



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 358 092**

51 Int. Cl.:

C12N 9/52 (2006.01)

C12N 15/57 (2006.01)

A23J 3/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04762907 .6**

96 Fecha de presentación : **08.10.2004**

97 Número de publicación de la solicitud: **1673451**

97 Fecha de publicación de la solicitud: **28.06.2006**

54 Título: **Variantes de proteasa.**

30 Prioridad: **10.10.2003 DK 2003 01494**
01.03.2004 DK 2004 00333

45 Fecha de publicación de la mención BOPI:
05.05.2011

45 Fecha de la publicación del folleto de la patente:
05.05.2011

73 Titular/es: **NOVOZYMES A/S**
Krogshøjvej 36
2880 Bagsværd, DK

72 Inventor/es: **De Maria, Leonardo;**
Andersen, Carsten;
Christensen, Lars, Lehmann, Hylling;
Lassen, Søren, Flensted y
Østergaard, Peter Rahbek

74 Agente: **Tomás Gil, Tesifonte Enrique**

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

Variantes de proteasa.

5 **Campo de la invención**

La presente invención se refiere a una estructura tridimensional de la proteasa nueva, al igual que variantes de una proteasa progenitora, en particular, variantes de propiedades enmendadas, tales como termoestabilidad mejorada y/o perfil de actividad de temperatura enmendada. La invención también se refiere a secuencias de ADN que codifican tales variantes, su producción en una célula huésped recombinante, al igual que métodos de uso de las variantes, en particular, en el campo de pienso para animales y detergentes. La invención además se refiere a métodos de generación y preparación de variantes de proteasa de propiedades enmendadas. Las proteasas progenitoras preferidas son proteasas de *Nocardiosis*, tales como proteasas que comprenden las partes de péptido maduro de SEC ID n.ºs: 2, 4, 6, 8, 10 y 21.

15 **Antecedentes de la invención**

Las secuencias de proteasa derivadas de cepas de *Nocardiosis* se describen en WO 88/03947, WO 01/58276 y DK 1996 00013 ("Proteasa 10", SEC ID n.ºs: 1-2).

JP 2003284571-A describe, como las SEC ID n.ºs: 2 y 1, la secuencia de aminoácidos y la secuencia de ADN correspondiente, respectivamente, de una proteasa derivada de *Nocardiosis* sp. TOA-1 (FERM P-18676). Las secuencias han sido introducidas en la base de datos GENESEQ como GENESEQP n.º ADF43564 y GENESEQN n.º ADF43563, respectivamente.

JP 2-255081-A describe una proteasa derivada de la cepa *Nocardiosis* sp. OPC-210 (FERM P-10508), no obstante, sin información de secuencia. La cepa ya no está disponible, puesto que el depósito fue retirado.

DD 200432 8 describe una preparación proteolítica derivada de la cepa *Nocardiosis dassonvillei* ZIMET 43647, no obstante, sin información de secuencia. La cepa parece ya no estar disponible.

Se describen secuencias de proteasa de *Nocardiosis* adicionales en PCT/DK04/000433 ("Proteasa 08" SEC ID n.ºs 9-10 en el presente documento); PCT/DK04/000434 ("Proteasa 11", SEC ID n.ºs: 5-6 en el presente documento); PCT/DK04/000432 ("Proteasa 18", SEC ID n.ºs 3-4 en el presente documento); y PCT/DK04/000435 ("Proteasa 35", SEC ID n.ºs 7-8 en el presente documento).

Es un objeto de la presente invención proporcionar proteasas alternativas, en particular, para el uso en pienso para animales y/o detergentes, en particular, variantes de proteasa mejoradas y nuevas, preferiblemente de propiedades enmendadas, tales como termoestabilidad mejorada y/o una temperatura óptima más alta o inferior.

40 **Resumen de la invención**

La presente invención se refiere a una variante de una proteasa progenitora, que comprende una sustitución en al menos una posición de al menos una región seleccionada del grupo de regiones que consisten en: 6-18, 22-28, 32-39, 42-58, 62-63, 66-76, 78-100, 103-106, 111-114, 118-131, 134-136, 139-141, 144-151, 155-156, 160-176, 179-181 y 184-188 donde

(a) la variante tiene actividad de la proteasa; y

50 (b) cada posición corresponde a una posición de aminoácidos 1 a 188 de SEC ID n.º 2 y

(c) la variante tiene un porcentaje de identidad con los aminoácidos 1 a 188 de SEC ID n.º 2 de al menos 60%.

La presente invención también se refiere a secuencias de ácidos nucleicos aisladas que codifican la variante de proteasa y a constructos de ácidos nucleicos, vectores y células huéspedes que comprenden las secuencias de ácidos nucleicos al igual que métodos para producir y usar las variantes de la proteasa.

Breve descripción de las figuras

60 La figura 1 es una alineación múltiple de la proteasa 10, la proteasa 18, la proteasa 11, la proteasa 35 y la proteasa 08 (las partes de péptido maduro de SEC ID n.ºs 2, 4, 6, 8 y 10, respectivamente), que también incluye una variante de la proteasa de la invención, es decir, la proteasa 22 (aminoácidos 1-188 de SEC ID n.º 21);

La figura 2 proporciona las coordenadas de la estructura tridimensional nueva de la proteasa 10 (aminoácidos 1 a 188 de SEC ID n.º 2) derivada de *Nocardiosis* sp. NRRL 18262.

Descripción detallada de la invención*Estructura tridimensional de la proteasa 10*

5 La estructura de la proteasa 10 fue resuelta conforme a los principios para los métodos cristalográficos de rayos X según se ofrecen, por ejemplo, en X-Ray Structure Determination, Stout, G.K. y Jensen, L.H., John Wiley & Sons, Inc. NY, 1989. Las coordenadas estructurales para la estructura cristalina con una resolución de 2.2 Å usando el método de sustitución isomorfa se presentan en la figura 2 en formato de PDB estándar (Protein Data Bank, Brookhaven National Laboratory, Brookhaven, CT). El archivo PDB de la figura 2 se refiere a la parte de péptido maduro de la proteasa 10
10 correspondiente a los residuos 1-188 de SEC ID n.º 2.

Dinámica Molecular (DM)

15 Las simulaciones de la Dinámica Molecular (DM) son una indicación de la movilidad de los aminoácidos en una estructura de proteína (véase McCammon, JA y Harvey, SC., (1987), "Dynamics of proteins and nucleic acids", Cambridge University Press). Tal dinámica de las proteínas es frecuentemente comparada con los factores B cristalográficos (véase Stout, GH y Jensen, LH, (1989), "X-ray structure determination", Wiley). Al ejecutar la simulación de DM a, por ejemplo, temperaturas diferentes, se simula la movilidad relacionada con la temperatura de los residuos. Las regiones que tienen la movilidad o la flexibilidad máxima (en este caso, fluctuaciones isotrópicas) se pueden sugerir
20 para mutagénesis aleatoria. Se entiende aquí que la alta movilidad encontrada en áreas determinadas de la proteína puede ser térmicamente mejorada substituyendo estos residuos.

Usando los programas CHARMM (Accelrys) y NAMD (Universidad de Illinois en Urbana-Champaign), la estructura de la proteasa 10 anteriormente descrita fue sometida a DM a 300 y 400 K. Empezando a partir de las coordenadas de la figura 2, se construyó hidrógeno y átomos pesados faltantes usando los procedimientos de CHARMM HBUILD e IC BUILD, respectivamente. Luego, la estructura fue minimizada usando el procedimiento de minimización de Gradientes conjugados CHARMM (CONJ) para un total de 200 pasos. La proteína luego fue colocada en una caja de 70 X 70 X 70 Ångstrom y disuelta con moléculas de agua TIP3. Se agregó un total de 11124 moléculas de agua y luego fueron minimizadas, manteniendo fijas las coordenadas de proteína, usando el procedimiento de minimización Adopted Basis Newton Raphson (ABNR) de CHARMM para 20000 pasos. El sistema luego fue calentado a la temperatura deseada a razón de 1 K cada 100 pasos usando el software NAMD. Después de un equilibrado de 50 picosegundos, se realizó una DM de conjunto NVE durante 1 nanosegundo. Ambos pasos se realizaron con el software NAMD. Se utilizó un corte de 12 Ångstrom para las interacciones no ligadas. Las condiciones límites periódicas se utilizaron después del paso de disolución y para todos los posteriores. La fluctuaciones isotrópicas de valor medio cuadrado (VMC)
35 se calcularon con el procedimiento COOR DYNA de CHARMM.

Las siguientes regiones para mutagénesis sugeridas surgen de simulaciones de DM: del residuo 160 al 170, del residuo 78 al 90, del residuo 43 al 50, del residuo 66 al 75, y del residuo 22 al 28.

Estrategia para preparar variantes

40 Las regiones de residuos de aminoácidos, al igual que las sustituciones de aminoácido individuales, fueron sugeridas para mutagénesis sobre la base de la estructura tridimensional de la figura 2 y la alineación de las cinco proteasas conocidas (las cinco filas superiores de la figura 1), principalmente, con el propósito de mejorar la termoestabilidad.

45 Las siguientes regiones fueron sugeridas, cf. reivindicación 1: 6-18, 22-28, 32-39, 42-58, 62-63, 66-76, 78-100, 103-106, 111-114, 118-131, 134-136, 139-141, 144-151, 155-156, 160-176, 179-181 y 184-188.

Al menos una de las siguientes posiciones de las regiones anteriores son preferiblemente sometidas a mutagénesis, cf. reivindicación 3, 6, 7, 8, 9, 10, 12, 13, 16, 17, 18, 22, 23, 24, 25, 26, 27, 28, 32, 33, 37, 38, 39, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 58, 62, 63, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 103, 105, 106, 111, 113, 114, 118, 120, 122, 124, 125, 127, 129, 130, 131, 134, 135, 136, 139, 140, 141, 144, 145, 146, 147, 148, 149, 150, 151, 155, 156, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 179, 180, 181, 184, 185, 186, 187 y/o 188.
55

Las variantes específicas contempladas se presentan en las reivindicaciones, es decir, las variantes de la proteasa 10, la proteasa 18, la proteasa 11, la proteasa 35, al igual que la proteasa 08 en las reivindicaciones 4 y 15; las variantes de la proteasa 10 en la reivindicación 16; las variantes de la proteasa 18 en la reivindicación 17; las variantes de la proteasa 11 en la reivindicación 18; la variantes de la proteasa 35 en la reivindicación 19 y las variantes de la proteasa 08 en la reivindicación 20.
60

Los diferentes conceptos subyacentes de la invención también se reflejan en las reivindicaciones de la siguiente manera: La estabilización por puentes de disulfuro en las reivindicaciones 5 y 6, la estabilización de prolina en las reivindicaciones 7-8, la sustitución de residuos neutros expuestos con residuos negativamente cargados en las reivindicaciones 9-10; la sustitución de residuos neutros expuestos con residuos cargados positivamente en las reivindicaciones 11-12; la sustitución de residuos pequeños con residuos más voluminosos en el interior de la proteína en la reivindicación 13; y las regiones propuestas para mutagénesis siguiendo las simulaciones de DM en la reivindicación 14.
65

ES 2 358 092 T3

El término “al menos uno” significa “uno o más”, es decir, por ejemplo en el contexto de las regiones: uno, dos, tres, cuatro, cinco, seis, siete, ocho nueve, diez, once, doce, trece, catorce, quince, dieciséis o diecisiete; o, en el contexto de posiciones o sustituciones: uno, dos, tres, cuatro, cinco, etcétera, hasta, por ejemplo, noventa.

5 En una forma de realización particular, el número de regiones propuesto y/o sometido a mutagénesis es al menos uno, dos, tres, cuatro, cinco, seis, siete, ocho nueve, diez, once, doce, trece, catorce, quince, dieciséis, o, al menos, diecisiete.

10 En otra forma de realización particular, el número de regiones propuesto y/o sometido a mutagénesis es no más de uno, dos, tres, cuatro, cinco, seis, siete, ocho nueve, diez, once, doce, trece, catorce, quince, dieciséis, o no más de diecisiete.

Polipéptidos que tienen actividad de la proteasa

15 Los polipéptidos que tienen actividad de la proteasa, o las proteasas, a veces también son llamados peptidasas, proteasas, hidrolasas peptídicas o enzimas proteolíticas. Las proteasas pueden ser del tipo exo que hidrolizan péptidos comenzando en cualquier extremo de las mismas, o del tipo endo que actúa internamente en cadenas de polipéptido (endopeptidasas). Las endopeptidasas muestran actividad en sustratos de péptido bloqueado N- y C-terminalmente que son pertinentes para la especificidad de la proteasa en cuestión.

20 El término “proteasa” se define en la presente como una enzima que hidroliza enlaces peptídicos. Esta definición de la proteasa también se aplica a la parte de proteasa de los términos “proteasa progenitora” y “variante de proteasa” según se utiliza en el presente documento. El término “proteasa” incluye cualquier enzima del grupo enzimático EC 3.4 (incluidas cada una de las trece subclases del mismo). El número EC se refiere a nomenclatura enzimática 1992 de NC-IUBBM, Academic Press, San Diego, California, incluidos los suplementos 1-5 publicados en Eur. J. Bio-chem. 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 habitualmente es suplementada y actualizada; véase, por ejemplo, la página web <http://www.chem.qmw.ac.uk/iubmb/enzyme/index.html>.

30 Las proteasas son clasificadas basándose en su mecanismo catalítico en los siguientes grupos: serina proteasas (S), cisteína proteasas (C), proteasas aspárticas (A), metaloproteasas (M), y proteasas (U) desconocidas o aún sin clasificar; véase Handbook of Proteolytic Enzymes, A. J. Barrett, N. D. Rawlings, J. F. Woessner (eds), Academic Press (1998), en particular, la parte de la introducción general.

35 En formas de realización particulares, las proteasas progenitoras y/o las variantes de proteasa de la invención y para el uso según la invención son seleccionadas del grupo que consiste en:

(a) proteasas del grupo enzimático EC 3.4;

40 (b) serina proteasas del grupo S del manual anterior;

(c1) serina proteasas de la familia de peptidasa S2A; y

45 (c2) serina proteasas de la familia de peptidasa S1E como se describe en Biochem. J. 290:205-218 (1993) y en la base de datos de proteasa MEROPS, publicación 6.20, 24 de marzo de 2003, (www.merops.ac.uk). La base de datos es descrita en Rawlings, N. D., O'Brien, E. A. & Barrett, A. J. (2002) MEROPS: the protease database. Nucleic Acids Res. 30, 343-346.

50 Para determinar si una proteasa determinada es una serina proteasa, y una proteasa de la familia S2A, se hace referencia al manual anterior y a los principios indicados en el mismo. Tal determinación puede llevarse a cabo para todos los tipos de proteasas, ya sean de origen natural o proteasas de tipo salvaje; o proteasas sintéticas o creadas genéticamente.

55 La actividad de la proteasa puede medirse usando cualquier ensayo en el que se emplee un sustrato, que incluya enlaces peptídicos pertinentes para la especificidad de la proteasa en cuestión. El pH del ensayo y la temperatura del ensayo igualmente deben ser adaptados a la proteasa en cuestión. Ejemplos de valores de pH del ensayo son 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. Ejemplos de ensayos de proteasa adecuados se describen en la parte experimental.

Proteasa progenitora

60 La proteasa progenitora es una proteasa de la cual la variante de proteasa es, o, puede ser derivada. Para los objetivos presentes, cualquier proteasa se puede usar como la proteasa progenitora, siempre que la variante de proteasa resultante sea homóloga a la proteasa 10, es decir, la proteasa derivada de *Nocardia* sp. NRRL 18262 y que comprenda los aminoácidos 1-188 de SEC ID n.º 2.

En una forma de realización particular, la proteasa progenitora también es homóloga a la proteasa 10.

En el presente contexto, homólogo significa que tiene una de al menos 60% con SEC ID n.º 2, es decir, aminoácidos 1-188 de la parte de péptido maduro de la proteasa 10. La homología se determina como se describe generalmente a continuación en la sección titulada Homología de aminoácidos.

- 5 La proteasa progenitora puede ser un polipéptido de tipo salvaje o de origen natural, o una variante alélica del mismo, o un fragmento del mismo que tiene actividad de la proteasa, en particular, una parte madura del mismo. También ser una variante del mismo y/o un polipéptido sintético o creado genéticamente.

- En una forma de realización particular, la proteasa progenitora de tipo salvaje 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 derivada de vii) la especie Nocardopsis, tal como *Nocardopsis alba*, *Nocardopsis antarctica*, *Nocardopsis composta*, *Nocardopsis dassonvillei*, *Nocardopsis exhalans*, *Nocardopsis halophila*, *Nocardopsis halotolerans*, *Nocardopsis kunsanensis*, *Nocardopsis listeri*, *Nocardopsis lucentensis*, *Nocardopsis metallicus*, *Nocardopsis prasina*, *Nocardopsis sp.*, *Nocardopsis synnemataformans*, *Nocardopsis trehalosi*, *Nocardopsis tropica*, *Nocardopsis umidischolae* o *Nocardopsis xinjiangensis*.

- Ejemplos de tales cepas son: *Nocardopsis alba* DSM 15647 (productor de tipo salvaje de la proteasa 08), *Nocardopsis dassonvillei* NRRL 18133 (productor de tipo salvaje de la proteasa M58-1 descrito en WO 88/03947), *Nocardopsis dassonvillei subesp. dassonvillei* DSM 43235 (productor de tipo salvaje de la proteasa 18), *Nocardopsis prasina* DSM 15648 (productor de tipo salvaje de la proteasa 11), *Nocardopsis prasina* DSM 15649 (productor de tipo salvaje de la proteasa 35), *Nocardopsis sp.* NRRL 18262 (productor de tipo salvaje de la proteasa 10), *Nocardopsis sp.* FERM P-18676 (descrito en JP 2003284571-A).

- Las cepas de estas especies son accesibles al público en un número de colecciones de cultivos, tal como American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau voor Schimmelcultures (CBS) y Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL), por ejemplo, la *Nocardopsis dassonvillei subesp. dassonvillei* DSM 43235 está disponible públicamente en DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Alemania).

- Además, tales polipéptidos se pueden identificar y obtener de otras fuentes incluidos los microorganismos o ADN aislado de la naturaleza (p. ej., tierra, abonos, agua, etc.) usando sondas adecuadas. Las técnicas para aislar microorganismos o ADN de hábitats naturales se conocen en la técnica. La secuencia de ácidos nucleicos puede entonces ser derivada seleccionando de manera similar una genoteca de ADNc o genómico de otro microorganismo. Una vez que una secuencia de ácidos nucleicos que codifica un polipéptido ha sido detectada con la/s sonda/s, la secuencia puede ser aislada o clonada utilizando técnicas conocidas por los expertos en la materia (véase, por ejemplo, Sambrook *et al.*, 1989, *supra*).

- La proteasa progenitora puede ser una parte madura de cualquiera de las secuencias de aminoácidos mencionadas anteriormente. Una parte madura significa una secuencia de aminoácidos madura y se refiere a esa parte de una secuencia de aminoácidos que permanece después de que una parte del péptido señal potencial y/o parte del propéptido ha sido cortada. Las partes maduras de cada una de las proteasas 08, 10, 11, 18, 22 y 35 se especifican en el listado de secuencias anexo.

- La proteasa progenitora también puede ser un fragmento de una secuencia de aminoácidos específica, es decir, un polipéptido que tiene uno o más aminoácidos delecionados del término amino y/o carboxilo de esta secuencia de aminoácidos. En una forma de realización, un fragmento contiene al menos 80, o al menos 90, o al menos 100, o al menos 110, o al menos 120, o al menos 130, o al menos 140, o al menos 150, o al menos 160, o al menos 170, o al menos 180, o al menos 185 residuos de aminoácidos.

- La proteasa progenitora también puede ser una variante alélica, alélica refiriéndose a la existencia de dos o más formas alternativas de un gen que ocupa la misma localización cromosómica. La variación alélica surge naturalmente a través de mutación, y puede resultar en polimorfismo dentro de poblaciones. Las mutaciones de genes pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos que tienen secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

- En otra forma de realización, la proteasa progenitora puede ser una proteasa creada genéticamente, por ejemplo, una variante del tipo salvaje o proteasas progenitoras naturales mencionadas anteriormente comprendiendo una sustitución, delección y/o inserción de uno o más aminoácidos. En otras palabras: la proteasa progenitora puede en sí misma ser una variante de proteasa, tal como la proteasa 22. La secuencia de aminoácidos de tal proteasa progenitora puede diferir de la secuencia de aminoácidos específica 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. Los cambios de aminoácidos pueden ser de una naturaleza menor o más importante. Los cambios de aminoácidos de una naturaleza más importante son, por ejemplo, aquellos que dan como resultado una proteasa variante de la presente invención con propiedades enmendadas. En otra forma de realización particular, los cambios de aminoácidos son de una naturaleza menor, es decir, sustituciones conservadoras de aminoácidos que no afectan significativamente el pliegue y/o la actividad de la proteína; delecciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; extensiones de carboxilo o amino terminal pequeñas, tales como un residuo de metionina amino terminal; un pequeño péptido de

enlace de hasta 20-25 residuos aproximadamente; o una pequeña extensión que facilita la purificación cambiando la carga neta u otra función, tal como un tracto de polihistidina, un epítipo antigénico o un dominio de enlace.

Ejemplos de sustituciones conservadoras están en el 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). Las sustituciones de aminoácido que generalmente no alteran la actividad específica se conocen en la técnica y son descritas, por ejemplo, por H. Neurath y R.L. Hill, 1979, en *The Proteins*, Academic Press, New York. Los intercambios que ocurren con mayor frecuencia 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 de forma inversa.

Otros ejemplos de proteasas progenitoras creadas genéticamente son proteasas sintéticas, diseñadas por el hombre, cuya existencia en la naturaleza es imprevista. EP 897985 describe un proceso de preparación de una proteína de consenso. Las proteasas redistribuidas son otros ejemplos de proteasas progenitoras sintéticas o creadas genéticamente, que 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. Incluido en el concepto de una proteasa sintética también está cualquier proteasa quimérica o híbrida, es decir, una proteasa que comprende una combinación de secuencias de aminoácidos parciales derivadas de al menos dos proteasas. La transposición de genes es descrita generalmente en, por ejemplo, WO 95/22625 y WO 96/00343. La recombinación de genes de proteasa puede realizarse independientemente de la secuencia específica de los progenitores por redistribución sintética como se describe en Ness, J.E. *et al*, en *Nature Biotechnology*, Vol. 20 (12), pág. 1251-1255, 2002. Los oligonucleótidos sintéticos se degeneraron en su secuencia de ADN para proporcionar la posibilidad de que todos los aminoácidos encontrados en el conjunto de proteasas progenitoras estén diseñados y los genes ensamblados según la referencia. La redistribución puede llevarse a cabo para la secuencia de longitud total o sólo para parte de la secuencia y luego combinada más tarde con el resto del gen para dar una secuencia de longitud total. Dos, tres, cuatro, cinco o los seis de las proteasas designadas proteasa 10, 18, 11, 35, 08 y 22 (SEC ID n.ºs 2, 4, 6, 8, 10 y 21, en particular, las partes maduras de las mismas) son ejemplos particulares de tales proteasas progenitoras que se pueden someter a redistribución como se ha descrito anteriormente, para proporcionar proteasas adicionales de la invención.

En otras formas de realización particulares, la proteasa progenitora comprende, o consiste en, respectivamente, la secuencia de aminoácidos específica, o una variante alélica de la misma; o un fragmento de la misma que tiene actividad de la proteasa.

En otras formas de realización particulares, la variante de proteasa de la invención no es idéntica a: (i) los aminoácidos 1-188 de SEC ID n.º 2, los aminoácidos 1-188 de SEC ID n.º 4, los aminoácidos 1-188 de SEC ID n.º 6, los aminoácidos 1-188 de SEC ID n.º 8, y los aminoácidos 1-188 de SEC ID n.º 10 (ii) los aminoácidos 1-188 de SEC ID n.º 2 (iii) los aminoácidos 1-188 de SEC ID n.º 2 con la sustitución T87A; (iv) los aminoácidos 1-188 de SEC ID n.º 4 (v) los aminoácidos 1-188 de SEC ID n.º 6 (vi) los aminoácidos 1-188 de SEC ID n.º 8 (vii) los aminoácidos 1-188 de SEC ID n.º 10 (viii) la proteasa derivada de *Nocardiopsis dassonvillei* NRRL 18133; (ix) la proteasa que tiene los aminoácidos 1 a 188 de SEC ID n.º 2 como se describe en JP 2003284571-A; (x) la proteasa que tiene la secuencia introducida en GENESEQP con n.º ADF43564; (xi) la proteasa descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 2, en particular la parte madura de la misma; (xii) la proteasa descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 4, en particular, la parte madura de la misma; (xiii) la proteasa descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 6, en particular, la parte madura de la misma; (xiv) la proteasa descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 8, en particular, la parte madura de la misma; (xv) la proteasa descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 10, en particular, la parte madura de la misma; (xvi) la proteasa descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 12, en particular, la parte madura de la misma; y/o (xvii) cualquier proteasa de la técnica anterior de un porcentaje de identidad con SEC ID n.º 2 de al menos 60%.

Taxonomía de los microorganismos

Las preguntas acerca de la taxonomía pueden ser resueltas consultando una base de datos de taxonomía, tal como el explorador de taxonomías NCBI que está disponible en el siguiente sitio de Internet: <http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html>, y/o consultando manuales de taxonomía. Para los objetivos presentes, la taxonomía es preferiblemente según el capítulo: The road map to the Manual de G.M. Garrity & J. G. Holt en *Bergey's Manual of Systematic Bacteriology*, 2001, segunda edición, volumen 1, David R. Bone, Richard W. Castenholz.

Homología de los aminoácidos

La presente invención se refiere a proteasas, es decir, proteasas progenitoras, y/o variantes de proteasa, con un cierto grado de identidad con los aminoácidos 1 a 188 de SEC ID n.º 2, tales proteasas variantes y/o progenitoras son designadas en adelante "proteasas homólogas".

Para objetivos 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 una alineación

de Needleman-Wunsch (es decir, una alineación global). El programa se usa para la alineación de polipéptido, al igual que secuencias de nucleótidos. La matriz de puntuación predeterminada BLOSUM50 se usa para alineamientos de polipéptidos, y la matriz predeterminada se usa para alineamientos de nucleótidos. 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.

“Align” es parte de la versión v20u6 del paquete de FASTA (véase W. R. Pearson and 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). Los alineamientos de proteínas de FASTA usan el algoritmo de Smith-Waterman sin ninguna limitación en tamaño de espacio (véase “Smith-Waterman algorithm”, T. F. Smith y M. S. Waterman (1981) J. Mol. Biol. 147:195-197).

Los alineamientos múltiples de secuencias proteicas pueden realizarse usando “ClustalW” (Tompson, J.D., Higgins, D.G. y Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22:4673-4680). Los alineamientos múltiples de secuencias proteicas pueden realizarse usando la alineación de proteínas como un molde, sustituyendo los aminoácidos con el codón correspondiente de la secuencia de ADN.

En formas de realización particulares, la proteasa homóloga tiene una secuencia de aminoácidos que tiene un grado de identidad con los aminoácidos 1 a 188 de SEC ID n.º 2 de al menos 60%, 62%, 64%, 66%, 68%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o de al menos 99% aproximadamente.

En formas de realización alternativas, la proteasa homóloga tiene una secuencia de aminoácidos que tiene un grado de identidad con SEC ID n.º 2 de al menos 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58% o al menos 59%.

En otra forma de realización particular, la proteasa progenitora, y/o la variante de proteasa, comprende una secuencia de aminoácidos madura que difiere de no más de setenta y cinco, setenta y cuatro, setenta y tres, setenta y dos, setenta y uno, setenta, sesenta y nueve, sesenta y ocho, sesenta y siete, sesenta y seis, sesenta y cinco, sesenta y cuatro, sesenta y tres, sesenta y dos, sesenta y uno, sesenta, cincuenta y nueve, cincuenta y ocho, cincuenta y siete, cincuenta y seis, cincuenta y cinco, cincuenta y cuatro, cincuenta y tres, cincuenta y dos, cincuenta y uno, cincuenta, cuarenta y nueve, cuarenta y ocho, cuarenta y siete, cuarenta y seis, cuarenta y cinco, cuarenta y cuatro, cuarenta y tres, cuarenta y dos, cuarenta y uno, cuarenta, treinta y nueve, treinta y ocho, treinta y siete, treinta y seis, treinta y cinco, treinta y cuatro, treinta y tres, treinta y dos, treinta y uno, treinta, veintinueve, veintiocho, veintisiete, veintiséis, veinticinco, veinticuatro, veintitrés, veintidós, veintiuno, diecinueve, dieciocho, diecisiete, dieciséis, quince, catorce, trece, doce, once, diez, nueve, ocho, siete, seis, cinco, cuatro, tres, de no más de dos aminoácidos, o sólo por un aminoácido de la secuencia de aminoácidos específica, por ejemplo, los aminoácidos 1 a 188 de SEC ID n.º 2.

En otra forma de realización particular, la proteasa progenitora, y/o la variante de proteasa, comprende una secuencia de aminoácidos madura que difiere de al menos setenta y cinco, setenta y cuatro, setenta y tres, setenta y dos, setenta y uno, setenta, sesenta y nueve, sesenta y ocho, sesenta y siete, sesenta y seis, sesenta y cinco, sesenta y cuatro, sesenta y tres, sesenta y dos, sesenta y uno, sesenta, cincuenta y nueve, cincuenta y ocho, cincuenta y siete, cincuenta y seis, cincuenta y cinco, cincuenta y cuatro, cincuenta y tres, cincuenta y dos, cincuenta y uno, cincuenta, cuarenta y nueve, cuarenta y ocho, cuarenta y siete, cuarenta y seis, cuarenta y cinco, cuarenta y cuatro, cuarenta y tres, cuarenta y dos, cuarenta y uno, cuarenta, treinta y nueve, treinta y ocho, treinta y siete, treinta y seis, treinta y cinco, treinta y cuatro, treinta y tres, treinta y dos, treinta y uno, treinta, veintinueve, veintiocho, veintisiete, veintiséis, veinticinco, veinticuatro, veintitrés, veintidós, veintiuno, diecinueve, dieciocho, diecisiete, dieciséis, quince, catorce, trece, doce, once, diez, nueve, ocho, siete, seis, cinco, cuatro, tres, por al menos dos aminoácidos, o por un aminoácido de la secuencia de aminoácidos específica, por ejemplo, los aminoácidos 1 a 188 de SEC ID n.º 2.

Hibridación de ácido nucleicos

En la alternativa, las proteasas homólogas progenitoras, al igual que las proteasas variantes, pueden definirse como codificadas por una secuencia de ácidos nucleicos que se hibrida bajo condiciones de astringencia muy baja, preferiblemente, condiciones de astringencia baja, más preferiblemente, condiciones de astringencia media, más preferiblemente, condiciones de astringencia media-alta, incluso más preferiblemente, condiciones de astringencia alta, y de forma más preferible, condiciones de astringencia muy alta con los nucleótidos 900 1466, o 900-1463, de SEC ID n.º 1 o una subsecuencia o una cadena complementaria de la misma (J. Sambrook, E.F. Fritsch y T. Maniatus, 1989, Molecular Cloning, A Laboratory Manual, 2ª Edición, Cold Spring Harbor, New York). Una subsecuencia puede ser de al menos 100 nucleótidos, o al menos 200, 300, 400, o al menos 500 nucleótidos. Por otra parte, la subsecuencia puede codificar un fragmento de polipéptido que tiene la actividad enzimática pertinente.

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

ES 2 358 092 T3

Para sondas largas de al menos 100 nucleótidos de longitud, el material portador es finalmente lavado tres veces durante 15 minutos cada vez usando 2 x SSC, SDS al 0,2% 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 media), más preferiblemente al menos a 60°C (astringencia medio alta), incluso más preferiblemente al menos a 65°C (astringencia alta) y de la forma más preferible, al menos a 70°C (astringencia muy alta).

Para sondas cortas que son de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia se definen como prehibridación, hibridación y lavado post-hibridación a 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 USA 48:1390) en 0,9 M NaCl, 0,09 M Tris-HCl pH 7,6, 6 mM EDTA, 0,5% NP-40, 1X solución de Denhardt, 1 mM pirofosfato de sodio, 1 mM fosfato monobásico de sodio, 0,1 mM ATP y 0,2 mg de ARN de levadura por ml siguiendo procedimientos de transferencia de Southern estándar.

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

Numeración de posición

En el presente contexto, la base para numerar las posiciones es aminoácidos 1 a 188 de SEC ID n.º 2, proteasa 10, partiendo con A1 y finalizando con T188, véase la figura 1. Una proteasa progenitora, al igual que una proteasa variante, puede comprender extensiones en comparación con SEC ID n.º 2, es decir, en el N-terminal, y/o las extremidades de C-terminales de las mismas. Los aminoácidos de tales extensiones, si los hay, se deben numerar como es usual en la técnica, es decir, para una extensión C-terminal: 189, 190, 191 etcétera, y para una extensión N-terminal -1; -2; -3, etcétera.

Alteraciones, tales como sustituciones, deleciones, inserciones

En el presente contexto, los siguientes son ejemplos de varias maneras en las que una variante de proteasa puede ser diseñada o derivada de una secuencia de aminoácidos progenitora: un aminoácido se pueden sustituir por otro aminoácido; un aminoácido pueden ser delecionado; un aminoácido puede ser insertado; al igual que cualquier combinación de cualquier número de tales alteraciones.

Para los objetivos presentes, el término sustitución tiene por objeto incluir cualquier número de cualquier tipo de tales alteraciones. Ésta es una definición razonable, porque, por ejemplo, una deleción se puede considerar una sustitución de un aminoácido, AA, en una posición determinada, nn, con nada, (). Tal sustitución puede ser designada: AAnn(). Asimismo, una inserción de sólo un aminoácido, BB, debajo de un aminoácido, AA, en una posición, nn, puede ser designada: ()nnaBB. Y si dos aminoácidos, BB y CC, son insertados abajo del aminoácido AA en la posición nn, esta sustitución (combinación de dos sustituciones) puede ser designada: ()nnaBB+()nnbCC. Los espacios creados de este modo entre los aminoácidos nn y nn+1 en la secuencia madre tienen letra minúscula o de subíndice asignadas a, b, c etc. al número de posición anterior, en este caso, nn. Se sigue un procedimiento de numeración similar cuando se alinea una nueva secuencia a la alineación múltiple de la figura 1, en caso de que se cree un espacio por la alineación entre los aminoácidos nn y nn+1: se asigna un número a cada posición del espacio: nna, nnb etc. Una coma (,) entre sustituyentes, como, por ejemplo, en la sustitución T129E, D,Y,Q significa “bien, o”, es decir que T129 se sustituye con E, o D, o Y, o Q. Un signo de más (+) entre las sustituciones, por ejemplo 129D+135P significa “y”, es decir, que estas dos únicas sustituciones se combinan en una y la misma variante de proteasa.

En el presente contexto, el término “una sustitución” significa al menos una sustitución. Al menos una significa uno o más, por ejemplo uno, o dos, o tres, o cuatro, o cinco, o seis, o siete, u ocho, o nueve, o diez, o doce, o catorce, o quince, o dieciséis, o dieciocho, o veinte, o veintidós o veinticuatro, o veinticinco, o veintiocho, o treinta, etcétera, para incluir, en principio, cualquier número de sustituciones. Las variantes de la invención, no obstante, todavía deben ser, por ejemplo, al menos 60% idénticas a SEC ID n.º 2. Este porcentaje se determina por el programa mencionado arriba. Las sustituciones se pueden aplicar a cualquier posición comprendida por cualquier región mencionada en la reivindicación 1, y también están incluidas las variantes que comprenden combinaciones de cualquier número y tipo de tales sustituciones. El término sustitución, según se utiliza en este caso, también incluye deleciones, al igual que extensiones, o inserciones, que pueden contribuir a la longitud de la secuencia correspondiente los aminoácidos 1 a 188 de SEC ID n.º 2.

Además, el término “una sustitución” incluye una sustitución en cualquiera de los otros diecinueve aminoácidos naturales, o en otros aminoácidos, tales como aminoácidos no naturales. Por ejemplo, una sustitución del aminoácido T en la posición 22 incluye cada una de las siguientes sustituciones: 22A, 22C, 22D, 22E, 22F, 22G, 22H, 22I, 22K, 22L, 22M, 22N, 22P, 22Q, 22R, 22S, 22V, 22W y 22Y. Esto es, a propósito, equivalente a la designación 22X, donde X designa cualquier aminoácido. Estas sustituciones también pueden ser designadas T22A, T22C, T22X, etc. Lo mismo se aplica por analogía a todas y cada una de las posiciones mencionadas en la presente, para incluir específicamente, en este caso, cualquiera de tales sustituciones.

Números de identificación de posición correspondientes

Para cada residuo de aminoácido en cada proteasa progenitora o variante de la invención, y/o para el uso según la invención, es posible asignar directamente y sin ambigüedad un residuo de aminoácido en la secuencia de los aminoácidos 1 a 188 de SEC ID n.º 2 a la cual corresponde. Se asigna el mismo número a los residuos correspondientes, por referencia a la secuencia de la proteasa 10.

Como aparece según la numeración de la figura 1, conjuntamente con la numeración del listado de secuencias, para cada residuo de aminoácido de cada una de la proteasas proteasa 10, proteasa 18, proteasa 11, proteasa 35, proteasa 08 y proteasa 22, el residuo de aminoácido correspondiente en SEC ID n.º 2 tiene el mismo número. Este número puede derivarse fácilmente de la figura 1. Al menos en el caso de estas seis proteasas, el número es el mismo que el número asignado a este residuo de aminoácido en el listado de secuencias para la parte madura de la proteasa respectiva.

