

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 763 577**

51 Int. Cl.:

C12N 15/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.05.2012 E 16178440 (0)**

97 Fecha y número de publicación de la concesión europea: **25.09.2019 EP 3118310**

54 Título: **Vectores de expresión para secreción mejorada de proteína**

30 Prioridad:

31.05.2011 DE 102011118032

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

29.05.2020

73 Titular/es:

**BASF SE (100.0%)
Carl-Bosch-Strasse 38
67056 Ludwigshafen am Rhein, DE**

72 Inventor/es:

**DEGERING, CHRISTIAN;
EGGERT, THORSTEN;
EVERS, STEFAN;
MAURER, KARL-HEINZ y
BONGAERTS, JOHANNES**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 763 577 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

Vectores de expresión para secreción mejorada de proteína

La invención se encuentra en el campo de la biotecnología, principalmente la síntesis microbiana de proteínas. La invención se refiere principalmente a vectores de expresión para la preparación de proteínas y propone, además, células anfitrionas que contienen vectores de expresión de este tipo. La invención se refiere, además, a procedimientos y usos de este tipo de vectores de expresión y células anfitrionas para la preparación de proteína.

Para la preparación de proteínas pueden emplearse células anfitrionas, principalmente microorganismos que expresan los genes de las proteínas de interés. El gen de una proteína de interés (transgén) por lo regular se introduce en las células anfitrionas de modo que es expresado por parte de estas. Con frecuencia se encuentra presente en un llamado vector de expresión junto con una o varias secuencias promotoras, por lo cual se hace posible la expresión del gen.

Para la producción biotecnológica a escala industrial, las células anfitrionas en cuestión se cultivan en fermentadores que están configurados de manera correspondiente a las propiedades metabólicas de las células. Durante el cultivo, las células anfitrionas metabolizan el sustrato suministrado y forman el producto deseado que habitualmente se separa de los organismos de producción después de terminar la fermentación y se purifica y/o se concentra de la pasta de fermentador y/o del medio de fermentación.

Fundamentalmente se desea obtener un rendimiento de producto tan alto como sea posible en la fermentación. El rendimiento de producto depende en este caso de varios factores; por ejemplo, las células anfitrionas forman, además del producto deseado propiamente, habitualmente una gran cantidad de otras sustancias en las cuales por lo regular no existe interés. Además, la expresión de un transgén y, por lo tanto, el rendimiento de producción depende esencialmente del sistema de expresión usado. A manera de ejemplo, la solicitud internacional de patente WO 91/02792 divulga la producción fermentativa mejorada de una proteasa alcalina a partir de *Bacillus lentus* en una cepa optimizada de *Bacillus licheniformis* bajo el control de secuencias reguladoras de gen de *Bacillus licheniformis*, principalmente del promotor de *Bacillus licheniformis*.

Para la producción industrial de proteínas, por ejemplo, de enzimas hidrolíticas, de preferencia se emplean tales células anfitrionas que son capaces de secretar grandes cantidades de la proteína en el sobrenadante de cultivo, lo cual hace redundante una ruptura de células complicada que es necesaria en el caso de la producción intracelular. Para este propósito, se emplean preferentemente aquellas células anfitrionas, por ejemplo, especies de *Bacillus*, que pueden cultivarse con nutrientes económicos en fermentaciones eficientes de densidad celular alta y son capaces de secretar varios gramos por litro de la proteína objetivo en el sobrenadante de cultivo. Habitualmente, la proteína que ha de ser secretada es expresada por vectores de expresión que han sido introducidos en la célula anfitriona y codifican la proteína que ha de ser secretada. La proteína expresada usualmente comprende un péptido señal (secuencia de señal) que lleva a cabo la exportación del mismo de las células anfitrionas. El péptido señal es habitualmente parte de la cadena de polipéptido traducida en la célula anfitriona, pero además puede disociarse a modo de post-traducción por parte de la proteína dentro o fuera de la célula anfitriona.

Precisamente para esta producción extracelular de proteínas heterólogas existen, no obstante, numerosos pasajes estrechos y una alta demanda correspondiente para optimizar los desarrollos de la secreción. Uno de estos pasajes estrechos es la selección del péptido señal que permite una exportación eficiente de la proteína efectivo fuera de la célula anfitriona. Fundamentalmente, los péptidos señal pueden recién combinarse con proteínas, principalmente enzimas. A manera de ejemplo, en la publicación de Brockmeier et al. (J. Mol. Biol. 362, páginas 393-402 (2006)) se describe la estrategia del cribado de una biblioteca de péptidos señal en el ejemplo de una cutinasa. Sin embargo, no cada péptido señal provoca también una exportación suficiente de la proteína en condiciones de fermentación, principalmente condiciones de fermentación industriales o a gran escala industrial.

La publicación WO2011015327A1 divulga un procedimiento para la preparación de nucleasas.

Las publicaciones WO9221760A1 y WO2007122175A1 divulgan variantes de la proteasa de *Bacillus lentus*.

La publicación WO2003054185A1 divulga nuevas proteasas de tipo silvestre del tipo subtilisina.

La publicación DE102007049830A1 divulga la preparación de proteasas con ayuda de permutación circular.

Por lo tanto, es objetivo de la invención mejorar la secreción de una proteína de una célula anfitriona y de esta manera incrementar el rendimiento del producto de proteína en una fermentación, en cuyo caso la proteína es una proteasa.

Es objeto de la invención un vector de expresión que comprende

a) una secuencia de promotor y

b) una secuencia de ácido nucleico que codifica para una proteína, en donde la proteína comprende un péptido señal y otra secuencia de aminoácidos y el péptido señal comprende una secuencia de aminoácidos, la cual es

idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 6 en al menos 80 %, y la otra secuencia de aminoácidos de la proteína que comprende la secuencia de aminoácidos de una proteasa.

De manera sorprendente se ha encontrado que mediante un vector de expresión que codifica para una proteína con un péptido señal de este tipo se logra una secreción mejorada de la proteína desde una célula anfitriona, la cual incluye el vector de expresión y expresa la secuencia de ácido nucleico b).

De esta manera, en configuraciones preferidas de la invención, es posible incrementar el rendimiento del producto de proteína en una fermentación.

Un vector de expresión es una secuencia de ácido nucleico que provoca que pueda expresarse la proteína en una célula anfitriona, principalmente un microorganismo. Este comprende la información genética, es decir aquella secuencia de ácido nucleico (gen) b) que codifica para la proteína.

La expresión de la secuencia de ácido nucleico es su traducción al producto o a los productos génicos codificados por esta secuencia; es decir a un polipéptido (proteína) o en varios polipéptidos (proteínas). Los términos polipéptido y proteína se usan en la presente solicitud como sinónimos. Por consiguiente, en el sentido de la presente invención, expresión designa la biosíntesis de ácido ribonucleico (ARN) y proteínas a partir de las informaciones genéticas. Por lo regular, la expresión comprende la transcripción, es decir la síntesis de un ácido ribonucleico mensajero (mRNA) ("messenger") por medio de la secuencia de ADN (ácido desoxirribonucleico) del gen y su traducción a la cadena correspondiente de polipéptido, la cual opcionalmente puede modificarse después de la traducción. La expresión de una proteína describe en consecuencia la biosíntesis de la misma a partir de las informaciones genéticas que se proporcionan según la invención en el vector de expresión.

