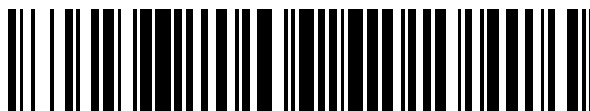


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 562 919**

51 Int. Cl.:

C12N 9/52

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.08.2008 E 08782631 (9)**

97 Fecha y número de publicación de la concesión europea: **18.11.2015 EP 2173873**

54 Título: **Proteína y secuencia de ADN que codifica una actividad del tipo de la subtilisina adaptada al frío**

30 Prioridad:

06.08.2007 US 954198 P

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

09.03.2016

73 Titular/es:

**UNIVERSITY OF CHILE (100.0%)
BEAUCHEF 861
SANTIAGO, CL**

72 Inventor/es:

**ASENJO, JUAN A.;
ANDREWS, BARBARA A.;
ACEVEDO, JUAN PABLO;
REYES, FERNANDO y
BURZIO, LUIS O.**

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 562 919 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Proteína y secuencia de ADN que codifica una actividad del tipo de la subtilisina adaptada al frío

Antecedentes de la invención

1. Campo de la invención

5 La presente invención se refiere a ácidos nucleicos purificados que codifican enzimas derivadas de bacterias antárticas (*Polaribacter sp.*) tales como proteinasas, que pueden ser una proteína, y a polipéptidos purificados que tienen una alta actividad proteolítica y pertenecen a la superfamilia de enzimas del tipo de la subtilisina (subtilasas). La presente invención también se refiere a una proteína que tiene actividad adaptada al frío, especialmente actividad específica en el intervalo de alrededor de 4-45°C, y que tienen actividad notable en el intervalo de 4 - 20°C. Además, la presente invención se refiere a un constructo de ADN que comprende una secuencia de ADN que codifica la proteasa del tipo de la subtilisina adaptada al frío, y una célula que incluye el constructo de ADN. Además, la presente invención se refiere a un método de preparación de la proteasa del tipo de la subtilisina adaptada al frío mediante el uso de técnicas de ADN recombinante.

2. Descripción del estado del arte

15 La familia de la serina proteasa del tipo de la subtilisina (S8) juega un papel en una multitud de diversos procesos celulares y metabólicos bacterianos, tales como la esporulación y la diferenciación, el recambio de la proteína, la maduración de enzimas y hormonas y el mantenimiento de la reserva de proteína celular. Otra función importante, especialmente para proteasas del tipo de la subtilisina extracelular, es la hidrólisis de las proteínas en ambientes celulares externos que le permite a la célula absorber y utilizar productos hidrolíticos.

20 Las serina proteasas se utilizan en numerosos y variados contextos industriales y con fines comerciales, incluyendo detergentes para ropa, procesamiento de alimentos, tratamiento del cuero, uso médico y productos para el cuidado de la piel. En los detergentes para ropa, se emplea la proteasa para descomponer compuestos orgánicos o poco solubles hasta formas más solubles que pueden disolverse más fácilmente en agua y detergente. Ejemplos de procesamiento de alimentos incluyen ablandamiento de carnes, preparación de hidrolizados de proteína y maduración del queso. En el caso de uso médico, las proteasas se aplican al tratamiento de quemaduras, heridas purulentas, forúnculos y abscesos profundos. Las proteasas pueden ser incluidas en el campo del cuidado de la piel para remover escamas en la superficie de la piel que se acumulan debido a un desequilibrio en la velocidad de descamación.

30 Las proteasas comunes utilizadas en algunas de estas aplicaciones se derivan de las células procariotas o eucariotas que se cultivan fácilmente para la fabricación industrial de sus enzimas. Por ejemplo, una especie común utilizada es *Bacillus* como se describe en la patente de los Estados Unidos No. 5.217.878. Alternativamente, la patente de los Estados Unidos No. 5.278.062, describe serina proteasas aisladas de un hongo, *Tritirachium album*, para su uso en composiciones detergentes para lavado de ropa. El advenimiento de la tecnología recombinante permite la expresión de proteínas de cualquier especie en un huésped adecuado para fabricación industrial. La mayoría de las proteasas disponibles comercialmente usadas en aplicaciones detergentes tienen temperaturas óptimas altas, por ejemplo 60°C. Las bacterias aisladas de ambientes fríos tales como el agua de mar de la Antártida son microorganismos psicrófilos y se espera que tengan enzimas adaptadas al frío.

40 Existen algunas enzimas con actividad del tipo de la subtilisina adaptadas al frío provenientes de microorganismos psicrófilos, por ejemplo: *Flavobacterium balustinum* (Morita, Y., Hasan, P., Sakaguchi, T., Murakami Y., Yokohama, K, Tamaya, E. (1998) Appl. Microbiol Biotechnol. 50: 669-675), *Bacillus TA41* (Davial, S., Feller, G, Narinx, E., Gerday, Ch (1994) J. Biol. Chem. 269: 17448 - 17453) y la cepa DY-A de *Pseudomonas* (Zeng, R., Zhang, R, Zhao, J., Lin, N. (2003) Extremophiles 7: 335 - 337). Todas estas proteínas tienen baja estabilidad a temperatura ambiente y en presencia de compuestos comunes presentes en los detergentes comerciales.

45 Por lo tanto, existe la necesidad de nuevas proteasas alternativas, en este caso proteasas del tipo de la subtilisina que funcionan a temperatura ambiente y a bajas temperaturas y en presencia de composiciones detergentes comerciales comunes.

Resumen de la invención

Una realización de la presente invención es un ácido nucleico aislado que comprende: una secuencia de nucleótidos idéntica a la SEQ ID NO: 1 o una variante degenerada de la misma; o una secuencia de nucleótidos que comprende posiciones idénticas 361-423, 466-528 y 1039-1101 de la SEQ ID NO: 1 o una variante degenerada de la misma.

50 Una realización de la presente invención es un ácido nucleico sustancialmente puro que comprende un ácido nucleico que codifica un polipéptido que tiene al menos aproximadamente 85% de homología (por ejemplo, identidad) con una

proteína del tipo de la subtilisina adaptada al frío derivada de *Polaribacter* o una proteína de referencia, tal como el polipéptido de la SEQ ID NO: 2, y más preferiblemente, al menos aproximadamente 90% de homología, y aún más preferiblemente, al menos aproximadamente 95% de homología. El nivel de homología (por ejemplo, identidad) se aplica a todas las realizaciones de la invención.

5 En ciertas realizaciones, el ácido nucleico sustancialmente puro comprende una variante de ácido nucleico modificada por ingeniería genética que codifica un polipéptido que difiere de una proteína de referencia o una proteína tal como la subtilisina adaptada al frío derivada de *Polaribacter* en no más de aproximadamente 30 sustituciones de aminoácidos, y más preferiblemente, no más de aproximadamente 20 sustituciones de aminoácidos. Preferiblemente, las sustituciones
10 modificadas por ingeniería genética provocan una sustitución conservadora en la secuencia de aminoácidos de una secuencia de referencia o una proteína adaptada al frío.

La invención incluye adicionalmente vectores de expresión capaces de reproducirse en una célula, tal como una célula eucariota o procariota, incluyendo el ácido nucleico de cualquiera de las reivindicaciones 1-3 operativamente enlazadas a una secuencia de control de la expresión así como a células transformadas que tienen tales vectores de expresión.

15 Aún otra realización de la invención es un vector capaz de reproducirse en una célula tal como una célula eucariota o procariota. El vector comprende un ácido nucleico que codifica un polipéptido que tiene al menos aproximadamente 85% de homología con una secuencia de referencia o una proteína del tipo de la subtilisina adaptada al frío derivada de *Polaribacter*, la SEQ ID NO: 2. El vector de la invención codifica para la expresión, intracelular o extracelular, de la proteína del tipo de la subtilisina adaptada al frío descrita en este documento.

20 Otra realización de la presente invención es un polipéptido que comprende una isoforma sustancialmente pura de una secuencia de referencia o una proteína del tipo de la subtilisina adaptada al frío derivada de *Polaribacter* o una variante de la misma modificada por ingeniería genética, y preferiblemente, un polipéptido que comprende la SEQ ID NO: 2.

La invención proporciona además una composición de limpieza o detergente que comprende el polipéptido o la proteína del tipo de la subtilisina adaptada al frío de la invención.

25 Aún otra forma de realización de la invención es un método de preparación de una enzima tal como una enzima del tipo de la subtilisina adaptada al frío, en donde la proteína tiene al menos aproximadamente 85% de homología con una secuencia de referencia o una proteína multifuncional derivada de *Polaribacter*. Tal método comprende:

1. La construcción de un vector de expresión quimérico recombinante, que comprende una secuencia de ácido nucleico de la presente invención, tal como la SEQ ID NO: 1.
- 30 2. La transformación de una célula huésped eucariota o procariota apropiada con el vector de expresión para expresar, de forma intracelular o extracelular, un ácido nucleico que codifica la proteína; y
3. El cultivo de la célula transformada en un cultivo y aislamiento de la proteína a partir de la célula transformada o el medio de cultivo.

Estos, junto con otros objetivos y ventajas que se harán posteriormente evidentes residen en la construcción y operación detallada como se describe y reivindica aquí más adelante en forma detallada.