Para una posición dada en otra proteasa - sea ésta una proteasa progenitora o una variante - siempre puede encontrarse una posición correspondiente de SEC ID n.º 2, de la siguiente manera:

La secuencia de aminoácidos de otra proteasa progenitora, o, a su vez, de una secuencia de aminoácidos de proteasa variante, es designada SEQ-X. Una posición correspondiente a la posición N de SEC ID n.º 2 se encuentra de la siguiente manera: la secuencia de aminoácidos de proteasa variante o progenitora SEQ-X se alinea con SEC ID n.º 2 como se especificó más arriba en la sección titulada Homología de los aminoácidos. De la alineación, la posición en la secuencia SEQ-X correspondiente a la posición N de SEC ID n.º 2 puede ser derivada de manera clara y sin ambigüedad, usando los principios descritos a continuación.

La SEQ-X es la parte madura de la proteasa en cuestión. En la alternativa, también puede incluir una parte de péptido señal, y/o una parte de propéptido, o puede ser un fragmento de la proteasa madura que tiene actividad de la proteasa, por ejemplo, un fragmento de la misma longitud que SEC ID n.º 2, y/o puede ser el fragmento que se extiende desde A1 a T188 cuando está alineado con SEC ID n.º 2 como se describe en este caso.

Región y posición

En el presente contexto, el término región significa al menos una posición de una secuencia de aminoácidos de proteasa progenitora. El término posición designa un residuo de aminoácido de tal secuencia de aminoácidos. En una forma de realización, región significa una o más posiciones sucesivas de la secuencia de aminoácidos de proteasa progenitora, por ejemplo, uno, dos, tres, cuatro, cinco, seis, siete, ocho, etc., hasta cualquier número de posiciones consecutivas de la secuencia. Por consiguiente, una región puede consistir en una posición solamente, o puede consistir en cualquier número de posiciones consecutivas, tales como, por ejemplo, la posición n.º 62 y 63 o la posición n.º 111, 112, 113 y 114. Para los objetivos presentes, estas dos regiones son designadas 62-63, y 111-114, respectivamente. Los límites de estas regiones o gamas están incluidos en la región.

Una región comprende específicamente todas y cada una de las posiciones que abarca. Por ejemplo, la región 111-114 específicamente comprende todas y cada una de las posiciones 111, 112, 113 y 114. Lo mismo se aplica por analogía a las otras regiones mencionadas en este caso.

Termoestabilidad

Para los objetivos presentes, el término termoestable según se aplica en el contexto de un polipéptido determinado, se refiere a la temperatura de fusión, T_m, de tal polipéptido, según se determina usando calorimetría por análisis diferencial (DSC por sus siglas en inglés) en 10 mM fosfato sódico, 50 mM cloruro sódico, pH 7, 0, usando una tasa de barrido constante de 1,5°C/min.

Las siguientes T_m fueron determinadas bajo las condiciones anteriores: 76,5°C (proteasa 10), 83,0°C (proteasa 18), 78,3°C (proteasa 08), 76,6°C (proteasa 35), 73,7°C (proteasa 11) y 83,5°C (proteasa 22).

Para un polipéptido termoestable, la T_m es al menos 83,1°C. En formas de realización particulares, la T_m es al menos 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o al menos 100°C.

En la alternativa, el término termoestable se refiere a una temperatura de fusión de al menos 73,8, o al menos 76,7°C, o al menos 78,4°C, preferiblemente al menos 74, 75, 76, 77, 78, 79, 80, 81, 82 o al menos 83°C, aún como se determina usando calorimetría por análisis diferencial con un pH de 7,0.

Para la determinación de la T_m, puede usarse una muestra del polipéptido con una pureza de al menos 90% (o 91, 92, 93, 94, 95, 96, 97 o 98%) como se determina por SDS-PAGE. Aun más, la muestra enzimática puede tener una concentración de entre 0,5 y 2,5 mg/ml de proteína (o entre 0,6 y 2,4, o entre 0,7 y 2,2, o entre 0,8 y 2,0 mg/ml de proteína), como se determina a partir de la absorbencia a 280 nm y sobre la base de un coeficiente de extinción calculado a partir de la secuencia de aminoácidos de la enzima en cuestión.

La calorimetría por análisis diferencial se desarrolla al pH deseado (p. ej. pH 5,5; 7,0; 3,0 o 2,5) y con un índice de calentamiento constante, por ejemplo de 1; 1,5; 2, 3, 4, 5, 6, 7, 8, 9 o 10°C/min.

En una forma de realización particular, la variante de proteasa de la invención es termoestable, preferiblemente más termoestable que la proteasa progenitora. En este contexto, las proteasas progenitoras preferidas son la proteasa 18, o la proteasa 10.

- 5 En otra forma de realización particular, un sobrenadante del cultivo de la variante de proteasa de la invención, apropiadamente diluido, muestra una actividad residual tras la incubación durante cuatro horas a 65°C en un tampón de 0,2M Na₂HPO₄, titulado con 0,1 M ácido cítrico a i) pH 6,0, o ii) pH 4,0, de al menos 20%, en relación con un control (congelado) no incubado. La actividad se mide usando el ensayo de Protazyme AK a pH 8, 5 y 37°C, como se describe en el ejemplo 2. En otras formas de realización particulares, la actividad residual es al menos 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, o al menos 77%.

Perfil de actividad de la temperatura

- 15 En una forma de realización particular, la variante de proteasa de la invención muestra un perfil de actividad de la temperatura enmendada en comparación con, por ejemplo, la proteasa 10 (o la proteasa 18, la proteasa 11, la proteasa 35 o la proteasa 08). Por ejemplo, la variante de proteasa de la invención puede mostrar una actividad relativa con pH 9 y 80°C de al menos 0,40, preferiblemente al menos 0,45; 0,50, 0,55; 0,60; 0,65; 0,70; 0,75; 0,80; 0,85; 0,90, o al menos 0,95. El término “relativa” se refiere a la actividad máxima medida para la proteasa en cuestión. Para la proteasa 22, la actividad es relativa a la actividad a 80°C que se fija en 1,000 (100%), y para la proteasa 10, la actividad a 70°C se fija en 1,000 (100%); véase el ejemplo 3. Como otro ejemplo, la variante de la proteasa de la invención muestra una actividad relativa con pH 9 y 90°C de al menos 0,10, preferiblemente al menos 0,15, 0,20; 0,25; 0,30 o de al menos 0,35. En una forma de realización particular, la actividad de la proteasa se mide usando el ensayo de Protazyme AK del ejemplo 1.

Variantes poco alergénicas

- En una forma de realización específica, las variantes de proteasa de la presente invención son (también) variantes poco alergénicas, diseñadas para invocar una respuesta inmunológica reducida al ser expuestas a animales, incluido el hombre. El término respuesta inmunológica debe ser entendido como cualquier reacción por el sistema inmunológico de un animal expuesto a la variante de proteasa. Un tipo de respuesta inmunológica es una respuesta alérgica que conduce a niveles aumentados de IgE en el animal expuesto. Las variantes poco alergénicas pueden ser preparadas usando técnicas conocidas en la técnica. Por ejemplo, la variante de proteasa se puede conjugar con partes de protección de fracciones de polímero o epítomos de la variante de proteasa implicados en una respuesta inmunológica. La conjugación con polímeros pueden implicar acoplamiento químico *in vitro* de polímero a la variante de proteasa, por ejemplo, como se describe en WO 96/17929, WO 98/30682, WO 98/35026 y/o WO 99/00489. La conjugación además o alternativamente puede implicar acoplamiento *in vivo* de polímeros a la variante de proteasa. Tal conjugación puede lograrse por creación genética de la secuencia de nucleótidos que codifica la variante de la proteasa, insertando secuencias de consenso que codifican sitios de glicosilación adicional en la variante de proteasa y expresando la variante de proteasa en un huésped capaz de glicosilar la variante de proteasa; véase, por ejemplo, WO 00/26354. Otro modo de conseguir variantes poco alergénicas es creando genéticamente la secuencia de nucleótidos que codifica la variante de proteasa para causar que las variantes de proteasa se auto-oligomericen, ocasionando que los monómeros de la variante de proteasa puedan proteger los epítomos de otros monómeros de la variante de proteasa y así reducir la antigenicidad de los oligómeros. Tales productos y su preparación se describen, por ejemplo, en WO 96/16177. Los epítomos implicados en una respuesta inmunológica pueden identificarse por varios métodos, tales como el método de exposición en fago descrito en WO 00/26230 y WO 01/83559, o el enfoque aleatorio descrito en EP 561907. Una vez que un epítopo ha sido identificado, su secuencia de aminoácidos se puede alterar para producir propiedades alteradas inmunológicas de la variante de proteasa por técnicas de manipulación génica conocidas tales como mutagénesis dirigida (véase, por ejemplo WO 00/26230, WO 00/26354 y/o WO 00/22103) y/o puede realizarse la conjugación de un polímero en proximidad suficiente al epítopo para que el polímero proteja al epítopo.

Constructos y secuencias de ácidos nucleicos

La presente invención también se refiere a secuencias de ácidos nucleicos que comprenden una secuencia de ácidos nucleicos que codifica una variante de proteasa de la invención.

- 55 El término “secuencia de ácidos nucleicos aislada” se refiere a una secuencia de ácidos nucleicos que es esencialmente libre de otra secuencia de ácidos nucleicos, por ejemplo, al menos aproximadamente 20% pura, preferiblemente al menos aproximadamente 40% pura, más preferiblemente al menos aproximadamente 60%, pura, incluso más preferiblemente al menos aproximadamente 80% pura, y de forma más preferible, al menos aproximadamente 90% pura como se determina por electroforesis de agarosa. Por ejemplo, una secuencia de ácidos nucleicos aislada se puede obtener por procedimientos de clonación estándares usados en ingeniería genética para recolocar la secuencia de ácidos nucleicos de su ubicación natural a un sitio diferente donde será reproducida. Los procedimientos de clonación pueden implicar escisión y aislamiento de un fragmento de ácido nucleico deseado que comprende 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 se replicarán copias múltiples o clones de la secuencia de ácidos nucleicos. La secuencia de ácidos nucleicos puede ser de origen de ADN genómico, ADNc, ARN, semisintético, sintético, o cualquier combinación de los mismos.

Las secuencias de ácidos nucleicos de la invención se pueden preparar por introducción de al menos una mutación en la secuencia codificante de proteasa progenitora o una subsecuencia de la misma, donde la secuencia de ácidos nucleicos mutante codifica una proteasa variante. La introducción de una mutación en la secuencia de ácidos nucleicos para intercambiar un nucleótido por otro nucleótido se puede realizar por mutagénesis dirigida usando cualquiera de los métodos conocidos en la técnica, por ejemplo, por mutagénesis dirigida, por mutagénesis aleatoria, o por mutagénesis aleatoria localizada, adicionada o dopada.

La mutagénesis aleatoria es adecuadamente realizada bien como mutagénesis aleatoria específica de la región o localizada en al menos tres partes del gen que se traducen en la secuencia de aminoácidos mostradas en cuestión, o en el gen entero. Cuando la mutagénesis se realiza por el uso de un oligonucleótido, el oligonucleótido puede ser dopado o adicionado con los tres nucleótidos no progenitores durante la síntesis del oligonucleótido en las posiciones que se van a cambiar. El dopaje o el adicionado se puede realizar de modo que se eviten los codones para aminoácidos indeseados. El oligonucleótido adicionado o dopado se puede incorporar en el ADN que codifica la enzima de proteasa por cualquier técnica, usando, por ejemplo, PCR, LCR o cualquier ADN polimerasa y ligasa según se considere apropiado.

Preferiblemente, el dopaje se realiza usando “dopaje aleatorio constante”, en el cual se define previamente el porcentaje de mutación y el tipo salvaje en cada posición. Además, el dopaje puede ser dirigido hacia una preferencia para la introducción de nucleótidos determinados, y así una preferencia para la introducción de uno o más residuos de aminoácidos específicos. El dopaje puede realizarse, por ejemplo, para permitir la introducción de 90% de tipo salvaje y 10% de mutaciones en cada posición. Una consideración adicional en la elección de un esquema de dopaje se basa en limitaciones estructurales de proteínas y genéticas.

La mutagénesis aleatoria puede ser localizada ventajosamente en una parte de la proteasa progenitora en cuestión. Esto puede ser ventajoso, por ejemplo, cuando las regiones determinadas de la enzima han sido identificadas para ser de importancia particular para una propiedad determinada de la enzima.

Los métodos alternativos para conseguir variantes de la invención incluyen transposición de genes, por ejemplo, como se describe en WO 95/22625 o en WO 96/00343, y el proceso de derivación de consenso como se describe en EP 897985 (véase la sección “Proteasa progenitora” para más detalles).

En formas de realización particulares, la secuencia de ácidos nucleicos de la invención no es idéntica a: (i) los nucleótidos 900-1466, o 900-1463, de SEC ID n.º 1, los nucleótidos 499-1062 de SEC ID n.º 3, los nucleótidos 496-1059 de SEC ID n.º 5, los nucleótidos 496-1059 de SEC ID n.º 7, y los nucleótidos 502-1065 de SEC ID n.º 9; (ii) los nucleótidos 900-1466 de SEC ID n.º 1; (iii) los nucleótidos 900-1463 de SEC ID n.º 1; (iv) los nucleótidos 900-1463 de SEC ID n.º 1 como se describen en DK 1996 00013; (v) los nucleótidos 499-1062 de SEC ID n.º 3 (vi) los nucleótidos 496-1059 de SEC ID n.º 5; (vii) los nucleótidos 496-1059 de la SEC ID n.º 7; (viii) los nucleótidos 502-1065 de SEC ID n.º 9; (xi) la secuencia de ácidos nucleicos que codifica la parte de péptido maduro de la proteasa derivada de *Nocardia* *dasdonvillei* NRRL 18133; (x) la secuencia de ácidos nucleicos que tiene SEC ID n.º 1 como se describe en JP 2003284571-A; (xi) la secuencia de ácidos nucleicos GENESEQN n.º ADF43563; (xii) la secuencia de ácidos nucleicos descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 1, en particular, la parte de codificación de péptido maduro de la misma; (xiii) la secuencia de ácido nucleico descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º: 3, en particular, la parte de codificación de péptido maduro de la misma; (xiv) la secuencia de ácido nucleico descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º: 5, en particular, la parte de codificación de péptido maduro de la misma; (xv) la secuencia de ácido nucleico descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 7, en particular, la parte de codificación de péptido maduro de la misma; (xvi) la secuencia de ácidos nucleicos descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º 9, en particular, la parte de codificación de péptido maduro de la misma; (xvii) la secuencia de ácido nucleico descrita en la solicitud de patente DK n.º 2004 00969 como SEC ID n.º: 11, en particular, la parte de codificación de péptido maduro de la misma; y/o (xviii) secuencias de ácidos nucleicos que codifican cualquier proteasa de la técnica anterior de al menos 60% de identidad con los aminoácidos 1 a 188 de SEC ID n.º 2.

Constructos de ácidos nucleicos

Un constructo de ácidos nucleicos comprende una secuencia de ácidos nucleicos de la presente invención operativamente vinculado 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. Se entenderá que la expresión incluye cualquier paso relacionado con la producción del polipéptido incluidos, entre otros, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

Vector de expresión

Una secuencia de ácidos nucleicos que codifica una variante de proteasa de la invención puede ser expresada usando un vector de expresión que incluye típicamente secuencias de control que codifican un promotor, operador, sitio de unión al ribosoma, señal de iniciación de traducción, y, opcionalmente, un gen represor o varios genes activadores.

El vector de expresión recombinante que lleva la secuencia de ADN que codifica una variante de proteasa de la invención puede ser cualquier vector que pueda convenientemente ser sometido a procedimientos de ADN recombi-

nante, y la elección del vector dependerá frecuentemente de la célula huésped en la cual será introducido. El vector puede ser uno que, al ser introducido en una célula huésped, se integre en el genoma de la célula huésped y se replique con el/los cromosoma/s en que ha sido integrado.

La variante de proteasa también puede ser coexpresada junto con al menos otra enzima de interés de pienso para animales, tal como una alfa-amilasa, una fitasa, una galactanasa, una xilanas, una endoglucanasa, una endo-1,3(4)-beta- glucanasa, una alfa-galactosidasa y/o una proteasa. Las enzimas se pueden coexpresar a partir 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 sólo un vector, los genes se pueden expresar a partir de uno o más promotores. Si se clona bajo la regulación de un promotor (di o multicistrónico), el orden en el cual se clonan los genes puede afectar los niveles de expresión de las proteínas. La variante de proteasa también puede ser expresada como una proteína de fusión, es decir, que el gen que codifica la variante de proteasa ha sido fusionada en marco al gen que codifica otra proteína. Esta proteína puede ser otra enzima o un dominio funcional de otra enzima.

Células huéspedes

La presente invención también se refiere a células huéspedes recombinantes, que comprenden una secuencia de ácidos nucleicos de la invención, lo cual se usa ventajosamente en la producción recombinante de los polipéptidos. Un vector que comprende 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 que se autoduplica. El término "célula huésped" comprende cualquier progenie de una célula madre que no es idéntica a la célula madre debido a mutaciones que ocurren durante la replicación. La elección de una célula huésped en gran parte depende del gen que codifica el polipéptido y su fuente.

La célula huésped puede ser un microorganismo unicelular, por ejemplo, una procariota, o un microorganismo no unicelular, por ejemplo, una célula eucariota, tal como una célula de animal, mamífero, insecto, planta o fúngica. Las células de animal preferidas son células de animal no humano.

En una forma de realización preferida, la célula huésped es una célula fúngica, o una célula de levadura, tal como una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* o *Yarrowia*. La célula huésped fúngica puede ser una célula fúngica filamentosa, tal como una célula de una especie de *Acremonium*, *Aspergillus*, *Fusarium*, *Hemicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium* o *Trichoderma*, entre otras. Son células unicelulares útiles las células bacterianas tales como bacterias gram positivas incluidas, entre otras, una célula de *Bacillus*, por ejemplo, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus stearothermophilus*, *Bacillus subtilis* y *Bacillus thuringiensis*, o una célula de *Streptomyces*, tal como *Streptomyces lividans* o *Streptomyces murinus*, o una célula de *Nocardia*, o células de bacterias de ácido láctico; o bacterias gram negativas tales como *E. coli* y *Pseudomonas* sp. Las bacterias del ácido láctico incluyen, entre otras, especies de los géneros *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Streptococcus*, *Pediococcus* y *Enterococcus*.

Métodos de producción

La presente invención también se refiere a métodos para producir una variante de proteasa de la presente invención que comprenden (a) cultivo de una célula huésped bajo condiciones propicias para la producción de la variante de proteasa; y (b) recuperación de la variante de proteasa.

En los métodos de producción de la presente invención, las células se cultivan en un medio nutritivo adecuado para la producción del polipéptido usando métodos conocidos en la técnica. Por ejemplo, la célula se puede cultivar por cultivo en matraz de agitación, fermentación a gran escala o pequeña escala (incluyendo fermentaciones por lote alimentado, discontinua, continua o en estado sólido) en fermentadores de laboratorio o industriales realizadas en un medio adecuado y bajo condiciones que permitan al polipéptido ser expresado y/o aislado. El cultivo se desarrolla en un medio nutritivo adecuado que comprende fuentes de nitrógeno y carbono y sales inorgánicas, usando procedimientos conocidos en la técnica. Se encuentran disponibles medios adecuados de proveedores comerciales o se pueden preparar según composiciones publicadas (p. ej., en catálogos de la American Type Culture Collection). Si la proteasa se segrega en el medio nutritivo, puede recuperarse directamente del medio. Si no es segregada, puede recuperarse de lisatos celulares.

La proteasa resultante se puede recuperar por métodos conocidos en la técnica. Por ejemplo, puede ser recuperarse del medio nutritivo por procedimientos convencionales incluidos, entre otros, centrifugado, filtración, extracción, secado por pulverización, evaporación o precipitación.

Las proteasas de la presente invención pueden purificarse mediante una variedad de procedimientos conocidos en la técnica incluidos, entre otros, cromatografía (p. ej., de intercambio iónico, de afinidad, hidrofóbica, de cromatografía de exclusión de tamaño), procedimientos electroforéticos (p. ej., isoelectroenfoque preparatorio), solubilidad diferencial (p. ej., precipitación de sulfato amónico), SDS-PAGE o extracción (véase, por ejemplo, *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

Plantas

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 que tiene actividad de la proteasa de la presente invención para expresar y producir el polipéptido en cantidades recuperables. El polipéptido puede recuperarse de la planta o parte de planta. De manera alternativa, la planta o parte de planta con el polipéptido recombinante se puede utilizar como tal para mejorar calidad de un alimento o pienso, por ejemplo, mejorando el valor nutritivo, la apetencia y las propiedades reológicas, o para destruir un factor antinutritivo.

En una forma de realización particular, el polipéptido es dirigido a las vacuolas de almacenamiento de endospermo en semillas. Este se puede obtener sintetizándolo como un precursor con un péptido señal adecuado; véase Horvath *et al* en PNAS, 15 de feb. de 2000, vol. 97, n.º 4, pág. 1914-1919.

La planta transgénica puede ser dicotiledónea o monocotiledónea o variantes de las mismas creadas genéticamente. Ejemplos de plantas monocotiledóneas son hierbas, tales como poa de prados (poa pratense, Poa), hierba forrajera tal como Festuca, Lolium, césped templado, tal como Agrostis, y cereales, por ejemplo, trigo, avena, centeno, cebada, arroz, sorgo, triticale (híbrido estabilizado de trigo (Triticum) y centeno (Secale) y mazorca (maíz). Ejemplos de plantas dicotiledóneas son tabaco, leguminosas, tales como girasol (helianteo), algodón (Gossipium), altramuces, patata, remolacha azucarera, guisante, moldura y semilla de soja y plantas de crucífero (familia Brassicaceae), tal como coliflor, semilla de colza y el organismo modelo estrechamente relacionado, la *Arabidopsis thaliana*. Las plantas de bajo fitato según se describe, por ejemplo, en la patente estadounidense n.º 5.689.054 y la patente estadounidense n.º 6.111.168, son ejemplos de plantas creadas genéticamente.

Ejemplos de plantas de dicotiledóneas son tabaco, leguminosas, tales como altramuces, patata, remolacha azucarera, guisante, moldura y semilla de soja, y plantas crucíferas (familia Brassicaceae), tales como coliflor, semilla de colza y el organismo modelo estrechamente relacionado, la *Arabidopsis thaliana*. Las plantas de bajo fitato según se describe, por ejemplo, en la patente estadounidense n.º 5.689.054 y la patente estadounidense n.º 6.111.168, son ejemplos de plantas creadas genéticamente. Ejemplos de partes de planta son tallo, callo, hojas, raíz, frutas, semilla y tubérculos, al igual que los tejidos individuales que comprenden estas partes, por ejemplo epidermis, mesófilo, parénquima, tejidos vasculares, meristemas. Los compartimentos de célula vegetal específicos, tales como cloroplasto, apoplasto, mitocondria, vacuola, peroxisomas y citoplasma también se consideran parte de la planta. Además, cualquier célula vegetal, cualquiera que sea el origen del tejido, se considera una parte de la planta. Asimismo, las partes de la planta tales como tejidos específicos y células aisladas para facilitar la utilización de la invención también se consideran partes de la planta, por ejemplo, embriones, endospermas, aleurona y revestimientos de semilla.

También se incluye dentro del campo de la presente invención la progenie de tales plantas, partes de la planta y células vegetales.

La planta transgénica o célula vegetal que expresa un polipéptido de la presente invención se puede construir conforme a métodos conocidos en la técnica. En resumen, la planta o célula vegetal se construye incorporando uno o más constructos de expresión que codifican un polipéptido de la presente invención en el genoma huésped de la planta y propagando la planta o célula vegetal modificada resultante en una planta o célula vegetal transgénica.

Convenientemente, el constructo de expresión es un constructo de ácidos nucleicos que comprende una secuencia de ácidos nucleicos que codifica un polipéptido de la presente invención operativamente enlazado con secuencias reguladoras apropiadas requeridas para la expresión de la secuencia de ácidos nucleicos en la planta o parte de la planta de elección. Además, el constructo de expresión puede comprender un marcador seleccionable útil para identificar células huéspedes en la cuales se ha integrado el constructo de expresión y las secuencias de ADN necesarias para la introducción del constructo en la planta en cuestión (esto último depende del método de introducción de ADN que se debe utilizar).

La elección de secuencias reguladoras, tales como secuencias promotoras y terminadoras y opcionalmente secuencias de tránsito o señal se determina, por ejemplo, basándose en cuándo, dónde y cómo desea expresarse el polipéptido. Por ejemplo, la expresión del gen que codifica un polipéptido de la presente invención puede ser constitutiva o inducible, o puede ser desarrollable, específica de fase o tejido, y el producto genético puede ser dirigido a un tejido o parte de planta específica tal como semillas u hojas. Las secuencias reguladoras son descritas, por ejemplo, por Tague *et al.*, 1988, Plant Physiology 86: 506.

Para la expresión constitutiva, pueden utilizarse los siguientes promotores: el promotor 35S-CaMV (Franck *et al.*, 1980, Cell 21: 285-294), la ubiquitina de maíz 1 (Christensen AH, Sharrock RA and Quail 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation), o el promotor de la actina 1 de arroz (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 de un órgano pueden ser, por ejemplo, un promotor de tejidos sumideros de almacenamiento tales como semillas, tubérculos de patata y frutas (Edwards & Coruzzi, 1990, Ann. Rev. Genet. 24: 275-303), de tejidos sumidero metabólicos tales como meristemas (Ito *et al.*, 1994, Plant Mol. Biol. 24: 863-878), un promotor específico de semilla tal como el promotor de glutelina, prolamina, globulina, o albúmina del 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

desconocido de Vicia faba (Conrad *et al.*, 1998, Journal of Plant Physiology 152: 708-711), un promotor de una proteína estructural de aceite de semilla (Chen *et al.*, 1998, Plant and Cell Physiology 39: 935-941), el promotor napA de una proteína de almacenamiento de *Brassica napus*, u otro promotor cualquiera específico de semillas conocido en la técnica, p. ej., como se describe en WO 91/14772. Además, el promotor puede ser un promotor específico de la hoja, tal como el promotor de rbcS de arroz o tomate (Kyojuka *et al.*, 1993, Plant Physiology 102: 991-1000, the chlorella virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, Plant Molecular Biology 26: 85-93), o el promotor aldP del gen del arroz (Kagaya *et al.*, 1995, Molecular and General Genetics 248: 668-674), un promotor inducible por herida tal 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 tales 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.

Un elemento intensificador del promotor también puede usarse para conseguir una expresión más alta de la enzima en la planta. Por ejemplo, el elemento intensificador del promotor 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*, revelan el uso del primer intrón del gen de actina 1 de arroz para mejorar la expresión.

Aun más, el uso de codón se puede optimizar para las especies de planta en cuestión para mejorar la expresión (véase Horvat *et al* mencionado anteriormente).

El gen de marcador seleccionable y cualquier otra parte del constructo de expresión se puede elegir a partir de los disponibles en la técnica.

El constructo de ácidos nucleicos se incorpora en el genoma de la planta según técnicas convencionales conocidas en la técnica, incluidas 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).

Actualmente, la transferencia de genes mediada por *Agrobacterium tumefaciens* es el método de elección para generar dicotiledóneas transgénicas (para una revisión, véase Hooykas and Schilperoort, 1992, Plant Molecular Biology 19: 15-38), y también puede utilizarse para la transformación de monocotiledóneas, aunque generalmente se prefieren otros métodos de transformación para estas plantas. Actualmente, el método de elección para generar monocotiledóneas transgénicas, complementando el enfoque del Agrobacterium, es el bombardeo de partículas (oro microscópico o partículas de tungsteno revestido con ADN transformante) de callos embrionarios o embriones en desarrollo (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 la transformación de monocotiledóneas se basa en la transformación de protoplasto como describen Omirulleh *et al.*, 1993, Plant Molecular Biology 21: 415-428.

Después de la transformación, se seleccionan los transformantes que tienen incorporado el constructo de expresión y se regeneran en plantas enteras según métodos bien conocidos en la técnica.

La presente invención también se refiere a métodos para producir un polipéptido de la presente invención que comprende (a) cultivar una planta transgénica o una célula vegetal que comprende una secuencia de ácidos nucleicos que codifica una variante de proteasa de la presente invención bajo condiciones propicias para la producción de la variante de proteasa; y (b) recuperar la variante de proteasa.

Animales como huéspedes de expresión

La presente invención también se refiere a un animal transgénico no humano y productos o elementos de los mismos, ejemplos de los cuales son fluidos biológicos tales como leche y sangre, órganos, carne y células animales. Las técnicas para expresar proteínas, por ejemplo en células mamíferas, se conocen en la técnica; véase por ejemplo el manual Protein Expression: A Practical Approach, Higgins and Hames (eds), Oxford University Press (1999) y los otros tres manuales en esta serie acerca de la transcripción de genes, la maduración del ARN y el tratamiento postraduccional. En términos generales, para preparar un animal transgénico, las células seleccionadas de un animal seleccionado se transforman con una secuencia de ácidos nucleicos que codifica una variante de proteasa de la presente invención para expresar y producir la variante de proteasa. La variante de proteasa puede ser recuperada del animal, por ejemplo, de la leche de animales hembra, o puede expresarse para beneficio del animal mismo, por ejemplo, para ayudar a la digestión del animal. A continuación se mencionan ejemplos de animales en la sección titulada Pienso para animales y aditivos para pienso.

Para producir un animal transgénico con el propósito de recuperar la variante de proteasa de la leche del animal, se puede insertar un gen que codifica la variante de proteasa en los ovarios fertilizados de un animal en cuestión, por ejemplo, usando un vector de expresión de transgenes que comprende un promotor de proteínas de la leche adecuado y el gen que codifica la variante de proteasa. El vector de expresión de transgenes se microinyecta en ovarios fertilizados y, preferiblemente, se integra permanentemente en el cromosoma. Una vez que el ovario comienza a crecer y dividirse, el embrión potencial se implanta en una madre sustituta y se identifica a los animales que llevan el transgen. El animal resultante puede después ser multiplicado por reproducción convencional. La variante de proteasa puede purificarse de la leche del animal; véase, 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.

En la alternativa, para producir un animal no humano transgénico que lleva en el genoma de sus células somáticas y/o germinales una secuencia de ácidos nucleicos incluido un constructo de transgen heterólogo que incluye un transgen que codifica la variante de proteasa, el transgen puede ser operativamente vinculado a una primera secuencia reguladora para la expresión específica de la glándula salival de la variante de proteasa, como se describe en WO 2000064247.

Pienso para animales y aditivos para pienso

Para los objetivos presentes, el término animal incluye todos los animales, incluidos los seres humanos. En una forma de realización particular, las variantes de proteasa y las composiciones de la invención se pueden usar como un aditivo de pienso para animales no humanos. Ejemplos de animales son no rumiantes y rumiantes, tales como oveja, cabras, caballos y ganado, por ejemplo, ganado bovino para carne, vacas y terneros jóvenes. En una forma de realización particular, el animal es un animal no rumiante. Los animales no rumiantes incluyen animales monogástricos, por ejemplo, cerdos o puercos (incluidos, entre otros, lechones, cerdos en crecimiento y puercas); aves tales como pavos, patos y pollo (incluidos, entre otros, pollos para asar, ponedoras); terneros jóvenes; y peces (incluidos, entre otros, salmón, trucha, tilapia, siluro y carpas; y crustáceos (incluidos, entre otros, gambas y cigalas).

El término pienso o composición de pienso significa cualquier compuesto, preparación, mezcla, o composición destinada o adecuada para la ingestión por parte de un animal. El pienso se puede administrar al animal antes, después, o simultáneamente con la dieta. Se prefiere la última opción.

La composición de la invención, cuando está destinada a la adición al pienso para animales, se puede designar un aditivo para pienso. Tal aditivo siempre comprende la variante de proteasa en cuestión, preferiblemente, en forma de líquido estabilizado o composiciones secas. El aditivo puede comprender otros componentes o ingredientes de pienso para animales. Las denominadas premezclas para pienso para animales son ejemplos particulares de tales aditivos para pienso. Las premezclas pueden contener la/s enzima/s en cuestión, y en adición al menos una vitamina y/o al menos un mineral.

Por consiguiente, en una forma de realización particular, además de los polipéptidos componentes, la composición de la invención puede comprender o contener al menos una vitamina liposoluble y/o al menos una vitamina hidrosoluble, y/o al menos un oligoelemento. También puede incluirse al menos un macromineral.

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.

Además, ingredientes aditivos para pienso opcionales son los agentes colorantes, por ejemplo, carotenoides tales como betacaroteno, astaxantina, y luteína; compuestos de aroma; estabilizadores; ácidos grasos poliinsaturados; especies generadoras de oxígeno reactivo; péptidos antimicrobianos y/o al menos una enzima adicional.

Componentes enzimáticos adicionales de la invención incluyen, por lo menos, un polipéptido que tiene amilasa, preferiblemente actividad de alfa-amilasa, y/o al menos un polipéptido que tiene actividad de xilanasa; y/o al menos un polipéptido que tiene actividad de endoglucanasa; y/o al menos un polipéptido que tiene actividad de endo-1,3 (4)-beta-glucanasa; y/o al menos un polipéptido que tiene actividad de fitasa; y/o al menos un polipéptido que tiene actividad de galactanasa; y/o al menos un polipéptido que tiene actividad de alfa-galactosidasa; y/o como mínimo otro polipéptido que tiene actividad de la proteasa (EC 3,4.-.-); y/o al menos un polipéptido que tiene actividad de fosfolipasa A1 (EC 3,1,1,32), fosfolipasa A2 (EC 3,1,1,4), lisofosfolipasa (EC 3.1.1.5), fosfolipasa C (EC 3.1.4.3) y/o fosfolipasa D (EC 3.1.4.4).

La actividad de alfa-amilasa puede ser medida como se conoce en la técnica, por ejemplo, usando un sustrato basado en almidón.

La actividad de xilanasa puede medirse usando cualquier ensayo, en el cual se emplee un sustrato, que incluye endoenlaces 1,4-beta-D-xilosídicos en xilanos. Están disponibles diferentes tipos de sustratos para la determinación de la actividad de xilanasa, por ejemplo, comprimidos de arabinoxilano de Xylazyme reticulados (de MegaZyme), o dispersiones de polvo insolubles y soluciones de arabinoxilano coloreado con azo.

La actividad de endoglucanasa puede determinarse usando cualquier ensayo de endoglucanasa conocido en la técnica. Por ejemplo, pueden aplicarse diferentes sustratos con betaglucano o de celulosa. Un ensayo de endoglucanasa

puede usar beta-glucano de AZCL-cebada o, preferiblemente (1) AZCL-HE-celulosa, o (2) Azo-CM-celulosa como un sustrato. En ambos casos, la degradación del sustrato es seguida espectrofotométricamente a OD₅₉₅ (véase el método Megazyme para AZCL-polisacáridos para el ensayo de endohidrolasas en <http://www.megazyme.com/booklets/AZCLPOL.pdf>).

La actividad de endo-1,3(4)-beta-glucanasa puede determinarse usando cualquier ensayo de endo-1,3(4)-beta-glucanasa conocido la técnica. Un sustrato preferido para las mediciones de actividad de endo-1,3(4)-beta-glucanasa es un sustrato de cebada de betaglucano reticulado azo-coloreado, donde las mediciones se basan en principios de determinación espectrofotométricos.

La actividad de la fitasa puede medirse usando cualquier ensayo adecuado, por ejemplo, el ensayo FYT descrito en el ejemplo 4 de WO 98/28408.

La galactanasa se puede evaluar, por ejemplo, con galactano de AZCL de Megazyme, y alfagalactosidasa se puede evaluar, por ejemplo, con pNP-alfa-galactósido.

Para evaluar estas actividades enzimáticas, deben adaptarse el pH y la temperatura del ensayo a la enzima en cuestión (preferiblemente, un pH cercano al pH óptimo y una temperatura cercana a la temperatura óptima). Un pH de ensayo preferido está en la gama de 2-10, preferiblemente 3-9, más preferiblemente pH 3 o 4 o 5 o 6 o 7 u 8, por ejemplo, pH 3 o pH 7. Una temperatura de ensayo preferida está en la gama de 20-90°C, preferiblemente 30-90°C, más preferiblemente 40-80°C, incluso más preferiblemente 40-70°C, preferiblemente 40 o 45 o 50°C. La actividad enzimática se define por referencia a ciegos apropiados, por ejemplo, un tampón ciego.

Ejemplos de péptidos antimicrobianos (AMP) son CAP18, leucocina A, tritricina, protegrina1, tanatina, defensina, lactoferrina, lactoferricina y ovispirina tal como novispirina (Robert Lehren, 2000), plectasinas y estatinas, incluidos los compuestos y polipéptidos descritos en WO 03/044049 y WO 03/048148, al igual que variantes o fragmentos de los anteriores que retienen actividad antimicrobiana.

Ejemplos de polipéptidos antifúngicos (AFP) son los péptidos de *Aspergillus giganteus* y *Aspergillus niger*, al igual que variantes y fragmentos de los mismos que retienen actividad antifúngica, según se describen en WO 94/01459 y WO 02/090384.

Ejemplos de ácidos grasos poliinsaturados son ácidos grasos poliinsaturado C18; C20 y C22, tales como ácido araquidónico, ácido docosohexaenoico, ácido eicosapentanoico y ácido gamma-linolénico.

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

Usualmente, las vitaminas liposolubles e hidrosolubles, al igual que los oligoelementos, forman parte de una denominada premezcla destinada a la adición al pienso, mientras que los macrominerales normalmente se adicionan de manera separada al pienso. Una premezcla enriquecida con una proteasa de la invención es un ejemplo de un aditivo de pienso de la invención.

En una forma de realización particular, el aditivo de pienso de la invención está destinado a ser incluido (o se prescribe que tiene que ser incluido) en pienso o dietas para animales a niveles de 0,01 a 10,0%; más particularmente 0,05 a 5,0%; o 0,2 a 1,0% (% significa g de aditivo por 100 g de pienso). Esto es así en particular para premezclas.

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

En la alternativa, el aditivo de pienso de la invención comprende al menos uno de los componentes individuales especificados en la tabla A de WO 01/58275. Al menos uno significa uno o más de uno, o dos, o tres, o cuatro, etcétera hasta trece, o hasta quince componentes individuales. Más específicamente, este al menos un componente individual se incluye en el aditivo de la invención en cantidad tal como para proporcionar una concentración en pienso en la gama indicada en la columna cuatro, o la columna cinco, o la columna seis de la tabla A.