Los vectores son elementos genéticos que se componen de ácidos nucleicos, de preferencia ácido desoxirribonucleico (ADN) y son conocidos por el especialista en el campo de la biotecnología. Al usar en bacterias estos son principalmente plásmidos especiales, es decir elementos genéticos circulares. A los vectores pueden pertenecer, por ejemplo, aquellos que se derivan de plásmidos bacterianos, de virus o de bacteriófagos o predominantemente vectores o plásmidos sintéticos que contienen elementos de muy diverso origen. Con los elementos genéticos adicionales presentes en cada caso, los vectores son capaces de establecerse en las células anfitrionas, en las cuales se han introducido preferiblemente mediante transformación, durante múltiples generaciones como unidades estables. A este respecto, es insignificante para el propósito de la invención si se establecen extracromosómicamente como unidades propias o son integradas en el cromosoma o AON cromosómico. Cuál de los numerosos sistemas se escoge depende del caso individual. Pueden ser decisivos, por ejemplo, el número de copias obtenible, los sistemas de selección disponibles, antibióticos, entre ellos ante todo las resistencias a antibióticos o la capacidad de cultivo de las células anfitrionas capaces de alojar vectores.

Los vectores de expresión pueden ser regulables, además, mediante modificaciones de las condiciones de cultivo como, por ejemplo, la densidad celular la adición de determinados compuestos. Un ejemplo de un compuesto de este tipo es el derivado de galactosa isopropil- β -D-tiogalactopiranosido (IPTG), que se usa como activador del operón de lactosa bacteriano (lac-operón).

Además, un vector de expresión comprende al menos una secuencia de ácido nucleico, de preferencia ADN, con una función de control para la expresión de la secuencia de ácido nucleico b) que codifica para la proteína (una denominada secuencia reguladora de gen). Una secuencia reguladora de gen es en este caso cualquier secuencia de ácido nucleico cuya presencia en la respectiva célula anfitriona afecta, de preferencia incrementa, la frecuencia de transcripción de la secuencia de ácido nucleico b), la cual codifica para la proteína. De preferencia se trata de una secuencia promotora ya que una secuencia de este tipo es esencial para la expresión de la secuencia de ácido nucleico b). Sin embargo, vector de expresión según la invención también puede comprender además otras secuencias reguladoras de gen, por ejemplo, una o varias secuencias potenciadoras. En el contexto de la invención, un vector de expresión comprende consecuentemente al menos una unidad funcional de la secuencia de ácido nucleico b) y un promotor (casete de expresión). No obstante, esta no tiene que estar presentes necesariamente en forma de unidad física. El promotor causa la expresión de la secuencia de ácido nucleico b) en la célula anfitriona. En el contexto de la presente invención, un vector de expresión también puede restringirse al casete de expresión puro del promotor y la secuencia de ácido nucleico b) que va a expresarse, en cuyo caso este casete de expresión puede estar presente integrado de modo extracromosómico o también cromosómico. Las configuraciones de este tipo de los vectores de expresión según la invención representan respectivamente una forma separada de realización de la invención.

Por consiguiente, la presencia de al menos un promotor es esencial para un vector de expresión según la invención. Por lo tanto, por un promotor se entiende una secuencia de ADN que hace posible la expresión regulada de un gen. Naturalmente, una secuencia promotora es un componente de un gen y frecuentemente se encuentra en su extremo 5'y, por lo tanto, antes de la región que codifica ARN. De preferencia, la secuencia de promotor se encuentra en un vector de expresión según la invención en dirección de 5' de la secuencia de ácido nucleico b) que codifica para la proteína. La propiedad más importante de un promotor es la interacción específica con al menos una proteína o polipéptido que se enlaza al ADN, que es un mediador del inicio de la transcripción del gen mediante una polimerasa de ARN y se denomina factor de transcripción. Frecuentemente participan varios factores de transcripción y/u otras proteínas en el inicio de la transcripción mediante una polimerasa de ARN. Por consiguiente, un promotor es

preferentemente una secuencia de ADN con actividad de promotor; es decir, una secuencia de ADN a la cual se enlaza al menos de manera transitoria un factor de transcripción para la iniciación de la transcripción de un gen. La potencia de un promotor es medible por medio de la frecuencia de transcripción del gen expresado; es decir, mediante la cantidad de las moléculas de ARN generadas por unidad de tiempo, principalmente moléculas de ARNm.

5 La secuencia de promotor (a) y la secuencia de ácido nucleico (b) se encuentran una detrás de la otra en el vector de expresión. Más preferiblemente, la secuencia del promotor (a) se encuentra antes de la secuencia de ácido nucleico (b) en la molécula de ácido nucleico (en la orientación 5' → 3'). Igualmente, de manera preferida, entre las dos secuencias de ácido nucleico (a) y (b) no se encuentran secuencias de ácido nucleico que disminuyan la frecuencia de transcripción de la secuencia de ácido nucleico (b) que codifica para la proteína. Todas las indicaciones anteriores se refieren a aquella hebra de ADN que incluye la secuencia de ácido nucleico (b) que codifica para la proteína (la hebra codificante) y no a la hebra de ADN codogénica correspondiente. A partir de la secuencia de ácido nucleico (b) que codifica para la proteína, por consiguiente, se encuentra la secuencia del promotor (a) de preferencia aún más corriente abajo, es decir, en dirección de 5', en esta hebra de ADN.

10 La secuencia de ácido nucleico b) codifica para la proteína que va a secretarse. Esta es aquella proteína (proteína objetivo) que debe prepararse con ayuda de un vector de expresión según la invención.

15 La proteína codificada por la secuencia de ácido nucleico b) comprende un péptido señal con una secuencia de aminoácidos que es idéntica en al menos 80 % a la secuencia de aminoácidos indicada en SEQ ID NO. 6. Se ha establecido que péptidos señal de este tipo provocan una secreción eficiente de la proteína que los contiene, principalmente de proteína recombinante. Cada vez más preferiblemente, el péptido señal comprende una secuencia de aminoácidos que es idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 6 en al menos 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % y de modo muy particularmente preferido en 100 %. De modo particularmente preferido, el péptido señal presenta una secuencia de aminoácidos que es idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 6 en al menos 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % y de modo muy particularmente preferido en 100 %.

Muy particularmente se prefieren respectivamente las secuencias idénticas en 100 %, de modo que un vector de expresión correspondientemente preferido se caracteriza porque el péptido señal codificado por la secuencia de ácido nucleico b) presenta una secuencia de aminoácidos según SEQ ID NO. 6. Secuencias de aminoácidos que codifican para péptidos señal de este tipo, particularmente preferidas, se indican en SEQ ID NO. 5.

30 Además, en lugar de los péptidos señal mencionados que permiten una secreción de la proteína, es posible usar secuencias estructuralmente homólogas a estas secuencias. Por una secuencia estructuralmente homóloga se entiende una secuencia de aminoácidos cuya sucesión de aminoácidos presenta un plegamiento espacial comparable como el de un péptido señal con la secuencia de aminoácidos según SEQ ID NO. 6. Este plegamiento espacial causa que sea reconocido por la célula anfitriona en forma de secuencia de señal secretoria y, por consiguiente, la proteína que contiene la secuencia de señal estructuralmente homóloga se retira de la célula anfitriona. De preferencia se efectúa una interacción con el sistema de translocación usado por la célula anfitriona. La secuencia de aminoácidos estructuralmente homóloga se enlaza, por lo tanto, de preferencia directamente o indirectamente con al menos un componente del sistema de translocación de la célula anfitriona. Por enlace directo se entiende una interacción directa, por enlace indirecto se entiende que la interacción puede efectuarse por medio de uno o de varios otros componentes, principalmente proteínas u otras moléculas que fungan como adaptador y, por consiguiente, tienen una función de puente entre la secuencia de aminoácidos estructuralmente homóloga y un componente del sistema de translocación de la célula anfitriona.