35 Descripción detallada de realizaciones de la invención

Aunque solamente se explican en detalle ciertas realizaciones de la invención, se entiende que la invención no está limitada en su alcance a los detalles expuestos en la siguiente descripción. La invención puede incluir otras realizaciones y de ser practicada o llevada a cabo de varias maneras. También, en la descripción de estas
40 realizaciones, se recurrirá a terminología específica con fines de claridad. Se entiende que cada término específico incluye todos los equivalentes técnicos que actúan de una forma similar para lograr un propósito similar.

Para los fines de esta solicitud, los términos enumerados a continuación tendrán el siguiente significado:

45 "Isoforma" se refiere a una variante de secuencia de origen natural de una proteína sustancialmente homóloga dentro del mismo organismo. Preferiblemente, la isoforma comparte al menos aproximadamente 85% de identidad, y más preferiblemente, al menos aproximadamente 90% de identidad con una de las siguientes secuencias de residuos de aminoácidos:

- residuos de aminoácidos 25-1129 de la SEQ ID NO: 2.
- residuos de aminoácidos 96-1129 de la SEQ ID NO: 2

- residuos de aminoácidos 96-870 de la SEQ ID NO: 2
- restos de aminoácidos 96-650 de la SEQ ID NO: 2.
- restos de aminoácidos 96-560 de la SEQ ID NO: 2.

- 5 "Proteína de actividad del tipo de la subtilisina adaptada al frío derivada de *Polaribacter*" se refiere a una proteína del tipo de la subtilisina adaptada al frío que tiene la misma secuencia que una proteína aislada de la cepa 3-17 de *Polaribacter sp.* y que tiene las propiedades de la proteína descritas en la sección titulada "Características preferidas de la proteína del tipo de la subtilisina adaptada al frío". La secuencia de aminoácidos incluida en la SEQ ID NO: 2 u otras isoformas de la misma o polipéptidos quiméricos de la misma son ejemplos de proteínas con actividad como la de la subtilisina adaptada al frío derivada de *Polaribacter*.
- 10 "Porcentaje de identidad de secuencia" se refiere al porcentaje de dos secuencias que se consideran idénticas u homólogas en el estado del arte. Para determinar el porcentaje de identidad de dos secuencias de aminoácidos, se alinean las secuencias con el fin de lograr una comparación óptima (por ejemplo, pueden introducirse huecos tanto en una primera como en una segunda secuencia de aminoácidos o de ácido nucleico para una alineación óptima y las secuencias no homólogas pueden pasarse por alto para efectos de comparación). La alineación con fines de determinar el porcentaje de identidad de secuencia de aminoácidos puede conseguirse de diversas maneras que están dentro del estado del arte, por ejemplo, utilizando software informático que se encuentra públicamente disponible, tal como el software BLAST-2 que se configura con sus parámetros predeterminados. Aquellos ordinariamente capacitados en la técnica pueden determinar los parámetros apropiados para medir la alineación, incluyendo cualquier algoritmo necesario para conseguir la máxima alineación sobre la longitud completa de las secuencias que se comparan. El método de alineación ClustalW (1.60) se utiliza en esta solicitud.
- 15 "Polipéptido" se refiere a un polímero compuesto de aminoácidos unidos entre sí para formar enlaces peptídicos, formando preferiblemente una preproteína, proproteína, proteína o fragmento de la misma.
- 20 "Preproteína" se refiere a un polipéptido que consiste de una secuencia señal, prorregiones, y una secuencia de la proteína procesada.
- 25 "Proteína" se refiere a un polipéptido que consiste de prorregiones y de la secuencia de la proteína procesada.
- "Método de desplazamiento sobre el genoma" se refiere a una técnica para el aislamiento de polinucleótidos de las regiones de secuencia desconocidas a cada lado de aquellas conocidas; se conocen colectivamente como técnicas de desplazamiento sobre el genoma o de desplazamiento sobre el cromosoma.
- 30 "Polinucleótido" se refiere a un polímero de ADN o ARN que puede ser mono o bicatenario, que contiene opcionalmente bases nucleotídicas sintéticas, no naturales o alteradas que pueden ser incorporadas en polímeros de ADN o ARN. El polinucleótido puede estar en la forma de un fragmento separado o como componente de un constructo de una secuencia mayor de nucleótidos.
- Purificación de una proteína nativa con actividad como la de la subtilisina adaptada al frío derivada de *Polaribacter*.
- 35 Las realizaciones de polipéptidos nativos de la invención son preferiblemente la proteasa producida por el uso de microorganismos que pertenecen al género *Polaribacter*. Un ejemplo de los microorganismos que tienen la capacidad de producir la proteasa de acuerdo con la presente invención es una cepa 3-17 de *Polaribacter sp.* que se aisló a partir de agua de mar recogida en la Base Frei Montalva (Latitud 62° 11" Sur, Longitud 58° 58" Oeste), Isla Rey Jorge, Antártica chilena. Esta cepa se caracteriza por la secuencia de ácido nucleico de su gen de ARN ribosomal 16S que es idéntica a la secuencia de la SEQ ID NO: 19.
- 40 Las condiciones para el cultivo de la cepa en esta invención pueden ser diversas, en la medida en que permitan una buena producción de la proteasa. Por ejemplo: se puede usar un medio sólido o líquido, un cultivo agitado o un cultivo en biorreactor aeróbico con agitación, diferentes fuentes de carbono (glucosa, trehalosa, fructosa, sacarosa, maltosa, almidón y oligosacárido de malta), diferentes fuentes de nitrógeno (peptona, extracto de levadura, extracto de malta, extracto de carne, polvo de soja, polvo de semilla de algodón, aminoácidos y nitratos), diferentes sales inorgánicas (magnesio, fosfato, calcio, sodio, potasio, hierro y manganeso) y otros nutrientes necesarios. También se pueden alterar en forma adecuada las condiciones de cultivo, tales como el pH y la temperatura. En la presente invención las condiciones preferidas son pH neutro y una temperatura de cultivo de aproximadamente 10°C.
- 45 La proteasa de la presente invención está preferiblemente presente en el sobrenadante del medio de cultivo, pero también está presente en las paredes celulares de las bacterias. La proteasa puede ser usada en cualquier forma tal como células bacterianas, como una enzima cruda obtenida a partir de las células bacterianas o del sobrenadante del medio de cultivo, o como una enzima extraída y purificada. Alternativamente, también se puede utilizar una proteasa
- 50

5 inmovilizada por un método conocido. Dado que la proteasa de la presente invención se encuentra principalmente en el medio extracelular, una solución de enzima cruda puede ser obtenida fácilmente mediante la eliminación de las células bacterianas con la ayuda de filtración o centrifugación. Esta enzima cruda se puede purificar además por un método de purificación conocido o combinación de métodos conocidos. Las realizaciones típicas de métodos de purificación adecuados se describen en los ejemplos presentados en el presente documento.

Polinucleótidos y polipéptidos

10 Las realizaciones de polinucleótidos de la invención son preferiblemente ácidos desoxirribonucleicos (ADN), tanto ácidos desoxirribonucleicos monocatenarios como bicatenarios, y más preferiblemente ácidos desoxirribonucleicos bicatenarios. Sin embargo, también pueden ser, sin limitación, ácidos ribonucleicos (ARN), así como moléculas bicatenarias híbridas de ARN:ADN.

15 La presente invención abarca polinucleótidos que codifican una proteína con actividad del tipo de la subtilisina adaptada al frío derivada de *Polaribacter*, ya sea ARN, ADN, o ADNc nativo o sintético, que codifican la proteína, o la cadena complementaria de la misma, incluyendo, por ejemplo, ácidos nucleicos encontrados en un organismo que expresa proteína del tipo de la subtilisina adaptada al frío. Para propósitos de expresión recombinante, se tienen en cuenta convenientemente las preferencias de uso de codones para el organismo en el que dicho ácido nucleico se expresa en el diseño de un ácido nucleico sintético que codifica proteína del tipo de la subtilisina adaptada al frío.

20 Las secuencias de ácido nucleico pueden ser mutadas adicionalmente, por ejemplo, para incorporar sitios de restricción útiles. Véase Sambrook y colaboradores, Molecular Cloning, a Laboratory Manual (Cold Spring Harbor Press, 1989). Tales sitios de restricción se pueden utilizar para crear "casetes", o regiones de secuencia de ácido nucleico que son fácilmente sustituidas usando enzimas de restricción y reacciones de ligación. Los casetes se pueden utilizar, por ejemplo, para sustituir secuencias sintéticas que codifican secuencias mutadas de aminoácidos de proteínas tales como la subtilisina adaptada al frío.

Las secuencias de ácido nucleico de la presente invención pueden codificar, por ejemplo, una de varias isoformas de una proteína con actividad como la de la subtilisina adaptada al frío derivada de *Polaribacter*.

