión simulada
10 g de sustrato de maizel-SBM (6:4), 41 ml de HCl (0,105M)	3,0	40°C	t=0 min	Mezcla
5 ml de HCl (0, 1 05M) / pepsina (3000 U/g de sustrato), 1 ml de proteasa (para proporcionar 100 mg de enzima proteasa por kg de sustrato)	3,0	40°C	t=30 min	Digestión gástrica
16 ml de H ₂ O	3,0	40°C	t= 1,0 hora	Digestión gástrica
7 ml de NaOH (0,39M)	6,8	40°C	t=1,5 horas	Digestión intestinal
5 ml de NaHCO ₃ (1 M) / pancreatina (8 mg/g de dieta)	6.8	40°C	t=2,0 horas	Digestión intestinal
Terminación de la incubación	7,0	40°C	t=6,0 horas	

Condiciones

[0259]

- Sustrato: 4 g de SBM, 6 g de maíz (premezclado)
- pH: 3,0 fase estomacal / 6,8-7,0 fase intestinal
- HCl: 0,105 M durante 1,5 horas (es decir, 30 min HCl-sustrato premezcla)

pepsina: 3000 U /g dieta durante 1 hora
 pancreatina: 8 mg/g dieta durante 4 horas
 temperatura: 40°C
 Réplicas: n

5 Soluciones

[0260] 0,39 M de NaOH

0,105 M de HCl

0,105 M de HCl con 6000 U de Pepsina por 5 ml

1 M de NaHCO₃ con 16 mg de pancreatina por ml

10 125 mM de NaAc-tampón, pH 6,0

Determinaciones de proteína enzimática

[0261] La cantidad de proteína de enzima proteasa (EP) se calcula basándose en los valores A280 y las secuencias de aminoácidos (composiciones de aminoácido) usando los principios perfilados en S.C.Gill & P.H. von Hippel, Analytical Biochemistry 182, 319-326, (1989).

15 Procedimiento experimental para modelo *in vitro*

[0262] El procedimiento experimental fue según el anterior perfil. El pH se midió a tiempo de 1, 2.5, y 5,5 horas.

Incubaciones se terminaron después de 6 horas y muestras de 30 ml se quitaron y se colocaron en el hielo antes de centrifugado (10000 x g, 10 min, 4°C). Los sobrenadantes se quitaron y se almacenaron a -20°C.

Análisis

20 [0263] Todas las muestras se analizaron para el contenido proteína digerida y solubilizada usando filtración en gel.

Estimación de proteína digerida y solubilizada

[0264] El contenido de proteína solubilizada en sobrenadantes de muestras digeridas *in vitro* se estimó por cuantificación de proteína cruda (CP) usando filtración en gel HPLC. Los sobrenadantes se descongelaron, se filtraron a través de filtros de policarbonato 0,45 µm y se diluyeron (1:50 V/V) con H₂O. Muestras diluidas se 25 cromatografiaron por HPLC usando una columna de Superdex Peptide PE (7.5 x 300 mm) de filtración en gel (Global). El eluyente usado para elución isocrática fue 50 mM de tampón de fosfato de sodio (pH 7,0) con 150 mM de NaCl. El volumen total de eluyente por ciclo fue 26 ml y la velocidad de flujo fue 0,4 ml/min. Perfiles de elución se registraron a 214 nm y el área total bajo los perfiles se determinó por integración. Para estimar contenido de proteína de áreas integradas, una curva de calibración (R²=0,9993) se realizó de una serie de dilución de una muestra de 30 maíz-SBM de referencia digerida *in vitro* con contenido de proteína conocido. La determinación de proteína en esta muestra de referencia se llevo a cabo usando el método Kjeldahl (*determination of % nitrogen; A.O.A.C. (1984) Official Methods of Analysis 14th ed., Washington DC*).

[0265] El contenido de proteína digerida se estimó integrando la cromatografía de área correspondiente a péptidos y aminoácidos con una masa molecular de 1500 Dalton o menos (Savoie,L.; Gauthier,S.F. *Dialysis Cell For The In-vitro Measurement Of Protein Digestibility. J. Food Sci. 1986, 51, 494-498*; Babinszky,L.; Van,D.M.J.M.; Boer,H.; Den,H.L.A. *An In-vitro Method for Prediction of The Digestible Crude Protein Content in Pig Feeds. J. Sci. Food Agr. 1990, 50, 173-178*; Boisen,S.; Eggum,B.O. *Critical Evaluation of In-vitro Methods for Estimating Digestibility in Simple-Stomach Animals. Nutrition Research Reviews 1991, 4, 141-162*). Para determinar la línea divisoria de 1500 dalton, la columna de filtración de gel se calibró usando citocromo C (Boehringer, Alemania), aprotinina, gastrina I y 40 sustancia P (Sigma Aldrich, EEUU), como estándares de masa molecular.

Resultados

[0266] Los resultados mostrados en la tabla 7 de debajo indican que la proteasa L2a, como proteasa 10, 45 significativamente aumentó el nivel de proteína digerible y soluble relativamente a la de blanco. Además, la proteasa L2a aparece por lo menos mejorar numéricamente el nivel de proteína digerible en comparación con la proteasa conocida 10.

Tabla 7: Proteína digerido y solubilizada cruda

Enzima	N	En relación a blanco					
		%digestible CP		CV%	%soluble CP		CV%
Blanco	10	100,0	a	5,5	100,0	a	4,4
Proteasa <u>L2a</u>	3	1161	b	0,7	107,2	b	1,1
Proteasa 10	5	112,1	b	1,0	110,2	b	0,6

Letras diferentes dentro de la misma columna indican diferencias significativas (ANOVA de 1-vía, prueba Tukey-Kramer, P<0,05). SD = desviación típica. %CV = coeficiente de varianza = (SD/medio valor) x 100%

Ejemplo 9: Alimento para animales y aditivos de alimento

- 5 [0267] Un aditivo de alimento comprendiendo proteasa L2a de la invención, en forma de una premezcla de vitaminas y minerales, está compuesto como se muestra en la tabla 8 de debajo. Las vitaminas y los carotenoides están disponibles comercialmente de DSM Nutritional Products. Todas las cantidades están en g/kg.

Tabla 8: Composición premezcla

Vitamina A	ROVIMIX A 500	4,00
Vitamina D3	ROVIMIX D3 500	1,00
Vitamina E	ROVIMIX E 50 Ads	8,00
Vitamina B2	ROVIMIX B2 80-SD	1,0
	CAROPHYLL Amarillo	10,0
	Cloruro de colina 50%, min.	300,0
Minerales	Óxido Mn	60,0
	Óxido Zn	12,0
	Monohidrato de sulfato Fe	20,0
	Óxido Cu	2,0
	Sulfato Co	0,2
Enzima	Proteasa L2a (proteína enzimática)	10,0
	Harinas de trigos	571,8

- 10 [0268] La premezcla de la tabla 8 se incluye en una dieta para estratos con una composición como se muestra en la tabla 9 de debajo. La cantidad de cada ingrediente se indica en % (p/p). La concentración en la dieta de la proteasa L2a es 100 mg de proteína de enzima proteasa por kg de la dieta.

Tabla 9: Dieta para estratos

Maíz	55,00
Trigo	10,00
Avena	7,50
Soja	20,00
Piedra caliza	7,50
Premezcla de la tabla 8	1,00

LISTADO DE SECUENCIAS

- 5 [0269]
 <110> Novozymes A/S
 <120> Proteasas
 <130> 10644-204-WO
 <160> 27
- 10 <170> Versión de patentIn 3.2
 <210> 1
 <211> 1146
 <212> ADN
 213> Nocardopsis dassonvillei subesp. dassonvillei DSM 43235
- 15 <220>
 <221> CDS
 <222> (1)..(1143)
 <220>
 <221> sig_peptide
- 20 <222> (1)..(87)
 <220>
 <221> mat_peptide
 <222> (568)..(1143)
 <400> 1
- 25