45 La determinación de la entidad de secuencias de ácido nucleico o de aminoácidos se efectúa por medio de una comparación de secuencias. Una comparación así se efectúa asignando sucesiones similares en las secuencias de nucleótidos o secuencias de aminoácidos, unas a otras. Esta comparación de secuencias se efectúa de preferencia con base en el algoritmo BLAST establecido en el estado de la técnica y usado habitualmente (cf., por ejemplo, Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool" [Herramienta de búsqueda de alineación local básica]. *J. Mol. Biol.* 215:403-410, y Altschul, Stephan F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Hheng Zhang, Webb Miller, y David J. Lipman (1997): "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs" [BLAST y PSI-BLAST calibrados: una nueva generación de programas de búsqueda en bases de datos de proteínas]; *Nucleic Acids Res.*, 25, páginas 3389-3402) y ocurre principalmente al asignar sucesiones similares de nucleótidos o aminoácidos en las secuencias de ácido nucleico o aminoácidos unas a otras. Una asignación tabular de las posiciones en cuestión se denomina alineación. Otro algoritmo disponible en el estado de la técnica es el algoritmo FASTA. Comparaciones de secuencias (alineaciones), principalmente comparaciones múltiples de secuencias, usualmente se crean usando programas de ordenador. Frecuentemente se usan, por ejemplo, las series Clustal (compárese, por ejemplo, Chenna et al (2003): Multiple sequence alignment with the Clustal series of programs. *Nucleic Acid Research* 31, 3497-3500), T-Coffee (compárese por ejemplo Notredame et al. (2000): T-Coffee: A novel method for multiple sequence alignments. *J. Mol. Biol.* 302, 205-217) o programas que se basan en estos programas o algoritmos. Para propósitos de la presente invención, las comparaciones y alineaciones de secuencias se crean preferiblemente usando el programa de

ordenador Vector NTI® Suite de 10.3 (Invitrogen Corporation, 1600 Faraday Avenue, Carlsbad, California, EUA) usando los parámetros estándares (por defecto) predefinidos.

Tal comparación permite hacer una declaración sobre la similitud de las secuencias comparadas entre sí. Usualmente se reporta en porcentaje de identidad; es decir, la fracción de los nucleótidos o residuos de aminoácidos idénticos en las mismas posiciones o posiciones correspondientes unas a otras en una alineación. El término tomado del modo más amplio de homología se refiere a sustituciones de aminoácidos conservadas en el caso de secuencias de aminoácidos; es decir, aminoácidos que tienen propiedades similares porque usualmente ejercen actividades o funciones similares dentro de la proteína. Por lo tanto, la similitud de las secuencias comparadas también puede indicarse como porcentaje de homología o porcentaje de similitud. Los valores de identidad y/u homología pueden reportarse a través de polipéptidos o genes enteros o sólo a través de regiones particulares. Las regiones homólogas o idénticas de diferentes secuencias de ácido nucleico o aminoácidos se definen, por lo tanto, por concordancias en las secuencias. A menudo tienen las mismas o similares funciones. Pueden ser pequeñas y comprender sólo algunos nucleótidos o aminoácidos. Dichas regiones pequeñas a menudo ejercen funciones esenciales para la actividad entera de la proteína. Por lo tanto, puede ser práctico referir las concordancias de secuencia sólo a regiones particulares, particularmente pequeñas. A menos que se indique lo contrario, los valores de identidad u homología en la presente solicitud se refieren, sin embargo, a la longitud total de las secuencias de ácido nucleico o aminoácidos indicadas respectivamente.

La proteína codificada por la secuencia de ácido nucleico b) comprende además otra secuencia de aminoácidos. Esta secuencia de aminoácidos es consecuentemente la propia secuencia de aminoácidos real de la proteína sin péptido señal. Preferiblemente, la secuencia de aminoácidos es una proteína madura. Por una proteína madura se entiende su forma tratada completamente, ya que es posible que una forma inmadura sea codificada por parte de un gen correspondiente y después de la traducción sea tratada para dar lugar a una forma madura. A manera de ejemplo, las formas inmaduras de la proteína pueden comprender péptidos y/o propéptidos señal o elongaciones en la posición N terminal y/o C terminal, los cuales ya no están presentes en la forma madura. Por ejemplo, las formas inmaduras de proteasas, principalmente subtilasas y entre éstas ante todo subtilisinas, comprenden un péptido señal y un propéptido que ya no están presentes en la forma madura de la proteasa. Como alternativa, la otra secuencia de aminoácidos es la secuencia de aminoácidos de una proteína inmadura que comprende un propéptido. Una configuración de este tipo también se toma en consideración principalmente para proteasas, principalmente subtilasas y entre éstas ante todo subtilisinas. En configuraciones particularmente preferidas, la otra secuencia de aminoácidos no comprende otro péptido señal. En el caso de configuraciones de este tipo según la invención solamente el péptido señal según la invención provoca, por consiguiente, la secreción de la proteína desde una célula anfitriona.

La otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa. Se divulgan además secuencias de aminoácidos de la proteína en las cuales la secuencia de aminoácidos comprende una amilasa, celulasa, hemicelulasa, mananasa, tanasa, xilanasas, xantanasas, xiloglucanasas, β -glucosidasas, una enzima que disocia pectina, carragenasa, perhidrolasa, oxidasa, oxidoreductasa o una lipasa. De modo muy particularmente preferido, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una subtilisina.

A manera de ejemplo, una de las enzimas mencionadas a continuación puede prepararse ventajosamente con un vector de expresión según la invención.

Entre las proteasas se prefieren subtilisinas. Ejemplos de estas son las subtilisinas BPN' y Carlsberg, las proteasas PB92, las subtilisinas 147 y 309, las proteasas alcalinas de *Bacillus lentus*, subtilisinas DY y las enzimas asignadas a las subtilasas, aunque ya no a las subtilisinas en el sentido más estrecho: termitasa, proteinasa K y las proteasas TW3 y TW7. La subtilisina Carlsberg se encuentra disponible en forma evolucionada bajo el nombre comercial Alcalase® de la compañía Novozymes A/S, Bagsvaerd, Dinamarca. Las subtilisinas 147 y 309 se venden bajo los nombres comerciales Esperase®, o bien Savinase® de la compañía Novozymes. De la proteasa de *Bacillus lentus* DSM 5483 se derivan las variantes de proteasa enunciadas más adelante bajo la denominación BLAP®. Otras proteasas preferidas son, además, por ejemplo, las enzimas listadas bajo la denominación PUR. Otras proteasas son además las enzimas disponibles bajo los nombres comerciales Durazym®, Relase®, Everlase®, Nafizym®, Natalase®, Kannase® y Ovozyme® de la compañía Novozymes, las enzimas disponibles bajo los nombres comerciales Purafect®, Purafect® OxP, Purafect® Prime, Excellase® y Properase® de la compañía Genencor, la enzima disponible bajo el nombre comercial Protosol® de la compañía Advanced Biochemicals Ltd., Thane, India, la enzima disponible bajo el nombre comercial Wuxi® de la compañía Wuxi Snyder Bioproducts Ltd., China, las enzimas disponibles bajo los nombres comerciales Proleather® y Protease P® de la compañía Amano Pharmaceuticals Ltd., Nagoya, Japón, y la enzima disponible bajo el nombre comercial Proteinase K-16 de la compañía Kao Corp., Tokio, Japón. Además, también se prefieren las proteasas de *Bacillus gibsonii* y *Bacillus pumilus*, las cuales se divulgan en las solicitudes internacionales de patente WO2008/086916 y WO2007/131656.