La presente invención también se refiere a composiciones de pienso para animales. Las composiciones de pienso o dietas para animales tienen un contenido relativamente alto de proteína. Las dietas de ave y cerdo se pueden caracterizar según se indica en la tabla B de WO 01/58275, columnas 2-3. Las dietas de peces se pueden caracterizar según se indica en la columna 4 de esta tabla B. Además, tales dietas de peces normalmente tienen un contenido de grasa cruda de 200-310 g/kg. WO 01/58275 corresponde a US 09/779334 que se incorpora a la presente a modo de referencia.

Una composición de pienso para animales según la invención tiene un contenido de proteína bruta de 50-800 g/kg, y, además, comprende al menos una variante de proteasa como se reivindica aquí.

Además, o en la alternativa (al contenido bruto de proteína indicado más arriba), la composición de pienso para animales de la invención tiene un contenido de energía metabolizable de 10-30 MJ/kg; y/o un contenido de calcio de

ES 2 358 092 T3

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.

En formas de realización particulares, el contenido de energía metabolizable, proteína cruda, calcio, fósforo, metionina, metionina más cisteína y/o lisina está dentro de cualquiera de las gamas 2, 3, 4 o 5 en la tabla B de WO 01/58275 (R. 2-5).

La 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 (Official Methods of Analysis 14th ed., Association of Official Analytical Chemists, Washington DC).

La energía metabolizable puede calcularse basándose en la publicación de NRC Nutrient requirements in swine, novena edición corregida 1988, subcomisión para la nutrición de puercos, comité sobre nutrición animal, junta de agricultura, consejo de investigación nacional. National Academy Press, Washington, D.C., pág. 2-6, y la tabla europea de valores de energía para ingredientes de pienso de ave, centro Spelderholt para investigación y extensión de aves, 7361 DA Beekbergen, Países Bajos. Grafisch bedrijf Ponsen & looijen bv, Wageningen. ISBN 90-71463-12-5.

El contenido dietético de calcio, fósforo y aminoácidos disponibles en dietas para animales completas es calculado basándose en tablas de pienso tales como Veevoedertabel 1997, gegevens over chemische samenstelling, verteerbaarheid in voederwaarde van voedermiddelen, Central Veevoederbureau, Runderweg 6, 8219 pk Lelystad. ISBN 90-72839-13-7.

En una forma de realización particular, la composición de pienso para animales de la invención contiene al menos una proteína. La proteína puede ser una proteína animal, tal como harina de carne y hueso, y/o harina de pescado; o, en una forma de realización particular, puede ser una proteína vegetal. El término proteínas vegetales según se utilizan en este caso se refiere a cualquier compuesto, composición, preparación o mezcla que incluye, al menos, una proteína derivada de u originada en un vegetal, incluidas las 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).

Las proteínas vegetales se pueden derivar de fuentes de proteína vegetal, tal como leguminosas y cereales, por ejemplo, materiales de plantas de las familias Fabaceae (Leguminosae), Cruciferaeae, Chenopodiaceae y Poaceae, tal como harina de soja, harina de lupino y harina de semilla de colza.

En una forma de realización particular, la fuente de proteína vegetal es material de una o más plantas de la familia Fabaceae, por ejemplo, la semilla de soja, altramuza, guisante o judía.

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 quinoa.

Otros ejemplos de fuentes de proteína vegetal son semilla de colza, semilla de girasol, semilla de algodón y repollo.

La semilla de soja es una fuente de proteína vegetal preferida.

Otros ejemplos de fuentes de proteína vegetal son cereales tales como cebada, trigo, centeno, avena, maíz (mazorca), arroz, triticale y sorgo.

En otras formas de realización particulares, la composición de pienso para animales de la invención contiene 0-80% maíz; y/o 0-80% sorgo; y/o 0-70% trigo; y/o 0-70% cebada; y/o 0-30% avena; y/o 0-40% harina de soja; y/o 0-25%, preferiblemente 0-10%, harina de pescado; 0-25% harina de carne y hueso; y/o 0-20% lactosuero.

Las dietas para animales pueden, por ejemplo, ser fabricadas como pienso triturado (no granulado) o pienso granulado. Típicamente, los ingredientes de pienso molidos y cantidades suficientes y de vitaminas y minerales esenciales se mezclan y se agregan según las especificaciones para las especies en cuestión. Las enzimas se pueden adicionar como formulaciones de enzima sólidas o líquidas. Por ejemplo, una formulación de enzima sólida típicamente es adicionada antes o durante el paso de mezcla; y una preparación enzimática líquida típicamente es adicionada después del paso de granulación. La enzima también puede ser incorporada en un aditivo o premezcla de pienso.

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

La variante de proteasa debería, por supuesto, ser aplicada en una cantidad eficaz, es decir, en una cantidad adecuada para mejorar la solubilización y/o mejorar el valor nutritivo de pienso. Actualmente se contempla que la enzima se administre en una o más de las siguientes cantidades (gamas de dosificación): 0,01-200, 0,01-100, 0,5-100, 1-50, 5-100, 10-100, 0,05-50 o 0,10-10. Todas estas gamas son en mg de proteína de enzima proteásica por kg (ppm) de pienso.

Para determinar los mg de proteína enzimática por kg de pienso, la proteasa se purifica de la composición de pienso, y la actividad específica de la proteasa purificada es determinada usando un ensayo pertinente (véase debajo

de actividad de la proteasa, sustratos y ensayos). La actividad de la proteasa de la composición de pienso como tal también es determinada usando el mismo ensayo, y basándose en estas dos determinaciones, se calcula la dosificación en mg de proteína enzimática por kg de pienso.

Los mismos principios se aplican para determinar los mg de proteína enzimática en aditivos de pienso. Por supuesto, si se encuentra disponible una muestra de la proteasa usada para preparar el aditivo de pienso o el pienso, la actividad específica se determina a partir de esta muestra (sin necesidad de purificar la proteasa de la composición de pienso ni el aditivo).

Composiciones de detergentes

La variante de proteasa de la invención se puede adicionar a, y así convertirse en, un componente de una composición de detergente.

La composición de detergente de la invención puede, por ejemplo, ser formulada como una composición de detergente a mano o a máquina de lavandería incluida una composición de aditivo de lavandería adecuada para pretratamiento de tejidos manchados y una composición de suavizante adicionado de enjuague, o ser formulado como una composición de detergente para el uso en operaciones de limpieza de superficies duras del hogar en general, o ser formulado para operaciones de lavado de la vajilla a mano o a máquina.

En un aspecto específico, la invención proporciona un aditivo de detergente que comprende la variante de proteasa de la invención. El aditivo de detergente, al igual que la composición de detergente, puede comprender una o más enzimas adicionales tal como otra proteasa, tal como proteasas alcalinas de *Bacillus*, una lipasa, una cutinasa, una amilasa, una carbohidrasa, una celulasa, una pectinasa, una mananasa, una arabinasa, una galactanasa, una xilanas, una oxidasa, por ejemplo, una lacasa, y/o una peroxidasa.

En general, las propiedades de la/s enzima/s elegidas debería ser compatible con el detergente seleccionado, (es decir, pH óptimo, compatibilidad con otros ingredientes no enzimáticos y enzimáticos, etc.), y la/s enzima/s deberían estar presentes en cantidades eficaces.

Las lipasas adecuadas incluyen aquellas de origen fúngico o bacteriano. Los mutantes químicamente modificados o creados genéticamente de proteína están incluidos. Ejemplos de lipasas útiles incluyen lipasas de *Humicola* (sinónimo *Thermomyces*), por ejemplo de *H. lanuginosa* (*T. Lanuginosus*) como se describe en EP 258068 y EP 305216 o de *H. insolens* como se describe en WO 96/13580, una lipasa de *Pseudomonas*, por ejemplo de *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 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. Las enzimas de lipasa preferidas disponibles comercialmente incluyen LipolaseTM y Lipolase UltraTM (Novozymes A/S).

Las amilasas adecuadas (beta y/o alfa) incluyen aquellas de origen fúngico o bacteriano. Los mutantes químicamente modificados o creados genéticamente de proteína están incluidos. Las amilasas incluyen, por ejemplo, alfa-amilasas obtenidas de *Bacillus*, por ejemplo, una cepa especial de *B. licheniformis*, descrita en más detalle en GB 1.296.839. Ejemplos de amilasas útiles son las variantes descritas en 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, 105, 106, 124, 128, 133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408, y 444. Ailasas disponibles comercialmente son NatalaseTM, SupramylTM, StainzymeTM, DuramylTM, TermamylTM, FungamylTM y BANTM (Novozymes A/S), RapidaseTM y PurastarTM (de Genencor International Inc.).

Las celulosas adecuadas incluyen aquellas de origen fúngico o bacteriano. Los mutantes químicamente modificados o creados genéticamente de proteína están incluidos. Las celulosas adecuadas incluyen celulosas de los géneros *Bacillus*, *Pseudomonas*, *Humicola*, *Fusarium*, *Thielavia*, *Acremonium*, por ejemplo, las celulosas fúngicas producidas de *Humicola insolens*, *Myceliophthora thermophila* y *Fusarium oxysporum* descritos en US 4.435.307, US 5.648.263, US 5.691.178, US 5.776.757 y WO 89/09259. Celulosas especialmente adecuadas son las celulosas neutras o alcalinas que tienen beneficios de cuidados del color. Ejemplos de tales celulosas son las celulosas descritas en EP 0 495257, EP 531372, WO 96/11262, WO 96/29397, WO 98/08940. Otros ejemplos son variantes de celulasa tales como las descritas en 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. Las celulosas disponibles comercialmente incluyen CelluzymeTM y CarezymeTM (Novozymes A/S), ClazinasTM y Puradax HATM (Genencor International Inc.), y KAC-500(B)TM (Kao Corporation).

Peroxidasas/oxidases adecuadas incluyen aquellas de origen fúngico, bacteriano o de planta. Los mutantes químicamente modificados o creados genéticamente de proteína están incluidos. Ejemplos de peroxidasas útiles incluyen peroxidasas de *Coprinus*, por ejemplo, de *C. cinereus*, y variantes de las mismas como las descritas en WO 93/24618, WO 95/10602 y WO 98/15257. Peroxidasas disponibles comercialmente incluyen GuardzymeTM (Novozymes).

ES 2 358 092 T3

La/s enzima/s detergente/s se puede/n incluir en una composición de detergente añadiendo aditivos separados que contengan una o más enzimas, o añadiendo un aditivo combinado que comprenda todas estas enzimas. Un aditivo de detergente de la invención, es decir, un aditivo separado o un aditivo combinado, se puede formular, por ejemplo, como un granulado, un líquido, un compuesto acuoso, etc. Las formulaciones de aditivo de detergente preferidas son
5 granulados, en particular, granulados no polvorientos, líquidos, en particular, líquidos estabilizados o mezclas.

Los granulados no polvorientos pueden producirse, por ejemplo, como se describe en US 4.106.991 y 4.661.452 y opcionalmente pueden ser revestidos por métodos conocidos en la técnica. Ejemplos de materiales de recubrimiento cerosos son productos de poli(etileno óxido) (polietilenglicol; PEG) con pesos molares medios de 1000 a 20000,
10 nonilfenoles etoxilados que tienen de 16 a 50 unidades de óxido de etileno; alcoholes grasos etoxilados en los cuales el alcohol contiene de 12 a 20 átomos de carbono y en el cual hay 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 adecuados para la aplicación por técnicas de lecho fluidificado se dan en GB 1483591. Las preparaciones enzimáticas líquidas pueden, por ejemplo, ser estabilizadas añadiendo un poliol tal como propilenglicol, un azúcar o
15 alcohol de azúcar, ácido láctico o ácido bórico según métodos establecidos. Las enzimas protegidas se pueden preparar según el método descrito en EP 238216.

La composición de 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, y típicamente
20 contiene hasta 70% de agua y 0-30% de solvente orgánico, o no acuoso.

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

Cuando está incluido, el detergente normalmente contiene de aproximadamente 1% a aproximadamente 40% de un tensioactivo aniónico tal como alquilbencenosulfonato lineal, alfa-olefinsulfonato, alquil sulfato (sulfato de alcohol graso), alcohol etoxisulfato, alcanosulfonato secundario, éster metílico de ácido graso alfa-sulfo, ácido alquenilsuccínico o de alquilo o jabón.
30

Cuando está incluido, el detergente normalmente contiene de aproximadamente 0,2% a aproximadamente 40% de un tensioactivo no iónico tal como alcohol etoxilato, nonilfenol, etoxilato alquilpoliglicósido, óxido de alquildimetilamina, monoetanolamida de ácido graso etoxilado, monoetanolamida de ácido graso, polihidroxi alquil amida de ácido graso, o derivados N-acilo N-alquilo de glucosamina ("glucamidas").
35

El detergente puede contener 0-65% de un constructor de detergente o agente complejante tal como zeolita, difosfato, trifosfato, fosfonato, carbonato, citrato, ácido nitrilotriacético, ácido etilenodiaminatetraacético, ácido dietileno-triaminopentaacético, ácido alquenilsuccínico o de alquilo, silicatos solubles o silicatos estratificados (p. ej. SKS-6 de Hoechst).
40

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

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 tetraacetiletilenodiamina o nonanoiloxibencenosulfonato. Alternativamente, el sistema blanqueante puede comprender peroxiácidos de, por ejemplo, el tipo amida, imida o sulfona.
45

La/s enzima/s de la composición de detergente de la invención puede/n ser estabilizada/s usando agentes estabilizantes convencionales, 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 ácido 4-formilfenil borónico y la composición se puede formular como se describe en, por ejemplo, WO 92/19709 y WO 92/19708.
50

El detergente también puede contener otros ingredientes de detergentes convencionales tales como, por ejemplo, acondicionadores de tejido incluidas 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.
55

Se contempla actualmente que en las composiciones detergentes cualquier enzima, en particular la enzima de la invención, se pueda adicionar en una cantidad correspondiente a 0,01-100 mg de proteína enzimática por litro de solución de lavado, preferiblemente 0,05-5 mg de proteína enzimática por litro de solución de lavado, en particular 0,1-1 mg de proteína enzimática por litro de solución de lavado.
60

La enzima de la invención puede adicionalmente ser incorporada en las formulaciones de detergente descritas en WO 97/07202.
65

Método para generar variantes de proteasa

La invención también se refiere a un método para generar una variante de proteasa de una propiedad mejorada. El método comprende los siguientes pasos:

(a) seleccionar una proteasa progenitora de al menos 60% de identidad con los aminoácidos 1 a 188 de SEC ID n.º 2;

(b) establecer una estructura tridimensional de la proteasa progenitora por modelado por homología usando la estructura de la figura 2 como un modelo; y/o alineando la proteasa progenitora según la alineación de la figura 1;

(c) proponiendo al menos una sustitución de aminoácido, por ejemplo:

(i) sometiendo la estructura tridimensional de (b) a simulaciones de DM a temperaturas aumentadas e identificando regiones en la secuencia de aminoácidos de la proteasa progenitora de alta movilidad (fluctuaciones isotrópicas);

(ii) introduciendo puentes disulfuros a modo de sustituciones de cisteína (C-C);

(iii) introduciendo sustituciones de prolina (P);

(iv) sustituyendo residuos de aminoácidos neutros expuestos con residuos de aminoácidos cargados negativamente (E,D);

(v) sustituyendo residuos de aminoácidos neutros expuestos con residuos de aminoácidos cargados positivamente (R,K);

(vi) sustituyendo residuos de aminoácidos pequeños dentro de la proteína con residuos de aminoácidos más voluminosos (W);

(vii) comparando por alineación por homología y/o modelado por homología según el paso (c)(i) al menos dos proteasas progenitoras relacionadas y transfiriendo las diferencias de residuo de aminoácido entre estas estructuras de proteasa, preferiblemente, de una estructura que tenga la propiedad mejorada a una estructura que no tenga esta propiedad mejorada;

(d) preparando una secuencia de ADN que codifica la proteasa progenitora excepto la inclusión de un codón de ADN de al menos una sustitución de aminoácido propuesta en los pasos (c)(ii)-(c)(vii), o sometiendo a la secuencia de ADN progenitora a mutagénesis aleatoria, dirigiendo al menos a una de las regiones identificadas en el paso (c)(i);

(e) expresando la secuencia de ADN obtenida en el paso (d) en una célula huésped, y

(h) seleccionando una célula huésped que exprese una variante de proteasa con una propiedad mejorada.

La invención además se refiere a un método para la producción de una variante de proteasa obtenible u obtenida por el método de generación de variantes de proteasa anteriormente descrito, que comprende (a) cultivar la célula huésped para producir un sobrenadante que comprenda la variante; y (b) recuperar la variante.

La invención también se refiere a secuencias de ácidos nucleicos aisladas que comprenden una secuencia de ácidos nucleicos que codifica la variante de proteasa obtenible según este método, al igual que métodos para producirla (a) cultivando la célula huésped para producir un sobrenadante que comprenda la variante; y (b) recuperando la variante; una planta transgénica, o parte de planta, capaz de expresarla, animales transgénicos no humanos, o productos, o elementos de los mismos, capaces de expresarla; piensos para animales, al igual que aditivos para pienso, que la comprendan; métodos para mejorar el valor nutritivo de un pienso para animales por uso de la misma; métodos para el tratamiento de proteínas, tales como proteínas vegetales, por uso de la misma; así como el uso de la misma (i) en pienso para animales; (ii) en la preparación de pienso para animales; (iii) para mejorar el valor nutritivo de pienso para animales; y/o (iv) para el tratamiento de proteínas; y/o en detergentes.

Forma de realización alternativa

En una forma de realización alternativa, se usa el término “alteración” en vez de “sustitución” como el término general para enmiendas en la molécula de proteasa. Esta forma de realización alternativa incluye cada una de las reivindicaciones formuladas como se ejemplifica a continuación para la reivindicación 1, y también específicamente incluye todo lo que se expresa en la presente, por ejemplo, las definiciones (aparte de la definición de sustitución), es decir, los diferentes aspectos, formas de realización particulares, etc.

ES 2 358 092 T3

Una variante de una proteasa progenitora, que comprende una alteración en al menos una posición de al menos una región seleccionada del grupo de regiones que consisten en:

6-18; 22-28; 32-39; 42-58; 62-63; 66-76; 78-100; 103-106; 111-114; 118-131; 134-136; 139-141; 144-151; 155-156; 160-176; 179-181; y 184-188; donde

(a) la o las alteraciones son independientemente

(i) una inserción de un aminoácido inmediatamente debajo de la posición,

(ii) una delección del aminoácido que ocupa la posición, y/o

(iii) una sustitución del aminoácido que ocupa la posición;

(b) la variante tiene actividad de la proteasa; y

(c) cada posición corresponde a una posición de SEC ID n.º 2, preferiblemente, los aminoácidos 1 a 188 de las mismas; y

(d) la variante tiene un porcentaje de identidad con SEC ID n.º 2, preferiblemente con los aminoácidos 1 a 188 de la misma, de al menos 60%.

El término, “variante de polipéptido”, “variante de proteína”, “variante de enzima” “variante de proteasa” o simplemente “variante” se refiere a un polipéptido de la invención que comprende una o más alteraciones, tales como sustitución(es), inserción(es), delección(es) y/o truncamientos de uno o más residuos de aminoácido específico en una o más posiciones específicas en el polipéptido.

El término, “polipéptido progenitor”, “proteína progenitora”, “enzima progenitora”, “enzima estándar”, “proteasa progenitora” o simplemente “progenitor/a” se refiere al polipéptido sobre el que se basó la variante. Este término también se refiere al polipéptido con el cual se compara y se alinea una variante.

El término, “biblioteca aleatorizada”, “biblioteca variante”, o simplemente “biblioteca” se refiere a una biblioteca de polipéptidos variantes. La diversidad en la biblioteca variante puede generarse por medio de mutagénesis de los genes que codifican las variantes en el nivel de tripletes de ADN, de manera que los codones individuales se diversifiquen, por ejemplo, usando cebadores de secuencia parcialmente aleatorizados en una reacción de PCR. Se han descrito diferentes técnicas mediante las cuales se puede crear una biblioteca combinatoria diversa diversificando diferentes posiciones de nucleótido en un gen y recombinándolos, por ejemplo donde estas posiciones están demasiado separadas para ser cubiertas por un único cebador oligonucleótido (dopado o adicionado). Estas técnicas incluyen el uso de recombinación *in vivo* de los segmentos de gen individuales diversificados como se describe en WO 97/07205 en la página 3, líneas 8 a 29 (Novozymes A/S). También incluyen el uso de técnicas de redistribución de ADN para crear una librería de genes de longitud total, donde se combinan diferentes segmentos de gen, y donde cada segmento se puede diversificar, por ejemplo, por mutagénesis adicionada, (Stemmer, Nature 370, pág. 389-391, 1994 y US 5.811.238; US 5.605.793; y US 5.830.721). Se puede usar un gen que codifica una “estructura” de proteína (polipéptido progenitor de tipo salvaje) como un polinucleótido modelo, y combinarlo con uno o más oligonucleótidos monocatenarios o bicatenarios como se describe en WO 98/41623 y en WO 98/41622 (Novozymes A/S). Los oligonucleótidos monocatenarios podrían ser parcialmente aleatorizados durante la síntesis. Los oligonucleótidos bicatenarios podrían ser productos de PCR que incorporan diversidad en una zona específica. En ambos casos, se puede diluir la diversidad con los segmentos correspondientes que codifican la secuencia de la proteína de esqueleto para limitar el número medio de cambios que se introducen.

Los métodos también han sido establecidos para diseñar de las proporciones de mezclas de nucleótido (A; C; T; G) para ser insertadas en posiciones de codón específicas durante la síntesis de polinucleótidos o de oligonucleótidos, para introducir un sesgo para aproximar una distribución de frecuencias deseada hacia un conjunto de uno o más aminoácidos deseados que serán codificados por los codones particulares. Puede ser de interés producir una biblioteca variante que comprenda permutaciones de varias modificaciones de aminoácido conocidas en diferentes lugares en la secuencia primaria del polipéptido. Éstas podrían introducirse post traduccionalmente o por sitios de modificación químicos, o podrían introducirse a través de mutaciones en los genes de codificación. Las modificaciones solas pueden haber demostrado ser provechosas previamente por una razón u otra (p. ej. disminución de la antigenicidad, o mejora de la actividad específica, el rendimiento, la estabilidad u otras características). En tales ejemplos, puede ser conveniente primero crear una biblioteca de combinaciones diversas de secuencias conocidas. Por ejemplo, si doce mutaciones individuales son conocidas, una podría combinar (como mínimo) doce segmentos del gen codificante de la proteína progenitora, donde cada segmento está presente en dos formas: una con, y una sin, la mutación deseada. Al variar las cantidades relativas de esos segmentos, podría diseñarse una biblioteca (de tamaño 212) para la cual el número medio de mutaciones por gen puede ser predicho. Esta puede ser una manera útil de combinar mutaciones, que en sí mismas dan algún efecto, pero no suficiente, sin recurrir a bibliotecas muy grandes, como es frecuentemente el caso cuando se usa “mutagénesis adicionada”. Otro modo de combinar estas “mutaciones conocidas” podría ser usando redistribución de familias de ADN oligomérico que codifican las mutaciones conocidas con fragmentos de la secuencia de tipo salvaje de longitud total.

ES 2 358 092 T3

Al describir las diferentes variantes contempladas o producidas según la invención, se utilizan varias nomenclaturas y convenios que están descritos en detalle más abajo. Un marco de referencia se define primero alineando el polipéptido variante con una enzima progenitora. Una enzima progenitora preferida es la proteasa 10 (aminoácidos 1 a 188 de SEC ID n.º 2). De ese modo, varias alteraciones serán definidas en relación a la secuencia de aminoácidos de los aminoácidos 1 a 188 de SEC ID n.º 2.

Una sustitución en una variante es indicada como:

Aminoácido original - posición - aminoácido sustituido;

Se utilizan los códigos de tres o una letra, incluidos los códigos Xaa y X para indicar cualquier residuo de aminoácido. Por consiguiente, la notación "T82S" o "Thr82Ser" significa que la variante comprende una sustitución de treonina con serina en la posición de aminoácido variante correspondiente al aminoácido en la posición 82 en la enzima progenitora, cuando los dos se alinean como se ha indicado anteriormente.

Donde el residuo de aminoácido original puede ser cualquier residuo de aminoácido, una notación rápida puede usarse a veces indicando sólo la posición y el aminoácido sustituido, por ejemplo:

Posición - aminoácido sustituido; o "82S".

Tal notación es particularmente pertinente en relación con las modificaciones en una serie de polipéptidos homólogos.

De forma similar, cuando la identidad del residuo o los residuos de aminoácido de sustitución es irrelevante:

Aminoácido original - posición; o "T82"

Cuando el/los aminoácido/s originales y el/los aminoácido/s sustituidos puede/n ser cualquier aminoácido, sólo se indica la posición, por ejemplo: "82".

Cuando el/los aminoácido/s originales y/o aminoácido/s sustituidos pueden comprender más de uno, pero no todos, los aminoácido/s, los aminoácidos son catalogados separados por comillas:

Aminoácidos originales - n.º de posición - aminoácidos sustituidos; o "T10E,D,Y".

A continuación se presentan varios ejemplos de esta nomenclatura:

La sustitución de treonina para histidina en la posición 91 es designada: "His91Thr" o "H91T"; o la sustitución de cualquier ácido de residuo de aminoácido para histidina en la posición 91 es designada: "His91Xaa" o "H91X" o "His91" o "H91".

Para una modificación donde el/los aminoácido/s originales y/o aminoácido/s sustituidos puede comprender más de uno, pero no todos los aminoácido/s, la sustitución de ácido glutámico, ácido aspártico o tirosina para treonina en la posición 10:

"Thr10Glu,Asp,Tyr" o "T10E,D,Y"; que indica las variantes específicas: "T10E", "T10D" y "T10Y".

Una delección de glicina en la posición 26 estará indicada por: "Gly26*" o "G26*".

De manera correspondiente, la delección de uno o más residuos de aminoácido, tal como la delección de glicina y glutamina en las posiciones 26 y 27 será designada "Gly26*+Gln27*" o "G26*+Q27*".

La inserción de un residuo de aminoácido adicional tal como, por ejemplo, una lisina después de G26 es indicada por: "Gly26GlyLys" o "G26GK"; o, cuando más de un residuo de aminoácido es insertado, tal como, por ejemplo, Lys y Ala después de G26, esto se indicará como: "Gly26GlyLysAla" o "G26GKA".

En tales casos, el/los residuo/s de aminoácido insertado/s se numeran por medio de la adición de letras minúsculas al número de posición del residuo de aminoácido que precede el/los residuo/s de aminoácido insertado/s. En el ejemplo anterior, las secuencias serían entonces:

Progenitor:	Variante:		
26	26	26a	26b
G	G	K	A

ES 2 358 092 T3

En casos en los cuales se inserta un residuo de aminoácido idéntico al residuo de aminoácido existente, es claro que surge la degeneración en la nomenclatura. Si, por ejemplo, va a insertarse una glicina después de la glicina en el ejemplo anterior, esto sería indicado por medio de "G26GG".

- 5 Suponiendo que una alanina estuviera presente en la posición 25, el mismo cambio real bien podría indicarse también como "A25AG":

		Progenitor:		Variante:		
10	Numeración I:	25	26	25	26	26a
	Secuencia:	A	G	A	G	G
	Numeración II:			25	25a	26

15

Tales ejemplos serán claros para el experto en la materia, y la indicación "G26GG" y las indicaciones correspondientes para este tipo de inserciones están de este modo destinadas a comprender tales indicaciones de degeneración equivalentes.

20

Por analogía, si los segmentos de secuencia de aminoácidos se repiten en el polipéptido progenitor y/o en la variante, será claro para el experto en la materia que están comprendidas las indicaciones de degeneración equivalentes, también cuando otras alteraciones aparte de las inserciones sean catalogadas tales como delecciones y/o sustituciones. Por ejemplo, la delección de dos aminoácidos consecutivos "AG" en la secuencia "AGAG" de la posición 194-197, puede escribirse como "A196*+G197*" o "A194*+G1956*":

25

		Progenitor:				Variante:	
	Numeración I:	194	195	196	197	194	195
30	Secuencia:	A	G	A	G	A	G
	Numeración II:					196	197

35

Las variantes que comprenden modificaciones múltiples se separan por signos de suma, por ejemplo: "Arg170Tyr+Gly195Glu" o "R170Y+G195E", que representan modificaciones en las posiciones 170 y 195 que sustituyen tirosina y ácido glutámico por arginina y glicina, respectivamente. Por tanto, "Tyr167Gly,Ala,Ser,Thr+Arg170Gly,Ala,Ser,Thr" designa las siguientes variantes: "Tyr167Gly+Arg170Gly", "Tyr167Gly+Arg170Ala", "Tyr167Gly+Arg170Ser", "Tyr167Gly+Arg170Thr", "Tyr167Ala+Arg170Gly", "Tyr167Ala+Arg170Ala", "Tyr167Ala+Arg170Ser", "Tyr167Ala+Arg170Thr", "Tyr167Ser+Arg170Gly", "Tyr167Ser+Arg170Ala", "Tyr167Ser+Arg170Ser", "Tyr167Ser+Arg170Thr", "Tyr167Thr+Arg170Gly", "Tyr167Thr+Arg170Ala", "Tyr167Thr+Arg170Ser" y "Tyr167Thr+Arg170Thr".

40

Esta nomenclatura es particularmente pertinente en relación a modificaciones destinadas a sustitución, inserción o delección de residuos de aminoácidos que tienen propiedades específicas comunes. Tales modificaciones se denominan modificación o modificaciones de aminoácido conservadora/s.

45

Distintas formas de realización

Estas son distintas formas de realización adicionales de la invención:

50

La variante de cualquiera de las reivindicaciones 1-16 y 18-20 que comprende al menos una de las siguientes sustituciones: T10Y, A24S, V51T, E53Q, T82S, A86Q, T87S, I96A, G118N, S122R, N130S, L186I.

55

La variante de cualquiera de las reivindicaciones 1-16 y 18-19 que comprende al menos una de las siguientes sustituciones: R38T; Q42G,P; R49T,Q; Q54N,R; A89S,T; H91S,T; N92S; S99A,Q; A120T; E125Q; T129Y,Q; M131L; T135N; Y147F; N151S; R165S; T166V,F; F171Y; V179L,L; preferiblemente, al menos una de las siguientes sustituciones: R38T; N92S; A120T; E125Q; M131 L; T135N; Y147F; N151S; R165S y/o F171Y.

60

La variante de cualquiera de las reivindicaciones 1-19 que comprende al menos una de las siguientes sustituciones: A25S, T44S, A62S, P95A, V100I, I114V, T176N, N180S, V184L, R185T.

La variante de cualquiera de las reivindicaciones 1-20 que tiene propiedades enmendadas, tal como una termostabilidad mejorada y/o una temperatura óptima más alta o más baja, tal como una Tm de al menos 83,1°C según se mide por calorimetría por análisis diferencial en 10 mM fosfato sódico, 50 mM cloruro sódico, pH 7,0.

65

La variante de cualquiera de las reivindicaciones 1-20 que deriva de una cepa del género *Nocardiopsis*, tal como *Nocardiopsis alba*, *Nocardiopsis antarctica*, *Nocardiopsis prasina*, *Nocardiopsis composta*, *Nocardiopsis dassonvillei*, *Nocardiopsis exhalans*, *Nocardiopsis halophila*, *Nocardiopsis halotolerans*, *Nocardiopsis kunsanensis*, *Nocardiopsis*

ES 2 358 092 T3

listeri, *Nocardiopsis lucentensis*, *Nocardiopsis metallicus*, *Nocardiopsis* sp., *Nocardiopsis synnemataformans*, *Nocardiopsis trehalosi*, *Nocardiopsis tropica*, *Nocardiopsis umidischolae* o *Nocardiopsis xinjiangensis*, preferiblemente *Nocardiopsis alha* DSM 15647, *Nocardiopsis dassonvillei* NRRL 18133, *Nocardiopsis dassonvillei subsp. dassonvillei* DSM 43235, *Nocardiopsis prasina* DSM 15648, *Nocardiopsis prasina* DSM 15649, *Nocardiopsis* sp. NRRL 18262, de forma más preferible, *Nocardiopsis* sp. FERM P-18676.

Una composición, tal como un aditivo para pienso, que comprende al menos una variante de proteasa de cualquiera de las reivindicaciones 1-20 y

(a) al menos una vitamina liposoluble;

(b) al menos una vitamina hidrosoluble y/o

(c) al menos un oligoelemento,

que además comprende, opcionalmente, al menos una enzima seleccionada del siguiente grupo de enzimas: amilasas, galactanasas, alfa-galactosidasas, xilanasas, endoglucanasas, endo-1,3(4)-beta-glucanasas, fitasas, fosfolipasas y otras proteasas; si se desea, que también comprenden al menos una amilasa y/o fosfolipasa.

La presente invención es descrita, además, por los siguientes ejemplos los cuales no deberían interpretarse como limitadores del ámbito de la invención.

Ejemplos

Ejemplo 1

Ensayos de proteasas

Ensayo de pNA

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

Temperatura: Temperatura ambiente (25°C)

Tampones de ensayo: 100 mM ácido succínico, 100 mM HEPES, 100 mM colinesterasas, 100 mM CABS, 1 mM CaCl₂, 150 mM KCl, 0,01% 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.

Se mezclan 20 µl de proteasa (diluida en Triton 100-X al 0,01%) con 100 µl de tampón de ensayo. El ensayo se comienza añadiendo 100 µl de sustrato de pNA (50 mg disueltos en 1,0 ml DMSO y además diluido 45x con Triton 100-X al 0,01%). El aumento en OD₄₀₅ se controla como una medida de la actividad de la proteasa.

Ensayo de Protazyme AK

Sustrato: tableta de Protazyme AK (caseína reticulada y coloreada; de Megazyme).

Temperatura: controlada (temperatura de ensayo).

Tampones de ensayo: 100 mM ácido succínico, 100 mM HEPES, 100 mM colinesterasas, 100 mM CABS, 1 mM CaCl₂, 150 mM KCl, Triton 100-X al 0,01% ajustados 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.

Una tableta de Protazyme AK se suspende en 2,0 ml de Triton 100-X al 0,01% por agitación suave. Se mezclan 500 µl de esta suspensión y 500 µl de tampón de ensayo en un tubo de Eppendorf y se coloca en hielo. Se añaden 20 µl de muestra de proteasa (diluido en Triton 100-X al 0,01%). El ensayo se inicia transfiriendo el 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 velocidad de agitación máxima (1400 r.p.m.). La incubación se detiene transfiriendo el tubo de nuevo al baño de hielo. Luego, el tubo se centrifuga en un centrifugador congelado durante unos minutos y se transfieren 200 µl de sobrenadante a una placa de microtitulación. OD₆₅₀ se lee como una medida de actividad de la proteasa. Un tampón ciego se incluye en el ensayo (en vez de enzima).

Ejemplo 2

Preparación y prueba de variantes de proteasa

Cuatro variantes de proteasa que comprenden la secuencia de aminoácidos de los aminoácidos 1 a 188 de SEC ID n.º 2 (proteasa 10) con las únicas sustituciones N47D, T127R, N92K y Q54R, respectivamente, fueron preparadas como se describe más abajo para la variante N47D.

ES 2 358 092 T3

La mutagénesis dirigida se efectuó usando el método “Mega-cebador” como describen Sarkar y Sommen, 1990 (BioTechniques 8: 404-407).

La variante N47D se construyó usando los siguientes cebadores, de los cuales el cebador R10WT-CL29 (SEC ID n.º 11) es específico del gen, y el cebador RSWT126 (SEC ID n.º 12) es mutagénico:

R10WT-CL29: 5' CCGATTATGGAGCGGATTGAACATGCG 3' (SEC ID n.º 11)

RSWT126: 5' GTGACCATCGGCGACGGCAGGGGCGTCTTCG 3' (SEC ID n.º 12),

para amplificar por PCR un fragmento de ADN de aproximadamente 469 pb del constructo descrito más abajo.

El constructo de ADN de proteasa 10 usado para la amplificación anterior fue un cassette de expresión (SEC ID n.º 13) para incorporación en el genoma de *Bacillus subtilis*. El constructo contiene una fusión de ADN que codifica la secuencia señal y el gen que codifica la proteína madura y pro de la proteasa 10 (SEC ID n.º 14), una construcción promotora y también el gen cat que confiere resistencia al cloranfenicol. Para facilitar la integración en el genoma por recombinación homóloga, las regiones flanqueantes de alrededor de 3 kb de genes endógenos de *Bacillus subtilis* fueron incorporados arriba y abajo de la secuencia de codificación de la proteasa 10.

El fragmento de 469 pb resultante fue purificado de un gel de agarosa (Sigma Aldrich cat. n.º A6877) y usado como un Mega-cebador junto con el cebador R10WT-CL39N (SEC ID n.º 15) en una segunda PCR realizada en el mismo molde.

R10WT-CL39N: 5' GGAGCTCTGAAAAAAGGAGAGGATAAAGAATGAA 3' (SEC ID n.º 15).

La construcción completa de aproximadamente 10 kb se hace *in vitro* por PCR de largo rango, usando los oligonucleótidos R10WT-CL28N (SEC ID n.º 16), R10WT-CL28C (SEC ID n.º 17), y el Sistema Expand Long Template PCR de Roche Applied Science (cat. n.º 11759060), según el manual del proveedor.

R10WT-CL28N: 5' GCGTTCGGATAATCGCGGTGACAATGCCG 3' (SEC ID n.º 16)

R10WT-CL28C: 5' TTCATGAGTCTGCGCCCTGAGATCCTCTG 3' (SEC ID n.º 17)

El fragmento de 1,2 kb aproximadamente resultante fue purificado y combinado en una reacción de PCR nueva usando el Expand Long Template PCR System con los fragmentos flanqueantes de la construcción hecha por dos reacciones de la PCR usando R10WT-2C-rev (SEC ID n.º 18) y R10WT-CL28C (SEC ID n.º 17); y RSWT001 (SEC ID n.º 19) y R10WT-CL28N (SEC ID n.º 16) como conjuntos de cebador. El fragmento de 10 kb resultante puede ser amplificado usando los cebadores R10WT-CL28N (SEC ID n.º 16) y R10WT-CL28C (SEC ID n.º 17), para aumentar el número de transformantes.