25 Esta proteína con actividad como la de la subtilisina adaptada al frío derivada de *Polaribacter* corresponde a una preproteína. La secuencia señal de preproteína es el segmento de la proteína que está presente en la proteína precursora en la célula bacteriana, pero que está ausente en la proteína después de la secreción en el ambiente extracelular. La secuencia señal corresponde a los residuos de aminoácidos 1-24 en la SEQ ID NO: 2: Met Lys Lys Arg Tyr Ile Asn Leu Leu Thr Ile Gly Val Phe Met Ile Ser Ala Phe Asn Met Asn Ala. Las secuencias de aminoácidos 30 restantes de los polipéptidos representan la proproteína. Las proproteínas, especialmente en proteasas extracelulares, están presentes preferiblemente en una forma inactiva o en una forma parcialmente activa y pueden transformarse en una proteína activa a través de autodigestión o extracción de las proregiones, aquellas proregiones corresponden a menudo a la región inicial, región final o ambas regiones de la secuencia de proproteína. Las proregiones de la proteína con actividad del tipo de la subtilisina adaptada al frío derivada de *Polaribacter* son:

- 35 1. residuos de aminoácidos 25-95 en la SEQ ID NO: 2.
2. residuos de aminoácidos 871-1129 en la SEQ ID NO: 2.
3. residuos de aminoácidos 651-870 en la SEQ ID NO: 2
4. residuos de aminoácidos 561-650 en la SEQ ID NO: 2

40 Las secuencias de aminoácidos restantes de estos polipéptidos (distintos de la secuencia señal y de los segmentos de proproteína) representan la proteína procesada.

Diversas realizaciones de la proteína con actividad del tipo de la subtilisina adaptada al frío derivada de *Polaribacter*, incluyen, pero no se limitan a, una secuencia de aminoácidos como se muestra en la SEQ ID NO: 2; así como las posiciones 96 a 1129 de la SEQ ID NO: 2, las posiciones 96 a 870 de la SEQ ID NO: 2, las posiciones 96- 650 de la SEQ ID NO: 2 y las posiciones 96 a 560 de la SEQ ID NO: 2 que podrían ser individualmente proteolíticamente activas. 45 Las realizaciones adicionales de la proteína con actividad del tipo de la subtilisina adaptada al frío derivada de *Polaribacter* comprenden secuencias de aminoácidos que forman parte de la "tríada catalítica" de la SEQ ID NO: 2, es decir, las posiciones 121-141, 156-176 y 347-367 de la SEQ ID NO: 2. Dicho de otro modo, tales formas de realización comprenden las secuencias de nucleótidos 361 - 423, 466 - 528, 1039 - 1101 de la SEQ ID NO: 1. Otras realizaciones de la proteína con actividad del tipo de la subtilisina adaptada al frío derivada de *Polaribacter* comprenden secuencias de aminoácidos que se reconocen como dominios del tipo 3 de fibronectina en la SEQ ID NO: 2, es decir, las posiciones 50 651 - 870 y 561 - 650 de la SEQ ID NO: 2. Tales realizaciones comprenden las secuencias de nucleótidos 1951-2610 y 1681-1950 de la SEQ ID NO: 1, respectivamente. Estos dominios de fibronectina podrían ser auxiliares para la eficiencia

proteolítica contra de algunos sustratos o auxiliares en el replegamiento y procesamiento de proteínas.

Preferiblemente, los ácidos nucleicos codificarán polipéptidos que tienen al menos 85% de homología, más preferiblemente, al menos 90% de homología, incluso más preferiblemente, al menos aproximadamente 95% de homología con una proteína de referencia o una proteína como la subtilisina adaptada al frío derivada de *Polaribacter*, tal como los polipéptidos de la SEQ ID NO: 2 u otras isoformas de origen natural.

La proteína procesada del polipéptido de la SEQ ID NO: 2 es aproximadamente 62% idéntica a la serina proteinasa tal como la subtilisina en el *Psychroflexus torquis* ATCC 700755 de acuerdo con la secuencia proporcionada por el GenBank (Mountain View, California), la adquisición de la base de datos no. ZP_01253004, y aproximadamente 43% idéntica a la serina proteasa del tipo de la subtilisina en el *Flavobacterium bacterium* BBFL7, de acuerdo con la secuencia proporcionada por el GenBank, la adquisición de la base de datos no. ZP_01202744, y aproximadamente 30% idéntica a la serina proteasa alcalina en el *Bacillus sp. Ksm-Kp43*, de acuerdo con la secuencia proporcionada por Protein Data Bank, la adquisición de la base de datos no. 1WMDA. Preferiblemente, los ácidos nucleicos que codifican polipéptidos que tienen actividad del tipo de la subtilisina adaptada al frío son aproximadamente menos de 80% idénticos a las proteinasas anteriormente identificadas de *Psychroflexus torquis* ATCC 700755, *Flavobacterium bacterium* BBFL7 o *Bacillus sp. Ksm-Kp43*.

La secuencia que codifica la proteína del tipo de la subtilisina adaptada al frío puede ser, por ejemplo, sustancialmente o totalmente sintética. Para los propósitos de expresión recombinante, las preferencias de uso de codones para el organismo en el que dicho ácido nucleico se va a expresar se consideran convenientemente en el diseño de un ácido nucleico sintético que codifica la proteína adaptada al frío. Dado que el código del ácido nucleico está degenerado, se pueden usar numerosas secuencias de ácido nucleico para crear la misma secuencia de aminoácidos. Esta "degeneración" o "redundancia" natural del código genético es bien conocida en la técnica. Por tanto, se apreciará que la secuencia de ácido nucleico mostrada en el listado de secuencias proporciona solamente un ejemplo dentro de un grupo grande pero definido de secuencias de ácido nucleico que codificaran los polipéptidos relevantes como se describe en el presente documento.

Los polipéptidos de la presente invención preferiblemente incluyen todos los polipéptidos codificados por los ácidos nucleicos que tienen una secuencia idéntica a la SEQ ID NO: 1 o sus variantes degeneradas de los mismos, y todos los polipéptidos que comprenden las secuencias de aminoácidos mostradas como:

a) los residuos de aminoácidos 25 -1129 de la SEQ ID NO: 2.

b) los residuos de aminoácidos 96-1129 de la SEQ ID NO: 2.

c) los residuos de aminoácidos 96-870 de la SEQ ID NO: 2

d) residuos de aminoácidos 96-650 de la SEQ ID NO : 2

e) los residuos de aminoácidos 96-560 de la SEQ ID NO: 2.

así como todas las variantes obvias de estos péptidos que están en el estado de la técnica para elaborar y utilizar. Además, los polipéptidos de acuerdo con la presente invención tienen preferiblemente al menos aproximadamente 85% de identidad de secuencia, también preferiblemente al menos aproximadamente 90% de identidad de secuencia, más preferiblemente al menos 95% de identidad de secuencia, también más preferiblemente al menos 96% de identidad de secuencia, incluso preferiblemente al menos 97% de identidad de secuencia, incluso más preferiblemente al menos aproximadamente 98% de identidad de secuencia, aún preferiblemente al menos 99% de identidad de secuencia con la secuencia de aminoácidos seleccionada de:

a) los residuos de aminoácidos 25-1129 de la SEQ ID NO: 2.

b) los residuos de aminoácidos 96-1129 de la SEQ ID NO: 2.

c) los residuos de aminoácidos 96-870 de la SEQ ID NO: 2.

d) los residuos de aminoácidos 96 -650 de SEQ ID NO: 2.

e) los residuos de aminoácidos 96-560 de la SEQ ID NO: 2.

Métodos de síntesis de polipéptidos

Vectores de expresión

La presente invención también se refiere a vectores de expresión recombinante que comprenden una secuencia de ácido nucleico de la presente invención, un promotor, y señales de detención de la transcripción y de la traducción. La secuencia de ácido nucleico de la presente invención puede ser expresada insertando la secuencia de ácido nucleico o un constructo de ácido nucleico que comprende una secuencia como se identifica en la SEQ ID NO: 1 en un vector apropiado para expresión. En la creación del vector de expresión, la secuencia de codificación, como la identificada en la SEQ ID NO: 1, se encuentra en el vector de modo que está unida operativamente con las secuencias de control apropiadas para expresión.

El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido) que puede ser convenientemente sometido a procedimientos de ADN recombinante y puede provocar la expresión de la secuencia de ácido nucleico. La elección del vector normalmente dependerá de la compatibilidad del vector con la célula huésped en la que se introduce el vector. Los vectores pueden ser plásmidos circulares lineales o cerrados.

El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica y su replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un mini cromosoma, o un cromosoma artificial. El vector puede contener cualquier medio para asegurar la autorreplicación. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma de la célula huésped y se replica junto con el(los) cromosoma(s). Además, se pueden utilizar un único vector o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total que se va a introducir en el genoma de la célula huésped, o un transposón.

Los vectores de la presente invención contienen preferiblemente uno o más marcadores seleccionables que permiten una selección fácil de células transformadas. Un marcador seleccionable es un gen cuyo producto de expresión proporciona resistencia biocida o viral, resistencia a metales pesados, prototrofia a auxótrofos. Los ejemplos de marcadores seleccionables bacterianos son los genes *dal* de *Bacillus subtilis* o *Bacillus licheniformis*, o marcadores que confieren resistencia a los antibióticos tales como resistencia a la ampicilina, kanamicina, cloranfenicol o a la tetraciclina. Los marcadores adecuados para células huéspedes de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1, y URA3. Los marcadores seleccionables para uso en una célula huésped de hongos filamentosos incluyen, por ejemplo, *amdS* (acetamidasa), *argB* (ornitina carbamoiltransferasa), *bar* (fosfotricina acetiltransferasa), *hygB* (higromicina fosfotransferasa), *niaD* (nitrato reductasa), *pyrG* (orotidina-5'-fosfato descarboxilasa), *sC* (sulfato adeniltransferasa), *trpC* (antranilato sintasa), así como equivalentes de los mismos. Los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygroscopicus* se prefieren para uso en una célula de *Aspergillus*.