Ejemplos de amilasas son las α -amilasas de *Bacillus licheniformis*, de *Bacillus amiloliquefaciens* o de *Bacillus stearothermophilus* así como principalmente incluso de sus desarrollos adicionales mejorados para el empleo en detergentes o productos de limpieza. La enzima de *Bacillus licheniformis* se encuentra disponible en la empresa Novozymes bajo el nombre Termamyl® y en la empresa Danisco/Genencor bajo el nombre de Purastar®ST. Productos perfeccionados de estas α -amilasas se encuentran disponibles en las empresas Novozymes bajo los nombres

comerciales Duramyl® y Termamyl®ultra, von dem Unternehmen Danisco/Genencor bajo el nombre Purastar®OxAm y en la empresa Daiwa Seiko Inc., Tokio, Japón como Keistase®. La α -amilasa de *Bacillus amiloliquefaciens* se vende por parte de la empresa Novozymes bajo el nombre BAN®, y las variantes derivadas de la α -amilasa de *Bacillus stearothermophilus* se venden igualmente bajo los nombres BSG® y Novamyl®, por la empresa Novozymes. Además, pueden mencionarse la α -amilasa de *Bacillus sp.* A 7-7 (DSM 12368) y la ciclodextrina-glucanotransferasa (CGTasa) de *Bacillus agaradherens* (DSM 9948). Igualmente pueden emplearse productos de fusión de todas las moléculas mencionadas. Además, también son adecuados los desarrollos de la α -amilasa de *Aspergillus niger* y *A. oryzae* disponibles bajo los nombres comerciales Fungamyl® de la empresa Novozymes. Otros productos comerciales ventajosos son, por ejemplo, las amilasas Powerase® de la empresa Danisco/Genencor y las amilasas Amylase-LT®, Stainzyme® y Stainzyme plus®, estas últimas de la empresa Novozymes. También pueden prepararse variantes de estas enzimas disponibles por medio de mutaciones puntuales. Otras amilasas preferidas se divulgan en las publicaciones internacionales WO 00/60060, WO 03/002711, WO 03/054177 y WO07/079938, a cuya divulgación se hace referencia expresamente y su contenido divulgado a este respecto se incluyen expresamente a la presente solicitud de patente. Las amilasas son, además, de preferencia α -amilasas.

Ejemplos de lipasas o cutinasas son las lipasas que pueden obtenerse originalmente de *Humicola lanuginosa* (*Thermomyces lanuginosus*), o bien lipasas perfeccionadas, principalmente aquellas con el intercambio de aminoácidos D96L. Estas se venden, por ejemplo, por parte de la compañía Novozymes bajo los nombres comerciales Lipolase®, Lipolase®Ultra, LipoPrime®, Lipozyme® y Lipex®. Además, pueden prepararse, por ejemplo, las cutinasas que han sido aisladas originalmente de *Fusarium solani pisi* y *Humicola insolens*. Por parte de la compañía Danisco/Genencor pueden prepararse, por ejemplo, las lipasas o cutinasas cuyas enzimas de partida hayan sido aisladas originalmente de *Pseudomonas mendocina* y *Fusarium solanii*. Como otros productos comerciales importantes pueden mencionarse las preparaciones vendidas originalmente por la compañía Gist-Brocades (ahora Danisco/Genencor) M1 Lipase® y Lipomax® y las enzimas vendidas por la compañía Meito Sangyo KK, Japón, bajo los nombres Lipase MY-30®, Lipase OF® y Lipase PL®; además, el producto Lumafast® de la compañía Danisco/Genencor.

Ejemplos de celulasas (endoglucanasas, EG) comprenden secuencias de la preparación de celulasa, rica en endoglucanasa (EG), fúngica o sus desarrollos que son ofrecidos por la compañía Novozymes bajo el nombre comercial Celluzyme®. Los productos disponibles en la compañía Novozymes Endolase® y Carezyme® se basan en la 50 kD-EG, o bien en la 43 kD-EG de *Humicola insolens* DSM 1800. Los productos comerciales que pueden prepararse en esta compañía son Cellusoft®, Renozyme® y Celluclean®. Además, son capaces de prepararse, por ejemplo, celulasas que se encuentran disponibles en la compañía AB Enzymes, Finlandia, bajo los nombres comerciales Ecostone® y Biotouch®, y que se basan al menos parcialmente en la 20 kD-EG de *Melanocarpus*. Otras celulasas de la compañía AB Enzymes son Econase® y Ecopulp®. Otras celulasas adecuadas son de *Bacillus sp.* CBS 670.93 y CBS 669.93, en donde las de *Bacillus sp.* CBS 670.93 se encuentran disponibles en la empresa Danisco/Genencor bajo el nombre comercial Puradax®. Otros productos comerciales capaces de prepararse en la compañía Danisco/Genencor son "Genencor detergent cellulase L" e IndiAge®Neutra.

También pueden prepararse variantes de estas enzimas que pueden obtenerse mediante mutaciones puntuales. Celulasas particularmente preferidas son variantes de celulasa de *Thielavia terrestris* que se divulgan en la publicación internacional WO 98/12307, celulasas de *Melanocarpus*, principalmente *Melanocarpus albomyces*, que se divulgan en la publicación internacional WO 97/14804, celulasas del tipo EGIII de *Trichoma reesei*, que se divulgan en la solicitud europea de patente EP 1 305 432 o variantes que pueden obtenerse a partir de estas, principalmente aquellas que se divulgan en la solicitudes europeas de patente EP 1240525 y EP 1305432, así como celulasas que se divulgan en las solicitudes internacionales de patente WO 1992006165, WO 96/29397 y WO 02/099091. Por lo tanto, se hace referencia expresamente a su divulgación respectiva y, por lo tanto, se incluye expresamente su contenido divulgado a este respecto en la presente solicitud de patente.