R10WT-2C-rev: 5' TAATCGCATGTTCAATCCGCTCCATAATCG 3' (SEC ID n.º 18)

RSWT001: 5' CCCAACGGTTTCTTCATTCTTTATCCTCTCCTTTTTTTCAGAGC 3' (SEC ID n.º 19)

Las células competentes de una cepa de proteasa y de amilasa baja de *Bacillus subtilis* (tal como cepa SHA273 descrita en WO92/11357 y WO95/10603) fueron transformadas con los fragmentos de la PCR resultantes respectivos, y se seleccionaron transformantes resistentes al cloranfenicol y se controlaron por secuenciación del ADN para verificar la presencia de la mutación correcta en el genoma.

Las células de constructos que albergan *Bacillus subtilis* que codifican la proteasa 10 y cada una de las cuatro variantes de las mismas fueron usadas para incubar matraces de agitación con contenido de medios ricos (PS-1: 100 g/L de sacarosa (Danisco cat.n.º 109-0429), 40 de g/L soja de corteza, 10 g/L Na₂HPO₄·12H₂O (Merck cat.n.º 6579), 0,1 ml/L de plurónico PE 6100 (BASF 102-3098)), y el cultivo se realizó durante cinco días a 30°C bajo agitación vigorosa.

Después del cultivo, los sobrenadantes fueron diluidos cuatro veces en un tampón Na₂HPO₄ 0,2M, titulados con 0,1M de ácido cítrico con pH 4,0 o pH 6,0, y fueron divididos en dos. Una mitad fue incubada durante cuatro horas a 65°C con el pH respectivo, después de lo cual fue congelada. La otra mitad fue congelada inmediatamente y sirvió como el control.

Antes de medir la actividad de la proteasa residual, las muestras fueron diluidas diez veces en 50 mM de tampón CHES-HEPES, pH 8,5. La actividad fue determinada usando una versión modificada del ensayo Protazyme AK del ejemplo 1, solubilizando una tableta del sustrato en 4 ml de tampón CHES-HEPES, pH 8,5, mezclando bajo agitación continua un ml de esta solución de sustrato con 20 ul de muestra de proteasa diluida, que fue luego incubada a 37°C. El sustrato debería tener la temperatura correcta antes de añadir la proteasa. Después de 15 minutos, la reacción fue detenida añadiendo 100 ul de NaOH 1 M y el sustrato insoluble fue precipitado por centrifugado a 15000 r.p.m. durante 3 minutos después de lo cual se midió la absorbencia a 650 nm. Los valores deberían estar por debajo de OD 3,0; alternativamente, la muestra de proteasa debería ser diluida más de diez veces antes de la medición de actividad.

ES 2 358 092 T3

La actividad residual relativa (%) se calcula dividiendo la actividad tras la incubación a 65°C con la actividad del control correspondiente. Los resultados de la tabla 1 a continuación muestra que las cuatro variantes son de una termoestabilidad mejorada en comparación con la proteasa 10.

TABLA 1

Actividad residual tras la incubación durante cuatro horas a 65°C

Proteasa	% actividad residual pH 6	% actividad residual pH 4
Proteasa 10 + N47D	44	68
Proteasa 10 + T127R	-	77
Proteasa 10 + N92K	-	55
Proteasa 10 + Q54R	52	67
Proteasa 10	19	41

Ejemplo 3

Variante de proteasa 22

Se diseñó una variante de proteasa designada "Proteasa 22" y que comprendía un número de sustituciones en trece de las diecisiete regiones especificadas en la reivindicación 1. Esta variante comprende las siguientes sustituciones en comparación con la parte madura de la proteasa 10 (aminoácidos 1-188 de SEC ID n.º 2): T10Y, A25S, R38T, Q42P, T44S, R49K, Q54R, V56I, A62S, T82S, S99A, G118Ns, S120T, S122R, E125Q, T129Y, N130S, M131L, R165S, T166A, F171Y, T176N, V179L, N180S, V184L y R185T.

La parte madura de la proteasa 22 son los aminoácidos 1-196 de SEC ID n.º 21. La secuencia de ADN correspondiente a SEC ID n.º 21 es SEC ID n.º 20.

La secuencia de ADN de SEC ID n.º 20 fue construida e introducida en un huésped de Bacillus para la expresión. La proteasa expresada fue purificada y caracterizada como una proteasa alfa-lítica (familia de peptidasa S1E y/o S2A).

Se midió la relación de actividad-temperatura de la proteasa 22 a pH 9, usando el ensayo Protazyme AK del ejemplo 1. La proteasa 10 se incluye con fines comparativos. Los resultados se muestran en la tabla 2 a continuación.

TABLA 2

Perfil de la temperatura con pH 9 de la proteasa 22 y la proteasa 10

Temperatura (°C)	Actividad relativa con pH 9	
	Proteasa 22	Proteasa 10
15	0.016	0.015
25	0.010	0.024
37	0.028	0.068
50	0.069	0.199
60	0.138	0.510
70	0.474	1.000
80	1.000	0.394
90	0.375	-

A partir de estos resultados, resulta que la proteasa 22 tiene una temperatura óptima más alta a pH 9 que la proteasa 10, es decir, alrededor de 80°C en comparación con alrededor de 70°C.

Se utilizó la calorimetría por análisis diferencial (DSC) para determinar la estabilidad de la temperatura con pH 7,0 de la proteasa 22 y la proteasa 10. Las proteasas purificadas fueron dializadas durante la noche a 4°C contra 10 mM de

ES 2 358 092 T3

fosfato sódico, 50 mM de cloruro sódico, pH 7,0 y se analizaron en un instrumento VP-DSC (Micro Cal) con una tasa de barrido constante de 1,5°C/min de 20 a 100°C. La manipulación de datos se realizó usando el software MicroCal Origin.

5 La desnaturalización resultante o temperaturas de fusión, las T_m , fueron: para la proteasa 22: 83.5°C; para la proteasa 10: 76,5°C.

La invención descrita y reivindicada en la presente no debe estar limitada en su alcance por las formas de realización específicas descritas en la presente, ya que estas formas de realización están destinadas a ser ilustraciones de diferentes
10 aspectos de la invención. De hecho, varias modificaciones de la invención además de las descritas y mostradas en la presente serán evidentes para los expertos en la técnica de la descripción precedente.

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Proteasa con un porcentaje de identidad de los aminoácidos 1 a 188 de una proteasa progenitora que consiste en SEC ID n.º 2 de al menos 60%, que exhibe una termoestabilidad mejorada en comparación con la proteasa progenitora por determinación de actividad residual tras la incubación durante cuatro horas a 65°C con pH 6 o pH 4 en un tampón Na₂HPO₄ 0,2M y que comprende al menos una de las siguientes sustituciones: N47D, Q54R, N92K, y/o T127R, donde cada posición corresponde a una posición de aminoácidos 1 a 188 de SEC ID n.º 2;
 con la condición de que la proteasa no sea:
 los aminoácidos 1-188 de SEC ID n.º 2, los aminoácidos 1-188 de SEC ID n.º 2 con la sustitución T87A, los aminoácidos 1-188 de SEC ID n.º 4, los aminoácidos 1-188 de SEC ID n.º 6, los aminoácidos 1-188 de SEC ID n.º 8, los aminoácidos 1-188 de SEC ID n.º 10, los aminoácidos 1-188 de SEC ID n.º 22, y no las partes maduras de las proteasas que tienen las secuencias de SEC ID n.º 23, SEC ID n.º 24, SEC ID n.º 25, SEC ID n.º 26, SEC ID n.º 27 y SEC ID n.º 28 y donde el grado de identidad entre secuencias se determina por el programa "align" usando la matriz de puntuación BLOSUM50, la penalización para el primer residuo de un espacio es -12 y las penalizaciones para otros residuos de un espacio son -2.
2. Proteasa según la reivindicación 1 que tiene la secuencia de aminoácidos 1 a 188 de SEC ID n.º 2 salvo al menos una sustitución seleccionada de los siguientes: N47D, Q54R, N92K y/o T127R.
3. Proteasa según la reivindicación 1 que tiene la secuencia de aminoácidos 1 a 188 de SEC ID n.º 4 salvo al menos una sustitución seleccionada de los siguientes: N47D, N54R, S92K y/o T127R.
4. Proteasa según la reivindicación 1 que tiene la secuencia de aminoácidos 1 a 188 de SEC ID n.º 6 salvo al menos una sustitución seleccionada de los siguientes: N47D, Q54R, N92K y/o T127R.
5. Proteasa según la reivindicación 1 que tiene la secuencia de aminoácidos 1 a 188 de SEC ID n.º 8 salvo al menos una sustitución seleccionada de los siguientes: N47D, Q54R, N92K y/o T127R.
6. Proteasa según la reivindicación 1 que tiene la secuencia de aminoácidos 1 a 188 de SEC ID n.º 10 salvo al menos una sustitución seleccionada de los siguientes: N47D, S92K y/o T127R.
7. Proteasa según la reivindicación 2 que se selecciona de los siguientes:
 N47D de SEC ID n.º 2,
 T127R de SEC ID n.º 2,
 N92K de SEC ID n.º 2 y
 Q54R de SEC ID n.º 2.
8. Secuencia de ácidos nucleicos aislada que comprende una secuencia de ácidos nucleicos que codifica la proteasa de cualquiera de las reivindicaciones 1-7.
9. Constructo de ácidos nucleicos que comprende la secuencia de ácidos nucleicos según la reivindicación 8 operativamente enlazada a una o más secuencias de control que dirigen la producción de la variante de proteasa en un huésped de expresión adecuado.
10. Vector de expresión recombinante que comprende el constructo de ácidos nucleicos según la reivindicación 9.
11. Célula huésped recombinante que comprende el constructo de ácidos nucleicos según la reivindicación 9 y/o el vector de expresión según la reivindicación 10.
12. Método para la producción de la proteasa según cualquiera de las reivindicaciones 1-7, el método comprendiendo:
 (a) cultivar la célula huésped según la reivindicación 11 para producir un sobrenadante que comprende la proteasa;
 y
 (b) recuperar la proteasa.
13. Planta transgénica, o parte de la planta, capaz de expresar una proteasa según cualquiera de las reivindicaciones 1-7.

ES 2 358 092 T3

14. Animal transgénico no humano, o productos, o elementos de los mismos, que sean capaces de expresar una proteasa según cualquiera de las reivindicaciones 1-7.

15. Aditivo de pienso que comprende al menos una proteasa según cualquiera de las reivindicaciones 1-7, y

(a) al menos una vitamina liposoluble;

(b) al menos una vitamina hidrosoluble y/o

(c) al menos un oligoelemento.

16. Composición de pienso para animales con un contenido de proteína bruta de 50 a 800 g/kg y que comprende la proteasa según cualquiera de las reivindicaciones 1-7.

17. Método para mejorar el valor nutritivo de un pienso para animales, donde la proteasa según cualquiera de las reivindicaciones 1-7, y/o la composición según cualquiera de las reivindicaciones 15 o 16 se añade al pienso.

18. Método para el tratamiento de proteínas, que comprende el paso de añadir la proteasa según cualquiera de las reivindicaciones 1-7 y/o la composición según cualquiera de las reivindicaciones 15-16 a al menos una proteína o fuente de proteína.

19. Uso de la proteasa según cualquiera de las reivindicaciones 1-7 y/o la composición según cualquiera de las reivindicaciones 15-16 (i) en pienso para animales; (ii) en la preparación de pienso para animales; (iii) para mejorar el valor nutritivo de pienso para animales; y/o (iv) para el tratamiento de proteínas.

20. Uso de la proteasa según cualquiera de las reivindicaciones 1-7 en detergentes.

Proteasa 10	1	50
Proteasa 18	ADIIGGLAYTMGGRC	SVGF
Proteasa 11	ADIIGGLAYTMGGRC	SVGF
Proteasa 35	ADIIGGLAYTMGGRC	SVGF
Proteasa 08	ADIIGGLAYTMGGRC	SVGF
Proteasa 22	ADIIGGLAYTMGGRC	SVGF
Proteasa 10	51	100
Proteasa 18	VFEQSVFPQNDAA	FVRG
Proteasa 11	VFEQSVFPQNDAA	FVRG
Proteasa 35	VFEQSVFPQNDAA	FVRG
Proteasa 08	VFEQSVFPQNDAA	FVRG
Proteasa 22	VFEQSVFPQNDAA	FVRG
Proteasa 10	101	150
Proteasa 18	CRSGSTTGWHCGTI	QARQ
Proteasa 11	CRSGSTTGWHCGTI	QARQ
Proteasa 35	CRSGSTTGWHCGTI	QARQ
Proteasa 08	CRSGSTTGWHCGTI	QARQ
Proteasa 22	CRSGSTTGWHCGTI	QARQ
Proteasa 10	151	188
Proteasa 18	TQAQGVTSGGSGN	CRTG
Proteasa 11	TQAQGVTSGGSGN	CRTG
Proteasa 35	TQAQGVTSGGSGN	CRTG
Proteasa 08	TQAQGVTSGGSGN	CRTG
Proteasa 22	TQAQGVTSGGSGN	CRTG

Fig. 1

ES 2 358 092 T3

ATOM	1	N	ALA	1	-18.517	32.531	28.661	1.00	8.90
ATOM	2	CB	ALA	1	-18.802	30.741	20.290	1.00	12.24
ATOM	3	CA	ALA	1	-19.308	31.313	28.965	1.00	10.86
ATOM	4	C	ALA	1	-20.783	31.666	29.080	1.00	12.18
ATOM	5	O	ALA	1	-21.113	32.695	29.712	1.00	12.73
ATOM	6	N	ASP	2	-21.722	30.930	28.510	1.00	12.01
ATOM	7	CA	ASP	2	-23.176	31.225	28.612	1.00	12.07
ATOM	8	C	ASP	2	-23.667	30.604	29.929	1.00	10.24
ATOM	9	O	ASP	2	-23.389	29.410	30.109	1.00	11.30
ATOM	10	CB	ASP	2	-23.995	30.629	27.422	1.00	12.43
ATOM	11	CG	ASP	2	-23.545	31.314	26.129	1.00	16.23
ATOM	12	OD1	ASP	2	-23.300	30.668	25.134	1.00	21.68
ATOM	13	OD2	ASP	2	-23.346	32.527	26.168	1.00	17.64
ATOM	14	N	ILE	3	-24.387	31.321	30.757	1.00	9.80
ATOM	15	CA	ILE	3	-24.850	30.687	32.027	1.00	8.80
ATOM	16	C	ILE	3	-26.252	30.135	31.768	1.00	7.97
ATOM	17	O	ILE	3	-27.160	30.953	31.648	1.00	8.91
ATOM	18	CB	ILE	3	-24.789	31.723	33.207	1.00	7.85
ATOM	19	CG1	ILE	3	-23.378	32.342	33.312	1.00	5.63
ATOM	20	CG2	ILE	3	-23.284	31.096	34.549	1.00	4.75
ATOM	21	CD1	ILE	3	-22.321	31.320	33.579	1.00	5.82
ATOM	22	N	ILE	4	-26.319	28.814	31.963	1.00	7.21
ATOM	23	CD1	ILE	4	-26.578	27.854	27.424	1.00	8.61
ATOM	24	CG1	ILE	4	-27.102	28.463	28.794	1.00	8.70
ATOM	25	CB	ILE	4	-27.272	27.363	29.888	1.00	7.20
ATOM	26	CG2	ILE	4	-28.446	26.419	29.544	1.00	6.35
ATOM	27	CA	ILE	4	-27.569	28.083	31.259	1.00	7.04
ATOM	28	C	ILE	4	-27.799	27.046	32.350	1.00	7.12
ATOM	29	O	ILE	4	-26.841	26.414	32.764	1.00	5.80
ATOM	30	N	GLY	5	-29.017	26.894	32.834	1.00	8.40
ATOM	31	CA	GLY	5	-29.415	25.958	33.863	1.00	5.51
ATOM	32	C	GLY	5	-29.031	24.550	33.483	1.00	6.74
ATOM	33	O	GLY	5	-29.222	24.181	32.306	1.00	6.02
ATOM	34	N	GLY	6	-28.492	23.787	34.436	1.00	5.32
ATOM	35	CA	GLY	6	-28.113	22.385	34.125	1.00	6.51
ATOM	36	C	GLY	6	-26.697	22.143	33.678	1.00	7.67
ATOM	37	O	GLY	6	-26.264	20.957	33.687	1.00	8.08
ATOM	38	N	LEU	7	-25.941	23.127	33.235	1.00	7.02
ATOM	39	CD2	LEU	7	-25.075	23.250	29.859	1.00	15.01
ATOM	40	CD1	LEU	7	-24.009	25.544	29.892	1.00	12.10
ATOM	41	CG	LEU	7	-24.823	24.494	30.662	1.00	11.57
ATOM	42	CB	LEU	7	-24.100	24.149	31.987	1.00	7.81
ATOM	43	CA	LEU	7	-24.543	22.889	32.774	1.00	7.23
ATOM	44	C	LEU	7	-23.543	22.624	33.891	1.00	8.17
ATOM	45	O	LEU	7	-23.779	23.055	35.054	1.00	8.83
ATOM	46	N	ALA	8	-22.480	21.931	33.560	1.00	7.85
ATOM	47	CB	ALA	8	-20.568	20.517	33.998	1.00	7.20
ATOM	48	CA	ALA	8	-21.436	21.658	34.583	1.00	6.67
ATOM	49	C	ALA	8	-20.554	22.867	34.856	1.00	8.14
ATOM	50	O	ALA	8	-20.241	23.793	34.058	1.00	7.62
ATOM	51	N	TYR	9	-20.078	22.906	36.110	1.00	6.90
ATOM	52	CA	TYR	9	-19.074	23.854	36.602	1.00	7.03
ATOM	53	C	TYR	9	-18.138	22.960	37.480	1.00	8.21
ATOM	54	O	TYR	9	-18.560	21.945	38.048	1.00	7.61

Fig. 2

Cont. Fig 2

ATOM	55	CB	TYR	9	-19.474	25.108	37.320	1.00	7.45
ATOM	56	CG	TYR	9	-20.138	24.925	38.664	1.00	9.14
ATOM	57	CD1	TYR	9	-19.401	24.898	39.853	1.00	8.86
ATOM	58	CD2	TYR	9	-21.559	24.816	38.673	1.00	8.56
ATOM	59	CE1	TYR	9	-20.044	24.756	41.062	1.00	7.11
ATOM	60	CE2	TYR	9	-22.214	24.696	39.930	1.00	8.78
ATOM	61	CE	TYR	9	-21.438	24.673	41.072	1.00	7.41
ATOM	62	OH	TYR	9	-22.115	24.537	42.348	1.00	8.28
ATOM	63	N	THR	10	-15.867	23.367	37.552	1.00	8.23
ATOM	64	CG2	THR	10	-15.380	21.144	36.171	1.00	21.31
ATOM	65	OG1	THR	10	-14.022	22.954	36.816	1.00	17.27
ATOM	66	CB	THR	10	-14.816	21.869	37.398	1.00	15.54
ATOM	67	CA	THR	10	-15.881	22.592	38.334	1.00	12.22
ATOM	68	C	THR	10	-15.190	23.495	39.381	1.00	12.59
ATOM	69	O	THR	10	-15.040	24.724	39.295	1.00	11.83
ATOM	70	N	MET	11	-14.719	22.854	40.422	1.00	13.86
ATOM	71	CE	MET	11	-18.117	21.521	42.992	0.70	10.20
ATOM	72	SD	MET	11	-16.384	21.817	43.260	0.70	13.92
ATOM	73	CG	MET	11	-16.351	23.607	42.742	0.70	8.87
ATOM	74	CB	MET	11	-14.945	24.074	42.557	0.70	13.60
ATOM	79	CA	MET	11	-14.003	23.423	41.576	1.00	14.74
ATOM	80	C	MET	11	-13.204	22.219	42.141	1.00	16.84
ATOM	81	O	MET	11	-13.132	22.126	43.360	1.00	18.29
ATOM	82	N	GLY	12	-12.650	21.380	41.252	1.00	17.41
ATOM	83	CA	GLY	12	-11.531	20.160	41.721	1.00	20.30
ATOM	84	C	GLY	12	-12.961	19.034	41.377	1.00	22.00
ATOM	85	O	GLY	12	-12.730	18.252	40.444	1.00	25.04
ATOM	86	N	GLY	13	-14.079	19.064	42.126	1.00	17.68
ATOM	87	CA	GLY	13	-15.219	18.171	41.900	1.00	15.21
ATOM	88	C	GLY	13	-16.127	18.873	40.846	1.00	15.26
ATOM	89	O	GLY	13	-15.681	19.862	40.228	1.00	14.61
ATOM	90	N	ARG	14	-17.370	18.410	40.637	1.00	12.77
ATOM	91	NH2	ARG	14	-20.479	14.276	37.036	1.00	17.15
ATOM	92	NH1	ARG	14	-21.587	16.075	36.340	1.00	13.21
ATOM	93	CZ	ARG	14	-20.415	15.529	36.584	1.00	16.51
ATOM	94	NE	ARG	14	-19.265	16.236	36.423	1.00	15.10
ATOM	95	CD	ARG	14	-19.240	17.643	36.031	1.00	15.25
ATOM	96	CG	ARG	14	-19.255	18.317	37.291	1.00	14.76
ATOM	97	CB	ARG	14	-18.333	18.056	38.435	1.00	11.53
ATOM	98	CA	ARG	14	-18.269	19.018	39.659	1.00	11.00
ATOM	99	C	ARG	14	-19.655	19.162	40.278	1.00	9.52
ATOM	100	O	ARG	14	-20.027	18.274	41.091	1.00	8.13
ATOM	101	N	CYS	15	-20.352	20.221	39.853	1.00	8.85
ATOM	102	CA	CYS	15	-21.782	20.417	40.285	1.00	6.14
ATOM	103	C	CYS	15	-22.455	21.027	39.084	1.00	6.53
ATOM	104	O	CYS	15	-21.754	21.176	38.036	1.00	8.10
ATOM	105	CB	CYS	15	-21.897	21.271	41.368	1.00	7.27
ATOM	106	SG	CYS	15	-21.795	20.241	43.088	1.00	8.70
ATOM	107	N	SER	16	-23.746	21.368	39.154	1.00	5.13
ATOM	108	CA	SER	16	-24.402	21.936	37.975	1.00	4.73
ATOM	109	C	SER	16	-24.969	23.294	38.269	1.00	6.73
ATOM	110	O	SER	16	-25.331	23.536	39.470	1.00	7.15
ATOM	111	CB	SER	16	-25.540	20.930	37.602	1.00	5.02
ATOM	112	OG	SER	16	-25.031	19.670	37.228	1.00	7.01

Fig. 2

Cont. Fig 2

ATOM	113	N	VAL	17	-25.177	24.129	37.276	1.00	5.71
ATOM	114	CA	VAL	17	-25.780	25.469	37.450	1.00	6.30
ATOM	115	C	VAL	17	-27.274	25.365	37.742	1.00	6.27
ATOM	116	O	VAL	17	-27.904	24.514	37.064	1.00	6.87
ATOM	117	CB	VAL	17	-25.589	26.211	36.113	1.00	4.80
ATOM	118	CG1	VAL	17	-26.252	27.572	36.079	1.00	2.00
ATOM	119	CG2	VAL	17	-24.108	26.435	35.892	1.00	5.56
ATOM	120	N	GLY	18	-27.836	26.136	38.622	1.00	8.84
ATOM	121	CA	GLY	18	-29.277	26.067	38.899	1.00	4.89
ATOM	122	C	GLY	18	-29.898	27.072	37.958	1.00	7.79
ATOM	123	O	GLY	18	-30.578	26.683	36.960	1.00	8.56
ATOM	124	N	PHE	19	-29.783	28.366	38.175	1.00	6.84
ATOM	125	CA	PHE	19	-30.368	29.391	37.291	1.00	8.35
ATOM	126	C	PHE	19	-29.457	30.625	37.254	1.00	8.67
ATOM	127	O	PHE	19	-28.889	30.984	38.285	1.00	7.20
ATOM	128	CB	PHE	19	-31.761	29.873	37.827	1.00	6.74
ATOM	129	CG	PHE	19	-32.786	28.779	38.033	1.00	8.40
ATOM	130	CD1	PHE	19	-33.490	28.300	36.918	1.00	9.81
ATOM	131	CD2	PHE	19	-32.921	28.194	39.301	1.00	8.27
ATOM	132	CE1	PHE	19	-34.414	27.241	37.060	1.00	7.45
ATOM	133	CE2	PHE	19	-33.804	27.129	39.460	1.00	7.36
ATOM	134	CZ	PHE	19	-34.541	26.662	38.347	1.00	8.66
ATOM	135	N	ALA	20	-29.375	31.284	36.114	1.00	8.32
ATOM	136	CB	ALA	20	-28.552	32.954	34.501	1.00	7.08
ATOM	137	CA	ALA	20	-28.577	32.514	35.976	1.00	6.99
ATOM	138	C	ALA	20	-29.347	33.558	36.793	1.00	8.39
ATOM	139	O	ALA	20	-30.614	33.548	36.744	1.00	6.62
ATOM	140	N	ALA	21	-28.653	34.461	37.453	1.00	6.45
ATOM	141	CB	ALA	21	-29.774	34.943	39.600	1.00	6.04
ATOM	142	CA	ALA	21	-29.305	35.514	38.244	1.00	9.09
ATOM	143	C	ALA	21	-28.267	36.598	38.599	1.00	10.25
ATOM	144	O	ALA	21	-27.048	36.412	38.434	1.00	10.39
ATOM	145	N	THR	22	-28.734	37.704	39.154	1.00	10.87
ATOM	146	CA	THR	22	-27.795	38.747	39.633	1.00	9.81
ATOM	147	C	THR	22	-28.044	38.773	41.139	1.00	12.70
ATOM	148	O	THR	22	-29.153	38.378	41.607	1.00	13.28
ATOM	149	CB	THR	22	-28.009	40.191	39.000	1.00	12.32
ATOM	150	OG1	THR	22	-29.443	40.520	39.201	1.00	17.96
ATOM	151	CG2	THR	22	-27.730	40.314	37.512	1.00	10.22
ATOM	152	N	ASN	23	-27.067	39.261	41.919	1.00	12.56
ATOM	153	ND2	ASN	23	-23.651	39.789	44.187	1.00	14.15
ATOM	154	OD1	ASN	23	-25.034	41.090	43.152	1.00	11.99
ATOM	155	CG	ASN	23	-34.917	40.044	43.825	1.00	13.60
ATOM	156	CE	ASN	23	-26.025	39.065	44.153	1.00	12.41
ATOM	157	CA	ASN	23	-27.308	39.367	43.381	1.00	14.31
ATOM	158	C	ASN	23	-27.947	40.754	43.600	1.00	15.71
ATOM	159	O	ASN	23	-28.252	41.558	42.664	1.00	13.68
ATOM	160	N	ALA	24	-28.043	41.088	44.883	1.00	16.68
ATOM	161	CB	ALA	24	-28.899	42.370	46.862	1.00	19.20
ATOM	162	CA	ALA	24	-28.626	42.344	45.371	1.00	20.06
ATOM	163	C	ALA	24	-27.831	43.543	44.936	1.00	22.43
ATOM	164	O	ALA	24	-28.408	44.658	44.795	1.00	24.93
ATOM	165	N	ALA	25	-26.556	43.412	44.648	1.00	23.59
ATOM	166	CA	ALA	25	-26.727	44.513	44.128	1.00	20.79

Fig. 2

Cont. Fig 2

ATOM	167	C	ALA	25	-25.765	44.542	42.613	1.00	21.34
ATOM	168	O	ALA	25	-25.018	45.388	42.040	1.00	24.26
ATOM	169	CB	ALA	25	-24.278	44.379	44.584	1.00	24.28
ATOM	170	N	GLY	26	-26.508	43.687	41.910	1.00	16.97
ATOM	171	CA	GLY	26	-26.453	43.763	40.456	1.00	15.31
ATOM	172	C	GLY	26	-25.320	43.024	39.803	1.00	13.98
ATOM	173	O	GLY	26	-25.158	43.168	38.560	1.00	15.90
ATOM	174	N	GLN	27	-24.594	42.196	40.523	1.00	13.26
ATOM	175	NE2	GLN	27	-19.688	42.607	42.319	1.00	23.84
ATOM	176	OE1	GLN	27	-21.306	41.674	43.669	1.00	19.08
ATOM	177	CD	GLN	27	-20.952	42.234	42.525	1.00	20.42
ATOM	178	CG	GLN	27	-21.934	42.487	41.519	1.00	16.87
ATOM	179	CB	GLN	27	-22.364	41.130	40.905	1.00	13.66
ATOM	180	CA	GLN	27	-23.488	41.430	39.904	1.00	11.98
ATOM	181	C	GLN	27	-24.023	40.113	39.345	1.00	11.67
ATOM	182	O	GLN	27	-24.829	39.428	39.949	1.00	11.63
ATOM	183	N	PRO	28	-23.539	39.714	38.197	1.00	10.36
ATOM	184	CG	PRO	28	-22.111	39.444	36.367	1.00	11.95
ATOM	185	CD	PRO	28	-22.544	40.519	37.403	1.00	10.61
ATOM	186	CB	PRO	28	-23.429	38.692	36.116	1.00	11.06
ATOM	187	CA	PRO	28	-23.977	38.497	37.537	1.00	9.08
ATOM	188	C	PRO	28	-23.418	37.248	38.194	1.00	9.80
ATOM	189	O	PRO	28	-22.278	37.282	38.749	1.00	9.50
ATOM	190	N	OLY	29	-24.245	36.179	38.101	1.00	6.29
ATOM	191	CA	GLY	29	-23.721	34.885	38.671	1.00	4.48
ATOM	192	C	GLY	29	-24.827	33.875	38.440	1.00	6.36
ATOM	193	O	GLY	29	-25.604	34.036	37.454	1.00	7.58
ATOM	194	N	PHE	30	-24.889	32.917	39.339	1.00	7.28
ATOM	195	CA	PHE	30	-25.971	31.691	39.292	1.00	7.77
ATOM	196	C	PHE	30	-26.232	31.306	40.703	1.00	6.83
ATOM	197	O	PHE	30	-25.281	31.334	41.532	1.00	8.49
ATOM	198	CB	PHE	30	-25.653	30.741	38.312	1.00	3.78
ATOM	199	CG	PHE	30	-24.384	29.955	38.483	1.00	5.51
ATOM	200	CD1	PHE	30	-24.299	28.836	39.311	1.00	5.54
ATOM	201	CD2	PHE	30	-23.251	30.336	37.752	1.00	8.26
ATOM	202	CE1	PHE	30	-23.126	28.108	39.451	1.00	8.21
ATOM	203	CE2	PHE	30	-21.996	29.661	37.898	1.00	6.20
ATOM	204	CZ	PHE	30	-21.971	28.509	38.739	1.00	7.61
ATOM	205	N	VAL	31	-27.413	30.739	40.862	1.00	4.88
ATOM	206	CA	VAL	31	-27.751	30.017	42.118	1.00	6.26
ATOM	207	C	VAL	31	-27.445	28.530	41.828	1.00	6.68
ATOM	208	O	VAL	31	-27.515	28.036	40.680	1.00	4.79
ATOM	209	CB	VAL	31	-29.141	30.296	42.666	1.00	7.03
ATOM	210	CG1	VAL	31	-29.230	31.765	43.136	1.00	11.46
ATOM	211	CG2	VAL	31	-30.190	29.902	41.646	1.00	8.54
ATOM	212	N	THR	32	-27.150	27.786	42.910	1.00	5.51
ATOM	213	CA	THR	32	-26.762	26.373	42.892	1.00	7.30
ATOM	214	C	THR	32	-26.833	25.866	44.356	1.00	8.85
ATOM	215	O	THR	32	-27.382	26.560	45.240	1.00	6.58
ATOM	216	CB	THR	32	-25.318	26.271	42.249	1.00	6.83
ATOM	217	CG1	THR	32	-24.927	24.904	42.030	1.00	6.22
ATOM	218	CG2	THR	32	-24.141	26.895	43.109	1.00	4.90
ATOM	219	N	ALA	33	-26.318	24.676	44.619	1.00	8.92
ATOM	220	CA	ALA	33	-26.313	24.007	45.928	1.00	9.53

Fig. 2

Cont. Fig 2

ATOM	221	C	ALA	33	-25.158	24.465	46.827	1.00	9.16
ATOM	222	O	ALA	33	-24.007	24.503	46.369	1.00	8.53
ATOM	223	CB	ALA	33	-26.294	22.473	45.773	1.00	7.77
ATOM	224	N	GLY	34	-25.408	24.724	48.076	1.00	7.16
ATOM	225	CA	GLY	34	-24.348	25.143	49.024	1.00	8.05
ATOM	226	C	GLY	34	-23.390	24.043	49.347	1.00	7.35
ATOM	227	O	GLY	34	-22.194	24.315	49.698	1.00	8.30
ATOM	228	N	HIS	35	-23.788	22.780	49.271	1.00	7.52
ATOM	229	CA	HIS	35	-22.821	21.714	49.636	1.00	6.58
ATOM	230	C	HIS	35	-21.744	21.601	48.560	1.00	8.95
ATOM	231	O	HIS	35	-20.702	20.945	48.747	1.00	8.96
ATOM	232	CB	HIS	35	-23.497	20.364	49.883	1.00	8.85
ATOM	233	CG	HIS	35	-23.991	19.599	48.686	1.00	6.87
ATOM	234	ND1	HIS	35	-25.305	19.481	48.321	1.00	8.56
ATOM	235	CD2	HIS	35	-23.326	18.872	47.769	1.00	5.55
ATOM	236	CE1	HIS	35	-25.414	18.744	47.228	1.00	7.54
ATOM	237	NE2	HIS	35	-24.217	18.313	46.906	1.00	8.64
ATOM	238	N	CYS	36	-21.930	22.183	47.376	1.00	8.15
ATOM	239	CA	CYS	36	-20.940	22.145	48.312	1.00	8.05
ATOM	240	C	CYS	36	-19.746	23.062	46.679	1.00	11.13
ATOM	241	O	CYS	36	-18.715	22.841	45.999	1.00	10.34
ATOM	242	CB	CYS	36	-21.518	22.598	44.977	1.00	5.97
ATOM	243	SG	CYS	36	-22.774	21.389	44.403	1.00	9.16
ATOM	244	N	GLY	37	-19.855	24.012	47.601	1.00	9.60
ATOM	245	CA	GLY	37	-18.632	24.821	47.862	1.00	8.52
ATOM	246	C	GLY	37	-18.653	25.793	48.998	1.00	11.72
ATOM	247	O	GLY	37	-19.923	26.350	49.153	1.00	12.57
ATOM	248	N	ARG	38	-17.807	26.044	49.767	1.00	9.44
ATOM	249	NH2	ARG	38	-13.066	27.478	54.730	0.00	41.69
ATOM	250	NH1	ARG	38	-14.258	25.862	55.765	0.00	42.03
ATOM	251	CZ	ARG	38	-13.968	26.494	54.619	0.00	41.38
ATOM	252	NE	ARG	38	-14.559	26.142	53.467	0.00	39.87
ATOM	253	CD	ARG	38	-15.763	25.165	53.320	0.00	37.50
ATOM	254	CG	ARG	38	-17.064	25.666	52.814	1.00	23.13
ATOM	255	CE	ARG	38	-16.502	26.727	51.799	1.00	14.90
ATOM	256	CA	ARG	38	-17.812	27.040	50.845	1.00	12.65
ATOM	257	C	ARG	38	-17.566	28.415	50.239	1.00	11.84
ATOM	258	O	ARG	38	-16.987	28.562	49.137	1.00	10.80
ATOM	259	N	VAL	39	-17.953	29.488	50.953	1.00	12.22
ATOM	260	CA	VAL	39	-17.723	30.881	50.483	1.00	11.55
ATOM	261	C	VAL	39	-16.224	31.002	50.149	1.00	11.46
ATOM	262	O	VAL	39	-15.406	30.540	50.944	1.00	12.65
ATOM	263	CB	VAL	39	-18.195	31.902	51.513	1.00	14.37
ATOM	264	CG1	VAL	39	-17.541	33.262	51.282	1.00	18.26
ATOM	265	CG2	VAL	39	-19.720	32.035	51.617	1.00	19.01
ATOM	266	N	GLY	40	-15.853	31.595	49.017	1.00	10.59
ATOM	267	CA	GLY	40	-14.467	31.715	48.658	1.00	9.21
ATOM	268	C	GLY	40	-13.962	30.690	47.689	1.00	11.61
ATOM	269	O	GLY	40	-12.904	30.954	47.066	1.00	13.99
ATOM	270	N	THR	41	-14.603	29.592	47.441	1.00	10.26
ATOM	271	CG2	THR	41	-14.886	26.180	45.550	1.00	8.90
ATOM	272	OG1	THR	41	-15.058	26.792	47.930	1.00	14.49
ATOM	273	CB	THR	41	-15.123	27.285	46.571	1.00	12.41
ATOM	274	CA	THR	41	-14.199	28.566	46.525	1.00	9.98