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

Para la integración en el genoma de la célula huésped, el vector puede depender de la secuencia de ácido nucleico que codifica el polipéptido o cualquier otro elemento del vector para la integración estable del vector en el genoma por recombinación homóloga o no homóloga. Alternativamente, el vector puede contener secuencias adicionales de ácidos nucleicos para dirigir la integración por recombinación homóloga en el genoma de la célula huésped. Las secuencias adicionales de ácidos nucleicos permiten que el vector se integre en el genoma de la célula huésped en una ubicación(es) precisa(s) en el cromosoma(s). Para aumentar la probabilidad de integración en una ubicación precisa, los elementos de integración deben contener preferiblemente un número suficiente de ácidos nucleicos, tal como 100 a 1.500 pares de bases, preferiblemente 400 a 1.500 pares de bases, y lo más preferiblemente 800 a 1.500 pares de bases, que son altamente homólogos con la secuencia objetivo correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos de integración pueden ser cualquier secuencia que sea homóloga con la secuencia objetivo en el genoma de la célula huésped. Además, los elementos de integración pueden ser secuencias de ácidos nucleicos codificadoras o no codificadoras. Por otra parte, el vector puede ser integrado en el genoma de la célula huésped por recombinación no homóloga.

Para la replicación autónoma, el vector puede comprender además un origen de replicación que le permite al vector replicarse autónomamente en la célula huésped en cuestión. Los ejemplos de orígenes bacterianos de replicación son los orígenes de replicación de los plásmidos pBR322, pUC19, pACYC177, y pACYC184 que permiten la replicación en *E. coli*, y pUBub110, pE194, pTA1060, y pAM.beta.1 que permite la replicación en *Bacillus*. Los ejemplos de orígenes de replicación para uso en una célula huésped de levadura son el origen de replicación de 2 micras, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6.

Se puede insertar más de una copia de una secuencia de ácido nucleico de la presente invención en la célula huésped para aumentar la producción del producto del gen. Un aumento en el número de copias de la secuencia de ácido nucleico se puede obtener mediante la integración de al menos una copia adicional de la secuencia en el genoma de la célula huésped o mediante la inclusión de un gen marcador seleccionable amplificable con la secuencia de ácido nucleico donde las células que contienen las copias amplificadas del gen marcador seleccionable, y por lo tanto las copias adicionales de la secuencia de ácido nucleico, se pueden seleccionar mediante el cultivo de las células en presencia del agente seleccionable apropiado.

Los procedimientos utilizados para ligar los elementos descritos anteriormente para construir los vectores de expresión recombinantes de la presente invención son, bien conocidos por la persona ordinariamente capacitada en el arte (véase, por ejemplo., Sambrook y colaboradores, 1989, ver más arriba)

Células huésped

5 La presente invención también se refiere a células huésped recombinantes, que comprenden un vector de la presente invención para la producción recombinante de los polipéptidos.

10 Se introduce un vector que comprende una secuencia de ácido nucleico de la presente invención en una célula huésped de manera que el vector se mantiene como un integrante cromosómico o como un vector extracromosómico autorreplicante como se describió anteriormente. La escogencia de una célula huésped, en gran medida, dependerá del gen que codifica al polipéptido y su fuente.

La célula huésped puede ser un microorganismo unicelular, por ejemplo, un organismo procariota y eucariota unicelular (levadura), o un organismo no unicelular, por ejemplo, un eucariota.

15 Las células unicelulares útiles son células bacterianas tales como bacterias gram positivas incluyendo, pero sin limitarse a, 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*, por ejemplo, *Streptomyces lividans* o *Streptomyces murinus*, o bacterias gram negativas tales como *E. coli* y *Pseudomonas* sp. En una realización preferida, la célula huésped bacteriana es una *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus* o *Bacillus subtilis*. En otra forma de realización preferida, la célula de *Bacillus* es un *Bacillus* alcalofílica.

20 La introducción de un vector en una célula huésped bacteriana puede, por ejemplo, lograrse mediante transformación de un protoplasto a través de electroporación o conjugación, utilizando células competentes.

25 La célula huésped puede ser una célula eucariota, tal como una célula de mamífero, insecto, planta, o de hongo. En una realización preferida, la célula huésped es una célula de hongo. "Hongos" como se utiliza aquí incluye los hongos de los filos *Ascomycota*, *Basidiomycota*, *Chytridiomycota*, y *Zygomycota*, así como los hongos *Oomycota* y todo los mitospórico.

En una realización más preferida, la célula huésped de hongo es una célula de levadura. "Levadura" como se usa aquí incluye levadura ascosporogénea (Endomicetos), levadura basidiosporógena, y levadura perteneciente a los hongos imperfectos (Blastomicetos).

30 En una realización incluso más preferida, la célula huésped de levadura es una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*.

En una realización más preferida, la célula huésped de levadura es una célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* o *Saccharomyces oviformis*.

35 Producción

La presente invención también se refiere a un método para producir un polipéptido de la invención, comprendiendo el método (a) cultivar una célula huésped recombinante como se describió anteriormente bajo condiciones conducentes a la producción del polipéptido, y (b) la recuperación del polipéptido de las células y / o del medio de cultivo.

40 En los métodos de producción de la presente invención, las células se cultivan en un medio nutriente adecuado para la producción del polipéptido usando métodos conocidos en el arte. Por ejemplo, la célula puede ser cultivada mediante cultivo en un matraz de agitación, fermentación a gran escala o pequeña escala (incluyendo fermentaciones continuas, discontinuas, intermitentes o de estado sólido) en fermentadores de laboratorio o industriales realizada en un medio adecuado y bajo condiciones que permiten expresar y/o aislar el polipéptido. El cultivo e lleva a cabo en un medio nutriente adecuado que comprende fuentes de carbono y de nitrógeno y sales inorgánicas, usando procedimientos conocidos en el arte. Los medios adecuados están disponibles a través de proveedores comerciales o se pueden preparar de acuerdo con composiciones publicadas. Si el polipéptido es secretado en el medio nutriente, el polipéptido puede ser recuperado directamente del medio. Si el polipéptido no es secretado, puede ser recuperado de los lisados celulares.

Los polipéptidos pueden ser detectados usando métodos conocidos en el arte que son específicos para los polipéptidos.

Estos métodos de detección pueden incluir el uso de anticuerpos específicos, la formación de un producto enzimático, o la desaparición de un sustrato enzimático.

5 El polipéptido resultante puede ser recuperado por métodos conocidos en el arte. Por ejemplo, el polipéptido puede ser recuperado del medio nutriente por medio de procedimientos convencionales incluyendo, por ejemplo, centrifugación, filtración, extracción, deshidratación por dispersión, evaporación, o precipitación.

10 Los polipéptidos de la presente invención pueden ser purificados por medio de una variedad de procedimientos conocidos en el arte incluyendo, por ejemplo, cromatografía (por ejemplo, de intercambio iónico, afinidad, hidrófoba, cromatografía de exclusión por tamaño), procedimientos electroforéticos (por ejemplo, enfoque isoelectrico preparativo), solubilidad diferencial (por ejemplo, precipitación con sulfato de amonio), SDS-PAGE, o extracción (véase, por ejemplo, Protein Purification, J.C. Janson y Lars Ryden, editores, VCH Publishers, Nueva York, 1989). Los polipéptidos de la presente invención pueden requerir de purificación adicional. Las técnicas se aplican según sea necesario, incluyendo, sin limitación, FPLC (Pharmacia, Uppsala, Suecia), HPLC (por ejemplo, usando filtración en gel, columnas de fase inversa o ligeramente hidrófobas).

Uso en detergente

15 El polipéptido de la invención puede ser añadido al y por lo tanto convertirse en un componente de una composición detergente, (Patente de los Estados Unidos No. 5.693.520).

20 La composición detergente de la invención puede ser formulada por ejemplo, como una composición detergente para lavado manual o de máquina, que incluye una composición aditiva para lavado adecuada para pretratamiento de tejidos manchados, y una composición suavizante de tejidos añadida como enjuague. Alternativamente, la composición detergente de la presente invención se puede utilizar en operaciones de limpieza generales de superficies duras en el hogar, o se puede formular para operaciones de lavado de vajillas en forma manual o en máquina.

25 En un aspecto específico, la invención proporciona un aditivo detergente que comprende la enzima de la invención. El aditivo detergente así como la composición detergente pueden comprender una o más de otras enzimas tales como una lipasa, una cutinasa, una amilasa, una carbohidrasa, una celulasa, una pectinasa, una mananasa, una arabinasa, una galactanasa, una xilanas, una oxidas, por ejemplo, una lacasa, y/o una peroxidasa.

En general, las propiedades de la(s) enzima(s) escogida(s) deben ser compatibles con el detergente seleccionado, (es decir pH óptimo, compatibilidad con otros ingredientes enzimáticos y no enzimáticos, etc.), y la(s) enzima(s) debe(n) estar presente(s) en la composición detergente.

Otros usos

30 Las proteínas del tipo de la subtilisina son ampliamente utilizadas en la industria del cuero. Es preferible el biotratamiento de cuero utilizando un enfoque enzimático, ya que ofrece varias ventajas (Varela H., Ferrari M.D., Belobradic L, Weyrauch R., Loperena M.L. (1996). World J Microbiol. Biotechnol. 12: 643-645)

La presente invención también puede ser utilizada en sistemas de limpieza y desinfección de lentes de contacto (véase por ejemplo, la Patente de los Estados Unidos No 6.358.897, la patente de los Estados Unidos No. 6.139.646).