Además, pueden prepararse otras enzimas que se agrupan bajo el término hemicelulasas. A estas pertenecen, por ejemplo, mananasas, xantanlianas, xantaninas, xiloglucanasas, xilanasas, pululaninas, enzimas que disocian pectina y β -glucanasas. La β -glucanasa obtenida de *Bacillus subtilis* se encuentra disponible bajo el nombre Cereflo® de la compañía Novozymes. Hemicelulasas particularmente preferidas son mananasas, las cuales se venden, por ejemplo, bajo los nombres comerciales Mannaway® de la empresa Novozymes o Purabrite® de la empresa Genencor. Entre las enzimas que disocian pectina también se cuentan enzimas con las denominaciones de pectinasa, pectato liasa, pectinesterasa, pectina desmetoxilasa, pectina metoxilasa, pectina metilesterasa, pectasa, pectina metilesterasa, pectinoesterasa, pectina pectilhidrolasa, pectina despolimerasa, endopoligalacturonasa, pectolasa, pectina hidrolasa, pectina-poligalacturonasa, endopoligalacturonasa, poli- α -1,4-galacturonida glicanohidrolasa, endogalacturonasa, endo-D-galacturonasa, galacturano 1,4- α -galacturonidasa, exopoligalacturonasa, poli(galacturonato) hidrolasa, exo-D-galacturonasa, exo-D-galacturonanasa, exopoli-D-galacturonasa, exopoli- α -galacturonosidasa, exopoligalacturonosidasa o exopoligalacturanosidasa. Ejemplos de enzimas adecuadas a este respecto se encuentran disponibles, por ejemplo, bajo los nombres Gamanase®, Pektinex AR®, X-Pekt® o Pektaway® de la empresa Novozymes, bajo el nombre Rohapect UF®, Rohapect TPL®, Rohapect PTE100®, Rohapect MPE®, Rohapect MA plus HC, Rohapect DA12L®, Rohapect 10L®, Rohapect B1L® de la empresa AB Enzymes y bajo el nombre Pyrolase® de la empresa Diversa Corp., San Diego, CA, EUA.

Además, también pueden prepararse oxidoreductasas, por ejemplo, oxidasas, oxigenasas, catalasas, peroxidasas, tales como halo-, cloro-, bromo-, lignina-, glucosa- o manganeso-peroxidasas, dioxigenasen o lacasas (fenoloxidasas, polifenoloxidasas). Como productos comerciales adecuados pueden nombrarse Denilite® 1 y 2 de la compañía Novozymes. Otras enzimas se divulgan en las solicitudes internacionales de patente WO 98/45398, WO 2005/056782, WO 2004/058961 así como WO 2005/124012.

En otra forma de realización de la invención no se encuentra presente la otra secuencia de aminoácidos naturalmente junto con el péptido señal en una cadena de polipéptido en un microorganismo. Por consiguiente, la proteína codificada por la secuencia de ácido nucleico b) es una proteína recombinante. No presente naturalmente significa, por lo tanto, que las dos secuencias de aminoácido no son componentes de una proteína propia del microorganismo. Por consiguiente, una proteína que comprende el péptido señal y la otra secuencia de aminoácidos no puede expresarse en el microorganismo por una secuencia de ácido nucleico que es parte del ADN cromosómico del microorganismo en su forma de tipo silvestre. Una proteína de este tipo y/o la secuencia de ácido nucleico que codifica respectivamente para esta, por consiguiente, no se encuentra presente en la forma de tipo silvestre del microorganismo y/o no puede aislarse de la forma de tipo silvestre del microorganismo. Ambas secuencias, el péptido señal y la otra secuencia de aminoácidos, tienen que estar asignadas, por lo tanto, a dos diferentes cadenas de polipéptido en una forma de tipo silvestre de un microorganismo, en tanto que ambas estén presentes o puedan estar presentes en general en la forma de tipo silvestre de un microorganismo. En el contexto de esta configuración de la invención, por lo tanto, el péptido señal y la otra secuencia de aminoácidos, o bien los ácidos nucleicos que codifican para la misma, se han combinado recientemente con ayuda de procedimientos de ingeniería genética y esta combinación de péptido señal y de la otra secuencia de aminoácidos no existe en la naturaleza. Por consiguiente, en la forma de tipo silvestre de un microorganismo no se encuentra presente una conexión de este tipo entre el péptido señal con la otra secuencia de aminoácidos, y más específicamente ni al nivel de ADN ni al nivel de proteína. Sin embargo, el péptido señal, como también la otra secuencia de aminoácidos o la secuencia de ácido nucleico que codifican respectivamente para la misma, pueden ser respectivamente de origen natural, aunque su combinación no existe en la naturaleza. No obstante, el péptido señal y la otra secuencia de aminoácidos pueden provenir, ellos mismos, del mismo microorganismo o también de diferentes microorganismos.

En una forma preferida de realización, un ácido nucleico según la invención se caracteriza porque es un ácido nucleico no natural. No natural significa que un ácido nucleico según la invención no puede aislarse en su forma de tipo silvestre de un organismo que existe en la naturaleza. Por lo tanto, principalmente y respecto a las bacterias de tipo silvestre, un ácido nucleico según la invención no es un ácido nucleico propio de una bacteria. Las secuencias (a) y (b) no provienen del mismo o de los mismos organismos, principalmente bacterias, sino que provienen de organismos diferentes, principalmente bacterias. Diferentes bacterias son, por ejemplo, bacterias que pertenecen a diferentes cepas o tipos o géneros.

En otra forma de realización de la invención, el vector de expresión se caracteriza porque el péptido señal se encuentra dispuesto de modo terminal N de la otra secuencia de aminoácidos en la proteína codificada por la secuencia de ácido nucleico b). La proteína codificada por la secuencia de ácido nucleico b) tiene, por lo tanto, la siguiente estructura: terminal N - péptido señal - (secuencia de aminoácidos adicional opcional) - otra secuencia de aminoácidos - terminal C. Una estructura de este tipo de la proteína que va a expresarse ha resultado ser particularmente ventajosa.

En otra forma de realización de la invención, el vector de expresión se caracteriza porque la proteína codificada por la secuencia de ácido nucleico b) comprende además una secuencia de conexión, la cual se encuentra dispuesta entre el péptido señal y la otra secuencia de aminoácidos de la proteína. La proteína codificada por la secuencia de ácido nucleico B) tiene, por lo tanto, la siguiente estructura: terminal N - péptido señal - secuencia de conexión (también "acoplamiento" o "espaciador") - otra secuencia de aminoácidos - terminal C. Una estructura de este tipo de la proteína que va a expresarse también ha resultado ser particularmente ventajosa. De preferencia, la secuencia de conexión presenta una longitud entre 1 y 50 aminoácidos, entre 2 y 25 aminoácidos, entre 2 y 15 aminoácidos, entre 3 y 10 aminoácidos y de modo particularmente preferido entre 3 y 5 aminoácidos. Un ejemplo de una secuencia de conexión particularmente preferida es la sucesión de aminoácidos alanina, ácido glutámico y fenilalanina (desde terminal N terminal C).

En otra forma de realización de la invención, el vector de expresión se caracteriza porque la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa, en donde la secuencia de aminoácidos de la proteasa es idéntica en al menos 80 % a SEQ ID NO. 7. La secuencia de aminoácidos de la proteasa es preferiblemente idéntica en al menos 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % y de modo muy particularmente preferido en 100 % es idéntica a SEQ ID NO. 7.

Como alternativa, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a SEQ ID NO. 8. La secuencia de aminoácidos de la proteasa es idéntica preferiblemente en al menos 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % y de modo muy particularmente preferido en 100 % a SEQ ID NO. 8.

Como alternativa, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a SEQ ID NO. 9. La secuencia de aminoácidos de la proteasa es idéntica

preferiblemente en al menos 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % y de modo muy particularmente preferido en 100 % idéntica a SEQ ID NO. 9.