ES 2 380 105 T3

atg	cga	ccc	tcc	ccc	gct	atc	tcc	gct	atc	ggc	acc	ggc	gca	ctc	45
Met	Arg	Pro	Ser	Pro	Ala	Ile	Ser	Ala	Ile	Gly	Thr	Gly	Ala	Leu	
				-185					-180					-175	
gcg	ttc	ggt	ctg	gcg	ttc	tcc	gtg	acg	ccg	ggc	gcc	agt	gcg	gcg	90
Ala	Phe	Gly	Leu	Ala	Phe	Ser	Val	Thr	Pro	Gly	Ala	Ser	Ala	Ala	
				-170					-165					-160	
acc	gta	ccg	gcc	gag	cca	gcg	agc	gag	gcc	cag	acg	atg	atg	gaa	135
Thr	Val	Pro	Ala	Glu	Pro	Ala	Ser	Glu	Ala	Gln	Thr	Met	Met	Glu	
				-155					-150					-145	
gcg	ctg	cag	aga	gac	ctc	ggc	ctc	acc	ccg	ctc	ggg	gcc	gag	gag	180
Ala	Leu	Gln	Arg	Asp	Leu	Gly	Leu	Thr	Pro	Leu	Gly	Ala	Glu	Glu	
				-140					-135					-130	
ctg	ctc	tcg	gcg	cag	gaa	gag	gcg	atc	gag	acc	gac	gcc	gag	gcc	225
Leu	Leu	Ser	Ala	Gln	Glu	Glu	Ala	Ile	Glu	Thr	Asp	Ala	Glu	Ala	
				-125					-120					-115	
acc	gag	gcc	gcg	gga	gcg	tcc	tac	ggc	ggc	tcc	ctg	ttc	gac	acc	270
Thr	Glu	Ala	Ala	Gly	Ala	Ser	Tyr	Gly	Gly	Ser	Leu	Phe	Asp	Thr	
				-110					-105					-100	
gag	acc	ctc	cag	ctc	acc	gtg	ctg	gtg	acc	gac	gcc	tcg	gcc	gtc	318
Glu	Thr	Leu	Gln	Leu	Thr	Val	Leu	Val	Thr	Asp	Ala	Ser	Ala	Val	Glu
				-95					-90					-85	
gcg	gtg	gag	gcc	acc	ggc	gcc	gag	gcc	acc	gtg	gtc	tca	cac	ggc	366
Ala	Val	Glu	Ala	Thr	Gly	Ala	Glu	Ala	Thr	Val	Val	Ser	His	Gly	Ala
			-80					-75					-70		
gag	ggc	ctg	gcc	gag	gtg	gtc	gac	gcg	ctc	gac	gag	acc	ggc	ggc	414
Glu	Gly	Leu	Ala	Glu	Val	Val	Asp	Ala	Leu	Asp	Glu	Thr	Gly	Gly	Arg
		-65					-60					-55			
gaa	ggg	gtc	gtc	ggc	tgg	tac	ccg	gac	gtg	gag	agc	gac	acc	gtc	462

Glu -50	Gly	Val	Val	Gly	Trp	Tyr -45	Pro	Asp	Val	Glu	Ser -40	Asp	Thr	Val	Val	
gtc Val -35	cag Gln	gtc Val	gcc Ala	gag Glu	ggc Gly -30	gcc Ala	agc Ser	gcc Ala	gac Asp	ggc Gly -25	ctc Leu	atc Ile	gag Glu	gcc Ala	gcg Ala -20	510
ggc Gly	gtg Val	gac Asp	ccc Pro	tcc Ser -15	gcc Ala	gtc Val	cgg Arg	gtg Val	gag Glu -10	gag Glu	acc Thr	agt Ser	gag Glu	act Thr -5	ccg Pro	558
cgc Arg	ctg Leu	tac Tyr -1	gcc Ala 1	gac Asp	atc Ile	gtc Val	ggc Gly 5	ggc Gly	gag Glu	gcg Ala	tac Tyr 10	tac Tyr 10	atg Met	ggc Gly	ggc Gly	606
gga Gly	cgc Arg 15	tgc Cys	tcg Ser	gtc Val	ggg Gly	ttc Phe 20	gcc Ala	gtg Val	acc Thr	gac Asp	ggc Gly 25	tcc Ser	ggc Gly	gcg Ala	ggc Gly	654
ggc Gly 30	ttc Phe	gtg Val	acg Thr	gcg Ala	ggc Gly 35	cac His	tgc Cys	ggc Gly	acc Thr	gtc Val 40	ggc Gly	acc Thr	ggc Gly	gcc Ala	gag Glu 45	702
agc Ser	tcc Ser	gac Asp	ggc Gly	agc Ser 50	ggc Gly	tcc Ser	gga Gly	acc Thr	ttc Phe 55	cag Gln	gag Glu	tcc Ser	gtc Val	ttc Phe 60	ccg Pro	750
ggc Gly	agc Ser	gac Asp	ggc Gly 65	gcc Ala	ttc Phe	gtc Val	gcg Ala	gcc Ala 70	acc Thr	tcc Ser	aac Asn	tgg Trp	aac Asn 75	gtg Val	acc Thr	798
aac Asn	ctg Leu	gtc Val 80	agc Ser	cgg Arg	tac Tyr	gac Asp	tcc Ser 85	ggc Gly	agc Ser	ccc Pro	cag Gln	gcg Ala 90	gtg Val	tcg Ser	ggt Gly	846
tcc Ser	agc Ser 95	cag Gln	gcc Ala	ccg Pro	gag Glu	ggc Gly 100	tcg Ser	gcg Ala	gtg Val	tgc Cys	gcg Arg 105	tcc Ser	ggc Gly	tcc Ser	acc Thr	894
acc Thr 110	ggc Gly	tgg Trp	cac His	tgc Cys	ggg Gly 115	acc Thr	atc Ile	gag Glu	gcc Ala	gcg Arg 120	ggc Gly	cag Gln	acg Thr	gtg Val	aac Asn 125	942
tac Tyr	ccg Pro	cag Gln	ggc Gly	acg Thr 130	gtc Val	cag Gln	gac Asp	ctg Leu	acc Thr 135	cgg Arg	acg Thr	gac Asp	gtg Val	tgc Cys 140	gcc Ala	990
gag Glu	ccc Pro	ggt Gly	gac Asp 145	tcc Ser	ggc Gly	ggc Gly	tcg Ser	ttc Phe 150	atc Ile	gcc Ala	ggt Gly	tcg Ser	cag Gln 155	gcc Ala	cag Gln	1038
ggc Gly	gtc Val 160	acc Thr	tcc Ser	ggc Gly	ggc Gly	tcg Ser	ggc Gly 165	aac Asn	tgc Cys	acc Thr	tcc Ser	ggc Gly 170	ggc Gly	acg Thr	acc Thr	1086
tac Tyr	tac Tyr 175	cag Gln	gag Glu	gtc Val	act Thr	ccc Pro 180	ctg Leu	ctg Leu	agc Ser	agc Ser	tgg Trp 185	ggg Gly	ctg Leu	tcc Ser	ctg Leu	1134
gtg Val 190	acc Thr	ggt Gly	tag													1146

<211> 381

<212> PRT

<213> *Nocardiopsis dassonvillei* subesp. *dassonvillei* DSM 43235

<400> 2

Met Arg Pro Ser Pro Ala Ile Ser Ala Ile Gly Thr Gly Ala Leu
 -185 -180 -175

Ala Phe Gly Leu Ala Phe Ser Val Thr Pro Gly Ala Ser Ala Ala
 -170 -165 -160

Thr Val Pro Ala Glu Pro Ala Ser Glu Ala Gln Thr Met Met Glu
 -155 -150 -145

Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Leu Gly Ala Glu Glu
 -140 -135 -130

Leu Leu Ser Ala Gln Glu Glu Ala Ile Glu Thr Asp Ala Glu Ala
 -125 -120 -115

Thr Glu Ala Ala Gly Ala Ser Tyr Gly Gly Ser Leu Phe Asp Thr
 -110 -105 -100

Glu Thr Leu Gln Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val Glu
 -95 -90 -85

Ala Val Glu Ala Thr Gly Ala Glu Ala Thr Val Val Ser His Gly Ala
 -80 -75 -70

Glu Gly Leu Ala Glu Val Val Asp Ala Leu Asp Glu Thr Gly Gly Arg
 -65 -60 -55

Glu Gly Val Val Gly Trp Tyr Pro Asp Val Glu Ser Asp Thr Val Val
 -50 -45 -40

Val Gln Val Ala Glu Gly Ala Ser Ala Asp Gly Leu Ile Glu Ala Ala
 -35 -30 -25 -20

Gly Val Asp Pro Ser Ala Val Arg Val Glu Glu Thr Ser Glu Thr Pro
 -15 -10 -5

Arg Leu Tyr Ala Asp Ile Val Gly Gly Glu Ala Tyr Tyr Met Gly Gly
 -1 1 5 10

Gly Arg Cys Ser Val Gly Phe Ala Val Thr Asp Gly Ser Gly Ala Gly
 15 20 25

Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Gly Ala Glu
 30 35 40 45

Ser Ser Asp Gly Ser Gly Ser Gly Thr Phe Gln Glu Ser Val Phe Pro
 50 55 60

Gly Ser Asp Gly Ala Phe Val Ala Ala Thr Ser Asn Trp Asn Val Thr

	65		70		75														
Asn	Leu	Val	Ser	Arg	Tyr	Asp	Ser	Gly	Ser	Pro	Gln	Ala	Val	Ser	Gly				
		80					85					90							
Ser	Ser	Gln	Ala	Pro	Glu	Gly	Ser	Ala	Val	Cys	Arg	Ser	Gly	Ser	Thr				
						100					105								
Thr	Gly	Trp	His	Cys	Gly	Thr	Ile	Glu	Ala	Arg	Gly	Gln	Thr	Val	Asn				
110					115					120					125				
Tyr	Pro	Gln	Gly	Thr	Val	Gln	Asp	Leu	Thr	Arg	Thr	Asp	Val	Cys	Ala				
				130					135					140					
Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Phe	Ile	Ala	Gly	Ser	Gln	Ala	Gln				
			145					150					155						
Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly	Asn	Cys	Thr	Ser	Gly	Gly	Thr	Thr				
		160					165					170							
Tyr	Tyr	Gln	Glu	Val	Thr	Pro	Leu	Leu	Ser	Ser	Trp	Gly	Leu	Ser	Leu				
	175					180					185								