Fig. 2

Cont. Fig 2

ATOM	275	C	THR	41	-14.360	29.148	45.121	1.00	11.13
ATOM	276	O	THR	41	-15.404	29.657	44.713	1.00	10.17
ATOM	277	N	GLN	42	-13.297	28.983	44.365	1.00	12.36
ATOM	278	NH2	GLN	42	-11.317	32.840	43.178	0.70	32.68
ATOM	279	OE1	GLN	42	-9.407	31.714	42.552	0.70	35.48
ATOM	280	CD	GLN	42	-10.512	31.791	43.080	0.70	30.29
ATOM	281	CG	GLN	42	-11.132	30.558	43.680	0.70	24.69
ATOM	282	CB	GLN	42	-11.801	29.729	42.592	0.70	18.80
ATOM	288	CA	GLN	42	-13.263	29.456	42.977	1.00	11.77
ATOM	289	C	GLN	42	-13.852	28.392	42.063	1.00	13.86
ATOM	290	O	GLN	42	-13.615	27.187	42.330	1.00	12.07
ATOM	291	N	VAL	43	-14.544	28.817	41.002	1.00	13.96
ATOM	292	CG3	VAL	43	-17.397	28.751	39.309	1.00	15.91
ATOM	293	CG1	VAL	43	-17.192	27.656	41.542	1.00	10.29
ATOM	294	CB	VAL	43	-16.657	27.731	40.117	1.00	14.08
ATOM	295	CA	VAL	43	-18.132	27.886	40.018	1.00	11.05
ATOM	296	C	VAL	43	-14.717	28.297	38.584	1.00	10.70
ATOM	297	O	VAL	43	-14.393	29.467	38.214	1.00	8.75
ATOM	298	N	THR	44	-14.726	27.266	37.757	2.00	8.19
ATOM	299	CG2	THR	44	-13.085	26.655	34.277	1.00	16.80
ATOM	300	OG1	THR	44	-12.054	27.080	36.423	1.00	14.72
ATOM	301	CB	THR	44	-13.232	26.536	35.773	1.00	7.96
ATOM	302	CA	THR	44	-14.459	27.387	36.328	1.00	8.95
ATOM	303	C	THR	44	-15.708	26.877	35.588	1.00	10.78
ATOM	304	O	THR	44	-16.155	25.743	35.947	1.00	9.00
ATOM	305	N	ILE	45	-16.219	27.660	34.626	1.00	10.77
ATOM	306	CA	ILE	45	-17.352	27.175	33.850	1.00	11.78
ATOM	307	C	ILE	45	-16.950	27.476	32.396	1.00	10.56
ATOM	308	O	ILE	45	-16.976	28.673	32.030	1.00	10.62
ATOM	309	CB	ILE	45	-18.743	27.767	34.312	1.00	8.73
ATOM	310	CG1	ILE	45	-19.767	27.359	33.217	1.00	14.25
ATOM	311	CG2	ILE	45	-18.635	29.300	34.483	1.00	14.91
ATOM	312	CD1	ILE	45	-21.239	27.351	33.717	1.00	17.74
ATOM	313	N	GLY	46	-16.588	26.493	31.623	1.00	11.49
ATOM	314	CA	GLY	46	-16.162	26.796	30.214	1.00	13.72
ATOM	315	C	GLY	46	-15.009	27.812	30.282	1.00	12.34
ATOM	316	O	GLY	46	-14.011	27.661	31.002	1.00	14.15
ATOM	317	N	ASN	47	-15.134	38.895	29.512	1.00	12.14
ATOM	318	ND2	ASN	47	-15.075	31.342	26.221	1.00	26.89
ATOM	319	OD1	ASN	47	-16.257	31.266	28.086	1.00	18.12
ATOM	320	CG	ASN	47	-15.180	31.045	27.520	1.00	19.67
ATOM	321	CB	ASN	47	-13.914	30.476	28.081	1.00	17.20
ATOM	322	CA	ASN	47	-14.106	29.967	29.522	1.00	14.86
ATOM	323	C	ASN	47	-14.409	31.129	30.484	1.00	14.20
ATOM	324	O	ASN	47	-13.929	32.264	30.367	1.00	16.69
ATOM	325	N	GLY	48	-15.234	30.900	31.476	1.00	11.79
ATOM	326	CA	GLY	48	-15.629	31.860	32.487	1.00	10.60
ATOM	327	C	GLY	48	-15.114	31.350	33.840	1.00	8.68
ATOM	328	O	GLY	48	-14.741	30.180	34.021	1.00	8.85
ATOM	329	N	ARG	49	-15.067	32.286	34.782	1.00	8.71
ATOM	330	NH2	ARG	49	-10.753	29.684	38.714	0.00	43.61
ATOM	331	NH1	ARG	49	-9.430	30.717	37.086	0.00	39.51
ATOM	332	CZ	ARG	49	-10.647	30.223	37.466	0.00	42.12
ATOM	333	NE	ARG	49	-11.803	30.174	36.714	0.00	38.98

Fig. 2

Cont. Fig 2

ATOM	334	CD	ARG	49	-11.625	30.774	35.625	1.00	24.44
ATOM	335	CG	ARG	49	-12.079	32.161	35.395	1.00	20.17
ATOM	336	CE	ARG	49	-13.190	32.537	36.355	1.00	13.13
ATOM	337	CA	ARG	49	-14.600	31.985	36.143	1.00	8.87
ATOM	338	C	ARG	49	-15.485	32.746	37.136	1.00	9.56
ATOM	339	O	ARG	49	-16.026	33.814	36.747	1.00	10.73
ATOM	340	N	GLY	50	-15.644	32.236	38.337	1.00	10.04
ATOM	341	CA	GLY	50	-16.416	32.933	39.372	1.00	6.68
ATOM	342	C	GLY	50	-15.985	32.438	40.758	1.00	7.46
ATOM	343	O	GLY	50	-15.035	31.654	40.838	1.00	8.57
ATOM	344	N	VAL	51	-16.755	32.885	41.756	1.00	9.13
ATOM	345	CG2	VAL	51	-15.047	33.181	45.219	1.00	9.31
ATOM	346	CG1	VAL	51	-16.061	34.936	43.780	1.00	9.54
ATOM	347	CB	VAL	51	-15.469	33.510	43.758	1.00	10.40
ATOM	348	CA	VAL	51	-16.439	32.474	43.145	1.00	8.97
ATOM	349	C	VAL	51	-17.739	32.363	43.951	1.00	8.46
ATOM	350	O	VAL	51	-18.657	33.166	43.726	1.00	8.01
ATOM	351	N	PHE	52	-17.778	31.394	44.846	1.00	7.71
ATOM	352	CD2	PHE	52	-20.510	28.198	46.287	1.00	7.34
ATOM	353	CE2	PHE	52	-20.952	27.038	45.614	1.00	12.48
ATOM	354	CZ	PHE	52	-20.103	26.415	44.672	1.00	11.92
ATOM	355	CE1	PHE	52	-18.857	26.983	44.355	1.00	8.56
ATOM	356	CD1	PHE	52	-18.454	28.151	45.012	1.00	7.30
ATOM	357	CG	PHE	52	-19.269	28.765	45.964	1.00	10.46
ATOM	358	CB	PHE	52	-18.820	30.015	46.703	1.00	9.22
ATOM	359	CA	PHE	52	-18.916	31.237	45.766	1.00	9.28
ATOM	360	C	PHE	52	-18.928	32.498	46.637	1.00	10.41
ATOM	361	O	PHE	52	-17.979	32.752	47.403	1.00	10.33
ATOM	362	N	GLU	53	-20.038	33.239	46.481	1.00	8.81
ATOM	363	OE2	GLU	53	-22.012	37.756	45.426	1.00	27.06
ATOM	364	OE1	GLU	53	-21.229	39.265	46.722	1.00	33.78
ATOM	365	CD	GLU	53	-21.338	38.087	46.413	1.00	28.63
ATOM	366	CG	GLU	53	-20.701	36.961	47.162	1.00	15.08
ATOM	367	CB	GLU	53	-20.818	35.612	46.441	1.00	10.91
ATOM	368	CA	GLU	53	-20.172	34.475	47.239	1.00	10.55
ATOM	369	C	GLU	53	-21.035	34.208	48.485	1.00	13.01
ATOM	370	O	GLU	53	-20.664	34.743	49.558	1.00	11.89
ATOM	371	N	GLN	54	-22.095	33.444	48.352	1.00	9.12
ATOM	372	NE2	GLN	54	-26.256	36.329	49.251	1.00	43.84
ATOM	373	OE1	GLN	54	-25.933	35.594	51.419	1.00	46.53
ATOM	374	CD	GLN	54	-25.586	35.741	50.241	1.00	40.72
ATOM	375	CG	GLN	54	-24.256	35.205	49.756	1.00	29.81
ATOM	376	CB	GLN	54	-24.346	33.707	49.555	1.00	17.82
ATOM	377	CA	GLN	54	-22.955	33.105	49.508	1.00	10.81
ATOM	378	C	GLN	54	-23.164	31.600	49.527	1.00	12.82
ATOM	379	O	GLN	54	-23.418	31.101	48.410	1.00	13.75
ATOM	380	N	SER	55	-23.074	30.926	50.665	1.00	10.86
ATOM	381	OG	SER	55	-22.169	27.372	50.204	1.00	14.02
ATOM	382	CB	SER	55	-21.595	28.781	50.228	1.00	9.88
ATOM	383	CA	SER	55	-23.280	29.470	50.637	1.00	11.20
ATOM	384	C	SER	55	-23.730	28.998	52.014	1.00	12.70
ATOM	385	O	SER	55	-23.084	29.298	53.011	1.00	11.84
ATOM	386	N	VAL	56	-24.824	28.274	52.086	1.00	11.38
ATOM	387	CA	VAL	56	-25.345	27.735	53.342	1.00	9.56

Fig. 2

Cont. Fig 2

ATOM	388	C	VAL	56	-25.516	26.223	53.218	1.00	11.47
ATOM	389	O	VAL	56	-26.250	25.756	52.302	1.00	10.80
ATOM	390	CB	VAL	56	-26.691	28.365	53.715	1.00	12.57
ATOM	391	CG1	VAL	56	-27.250	27.561	54.895	1.00	13.65
ATOM	392	CG2	VAL	56	-26.842	29.809	54.111	1.00	14.61
ATOM	393	N	PHE	57	-24.903	25.475	54.116	1.00	9.44
ATOM	394	CA	PHE	57	-25.035	24.030	54.173	1.00	10.86
ATOM	395	C	PHE	57	-24.351	23.597	55.503	1.00	12.68
ATOM	396	O	PHE	57	-23.200	24.057	55.632	1.00	14.31
ATOM	397	CB	PHE	57	-24.383	23.289	52.962	1.00	9.20
ATOM	398	CG	PHE	57	-24.530	21.797	53.071	1.00	7.41
ATOM	399	CD1	PHE	57	-23.489	20.999	53.547	1.00	10.87
ATOM	400	CD2	PHE	57	-25.748	21.211	52.762	1.00	10.76
ATOM	401	CE1	PHE	57	-23.654	19.621	53.690	1.00	16.75
ATOM	402	CE2	PHE	57	-25.948	19.851	52.869	1.00	9.80
ATOM	403	CE	PHE	57	-24.897	19.033	53.326	1.00	17.07
ATOM	404	N	PRO	58	-24.888	22.759	56.355	1.00	11.76
ATOM	405	CA	PRO	58	-26.182	22.082	56.294	1.00	11.15
ATOM	406	C	PRO	58	-27.302	22.967	56.755	1.00	9.63
ATOM	407	O	PRO	58	-27.072	24.207	56.657	1.00	10.82
ATOM	408	CB	PRO	58	-25.955	20.768	57.043	1.00	11.03
ATOM	409	CG	PRO	58	-24.947	21.178	58.078	1.00	12.71
ATOM	410	CD	PRO	58	-24.125	22.322	57.531	1.00	12.73
ATOM	411	N	GLY	59	-28.466	22.432	57.103	1.00	10.97
ATOM	412	CA	GLY	59	-29.594	23.338	57.495	1.00	11.26
ATOM	413	C	GLY	59	-30.330	23.680	56.200	1.00	10.98
ATOM	414	O	GLY	59	-31.477	23.240	56.091	1.00	11.30
ATOM	415	N	ASN	60	-29.767	24.482	55.291	1.00	10.66
ATOM	416	CA	ASN	60	-30.400	24.729	53.962	1.00	8.17
ATOM	417	C	ASN	60	-29.377	24.099	52.981	1.00	9.56
ATOM	418	O	ASN	60	-28.346	23.532	53.474	1.00	8.76
ATOM	419	CB	ASN	60	-30.598	26.175	53.595	1.00	9.50
ATOM	420	CG	ASN	60	-31.369	26.934	54.664	1.00	12.19
ATOM	421	OD1	ASN	60	-30.872	27.984	55.082	1.00	16.81
ATOM	422	ND2	ASN	60	-32.478	26.340	55.060	1.00	15.11
ATOM	423	N	ASP	61	-29.582	24.193	51.661	1.00	7.15
ATOM	424	CA	ASP	61	-28.544	23.661	50.701	1.00	7.83
ATOM	425	C	ASP	61	-28.598	24.692	49.547	1.00	8.69
ATOM	426	O	ASP	61	-29.213	24.393	48.519	1.00	7.71
ATOM	427	CB	ASP	61	-28.818	22.216	50.313	1.00	4.48
ATOM	428	CG	ASP	61	-27.637	21.575	49.640	1.00	6.18
ATOM	429	OD1	ASP	61	-27.591	20.419	49.245	1.00	6.86
ATOM	430	OD2	ASP	61	-26.622	22.316	49.431	1.00	8.20
ATOM	431	N	ALA	62	-28.041	25.868	49.751	1.00	7.72
ATOM	432	CB	ALA	62	-29.258	27.857	49.385	1.00	8.64
ATOM	433	CA	ALA	62	-28.134	26.967	48.775	1.00	8.98
ATOM	434	C	ALA	62	-26.880	27.802	48.618	1.00	9.19
ATOM	435	O	ALA	62	-26.114	27.939	49.618	1.00	10.25
ATOM	436	N	ALA	63	-26.667	28.360	47.434	1.00	7.69
ATOM	437	CA	ALA	63	-25.476	29.189	47.173	1.00	7.50
ATOM	438	C	ALA	63	-25.668	30.110	43.987	1.00	7.33
ATOM	439	O	ALA	63	-26.526	29.843	45.116	1.00	6.36
ATOM	440	CB	ALA	63	-24.247	28.263	46.886	1.00	4.57
ATOM	441	N	PHE	64	-24.859	31.172	45.985	1.00	8.23

Fig. 2

Cont. Fig 2

ATOM	442	CD2	PHE	64	-24.221	35.364	43.942	1.00	6.73
ATOM	443	CH2	PHE	64	-24.051	36.237	42.841	1.00	7.23
ATOM	444	CZ	PHE	64	-25.063	36.205	41.842	1.00	6.08
ATOM	445	CE1	PHE	64	-26.166	35.352	41.959	1.00	10.30
ATOM	446	CD1	PHE	64	-26.263	34.453	43.038	1.00	10.04
ATOM	447	CG	PHE	64	-25.287	34.477	44.027	1.00	6.69
ATOM	448	CB	PHE	64	-25.415	33.537	45.240	1.00	6.42
ATOM	449	CA	PHE	64	-24.840	32.152	44.922	1.00	7.87
ATOM	450	C	PHE	64	-23.351	32.196	44.518	1.00	10.06
ATOM	451	O	PHE	64	-22.454	32.478	45.362	1.00	10.39
ATOM	452	N	VAL	65	-23.080	31.952	43.234	1.00	8.61
ATOM	453	CA	VAL	65	-21.722	32.028	42.662	1.00	8.62
ATOM	454	C	VAL	65	-21.686	33.327	41.831	1.00	8.89
ATOM	455	O	VAL	65	-22.514	33.548	40.948	1.00	7.44
ATOM	456	CB	VAL	65	-21.338	30.840	41.722	1.00	11.34
ATOM	457	CG1	VAL	65	-20.018	31.055	40.967	1.00	10.37
ATOM	458	CG2	VAL	65	-21.333	29.530	42.493	1.00	9.94
ATOM	459	N	ARG	66	-20.744	34.213	42.094	1.00	6.55
ATOM	460	NH2	ARG	66	-16.111	39.098	43.470	1.00	18.29
ATOM	461	NH1	ARG	66	-17.309	39.979	41.747	1.00	18.67
ATOM	462	CZ	ARG	66	-17.271	39.260	42.832	1.00	18.74
ATOM	463	NE	ARG	66	-18.308	38.673	43.409	1.00	22.97
ATOM	464	CD	ARG	66	-19.672	38.751	42.950	1.00	23.13
ATOM	465	CG	ARG	66	-19.916	37.827	41.797	1.00	18.33
ATOM	466	CE	ARG	66	-19.949	36.464	42.422	1.00	12.11
ATOM	467	CA	ARG	66	-20.545	35.475	41.415	1.00	8.00
ATOM	468	C	ARG	66	-19.501	35.310	40.305	1.00	7.85
ATOM	469	O	ARG	66	-18.447	34.738	40.557	1.00	8.51
ATOM	470	N	GLY	67	-19.828	35.857	39.132	1.00	8.23
ATOM	471	CA	GLY	67	-18.921	35.779	37.962	1.00	3.28
ATOM	472	C	GLY	67	-17.838	36.834	38.114	1.00	7.11
ATOM	473	O	GLY	67	-18.123	38.019	38.332	1.00	9.23
ATOM	474	N	THR	68	-16.585	36.418	37.933	1.00	6.19
ATOM	475	CA	THR	68	-15.407	37.292	37.994	1.00	6.13
ATOM	476	C	THR	68	-14.784	37.468	36.611	1.00	10.55
ATOM	477	O	THR	68	-13.939	38.358	36.375	1.00	9.20
ATOM	478	CB	THR	68	-14.366	36.832	39.071	1.00	13.02
ATOM	479	CG1	THR	68	-13.865	35.569	38.579	1.00	9.29
ATOM	480	CG2	THR	68	-14.870	36.773	40.522	1.00	11.87
ATOM	481	N	SER	69	-15.205	36.672	35.618	1.00	12.02
ATOM	482	CA	SER	69	-14.736	36.796	34.233	1.00	11.62
ATOM	483	C	SER	69	-15.660	36.069	33.264	1.00	12.37
ATOM	484	O	SER	69	-15.911	34.868	33.480	1.00	11.23
ATOM	485	CB	SER	69	-13.337	36.184	34.027	1.00	12.32
ATOM	486	CG	SER	69	-12.823	36.389	32.763	1.00	15.07
ATOM	487	N	ASN	70	-16.100	36.767	32.236	1.00	12.58
ATOM	488	ND2	ASN	70	-16.655	36.201	28.330	1.00	20.03
ATOM	489	OD1	ASN	70	-15.425	34.370	28.194	1.00	19.07
ATOM	490	CG	ASN	70	-15.954	35.243	28.878	1.00	14.54
ATOM	491	CB	ASN	70	-15.848	35.335	30.379	1.00	8.27
ATOM	492	CA	ASN	70	-16.894	36.195	31.131	1.00	11.17
ATOM	493	C	ASN	70	-18.166	35.424	31.400	1.00	10.53
ATOM	494	O	ASN	70	-18.343	34.319	30.817	1.00	14.17
ATOM	495	N	PHE	71	-19.048	35.946	32.209	1.00	8.98

Fig. 2

Cont. Fig 2

ATOM	496	CD2	PHE	71	-21.139	34.054	36.002	1.00	5.72
ATOM	497	CE2	PHE	71	-20.735	33.063	36.900	1.00	7.44
ATOM	498	CZ	PHE	71	-19.483	32.395	36.669	1.00	10.35
ATOM	499	CE1	PHE	71	-18.713	32.681	35.550	1.00	10.41
ATOM	500	CD1	PHE	71	-19.159	33.681	34.652	1.00	9.67
ATOM	501	CG	PHE	71	-20.350	34.372	34.873	1.00	5.76
ATOM	502	CB	PHE	71	-20.839	35.417	33.897	1.00	7.11
ATOM	503	CA	PHE	71	-20.346	35.257	32.435	1.00	10.26
ATOM	504	C	PHE	71	-21.337	35.958	31.503	1.00	11.61
ATOM	505	O	PHE	71	-21.429	37.207	31.537	1.00	14.68
ATOM	506	N	THR	72	-22.071	35.275	30.691	1.00	11.03
ATOM	507	CA	THR	72	-23.133	35.902	29.860	1.00	9.26
ATOM	508	C	THR	72	-24.405	35.261	30.425	1.00	9.51
ATOM	509	O	THR	72	-24.562	34.027	30.269	1.00	11.35
ATOM	510	CB	THR	72	-23.036	35.596	28.328	1.00	25.61
ATOM	511	CG1	THR	72	-21.768	36.189	27.928	1.00	15.86
ATOM	512	CG2	THR	72	-24.177	36.178	27.494	1.00	14.54
ATOM	513	N	LEU	73	-25.282	36.028	31.020	1.00	11.12
ATOM	514	CA	LEU	73	-26.502	35.479	31.608	1.00	9.31
ATOM	515	C	LEU	73	-27.504	35.184	30.484	1.00	9.04
ATOM	516	O	LEU	73	-27.563	35.896	29.460	1.00	8.44
ATOM	517	CB	LEU	73	-27.046	36.412	32.687	1.00	8.84
ATOM	518	CG	LEU	73	-26.215	36.869	33.883	1.00	13.85
ATOM	519	CD1	LEU	73	-27.074	37.521	34.968	1.00	12.85
ATOM	520	CD2	LEU	73	-25.471	35.702	34.530	1.00	9.85
ATOM	521	N	THR	74	-28.295	34.161	30.737	1.00	7.42
ATOM	522	CG2	THR	74	-27.696	32.554	28.200	1.00	3.40
ATOM	523	CG1	THR	74	-29.174	31.283	29.642	1.00	9.42
ATOM	524	CB	THR	74	-29.063	32.532	28.922	1.00	7.73
ATOM	525	CA	THR	74	-29.389	33.735	29.859	1.00	7.40
ATOM	526	C	THR	74	-30.600	33.347	30.732	1.00	10.05
ATOM	527	O	THR	74	-30.473	33.184	31.959	1.00	7.44
ATOM	528	N	ASN	75	-31.775	33.201	30.069	1.00	8.53
ATOM	529	ND2	ASN	75	-36.021	33.747	28.974	1.00	14.07
ATOM	530	CD1	ASN	75	-33.929	33.279	28.252	1.00	14.16
ATOM	531	CG	ASN	75	-34.723	33.826	29.157	1.00	15.69
ATOM	532	CB	ASN	75	-34.178	33.518	30.570	1.00	11.17
ATOM	533	CA	ASN	75	-32.924	32.729	30.849	1.00	9.39
ATOM	534	C	ASN	75	-33.156	31.252	30.465	1.00	10.95
ATOM	535	O	ASN	75	-34.322	30.835	30.620	1.00	13.03
ATOM	536	N	LEU	76	-32.177	30.516	29.993	1.00	7.99
ATOM	537	CD2	LEU	76	-32.993	29.592	26.412	1.00	11.04
ATOM	538	CD1	LEU	76	-30.530	29.917	26.112	1.00	13.41
ATOM	539	CG	LEU	76	-31.687	29.987	27.082	1.00	11.12
ATOM	540	CB	LEU	76	-31.416	29.013	28.250	1.00	10.63
ATOM	541	CA	LEU	76	-32.315	29.158	29.530	1.00	8.79
ATOM	542	C	LEU	76	-31.876	28.059	30.512	1.00	8.97
ATOM	543	O	LEU	76	-31.038	28.290	31.385	1.00	6.01
ATOM	544	N	VAL	77	-32.529	26.936	30.323	1.00	8.54
ATOM	545	CA	VAL	77	-32.285	25.689	31.062	1.00	8.05
ATOM	546	C	VAL	77	-32.125	24.578	29.973	1.00	9.00
ATOM	547	O	VAL	77	-33.126	24.380	29.178	1.00	6.78
ATOM	548	CB	VAL	77	-33.397	25.290	32.052	1.00	8.09
ATOM	549	CG1	VAL	77	-33.049	23.911	32.715	1.00	5.84

Fig. 2

Cont. Fig 2

ATOM	550	CG2	VAL	77	-33.575	26.350	33.142	1.00	10.07
ATOM	551	N	SER	78	-31.017	23.871	29.974	1.00	8.67
ATOM	552	OG	SER	78	-29.161	21.344	27.899	1.00	9.85
ATOM	553	CB	SER	78	-29.355	22.409	28.857	1.00	12.29
ATOM	554	CA	SER	78	-30.831	22.814	28.929	1.00	8.23
ATOM	555	C	SER	78	-31.710	21.583	29.189	1.00	9.82
ATOM	556	O	SER	78	-31.759	21.055	30.305	1.00	7.69
ATOM	557	N	ARG	79	-32.337	21.082	28.116	1.00	9.80
ATOM	558	NH2	ARG	79	-37.443	22.028	24.212	1.00	31.34
ATOM	559	NH1	ARG	79	-38.500	22.780	26.184	1.00	27.56
ATOM	560	CZ	ARG	79	-37.601	21.991	25.584	1.00	28.07
ATOM	561	NE	ARG	79	-36.785	21.131	26.161	1.00	21.89
ATOM	562	CD	ARG	79	-36.718	20.936	27.589	1.00	18.78
ATOM	563	CG	ARG	79	-35.265	21.118	27.992	1.00	9.14
ATOM	564	CB	ARG	79	-34.499	19.957	27.396	1.00	7.07
ATOM	565	CA	ARG	79	-33.163	19.879	28.166	1.00	10.01
ATOM	566	C	ARG	79	-32.272	18.740	27.659	1.00	11.09
ATOM	567	O	ARG	79	-32.799	17.622	27.476	1.00	11.76
ATOM	568	N	TYR	80	-30.980	18.901	27.456	1.00	10.89
ATOM	569	OH	TYR	80	-25.669	19.608	31.411	1.00	13.70
ATOM	570	CD2	TYR	80	-28.969	18.626	30.175	1.00	8.55
ATOM	571	CE2	TYR	80	-28.018	19.277	30.962	1.00	9.07
ATOM	572	CZ	TYR	80	-26.667	19.006	30.683	1.00	13.15
ATOM	573	CE1	TYR	80	-26.290	18.103	29.673	1.00	13.12
ATOM	574	CD1	TYR	80	-27.305	17.481	28.921	1.00	12.00
ATOM	575	CG	TYR	80	-28.646	17.742	29.153	1.00	10.87
ATOM	576	CB	TYR	80	-29.686	17.010	28.331	1.00	11.71
ATOM	577	CA	TYR	80	-30.100	17.809	27.036	1.00	13.19
ATOM	578	C	TYR	80	-30.669	16.889	25.939	1.00	15.57
ATOM	579	O	TYR	80	-31.074	17.414	24.848	1.00	16.53
ATOM	580	N	ASN	81	-30.732	15.583	26.157	1.00	15.67
ATOM	581	ND2	ASN	81	-30.251	12.075	23.672	0.50	23.42
ATOM	582	OD1	ASN	81	-31.553	10.851	24.930	0.50	22.68
ATOM	583	CG	ASN	81	-31.012	11.924	24.769	0.50	24.69
ATOM	584	CB	ASN	81	-30.915	13.117	25.699	0.50	20.23
ATOM	589	CA	ASN	81	-31.169	14.536	25.163	1.00	17.67
ATOM	590	C	ASN	81	-32.611	14.689	24.807	1.00	18.70
ATOM	591	O	ASN	81	-33.067	14.207	23.731	1.00	20.41
ATOM	592	N	THR	82	-33.405	15.385	25.621	1.00	15.25
ATOM	593	CG2	THR	82	-36.933	16.845	26.136	1.00	15.50
ATOM	594	OG1	THR	82	-35.944	14.838	27.235	1.00	15.80
ATOM	595	CB	THR	82	-35.663	16.070	26.495	1.00	13.76
ATOM	596	CA	THR	82	-34.787	15.661	25.275	1.00	16.61
ATOM	597	C	THR	82	-34.775	16.712	24.128	1.00	19.57
ATOM	598	O	THR	82	-35.725	16.765	23.314	1.00	20.18
ATOM	599	N	GLY	83	-33.765	17.555	23.973	1.00	17.44
ATOM	600	CA	GLY	83	-33.611	18.561	22.941	1.00	16.21
ATOM	601	C	GLY	83	-34.082	19.964	23.286	1.00	12.64
ATOM	602	O	GLY	83	-35.127	20.225	23.908	1.00	12.83
ATOM	603	N	GLY	84	-33.281	20.926	22.859	1.00	11.84
ATOM	604	CA	GLY	84	-33.604	22.331	23.082	1.00	11.09
ATOM	605	C	GLY	84	-33.492	22.741	24.541	1.00	9.78
ATOM	606	O	GLY	84	-32.796	22.177	25.384	1.00	10.83
ATOM	607	N	TYR	85	-34.104	23.875	24.825	1.00	10.88

Fig. 2

Cont. Fig 2

ATOM	606	OH	TYR	85	-28.004	24.555	24.203	1.00	18.98
ATOM	609	CD2	TYR	85	-31.561	25.395	24.068	1.00	12.07
ATOM	610	CE2	TYR	85	-30.266	25.023	23.693	1.00	13.73
ATOM	611	CZ	TYR	85	-29.263	24.890	24.644	1.00	15.05
ATOM	612	CE1	TYR	85	-29.519	25.081	26.012	1.00	12.94
ATOM	613	CD1	TYR	85	-30.826	25.421	26.395	1.00	12.96
ATOM	614	CG	TYR	85	-31.850	25.570	25.445	1.00	11.64
ATOM	615	CB	TYR	85	-33.264	25.866	25.882	1.00	10.50
ATOM	616	CA	TYR	85	-34.079	24.539	26.116	1.00	10.86
ATOM	617	C	TYR	85	-35.418	25.002	26.674	1.00	13.82
ATOM	618	O	TYR	85	-36.268	25.491	25.854	1.00	14.41
ATOM	619	N	ALA	86	-35.569	24.891	27.969	1.00	9.86
ATOM	620	CB	ALA	86	-37.046	24.630	29.971	1.00	9.75
ATOM	621	CA	ALA	86	-36.735	25.425	28.695	1.00	11.67
ATOM	622	C	ALA	86	-36.361	26.918	28.958	1.00	11.87
ATOM	623	O	ALA	86	-35.188	27.341	28.972	1.00	9.92
ATOM	624	N	THR	87	-37.345	27.829	29.131	1.00	10.43
ATOM	625	CG2	THR	87	-36.841	29.875	26.861	1.00	17.11
ATOM	626	OG1	THR	87	-38.959	29.934	28.083	1.00	17.19
ATOM	627	CB	THR	87	-37.539	30.174	28.200	1.00	17.24
ATOM	628	CA	THR	87	-37.057	29.241	29.379	1.00	10.64
ATOM	629	C	THR	87	-37.640	29.593	30.724	1.00	11.34
ATOM	630	O	THR	87	-38.696	29.055	31.041	1.00	13.09
ATOM	631	N	VAL	88	-37.001	30.448	31.521	1.00	10.66
ATOM	632	CA	VAL	88	-37.441	30.863	32.856	1.00	9.45
ATOM	633	C	VAL	88	-38.255	32.155	32.696	1.00	11.88
ATOM	634	O	VAL	88	-37.698	33.094	32.136	1.00	10.63
ATOM	635	CB	VAL	88	-36.246	31.053	33.821	1.00	7.00
ATOM	636	CG1	VAL	88	-36.652	31.636	35.192	1.00	5.32
ATOM	637	CG2	VAL	88	-35.478	29.746	33.985	1.00	9.05
ATOM	638	N	ALA	89	-39.467	32.233	33.201	1.00	10.89
ATOM	639	CB	ALA	89	-41.539	33.201	32.260	1.00	9.11
ATOM	640	CA	ALA	89	-40.325	33.440	33.132	1.00	11.01
ATOM	641	C	ALA	89	-40.690	33.918	34.548	1.00	11.13
ATOM	642	O	ALA	89	-41.242	35.028	34.655	1.00	13.90
ATOM	643	N	GLY	90	-40.351	33.181	35.617	1.00	9.54
ATOM	644	CA	GLY	90	-40.632	33.592	36.976	1.00	8.75
ATOM	645	C	GLY	90	-40.322	32.409	37.921	1.00	10.19
ATOM	646	O	GLY	90	-39.660	31.477	37.500	1.00	9.62
ATOM	647	N	HIS	91	-40.857	32.540	39.135	1.00	11.81
ATOM	648	CD2	HIS	91	-40.073	33.404	43.018	1.00	18.32
ATOM	649	NE2	HIS	91	-40.052	34.763	43.177	1.00	17.97
ATOM	650	CE1	HIS	91	-39.622	35.366	42.127	1.00	17.17
ATOM	651	ND1	HIS	91	-39.350	34.411	41.259	1.00	16.07
ATOM	652	CG	HIS	91	-39.605	33.184	41.765	1.00	16.04
ATOM	653	CB	HIS	91	-39.411	31.883	41.060	1.00	11.82
ATOM	654	CA	HIS	91	-40.637	31.530	40.180	1.00	10.58
ATOM	655	C	HIS	91	-41.854	31.229	41.025	1.00	12.38
ATOM	656	O	HIS	91	-41.723	31.032	42.248	1.00	13.11
ATOM	657	N	ASN	92	-43.013	31.126	40.369	1.00	12.16
ATOM	658	ND2	ASN	92	-46.608	32.760	41.360	1.00	44.37
ATOM	659	OD1	ASN	92	-45.564	33.225	39.385	1.00	40.52
ATOM	660	CG	ASN	92	-45.903	32.418	40.259	1.00	35.33
ATOM	661	CB	ASN	92	-45.524	30.938	40.252	1.00	23.82

Fig. 2

Cont. Fig 2

ATOM	662	CA	ASN	92	-44.261	30.746	41.119	1.00	13.35
ATOM	663	C	ASN	92	-44.154	29.259	41.493	1.00	12.66
ATOM	664	O	ASN	92	-43.930	28.437	40.582	1.00	11.76
ATOM	665	N	GLN	93	-44.364	28.935	42.749	1.00	12.08
ATOM	666	NH2	GLN	93	-42.340	27.432	47.693	1.00	21.50
ATOM	667	OH1	GLN	93	-44.559	27.501	47.458	1.00	31.26
ATOM	668	CD	GLN	93	-43.422	27.632	46.972	1.00	27.63
ATOM	669	CG	GLN	93	-43.291	28.002	45.511	1.00	20.72
ATOM	670	CB	GLN	93	-44.409	27.379	44.700	1.00	15.24
ATOM	671	CA	GLN	93	-44.263	27.516	43.171	1.00	12.74
ATOM	672	C	GLN	93	-45.394	26.705	42.565	1.00	13.82
ATOM	673	O	GLN	93	-46.572	27.162	42.672	1.00	15.94
ATOM	674	N	ALA	94	-45.166	25.549	42.048	1.00	11.74
ATOM	675	CA	ALA	94	-46.178	24.676	41.481	1.00	11.41
ATOM	676	C	ALA	94	-46.815	23.934	42.649	1.00	14.20
ATOM	677	O	ALA	94	-46.120	23.587	43.637	1.00	14.08
ATOM	678	CB	ALA	94	-45.495	23.704	40.529	1.00	6.58
ATOM	679	N	PRO	95	-48.112	23.645	42.551	1.00	14.96
ATOM	680	CG	PRO	95	-50.259	23.275	41.661	1.00	15.96
ATOM	681	CD	PRO	95	-48.954	24.020	41.412	1.00	15.38
ATOM	682	CE	PRO	95	-50.261	22.810	43.087	1.00	13.99
ATOM	683	CA	PRO	95	-48.815	22.843	43.571	1.00	13.77
ATOM	684	C	PRO	95	-48.308	21.414	43.670	1.00	12.85
ATOM	685	O	PRO	95	-47.789	20.722	42.764	1.00	13.21
ATOM	686	N	ILE	96	-48.439	20.860	44.892	1.00	9.64
ATOM	687	CD1	ILE	96	-45.305	19.598	47.937	1.00	20.39
ATOM	688	CG1	ILE	96	-47.785	19.940	47.690	1.00	17.20
ATOM	689	CB	ILE	96	-48.425	19.023	46.634	1.00	13.72
ATOM	690	CG2	ILE	96	-48.131	17.530	46.952	1.00	17.91
ATOM	691	CA	ILE	96	-48.058	19.444	45.141	1.00	11.51
ATOM	692	C	ILE	96	-48.841	18.627	44.138	1.00	13.14
ATOM	693	O	ILE	96	-50.052	18.979	43.880	1.00	15.21
ATOM	694	N	GLY	97	-48.332	17.575	43.828	1.00	10.18
ATOM	695	CA	GLY	97	-49.020	16.783	42.537	1.00	9.00
ATOM	696	C	GLY	97	-48.645	17.200	41.126	1.00	11.33
ATOM	697	O	GLY	97	-48.867	15.374	40.221	1.00	12.32
ATOM	698	N	SER	98	-48.108	18.393	40.935	1.00	10.97
ATOM	699	OG	SER	98	-48.122	21.262	40.073	0.70	15.96
ATOM	700	CB	SER	98	-47.149	20.355	39.724	0.70	13.01
ATOM	703	CA	SER	98	-47.643	18.918	39.637	1.00	11.70
ATOM	704	C	SER	98	-46.376	18.198	39.140	1.00	11.01
ATOM	705	O	SER	98	-45.567	17.708	39.906	1.00	12.53
ATOM	706	N	SER	99	-45.203	18.149	37.825	1.00	8.54
ATOM	707	OG	SER	99	-45.372	18.423	34.957	2.00	12.93
ATOM	708	CB	SER	99	-45.157	17.258	35.747	1.00	7.78
ATOM	709	CA	SER	99	-45.010	17.562	37.226	1.00	8.54
ATOM	710	C	SER	99	-43.921	18.659	37.391	1.00	7.90
ATOM	711	O	SER	99	-44.195	19.884	37.334	1.00	10.31
ATOM	712	N	VAL	100	-42.675	18.231	37.384	1.00	9.21
ATOM	713	CA	VAL	100	-41.468	19.082	37.505	1.00	5.59
ATOM	714	C	VAL	100	-40.375	18.343	36.773	1.00	5.35
ATOM	715	O	VAL	100	-40.380	17.108	36.785	1.00	9.03
ATOM	716	CB	VAL	100	-41.112	19.395	38.979	1.00	5.88
ATOM	717	CG1	VAL	100	-40.630	18.114	39.670	1.00	8.61