5 Como alternativa, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a SEQ ID NO. 10 y en la posición 99 en el listado según SEQ ID NO. 10 comprende el aminoácido ácido glutámico (E) o ácido aspártico (D). La secuencia de aminoácidos de la proteasa preferiblemente es idéntica en al menos 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % y de modo muy particularmente preferido idéntica a SEQ ID NO. 10 en las posiciones 1 a 98 y 100 a 269 en el listado según SEQ ID NO. 10.

10 Como alternativa, la otra secuencia de aminoácidos de la proteína comprende una secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a SEQ ID NO. 10 y en la posición 99 en el listado según SEQ ID NO. 10 presenta el aminoácido ácido glutámico (E) o ácido aspártico (D) y, además, en el listado según SEQ ID NO. 10 presenta al menos uno de los siguientes aminoácidos:

- (a) treonina en la posición 3 (3T),
- (b) isoleucina en la posición 4 (4I),
- 15 (c) alanina, treonina o arginina en la posición 61 (61A, 61T o 61R),
- (d) ácido aspártico o ácido glutámico en la posición 154 (154D o 154E),
- (e) prolina en la posición 188 (188P),
- (f) metionina en la posición 193 (193M),
- (g) isoleucina en la posición 199 (199I),
- 20 (h) ácido aspártico, ácido glutámico o glicina en la posición 211 (211D, 211E o 211G),
- (i) combinaciones de los aminoácidos (a) a (h).

25 La secuencia de aminoácidos de esta proteasa es preferiblemente idéntica en al menos 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % y de modo muy particularmente preferido idéntica a SEQ ID NO. 10 en todas las posiciones no modificadas o no previstas para una modificación. De modo muy particularmente preferido, la otra secuencia de aminoácidos de la proteína comprende, por lo tanto, la secuencia de aminoácidos de una proteasa que presenta una secuencia de aminoácidos modificada frente a SEQ ID NO. 10 en al menos dos posiciones, en donde la primera modificación en el listado según SEQ ID NO. 10 ácido glutámico está en la posición 99 y la segunda modificación en el listado según SEQ ID NO. 10 se selecciona del grupo compuesto por:

- 30 (a) treonina en la posición 3 (3T),
- (b) isoleucina en la posición 4 (4I),
- (c) alanina, treonina o arginina en la posición 61 (61A, 61T o 61R),
- (d) ácido aspártico o ácido glutámico en la posición 154 (154D o 154E),
- (e) prolina en la posición 188 (188P),
- 35 (f) metionina en la posición 193 (193M),
- (g) isoleucina en la posición 199 (199I),
- (h) ácido aspártico, ácido glutámico o glicina en la posición 211 (211D, 211E o 211G),
- (i) combinaciones de los aminoácidos (a) a (h).

40 Igualmente, de manera muy particularmente preferida, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa que presenta una secuencia de aminoácidos modificada en al menos dos posiciones frente a SEQ ID NO. 10, en donde la primera modificación en el listado según SEQ ID NO. 10 ácido aspártico está en la posición 99 y la siguiente modificación en el listado según SEQ ID NO. 10 se selecciona del grupo compuesto por:

- (a) treonina en la posición 3 (3T),
- 45 (b) isoleucina en la posición 4 (4I),

- (c) alanina, treonina o arginina en la posición 61 (61A, 61T o 61R),
- (d) ácido aspártico o ácido glutámico en la posición 154 (154D o 154E),
- (e) prolina en la posición 188 (188P),
- (f) metionina en la posición 193 (193M),
- 5 (g) isoleucina en la posición 199 (199I),
- (h) ácido aspártico, ácido glutámico o glicina en la posición 211 (211D, 211E o 211G),
- (i) combinaciones de los aminoácidos (a) a (h).

Se ha establecido que incluso las proteasas antes mencionadas pueden prepararse de modo particularmente ventajoso con vectores de expresión según la invención. Para este tipo de formas de realización de la invención se ha encontrado que con tales combinaciones de péptidos señal y subtilisinas pueden lograrse rendimientos de producto particularmente buenos en una fermentación. A este respecto se indican en la secuencia de aminoácidos de las proteasas maduras, es decir los productos tratados terminados. En un vector de expresión según la invención, a este respecto también pueden estar comprendidas otras secuencias de la proteasa inmadura, principalmente, a manera de ejemplo, propéptidos. En tal caso, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de la proteasa y del propéptido. Otra configuración de la invención se caracteriza, por consiguiente, porque la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa, principalmente una proteasa como se ha descrito antes, junto con un propéptido o su propéptido.

En general es válido que en la otra secuencia de aminoácidos de la proteína no tiene que comprender solamente la secuencia de aminoácidos de una proteína madura; por el contrario, pueden estar comprendidas otras secuencias de aminoácidos como, por ejemplo, propéptidos de esta secuencia de aminoácidos. Esto es válido no solamente para proteasas, sino para todas las proteínas, principalmente incluso todos los tipos de enzimas.

Ácidos nucleicos y vectores de expresión según la invención pueden generarse mediante procedimientos conocidos por se para modificar ácidos nucleicos. Tales se representan, por ejemplo, en manuales correspondientes como el de Fritsch, Sambrook y Maniatis, "Molecular cloning: a laboratory manual", Cold Spring Harbour Laboratory Press, New York, 1989, que es conocido por el especialista en el campo de la biotecnología. Ejemplo de tales procedimientos son la síntesis química o la reacción en cadena de polimerasa (PCR), opcionalmente en conexión con otros procedimientos estándar de la biología molecular y/o química o bioquímica.

Con todos los objetos de la invención y formas de realización mencionados, como otros objetos de la invención, se asocian células anfitrionas no humanas que incluyen vectores según la invención, procedimientos de preparación en los cuales se emplean células anfitrionas correspondientes, así como modificaciones de los vectores o células anfitrionas correspondientes. Por lo tanto, las anteriores descripciones se refieren a estos objetos de la invención de manera correspondiente.

Otro objeto de la invención es una célula anfitriona no humana que contiene un vector de expresión según la invención. Un vector de expresión según la invención se introduce preferiblemente en la célula anfitriona mediante su transformación. De acuerdo con la invención, esto se efectúa transformando un vector según la invención en un microorganismo que después representa una célula anfitriona según la invención. Como alternativa, los componentes individuales, es decir partes o fragmentos de ácido nucleico, por ejemplo, los componentes (a) y/o (b) de un vector según la invención también pueden introducirse en una célula anfitriona de modo que la célula anfitriona resultante luego contenga un vector según la invención. Este procedimiento es adecuado particularmente si la célula anfitriona contiene ya no o varios componentes de un vector según la invención y los otros componentes se complementan luego de manera correspondiente. En el estado de la técnica están establecidos procedimientos para la transformación de células y son suficientemente conocidos por el especialista. Como células anfitrionas son adecuadas principalmente todas las células, es decir células procariotas o eucariotas. Se prefieren aquellas células anfitrionas que pueden manejarse genéticamente de modo ventajoso, por ejemplo, con respecto a la transformación con el vector y su establecimiento estable, por ejemplo, hongos o bacterias unicelulares. Además, las células anfitrionas preferidas se caracterizan por una buena capacidad de manejo microbiológico y biotecnológico. Esto se refiere, por ejemplo, a una fácil capacidad de cultivo, altas tasas de crecimiento, bajos requisitos para los medios de fermentación y buenas tasas de producción y secreción para proteínas extrañas. Por la abundancia de diferentes sistemas disponibles en el estado de la técnica, frecuentemente los sistemas de expresión óptimos tienen que determinarse experimentalmente para el caso individual.