Val Thr Gly
190

<210> 3

<211> 1152 <212> ADN

<213> Nocardiosis prasina DSM 15649

5 <220>

<221> CDS

<222> (1)..(1149)

<220>

<221> sig_peptide

10 <222> (1)..(87)

<220>

<221> mat_peptide

<222> (574)..(1149)

<400> 3

ES 2 380 105 T3

atg	cga	ccc	tcc	ccc	gtc	atc	tcc	gcg	atc	ggc	acg	gga	gca	ctg	45
Met	Arg	Pro	Ser	Pro	Val	Ile	Ser	Ala	Ile	Gly	Thr	Gly	Ala	Leu	
	-190					-185					-180				
gcc	ttc	ggg	ctc	gcg	ctc	tcg	gtc	gcg	ccc	ggc	gcc	tcc	gcc	gtc	90
Ala	Phe	Gly	Leu	Ala	Leu	Ser	Val	Ala	Pro	Gly	Ala	Ser	Ala	Val	
	-175					-170					-165				
acc	gca	ccc	acc	gag	ccc	gcg	ccc	cag	ggc	gag	gcg	gcc	acc	atg	135
Thr	Ala	Pro	Thr	Glu	Pro	Ala	Pro	Gln	Gly	Glu	Ala	Ala	Thr	Met	
	-160					-155					-150				
cag	gaa	gcg	ctt	gag	agg	gac	ttc	ggc	ctc	acc	ccg	ttc	gag	gcc	180
Gln	Glu	Ala	Leu	Glu	Arg	Asp	Phe	Gly	Leu	Thr	Pro	Phe	Glu	Ala	

-145						-140						-135						
gaa gac	ctg ctc	gaa gcc	cag	aat gac	gct ctc	ggg	atc gac	acg										225
Glu Asp	Leu Leu	Glu Ala	Gln	Asn Asp	Ala Leu	Gly	Ile Asp	Thr										
-130			-125			-120												
gcg gcg	gcc aag	gcc gcc	ggt	gac gcc	tac gcg	ggc	tcc gtg	ttc										270
Ala Ala	Ala Lys	Ala Ala	Gly	Asp Ala	Tyr Ala	Gly	Ser Val	Phe										
-115			-110			-105												
gac acc	gac acc	ctg gaa	ctg	acc gtc	ctg ctc	acg gac	gcc gga	gcc										318
Asp Thr	Asp Thr	Leu Glu	Leu Thr	Val Leu	Leu Thr	Asp Ala	Gly Ala											
-100			-95			-90												
gtg tcg	gac gtc	gag gcc	acc	ggc gcc	ggg	acc	gaa ctg	gtc tcg	tac									366
Val Ser	Asp Val	Glu Ala	Thr	Gly Ala	Gly Thr	Glu	Leu Val	Ser Tyr										
-85		-80		-75														
ggc acc	gag ggc	ctg gcg	gag	atc atg	gac gag	ctc gac	gca gcc	ggc										414
Gly Thr	Glu Gly	Leu Ala	Glu Ile	Met Asp	Glu Leu	Asp Ala	Ala Ala	Gly										
		-65		-60														
gcc cag	ccg ggt	gtc gtc	ggc	tgg tac	ccg gac	ctc gcc	ggc gac	acc										462
Ala Gln	Pro Gly	Val Val	Gly Trp	Tyr Pro	Asp Leu	Ala Ala	Gly Asp	Thr										
		-50		-45			-40											
gtc gtc	atc gag	gcc acc	gac	acc tcc	gag gcc	cag agc	ttc gtc	gag										510
Val Val	Ile Glu	Ala Thr	Asp Thr	Ser Ser	Glu Ala	Gln Ser	Phe Val	Glu										
	-35		-30			-25												
gcc gcg	ggc gtg	gac tcc	tcc	gcc gtc	cag gtg	gag cag	acc gac	gag										558
Ala Ala	Gly Val	Asp Ser	Ser Ala	Val Val	Gln Val	Glu Gln	Thr Asp	Glu										
-20			-15		-10													
gcg ccg	cag ctg	tac gcc	gac	atc gtc	ggc ggt	gac gcc	tac tac	atg										606
Ala Pro	Gln Leu	Tyr Ala	Asp Ile	Val Val	Gly Gly	Asp Ala	Tyr Tyr	Met										
-5		-1	1		5		10											
ggc ggc	ggg cgc	tgc tgc	gtc gga	ttc gcg	gtc acc	gac agt	tcc ggc	ggc										654
Gly Gly	Gly Arg	Cys Ser	Val Gly	Phe Ala	Val Thr	Asp Ser	Ser Ser	Gly										
	15		20		25													
aac gac	gga ttc	gtg acg	gcc ggc	cac tgc	ggc acc	gac gtc	ggc acc	tcc										702
Asn Asp	Gly Phe	Val Thr	Ala Ala	Gly His	Cys Gly	Thr Val	Gly Thr	Ser										
	30		35		40													
gcc gac	agc gag	gac ggc	agc ggc	tcc ggt	gtg ttc	gag gag	tcc atc											750
Ala Asp	Ser Glu	Asp Thr	Gly Ser	Gly Ser	Val Val	Phe Glu	Glu Ile											
45			50		55													
ttc ccg	ggc aac	gac gcg	gcc ttc	gtc agt	tcg acc	gac tcc	aac tgg	acc										798
Phe Pro	Gly Asn	Asp Thr	Ala Ala	Phe Val	Ser Ser	Thr Thr	Asn Trp	Thr										
60			65		70		75											
gtc acc	aac ctg	gtc aac	atg tac	agc tgc	ggc ggt	ggc acc	cag tcc	gtc										846
Val Thr	Asn Leu	Val Asn	Met Tyr	Ser Ser	Gly Gly	Gly Thr	Gln Ser	Val										
		80		85			90											
ggc ggc	tcc agc	cag gcc	ccg gtc	ggc gcg	gcc gtc	tgc cgt	tcc ggc											894
Gly Gly	Ser Ser	Gln Ala	Pro Val	Gly Ala	Ala Val	Cys Arg	Ser Gly											
		95		100		105												
tcc acc	acg ggc	tgg cac	tgc ggg	tcc atc	gag gcc	cgc ggg	cag tcg											942
Ser Thr	Thr Gly	Trp His	Cys Gly	Ser Ser	Ile Glu	Ala Arg	Gln Ser											
	110		115		120													
gtg agc	tac ccg	gag gcc	acc gtc	acc gac	atg acc	cgt acc	gac gtc											990
Val Ser	Tyr Pro	Glu Gly	Thr Thr	Val Thr	Met Thr	Arg Thr	Asp Val											

125		130		135		
tgc gcc gag ccc ggc gac tcc ggc ggt tgc ttc atc gcc gac gac cag						1038
Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ala Asp Asp Gln						155
140		145		150		
gcc cag ggc atg acc tgc ggc ggc tcc ggc aac tgc tcc tcc ggt ggt						1086
Ala Gln Gly Met Thr Ser Gly Gly Ser Gly Asn Cys Ser Ser Gly Gly						170
		160		165		
acc acg tac tac cag gag gtc ggc ccg gcg ctg agc acc tgg aac ctc						1134
Thr Thr Tyr Tyr Gln Glu Val Gly Pro Ala Leu Ser Thr Trp Asn Leu						185
		175		180		
agc ctc gtc acc agc tag						1152
Ser Leu Val Thr Ser						
		190				

<210> 4

<211> 383

<212> PRT

5 <213> Nocardioopsis prasina DSM 15649

<400> 4

Met Arg Pro Ser Pro Val Ile Ser Ala Ile Gly Thr Gly Ala Leu
 -190 -185 -180

Ala Phe Gly Leu Ala Leu Ser Val Ala Pro Gly Ala Ser Ala Val
 -175 -170 -165

Thr Ala Pro Thr Glu Pro Ala Pro Gln Gly Glu Ala Ala Thr Met
 -160 -155 -150

Gln Glu Ala Leu Glu Arg Asp Phe Gly Leu Thr Pro Phe Glu Ala
 -145 -140 -135

Glu Asp Leu Leu Glu Ala Gln Asn Asp Ala Leu Gly Ile Asp Thr
 -130 -125 -120

Ala Ala Ala Lys Ala Ala Gly Asp Ala Tyr Ala Gly Ser Val Phe
 -115 -110 -105

Asp Thr Asp Thr Leu Glu Leu Thr Val Leu Leu Thr Asp Ala Gly Ala
 -100 -95 -90

Val Ser Asp Val Glu Ala Thr Gly Ala Gly Thr Glu Leu Val Ser Tyr
 -85 -80 -75 -70

Gly Thr Glu Gly Leu Ala Glu Ile Met Asp Glu Leu Asp Ala Ala Gly
 -65 -60 -55

Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Leu Ala Gly Asp Thr
 -50 -45 -40