Fig. 2

Cont. Fig 2

ATOM	718	CG2	VAL	100	-40.142	20.579	39.127	1.00	5.24
ATOM	719	N	CYS	101	-39.423	19.055	36.168	1.00	5.24
ATOM	720	CA	CYS	101	-38.304	18.494	35.437	1.00	4.35
ATOM	721	C	CYS	101	-36.989	18.996	36.086	1.00	5.37
ATOM	722	O	CYS	101	-36.984	20.152	36.529	1.00	8.17
ATOM	723	CB	CYS	101	-38.312	18.824	33.935	1.00	5.99
ATOM	724	SG	CYS	101	-39.723	18.001	33.063	1.00	8.35
ATOM	725	N	ARG	102	-35.982	18.175	36.084	1.00	6.08
ATOM	726	CA	ARG	102	-34.649	18.527	36.587	1.00	7.57
ATOM	727	C	ARG	102	-33.605	18.534	35.462	1.00	8.71
ATOM	728	O	ARG	102	-33.604	17.647	34.598	1.00	8.00
ATOM	729	CB	ARG	102	-34.261	17.489	37.655	1.00	4.69
ATOM	730	CG	ARG	102	-32.859	17.784	38.286	1.00	3.27
ATOM	731	CD	ARG	102	-32.303	16.653	39.077	1.00	6.25
ATOM	732	NE	ARG	102	-32.250	15.402	38.343	1.00	7.04
ATOM	733	CZ	ARG	102	-31.471	15.083	37.294	1.00	9.62
ATOM	734	NH1	ARG	102	-31.692	13.864	36.773	1.00	9.55
ATOM	735	NH2	ARG	102	-30.608	15.953	36.805	1.00	5.91
ATOM	736	N	SER	103	-32.662	19.454	35.404	1.00	7.27
ATOM	737	CA	SER	103	-31.567	19.475	34.435	1.00	4.96
ATOM	738	C	SER	103	-30.259	19.365	35.228	1.00	6.42
ATOM	739	O	SER	103	-30.059	20.177	36.162	1.00	6.49
ATOM	740	CB	SER	103	-31.571	20.781	33.624	1.00	5.59
ATOM	741	OG	SER	103	-30.581	20.673	32.575	1.00	7.71
ATOM	742	N	GLY	104	-29.359	18.411	34.965	1.00	7.52
ATOM	743	CA	GLY	104	-28.071	18.282	35.661	1.00	5.92
ATOM	744	C	GLY	104	-27.031	17.745	34.686	1.00	6.41
ATOM	745	O	GLY	104	-27.354	17.083	33.665	1.00	7.09
ATOM	746	N	SER	105	-25.757	18.002	34.912	1.00	7.88
ATOM	747	OG	SER	105	-22.953	18.433	35.504	1.00	11.75
ATOM	748	CB	SER	105	-23.430	18.504	34.179	1.00	7.00
ATOM	749	CA	SER	105	-24.638	17.559	34.049	1.00	10.37
ATOM	750	C	SER	105	-24.235	16.085	34.102	1.00	10.06
ATOM	751	O	SER	105	-23.505	15.679	33.176	1.00	10.75
ATOM	752	N	THR	106	-24.719	15.248	35.018	1.00	9.74
ATOM	753	CA	THR	106	-24.403	13.811	35.029	1.00	9.92
ATOM	754	C	THR	106	-25.458	13.046	34.238	1.00	8.56
ATOM	755	O	THR	106	-25.079	12.174	33.464	1.00	10.42
ATOM	756	CB	THR	106	-24.322	13.103	36.435	1.00	10.71
ATOM	757	OG1	THR	106	-23.436	13.978	37.167	1.00	10.45
ATOM	758	CG2	THR	106	-22.782	11.671	36.508	1.00	6.76
ATOM	759	N	THR	107	-26.723	13.319	34.467	1.00	7.87
ATOM	760	CA	THR	107	-27.804	12.599	33.831	1.00	5.96
ATOM	761	C	THR	107	-28.634	13.349	32.838	1.00	7.88
ATOM	762	O	THR	107	-29.531	12.664	32.266	1.00	8.67
ATOM	763	CB	THR	107	-28.695	11.935	34.965	1.00	8.85
ATOM	764	OG1	THR	107	-29.241	13.086	35.695	1.00	7.24
ATOM	765	CG2	THR	107	-27.869	11.012	35.858	1.00	8.14
ATOM	766	N	GLY	108	-28.523	14.609	32.643	1.00	7.16
ATOM	767	CA	GLY	108	-29.341	15.323	31.611	1.00	8.94
ATOM	768	C	GLY	108	-30.668	15.756	32.223	1.00	10.50
ATOM	769	O	GLY	108	-30.722	16.175	33.436	1.00	8.98
ATOM	770	N	TRP	109	-31.691	15.642	31.408	1.00	6.52
ATOM	771	CD2	TRP	109	-35.413	18.684	30.968	1.00	5.39

Fig. 2

Cont. Fig 2

ATOM	772	CH3	TRP	109	-34.790	19.728	31.656	1.00	5.67
ATOM	773	CZ3	TRP	109	-35.588	20.864	31.900	1.00	10.11
ATOM	774	CH2	TRP	109	-36.931	20.950	31.533	1.00	7.46
ATOM	775	CZ2	TRP	109	-37.555	19.912	30.864	1.00	5.36
ATOM	776	CH2	TRP	109	-36.762	18.789	30.603	1.00	8.40
ATOM	777	NE1	TRP	109	-37.097	17.619	29.964	1.00	11.33
ATOM	778	CD1	TRP	109	-35.996	16.771	29.909	1.00	8.89
ATOM	779	CG	TRP	109	-34.928	17.409	30.506	1.00	7.87
ATOM	780	CB	TRP	109	-33.558	16.856	30.594	1.00	4.08
ATOM	781	CA	TRP	109	-33.070	16.106	31.803	1.00	6.13
ATOM	782	C	TRP	109	-34.031	15.013	32.176	1.00	7.55
ATOM	783	O	TRP	109	-34.244	14.017	31.372	1.00	8.28
ATOM	784	N	HIS	110	-34.566	15.040	33.380	1.00	6.29
ATOM	785	CD2	HIS	110	-32.409	12.573	33.819	1.00	8.99
ATOM	786	NE2	HIS	110	-31.845	11.558	33.121	1.00	10.03
ATOM	787	CE1	HIS	110	-32.751	10.595	32.981	1.00	12.94
ATOM	788	ND1	HIS	110	-33.856	11.027	33.562	1.00	12.68
ATOM	789	CG	HIS	110	-33.678	12.246	34.128	1.00	9.92
ATOM	790	CB	HIS	110	-34.761	13.004	34.840	1.00	8.94
ATOM	791	CA	HIS	110	-35.487	13.971	33.861	1.00	7.82
ATOM	792	C	HIS	110	-36.648	14.637	34.584	1.00	7.74
ATOM	793	O	HIS	110	-36.444	15.708	35.190	1.00	8.03
ATOM	794	N	CYS	111	-37.832	13.990	34.506	1.00	7.77
ATOM	795	CA	CYS	111	-39.052	14.589	35.065	1.00	7.64
ATOM	796	C	CYS	111	-39.864	13.660	35.952	1.00	9.31
ATOM	797	O	CYS	111	-39.559	12.451	35.928	1.00	11.03
ATOM	798	CB	CYS	111	-39.988	15.100	33.925	1.00	9.61
ATOM	799	SG	CYS	111	-39.150	16.153	32.711	1.00	9.12
ATOM	800	N	GLY	112	-40.828	14.245	36.638	1.00	8.81
ATOM	801	CA	GLY	112	-41.625	13.408	37.597	1.00	8.51
ATOM	802	C	GLY	112	-42.521	14.315	38.399	1.00	10.25
ATOM	803	O	GLY	112	-42.794	15.437	37.941	1.00	12.04
ATOM	804	N	THR	113	-42.979	13.969	39.595	1.00	10.08
ATOM	805	CA	THR	113	-43.870	14.809	40.372	1.00	8.62
ATOM	806	C	THR	113	-43.359	15.387	41.685	1.00	9.71
ATOM	807	O	THR	113	-42.441	14.786	42.268	1.00	10.48
ATOM	808	CB	THR	113	-45.340	13.996	40.648	1.00	17.08
ATOM	809	CG1	THR	113	-44.930	12.755	41.340	1.00	19.10
ATOM	810	CG2	THR	113	-46.004	13.705	39.362	1.00	15.24
ATOM	811	N	ILE	114	-43.963	16.488	42.073	1.00	8.35
ATOM	812	CA	ILE	114	-43.662	17.108	43.365	1.00	9.73
ATOM	813	C	ILE	114	-44.554	16.338	44.383	1.00	13.59
ATOM	814	O	ILE	114	-45.816	16.336	44.198	1.00	12.54
ATOM	815	CB	ILE	114	-44.008	18.621	43.384	1.00	10.58
ATOM	816	CG1	ILE	114	-43.089	19.319	42.341	1.00	12.00
ATOM	817	CG2	ILE	114	-43.864	19.215	44.814	1.00	12.84
ATOM	818	CD1	ILE	114	-43.555	20.750	42.065	1.00	10.62
ATOM	819	N	GLN	115	-43.977	15.668	45.379	1.00	12.39
ATOM	820	NE2	GLN	115	-43.951	10.378	45.759	1.00	31.00
ATOM	821	OE1	GLN	115	-42.096	11.415	46.407	1.00	34.76
ATOM	822	CD	GLN	115	-43.243	11.496	45.943	1.00	30.30
ATOM	823	CG	GLN	115	-43.993	12.758	45.524	1.00	18.83
ATOM	824	CB	GLN	115	-44.077	13.606	46.811	1.00	11.43
ATOM	825	CA	GLN	115	-44.732	14.936	46.396	1.00	11.89

Fig. 2

Cont. Fig 2

ATOM	826	C	GLN	115	-44.956	15.640	47.693	1.00	11.61
ATOM	827	O	GLN	115	-46.105	15.494	48.196	1.00	14.96
ATOM	828	N	ALA	116	-44.037	16.356	48.292	1.00	12.17
ATOM	829	CA	ALA	116	-44.291	16.976	49.618	1.00	10.65
ATOM	830	C	ALA	116	-43.263	18.055	49.863	1.00	9.88
ATOM	831	O	ALA	116	-42.162	17.978	49.326	1.00	13.41
ATOM	832	CB	ALA	116	-44.101	15.894	50.689	1.00	10.10
ATOM	833	N	ARG	117	-43.386	19.020	50.636	1.00	9.81
ATOM	834	NH2	ARG	117	-43.870	23.560	45.115	1.00	13.10
ATOM	835	NH1	ARG	117	-45.496	23.020	46.568	1.00	17.40
ATOM	836	CZ	ARG	117	-44.191	23.340	46.386	1.00	16.37
ATOM	837	NE	ARG	117	-43.268	23.465	47.330	1.00	14.91
ATOM	838	CD	ARG	117	-43.450	23.251	48.783	1.00	13.89
ATOM	839	CG	ARG	117	-43.323	21.801	49.092	1.00	10.78
ATOM	840	CE	ARG	117	-43.392	21.470	50.596	1.00	10.30
ATOM	841	CA	ARG	117	-42.725	20.137	50.983	1.00	10.62
ATOM	842	C	ARG	117	-42.465	20.023	52.496	1.00	12.42
ATOM	843	O	ARG	117	-43.122	19.229	53.201	1.00	14.36
ATOM	844	N	GLY	118	-41.566	20.803	52.999	1.00	10.94
ATOM	845	CA	GLY	118	-41.246	20.891	54.430	1.00	14.46
ATOM	846	C	GLY	118	-40.590	19.675	55.005	1.00	13.77
ATOM	847	O	GLY	118	-40.761	19.531	56.229	1.00	14.95
ATOM	848	N	GLN	119	-39.874	18.886	54.215	1.00	11.83
ATOM	849	NE2	GLN	119	-42.518	16.522	54.159	1.00	27.14
ATOM	850	OE1	GLN	119	-41.331	14.597	53.939	1.00	29.04
ATOM	851	CD	GLN	119	-41.505	15.779	53.683	1.00	26.05
ATOM	852	CG	GLN	119	-40.511	16.511	52.791	1.00	23.83
ATOM	853	CE	GLN	119	-39.161	16.633	53.499	1.00	15.28
ATOM	854	CA	GLN	119	-39.228	17.649	54.654	1.00	12.26
ATOM	855	C	GLN	119	-37.819	17.866	55.191	1.00	13.13
ATOM	856	O	GLN	119	-37.023	18.655	54.674	1.00	11.96
ATOM	857	N	SER	120	-37.520	17.136	56.272	1.00	12.77
ATOM	858	OG	SER	120	-36.874	16.674	59.024	0.50	13.43
ATOM	859	CB	SER	120	-36.074	17.591	58.330	0.50	15.06
ATOM	862	CA	SER	120	-36.147	17.234	56.873	1.00	11.79
ATOM	863	C	SER	120	-35.513	15.938	56.438	1.00	14.19
ATOM	864	O	SER	120	-36.167	14.855	56.352	1.00	14.17
ATOM	865	N	VAL	121	-34.228	16.035	56.037	1.00	14.53
ATOM	866	CG2	VAL	121	-34.392	15.445	53.235	1.00	20.02
ATOM	867	CG1	VAL	121	-32.537	13.814	53.494	1.00	22.53
ATOM	868	CB	VAL	121	-33.176	15.085	54.041	1.00	19.23
ATOM	869	CA	VAL	121	-33.466	14.920	55.565	1.00	14.60
ATOM	870	C	VAL	121	-32.106	14.892	56.248	1.00	17.85
ATOM	871	O	VAL	121	-31.399	15.890	56.335	1.00	16.78
ATOM	872	N	SER	122	-31.749	13.694	56.677	1.00	18.77
ATOM	873	OG	SER	122	-31.320	13.436	59.447	1.00	35.50
ATOM	874	CB	SER	122	-30.306	12.902	58.611	1.00	26.24
ATOM	875	CA	SER	122	-30.397	13.594	57.262	1.00	20.95
ATOM	876	C	SER	122	-29.504	12.911	56.238	1.00	20.78
ATOM	877	O	SER	122	-29.704	11.746	55.840	1.00	25.20
ATOM	878	N	TYR	123	-28.548	13.651	55.794	1.00	19.04
ATOM	879	CA	TYR	123	-27.479	13.164	54.927	1.00	21.72
ATOM	880	C	TYR	123	-26.475	12.613	55.993	1.00	25.29
ATOM	881	O	TYR	123	-26.521	13.015	57.187	1.00	26.53

Fig. 2

Cont. Fig 2

ATOM	882	CB	TYR	123	-26.981	14.342	54.135	1.00	18.93
ATOM	883	CG	TYR	123	-27.515	14.920	53.100	1.00	18.61
ATOM	884	CD1	TYR	123	-27.849	16.273	52.784	1.00	15.70
ATOM	885	CD2	TYR	123	-28.840	14.144	52.381	1.00	21.00
ATOM	886	CE1	TYR	123	-28.658	16.844	51.808	1.00	15.37
ATOM	887	CE2	TYR	123	-29.712	14.700	51.423	1.00	18.34
ATOM	888	CZ	TYR	123	-29.581	16.067	51.133	1.00	16.62
ATOM	889	OH	TYR	123	-30.390	16.593	50.172	1.00	15.25
ATOM	890	N	PRO	124	-25.578	11.721	55.647	1.00	28.48
ATOM	891	CG	PRO	124	-24.105	10.334	54.354	1.00	32.40
ATOM	892	CD	PRO	124	-25.391	11.155	54.297	1.00	30.53
ATOM	893	CE	PRO	124	-23.748	10.168	55.828	1.00	32.32
ATOM	894	CA	PRO	124	-24.583	11.183	56.598	1.00	31.44
ATOM	895	C	PRO	124	-23.732	12.285	57.226	1.00	32.68
ATOM	896	O	PRO	124	-23.355	12.124	58.408	1.00	33.64
ATOM	897	N	GLU	125	-23.417	13.329	56.485	1.00	32.09
ATOM	898	CA	GLU	125	-22.646	14.515	56.843	1.00	32.45
ATOM	899	C	GLU	125	-23.410	15.484	57.766	1.00	32.22
ATOM	900	O	GLU	125	-22.932	15.954	58.861	1.00	33.87
ATOM	901	CB	GLU	125	-22.144	15.220	55.588	1.00	28.10
ATOM	902	CG	GLU	125	-22.899	15.278	54.299	1.00	34.56
ATOM	903	CD	GLU	125	-23.341	14.605	53.384	0.00	53.35
ATOM	904	OE1	GLU	125	-23.517	15.100	52.231	0.00	59.96
ATOM	905	OE2	GLU	125	-23.156	13.352	53.591	0.00	58.46
ATOM	906	N	GLY	126	-24.666	15.793	57.397	1.00	31.75
ATOM	907	CA	GLY	126	-25.549	16.682	58.182	1.00	28.35
ATOM	908	C	GLY	126	-27.017	16.696	57.720	1.00	25.07
ATOM	909	O	GLY	126	-27.393	16.107	56.682	1.00	25.93
ATOM	910	N	THR	127	-27.811	17.417	58.486	1.00	19.30
ATOM	911	CG2	THR	127	-31.485	17.638	59.616	1.00	15.49
ATOM	912	OG1	THR	127	-29.548	16.190	60.199	1.00	22.50
ATOM	913	CB	THR	127	-29.964	17.487	59.664	1.00	14.93
ATOM	914	CA	THR	127	-29.242	17.558	58.256	1.00	13.91
ATOM	915	C	THR	127	-29.689	18.791	57.491	1.00	11.25
ATOM	916	O	THR	127	-29.203	19.885	57.803	1.00	10.59
ATOM	917	N	VAL	128	-30.649	18.626	56.584	1.00	11.18
ATOM	918	CA	VAL	128	-31.201	19.747	55.791	1.00	9.77
ATOM	919	C	VAL	128	-32.688	19.787	56.164	1.00	8.45
ATOM	920	O	VAL	128	-33.182	18.697	56.393	1.00	11.45
ATOM	921	CB	VAL	128	-30.956	19.633	54.298	1.00	7.25
ATOM	922	CG1	VAL	128	-29.466	19.790	54.013	1.00	11.55
ATOM	923	CG2	VAL	128	-31.377	18.285	53.735	1.00	9.36
ATOM	924	N	THR	129	-33.266	20.950	56.248	1.00	9.49
ATOM	925	CG2	THR	129	-34.182	21.202	59.125	1.00	18.71
ATOM	926	OG1	THR	129	-34.282	23.173	57.758	1.00	14.40
ATOM	927	CB	THR	129	-34.884	21.913	57.959	1.00	8.96
ATOM	928	CA	THR	129	-34.680	21.088	56.613	1.00	9.66
ATOM	929	C	THR	129	-35.407	21.804	58.487	1.00	10.11
ATOM	930	O	THR	129	-34.796	22.467	54.615	1.00	10.92
ATOM	931	N	ASN	130	-36.709	21.619	55.563	1.00	10.84
ATOM	932	ND2	ASN	130	-38.570	25.717	53.765	1.00	35.10
ATOM	933	OD1	ASN	130	-39.854	23.969	54.369	1.00	28.19
ATOM	934	CG	ASN	130	-38.739	24.527	54.362	1.00	30.11
ATOM	935	CE	ASN	130	-37.496	23.827	54.952	1.00	17.09

Fig. 2

Cont. Fig 2

ATOM	936	CA	ASN	130	-37.620	22.299	54.616	1.00	12.34
ATOM	937	C	ASN	130	-37.388	21.988	53.136	1.00	11.59
ATOM	938	O	ASN	130	-37.557	22.883	52.268	1.00	11.46
ATOM	939	N	MET	131	-37.056	20.742	52.880	1.00	9.76
ATOM	940	CB	MET	131	-33.020	20.248	50.187	1.00	16.57
ATOM	941	SD	MET	131	-33.597	21.082	51.529	1.00	20.94
ATOM	942	CG	MET	131	-34.411	19.573	52.379	1.00	9.59
ATOM	943	CE	MET	131	-35.664	19.199	51.580	1.00	5.73
ATOM	944	CR	MET	131	-36.732	20.297	51.529	1.00	9.93
ATOM	945	C	MET	131	-38.007	19.797	50.840	1.00	9.00
ATOM	946	O	MET	131	-38.962	19.372	51.519	1.00	10.48
ATOM	947	N	THR	132	-37.995	19.869	49.527	1.00	8.82
ATOM	948	CA	THR	132	-39.129	19.393	48.710	1.00	5.72
ATOM	949	C	THR	132	-38.769	18.054	48.108	1.00	7.44
ATOM	950	O	THR	132	-37.719	17.850	47.428	1.00	9.83
ATOM	951	CB	THR	132	-39.497	20.516	47.675	1.00	6.44
ATOM	952	CG1	THR	132	-39.851	21.700	48.434	1.00	9.35
ATOM	953	CG2	THR	132	-40.681	20.100	46.737	1.00	7.59
ATOM	954	N	ARG	133	-39.640	17.061	48.306	1.00	4.85
ATOM	955	NH2	ARG	133	-42.116	12.875	51.526	1.00	39.69
ATOM	956	NH1	ARG	133	-43.846	11.985	50.304	1.00	32.03
ATOM	957	CZ	ARG	133	-42.539	12.272	50.397	1.00	36.90
ATOM	958	NH	ARG	133	-41.666	11.948	49.418	1.00	28.98
ATOM	959	CD	ARG	133	-40.253	12.204	49.398	1.00	20.06
ATOM	960	CG	ARG	133	-39.832	13.218	48.354	1.00	11.77
ATOM	961	CE	ARG	133	-40.079	14.660	48.801	1.00	6.59
ATOM	962	CA	ARG	133	-39.443	15.674	47.833	1.00	6.76
ATOM	963	C	ARG	133	-40.092	15.457	46.455	1.00	7.46
ATOM	964	O	ARG	133	-41.227	15.963	46.241	1.00	9.38
ATOM	965	N	THR	134	-39.360	14.793	45.552	1.00	9.68
ATOM	966	CG2	THR	134	-39.353	17.060	43.521	1.00	7.44
ATOM	967	OG1	THR	134	-38.110	15.030	42.788	1.00	8.17
ATOM	968	CE	THR	134	-39.392	15.603	43.141	1.00	8.74
ATOM	969	CA	THR	134	-39.921	14.565	44.202	1.00	8.24
ATOM	970	C	THR	134	-39.576	13.135	43.785	1.00	8.06
ATOM	971	O	THR	134	-38.694	12.518	44.396	1.00	9.92
ATOM	972	N	THR	135	-40.301	12.622	42.770	1.00	7.24
ATOM	973	CG2	THR	135	-42.463	10.941	42.733	1.00	12.10
ATOM	974	OG1	THR	135	-41.763	11.403	40.497	1.00	11.13
ATOM	975	CE	THR	135	-41.362	10.601	41.650	1.00	10.58
ATOM	976	CA	THR	135	-40.037	11.294	42.228	1.00	8.03
ATOM	977	C	THR	135	-38.974	11.340	41.147	1.00	7.98
ATOM	978	O	THR	135	-38.731	10.327	40.471	1.00	9.10
ATOM	979	N	VAL	136	-38.326	12.452	40.852	1.00	10.38
ATOM	980	CG2	VAL	136	-38.483	14.754	39.130	1.00	6.65
ATOM	981	CG1	VAL	136	-36.094	14.300	38.374	1.00	10.08
ATOM	982	CE	VAL	136	-37.136	14.096	39.495	1.00	9.89
ATOM	983	CA	VAL	136	-37.296	12.595	39.813	1.00	9.10
ATOM	984	C	VAL	136	-35.990	11.927	40.235	1.00	9.23
ATOM	985	O	VAL	136	-35.640	12.113	41.414	1.00	11.00
ATOM	986	N	CYS	137	-35.273	11.188	39.408	1.00	7.90
ATOM	987	CA	CYS	137	-33.997	10.554	39.780	1.00	8.61
ATOM	988	C	CYS	137	-32.841	11.600	39.681	1.00	9.07
ATOM	989	O	CYS	137	-33.024	12.659	39.015	1.00	9.87

Fig. 2

Cont. Fig 2

ATOM	990	CB	CYS	137	-33.702	9.452	38.761	1.00	8.35
ATOM	991	SG	CYS	137	-33.425	10.013	37.031	1.00	11.68
ATOM	992	N	ALA	138	-31.687	11.330	40.300	1.00	7.09
ATOM	993	CB	ALA	138	-30.467	13.454	41.104	1.00	8.40
ATOM	994	CA	ALA	138	-30.503	12.207	40.220	1.00	8.25
ATOM	995	C	ALA	138	-29.294	11.352	40.568	1.00	8.58
ATOM	996	O	ALA	138	-29.409	10.287	41.211	1.00	10.28
ATOM	997	N	GLU	139	-28.105	11.803	40.178	1.00	9.17
ATOM	998	OE2	GLU	139	-27.880	7.597	37.175	1.00	10.81
ATOM	999	OE1	GLU	139	-25.854	8.348	37.135	1.00	10.87
ATOM	1000	CD	GLU	139	-26.946	8.270	37.673	1.00	12.62
ATOM	1001	CG	GLU	139	-27.289	9.053	38.912	1.00	10.79
ATOM	1002	CE	GLU	139	-26.414	10.298	39.196	1.00	8.05
ATOM	1003	CA	GLU	139	-26.818	11.099	40.431	1.00	9.02
ATOM	1004	C	GLU	139	-25.776	12.156	40.775	1.00	9.85
ATOM	1005	O	GLU	139	-25.966	13.362	40.507	1.00	9.59
ATOM	1006	N	PRO	140	-24.577	11.769	41.394	1.00	11.00
ATOM	1007	CG	PRO	140	-23.255	10.606	42.890	1.00	10.81
ATOM	1008	CD	PRO	140	-24.305	10.376	41.788	1.00	10.23
ATOM	1009	CE	PRO	140	-22.595	11.921	42.569	1.00	13.07
ATOM	1010	CA	PRO	140	-23.611	12.741	41.794	1.00	10.28
ATOM	1011	C	PRO	140	-23.095	13.561	40.611	1.00	9.93
ATOM	1012	O	PRO	140	-22.846	13.081	39.498	1.00	11.30
ATOM	1013	N	GLY	141	-23.015	14.865	40.884	1.00	7.17
ATOM	1014	CA	GLY	141	-22.596	15.857	39.927	1.00	7.59
ATOM	1015	C	GLY	141	-23.845	16.611	39.448	1.00	6.72
ATOM	1016	O	GLY	141	-23.742	17.715	38.907	1.00	7.49
ATOM	1017	N	ASP	142	-25.050	16.077	39.671	1.00	7.35
ATOM	1018	CA	ASP	142	-26.325	16.710	39.286	1.00	6.21
ATOM	1019	C	ASP	142	-26.662	17.752	40.369	1.00	5.40
ATOM	1020	O	ASP	142	-27.522	18.627	40.062	1.00	3.95
ATOM	1021	CE	ASP	142	-27.497	15.784	39.058	1.00	7.83
ATOM	1022	CG	ASP	142	-27.531	14.867	37.860	1.00	7.53
ATOM	1023	OD1	ASP	142	-28.075	13.736	37.908	1.00	7.92
ATOM	1024	OD2	ASP	142	-27.048	15.355	36.863	1.00	6.85
ATOM	1025	N	SER	143	-26.112	17.668	41.546	1.00	5.60
ATOM	1026	CA	SER	143	-26.422	18.670	42.589	1.00	5.95
ATOM	1027	C	SER	143	-26.287	20.098	42.102	1.00	7.22
ATOM	1028	O	SER	143	-25.328	20.407	41.346	1.00	7.03
ATOM	1029	CE	SER	143	-25.451	19.527	43.777	1.00	7.47
ATOM	1030	CG	SER	143	-25.764	17.255	44.327	1.00	11.26
ATOM	1031	N	GLY	144	-27.206	20.956	42.571	1.00	5.35
ATOM	1032	CA	GLY	144	-27.301	22.870	42.249	1.00	5.48
ATOM	1033	C	GLY	144	-28.051	22.665	40.945	1.00	6.00
ATOM	1034	O	GLY	144	-28.334	23.858	40.698	1.00	7.24
ATOM	1035	N	GLY	145	-28.295	21.671	40.140	1.00	5.00
ATOM	1036	CA	GLY	145	-28.959	21.818	38.828	1.00	5.08
ATOM	1037	C	GLY	145	-30.400	22.272	38.981	1.00	6.80
ATOM	1038	O	GLY	145	-31.096	22.093	40.013	1.00	8.09
ATOM	1039	N	SER	146	-31.016	22.823	37.953	1.00	4.98
ATOM	1040	CA	SER	146	-32.375	23.344	37.961	1.00	4.54
ATOM	1041	C	SER	146	-33.561	22.389	38.160	1.00	5.95
ATOM	1042	O	SER	146	-33.513	21.305	37.566	1.00	6.69
ATOM	1043	CE	SER	146	-32.609	23.870	36.500	1.00	4.69

Fig. 2

Cont. Fig 2

ATOM	1044	OG	SER	146	-31.487	24.440	35.945	1.00	7.80
ATOM	1045	N	TYR	147	-34.584	22.813	38.888	1.00	4.96
ATOM	1046	OH	TYR	147	-34.686	16.409	41.718	1.00	15.17
ATOM	1047	CD3	TYR	147	-36.809	19.212	40.614	1.00	8.25
ATOM	1048	CE2	TYR	147	-36.419	17.910	40.955	1.00	10.83
ATOM	1049	CE	TYR	147	-35.130	17.695	41.413	1.00	13.00
ATOM	1050	CE1	TYR	147	-34.196	18.722	41.524	1.00	8.41
ATOM	1051	CD1	TYR	147	-34.601	20.028	41.200	1.00	7.28
ATOM	1052	CG	TYR	147	-35.885	20.262	40.762	1.00	6.98
ATOM	1053	CB	TYR	147	-36.258	21.677	40.422	1.00	5.48
ATOM	1054	CA	TYR	147	-35.875	22.091	39.028	1.00	5.55
ATOM	1055	C	TYR	147	-36.829	23.167	38.409	1.00	6.41
ATOM	1056	O	TYR	147	-36.859	24.306	38.918	1.00	6.54
ATOM	1057	N	ILE	148	-37.559	22.845	37.365	1.00	5.67
ATOM	1058	CA	ILE	148	-38.454	23.821	36.710	1.00	6.64
ATOM	1059	C	ILE	148	-39.776	23.154	36.317	1.00	6.20
ATOM	1060	O	ILE	148	-39.743	21.993	35.834	1.00	6.60
ATOM	1061	CB	ILE	148	-37.661	24.408	35.475	1.00	8.25
ATOM	1062	CG1	ILE	148	-38.445	25.567	34.843	1.00	10.19
ATOM	1063	CG2	ILE	148	-37.269	23.309	34.443	1.00	11.02
ATOM	1064	CD1	ILE	148	-37.739	26.438	33.738	1.00	11.12
ATOM	1065	N	SER	149	-40.878	23.912	36.472	1.00	6.67
ATOM	1066	OG	SER	149	-42.874	22.350	38.136	1.00	15.36
ATOM	1067	CB	SER	149	-43.209	23.415	37.231	1.00	9.17
ATOM	1068	CA	SER	149	-42.219	23.396	36.098	1.00	7.08
ATOM	1069	C	SER	149	-42.712	24.364	35.000	1.00	7.85
ATOM	1070	O	SER	149	-43.087	25.496	35.341	1.00	8.98
ATOM	1071	N	GLY	150	-42.632	23.897	33.754	1.00	10.00
ATOM	1072	CA	GLY	150	-43.066	24.798	32.663	1.00	11.86
ATOM	1073	C	GLY	150	-41.990	25.883	32.482	1.00	8.97
ATOM	1074	O	GLY	150	-40.850	25.591	32.148	1.00	12.79
ATOM	1075	N	THR	151	-42.463	27.096	32.745	1.00	9.09
ATOM	1076	CG2	THR	151	-42.428	29.087	30.347	1.00	13.40
ATOM	1077	OG1	THR	151	-43.398	29.844	32.497	1.00	13.13
ATOM	1078	CB	THR	151	-42.170	29.426	31.816	1.00	11.88
ATOM	1079	CA	THR	151	-41.527	28.257	32.661	1.00	10.61
ATOM	1080	C	THR	151	-41.196	28.758	34.085	1.00	9.76
ATOM	1081	O	THR	151	-40.553	29.810	34.179	1.00	10.13
ATOM	1082	N	GLN	152	-41.628	28.099	35.157	1.00	6.87
ATOM	1083	CA	GLN	152	-41.440	28.600	36.494	1.00	7.20
ATOM	1084	C	GLN	152	-40.304	27.909	37.266	1.00	8.56
ATOM	1085	O	GLN	152	-40.488	26.681	37.461	1.00	12.38
ATOM	1086	CB	GLN	152	-42.770	28.493	37.286	1.00	6.03
ATOM	1087	CG	GLN	152	-43.935	29.238	36.607	1.00	9.11
ATOM	1088	CD	GLN	152	-43.668	30.705	36.459	1.00	7.44
ATOM	1089	OE1	GLN	152	-43.411	31.422	37.422	1.00	11.18
ATOM	1090	NE2	GLN	152	-43.686	31.229	35.247	1.00	15.32
ATOM	1091	N	ALA	153	-39.288	28.691	37.896	1.00	7.16
ATOM	1092	CA	ALA	153	-38.166	28.065	38.442	1.00	6.99
ATOM	1093	C	ALA	153	-38.696	27.573	39.786	1.00	6.78
ATOM	1094	O	ALA	153	-39.432	28.329	40.450	1.00	8.03
ATOM	1095	CB	ALA	153	-37.062	29.133	38.567	1.00	6.81
ATOM	1096	N	GLN	154	-38.383	26.349	40.198	1.00	4.41
ATOM	1097	CA	GLN	154	-38.827	25.782	41.460	1.00	6.90

Fig. 2

Cont. Fig 2

ATOM	1098	C	GLN	154	-37.692	25.687	42.494	1.00	7.22
ATOM	1099	O	GLN	154	-37.931	26.098	43.640	1.00	7.22
ATOM	1100	CB	GLN	154	-39.459	24.374	41.221	1.00	8.12
ATOM	1101	CG	GLN	154	-40.644	24.402	40.211	1.00	6.24
ATOM	1102	CD	GLN	154	-41.732	25.321	40.671	1.00	9.04
ATOM	1103	OE1	GLN	154	-42.271	25.234	41.795	1.00	9.56
ATOM	1104	NE2	GLN	154	-42.164	26.267	39.859	1.00	4.50
ATOM	1105	N	GLY	155	-36.547	25.153	42.078	1.00	7.56
ATOM	1106	CA	GLY	155	-35.475	24.954	43.098	1.00	7.31
ATOM	1107	C	GLY	155	-34.202	24.370	42.501	1.00	7.69
ATOM	1108	O	GLY	155	-34.029	24.335	41.280	1.00	7.17
ATOM	1109	N	VAL	156	-33.252	24.073	43.370	1.00	6.91
ATOM	1110	CA	VAL	156	-31.925	23.515	42.968	1.00	7.24
ATOM	1111	C	VAL	156	-31.760	22.136	43.631	1.00	5.91
ATOM	1112	O	VAL	156	-32.096	21.942	44.815	1.00	6.78
ATOM	1113	CB	VAL	156	-30.786	24.527	43.154	1.00	6.19
ATOM	1114	CG1	VAL	156	-31.048	25.862	42.407	1.00	5.61
ATOM	1115	CG2	VAL	156	-30.409	24.754	44.616	1.00	8.17
ATOM	1116	N	THR	157	-31.186	21.164	42.911	1.00	3.82
ATOM	1117	CG2	THR	157	-30.184	17.395	42.515	1.00	4.16
ATOM	1118	OG1	THR	157	-30.991	19.138	41.073	1.00	8.01
ATOM	1119	CB	THR	157	-30.243	18.893	42.293	1.00	3.37
ATOM	1120	CA	THR	157	-30.971	19.799	43.383	1.00	4.97
ATOM	1121	C	THR	157	-30.083	19.754	44.627	1.00	5.89
ATOM	1122	O	THR	157	-28.979	20.281	44.589	1.00	5.74
ATOM	1123	N	SER	158	-30.588	19.070	45.635	1.00	5.99
ATOM	1124	CA	SER	158	-29.830	18.914	46.876	1.00	7.64
ATOM	1125	C	SER	158	-29.316	17.473	46.969	1.00	10.26
ATOM	1126	O	SER	158	-28.087	17.229	47.132	1.00	10.04
ATOM	1127	CB	SER	158	-30.619	19.304	48.134	1.00	8.30
ATOM	1128	OG	SER	158	-29.853	18.975	49.296	1.00	9.39
ATOM	1129	N	GLY	159	-30.150	16.443	46.900	1.00	9.83
ATOM	1130	CA	GLY	159	-29.635	15.060	47.040	1.00	9.90
ATOM	1131	C	GLY	159	-30.756	14.048	47.006	1.00	12.82
ATOM	1132	O	GLY	159	-31.878	14.478	46.620	1.00	13.21
ATOM	1133	N	GLY	160	-30.510	12.792	47.352	1.00	12.00
ATOM	1134	CA	GLY	160	-31.646	11.846	47.257	1.00	12.56
ATOM	1135	C	GLY	160	-31.091	10.410	47.219	1.00	15.35
ATOM	1136	O	GLY	160	-29.988	10.197	47.741	1.00	15.70
ATOM	1137	N	SER	161	-31.869	9.497	46.679	1.00	13.10
ATOM	1138	OG	SER	161	-33.410	7.381	47.752	1.00	18.73
ATOM	1139	CB	SER	161	-31.996	7.324	47.933	1.00	16.84
ATOM	1140	CA	SER	161	-31.379	8.089	46.671	1.00	14.66
ATOM	1141	C	SER	161	-31.670	7.448	45.325	1.00	13.50
ATOM	1142	O	SER	161	-32.491	8.066	44.640	1.00	13.10
ATOM	1143	N	GLY	162	-31.078	6.310	45.040	1.00	11.84
ATOM	1144	CA	GLY	162	-31.318	5.641	43.734	1.00	12.53
ATOM	1145	C	GLY	162	-30.487	6.331	42.672	1.00	12.12
ATOM	1146	O	GLY	162	-29.545	7.088	42.957	1.00	11.79
ATOM	1147	N	ASN	163	-30.786	6.068	41.407	1.00	10.66
ATOM	1148	CA	ASN	163	-30.058	6.588	40.269	1.00	8.29
ATOM	1149	C	ASN	163	-31.033	6.671	39.088	1.00	11.02
ATOM	1150	O	ASN	163	-32.220	6.293	39.233	1.00	10.33
ATOM	1151	CB	ASN	163	-28.827	5.741	39.950	1.00	10.88