35 Además, la presente invención puede ser utilizada en tratamientos tópicos, en el tratamiento de cicatrices de la piel y en infecciones de la piel (Jung H. (1998) Facial Plast Surg. 14 (4): 255-257).

En la industria de alimentos, por ejemplo, se han utilizado aminopeptidasas para aumentar la velocidad en el proceso de maduración de los quesos. Se han utilizado muchas proteasas en la preparación de hidrolizados de proteínas de alto valor nutricional.

40 Una de las áreas menos exploradas para el uso de las proteasas es la industria de la seda y se han solicitado muy pocas patentes que describen el uso de las proteasas para el desgomado de la seda (véase, por ejemplo, la patente de los Estados Unidos No. 6.080.689).

45 Recientemente, el uso de proteasa alcalina en el manejo de los residuos de diferentes industrias de procesamiento de alimentos y actividades del hogar abrieron una nueva era en el uso de proteasas en el manejo de los residuos (Dalev P. G. (1994). Bioresour Technol. 48: 265-267).

Otro sector de aplicación menos conocido es el uso de las proteasas para limpieza de sistemas de membrana (patente de los Estados Unidos No. 6.387.189). Las proteasas de la presente invención pueden ser utilizadas en una cualquiera o más de las aplicaciones mencionadas aquí.

Materiales y métodos

La presente invención se ejemplifica adicionalmente mediante los siguientes ejemplos no limitantes.

Ejemplo 1

Purificación de proteína y análisis espectrométrico de masas

5 En resumen, se cultivó la cepa 3-17 de *Polaribacter* en un fermentador aireado durante 48 horas a 10°C. Después de la precipitación del cultivo, se añadió (NH₄)₂SO₄ 1 M a la fracción sobrenadante y se incubó con resina de flujo rápido Fenil-Sefarosa (Pharmacia Biotech) a 4°C durante 2 h. Se empacó la resina con proteasa absorbida en una columna cromatográfica y luego se la sometió a cromatografía de interacción hidrófoba. Se utilizó Tris / HCl 0,05 M, CaCl₂ 2 mM, (NH₄)SO₄ 1 M pH 7,5 como regulador de equilibrio y se utilizó Tris / HCl 0,05 M, CaCl₂ 2 mM, 2-Propanol al 7,5%, como regulador de elución. Se desalinizó la fracción cromatográfica con actividad de proteasa en una columna Sephadex G-25, y se sometió a cromatografía de afinidad con bacitracina-Sefarosa, que fue realizada mediante el método de Stepanov y Rudenskaya (Stepanov V.M., Rudenskaya G.N. (1983). J Appl Biochem. 5: 420-428). A esta proteína purificada, se le realizó un ensayo de actividad del tipo de la subtilisina a 20°C, se analizó la proteína en la fracción con la mayor actividad de proteasa por espectrometría de masas LC / MS / MS LCQ Deca XP (trampa de iones) por parte del CHUL Research Center, Quebec, y a partir de este análisis se calculó la secuencia de algunos péptidos a partir de sus espectros MS / MS (SEQ ID NOS: 3, 4, 5).

Ejemplo 2

Clonación del ADN de la proteína caracterizada de la invención

Manipulación de ácido nucleico

20 Se aislaron bacterias psicrófilas del agua de mar recolectada en la Base Frei Montalva (Latitud 62° 11" Sur, longitud 58° 58" Oeste), isla Rey Jorge, Antártico Chileno. Las células para manipulación de ADN fueron recolectadas por centrifugación a 4°C después de 4 días de cultivo en medio marino 2216 (BD 279110) con agitación a 4°C. La manipulación del ADN se realizó como se describe en Sambrook y colaboradores. Molecular Cloning, a Laboratory Manual (Cold Spring Harbor Press, 1989). Se purificaron los productos de la PCR a partir de gel de agarosa después de electroforesis mediante QIAEXII suministrado por QIAGEN Inc. (CA, EE.UU.). Se clonaron los productos purificados de la PCR en el vector pGEM-T Easy (Promega, WI, EE.UU.), y secuenciados por Macrogen (Corea). Los cebadores y las enzimas de restricción fueron suministradas por Invitrogen (CA, EE.UU.) y New England Biolabs (MA, EE.UU.), respectivamente. La Taq polimerasa y la Elongasa fueron adquiridas a través de Promega e Invitrogen, respectivamente.

30 Diseño del cebador degenerado

Se diseñaron tres cebadores degenerados (SEQ ID NOS: 6, 7, 8) con base en la secuencia de aminoácidos previamente calculada de los péptidos (SEQ ID NOS: 3, 4, 5). El cebador sentido SEQ ID NO: 6 y el cebador antisentido SEQ ID NO: 7 fueron capaces de amplificar una región central del gen que codifica las proteínas del tipo de la subtilisina (SEQ ID NO: 9). Las amplificaciones se clonaron en un sistema pGEM-T (Promega) y los clones seleccionados fueron secuenciados automáticamente. Para completar el resto del gen que codifica la proteína del tipo de la subtilisina, se implementó un nuevo método de desplazamiento sobre el genoma.

Método de desplazamiento sobre el Genoma

1) Construcción de un casete de oligonucleótido:

40 Se construyó un adaptador de AdaptT de un casete de oligonucleótido bicatenario mediante hibridación de los dos cebadores no fosforilados AdaptF y AdaptR (SEQ ID NOS: 10, 11). La hibridación se realizó mediante el calentamiento de los cebadores (10 µM) en un baño de agua en ebullición, y luego se enfrió lentamente hasta temperatura ambiente. Este casete tiene una timidina que sobresale en 3'.

2) Construcción de bibliotecas de casetes de oligonucleótidos:

45 Para la construcción de fragmentos de ADN enlazados al AdaptT del casete de oligonucleótidos, se digirió 1 µg de ADN genómico con 10 unidades de actividad de una enzima de restricción (HindIII, XbaI, EcoRV, EcoRI, Sau3AI, PvuII) y 2 µl de regulador de reacción de la enzima 10X en 20 µl de volumen de reacción. Para completar el extremo recesivo 3' de los fragmentos y añadir una adenina que sobresale en 3', se incubaron 500 ng del ADN digerido y purificado con 5 unidades de actividad de la Taq ADN polimerasa, 1 µl de mezcla dNTP 10 mM (dATP, dTTP, dGTP y dCTP) y 5 µl de

regulador ADN polimerasa termofílico 10X en un volumen total de 50 µl, en 70°C durante 45 min. Se incubaron 7 µl de esta mezcla con 15 µmoles del casete de oligonucleótidos AdaptT, 1 unidad de la T4 ADN ligasa (Invitrogen) y 2 µl de regulador de ligasa 5x, en un volumen total de 10 µl. Se incubó la reacción de ligación a 16°C durante la noche.

3) Primera ronda de PCR

- 5 Se realizó la reacción de amplificación en un volumen de 50 µl con regulador de mezcla Elongasa 1X, MgCl₂ 1,9 mM, dNTP 0,2 mM, primer cebador específico 0,5 µM (diseñado a partir de la secuencia conocida del gen objetivo (SEQ ID NO: 9); debe ser un cebador directo para amplificar el extremo 3' (SEQ ID NO: 13) o un cebador inverso para amplificar el extremo 5' (SEQ ID NO: 14)), 5 µl del ADN ligado diluido 10 veces y 1 µl de Elongasa. Las condiciones de los ciclos térmicos fueron: 1 ciclo a 94°C durante 1 min, 20 ciclos de 94°C durante 32 segundos y 68°C durante 5 min, y 1 ciclo adicional final a 70°C durante 7 min. Las reacciones se llevaron a cabo en un gradiente Eppendorf Master Cycler (HA, GE). Se diluyó el producto de la PCR 10 veces y se utilizaron 3 µl como molde de ADN para la segunda PCR.

4) Segunda ronda de PCR

- 15 Se realizó la segunda reacción de amplificación en un volumen total de 50 µl de regulador de mezcla de Elongasa 1X, donde las concentraciones finales fueron: MgCl₂ 1,9 mM, dNTP 0,2 mM, segundo cebador específico 0,5 µM (diseñado a partir de la secuencia conocida del gen objetivo (SEQ ID NO: 9); debe ser un cebador directo para amplificar el extremo 3' (SEQ ID NO: 15) o un cebador inverso para amplificar la 5' (SEQ ID NO: 16)), un cebador específico del casete de oligonucleótido 0,2 µM AdaptF2 (SEQ ID NO: 12), 3 µl del producto diluido de la primera PCR y 1 µl de Elongasa. Las condiciones de los ciclos térmicos fueron: 1 ciclo a 94°C durante 1 min, 35 ciclos de 94°C durante 32 segundos y 68°C durante 5 min, y 1 ciclo adicional final a 70°C durante 7 min.

- 20 5) Construcción de la secuencia completa de ácido nucleico:

- 25 Las amplificaciones de la segunda PCR (5' y 3') fueron clonadas en un sistema pGEM-T (Promega) y se secuenciaron automáticamente los clones seleccionados. Mediante la superposición de las secuencias previamente amplificadas, fue posible obtener la secuencia completa de nucleótidos del gen que codifica la proteína del tipo de la subtilisina. Con el fin de estar seguros acerca de la secuencia correcta del gen, se diseñaron dos cebadores a partir de los extremos del gen y se realizó otra amplificación se realizó utilizando una ADN polimerasa de alta fidelidad. Se clonaron las amplificaciones en un sistema pGEM-T (Promega) y se secuenciaron automáticamente los clones seleccionados.