Val Val Ile Glu Ala Thr Asp Thr Ser Glu Ala Gln Ser Phe Val Glu
 -35 -30 -25

Ala Ala Gly Val Asp Ser Ser Ala Val Gln Val Glu Gln Thr Asp Glu
 -20 -15 -10

Ala Pro Gln Leu Tyr Ala Asp Ile Val Gly Gly Asp Ala Tyr Tyr Met
 -5 -1 1 5 10

Gly Gly Gly Arg Cys Ser Val Gly Phe Ala Val Thr Asp Ser Ser Gly
 15 20 25

Asn Asp Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Ser
 30 35 40

Ala Asp Ser Glu Asp Gly Ser Gly Ser Gly Val Phe Glu Glu Ser Ile
 45 50 55

Phe Pro Gly Asn Asp Ala Ala Phe Val Ser Ser Thr Ser Asn Trp Thr
 60 65 70 75

Val Thr Asn Leu Val Asn Met Tyr Ser Ser Gly Gly Thr Gln Ser Val
 80 85 90

Gly Gly Ser Ser Gln Ala Pro Val Gly Ala Ala Val Cys Arg Ser Gly
 95 100 105

Ser Thr Thr Gly Trp His Cys Gly Ser Ile Glu Ala Arg Gly Gln Ser
 110 115 120

Val Ser Tyr Pro Glu Gly Thr Val Thr Asp Met Thr Arg Thr Asp Val
 125 130 135

Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ala Asp Asp Gln
 140 145 150 155

Ala Gln Gly Met Thr Ser Gly Gly Ser Gly Asn Cys Ser Ser Gly Gly
 160 165 170

Thr Thr Tyr Tyr Gln Glu Val Gly Pro Ala Leu Ser Thr Trp Asn Leu
 175 180 185

Ser Leu Val Thr Ser
 190

<211> 1152

<212> ADN

<213> Nocardiosis prasina DSM 14010

<220>

5 <221> CDS

<222> (1)..(1149)

<220>

<221> six_peptides

<222> (1)..(87)

10 <220>

<221> mat_peptide

<222> (574)..(1149)

<400> 5

atg Met	cga Arg	ccc Pro	tcc Ser	ccc Pro	gtc Val	atc Ile	tcc Ser	gcg Ala	atc Ile	ggc Gly	acg Thr	gga Gly	gcg Ala	ctg Leu	45	
	-190					-185					-180					
gcc Ala	ttc Phe	ggg Gly	ctc Leu	gcg Ala	ctc Leu	tcg Ser	gtc Val	gct Ala	ccc Pro	ggc Gly	gcc Ala	tcc Ser	gcc Ala	gtg Val	90	
	-175					-170					-165					
acc Thr	gcc Ala	ccc Pro	gcc Ala	gag Glu	ccc Pro	tcg Ser	ccc Pro	cag Gln	ggc Gly	gag Glu	gcg Ala	acc Thr	acc Thr	atg Met	135	
	-160					-155					-150					
cag Gln	gaa Glu	gcg Ala	ctt Leu	gag Glu	agg Arg	gac Asp	ttc Phe	ggc Gly	ctc Leu	acc Thr	ccg Pro	ttc Phe	gag Glu	gcc Ala	180	
	-145					-140					-135					
gac Asp	gac Asp	ctg Leu	ctc Leu	gaa Glu	gcc Ala	cag Gln	aag Lys	gag Glu	gcc Ala	ctc Leu	ggg Gly	atc Ile	gac Asp	acg Thr	225	
	-130					-125					-120					
gcg Ala	gcg Ala	gcc Ala	gag Glu	gcc Ala	gcc Ala	ggc Gly	gac Asp	gcc Ala	tac Tyr	gcg Ala	ggc Gly	tcc Ser	gtg Val	ttc Phe	270	
	-115					-110					-105					
gac Asp	acc Thr	gac Asp	acc Thr	ctg Leu	gaa Glu	ctg Leu	acc Thr	gtc Val	ctg Leu	ctc Leu	acg Thr	gac Asp	ggc Gly	ggc Gly	ccg Pro	318
	-100					-95					-90					
gcg Ala	tcg Ser	gac Asp	gtc Val	gag Glu	gcc Ala	gcc Ala	ggc Gly	gcc Ala	gag Glu	acc Thr	tcg Ser	gtg Val	gtc Val	tcc Ser	cac His	366
	-85				-80				-75						-70	
ggc Gly	acc Thr	gac Asp	ggc Gly	ctg Leu	gcg Ala	gcg Ala	atc Ile	atg Met	gac Asp	gag Glu	ctc Leu	gac Asp	gcg Ala	gtc Val	ggc Gly	414
				-65					-60					-55		
gcc Ala	cag Gln	ccg Pro	ggt Gly	gtc Val	gtc Val	ggc Gly	tgg Trp	tac Tyr	ccc Pro	gac Asp	ctc Leu	gcc Ala	agc Ser	gac Asp	acg Thr	462
			-50					-45					-40			
gtg Val	gtc Val	gtc Val	gag Glu	gcc Ala	acc Thr	gac Asp	gcg Ala	tcc Ser	gac Asp	gcc Ala	cag Gln	ggc Gly	ttc Phe	atc Ile	gag Glu	510
		-35					-30					-25				
gcc Ala	gcc Ala	ggc Gly	gtg Val	gac Asp	tcc Ser	tcc Ser	gcc Ala	gtc Val	cag Gln	gtg Val	gag Glu	gag Glu	acc Thr	gac Asp	gag Glu	558
	-20					-15					-10					
tcg Ser	ccc Pro	gag Glu	ctg Leu	tac Tyr	gcc Ala	gac Asp	atc Ile	gtc Val	ggc Gly	ggc Gly	gac Asp	gcc Ala	tac Tyr	tac Tyr	atg Met	606
	-5			-1	1				5				10			
ggc Gly	ggc Gly	gga Gly	cgc Arg	tgc Cys	tcg Ser	gtg Val	ggc Gly	ttc Phe	gcg Ala	gcc Ala	acc Thr	gac Asp	agc Ser	gcg Ala	ggc Gly	654
			15					20					25			
aac Asn	gac Asp	gga Gly	ttc Phe	gtg Val	acg Thr	gcc Ala	ggc Gly	cac His	tgc Cys	ggc Gly	acc Thr	gtc Val	ggc Gly	acc Thr	tcc Ser	702
		30					35					40				

gcc Ala	gac Asp 45	agc Ser	gag Glu	gac Asp	ggc Gly	agc Ser 50	ggc Gly	tcc Ser	ggt Gly	gtg Val	ttc Phe 55	gag Glu	gag Glu	tcg Ser	atc Ile	750
ttc Phe 60	ccg Pro	ggc Gly	aac Asn	gac Asp	gcc Ala 65	gcc Ala	ttc Phe	gtc Val	cgg Arg	tcc Ser 70	acg Thr	tcc Ser	aac Asn	tgg Trp	acc Thr 75	798
gtc Val	acc Thr	aac Asn	ctg Leu	gtc Val 80	aac Asn	atg Met	tac Tyr	agc Ser	tcc Ser 85	ggc Gly	ggc Gly	acc Thr	cag Gln	tcc Ser 90	gtc Val	846
ggc Gly	ggc Gly	tcc Ser	acc Thr 95	cag Gln	gcc Ala	ccg Pro	gtc Val	ggc Gly 100	gcg Ala	gcc Ala	gtg Val	tgc Cys	cgc Arg 105	tcc Ser	ggt Gly	894
tcc Ser	acc Thr	acg Thr 110	ggc Gly	tgg Trp	cac His	tgc Cys	ggc Gly 115	acc Thr	atc Ile	gag Glu	gcc Ala	cga Arg 120	ggc Gly	cag Gln	tcg Ser	942
gtg Val	agc Ser 125	tac Tyr	ccg Pro	gag Glu	ggc Gly	acc Thr 130	gtc Val	aac Asn	gac Asp	atg Met	acc Thr 135	cgg Arg	acc Thr	aac Asn	gtg Val	990
tgc Cys 140	gcc Ala	gag Glu	ccc Pro	ggc Gly	gac Asp 145	tcc Ser	ggc Gly	ggt Gly	tcg Ser	ttc Phe 150	atc Ile	tcc Ser	gac Asp	gac Asp	cag Gln 155	1038
gcc Ala	cag Gln	ggc Gly	atg Met	acc Thr 160	tcg Ser	ggc Gly	ggc Gly	tcc Ser	ggc Gly 165	aac Asn	tgc Cys	acc Thr	tcc Ser	ggt Gly 170	ggt Gly	1086
acg Thr	acg Thr	tac Tyr	tac Tyr 175	cag Gln	gag Glu	gtc Val	ggc Gly	ccg Pro 180	gcg Ala	ctg Leu	agc Ser	acc Thr	tgg Trp 185	aac Asn	ctc Leu	1134
agc Ser	ctc Leu	gtc Val 190	acg Thr	agc Ser	tag											1152