Fig. 2

Cont. Fig 2

ATOM	1152	CG	ASN	163	-29.238	4.312	39.578	1.00	14.23
ATOM	1153	OD1	ASN	163	-29.966	3.979	38.660	1.00	11.31
ATOM	1154	ND2	ASN	163	-28.649	3.362	40.334	1.00	18.46
ATOM	1155	N	CYS	164	-30.499	7.132	37.956	1.00	10.86
ATOM	1156	CA	CYS	164	-31.420	7.271	36.806	1.00	11.83
ATOM	1157	C	CYS	164	-31.687	6.006	35.998	1.00	13.69
ATOM	1158	O	CYS	164	-32.428	6.145	34.993	1.00	15.15
ATOM	1159	CB	CYS	164	-31.100	8.500	35.971	1.00	10.31
ATOM	1160	SG	CYS	164	-31.448	10.097	36.795	1.00	9.57
ATOM	1161	N	ARG	165	-31.110	4.919	36.364	1.00	12.68
ATOM	1162	NH2	ARG	165	-26.089	2.173	38.339	1.00	59.15
ATOM	1163	NH1	ARG	165	-25.617	4.096	37.196	1.00	58.35
ATOM	1164	CZ	ARG	165	-26.258	2.909	37.212	1.00	55.39
ATOM	1165	NE	ARG	165	-27.054	2.356	36.310	1.00	47.93
ATOM	1166	CD	ARG	165	-27.631	2.632	35.037	1.00	40.44
ATOM	1167	CG	ARG	165	-28.933	3.381	34.944	1.00	31.56
ATOM	1168	CB	ARG	165	-30.065	2.765	35.785	1.00	18.33
ATOM	1169	CA	ARG	165	-31.324	3.621	35.703	1.00	17.34
ATOM	1170	C	ARG	165	-32.498	2.928	36.433	1.00	14.57
ATOM	1171	O	ARG	165	-33.499	2.588	35.782	1.00	15.39
ATOM	1172	N	THR	166	-32.347	2.784	37.751	1.00	12.62
ATOM	1173	CG2	THR	166	-31.557	0.620	39.350	1.00	18.19
ATOM	1174	OG1	THR	166	-32.296	2.562	40.679	1.00	17.67
ATOM	1175	CB	THR	166	-32.716	1.474	39.795	1.00	16.75
ATOM	1176	CA	THR	166	-33.407	2.140	38.540	1.00	13.18
ATOM	1177	C	THR	166	-34.528	3.049	39.012	1.00	15.18
ATOM	1178	O	THR	166	-35.581	2.528	39.436	1.00	16.50
ATOM	1179	N	GLY	167	-34.296	4.347	39.040	1.00	13.17
ATOM	1180	CA	GLY	167	-35.253	5.345	39.536	1.00	13.42
ATOM	1181	C	GLY	167	-34.815	5.663	40.997	1.00	14.50
ATOM	1182	O	GLY	167	-33.957	4.993	41.596	1.00	13.35
ATOM	1183	N	GLY	168	-35.330	6.773	41.562	1.00	14.99
ATOM	1184	CA	GLY	168	-34.923	7.118	42.940	1.00	12.40
ATOM	1185	C	GLY	168	-35.852	8.241	43.371	1.00	15.29
ATOM	1186	O	GLY	168	-36.909	8.509	42.754	1.00	14.71
ATOM	1187	N	THR	169	-35.422	8.845	44.471	1.00	14.65
ATOM	1188	CG2	THR	169	-37.341	10.701	47.345	1.00	11.85
ATOM	1189	OG1	THR	169	-37.698	8.540	46.281	1.00	16.91
ATOM	1190	CB	THR	169	-36.711	9.582	46.529	1.00	14.83
ATOM	1191	CA	THR	169	-36.177	9.973	45.077	1.00	13.13
ATOM	1192	C	THR	169	-35.165	11.121	45.248	1.00	11.43
ATOM	1193	O	THR	169	-34.007	10.807	45.611	1.00	11.62
ATOM	1194	N	THR	170	-35.621	12.315	44.946	1.00	8.52
ATOM	1195	CG2	THR	170	-38.232	12.956	43.075	1.00	6.44
ATOM	1196	OG1	THR	170	-35.355	14.161	43.068	1.00	19.35
ATOM	1197	CB	THR	170	-34.090	13.942	43.798	1.00	12.63
ATOM	1198	CA	THR	170	-34.667	13.406	45.165	1.00	8.99
ATOM	1199	C	THR	170	-35.363	14.555	45.880	1.00	11.07
ATOM	1200	O	THR	170	-36.582	14.758	45.736	1.00	12.91
ATOM	1201	N	PHE	171	-34.531	15.291	46.609	1.00	9.07
ATOM	1202	CD2	PHE	171	-36.450	15.430	50.154	1.00	12.37
ATOM	1203	CE2	PHE	171	-37.017	14.281	50.750	1.00	12.72
ATOM	1204	CE	PHE	171	-36.332	13.053	50.719	1.00	14.13
ATOM	1205	CE1	PHE	171	-35.064	12.901	50.136	1.00	11.85

Fig. 2

Conf. Fig 2

ATOM	1206	CD1	PHE	171	-34.501	14.041	49.547	1.00	12.01
ATOM	1207	CG	PHE	171	-35.187	15.252	49.538	1.00	14.42
ATOM	1208	CB	PHE	171	-34.500	16.430	48.855	1.00	8.65
ATOM	1209	CA	PHE	171	-34.946	16.484	47.353	1.00	10.31
ATOM	1210	C	PHE	171	-34.276	17.746	46.736	1.00	9.69
ATOM	1211	O	PHE	171	-33.096	17.749	46.336	1.00	10.28
ATOM	1212	N	TYR	172	-35.022	18.818	46.721	1.00	6.76
ATOM	1213	OH	TYR	172	-40.405	22.042	43.789	1.00	10.33
ATOM	1214	CD2	TYR	172	-37.368	19.978	44.053	1.00	7.13
ATOM	1215	CE2	TYR	172	-38.680	20.393	43.754	1.00	9.56
ATOM	1216	CZ	TYR	172	-39.128	21.645	44.088	1.00	11.06
ATOM	1217	CE1	TYR	172	-38.255	22.544	44.740	1.00	7.89
ATOM	1218	CD1	TYR	172	-36.943	22.125	44.984	1.00	5.66
ATOM	1219	CG	TYR	172	-36.496	20.849	44.693	1.00	5.21
ATOM	1220	CB	TYR	172	-35.049	20.462	44.892	1.00	6.84
ATOM	1221	CA	TYR	172	-34.465	20.077	46.256	1.00	7.19
ATOM	1222	C	TYR	172	-34.711	21.217	47.245	1.00	6.71
ATOM	1223	O	TYR	172	-35.673	21.172	48.064	1.00	8.96
ATOM	1224	N	GLN	173	-33.781	22.185	47.157	1.00	5.53
ATOM	1225	NE2	GLN	173	-33.746	26.714	50.033	1.00	6.96
ATOM	1226	OE1	GLN	173	-32.289	25.169	50.891	1.00	10.23
ATOM	1227	CD	GLN	173	-32.795	25.753	49.933	1.00	9.42
ATOM	1228	CG	GLN	173	-32.411	25.493	48.482	1.00	3.75
ATOM	1229	CB	GLN	173	-32.463	24.037	48.054	1.00	8.35
ATOM	1230	CA	GLN	173	-33.883	23.428	47.962	1.00	5.17
ATOM	1231	C	GLN	173	-34.741	24.402	47.187	1.00	6.61
ATOM	1232	O	GLN	173	-34.469	24.693	45.967	1.00	7.86
ATOM	1233	N	GLU	174	-35.814	24.921	47.782	1.00	6.75
ATOM	1234	OE2	GLU	174	-40.122	25.837	46.396	1.00	8.25
ATOM	1235	OE1	GLU	174	-40.521	23.919	47.243	1.00	10.88
ATOM	1236	CD	GLU	174	-39.899	24.969	47.265	1.00	10.25
ATOM	1237	CG	GLU	174	-38.863	25.164	48.362	1.00	8.85
ATOM	1238	CB	GLU	174	-37.861	26.313	48.083	1.00	9.36
ATOM	1239	CA	GLU	174	-36.686	25.892	47.108	1.00	7.86
ATOM	1240	C	GLU	174	-35.933	27.178	46.774	1.00	8.69
ATOM	1241	O	GLU	174	-35.082	27.712	47.515	1.00	9.98
ATOM	1242	N	VAL	175	-36.198	27.769	45.591	1.00	8.61
ATOM	1243	CG2	VAL	175	-34.568	29.950	43.032	1.00	11.90
ATOM	1244	CG1	VAL	175	-36.989	29.512	42.968	1.00	13.78
ATOM	1245	CB	VAL	175	-35.652	29.062	43.589	1.00	13.52
ATOM	1246	CA	VAL	175	-35.605	29.022	45.144	1.00	8.03
ATOM	1247	C	VAL	175	-36.196	30.221	45.869	1.00	8.45
ATOM	1248	O	VAL	175	-35.453	31.199	46.161	1.00	8.96
ATOM	1249	N	THR	176	-37.454	30.220	46.287	1.00	9.06
ATOM	1250	CG2	THR	176	-40.371	32.282	47.869	1.00	18.26
ATOM	1251	CG1	THR	176	-40.343	30.692	46.091	1.00	18.82
ATOM	1252	CB	THR	176	-39.648	31.030	47.350	1.00	15.35
ATOM	1253	CA	THR	176	-38.133	31.355	46.965	1.00	10.08
ATOM	1254	C	THR	176	-37.370	32.053	48.082	1.00	12.75
ATOM	1255	O	THR	176	-37.203	33.295	48.105	1.00	12.78
ATOM	1256	N	PRO	177	-36.827	31.275	49.019	1.00	13.60
ATOM	1257	CA	PRO	177	-36.059	31.831	50.137	1.00	14.56
ATOM	1258	C	PRO	177	-34.832	32.550	49.634	1.00	14.53
ATOM	1259	O	PRO	177	-34.405	33.537	50.205	1.00	14.33

Fig. 2

Cont. Fig 2

ATOM	1260	CB	PRO	177	-35.684	30.599	50.967	1.00	15.59
ATOM	1261	CG	PRO	177	-36.607	29.488	50.587	1.00	15.55
ATOM	1262	CD	PRO	177	-37.028	29.818	49.153	1.00	14.17
ATOM	1263	N	MET	178	-34.177	32.085	48.557	1.00	11.26
ATOM	1264	CE	MET	178	-21.755	28.533	46.007	1.00	19.72
ATOM	1265	SD	MET	178	-30.708	29.927	46.237	1.00	22.73
ATOM	1266	CG	MET	178	-31.639	30.737	47.651	1.00	17.28
ATOM	1267	CB	MET	178	-32.343	31.919	46.980	1.00	9.92
ATOM	1268	CA	MET	178	-32.991	32.789	48.077	1.00	11.19
ATOM	1269	C	MET	178	-33.372	34.163	47.572	1.00	14.99
ATOM	1270	O	MET	178	-32.631	35.161	47.724	1.00	16.24
ATOM	1271	N	VAL	179	-34.492	34.273	46.870	1.00	12.66
ATOM	1272	CG2	VAL	179	-35.871	34.516	44.214	1.00	11.56
ATOM	1273	CG1	VAL	179	-36.995	36.540	45.154	1.00	16.31
ATOM	1274	CB	VAL	179	-36.242	35.254	45.495	1.00	14.06
ATOM	1275	CA	VAL	179	-34.973	35.549	46.347	1.00	14.10
ATOM	1276	C	VAL	179	-35.411	36.462	47.516	1.00	17.87
ATOM	1277	O	VAL	179	-35.152	37.674	47.535	1.00	18.01
ATOM	1278	N	ASN	180	-36.139	35.872	48.451	1.00	17.87
ATOM	1279	ND2	ASN	180	-39.996	35.004	49.792	1.00	34.66
ATOM	1280	OD1	ASN	180	-39.173	36.590	48.442	1.00	26.20
ATOM	1281	CG	ASN	180	-39.030	35.860	49.435	1.00	27.22
ATOM	1282	CB	ASN	180	-37.798	35.850	50.334	1.00	22.46
ATOM	1283	CA	ASN	180	-36.683	36.628	49.576	1.00	21.81
ATOM	1284	C	ASN	180	-35.663	37.122	50.588	1.00	22.21
ATOM	1285	O	ASN	180	-35.786	38.276	51.007	1.00	25.30
ATOM	1286	N	SER	181	-34.775	36.258	50.952	1.00	20.13
ATOM	1287	OG	SER	181	-34.362	34.891	53.477	1.00	46.23
ATOM	1288	CB	SER	181	-33.268	35.230	52.620	1.00	29.26
ATOM	1289	CA	SER	181	-33.740	36.530	51.921	1.00	20.62
ATOM	1290	C	SER	181	-32.474	37.109	51.355	1.00	20.10
ATOM	1291	O	SER	181	-31.914	37.915	52.104	1.00	20.08
ATOM	1292	N	TRP	182	-32.042	36.665	50.186	1.00	17.06
ATOM	1293	CD2	TRP	182	-28.744	33.771	49.582	1.00	20.91
ATOM	1294	CE3	TRP	182	-28.470	33.240	48.331	1.00	15.11
ATOM	1295	CZ3	TRP	182	-27.853	31.999	48.269	1.00	15.36
ATOM	1296	CH2	TRP	182	-27.529	31.313	49.436	1.00	15.65
ATOM	1297	CZ2	TRP	182	-27.769	31.806	50.713	1.00	19.64
ATOM	1298	CE2	TRP	182	-28.381	33.057	50.742	1.00	24.59
ATOM	1299	NE1	TRP	182	-28.798	33.820	51.828	1.00	27.77
ATOM	1300	CD1	TRP	182	-29.323	34.986	51.373	1.00	29.42
ATOM	1301	CG	TRP	182	-29.377	35.009	50.004	1.00	22.91
ATOM	1302	CB	TRP	182	-29.884	36.098	49.127	1.00	20.07
ATOM	1303	CA	TRP	182	-30.771	37.210	49.681	1.00	14.46
ATOM	1304	C	TRP	182	-30.992	38.306	48.680	1.00	14.45
ATOM	1305	O	TRP	182	-30.007	38.974	48.321	1.00	16.23
ATOM	1306	N	GLY	183	-32.203	38.445	48.182	1.00	13.61
ATOM	1307	CA	GLY	183	-32.431	39.519	47.179	1.00	13.59
ATOM	1308	C	GLY	183	-31.864	39.105	45.800	1.00	12.98
ATOM	1309	O	GLY	183	-31.478	40.015	45.005	1.00	11.51
ATOM	1310	N	VAL	184	-31.805	37.788	45.559	1.00	12.12
ATOM	1311	CG2	VAL	184	-31.741	34.892	44.375	1.00	20.41
ATOM	1312	CG1	VAL	184	-29.458	35.739	44.931	1.00	20.83
ATOM	1313	CB	VAL	184	-30.729	35.966	44.088	1.00	14.52

Fig. 2

ES 2 358 092 T3

LISTA DE SECUENCIAS

<110> Novozymes A/S

5 <120> Variantes de proteasa

<130> 10508

10 <160> 28

<170> PatentIn version 3.2

15 <210> 1

<211> 1596

<212> ADN

20 <213> *Nocardiosis* sp. NRRL 18262 (“Proteasa 10”)

<220>

<221> CDS

25 <222> (318)..(1463)

<220>

<221> sig-péptido

30 <222> (318)..(404)

<220>

<221> mat_péptido

35 <222> (900)..(1463)

40

45

50

55

60

65

ES 2 358 092 T3

<400> 1

5	acgttttgga cgggtaccgg tgtccgcgtg tggccagaat gcccccttgc gacaggggac	60
	ggattcgggc ggtagcgcat cgactccgac aaccgcgagg tggccgttcg cgtcgccacg	120
	ttctgcgacc gtcatgcgac ccacatcgg gtgacccac cgagctctga atggtccacc	180
10	gttctgacgg tcttccctc accaaaacgt gcacctatgg ttaggacgtt gttaccgaa	240
	tgtctcggtg aacgacaggg gccggacggg attcggcccc gatcccccggt tgcaccccc	300
	aggagagtag ggacccc atg cga ccc tcc ccc gtt gtc tcc gcc atc ggt	350
15	Met Arg Pro Ser Pro Val Val Ser Ala Ile Gly	
	-190 -185	
	acg gga gcg ctg gcc ttc ggt ctg gcg ctg tcc ggt acc ccg ggt	395
	Thr Gly Ala Leu Ala Phe Gly Leu Ala Leu Ser Gly Thr Pro Gly	
	-180 -175 -170	
20	gcc ctc gcg gcc acc gga gcg ctc ccc cag tca ccc acc ccg gag	440
	Ala Leu Ala Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu	
	-165 -160 -155	
	gcc gac gcg gtc tcc atg cag gag gcg ctc cag cgc gac ctc gac	485
25	Ala Asp Ala Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp	
	-150 -145 -140	
	ctg acc tcc gcc gag gcc gag gag ctg ctg gcc gcc cag gac acc	530
	Leu Thr Ser Ala Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr	
	-135 -130 -125	
30	gcc ttc gag gtc gac gag gcc gcg gcc gag gcc gcc ggg gac gcc	575
	Ala Phe Glu Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala	
	-120 -115 -110	
35	tac ggc ggc tcc gtc ttc gac acc gag agc ctg gaa ctg acc gtc ctg	623
	Tyr Gly Gly Ser Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu	
	-105 -100 -95	

ES 2 358 092 T3

	gtc acc gat gcc gcc gcc gtc gag gcc gtg gag gcc acc ggc gcc ggg	671
	Val Thr Asp Ala Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly	
	-90 -85 -80	
5	acc gag ctg gtc tcc tac ggc atc gac ggt ctc gac gag atc gtc cag	719
	Thr Glu Leu Val Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln	
	-75 -70 -65	
10	gag ctc aac gcc gcc gac gcc gtt ccc ggt gtg gtc ggc tgg tac ccg	767
	Glu Leu Asn Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro	
	-60 -55 -50 -45	
	gac gtg gcg ggt gac acc gtc gtc ctg gag gtc ctg gag ggt tcc gga	815
	Asp Val Ala Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly	
	-40 -35 -30	
15	gcc gac gtc agc ggc ctg ctc gcg gac gcc ggc gtg gac gcc tcg gcc	863
	Ala Asp Val Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala	
	-25 -20 -15	
20	gtc gag gtg acc acg agc gac cag ccc gag ctc tac gcc gac atc atc	911
	Val Glu Val Thr Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile	
	-10 -5 -1 1	
25	ggt ggt ctg gcc tac acc atg ggc ggc cgc tgt tcg gtc ggc ttc gcg	959
	Gly Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala	
	5 10 15 20	
	gcc acc aac gcc gcc ggt cag ccc ggg ttc gtc acc gcc ggt cac tgc	1007
	Ala Thr Asn Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys	
	25 30 35	
30	ggc cgc gtg ggc acc cag gtg acc atc ggc aac ggc agc ggc gtc ttc	1055
	Gly Arg Val Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe	
	40 45 50	
35	gag cag tcc gtc ttc ccc ggc aac gac gcg gcc ttc gtc cgc ggt acg	1103
	Glu Gln Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr	
	55 60 65	
	tcc aac ttc acg ctg acc aac ctg gtc agc cgc tac aac acc ggc ggg	1151
	Ser Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly	
	70 75 80	
40	tac gcc acg gtc gcc ggt cac aac cag gcc ccc atc ggc tcc tcc gtc	1199
	Tyr Ala Thr Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val	
	85 90 95 100	
45	tgc cgc tcc ggc tcc acc acc ggt tgg cac tgc ggc acc atc cag gcc	1247
	Cys Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala	
	105 110 115	
	cgc ggc cag tcg gtg agc tac ccc gag ggc acc gtc acc aac atg acc	1295
	Arg Gly Gln Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr	
	120 125 130	
50	cgg acc acc gtg tgc gcc gag ccc ggc gac tcc ggc ggc tcc tac atc	1343
	Arg Thr Thr Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile	
	135 140 145	
55	tcc ggc acc cag gcc cag ggc gtg acc tcc ggc ggc tcc ggc aac tgc	1391
	Ser Gly Thr Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys	
	150 155 160	
60	cgc acc ggc ggg acc acc ttc tac cag gag gtc acc ccc atg gtg aac	1439
	Arg Thr Gly Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn	
	165 170 175 180	
65	tcc tgg ggc gtc cgt ctc cgg acc tgatccccgc ggttcaggc ggaccgacgg	1493

ES 2 358 092 T3

Ser Trp Gly Val Arg Leu Arg Thr
185

tcgtgacctg agtaccaggc gtccccgccg cttccagcgg cgtccgcacc ggggtgggac 1553

cgggcgtggc caccgcccc cccgtgaccg gaccgcccgg cta 1596

<210> 2

10 <211> 382

<212> PRT

<213> *Nocardiopsis* sp. NRRL 18262 ("Proteasa 10")

15 <400> 2

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

20 Ala Phe Gly Leu Ala Leu Ser Gly Thr Pro Gly Ala Leu Ala Ala
-175 -170 -165

25 Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
-160 -155 -150

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala
-145 -140 -135

30 Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val
-130 -125 -120

35 Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser
-115 -110 -105

40 Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala
-100 -95 -90

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

45 Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala
-70 -65 -60

50 Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly
-55 -50 -45

Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser
-40 -35 -30 -25

55 Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr
-20 -15 -10

60 Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala
-5 -1 1 5

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

65

ES 2 358 092 T3

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

40 <210> 3
 <211> 1065
 <212> ADN
 45 <213> *Nocardiopsis dassonvillei subespecie dassonvillei* DSM 43235 ("Proteasa 18")

<220>
 <221> CDS
 50 <222> (1)..(1062)

<220>
 <221> mat_péptido
 55 <222> (499)..(1062)

<400> 3
 60 gct ccg gcc ccc gtc ccc cag acc ccc gtc gcc gac gac agc gcc 45
 Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala
 -165 -160 -155
 gcc agc atg acc gag gcg ctc aag cgc gac ctc gac ctc acc tcg 90
 65 Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser
 -150 -145 -140

ES 2 358 092 T3

	gcc gag gcc gag gag ctt ctc tcg gcg cag gaa gcc gcc atc gag	135
	Ala Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu	
	-135 -130 -125	
5	acc gac gcc gag gcc acc gag gcc gcg ggc gag gcc tac ggc ggc	180
	Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly	
	-120 -115 -110	
10	tca ctg ttc gac acc gag acc ctc gaa ctc acc gtg ctg gtc acc gac	228
	Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp	
	-105 -100 -95	
15	gcc tcc gcc gtc gag gcg gtc gag gcc acc gga gcc cag gcc acc gtc	276
	Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val	
	-90 -85 -80 -75	
	gtc tcc cac ggc acc gag ggc ctg acc gag gtc gtg gag gac ctc aac	324
	Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn	
	-70 -65 -60	
20	ggc gcc gag gtt ccc gag agc gtc ctc ggc tgg tac ccg gac gtg gag	372
	Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu	
	-55 -50 -45	
25	agc gac acc gtc gtg gtc gag gtg ctg gag ggc tcc gac gcc gac gtc	420
	Ser Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val	
	-40 -35 -30	
	gcc gcc ctg ctc gcc gac gcc ggt gtg gac tcc tcc tcg gtc cgg gtg	468
	Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val	
	-25 -20 -15	
30	gag gag gcc gag gag gcc ccg cag gtc tac gcc gac atc atc ggc ggc	516
	Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly	
	-10 -5 -1 1 5	
35	ctg gcc tac tac atg ggc ggc cgc tgc tcc gtc ggc ttc gcc gcg acc	564
	Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr	
	10 15 20	
40	aac agc gcc ggt cag ccc ggt ttc gtc acc gcc ggc cac tgc ggc acc	612
	Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr	
	25 30 35	
	gtc ggc acc ggc gtg acc atc ggc aac ggc acc ggc acc ttc cag aac	660
	Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn	
	40 45 50	
45	tcg gtc ttc ccc ggc aac gac gcc gcc ttc gtc cgc ggc acc tcc aac	708
	Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn	
	55 60 65 70	
50	ttc acc ctg acc aac ctg gtc tcg cgc tac aac tcc ggc ggc tac cag	756
	Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln	
	75 80 85	
	tcg gtg acc ggt acc agc cag gcc ccg gcc ggc tcg gcc gtg tgc cgc	804
	Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg	
	90 95 100	
55	tcc ggc tcc acc acc ggc tgg cac tgc ggc acc atc cag gcc cgc aac	852
	Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn	
	105 110 115	
60	cag acc gtg cgc tac ccg cag ggc acc gtc tac tcg ctc acc cgc acc	900
	Gln Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr	
	120 125 130	

ES 2 358 092 T3

	aac gtc tgc gcc gag ccc ggc gac tcc ggc ggt tgc ttc atc tcc ggc	948
	Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly	
	135 140 145 150	
5	tcg cag gcc cag ggc gtc acc tcc ggc ggc tcc ggc aac tgc tcc gtc	996
	Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Val	
	155 160 165	
10	ggc ggc acg acc tac tac cag gag gtc acc ccg atg atc aac tcc tgg	1044
	Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp	
	170 175 180	
15	ggt gtc agg atc cgg acc taa	1065
	Gly Val Arg Ile Arg Thr	
	185	

<210> 4

<211> 354

<212> PRT

<213> *Nocardiopsis dassonvillei subespecie dassonvillei* DSM 43235 ("Proteasa 18")

<400> 4

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

ES 2 358 092 T3

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

<210> 5
 45 <211> 1062
 <212> ADN
 <213> *Nocardiosis prasina* DSM 15648 ("Proteasa 11")
 50 <220>
 <221> CDS
 <222> (1)..(1059)
 55 <220>
 <221> mat_péptido
 <222> (496)..(1059)
 60 <400> 5

gac acc gga ccg ctc ccc cag tca ccc acc ccg gag gcc gac gcc	45
Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala	
-165 -160 -155	
gtc tcc atg cag gag gcc ctc cag cgc gac ctc ggc ctg acc ccg	90
Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro	

ES 2 358 092 T3

	-150	-145	-140	
5	ctt gag gcc gat gaa ctg ctg gcc gcc cag gac acc gcc ttc gag Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu -135 -130 -125	135		
	gtc gac gag gcc gcg gcc gcg gcc gcc ggg gac gcc tac ggc ggc Val Asp Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly -120 -115 -110	180		
10	tcc gtc ttc gac acc gag acc ctg gaa ctg acc gtc ctg gtc acc gac Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp -105 -100 -95	228		
15	gcc gcc tcg gtc gag gct gtc gag gcc acc ggc gcg ggt acc gaa ctc Ala Ala Ser Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu -85 -80 -75	276		
20	gtc tcc tac ggc atc gag ggc ctc gac gag atc atc cag gat ctc aac Val Ser Tyr Gly Ile Glu Gly Leu Asp Glu Ile Ile Gln Asp Leu Asn -70 -65 -60	324		
	gcc gcc gac gcc gtc ccc ggc gtc gtc ggc tgg tac ccg gac gtc gcg Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala -55 -50 -45	372		
25	ggt gac acc gtc gtc ctg gag gtc ctg gag ggt tcc gga gcc gac gtc Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val -40 -35 -30	420		
30	agc ggc ctg ctc gcc gac gcc ggc gtc gac gcc tcg gcc gtc gag gtc Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val -25 -20 -15 -10	468		
	acc agc agt ggc cag ccc gag ctc tac gcc gac atc atc ggc ggt ctg Thr Ser Ser Ala Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu -5 -1 1 5	516		
35	gcc tac acc atg ggc ggc cgc tgt tcg gtc gga ttc gcg gcc acc aac Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Thr Asn 10 15 20	564		
40	gcc gcc ggt cag ccc gga ttc gtc acc gcc ggt cac tgt ggc cgc gtc Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val 25 30 35	612		
	ggc acc cag gtc agc atc ggc aac ggc cag ggc gtc ttc gag cag tcc Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser 40 45 50 55	660		
45	atc ttc ccg ggc aac gac gcc gcc ttc gtc cgc ggc acg tcc aac ttc Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe 60 65 70	708		
50	acg ctg acc aac ctg gtc agc cgc tac aac acc ggc ggt tac gcc acc Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr 75 80 85	756		
55	gtc gcc ggc cac aac cag gcg ccc atc ggc tcc tcc gtc tgc cgc tcc Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser 90 95 100	804		
	ggc tcc acc acc ggc tgg cac tgc ggc acc atc cag gcc cgc ggc cag Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln 105 110 115	852		
60	tcg gtc agc tac ccc gag ggc acc gtc acc aac atg acc cgg acc acc Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr 120 125 130 135	900		
65				

ES 2 358 092 T3

5

10

20

25

30

35

40

45

45

50

55

60

65

ES 2 358 092 T3

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

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

10 Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser
40 45 50 55

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

15 Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
75 80 85

20 Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
105 110 115

25 Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
120 125 130 135

30 Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
155 160 165

35 Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
170 175 180

Val Arg Leu Arg Thr
185

40 <210> 7

<211> 1062

<212> ADN

45 <213> *Nocardiopsis prasina* DSM 15649 ("Proteasa 35")

<220>

<221> CDS

50 <222> (1)..(1059)

<220>

<221> mat_péptido

55 <222> (496)..(1059)

<400> 7

60 gcc acc gga cca ctc ccc cag tca ccc acc ccg gag gcc gac gcc 45
Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
-165 -160 -155

65 gtc tcc atg cag gag gcg ctc cag cgc gac ctc ggc ctg acc ccg 90
Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro

ES 2 358 092 T3

	-150	-145	-140	
5	ctt gag gcc gat gaa ctg ctg gcc gcc cag gac acc gcc ttc gag Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu -135 -130 -125	135		
10	gtc gac gag gcc gcg gcc gag gcc gcc ggt gac gcc tac gcc gcc Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly -120 -115 -110	180		
15	tcc gtc ttc gac acc gag acc ctg gaa ctg acc gtc ctg gtc acc gac Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp -105 -100 -95	228		
20	tcc gcc gcg gtc gag gcc gtc gag gcc acc gcc gcc gcc acc gaa ctg Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu -85 -80 -75	276		
25	gtc tcc tac gcc atc acg gcc ctc gac gag atc gtc gag gag ctc aac Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn -70 -65 -60	324		
30	gcc gcc gac gcc gtt ccc gcc gtc gtc gcc tgg tac ccg gac gtc gcc Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala -55 -50 -45	372		
35	ggt gac acc gtc gtc ctg gag gtc ctg gag ggt tcc gcc gcc gac gtc Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val -40 -35 -30	420		
40	ggc gcc ctg ctc gcc gac gcc gcc gtc gag gac gcc tcc gcc gtc gag gtc Gly Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val -25 -20 -15 -10	468		
45	acc acc acc gag cag ccc gag ctg tac gcc gac atc atc gcc ggt ctg Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu -5 -1 1 5	516		
50	gcc tac acc atg gcc gcc cgc tgt tcc gtc gcc ttc gcc gcc acc aac Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn 10 15 20	564		
55	gcc gcc ggt cag ccc gcc ttc gtc acc gcc ggt cac tgt gcc cgc gtc Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val 25 30 35	612		
60	ggc acc cag gtc acc atc gcc aac gcc cgg gcc gtc ttc gag cag tcc Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser 40 45 50 55	660		
65	atc ttc ccg gcc aac gac gcc gcc ttc gtc cgc gga acg tcc aac ttc Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe 60 65 70	708		
70	acg ctg acc aac ctg gtc agc cgc tac aac acc gcc gcc tac gcc acc Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr 75 80 85	756		
75	gtc gcc ggt cac aac cag gcc ccc atc gcc tcc tcc gtc tgc cgc tcc Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser 90 95 100	804		
80	ggc tcc acc acc ggt tgg cac tgc gcc acc atc cag gcc cgc gcc cag Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln 105 110 115	852		
85	tcc gtc agc tac ccc gag gcc acc gtc acc aac atg acc cgg acc acc Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr 120 125 130 135	900		

ES 2 358 092 T3

gtg tgc gcc gag ccc ggc gac tcc ggc ggc tcc tac atc tcc ggc aac 948
Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn 150
140 145 150

5 cag gcc cag ggc gtc acc tcc ggc ggc tcc ggc aac tgc cgc acc ggc 996
Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly 165
155 160 165

10 ggg acc acc ttc tac cag gag gtc acc ccc atg gtc aac tcc tgg ggc 1044
Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly 180
170 175 180

gtc cgt ctc cgg acc taa 1062
Val Arg Leu Arg Thr 185

<210> 8

<211> 353

20 <212> PRT

<213> *Nocardiopsis prasina* DSM 15649 ("Proteasa 35")

<400> 8

25

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
-165 -160 -155

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

35 Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
-135 -130 -125

40 Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly
-120 -115 -110

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

50 Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
-85 -80 -75

55 Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn
-70 -65 -60

60 Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
-55 -50 -45

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

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

Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
-5 -1 1 5

ES 2 358 092 T3

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

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
5 25 30 35

Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser
10 40 45 50 55

Ile 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 Thr Gly Gly Tyr Ala Thr
15 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
20 90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
25 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
30 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
35 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
170 175 180

Val Arg Leu Arg Thr
185

<210> 9

45 <211> 1068

<212> ADN

<213> *Nocardiopsis alba* DSM 15647 ("Proteasa 08")

50 <220>

<221> CDS

<222> (1)..(1065)

55 <220>

<221> mat_péptido

<222> (502)..(1065)

60 <400> 9

ggc acc ggc ccc ctc ccc cag tcc ccc acc ccg gat gaa gcc gag 45
Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu
-165 -160 -155

ggc acc acc atg gtc gag gcc ctc cag cgc gac ctc ggc ctg tcc 90
Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser

ES 2 358 092 T3

		-150		-145		-140			
5		ccc tct cag gcc gac gag ctc ctc gag gcg cag gcc gag tcc ttc	135						
		Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe							
		-135		-130		-125			
10		gag atc gac gag gcc gcc acc gcg gcc gca gcc gac tcc tac gcc	180						
		Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Asp Ser Tyr Gly							
		-120		-115		-110			
15		ggc tcc atc ttc gac acc gac agc ctc acc ctg acc gtc ctg gtc acc	228						
		Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr							
		-105		-100		-95			
20		gac gcc tcc gcc gtc gag gcg gtc gag gcc gcc gcc gcc gag gcc aag	276						
		Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys							
		-90		-85		-80			
25		gtg gtc tcg cac gcc atg gag gcc ctg gag gag atc gtc gcc gac ctg	324						
		Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu							
		-75		-70		-65			
30		aac gcg gcc gac gct cag ccc gcc gtc gtg gcc tgg tac ccc gac atc	372						
		Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile							
		-55		-50		-45			
35		cac tcc gac acg gtc gtc ctc gag gtc ctc gag gcc tcc ggt gcc gac	420						
		His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp							
		-40		-35		-30			
40		gtg gac tcc ctg ctc gcc gac gcc ggt gtg gac acc gcc gac gtc aag	468						
		Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys							
		-25		-20		-15			
45		gtg gag agc acc acc gag cag ccc gag ctg tac gcc gac atc atc gcc	516						
		Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Tyr Ala Asp Ile Ile Gly							
		-10		-5		-1			
50		ggt ctc gcc tac acc atg ggt ggg cgc tgc tgc gtc gcc ttc gcg gcc	564						
		Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala							
		10		15		20			
55		acc aac gcc tcc gcc cag ccc ggg ttc gtc acc gcc gcc cac tgc gcc	612						
		Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly							
		25		30		35			
60		acc gtc gcc acc ccg gtc agc atc gcc aac gcc cag gcc gtc ttc gag	660						
		Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu							
		40		45		50			
65		cgt tcc gtc ttc ccc gcc aac gac tcc gcc ttc gtc cgc gcc acc tcg	708						
		Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser							
		55		60		65			
70		aac ttc acc ctg acc aac ctg gtc agc cgc tac aac acc ggt ggt tac	756						
		Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr							
		70		75		80			
75		gcg acc gtc tcc gcc tcc tcg cag gcg gcg atc gcc tcg cag atc tgc	804						
		Ala Thr Val Ser Gly Ser Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys							
		90		95		100			
80		cgt tcc gcc tcc acc acc gcc tgg cac tgc gcc acc gtc cag gcc cgc	852						
		Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg							
		105		110		115			
85		ggc cag acg gtg agc tac ccc cag gcc acc gtg cag aac ctg acc cgc	900						
		Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg							
		120		125		130			

ES 2 358 092 T3

acc aac gtc tgc gcc gag ccc ggt gac tcc ggc ggc tcc ttc atc tcc 948
 Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser
 135 140 145
 5 ggc agc cag gcc cag ggc gtc acc tcc ggt ggc tcc ggc aac tgc tcc 996
 Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser
 150 155 160 165
 10 ttc ggt ggc acc acc tac tac cag gag gtc aac ccg atg ctg agc agc 1044
 Phe Gly Gly Thr Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser
 170 175 180
 15 tgg ggt ctg acc ctg cgc acc tga 1068
 Trp Gly Leu Thr Leu Arg Thr
 185

<210> 10
 <211> 355
 20 <212> PRT
 <213> *Nocardiopsis alba* DSM 15647 ("Proteasa 08")

<400> 10

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

ES 2 358 092 T3

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

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

10 Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu
40 45 50

15 Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser
55 60 65

Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr
70 75 80 85

20 Ala Thr Val Ser Gly Ser Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys
90 95 100

25 Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg
105 110 115

Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg
120 125 130

30 Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser
135 140 145

35 Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser
150 155 160 165

Phe Gly Gly Thr Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser
170 175 180

40 Trp Gly Leu Thr Leu Arg Thr
185

45 <210> 11
<211> 27
<212> ADN
50 <213> Artificial

<220>
55 <223> Cebador

<400> 11

ccgattatgg agcggattga acatgcg

60 <210> 12
<211> 31
<212> ADN
65 <213> Artificial

ES 2 358 092 T3

<220>
 <223> Cebador

5 <400> 12

gtgaccatcg gcgacggcag gggcgtcttc g

10 <210> 13
 <211> 10172
 <212> ADN
 <213> Artificial

15 <220>
 <223> Cassette de expresión

20 <220>
 <221> fuente
 <222> (1)..(3323)
 <223> Secuencia de genoma de *Bacillus subtilis* incluyendo genes yfmH-yfmD-yfmC-yfmB-yfmA