Ejemplo 3

Construcción del vector de expresión y expresión de la proteína de la invención.

- 30 Con las secuencias obtenidas con el nuevo método de desplazamiento sobre el genoma, se diseñaron dos cebadores (SEQ ID NOS: 17, 18) y se llevó a cabo una PCR final. Además, se completó la clonación y secuenciación del gen completo que codifica la proteína purificada de la presente invención. Se analizaron ocho clones, que representan al mismo gen (SEQ ID NO: 1). Se expresó la proteína del tipo de la subtilisina adaptada al frío en *E. coli* utilizando los sitios de *NcoI* y *Xho I* de un vector pET22b suministrado por Novagen. El vector pET coloca la proteína recombinante bajo el control de las señales de transcripción y traducción del bacteriófago T7. Una vez establecido en un huésped sin expresión, *E. coli* DH5α, se transfirió luego el plásmido a un huésped de expresión, *E. coli* BL21 (DE3) pLys S que tienen una copia cromosómica del gen de la T7 polimerasa bajo el control de lacUV5. Se indujo la expresión mediante la adición de IPTG.

Secuencias

- 40 Las secuencias de ácido nucleico descritas aquí, y en consecuencia las secuencias de proteínas derivadas de las mismas, han sido cuidadosamente secuenciadas. Sin embargo, aquellos ordinariamente capacitados en el arte se darán cuenta que la tecnología de secuenciación de ácido nucleico puede ser susceptible a algún error inadvertido. Aquellos ordinariamente capacitados en las materias relevantes son capaces de validar o corregir estas secuencias con base en la descripción amplia suministrada aquí de los métodos de aislamiento de las secuencias de ácido nucleico en cuestión y tales modificaciones que se encuentran fácilmente disponibles mediante la presente divulgación son abarcadas por la presente invención. Además, se cree que aquellas secuencias reportadas aquí definen macromoléculas biológicas funcionales en la invención ya sea que estudios posteriores de clarificación identifiquen o no errores en la secuenciación.

Listado de secuencias

<110> Asenjo, Juan A

Andrews, Barbara A

Acevedo, Juan Pablo

Reyes, Fernando

Burzio, Luis

5 <120> PROTEÍNA Y SECUENCIA DE ADN QUE CODIFICA UNA ACTIVIDAD DEL TIPO DE LA SUBTILISINA ADAPTADA AL FRÍO

<130> UOCH-1001USP

<160> 19

<170> PatentIn versión 3.4

10 <210> 1

<211> 3390

<212> ADN

<213> Cepa 3-17 de Polaribacter sp.

<400> 1

```

atgaaaaaaaaa ggtacattaa tttacttctt acaattggag tttttatgat ttctgctttc      60
aacatgaatg ctcaaaaaca agaagaatta acaaaaatta gcagtaagta caatcaagaa      120
aaacttacta cggttaaaaaa tgatttttaa cagaaggctt ctttagataa acaaaatgca      180
attacaattg caaagagtaa aggatggaaa actagattta ccaataagaa aggtgaatta      240
ttagaaattc aaaaagtagt aaatggaaaa ccaatttatt ataccacttt taatgttgca      300
gccgcaaaat ctacaagaac aaatcattta aacaacgggtg gttcttttagg cttaaatttg      360
atggggcaaa atatgactgc tcatgtttgg gatggcggat tagcaagagc atctcaccaa      420
gaatatgatg gtgctggtgg tacaatatga ttctctattg gagatggcac aacagcttta      480
cactaccatt ctgctcacgt aacaggtaca attatggctt ctggtgttgt tgcaaatgca      540
aaaggaatgg cgcctcatgc aagtgctgtt gggtatgatt ggaataatga cacttctgaa      600
gctataaacg cagcttcaaa cggaatgtta gtttctaate attcttatgg ttttgctaca      660
agaaatgcac aaggccaacc tcaacttcca gattattatt ttggagggtta cattacagac      720
tctagagatt gggataacat tatgtttaat gcaccaaact atttaatggt tggtgcagca      780
ggaaatgatg gaaatgataa ttctgctaata ggtgctccat tagctggaaa ttcttcttat      840
gacaaattat ctggtcatgc aactgcaaaa aacggtatgg ttgttgccaa cgcaaatgat      900
gctaatatag atgtaaattg aaacctgctt tccgttacta taaattcttc tagtagtgaa      960
ggaccaacag atgattaccg tattaaacca gatattactg gaaacggaac atctgtatat      1020
tctacttatt cgtctagtaa tacagcttat aatagtacta ctggtacttc tatggcatcg      1080
ccaaatggtg cgggtacact attaatTTTA caacaacatg ctaacaatgt tagagggttcg      1140

```

15

ES 2 562 919 T3

tttattaaag	cttcaacttt	aaaaggaatt	gctttacata	ctgcagatga	cgcaggttct	1200
aatggaccag	atgcaatttt	tggtctggga	ttaatgaatg	ctaaaagagc	tgctgtagca	1260
attactcaaa	acggtactga	atctaagatt	gaagaactaa	ctttatctag	cagacaaacg	1320
tatcaaatta	ctgtagatgc	tgatggagtt	aatgatttaa	tggtctctat	ttcttggaca	1380
gatagagctg	gaactgcaac	tactacagca	aattcaagta	ctgctgtttt	agtaaagtat	1440
ttagatatta	gagtttctaa	aaacggaaca	acctatactc	cttggagatt	aacaggagta	1500
acaacaaatg	gaaaagggtga	taatactgta	gatccttatg	aaagagttga	tgttgctaac	1560
gcttcaggaa	cttatacaat	tactgtaact	cataaagggt	ctttaactgg	cggaagtcaa	1620
aattattcat	taattgtaac	tggttttagca	ggaacgcctg	ttgtatgtaa	tgcaacaatc	1680
ccttctaate	ttactgttga	tgaatctgga	gcttcaaccg	caactgtttc	ttggaatact	1740
gtcgctggaa	cttcttacga	ttttagatac	cgcaaaacag	gcacgtcaac	ctggacaact	1800
tctgctgttg	cggaacttc	ggtttcttta	acaggattat	ctactcaaac	ttcttatcaa	1860
actcaagtaa	gaagcaaag	tcctaataat	tctacctcag	catattctag	tgctgttagt	1920
ttacaactt	cagatgttca	gcttaattat	tgtgcttcaa	acggaaatag	tgttgcagat	1980
gaatacataa	gtaaagtagt	tcttggaggt	ataaacaata	caaccggagc	ttcatcaagc	2040
ggatacgtg	attacacctc	tcaatctaca	agtttaacga	aaggagtttc	ttcaacaatt	2100
acaattaccc	caacctgggc	aggagcttca	tataacgaag	gttatgctgt	atattattgac	2160
tataataagg	atggtgattt	tacagataat	ggagaaaccg	tttggaacaa	aatagcttct	2220
aaaacaaaac	ctgttagcgg	ttcatttact	gtgccaacat	ctgcaactac	aggagcaact	2280
agaatgcgtg	tagtaatgca	atacaatacc	gtacctgctg	cttgtggaac	ctataattat	2340
ggtgaaacag	aagattatac	tgtaaacata	accggaagta	gtgcagatac	aatagcacca	2400
actgcgccta	caaagtgtac	agcttcagct	attacceaaa	ctacggctac	attatcttgg	2460
acagcatcta	cagataacgt	aggagtagca	ggatacgaga	tatttagtaa	cggaacaagt	2520
gttgaaccg	taacagcaac	ttctgctaac	ataactgggt	taacagcaaa	tacttcatac	2580
tcatatacta	taaaagctaa	tgatgcagca	gagaacacat	ctaactcaag	taatagtgtg	2640
tcattttaca	cattaggaag	tacgctagta	tattgttcct	ctaaaggaaa	tagagtaact	2700
tatgaatgga	tcgattatgt	gagttttgga	ggaatgacaa	atacaactgc	agcaaacgca	2760
ggatatggag	attttacttc	aaaaacagca	acagtatcta	aagggaagtga	taaccaactg	2820
ataataagtg	caggttttgc	aagtactgca	tatacagaac	attgggcagt	ttggatcgat	2880
tttaatcaaa	acggaacttt	tgaagaaagc	gaaaaagtta	cttctggttc	ttcttctagt	2940
gcggctaatt	taactgcaac	tatttcaatt	ccttcttcag	ctaatactgg	tcaacaaga	3000

ES 2 562 919 T3

```

atgCGTgTtt caatgaaata caatagtgcg caaacagctt gtgaaacatt ttctgatgga 3060
gaagtagaag actacacagt aaatattaca aacgctacag caaattatac tacatttatt 3120
aatactaatt ctaaaaatga attaggaaat gaaagtaaag cattcgattt tacagtatat 3180
cctaaccctg taaaaggaac tgttttaaac attcacttaa atgatgctag agaagttaac 3240
tttgcaatta caaacatggt agggcaaacc ttaaaaagt gtattttaac aaaacaacct 3300
atagatgtta gtactattaa aacaggtggt tacatgttag aaataactga tggacaaaag 3360
tctgttggtta aaaaattcgt tagacaataa 3390

```

<210> 2

<211> 1129

<212> PRT

5 <213> Cepa 3-17 de Polaribacter sp.