<210> 6

<211> 383

<212> PRT

5 <213> Nocardiosis prasina DSM 14010

<400> 6

Met Arg Pro Ser Pro Val Ile Ser Ala Ile Gly Thr Gly Ala Leu
 -190 -185 -180

Ala Phe Gly Leu Ala Leu Ser Val Ala Pro Gly Ala Ser Ala Val
 -175 -170 -165

Thr Ala Pro Ala Glu Pro Ser Pro Gln Gly Glu Ala Thr Thr Met
 -160 -155 -150

Gln Glu Ala Leu Glu Arg Asp Phe Gly Leu Thr Pro Phe Glu Ala
 -145 -140 -135

Asp Asp Leu Leu Glu Ala Gln Lys Glu Ala Leu Gly Ile Asp Thr
 -130 -125 -120

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Ala Gly Ser Val Phe
 -115 -110 -105

Asp Thr Asp Thr Leu Glu Leu Thr Val Leu Leu Thr Asp Gly Gly Pro
 -100 -95 -90

Ala Ser Asp Val Glu Ala Ala Gly Ala Glu Thr Ser Val Val Ser His
 -85 -80 -75 -70

Gly Thr Asp Gly Leu Ala Ala Ile Met Asp Glu Leu Asp Ala Val Gly
 -65 -60 -55

Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Leu Ala Ser Asp Thr
 -50 -45 -40

Val Val Val Glu Ala Thr Asp Ala Ser Asp Ala Gln Gly Phe Ile Glu
 -35 -30 -25

Ala Ala Gly Val Asp Ser Ser Ala Val Gln Val Glu Glu Thr Asp Glu
 -20 -15 -10

Ser Pro Glu Leu Tyr Ala Asp Ile Val Gly Gly Asp Ala Tyr Tyr Met
 -5 -1 1 5 10

Gly Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asp Ser Ala Gly
 15 20 25

Asn Asp Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Ser
 30 35 40

Ala Asp Ser Glu Asp Gly Ser Gly Ser Gly Val Phe Glu Glu Ser Ile
 45 50 55

Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Ser Thr Ser Asn Trp Thr
 60 65 70 75

Val Thr Asn Leu Val Asn Met Tyr Ser Ser Gly Gly Thr Gln Ser Val
 80 85 90

Gly Gly Ser Thr Gln Ala Pro Val Gly Ala Ala Val Cys Arg Ser Gly
 95 100 105

Ser Thr Thr Gly Trp His Cys Gly Thr Ile Glu Ala Arg Gly Gln Ser
 110 115 120

Val Ser Tyr Pro Glu Gly Thr Val Asn Asp Met Thr Arg Thr Asn Val
 125 130 135

Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Asp Asp Gln
 140 145 150 155

Ala Gln Gly Met Thr ser Gly Gly ser Gly Asn Cys Thr Ser Gly Gly
160 165 170

Thr Thr Tyr Tyr Gln Glu Val Gly Pro Ala Leu Ser Thr Trp Asn Leu
175 180 185

ser Leu Val Thr ser
190

<210> 7

<211> 1155

<212> ADN

5 <213> Nocardiosis sp. DSM 16424

<220>

<221> CDS

<222> (1)..(1152)

<220>

10 <221> sig_peptide

<222> (1)..(87)

<220>

<221> mat_peptide

<222> (586)..(1152)

15 <400> 7

atg Met -195	aga Arg	ccc Pro	tcc Ser	acc Thr	atc Ile -190	gcc Ala	tcc Ser	gcc Ala	gtc Val	ggc Gly -185	aca Thr	gga Gly	gca Ala	ctg Leu	45	
gcc Ala -180	ttc Phe	ggt Gly	ctg Leu	gca Ala	ctg Leu -175	tcc Ser	atg Met	gcc Ala	ccc Pro	gga Gly -170	gcc Ala	ctc Leu	gcg Ala	gcg Ala	90	
ccc Pro -165	ggc Gly	ccc Pro	gtc Val	ccc Pro	cag Gln -160	acc Thr	ccc Pro	gtc Val	gcc Ala	gac Asp -155	gac Asp	agc Ser	gcc Ala	gcc Ala	135	
agc Ser -150	atg Met	acc Thr	gaa Glu	gcg Ala	ctc Leu -145	aag Lys	cgt Arg	gac Asp	ctc Leu	aac Asn -140	ctc Leu	tcc Ser	tcg Ser	gcc Ala	180	
gag Glu -135	gcc Ala	gag Glu	gag Glu	ctg Leu	ctc Leu -130	tcg Ser	gcg Ala	cag Gln	gaa Glu	gcc Ala -125	gcg Ala	atc Ile	gag Glu	acc Thr	225	
gac Asp -120	gcc Ala	gag Glu	gcc Ala	gcc Ala	gag Glu -115	gcc Ala	gcg Ala	gga Gly	gag Glu	gcc Ala -110	tac Tyr	ggc Gly	ggc Gly	tcc Ser	270	
ctg Leu -105	ttc Phe	gac Asp	acc Thr	gaa Glu	acc Thr -100	ctc Leu	gaa Glu	ctc Leu	acc Thr	gtg Val -95	ctg Leu	gtg Val	acc Thr	gac Asp	acc Thr -90	318
acg Thr	gcc Ala	gtc Val	gac Asp	gcg Ala	gtc Val -85	gag Glu	gcc Ala	acc Thr	gga Gly -80	gcc Ala	gag Glu	gcc Ala	acc Thr	gtg Val -75	gtc Val	366
acc Thr	cac His	ggc Gly -70	acc Thr	gac Asp	ggc Gly	ctg Leu	gcc Ala	gag Glu	gtc Val	gtg Val	gag Glu	gac Asp	ctc Leu	aac Asn	agc Ser	414

gcc Ala	gac Asp	gcc Ala -55	ccg Pro	gcg Ala	ggc Gly	gtc Val	ctc Leu -50	ggc Gly	tgg Trp	tac Tyr	ccc Pro	gac Asp -45	atg Met	gag Glu	agc Ser	462
gac Asp	acc Thr -40	gtg Val	gtg Val	gtc Val	gag Glu	gtg Val -35	ctg Leu	gag Glu	ggc Gly	tcc Ser	gac Asp -30	gcc Ala	gac Asp	gtc Val	gcc Ala	510
gcc Ala -25	ctg Leu	ctc Leu	gcc Ala	gac Asp	gcc Ala -20	ggc Gly	gtg Val	gac Asp	gcc Ala	tcc Ser -15	gcc Ala	gtc Val	cgg Arg	gtg Val	gag Glu -10	558
gag Glu	gcg Ala	gag Glu	gag Glu	gtc Val -5	ccg Pro	cag Gln	gtc Val	tac Tyr -1	gcc Ala 1	aac Asn	atc Ile	atc Ile	ggc Gly 5	ggc Gly	ctg Leu	606
gcc Ala	tac Tyr	acc Thr 10	atg Met	ggc Gly	gga Gly	cgc Arg	tgc Cys 15	tcc Ser	gtc Val	ggc Gly	ttc Phe	gcg Ala 20	gcg Ala	acc Thr	aac Asn	654
agc Ser	gcc Ala 25	gga Gly	cag Gln	ccc Pro	ggt Gly	ttc Phe 30	gtg Val	acg Thr	gcg Ala	ggc Gly	cac His 35	tgc Cys	ggc Gly	acc Thr	gtc Val	702
ggc Gly 40	acc Thr	gcc Ala	gtg Val	acc Thr	atc Ile 45	ggc Gly	gac Asp	ggc Gly	cgc Arg	ggc Gly 50	gtc Val	ttc Phe	gag Glu	cgc Arg	tcg Ser 55	750
gtc Val	ttc Phe	ccc Pro	ggc Gly	aac Asn 60	gac Asp	gcc Ala	gcc Ala	ttc Phe 65	gtc Val	cgc Arg	ggc Gly	acc Thr	tcc Ser	aac Asn 70	ttc Phe	798
acc Thr	ctg Leu	acc Thr	aac Asn 75	ctg Leu	gtc Val	tcc Ser	cgc Arg	tac Tyr 80	aac Asn	tcc Ser	ggc Gly	ggc Gly	cac His 85	cag Gln	gcg Ala	846
gtg Val	acc Thr	ggc Gly 90	acc Thr	agc Ser	cag Gln	gcc Ala	ccg Pro 95	gcc Ala	ggc Gly	tcg Ser	gcc Ala	gtc Val 100	tgc Cys	cgc Arg	tcc Ser	894
ggc Gly	tcc Ser 105	acc Thr	acc Thr	ggc Gly	tgg Trp	cac His 110	tgc Cys	ggc Gly	acc Thr	atc Ile	cag Gln 115	gcc Ala	cgc Arg	aac Asn	cag Gln	942
acc Thr 120	gtg Val	cgc Arg	tac Tyr	ccg Pro	cag Gln 125	ggc Gly	acc Thr	gtc Val	aac Asn	gcg Ala 130	ctc Leu	acc Thr	cgc Arg	acc Thr	aac Asn 135	990
gtg Val	tgc Cys	gcc Ala	gag Glu	ccc Pro 140	ggt Gly	gac Asp	tcc Ser	ggc Gly	ggc Gly 145	tcg Ser	ttc Phe	atc Ile	tcc Ser	ggc Gly 150	tcg Ser	1038
cag Gln	gcc Ala	cag Gln	ggc Gly 155	gtc Val	acc Thr	tcc Ser	ggc Gly 160	ggc Gly	tcc Ser	ggc Gly	aac Asn	tgc Cys	tcc Ser 165	ttc Phe	ggc Gly	1086
ggc Gly	acg Thr	acc Thr 170	tac Tyr	tac Tyr	cag Gln	gag Glu	gtc Val 175	gcc Ala	ccg Pro	atg Met	atc Ile	aac Asn 180	tcc Ser	tgg Trp	ggc Gly	1134
gtt Val	cgc Arg 185	atc Ile	cgc Arg	acc Thr	agc Ser	tga										1155