Otras formas preferidas de realización representan aquellas células anfitrionas que debido a los elementos de regulación genéticos que se hacen disponibles, por ejemplo, en el vector, pero pueden estar presentes también en estas células desde el principio, son regulables en cuanto a su actividad. A manera de ejemplo, mediante la adición controlada de compuestos químicos que sirven como activadores, modificando las condiciones de cultivo o al lograr una densidad celular determinada, éstas pueden estimularse para la expresión. Esto hace posible una producción económica de las proteínas.

Las células anfitrionas preferidas son procariotas o bacterianas. Las bacterias se caracterizan por tiempos breves de generación y bajas demandas para las condiciones de cultivo. De esta manera pueden establecerse procedimientos rentables. Además, el especialista en el caso de bacterias en la tecnología de fermentación dispone de una abundante experiencia. Para una producción especial, por las más diversas razones, como la fuente de nutrientes, tasas de formación de producto, demanda de tiempo, etc., que pueden determinarse experimentalmente en un caso individual, pueden ser adecuadas bacterias gram-negativas o gram-positivas

En el caso de bacterias gram-negativas como, por ejemplo, *Escherichia coli* se secreta una gran cantidad de polipéptidos en el espacio periplasmático, es decir en el compartimiento entre las dos membranas que encierran las células. Esto puede ser ventajoso para aplicaciones especiales. Además, también pueden configurarse bacterias gram-negativas de modo que expulsen los polipéptidos expresados no sólo en el espacio periplasmático, sino en el medio que rodea a la bacteria. Por el contrario, las bacterias gram-positivas, por ejemplo, bacilos o actinomicetos u otros representantes de los actinomicetales no poseen membrana exterior, de modo que las proteínas secretadas se liberan inmediatamente en el medio que circunda las bacterias, por lo regular el medio nutriente, del cual pueden purificarse los polipéptidos expresados. Estos pueden aislarse directamente del medio o seguir tratándose. Además, las bacterias gram-positivas están relacionadas con la mayoría de organismos de origen, o son idénticas a estos, y forman generalmente incluso enzimas comparables de modo que disponen de un uso de codón similar y su aparato de síntesis de proteína está orientado naturalmente de manera correspondiente.

Por uso de codón se entiende la traducción del código genético en aminoácidos; es decir, cuál orden de nucleótidos (triplete o triplete de base) codifica para cuáles aminoácidos o para cuál función, por ejemplo, inicio y final de la región que va a traducirse, sitios de enlace para diferentes proteínas, etc. De esta manera cualquier organismo, principalmente cualquier cepa de producción, posee un uso de codón determinado. Puede llegarse a pasajes estrechos en la biosíntesis de proteína si los codones que se encuentran en el ácido nucleico transgénico en la célula anfitriona se enfrentan a una cantidad comparativamente baja de ARNts cargados. Los codones sinónimos codifican, por el contrario, para los mismos aminoácidos y pueden traducirse mejor dependiendo del anfitrión. Esta transcripción opcionalmente necesaria depende, por lo tanto, de la elección del sistema de expresión. Principalmente en caso de muestras de organismos desconocidos, eventualmente no cultivables, puede ser necesaria una adaptación correspondiente.

La presente invención es, en principio, aplicable a todos los microorganismos, principalmente a todos los microorganismos ser rentables, de modo particularmente preferido a aquellos del género *Bacillus*, y conduce a que, empleando tales microorganismos en calidad de organismos de producción, puede realizarse un incremento elevado del producto en una fermentación. Tales microorganismos representan células anfitrionas preferidas en el sentido de la invención.

En otra forma de realización de la invención, la célula anfitriona se caracteriza por lo tanto porque es una bacteria, preferiblemente una que se selecciona del grupo de los géneros de *Escherichia*, *Klebsiella*, *Bacillus*, *Staphylococcus*, *Corynebacterium*, *Arthrobacter*, *Streptomyces*, *Stenotrophomonas* y *Pseudomonas*, más preferiblemente una que se selecciona del grupo de *Escherichia coli*, *Klebsiella planticola*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus amiloliquefaciens*, *Bacillus subtilis*, *Bacillus alcalophilus*, *Bacillus globigii*, *Bacillus gibsonii*, *Bacillus clausii*, *Bacillus halodurans*, *Bacillus pumilus*, *Staphylococcus carnosus*, *Corynebacterium glutamicum*, *Arthrobacter oxidans*, *Streptomyces lividans*, *Streptomyces coelicolor* y *Stenotrophomonas maltophilia*. Muy particularmente se prefiere *Bacillus licheniformis*.

La célula anfitriona también puede ser, no obstante, una célula eucariota que se caracteriza porque posee un núcleo celular. Otro objeto de la invención representa, por lo tanto, una célula anfitriona que se caracteriza porque posee un núcleo celular.

En contraste con las células procariotas, las células eucariotas son capaces de modificar la proteína formada de modo post-traducción. Ejemplos de estas son hongos tales como actinomicetos o levaduras tales como *Saccharomyces* o *Kluyveromyces*. Esto puede ser particularmente ventajoso, por ejemplo, si las proteínas han de experimentar modificaciones específicas en relación con su síntesis, las cuales hacen posibles sistemas de este tipo. A las modificaciones que realizan los sistemas eucariotas particularmente en relación con la síntesis de proteínas pertenecen, por ejemplo, el enlazamiento de compuestos de bajo peso molecular como anclas de membrana u oligosacáridos. Este tipo de modificaciones de oligosacáridos pueden ser deseables, por ejemplo, para reducir la alergenicidad de una proteína expresada. Una coexpresión con las enzimas formadas naturalmente por este tipo de células tales como, por ejemplo, celulasas también puede ser ventajosa. Además, a manera de ejemplo, pueden ser particularmente adecuados los sistemas de expresión fúngicos termofílicos para la expresión de variantes resistentes a la temperatura.

Como productos que se forman durante la fermentación, en el contexto de la invención se consideran proteínas que se codifican por parte de la secuencia de ácido nucleico (b), principalmente aquellas tal como se han descrito anteriormente. En este caso se trata de proteasas y de manera muy particularmente preferida de subtilisinas.

Las células anfitrionas pueden modificarse además con respecto de sus requisitos para las condiciones de cultivo, pueden presentar otros marcadores de selección o adicionales. Tales células anfitrionas pueden ser principalmente

aquellas que expresan varias proteínas o enzimas. Preferiblemente secretan estas en el medio que rodea las células anfitrionas.

5 Las células anfitrionas según la invención se cultivan y se fermentan de manera conocida, por ejemplo, en sistemas discontinuos o continuos. En el primer caso se inocula un medio nutriente adecuado con las células anfitrionas y el producto se cosecha del medio después de un lapso de tiempo que ha de determinarse experimentalmente. Las fermentaciones continuas se caracterizan por lograr un equilibrio de flujo en el cual, durante un lapso de tiempo comparativamente largo, se mueren parcialmente las células, pero también se reproducen y, simultáneamente, puede retirarse producto del medio.

10 Preferiblemente se usan células anfitrionas según la invención para preparar proteínas que se codifican por parte de la secuencia de ácido nucleico (b). Otro objeto de la invención es, por lo tanto, un procedimiento para la preparación de una proteasa que comprende

a) cultivar una célula anfitriona según la invención

b) aislar la proteasa del medio de cultivo o de la célula anfitriona.