<400> 2

```

Met Lys Lys Arg Tyr Ile Asn Leu Leu Leu Thr Ile Gly Val Phe Met
1           5           10           15

Ile Ser Ala Phe Asn Met Asn Ala Gln Lys Gln Glu Glu Leu Thr Lys
20           25           30

Ile Ser Ser Lys Tyr Asn Gln Glu Lys Leu Thr Thr Leu Lys Asn Asp
35           40           45

Phe Lys Gln Lys Ala Ser Leu Asp Lys Gln Asn Ala Ile Thr Ile Ala
50           55           60

Lys Ser Lys Gly Trp Lys Thr Arg Phe Thr Asn Lys Lys Gly Glu Leu
65           70           75           80

Leu Glu Ile Gln Lys Val Val Asn Gly Lys Pro Ile Tyr Tyr Thr Thr
85           90           95

Phe Asn Val Ala Ala Ala Lys Ser Thr Arg Thr Asn His Leu Asn Asn
100          105          110

Gly Gly Ser Leu Gly Leu Asn Leu Met Gly Gln Asn Met Thr Ala His
115          120          125

Val Trp Asp Gly Gly Leu Ala Arg Ala Ser His Gln Glu Tyr Asp Gly
130          135          140

Ala Gly Gly Thr Asn Arg Phe Ser Ile Gly Asp Gly Thr Thr Ala Leu
145          150          155          160

```

ES 2 562 919 T3

His	Tyr	His	Ser	Ala	His	Val	Thr	Gly	Thr	Ile	Met	Ala	Ser	Gly	Val	165	170	175	
Val	Ala	Asn	Ala	Lys	Gly	Met	Ala	Pro	His	Ala	Ser	Ala	Val	Gly	Tyr	180	185	190	
Asp	Trp	Asn	Asn	Asp	Thr	Ser	Glu	Ala	Ile	Asn	Ala	Ala	Ser	Asn	Gly	195	200	205	
Met	Leu	Val	Ser	Asn	His	Ser	Tyr	Gly	Phe	Ala	Thr	Arg	Asn	Ala	Gln	210	215	220	
Gly	Gln	Pro	Gln	Leu	Pro	Asp	Tyr	Tyr	Phe	Gly	Gly	Tyr	Ile	Thr	Asp	225	230	235	240
Ser	Arg	Asp	Trp	Asp	Asn	Ile	Met	Phe	Asn	Ala	Pro	Asn	Tyr	Leu	Met	245	250	255	
Val	Val	Ala	Ala	Gly	Asn	Asp	Gly	Asn	Asp	Asn	Ser	Ala	Asn	Gly	Ala	260	265	270	
Pro	Leu	Ala	Gly	Asn	Ser	Ser	Tyr	Asp	Lys	Leu	Ser	Gly	His	Ala	Thr	275	280	285	
Ala	Lys	Asn	Gly	Met	Val	Val	Ala	Asn	Ala	Asn	Asp	Ala	Asn	Ile	Asp	290	295	300	
Val	Asn	Gly	Asn	Leu	Leu	Ser	Val	Thr	Ile	Asn	Ser	Ser	Ser	Ser	Glu	305	310	315	320
Gly	Pro	Thr	Asp	Asp	Tyr	Arg	Ile	Lys	Pro	Asp	Ile	Thr	Gly	Asn	Gly	325	330	335	
Thr	Ser	Val	Tyr	Ser	Thr	Tyr	Ser	Ser	Ser	Asn	Thr	Ala	Tyr	Asn	Ser	340	345	350	
Thr	Thr	Gly	Thr	Ser	Met	Ala	Ser	Pro	Asn	Val	Ala	Gly	Thr	Leu	Leu	355	360	365	
Ile	Leu	Gln	Gln	His	Ala	Asn	Asn	Val	Arg	Gly	Ser	Phe	Ile	Lys	Ala	370	375	380	
Ser	Thr	Leu	Lys	Gly	Ile	Ala	Leu	His	Thr	Ala	Asp	Asp	Ala	Gly	Ser	385	390	395	400

ES 2 562 919 T3

Asn	Gly	Pro	Asp	Ala	Ile	Phe	Gly	Trp	Gly	Leu	Met	Asn	Ala	Lys	Arg	405	410	415	
Ala	Ala	Val	Ala	Ile	Thr	Gln	Asn	Gly	Thr	Glu	Ser	Lys	Ile	Glu	Glu	420	425	430	
Leu	Thr	Leu	Ser	Ser	Arg	Gln	Thr	Tyr	Gln	Ile	Thr	Val	Asp	Ala	Asp	435	440	445	
Gly	Val	Asn	Asp	Leu	Met	Ala	Ser	Ile	Ser	Trp	Thr	Asp	Arg	Ala	Gly	450	455	460	
Thr	Ala	Thr	Thr	Thr	Ala	Asn	Ser	Ser	Thr	Ala	Val	Leu	Val	Asn	Asp	465	470	475	480
Leu	Asp	Ile	Arg	Val	Ser	Lys	Asn	Gly	Thr	Thr	Tyr	Thr	Pro	Trp	Arg	485	490	495	
Leu	Thr	Gly	Val	Thr	Thr	Asn	Gly	Lys	Gly	Asp	Asn	Thr	Val	Asp	Pro	500	505	510	
Tyr	Glu	Arg	Val	Asp	Val	Ala	Asn	Ala	Ser	Gly	Thr	Tyr	Thr	Ile	Thr	515	520	525	
Val	Thr	His	Lys	Gly	Ser	Leu	Thr	Gly	Gly	Ser	Gln	Asn	Tyr	Ser	Leu	530	535	540	
Ile	Val	Thr	Gly	Leu	Ala	Gly	Thr	Pro	Val	Val	Cys	Asn	Ala	Thr	Ile	545	550	555	560
Pro	Ser	Asn	Leu	Thr	Val	Asp	Glu	Ser	Gly	Ala	Ser	Thr	Ala	Thr	Val	565	570	575	
Ser	Trp	Asn	Thr	Val	Ala	Gly	Thr	Ser	Tyr	Asp	Phe	Arg	Tyr	Arg	Lys	580	585	590	
Thr	Gly	Thr	Ser	Thr	Trp	Thr	Thr	Ser	Ala	Val	Ala	Gly	Thr	Ser	Val	595	600	605	
Ser	Leu	Thr	Gly	Leu	Ser	Thr	Gln	Thr	Ser	Tyr	Gln	Thr	Gln	Val	Arg	610	615	620	
Ser	Lys	Cys	Pro	Asn	Asn	Ser	Thr	Ser	Ala	Tyr	Ser	Ser	Ala	Val	Ser	625	630	635	640

ES 2 562 919 T3

Phe Thr Thr Ser Asp Val Gln Leu Asn Tyr Cys Ala Ser Asn Gly Asn
 645 650 655
 Ser Val Ala Asp Glu Tyr Ile Ser Lys Val Val Leu Gly Gly Ile Asn
 660 665 670
 Asn Thr Thr Gly Ala Ser Ser Ser Gly Tyr Ala Asp Tyr Thr Ser Gln
 675 680 685
 Ser Thr Ser Leu Thr Lys Gly Val Ser Ser Thr Ile Thr Ile Thr Pro
 690 695 700
 Thr Trp Ala Gly Ala Ser Tyr Asn Glu Gly Tyr Ala Val Phe Ile Asp
 705 710 715 720
 Tyr Asn Lys Asp Gly Asp Phe Thr Asp Asn Gly Glu Thr Val Trp Thr
 725 730 735
 Lys Ile Ala Ser Lys Thr Lys Pro Val Ser Gly Ser Phe Thr Val Pro
 740 745 750
 Thr Ser Ala Thr Thr Gly Ala Thr Arg Met Arg Val Val Met Gln Tyr
 755 760 765
 Asn Thr Val Pro Ala Ala Cys Gly Thr Tyr Asn Tyr Gly Glu Thr Glu
 770 775 780
 Asp Tyr Thr Val Asn Ile Thr Gly Ser Ser Ala Asp Thr Ile Ala Pro
 785 790 795 800
 Thr Ala Pro Thr Asn Val Ser Ala Ser Ala Ile Thr Gln Thr Thr Ala
 805 810 815
 Thr Leu Ser Trp Thr Ala Ser Thr Asp Asn Val Gly Val Ala Gly Tyr
 820 825 830
 Glu Ile Phe Ser Asn Gly Thr Ser Val Gly Thr Val Thr Ala Thr Ser
 835 840 845
 Ala Asn Ile Thr Gly Leu Thr Ala Asn Thr Ser Tyr Ser Tyr Thr Ile
 850 855 860
 Lys Ala Asn Asp Ala Ala Glu Asn Thr Ser Asn Ser Ser Asn Ser Val
 865 870 875 880
 Ser Phe Thr Thr Leu Gly Ser Thr Leu Val Tyr Cys Ser Ser Lys Gly

ES 2 562 919 T3

[illegible]

Gln

<210> 3

<211> 13

<212> PRT

5 <213> Cepa 3-17 de Polaribacter sp.

<220>

<221> CARACTERÍSTICA NUEVA

<222> (1) .. (4)

<223> Xaa es Leu o Ile

10 <400> 3

Xaa	Xaa	Xaa	Xaa	Leu	Gln	Gln	His	Ala	Asn	Asn	Val	Arg
1				5					10			

<210> 4

<211> 8

<212> PRT

15 <213> Cepa 3-17 de Polaribacter sp.