<211> 384

<212> PRT

<213> Nocardiosis sp. DSM 16424

<400> 8

Met Arg Pro Ser Thr Ile Ala Ser Ala Val Gly Thr Gly Ala Leu
 -195 -190 -185

Ala Phe Gly Leu Ala Leu Ser Met Ala Pro Gly Ala Leu Ala Ala
 -180 -175 -170

Pro Gly Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala Ala
 -165 -160 -155

Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asn Leu Ser Ser Ala
 -150 -145 -140

Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu Thr
 -135 -130 -125

Asp Ala Glu Ala Ala Glu Ala Ala Gly Glu Ala Tyr Gly Gly Ser
 -120 -115 -110

Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Thr
 -105 -100 -95 -90

Thr Ala Val Asp Ala Val Glu Ala Thr Gly Ala Glu Ala Thr Val Val
 -85 -80 -75

Thr His Gly Thr Asp Gly Leu Ala Glu Val Val Glu Asp Leu Asn Ser
 -70 -65 -60

Ala Asp Ala Pro Ala Gly Val Leu Gly Trp Tyr Pro Asp Met Glu Ser
 -55 -50 -45

Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val Ala
 -40 -35 -30

Ala Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Arg Val Glu
 -25 -20 -15 -10

Glu Ala Glu Glu Val Pro Gln Val Tyr Ala Asn Ile Ile Gly Gly Leu
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
 10 15 20

Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val
 25 30 35

Gly Thr Ala Val Thr Ile Gly Asp Gly Arg Gly Val Phe Glu Arg Ser
 40 45 50 55

Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
60 65 70
Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly His Gln Ala
75 80 85
Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg Ser
90 95 100
Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln
105 110 115
Thr Val Arg Tyr Pro Gln Gly Thr Val Asn Ala Leu Thr Arg Thr Asn
120 125 130 135
Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly Ser
140 145 150
Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Phe Gly
155 160 165
Gly Thr Thr Tyr Tyr Gln Glu Val Ala Pro Met Ile Asn Ser Trp Gly
170 175 180
Val Arg Ile Arg Thr Ser
185

<210> 9

<211> 1152

<212> ADN

5 <213> Nocardiosis alkaliphila DSM 44657

<220>

<221> CDS

<222> (1)..(1149)

<220>

10 <221> sig_peptide

<222> (1)..(87)

<220>

<221> mat_peptide

<222> (586)..(1149)

15 <400> 9

ES 2 380 105 T3

atg	cga	ccc	tcc	ccc	gtt	gtc	tcc	gcc	ata	ggg	aca	gga	gcc	ttg	45
Met	Arg	Pro	Ser	Pro	Val	Val	Ser	Ala	Ile	Gly	Thr	Gly	Ala	Leu	
-195					-190					-185					
gcc	ttc	ggc	ctg	gct	ctg	ggc	act	tcc	ccc	gcg	gcc	atc	gcc	gcc	90
Ala	Phe	Gly	Leu	Ala	Leu	Gly	Thr	Ser	Pro	Ala	Ala	Ile	Ala	Ala	
-180					-175					-170					
ccc	gcc	ccc	cag	tcc	ccc	gac	acc	gaa	acg	cag	gcc	gag	gcc	gtc	135
Pro	Ala	Pro	Gln	Ser	Pro	Asp	Thr	Glu	Thr	Gln	Ala	Glu	Ala	Val	
-165					-160					-155					
acc	atg	gcc	gaa	gcc	ctc	caa	cgc	gat	ctc	ggg	ctg	tcc	tcc	tcc	180

Thr Met Ala Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser
 -150 -145 -140
 gag gcc acc gaa ctc ctc gcc gca cag gcc gag gcg ttc gag gtc 225
 Glu Ala Thr Glu Leu Leu Ala Ala Gln Ala Glu Ala Phe Glu Val
 -135 -130 -125
 gac gag gcc gcc acc gag gcc gcc gcc gac gcc tac ggc ggc tcc 270
 Asp Glu Ala Ala Thr Glu Ala Ala Ala Asp Ala Tyr Gly Gly Ser
 -120 -115 -110
 ctc ttc gac acc gac agc ctc gaa ctg acc gtg ctg gtc acc gac agc 318
 Leu Phe Asp Thr Asp Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ser
 -105 -100 -95 -90
 gcc gcc gtc gac gcg gtc gag gcc acc ggc gcc aag gcc gag gtc gtc 366
 Ala Ala Val Asp Ala Val Glu Ala Thr Gly Ala Lys Ala Glu Val Val
 -85 -80 -75
 gac cac ggt atc gag ggc ctc gag gag atc gtc gac gaa ctc aac gag 414
 Asp His Gly Ile Glu Gly Leu Glu Glu Ile Val Asp Glu Leu Asn Glu
 -70 -65 -60
 tcc aac gcc aag tcg ggc gtc gtc ggt tgg tac ccc gac gtg gcc ggt 462
 Ser Asn Ala Lys Ser Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly
 -55 -50 -45
 gac acg gtc gtc ctg gag gtc atg gaa ggc tcc gag gcc gac gtg gac 510
 Asp Thr Val Val Leu Glu Val Met Glu Gly Ser Glu Ala Asp Val Asp
 -40 -35 -30
 gcc ctg ctc gcc gag acc ggg gtc gac gcc gcc gac gtc acg gtg gag 558
 Ala Leu Leu Ala Glu Thr Gly Val Asp Ala Ala Asp Val Thr Val Glu
 -25 -20 -15 -10
 acc acc acc gag cag ccc gag ctc tac gcc gac atc atc ggt ggc ctg 606
 Thr Thr Thr Glu Gln Pro Glu Leu Tyr Tyr Ala Asp Ile Ile Gly Gly Leu
 -5 -1 1 5
 gcc tac acc atg ggc gga cgt tgc tcg gtc ggc ttc gcc gcc acc aac 654
 Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
 10 15 20
 tcc tcc ggc cag ccc gga ttc gtc acc gcc ggc cac tgc ggc agt gtc 702
 Ser Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Ser Val
 25 30 35
 ggc acc ggc gtc acc atc ggt aac ggc cgg ggc gtc ttc gag cgt tcc 750
 Gly Thr Gly Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Arg Ser
 40 45 50 55
 atc ttc ccg ggc aac gac gcc gcc ttc gtc cgt ggc acg tcc aac ttc 798
 Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
 60 65 70
 acc ctg acc aac ctg gtc agc cgc tac aac tcc ggc ggc tac gcc acg 846
 Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr
 75 80 85
 gtg tcc ggg tcc tcc gcg gcc ccg atc ggc tcc cag gtg tgc cgc tcc 894
 Val Ser Gly Ser Ser Ala Ala Pro Ile Gly Ser Gln Val Cys Arg Ser
 90 95 100
 ggc tcc acc acc ggc tgg cac tgc ggc acc atc cag gcc cgc aac cag 942
 Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln
 105 110 115
 acg gtg cgc tac ccg cag ggc acc gtc cag gcc ctg acc cgc acc agc 990

ES 2 380 105 T3

Thr 120	Val	Arg	Tyr	Pro	Gln 125	Gly	Thr	Val	Gln	Ala 130	Leu	Thr	Arg	Thr	Ser 135	
gtg Val	tgc Cys	gcc Ala	gag Glu	ccc Pro 140	ggt Gly	gac Asp	tcc Ser	ggt Gly	ggt Gly 145	tcc Ser	ttc Phe	atc Ile	tcc Ser	ggc Gly 150	agc Ser	1038
cag Gln	gcc Ala	cag Gln	ggc Gly 155	gtc Val	acc Thr	tcc Ser	ggt Gly	ggc Gly 160	tcg Ser	ggc Gly	aac Asn	tgc Cys	cgc Arg 165	acc Thr	ggt Gly	1086
ggc Gly	acg Thr	acc Thr 170	tac Tyr	tac Tyr	cag Gln	gag Glu	gtc Val 175	aac Asn	ccc Pro	atg Met	ctc Leu	aac Asn 180	agc Ser	tgg Trp	ggc Gly	1134
ctg Leu	cgt Arg 185	ctg Leu	cgc Arg	acc Thr	tga											1152