25 <220>
 <221> misc_recomb
 <222> (3561)..(4208)
 <223> Gen cat que proporciona resistencia al cloranfenicol

30 <220>
 <221> promotor
 <222> (4523)..(5633)
 <223> Triple promotor PamyL-scBAN-CryIIIa incluyendo la secuencia estabilizante mRND

35 <220>
 <221> péptido señal
 <222> (5658)..(5738)

40 <220>
 <221> péptido señal
 <222> (5658)..(5738)

45 <220>
 <221> CDS
 <222> (5658)..(6797)

50 <220>
 <221> mat_péptido
 <222> (6234)..(6797)

55 <220>
 <221> fuente
 <222> (6839)..(7540)
 <223> Parte de gen de pectato liasa de *Bacillus subtilis*

60 <220>
 <221> fuente
 <222> (7541)..(10172)
 <223> ADN de genoma de *Bacillus subtilis* incluidos genes de yfls-citM

65

31

ES 2 358 092 T3

<400> 13

5	gagtatcgcc agtaaggggc gtttttggtt tctggttgtt ttcttcattt caggtttcgc	60
	ccttttccttg ccaaatataa gaaaacggc gttccgataa tcgcggtgacaatgccgacc	120
	ggtgattcat aaggaaatgc aatccatctg gccagaacat ctgcgtacac cagcnaaatg	180
10	gcaccgaaca gtgccgaaaa cggaaacacg tattgataat gttctccgat cagcttgagg	240
	acaatatgag ggacgagcag cccgacaag ccaatcggcc cggcgacggc tacgnaagcg	300
	ccggaagaa ttaaaataat caaactgac agaatcctga tgccgttcat attttgtcca	360
15	agcccttttg ctgtttcgtc tccgagaccg agaacagaaa cagaaccgga aatcacgagg	420
	gcaagcccgatgccaatgac agaaaaagga gcgatggta tgacgtcctg ccagttgctg	480
	ccgtcgattg cgcctgtcat ccagtcaga acatcctcac ctgactcatt taaaataatg	540
20		
25		
30		
35		
40		
45		
50		
55		
60		
65		

ES 2 358 092 T3

atggcctgtg tcatagagga gaggaacaag tgcacggcca ttcttgacag cgccagcttg 600
 acaggcgta ttcgcgcgga tgaggcaatc atatacacia tcgcgcgcgc tgctgccgca 660
 5 cccgcanaag cgaatataac agatgaatag ggcgatgccg gcagaatgac gagagaagca 720
 acaacaaaaa gcgatgcacc cgcattcaca ccgaaaattt ggggtgaagc cagaggattt 780
 ctggtcatag cctgcatcag cgcacctgtc acagctaggc tggcgccgac aaaaacgccg 840
 10 attaatgtgc ggggaaggcg aagagtagag atgatgagct gttcctttga accgtcccat 900
 acaaaaagat atttcaatga atctatgatg ctgatgtctg aggcctctac tgaagattc 960
 agcccaagcc caaatataaa aatmatcagt gcaatgataa acatcatcag tcttgatgat 1020
 15 gagcgccgtt tggctgaatg atacaacagt ctcacttctt tactgcgtct ggttgcaaaa 1080
 acgaagaagc aaggattccc ctgccttctc atttgtccta tttattatac acttttttaa 1140
 gcacatcttt ggcgcttggt tcaactagact tgatgcctct gaatcttgct caagtgtcac 1200
 20 ggtccgcata atagacttgt ccatttttca ccgctttgag atttttccag agcgggttcg 1260
 ttttccactc atctacaatg gttttgcctt cgttggctga gatgaacaaa atatcaggat 1320
 25 cgattttgct caattgctca aggcagacct cttgataggc gttatctgac ttcacagcgt 1380
 gtgtaaagcc tagcatttta aagattttct cgtcatagga tgatgatgta tgaagctgga 1440
 aggaatccgc tcttgcaacg ccgagaacga tgttgcgggt ttcattcttc ggaagtccg 1500
 30 ctttttagatc gttgatgact tttttgtgct cggcaagcct ttcttttctt tcatcttctt 1560
 tatttaaatgc ttttagcaatg gtcgtaaagc tgtcgatcgt ttcgtcatat gtcgcttcac 1620
 ggctttttta ttcantcgtc ggggcgattt ttttcagctg tttatnaatg tttttatggc 1680
 35 gctcagcgtc agcgatgatt aaatcaggct tcaaggaaact gatgacctca agattgggtt 1740
 cgctgcgtgt gcctacagat gtgtaatcaa tggagctgcc gacaagcttt ttaatcatat 1800
 40 cttttttggt gtcactctgc atgcccaccg gcgtaatgcc gagattgtga acggcatcca 1860
 agaatgaag ctcaagcaca accaccgct taggtgtgcc gcttactgtc gtttttctt 1920
 cttcgtcatg gatcactctg gaatccttag actcgtttt gccgcttccg ttgttattct 1980
 45 ggcttgatga acagccggat acaatgaggc aggcgagcaa taaaacactc atgatggcaa 2040
 tcaacttggt agaatagggt cgcattgtcat tcttctttt ttcagattta gtaatgagaa 2100
 tcattatcac atgtaacact ataatagcat ggcttatcat gtcaatattt ttttagtaaa 2160
 50 gaaagctgcg tttttactgc tttctcatga aagcatcatc agacacaaat aagtggatg 2220
 cagcgttacc gtgtcttcga gacaaaaacg catgggcgtt ggcttttagag gtttcgaaca 2280
 tatcagcagt gacataagga aggagagtgc tgagataacc ggacaatttc ttttctattt 2340
 55 catctgttag tgcaaatcca atgtcgccga tattcatgat aatcgagaaa acaagtcga 2400
 tatcgatatg aaaatgttcc tcggcaaaaa ccgcaagctc gtgaatttct ggtgaacatc 2460
 60 cggcacgctt atggaaaatc tgtttgacta aatcactcac aatccaagca ttgtattgct 2520
 gtcttggtga aaagtattgc attagacata cctcctgctc gtaaggataa aggcagcgtt 2580

ES 2 358 092 T3

	tcattggtcgt	gtgctccgtg	cagcggcctc	tccttaattt	tgatttttct	gaaaataggt	2640
	cccgttccta	tcactttacc	atggacggaa	aacaaatagc	tactaccatt	cctcctgttt	2700
5	ttctcttcaa	tggtctggaa	tctgtttcag	gtacagacga	tcgggtatga	aagaaatata	2760
	gaaaacatga	aggaggaata	tcgacatgaa	accagttgta	aaagagtata	caaatgacga	2820
	acagctcatg	aagatgtag	aggaattgca	gaaaatgggt	gttgcgaaag	aggatgtata	2880
10	cgtcttagct	cacgacgatg	acagaacgga	acgcctggct	gacaacacga	acgccaacac	2940
	gatcggagcc	aaagaaacag	gtttcaagca	cgcggtggga	aatatcttca	ataaaaagg	3000
	agacgagctc	cgcaataaaa	ttcacgaaat	cggtttttct	gaagatgaag	ccgtccaatt	3060
15	tgaaaaacgc	ttagatgaag	gaaaagtgtc	tctctttgtg	acagataacg	aaaaagtga	3120
	agcttgggca	taaagcaagg	aaaaaaccaa	aaggccaatg	tcggcctttt	ggtttttttg	3180
	cggtctttgc	gggtgggattt	tgcaaatgca	cgcaatagga	tagcggaaca	tttcgggttc	3240
20	tgaaatgtccc	tcaatttgct	atttatatttt	tgtgataaat	tggaataaaa	tctcacaaaa	3300
	tagaaaatgg	gggtacatag	tgcccatcat	ggccagctag	catgcacatg	ggatctggga	3360
25	ccaataataa	tgactagaga	agaaagaatg	aagattgttc	atgaatttaa	ggaacgaata	3420
	ttggataaag	tggggtattt	ttaaaatata	tatttatgtt	acagtaatat	tgacttttaa	3480
	aaaaggattg	attctaagaa	gaaagcagac	aagtaagcct	cctaatttca	ctttagataa	3540
30	aaatttagga	ggcatatcaa	atgaacttta	ataaaattga	tttagacaat	tggaagagaa	3600
	aagagatatt	taatcattat	ttgaaccaac	aaacgacttt	tagtataacc	acagaaattg	3660
	atattagtgt	tttataccga	aacataaaaac	aagaaggata	taaatrttac	cctgcattta	3720
35	ttttcttagt	gacaaggggtg	ataaactcaa	atcacagctt	tagaactggg	tacaatagcg	3780
	acggagagtt	aggttatttg	gataagttag	agccacttta	tacaattttt	gatgggtgat	3840
40	ctaaaacatt	ctctgggtatt	tggaactcctg	taaagaatga	cttcaaaagag	ttttatgatt	3900
	tataaccttc	tgatgtagag	aaatataatg	gttcggggaa	attgtttccc	aaaacaccta	3960
	tacctgaaaa	tgctttttct	ctttctatta	ttccatggac	ttcattttact	gggttttaact	4020
45	taaatatcaa	taataatagt	aattaccttc	taccatttat	tacagcagga	aaattcatta	4080
	ataaagptaa	ttcaatatat	ttaccgctat	ctttacaggt	acatcattct	gtttgtgatg	4140
	gttatcatgc	aggattgttt	atgaactcta	ttcaggnaatt	gtcagatagg	cctaattgact	4200
50	ggcttttata	atatgagata	atgccgactg	tactttttac	agtcggtttt	ctaacgatac	4260
	atataataggt	acgaaaaagc	aacttttttt	gcgcttaaaa	ccagtcatac	caataactta	4320
	agggttaacta	gcctcgccgg	aaagagcgaa	aatgcctcac	atttgtgcca	cctaaaaagg	4380
55	agcgattttac	atatgagtta	tgcatgttgt	agaatgcnaa	aagtgaatc	agctggacta	4440
	aaaggcatgg	catgccttcg	atagtttatt	aatatttagtg	gagctcagtg	agagcgaagc	4500
	gaacacttga	tttttttaatt	ttctatcttt	tataggtcat	tagagtatac	ttattttgtcc	4560
60	tataaactat	ttagcagcat	aatagattta	ttgaataggt	catttaagtt	gagcatatta	4620
	ggggaggaaa	atcttgagga	aatatttgaa	gaacccgagg	atctagatca	ggtaccgcaa	4680

65

ES 2 358 092 T3

	cggtcgcaga tgctgctgaa gagattatta aaaagctgaa agcaaaaaggc tatcaattgg	4740
	taactgtatc tcagcttgaa gaagtgaaga agcagagagg ctattgaata aatgagtaga	4800
5	aagcgccata tcggcgcttt tcttttggaa gaaaatatag ggaaaatggc acttggttaa	4860
	aattcggaat atttatacaa tatcatatgt atcacattga aaggaggggc ctgctgtcca	4920
	gactgtccgc tgtgtaaaaa aaaggaataa agggggggtg acattatttt actgatatgt	4980
10	ataatataat ttgtataaga aaatggaggg gccctcgaaa cgtaagatga aaccttagat	5040
	aaaagtgcct tttttgttgc aattgaagaa ttattaatgt taagcttaat taaagataat	5100
	atctttgaat tgaacgccc ctcaaaagta agaactacaa aaaaagaara cgttatatag	5160
15	aaatatgttt gaaccttctt cagattacaa atatatctgg acggactcta cctcaaatgc	5220
	ttatctaatc atagaaatgc atacaagcac aaccttgaaa atttgaaaat ataactacca	5280
	atgaacttgt tcatgtgaat tatcgtgta ttttaatttc tcaattcaat atataatag	5340
20	ccaatacatt gttacaagta gaaattaaga cacccttgat agccttacta tacctaacat	5400
	gatgtagtat taaatgaata tgaatatata tttatgataa gaagcgactt atttataatc	5460
25	attacataat tttctattgg aatgattaag attccaatag aatagtgtat aaattattta	5520
	tcttgaaagg agggatgcct aaaaacgaag aacattaaaa acatataattt gcaccgtcta	5580
	atggarttat gaaaaatcat tttatcagtt tgaamattat gtattatgga gctctgaaa	5640
30	aaaggagagg ataaaga atg aag aaa ccg ttg ggg aaa att gtc gca agc	5690
	Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser	
	-190 -185	
	acc gca cta ctc att tct gtt gct ttt agt tca tcc atc gca tcc	5735
35	Thr Ala Leu Leu Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser	
	-180 -175 -170	
	gct gcc acc gga gcg ctc ccc cag tca ccc acc ccg gag gcc gac	5780
	Ala Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp	
	-165 -160 -155	
40	gcg gtc tcc atg cag gag gcg ctc cag cgc gac ctc gac ctg acc	5825
	Ala Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr	
	-150 -145 -140	
	tcc gcc gag gcc gag gag ctg ctg gcc gcc cag gac acc gcc ttc	5870
45	Ser Ala Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe	
	-135 -130 -125	
	gag gtc gac gag gcc gcg gcc gag gcc gcc ggg gac gcc tac ggc	5915
50	Glu Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly	
	-120 -115 -110	
	ggc tcc gtc ttc gac acc gag agc ctg gaa ctg acc gtc ctg gtc acc	5963
	Gly Ser Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr	
	-105 -100 -95	
55	gat gcc gcc gcg gtc gag gcc gtg gag gcc acc ggc gcc ggg acc gag	6011
	Asp Ala Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu	
	-90 -85 -80 -75	
	ctg gtc tcc tac ggc atc gac ggt ctc gac gag atc gtc cag gag ctc	6059
60	Leu Val Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu	
	-70 -65 -60	
65		

ES 2 358 092 T3

	aac gcc gcc gac gcc gtt ccc ggt gtg gtc ggc tgg tac ccg gac gtg Asn Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val	6107
5	gcg ggt gac acc gtc gtc ctg gag gtc ctg gag ggt tcc gga gcc gac Ala Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp	6155
10	gtc agc ggc ctg ctc gcc gac gcc ggc gtg gac gcc tcc gcc gtc gag Val Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu	6203
15	gtg acc acg agc gac cag ccc gag ctc tac gcc gac atc atc ggt ggt Val Thr Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly	6251
20	ctg gcc tac acc atg ggc ggc cgc tgt tcc gtc ggc ttc gcc gcc acc Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr	6299
25	aac gcc gcc ggt cag ccc ggg ttc gtc acc gcc ggt cac tgc ggc cgc Asn Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg	6347
30	gtg ggc acc cag gtg acc atc ggc aac ggc agg ggc gtc ttc gag cag Val Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln	6395
35	tcc gtc ttc ccc ggc aac gac ggc ggc ttc gtc cgc ggt acg tcc aac Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn	6443
40	ttc acg ctg acc aac ctg gtc agc cgc tac aac acc ggc ggg tac gcc Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala	6491
45	acg gtc gcc ggt cac aac cag gcc ccc atc ggc tcc tcc gtc tgc cgc Thr Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg	6539
50	tcc ggc tcc acc acc ggt tgg cac tgc ggc acc atc cag gcc cgc ggc Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly	6587
55	cag tcc gtg agc tac ccc gag ggc acc gtc acc aac atg acc cgg acc Gln Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr	6635
60	acc gtg tgc gcc gag ccc ggc gac tcc ggc ggc tcc tac atc tcc ggc Thr Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly	6683
65	acc cag gcc cag ggc gtg acc tcc ggc ggc tcc ggc aac tgc cgc acc Thr Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr	6731
70	ggc ggg acc acc ttc tac cag gag gtc acc ccc atg gtg aac tcc tgg Gly Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp	6779
75	ggc gtc cgt ctc cgg acc taatcgcatg ttcaatccgc tccataatcg Gly Val Arg Leu Arg Thr	6827
80	gtcgacgcgg cggttcgcgt cccgacagca catcaccgaa atattatgga agaaaatata	6887
85	agcaccatga cggccaaacg gatgcrtcca acggtgctaa ctatatcacg atgtccctaca	6947
90	actattatca cgatcatgat aaagctcca ttttcggatc aagtgacagc aaacctccg	7007

ES 2 358 092 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

atgacggcaa attaaaaatt acgctgcac ataaccgcta taaaaatatt gtccagcgcg 7067
 cgccgagagt ccgcttcggg caagtgcacg tatacaacaa ctattatgaa ggaagcaca 7127
 gctcttcaag ttatcctttt agctatgcat ggggaatcgg aaagtcatct aaatcttatg 7187
 cccaaacaa tgtcattgac gtaccgggac tgtcagctgc taaaacgac agcgtattca 7247
 gcgggggaac ggcttttatat gactccggca cgttgcctga cggcacacag atcaacgcac 7307
 cggctgcaaa cgggctgagc tcttctgtcg gctggacgcc gtctctgcat ggatcgattg 7367
 atgcttctgc taatgtgaaa tcaaatgtta taaatcangc ggggtcgggg aaattaaatt 7427
 aagaagtgaa aaacacacaa ggggtgctaac ctttgtgttt tttattaat taaaatgttt 7487
 attacttag ttaaggagta gaatggaaaa ggggacggg aaacaagtat ataggaggag 7547
 acctatttat ggcttcagaa aaagacgcag gaaaacagtc agcagtaaag cttgttccat 7607
 tgcttattac tgtcgtctgt ggactaatca tctggtttat tcccgcctcg tccggacttg 7667
 aacctaaagc ttggcatttg tttgcgattt ttgtcgcaac aattatcggc tttatctcca 7727
 agcccttgcc aatgggtgca attgcaattt ttgcattggc ggttactgca ctaactggaa 7787
 cactatcaat tgaggataca ttaagcggat tcgggaataa gaccatttgg cttatcgta 7847
 tcgcattctt tatttcccg ggatttatca aaaccggctt cgggtcgaga atttcgtatg 7907
 ttttcgttca gaatttcga aaaaaaccc ttggactttc ctattcactg ctattcagtg 7967
 atttaatact ttcacctgct attccaagta atacggcgcg tgcaggaggc attatatttc 8027
 ctattatcag atcattatcc gaacatttcg gatcaagccc ggcaaatgga acagagagaa 8087
 aaatcgggtgc attcttatta aaaccgggtt ttcaggggaa tctgacaca tctgctatgt 8147
 tcctgacagc gatggcgcg maccgctga ttgccaagct ggcccatgat gtcgcagggg 8207
 tggacttaac atggacaagc tgggcaattg ccgcgattgt accgggactt gtaagcttaa 8267
 tcctcncgcc gcttgtgatt tacaaactgt atccgccgga aatcaangaa acaccggatg 8327
 cggcgaaaat cgcaacagaa aaactgaag aatggggacc gttcaaaaa tcggagcttt 8387
 ccmgtggttat cgtgtttctt ttggtgcttg tgcgttgat ttttggcggc agcttcaaca 8447
 tcgacgctac cacaaccgca ttgatcgggt tggccgttct cttattatca caagttctga 8507
 cttgggatga tatcaagaa gaacaggcg cttgggatac gtcacttgg tttgcggcgc 8567
 ttgtcatgct cgccaacttc ttgaatgaat taggcattgt gtcttgggtc agtaatgcc 8627
 tgaaatcatc cgtatcaggg ttctcttgga ttgtggcatt catcatttta attgttgtgt 8687
 attattactc tcactatttc tttgcaagt cgacagccca catcagtgc atgtattcag 8747
 catcttggc tgcgtcgtg gcagcggcg caccgccgt ttagcagcg ctgagcctcg 8807
 cgttcatcag caacctgttc gggtaacga ctcactacg tctggagcg gctccggtct 8867
 tcttcggagc aggctacat ccgcaaggca aatgggtggt catcggattt atcctgtcga 8927
 ttgttcatat catcgtatgg cttgtgatcg gcggtattat gtggaaagta ctaggaaat 8987
 ggtagaaga aaaggcaga cgcggtctgc ctttttttat tttactcct tcgtaagaa 9047

ES 2 358 092 T3

	atggattttg	aaaaatgaga	aaattccctg	tgaaaaatgg	tatgatctag	gtagaaggga	9107
	cggctggtgc	tgtggtgaaa	aagcggttcc	atttttccct	gcaaacaaaa	ataatggggc	9167
5	tgattgcggc	tctgctggtc	tttgtcattg	gtgtgctgac	cattacgtta	gccgttcagc	9227
	atacacaggg	agaacggaga	caggcagagc	agctggcggg	tcaaacggcg	agaaccattt	9287
	cctatatgcc	gccggttaaa	gagctcattg	agagaaaaga	cggacatgcg	gctcagacgc	9347
10	aagaggtcat	tgaacaaatg	aaagaacaga	ctgggtgcgt	tgccatttat	gttttgaccg	9407
	aaaaaggaga	cattcgcagc	gcctctggaa	aaagcggatt	aaagaaactg	gagcgcagca	9467
	gagaaatfff	gtttggcggg	tcgcatgttt	ctgaaacaaa	agcggatgga	cgaagagtga	9527
15	tcagagggag	cgcgccgatt	ataaaagaac	agaagggata	cagccaagtg	atcggcagcg	9587
	tgtctgttga	ttttctgcaa	acggagacag	agcaaagcat	caaaaagcat	ttgagaaatt	9647
	tgagtgtgat	tgctgtgctt	gtactgctgc	tcggatttat	tggcgcgcgc	gtgctggcga	9707
20	aaagcatcag	aaaggatacg	ctcgggcttg	aaccgcattga	gacgcgggct	ctatatctgt	9767
	agaggaacgc	aatgcttttc	gcgattcgag	aagggatttat	tgccaccant	cgtgaaggcg	9827
	tcgtcaccat	gatgaacgta	tcggcggcgc	agatgctgaa	gctgcccagag	cctgtgatcc	9887
25	atcttccctat	agatgacgtc	atgccgggag	cagggctgat	gtctgtgctt	gaaaaaggag	9947
	aaatgctgcc	gaaccaggaa	gtaagcgtca	acgatcaagt	gtttattatc	aatacgaaag	10007
30	tgatgnaatca	agggggggcag	gcgtatggga	ttgtcgtcag	cttcaggggag	aaaacagagc	10067
	tgaagaagct	gatcgacaca	ttgacagagg	ttcgcaaaata	ttcagaggat	ctcagggcgc	10127
	agactcatga	atcttcnaat	aagcttttatg	cgatttttagg	gctgc		10172

<210> 14

<211> 380

40 $\langle 212 \rangle$ PRT

<213> Artificial

 $\langle 220 \rangle$

45 <223> Constructo sintético

<400> 14

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

55 Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly
-175 -170 -165

Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met
-160 -155 -150

Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala
-145 -140 -135

65 Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
 -130 -125 -120

ES 2 358 092 T3

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

ES 2 358 092 T3

	155	160	165	170
5	Phe Tyr Gln Glu Val	Thr Pro Met Val	Asn Ser Trp Gly Val	Arg Leu
	175		180	185
10	Arg Thr			
	<210> 15			
15	<211> 35			
	<212> ADN			
	<213> Artificial			
20	<220>			
	<223> Cebador			
	<400> 15			
25	ggagctctga aaaaaaggag aggataaaga atgaa			35
	<210> 16			
30	<211> 29			
	<212> ADN			
	<213> Artificial			
35	<220>			
	<223> Cebador			
	<400> 16			
40	gcgttcgat aatcgcggtg acaatgccg			29
	<210> 17			
45	<211> 29			
	<212> ADN			
	<213> Artificial			
50	<220>			
	<223> Cebador			
	<400> 17			
55	tccatgagtc tgcgccctga gatcctctg			29
	<210> 18			
60	<211> 30			
	<212> ADN			
	<213> Artificial			
65	<220>			
	<223> Cebador			

ES 2 358 092 T3

<400> 18		
	taatcgcatg ttcaatccgc tccataatcg	30
5	<210> 19	
	<211> 44	
	<212> ADN	
10	<213> Artificial	
	<220>	
	<223> Cebador	
15	<400> 19	
	cccaacgggtt tcttcattct ttatcctctc cttttttca gage	44
20	<210> 20	
	<211> 1164	
	<212> ADN	
25	<213> Artificial	
	<220>	
	<223> Proteasa 22	
30	<220>	
	<221> CDS	
35	<222> (1)..(1164)	
	<220>	
	<221> Péptido señal	
40	<222> (1)..(81)	
	<220>	
	<221> mat_péptido	
45	<222> (577)..(1164)	
50		
55		
60		
65		

ES 2 358 092 T3

<400> 20

5	atg aaa aaa ccg ctg gga aaa att gtc gca agc aca gca ctt ctt Met Lys Lys -190 Pro Leu Gly Lys -185 Val Ala Ser Thr Ala Leu Leu -180	45
10	att tca gtg gca ttt agc tca tct att gca tca gca gct aca gga Ile Ser Val -175 Ala Phe Ser Ser -170 Ile Ala Ser Ala Val Thr Gly -165	90
15	gca tta ccg cag tct ccg aca ccg gaa gca gat gca gtc tca atg Ala Leu Pro -160 Gln Ser Pro Thr -155 Glu Ala Asp Ala Val Ser Met -150	135
20	caa gaa gca ctg caa aga gat ctt gat ctt aca tca gca gaa gca Gln Glu Ala -145 Leu Gln Arg Asp -140 Asp Leu Thr Ser Ala Glu Ala -135	180
25	gaa gaa ctt ctt gct gca caa gat aca gca ttt gaa gtg gat gaa Glu Glu Leu -130 Leu Ala Ala Gln Asp -125 Thr Ala Phe Glu Val Asp Glu -120	225
30	gca gcg gca gaa gca gca gga gat gca tat ggc ggc tca gtt ttt Ala Ala Ala -115 Glu Ala Ala Gly Asp -110 Ala Tyr Gly Gly Ser Val Phe -105	270
35	gat aca gaa tca ctt gaa ctt aca gtt ctt gtt aca gat gca gca gca Asp Thr Glu -100 Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala -90	318
40	gct gaa gca gtt gaa gca aca gga gca gga aca gta ctt gtt tca tat Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr -85 -80 -75	366
45	gga att gat ggc ctt gat gaa att gtt caa gaa ctg aat gca gct gat Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp -70 -65 -60 -55	414
50	gct gtt ccg ggc gtt gtt ggc tgg tat ccg gat gtt gct gga gat aca Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr -50 -45 -40	462
55	gtt gtc ctt gaa gtt ctt gaa gga tca ggc gca gat gtt tca ggc ctg Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu -35 -30 -25	510
60	ctg gca gac gca gga gtc gat gca tca gca gtt gaa gtt aca aca tca	558

ES 2 358 092 T3

	Leu	Ala	Asp	Ala	Gly	Val	Asp	Ala	Ser	Ala	Val	Glu	Val	Thr	Thr	Ser	
			-20					-15					-10				
5	gat	caa	ccg	gaa	ctt	tat	gca	gat	att	att	ggc	ggc	ctg	gca	tat	tat	606
	Asp	Gln	Pro	Glu	Leu	Tyr	Ala	Asp	Ile	Ile	Gly	Gly	Leu	Ala	Tyr	Tyr	
		-5			-1	1					5				10		
10	atg	ggc	ggc	aga	tgc	agc	gtt	ggc	ttt	gca	gca	aca	aat	gca	tca	ggc	654
	Met	Gly	Gly	Arg	Cys	Ser	Val	Gly	Phe	Ala	Ala	Thr	Asn	Ala	Ser	Gly	
					15					20					25		
15	caa	ccg	ggc	ttt	gtt	aca	gca	ggc	cat	tgc	ggc	aca	gtt	ggc	aca	cca	702
	Gln	Pro	Gly	Phe	Val	Thr	Ala	Gly	His	Cys	Gly	Thr	Val	Gly	Thr	Pro	
				30					35					40			
20	gtt	tca	att	ggc	aat	ggc	aaa	ggc	gtt	ttt	gaa	cga	agc	att	ttt	ccg	750
	Val	Ser	Ile	Gly	Asn	Gly	Lys	Gly	Val	Phe	Glu	Arg	Ser	Ile	Phe	Pro	
			45					50					55				
25	ggc	aat	gat	tca	gca	ttt	gtt	aga	ggc	aca	tca	aat	ttt	aca	ctt	aca	798
	Gly	Asn	Asp	Ser	Ala	Phe	Val	Arg	Gly	Thr	Ser	Asn	Phe	Thr	Leu	Thr	
		60					65					70					
30	aat	ctg	gtt	tca	aga	tat	aat	tca	ggc	ggc	tat	gca	aca	gtt	gca	ggc	846
	Asn	Leu	Val	Ser	Arg	Tyr	Asn	Ser	Gly	Gly	Tyr	Ala	Thr	Val	Ala	Gly	
		75				80					85				90		
35	cat	aat	caa	gca	ccg	att	ggc	tca	gca	gtt	tgc	aga	tca	ggc	tca	aca	894
	His	Asn	Gln	Ala	Pro	Ile	Gly	Ser	Ala	Val	Cys	Arg	Ser	Gly	Ser	Thr	
					95					100					105		
40	aca	ggc	tgg	cat	tgc	ggc	aca	att	caa	gca	aga	aat	caa	aca	gtt	agg	942
	Thr	Gly	Trp	His	Cys	Gly	Thr	Ile	Gln	Ala	Arg	Asn	Gln	Thr	Val	Arg	
				110					115					120			
45	tat	ccg	caa	ggc	aca	gtt	tat	agt	ctg	aca	aga	aca	aca	gtt	tgt	gca	990
	Tyr	Pro	Gln	Gly	Thr	Val	Tyr	Ser	Leu	Thr	Arg	Thr	Thr	Val	Cys	Ala	
			125					130					135				
50	gaa	ccg	ggc	gat	tca	ggc	ggc	tca	tat	att	agc	ggc	act	caa	gca	caa	1038
	Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Tyr	Ile	Ser	Gly	Thr	Gln	Ala	Gln	
		140					145					150					
55	ggc	gtt	aca	tca	ggc	ggc	tca	ggc	aat	tgc	agt	gct	ggc	ggc	aca	aca	1086
	Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly	Asn	Cys	Ser	Ala	Gly	Gly	Thr	Thr	
		155				160					165					170	
60	tat	tac	caa	gaa	gtt	aat	ccg	atg	ctt	agt	tca	tgg	ggc	ctt	aca	ctt	1134
	Tyr	Tyr	Gln	Glu	Val	Asn	Pro	Met	Leu	Ser	Ser	Trp	Gly	Leu	Thr	Leu	
					175					180					185		
65	aga	aca	caa	tcg	cat	gtt	caa	tcc	gct	cca							1164
	Arg	Thr	Gln	Ser	His	Val	Gln	Ser	Ala	Pro							
				190					195								

<210> 21

55 <211> 388

<212> PRT

<213> Artificial

60 <220>

<223> Constructo sintético

65

ES 2 358 092 T3

<400> 21

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

ES 2 358 092 T3

His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr
 95 100 105
 5 Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg
 110 115 120
 10 Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala
 125 130 135
 Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln
 140 145 150
 15 Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr
 155 160 165 170
 20 Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu
 175 180 185
 Arg Thr Gln Ser His Val Gln Ser Ala Pro
 190 195
 25
 <210> 22
 <211> 188
 <212> PRT
 30 <213> *Nocardiopsis* sp. TOA-1
 <400> 22
 35 Ala Asp Ile Ile Gly Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser
 1 5 10 15
 Val Gly Phe Ala Ala Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr
 20 25 30
 40 Ala Gly His Cys Gly Ser Val Gly Thr Gln Val Ser Ile Gly Asn Gly
 35 40 45
 Arg Gly Val Phe Glu Arg Ser Val Phe Pro Gly Asn Asp Ala Ala Phe
 50 55 60
 45 Val Arg Gly Thr Ser Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr
 65 70 75 80
 Asn Ser Gly Gly Tyr Ala Thr Val Ser Gly Ser Ser Thr Ala Pro Ile
 85 90 95
 50 Gly Ser Gln Val Cys Arg Ser Gly Ser Thr Thr Gly Trp Tyr Cys Gly
 100 105 110
 Thr Ile Gln Ala Arg Asn Gln Thr Val Ser Tyr Pro Gln Gly Thr Val
 115 120 125
 55 His Ser Leu Thr Arg Thr Ser Val Cys Ala Glu Pro Gly Asp Ser Ala
 130 135 140
 Gly Ser Phe Ile Ser Gly Thr Gln Ala Gln Gly Val Thr Ser Gly Gly
 145 150 155 160
 60 Ser Gly Asn Cys Arg Thr Gly Gly Thr Thr Phe Tyr Gln Glu Val Asn
 165 170 175
 65 Pro Met Leu Asn Ser Trp Asn Leu Arg Leu Arg Thr
 180 185 188

ES 2 358 092 T3

<210> 23

<211> 381

<212> PRT

5 <213> *Nocardiopsis dassonvillei subesp. dassonvillei* DSM 43235

<400> 23

10

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

15

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

20

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

25

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

30

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

35

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

40

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

45

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

50

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

55

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

60

65

ES 2 358 092 T3

Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Gly Ala Glu
 30 35 40 45
 5 Ser Ser Asp Gly Ser Gly Ser Gly Thr Phe Gln Glu Ser Val Phe Pro
 50 55 60
 10 Gly Ser Asp Gly Ala Phe Val Ala Ala Thr Ser Asn Trp Asn Val Thr
 65 70 75
 15 Asn Leu Val Ser Arg Tyr Asp Ser Gly Ser Pro Gln Ala Val Ser Gly
 80 85 90
 20 Ser Ser Gln Ala Pro Glu Gly Ser Ala Val Cys Arg Ser Gly Ser Thr
 95 100 105
 25 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
 30 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
 35 Tyr Tyr Gln Glu Val Thr Pro Leu Leu Ser Ser Trp Gly Leu Ser Leu
 175 180 185
 40 Val Thr Gly
 190

<210> 24

45 <211> 383

<212> PRT

<213> *Nocardiopsis prasina* DSM 15649

50

55

60

65

ES 2 358 092 T3

<400> 24

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

ES 2 358 092 T3

Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ala Asp Asp Gln
 140 145 150 155
 5 Ala Gln Gly Met Thr Ser Gly Gly Ser Gly Asn Cys Ser Ser Gly Gly
 160 165 170
 10 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
 15
 <210> 25
 <211> 383
 20 <212> PRT
 <213> *Nocardiopsis prasina* DSM 14010
 <400> 25
 25 Met Arg Pro Ser Pro Val Ile Ser Ala Ile Gly Thr Gly Ala Leu
 -190 -185 -180
 30 Ala Phe Gly Leu Ala Leu Ser Val Ala Pro Gly Ala Ser Ala Val
 -175 -170 -165
 35 Thr Ala Pro Ala Glu Pro Ser Pro Gln Gly Glu Ala Thr Thr Met
 -160 -155 -150
 40 Gln Glu Ala Leu Glu Arg Asp Phe Gly Leu Thr Pro Phe Glu Ala
 -145 -140 -135
 45 Asp Asp Leu Leu Glu Ala Gln Lys Glu Ala Leu Gly Ile Asp Thr
 -130 -125 -120
 50 Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Ala Gly Ser Val Phe
 -115 -110 -105
 55 Asp Thr Asp Thr Leu Glu Leu Thr Val Leu Leu Thr Asp Gly Gly Pro
 -100 -95 -90
 60 Ala Ser Asp Val Glu Ala Ala Gly Ala Glu Thr Ser Val Val Ser His
 -85 -80 -75 -70
 65 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

ES 2 358 092 T3

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

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

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

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

20 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

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

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

35 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

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

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

50 Ser Leu Val Thr Ser
190

<210> 26

<211> 384

<212> PRT

<213> *Nocardiopsis* sp. DSM 16424

ES 2 358 092 T3

<400> 26

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

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

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

 20 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

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

 30 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

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

 40 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

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

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

 55 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

 60 Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly His Gln Ala
 75 80 85

 65 Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg Ser

ES 2 358 092 T3

90

95

100

5 Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln
105 110 115

10 Thr Val Arg Tyr Pro Gln Gly Thr Val Asn Ala Leu Thr Arg Thr Asn
120 125 130 135

15 Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly Ser
140 145 150

20 Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Phe Gly
155 160 165

25 Gly Thr Thr Tyr Tyr Gln Glu Val Ala Pro Met Ile Asn Ser Trp Gly
170 175 180

30 Val Arg Ile Arg Thr Ser
185

35

<210> 27

<211> 383

<212> PRT

30

<213> *Nocardiosis alkalifila* DSM 44657

<400> 27

35

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

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

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

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

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

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

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

60

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

65 Asp His Gly Ile Glu Gly Leu Glu Glu Ile Val Asp Glu Leu Asn Glu

ES 2 358 092 T3

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

60

<210> 28

<211> 384

65 <212> PRT

<213> *Nocardiopsis lucentensis* DSM 44048

ES 2 358 092 T3

<400> 28

5 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
 10 Thr Val Pro Thr Pro Gln Ala Pro Val Pro Asp Asp Glu Ala Thr
 -165 -160 -155
 15 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
 20 Asp Glu Glu Ala Thr Glu Ala Ala Gly Ala Ala Tyr Gly Gly Ser
 -120 -115 -110
 25 Leu Phe Asp Thr Glu Thr His Glu Leu Thr Val Leu Val Thr Asp Val
 -105 -100 -95 -90
 30 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
 35 Thr Asp Ala Gly Ser Glu Val Val Gly Trp Tyr Pro Asp Val Thr Ser
 -55 -50 -45
 40 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
 45 Glu Val Ser Glu Gln Pro Gln Thr Tyr Ala Asn Ile Ile Gly Gly Leu
 -5 -1 1 5
 50 Ala Tyr Tyr Met Ser Ser Gly Gly Arg Cys Ser Val Gly Phe Pro Ala
 10 15 20
 55 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

ES 2 358 092 T3

Arg Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser
 60 65 70
 5 Asn Phe Thr Leu Tyr Asn Leu Val Tyr Arg Tyr Ser Gly Tyr Gln Thr
 75 80 85
 10 Val Thr Gly Ser Asn Ala Ala Pro Ile Gly Ser Ser Ile Cys Arg Ser
 90 95 100
 15 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
 20 Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly Thr
 140 145 150
 25 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
 30 Val Arg Leu Arg Thr Ser
 185