<220>

<221> CARACTERÍSTICA NUEVA

<222> (8) .. (8)

<223> Xaa es Leu o Ile

20 <400> 4

Asn	Ala	Ser	Gly	Thr	Tyr	Thr	Xaa
1				5			

<210> 5

<211> 7

<212> PRT

25 <213> Cepa 3-17 de Polaribacter sp.

<400> 5

Asn	Gly	Thr	Thr	Tyr	Thr	Pro
1				5		

<210> 6

<211> 27

<212> ADN

<213> sintética

<400> 6

5 wtwcaacaac atgcwaayaa ygtwaga

<210> 7

<211> 21

<212> ADN

<213> sintética

10 <400> 7

wgtrtawgtw ccwgawgcrt t 21

<210> 8

<211> 21

<212> ADN

15 <213> sintética

<220>

<221> Característica nueva

<222> (6) .. (6)

<223> n es a, c, g, o t

20 <220>

<221> Característica nueva

<222> (9) .. (9)

<223> n es a, c, g, o t

<220>

25 <221> Característica nueva

<222> (12) .. (12)

<223> n es a, c, g, o t

<220>

<221> Característica nueva

30 <222> (18) .. (18)

<223> n es a, c, g, o t

<220>

<221> Característica nueva

<222> (21) .. (21)

<223> n es a, c, g, o t

5 <400> 8

aayggnacna cntayacncc n 21

<210> 9

<211> 471

<212> ADN

10 <213> Cepa 3-17 de Polaribacter sp.

<400> 9

ttacaacaac atgctaacaa tgtagaggt tcgtttatta aagcttcaac tttaaaagga 60

attgctttac atactgcaga tgacgcaggt tctaattggac cagatgcaat ttttggtgg 120

ggattaatga atgctaaaag agctgctgta gcaattactc aaaacggtac tgaatctaag 180

attgaagaac taactttatc tagcagacaa acgtatcaaa ttactgtaga tgctgatgga 240

gttaatgatt taatggcttc tatttcttgg acagatagag ctggaactgc aactactaca 300

gcaaattcaa gtactgctgt tttagtaaatt gatttagata ttagagtttc taaaaacgga 360

acaacctata ctcttggag attaacagga gtaacaacaa atggaaaagg tgataatact 420

gtagatcctt atgaaagagt tgatgttgct aacgcttcag gaacttatac a 471

<210> 10

15 <211> 29

<212> ADN

<213> sintética

<400> 10

ctaggccacg cgtagactag tactagctt 29

20 <210> 11

<211> 28

<212> ADN

<213> sintética

<400> 11

	agctagtact agtcgacgcg tggcctag	28
	<210> 12	
	<211> 23	
	<212> ADN	
5	<213> sintética	
	<400> 12	
	cacgcgtcga ctagtactag ctt	23
	<210> 13	
	<211> 26	
10	<212> ADN	
	<213> sintética	
	<400> 13	
	ctaatggacc agatgcaatt ttggc	26
	<210> 14	
15	<211> 31	
	<212> ADN	
	<213> sintética	
	<400> 14	
	gctgtagtag ttgcagttcc agctctatct g	31
20	<210> 15	
	<211> 32	
	<212> ADN	
	<213> sintética	
	<400> 15	
25	gaaagagttg atgttgctaa cgcttcagga ac	32
	<210> 16	
	<211> 29	
	<212> ADN	
	<213> sintética	
30	<400> 16	
	ccattagaac ctgcgtcatc tgcagtatg	29

<210> 17

<211> 35

<212> ADN

<213> sintética

5 <400> 17

catatgaaaa aaaggtacat taatttactt cttac 35

<210> 18

<211> 32

<212> ADN

10 <213> sintética

<400> 18

ctcgagttgt ctaacgaatt tttaacaac ag 32

<210> 19

<211> 897

15 <212> ADN

<213> Cepa 3-17 de Polaribacter sp.

<400> 19

caatggagga gactctgatc cagccatgcc gcgtgtagga agaatgccct atgggttgta	60
aactactttt atacaggaag aaacactagt atgtatacta gcttgacggg actgtaagaa	120
taaggaccgg ctaactccgt gccagcagcc gcggaatac ggaggggtcca agcggttatcc	180
ggaatcattg ggtttaaagg gtccgcaggc ggtcaattaa gtcagagggtg aaatcccata	240
gcttaactat ggaactgcct ttgatactgg ttgacttgag tcatatggaa gtagatagaa	300
tgtgtagtgt agcgggtgaaa tgcatagata ttacacagaa taccgattgc gaaggcagtc	360
tactacgtat gtactgacgc tgagggacga aagcgtgggg agcgaacagg attagatacc	420
ctggtagtcc acgccgtaaa cgatggatac tagttggttg gatttatctc agtgactaag	480
cgaaagtgat aagtatccca cctggggagt acggtcgcaa gactgaaact caaaggaatt	540
gacggggggcc cgcacaagcg gtggagcatg tggtttaatt cgatgatacg cgaggaacct	600
taccagggct taaatgtagt ctgacagctt tagagataga gttttcttcg gacagattac	660
aagggtgctgc atgggtgtcg tcagctcgtg ccgtgagggtg tcagggttaag tcctataacg	720
agcgcaaccc ctgtcgttag ttgccagcat gttatgatgg ggactctaac gagactgcct	780
acgcaagtag agaggaaggt ggggatgacg tcaaatacatc acggccctta cgtcctgggc	840
cacacacgtg ctacaatggg atggacaatg agcagccatc ggcaacagag agcgaat	897

REIVINDICACIONES

1. Un ácido nucleico aislado que comprende:
 - a) una secuencia de nucleótidos idéntica a la SEQ ID NO: 1 o una variante degenerada de la misma; o
 - 5 b) una secuencia de nucleótidos que comprende posiciones idénticas 361-423, 466-528 y 1039-1101 de SEQ ID NO: 1 o variante degenerada de la misma.
2. El ácido nucleico aislado como se reivindica en la reivindicación 1, que comprende la secuencia de nucleótidos como se muestra en:
 - a) posiciones 73-3390 de la SEQ ID NO: 1 o una variante degenerada de la misma; o
 - 10 b) posiciones 286-3390 de la SEQ ID NO: 1 o una variante degenerada de la misma; o
 - c) posiciones 286-2610 de la SEQ ID NO: 1 o una variante degenerada de la misma; o
 - d) posiciones 286-1950 de la SEQ ID NO: 1 o una variante degenerada de la misma; o
 - e) posiciones 286-1680 de la SEQ ID NO: 1 o una variante degenerada de la misma.
- 15 3. Un ácido nucleico aislado que comprende una secuencia que codifica un polipéptido, cuya secuencia de aminoácidos es al menos 85% idéntica a la SEQ ID NO: 2.
4. El ácido nucleico de acuerdo con la reivindicación 3, en donde la secuencia de aminoácidos del polipéptido es al menos 90% idéntica a la SEQ ID NO: 2.
5. Un vector de expresión que comprende el ácido nucleico como el reivindicado en cualquiera de las reivindicaciones 1-2 operativamente enlazado a una secuencia de control de la expresión.
- 20 6. Un vector de expresión que comprende el ácido nucleico de la reivindicación 3 operativamente enlazado a una secuencia de control de la expresión.
7. Una célula transformada que comprende al vector de expresión como se reivindica en cualquiera de las reivindicaciones 5-6.
8. Un polipéptido codificado por el ácido nucleico como se reivindica en cualquiera de las reivindicaciones 1-2.
- 25 9. El polipéptido como se reivindica en la reivindicación 8, que comprende la secuencia de aminoácidos dada por los siguientes residuos de aminoácidos:
 - a) residuos de aminoácidos 25-1129 de la SEQ ID NO: 2;
 - b) residuos de aminoácidos 96-1129 de la SEQ ID NO: 2;
 - c) residuos de aminoácidos 96-870 de la SEQ ID NO: 2;
 - 30 d) residuos de aminoácidos 96-650 de la SEQ ID NO: 2; o
 - e) residuos de aminoácidos 96-560 de la SEQ ID NO: 2.
10. Un método de preparación de una proteína que comprende las etapas de: transformar una célula con un vector de expresión como se reivindica en cualquiera de las reivindicaciones 5 y 6, cultivando la célula transformada en un cultivo; y el aislamiento de la proteína codificada en el vector como se reivindica en cualquiera de las reivindicaciones 5 y 6.
- 35 11. Una composición de limpieza o detergente que contiene los polipéptidos como se reivindica en cualquiera de las reivindicaciones 8 y 9.

12. El polipéptido como se reivindica en la reivindicación 9, que tiene una secuencia de aminoácidos que tiene al menos 85%, 90%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia con la secuencia de aminoácidos de uno de los siguientes residuos de aminoácido:

- 5
- a) residuos de aminoácidos 25-1129 de la SEQ ID NO: 2;
 - b) residuos de aminoácidos 96-1129 de la SEQ ID NO: 2;
 - c) residuos de aminoácidos 96-870 de la SEQ ID NO: 2;
 - d) residuos de aminoácidos 96-650 de la SEQ ID NO: 2; o
 - e) residuos de aminoácidos 96-560 de la SEQ ID NO: 2.