<210> 10

<211> 383

<212> PRT

5 <213> Nocardiosis alkaliphila DSM 44657

<400> 10

Met Arg Pro Ser Pro Val Val Ser Ala Ile Gly Thr Gly Ala Leu
 -195 -190 -185

Ala Phe Gly Leu Ala Leu Gly Thr Ser Pro Ala Ala Ile Ala Ala
 -180 -175 -170

Pro Ala Pro Gln Ser Pro Asp Thr Glu Thr Gln Ala Glu Ala Val
 -165 -160 -155

Thr Met Ala Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser
 -150 -145 -140

Glu Ala Thr Glu Leu Leu Ala Ala Gln Ala Glu Ala Phe Glu Val
 -135 -130 -125

Asp Glu Ala Ala Thr Glu Ala Ala Ala Asp Ala Tyr Gly Gly Ser
 -120 -115 -110

Leu Phe Asp Thr Asp Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ser
 -105 -100 -95 -90

Ala Ala Val Asp Ala Val Glu Ala Thr Gly Ala Lys Ala Glu Val Val
 -85 -80 -75

Asp His Gly Ile Glu Gly Leu Glu Glu Ile Val Asp Glu Leu Asn Glu
 -70 -65 -60

Ser Asn Ala Lys Ser Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly
 -55 -50 -45

Asp Thr Val Val Leu Glu Val Met Glu Gly Ser Glu Ala Asp Val Asp

ES 2 380 105 T3

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

<210> 11
 <211> 1155
 <212> ADN

5 <213> Nocardiosis lucentensis DSM 44048

<220>

<221> CDS

<222> (1)..(1152)

<220>

5 <221> sig_peptide

<220>

<221> mat peptide

<222> (586)..(1152)

<400> 11

atg Met -195	cga Arg	ccc Pro	tcc Ser	ccc Pro	gtt Val -190	atc Ile	tcc Ser	gcc Ala	cta Leu	gga Gly -185	acc Thr	ggc Gly	gcc Ala	ctc Leu	45	
gcc Ala -180	ttc Phe	gga Gly	ctg Leu	gtc Val	atc Ile -175	acc Thr	atg Met	gcc Ala	ccg Pro	ggc Gly -170	gtg Val	aac Asn	gcc Ala	gga Gly	90	
acc Thr -165	gta Val	ccc Pro	acc Thr	ccc Pro	cag Gln -160	gcc Ala	ccc Pro	gtc Val	ccc Pro	gac Asp -155	gac Asp	gag Glu	gcc Ala	acc Thr	135	
acc Thr -150	atg Met	ctc Leu	gaa Glu	gcc Ala	atg Met -145	gag Glu	agg Arg	gat Asp	ctc Leu	gac Asp -140	ctc Leu	acc Thr	ccg Pro	ttc Phe	180	
gag Glu -135	gcc Ala	gag Glu	gaa Glu	ctc Leu	ttc Phe -130	gag Glu	gca Ala	cag Gln	gaa Glu	gag Glu -125	gcc Ala	atc Ile	gac Asp	ctc Leu	225	
gac Asp -120	gag Glu	gag Glu	gcc Ala	acc Thr	gaa Glu -115	gcg Ala	gcc Ala	ggt Gly	gcg Ala	gcc Ala -110	tac Tyr	ggc Gly	ggt Gly	tcg Ser	270	
ctc Leu -105	ttc Phe	gac Asp	acc Thr	gaa Glu	acc Thr -100	cac His	gaa Glu	ctc Leu	acc Thr	gtc Val -95	ctg Leu	gtg Val	acc Thr	gac Asp	gtc Val -90	318
gac Asp	gcg Ala	gtc Val	gag Glu	gcc Ala -85	gtg Val	gag Glu	gcc Ala	acc Thr	gga Gly -80	gcc Ala	gcc Ala	gcc Ala	gag Glu	gtc Val -75	gtc Val	366
tcc Ser	cac His	ggc Gly	tcc Ser	gac Asp	ggt Gly -70	ctg Leu	gcc Ala	gac Asp	atc Ile	gtc Val	gag Glu	gac Asp	ctc Leu	aac Asn	gcc Ala	414
acc Thr	gac Asp	gcc Ala -55	ggc Gly	agc Ser	gag Glu	gtc Val	gtg Val	ggc Gly	tgg Trp	tac Tyr	ccc Pro	gac Asp -45	gtc Val	acc Thr	agc Ser	462
gac Asp -40	agc Ser	gtg Val	gtc Val	gtc Val	gag Glu	gtg Val	gtc Val	gag Glu	ggc Gly	tcc Ser	gac Asp -30	gtc Val	gac Asp	gtc Val	gac Asp	510
tcc Ser -25	atc Ile	gtc Val	gag Glu	ggc Gly	acg Thr -20	ggc Gly	gtc Val	gac Asp	ccg Pro	gcg Ala	gtc Val	atc Ile	gag Glu	gtc Val	cag Gln -10	558
gag Glu	gtc Val	tcc Ser	gaa Glu	cag Gln -5	cct Pro	cag Gln	acc Thr	tac Tyr -1	gcc Ala 1	aac Asn	atc Ile	atc Ile	ggc Gly 5	ggc Gly	ctg Leu	606
gcc Ala	tac Tyr 10	tac Tyr	atg Met	agc Ser	tcg Ser	ggc Gly 15	ggc Gly	cgc Arg	tgc Cys	tcg Ser	gtc Val	ggc Gly 20	ttc Phe	ccc Pro	gcc Ala	654
acc Thr	aac Asn	agc Ser	tcc Ser	ggc Gly	cag Gln	ccg Pro	ggc Gly	ttc Phe	gtc Val	acg Thr	gcg Ala	ggc Gly	cac His	tgc Cys	ggc Gly	702

25		30		35															
acc	gtc	ggc	acc	ggc	gtc	acc	atc	ggc	aac	ggc	cgc	ggc	acc	ttc	gag				750
Thr	Val	Gly	Thr	Gly	Val	Thr	Ile	Gly	Asn	Gly	Arg	Gly	Thr	Phe	Glu				
40				45				50						55					
cgc	tcc	gtg	ttc	ccc	ggc	aac	gac	gcc	gcc	ttc	gtc	cga	ggc	acg	tcc				798
Arg	Ser	Val	Phe	Pro	Gly	Asn	Asp	Ala	Ala	Phe	Val	Arg	Gly	Thr	Ser				
				60				65						70					
aac	ttc	acg	ctg	tac	aac	ctc	gtc	tac	cgc	tac	agc	ggc	tac	cag	acc				846
Asn	Phe	Thr	Leu	Tyr	Asn	Leu	Val	Tyr	Arg	Tyr	Ser	Gly	Tyr	Gln	Thr				
			75					80					85						
gtg	acg	ggc	agc	aac	gcc	gcc	ccg	atc	ggc	tcg	tcc	atc	tgc	cgt	tcc				894
Val	Thr	Gly	Ser	Asn	Ala	Ala	Pro	Ile	Gly	Ser	Ser	Ile	Cys	Arg	Ser				
		90					95					100							
ggt	tcc	acc	acc	ggc	tgg	cac	tgc	ggc	acc	atc	cag	gcc	cgc	aac	cag				942
Gly	Ser	Thr	Thr	Gly	Trp	His	Cys	Gly	Thr	Ile	Gln	Ala	Arg	Asn	Gln				
	105					110					115								
acc	gtc	cgg	tac	ccg	cag	ggc	acc	gtc	tac	tac	ctg	acc	cgt	acc	aac				990
Thr	Val	Arg	Tyr	Pro	Gln	Gly	Thr	Val	Tyr	Tyr	Leu	Thr	Arg	Thr	Asn				
					125					130					135				
gtg	tgc	gcc	gag	ccc	ggc	gac	tcc	gga	ggc	tcc	ttc	atc	tcc	gga	acg				1038
Val	Cys	Ala	Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Phe	Ile	Ser	Gly	Thr				
				140					145					150					
cag	gcc	cag	ggc	atg	acc	tcc	ggc	ggc	tcc	ggc	aac	tgc	agc	agc	ggt				1086
Gln	Ala	Gln	Gly	Met	Thr	Ser	Gly	Gly	Ser	Gly	Asn	Cys	Ser	Ser	Gly				
			155					160					165						
ggc	acc	acc	ttc	tac	cag	gag	gtg	gac	ccg	gtg	gag	agc	gcc	tgg	ggc				1134
Gly	Thr	Thr	Phe	Tyr	Gln	Glu	Val	Asp	Pro	Val	Glu	Ser	Ala	Trp	Gly				
		170					175					180							
gtg	cga	ctg	cgc	acc	agc	tag													1155
Val	Arg	Leu	Arg	Thr	Ser														
	185																		

<210> 12

<211> 384

<212> PRT

5 <213> Nocardiosis lucentensis DSM 44048

<400> 12

Met Arg Pro Ser Pro Val Ile Ser Ala Leu Gly Thr Gly Ala Leu
 -195 -190 -185

Ala Phe Gly Leu Val Ile Thr Met Ala Pro Gly Val Asn Ala Gly
 -180 -175 -170

Thr Val Pro Thr Pro Gln Ala Pro Val Pro Asp Asp Glu Ala Thr
 -165 -160 -155

Thr Met Leu Glu Ala Met Glu Arg Asp Leu Asp Leu Thr Pro Phe
 -150 -145 -140

Glu Ala Glu Glu Leu Phe Glu Ala Gln Glu Glu Ala Ile Asp Leu
 -135 -130 -125

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

Gln Ala Gln Gly Met Thr ser Gly Gly Ser Gly Asn Cys Ser ser Gly
 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Asp Pro Val Glu Ser Ala Trp Gly
 170 175 180

Val Arg Leu Arg Thr Ser
 185

<210> 13

<211> 81

<212> ADN

5 <213> Bacillus clausii

<220>

<221> sig_peptide

<222> (1)..(81)

<400> 13

atgaagaaac cgttggggaa aattgtcgca agcaccgcac tactcatttc tgttgctttt 60

10 agttcatcga tcgcatcggc t 81

<210> 14

<211> 49

<212> ADN

<213> sintético

15 <220>

<221> misc_feature

<223> Cebador

<400> 14

gcttttagtt catcgatcgc atcggctgcg accgtaccgg ccgagccag ???49

20 <210> 15

<211> 46

<212> ADN

<213> sintético

<220>

25 <221> misc_feature

<223> Cebador

<400> 15

ggagcggatt gaacatgcca ttactaacgg gtcaccaggg acagcc ??? 46

<210> 16

<211> 40
<212> ADN
<213> sintético
<220>
5 <221> misc-feature
<223> Cebador
<400> 16
gttcatcgat cgcacggct gtcaccgcac ccaccgagcc ???40
<210> 17
10 <211> 45
<212> ADN
<213> sintético
<220>
<221> misc_feature
15 <223> Cebador
<400> 17
ggagcggatt gaacatcga ttagctggtg acgaggctga ggttc ???45
<210> 18
<211> 38
20 <212> ADN
<213> sintético
<220>
<221> misc_feature
<223> Cebador
25 <400> 18
gttcatcgat cgcacggct gtgaccgccc ccgcccag ???38
<210> 19
<211> 45
<212> ADN
30 <213> sintético
<220>
<221> misc_feature
<223> Cebador
<400> 19
35 ggagcggatt gaacatcga ttagctcgtg acgaggctga ggttc ???45
<210> 20
<211> 41

<212> ADN
<213> sintético
<220> <221> misc_feature
<223> Cebador
5 <400> 20
gttcatcgat cgcacggct gcgcccggcc ccgtcccca g ???41
<210> 21
<211> 41
<212> ADN
10 <213> sintético
<220>
<221> misc_feature
<223> Cebador
<400> 21
15 ggagcggatt gaacatgca tcagctggtg cggatgcgaa c ???41
<210> 22
<211> 37
<212> ADN
<213> sintético
20 <220>
<221> mis_feature
<223> Cebador
<400> 22
gttcatcgat cgcacggct gccccggcc cccagtc ???37
25 <210> 23
<211> 44
<212> ADN
<213> sintético
<220>
30 <221> misc_feature
<223> Cebador
<400> 23
ggagcggatt gaacatgca ttaggtgagc agacgcaggc ccca ???44
<210> 24
35 <211> 42
<212> ADN
<213> sintético

<220>

<221> misc-feature

<223> Cebador

<400> 24

5 gttcatcgat cgcacggct ggaaccgtac ccacccccca gg ???42

<210> 25

<211> 41

<212> ADN

<213> sintético

10 <220>

<221> mis_feature

<223> Cebador

<400> 25

ggagcggatt gaacatgca ttagctggtg cgcagtcgca c ???41

15 <210> 26

<211> 9

<212> PRT

<213> Nocardiosis dassonvillei subesp. dassonvillei DSM 43235

<220>

20 <221> MISC_FEATURE

<223> N-terminal

<400> 26

Ala Asp Ile Val Gly Gly Glu Ala Tyr
1 5

<210> 27

25 <211> 10

<212> PRT

<213> Nocardiosis sp. DSM 16424

<220>

<221> MISC_FEATURE

30 <223> N_terminal

<400> 27

Ala Asn Ile Ile Gly Gly Leu Ala Tyr Thr
1 5 10

REIVINDICACIONES

1. Polipéptido aislado con actividad de proteasa, seleccionado del grupo que consiste en:
 - (a) un polipéptido con una secuencia de aminoácidos que tiene un grado de identidad para aminoácidos 1-192 de la SEC ID nº: 2 de al menos un 80%;
 - 5 (b) un fragmento de (a) que tiene actividad de proteasa.
2. Polipéptido según la reivindicación 1 que comprende aminoácidos 1-192 de la SEC ID nº: 2
3. Secuencia de ácidos nucleicos aislada comprendiendo una secuencia de ácidos nucleicos que codifica un polipéptido con actividad de proteasa, y que
 - 10 (a) codifica el polipéptido de cualquiera de las reivindicaciones 1 o 2;
 - (b) tiene un grado de identidad para nucleótidos 568-1143 de la SEC ID nº: 1 de al menos un 80%.
4. Secuencia de ácidos nucleicos según la reivindicación 3 que comprende nucleótidos 568-1143 de la SEC ID nº: 1.
5. Construcción de ácidos nucleicos comprendiendo la secuencia de ácidos nucleicos de cualquiera de las reivindicaciones 3-4 unida operativamente a una o más secuencias de control que dirigen la producción del polipéptido en un huésped de expresión adecuado.
 - 15 6. Vector de expresión recombinante comprendiendo la construcción de ácidos nucleicos según la reivindicación 5.
 7. Célula huésped recombinante comprendiendo el vector según la reivindicación 6.
 8. Planta transgénica, o parte de planta, capaz de expresar el polipéptido de cualquiera de las reivindicaciones 1-2.
 9. Animal transgénico no humano, o productos o elementos de este, siendo capaz de expresar el polipéptido de cualquiera de las reivindicaciones 1-2.
 - 20 10. Método para la producción de un polipéptido de cualquiera de las reivindicaciones 1-2, el método comprendiendo
 - (a) cultivo de una célula huésped recombinante según la reivindicación 7 para producir un sobrenadante comprendiendo el polipéptido; y
 - (b) recuperación del polipéptido.
 - 25 11. Método para la producción de un polipéptido de cualquiera de las reivindicaciones 1-2, el método comprendiendo
 - (a) cultivo de una cepa de *Nocardiosis dassonvillei* subesp. *dassonvillei* DSM 43235 y
 - (b) recuperación del polipéptido.
 12. Uso de al menos una proteasa de cualquiera de las reivindicaciones 1-2
 - 30 (i) en alimento para animales;
 - (ii) en aditivos de alimento;
 - (iii) en la preparación de una composición para uso en el alimento para animales;
 - (iv) para mejorar el valor nutritivo de un alimento para animales;
 - (v) para aumentar proteína soluble y/o digerible en el alimento para animales;
 - 35 (vi) para aumentar el grado de hidrólisis de proteínas en dietas para animales; y/o
 - (vii) para el tratamiento de proteínas.
 13. Método para mejorar el valor nutritivo de un alimento para animales, donde al menos una proteasa de cualquiera de las reivindicaciones 1-2 se añade al alimento.
 14. Aditivo de alimento comprendiendo
 - 40 (a) al menos una proteasa de cualquiera de las reivindicaciones 1-2 y

(b) al menos una vitamina liposoluble, y/o

(c) al menos una vitamina hidrosoluble, y/o

(d) al menos un oligoelemento.

5 15. Aditivo de alimento según la reivindicación 14, que comprende además amilasa; fitasa; xilanasas; galactanasas; alfa-galactosidasas; proteasa, fosfolipasa; y/o beta-glucanasa.

16. Alimento para animales con un contenido bruto de proteína de 50 a 800 g/kg y comprendiendo al menos una proteasa de cualquiera de las reivindicaciones 1-2.

17. Método para el tratamiento de proteínas, comprendiendo el paso de adición de al menos una proteasa de cualquiera de las reivindicaciones 1-2 a al menos una proteína o fuente de proteína.

10 18. Método según la reivindicación 17, donde semilla de soja se incluye entre al menos una fuente de proteína.

19. Uso de al menos una proteasa de cualquiera de las reivindicaciones 1-2 en detergentes.