15 Este objeto de la invención comprende preferiblemente procedimientos de fermentación. Los procedimientos de fermentación son conocidos del Estado de la técnica y representan la etapa de producción propiamente dicha a escala industrial, seguida por lo regular por un procedimiento de purificación adecuado del producto preparado, por ejemplo, de la proteína. Todos los procedimientos de fermentación que se basan en un procedimiento correspondiente para la preparación de una proteasa representan formas de realización de este objeto de la invención.

20 En este caso las condiciones respectivamente óptimas para los procedimientos de preparación, principalmente las condiciones óptimas de cultivo para las células anfitrionas usadas, tienen que determinarse experimentalmente según el conocimiento del especialista, por ejemplo, con respecto al volumen de fermentación y/o composición del medio y/o suministro de oxígeno y/o velocidad de agitación. Principalmente se toman en consideración los procedimientos de fermentación que se caracterizan porque la fermentación se realiza por medio de una estrategia de alimentación continua. En este caso, se alimentan continuamente los componentes del medio que se consumen por el cultivo en curso; también se habla de una estrategia de alimentación continua. De esta manera pueden lograrse incrementos considerables tanto en la densidad celular, como también en la masa celular masa seca y/o ante todo la actividad de la proteína de interés, de preferencia la enzima.

Además, la fermentación también puede configurarse de modo que se retiren los metabolitos no deseados mediante filtración o se neutralicen adicionando amortiguadores (reguladores de pH) o contraiones respectivamente apropiados.

30 La proteína preparada puede cosecharse del medio de fermentación. Un procedimiento de fermentación de este tipo es ventajoso frente a un aislamiento del polipéptido de la célula anfitriona, es decir un tratamiento de producto a partir de la masa celular (masa seca). Según la invención, con los péptidos señal se ponen a disposición marcadores de secreción adecuados a este respecto.

35 Todos los hechos descritos anteriormente pueden combinarse en procedimientos para preparar proteínas. Es concebible a este respecto una gran cantidad de posibilidades de combinación de etapas procedimentales. El procedimiento óptimo tiene que determinarse para cada caso individual en concreto.

Otro objeto de la invención es el uso de un vector de expresión según la invención o de una célula anfitriona según la invención para la preparación de una proteasa.

40 Todos los hechos, objetos y formas de realización que ya se han descrito anteriormente también son aplicables a estos objetos de la invención. Por lo tanto, en este sitio se hace referencia expresamente a la divulgación del sitio correspondiente, con la indicación de que esta divulgación también es válida para los usos según la invención (uso del vector y de la célula anfitriona).

Ejemplos:

45 Todas las etapas de trabajo de biología molecular siguen procedimientos estándar como se indican, por ejemplo, en el manual de Fritsch, Sambrook y Maniatis "Molecular cloning: a laboratory manual", Cold Spring Harbour Laboratory Press, Nueva York, 1989, o en obras comparables correspondientes. Las enzimas y los kits se emplearon según las indicaciones del fabricante correspondiente.

Ejemplo 1: Preparación de vectores de expresión según la invención

50 El plásmido pBSMuL3 (Brockmeier et al., 2006) fue acortado mediante digestión de restricción de SmaI y religación subsiguiente alrededor de la fracción de *E. coli*. El plásmido resultante, pBSMuL5 (cf. Figura 1) sirvió como vector para la clonación de las proteasas, que incluyen pro péptido en los sitios de restricción EcoRI y BamHI. Para esto se amplificaron los genes de la proteasa según SEQ ID NO. 8 con los cebadores según SEQ ID NO. 11 y SEQ ID NO. 12, y de la proteasa alcalina según SEQ ID NO. 9 con los cebadores según SEQ ID NO. 13 y SEQ ID NO. 14. Los plásmidos resultantes sirvieron como vectores para la clonación de los péptidos señal en los sitios de restricción HindIII

y EcoRI. El fragmento de ADN del péptido señal de control SubC (*B. licheniformis*, NCBI ("National Center for Biotechnology Information") número de acceso ("Accession Number"): X91260.1) como "benchmark" (punto de referencia) fue amplificado con ayuda de los cebadores según SEQ ID NO. 15 y SEQ ID NO. 16 y clonado respectivamente en los sitios de restricción HindIII y EcoRI de los plásmidos, de modo que surgieron plásmidos con una secuencia de ácido nucleico b) que codifican para una proteína con el péptido señal SubC en conexión con una proteasa según SEQ ID NO. 8 (plásmido 1) o SEQ ID NO. 9 (plásmido 2). Estos plásmidos sirvieron a continuación como control o "benchmark". Fueron amplificados el fragmento de ADN del péptido señal según SEQ ID NO. 2 con ayuda de los cebadores según SEQ ID NO. 19 y SEQ ID NO. 20, el fragmento de ADN del péptido señal según SEQ ID NO. 4 con los cebadores según SEQ ID NO. 17 y SEQ ID NO. 18, y el fragmento de ADN del péptido señal según SEQ ID NO. 6 con los cebadores según SEQ ID NO. 21 y SEQ ID NO. 22. Mientras se clonaban los fragmentos de ADN de los péptidos señal según SEQ ID NO. 2 y 4 respectivamente en el vector que codifica para una proteasa según SEQ ID NO. 8 (plásmidos 3 y 4), el fragmento de ADN del péptido señal según SEQ ID NO. 6 se insertó en el vector que codifica para una proteasa según SEQ ID NO. 9 (plásmido 5). A causa de la clonación, en este caso entre la secuencia de ADN del respectivo péptido señal y la secuencia de ADN del propéptido de la proteasa respectiva fue introducida una secuencia de 9 nucleótidos que codifica para la sucesión de aminoácidos AEF (cf. Figura 1). Esta llamada secuencia de conexión contiene la secuencia de reconocimiento de la endonucleasa de restricción EcoRI.

Todos los oligonucleótidos usados como cebadores son listados en la siguiente tabla 1:

Tabla 1:

Denominación	Secuencia de nucleótidos (en orientación 5' →3'; lo subrayado señala los sitios de restricción)	Sitios de restricción
SEQ ID NO. 11	ATATGAATTCGCTGAGGAAGCAAAGAAAA	EcoRI
SEQ ID NO. 12	ATATGGATCCCTAGCGTGTGCCGCTTCTGC	BamHI
SEQ ID NO. 13	ATATGAATTCGCTGAGGAAGCAAAGAAAA	EcoRI
SEQ ID NO. 14	ATATGGATCCCTAGCGCGTTGCTGCATCTGC	BamHI
SEQ ID NO. 15	ATATAAGCTTAAGGAGGATATTATGATGAGGAAAAAGAGT TTT	HindIII
SEQ ID NO. 16	ATATGAATTCAGCTGCAGAAGCGGAATCGCTGAA	EcoRI
SEQ ID NO. 17	ATATAAGCTTAAGGAGGATATTATGAAAAACTATTCAAAA CC	HindIII
SEQ ID NO. 18	ATATGAATTCAGCAGCCGCCGAGATTGTGAGAA	EcoRI
SEQ ID NO. 19	ATATAAGCTTAAGGAGGATATTATGGCGAAACCACTATCA AAA	HindIII
SEQ ID NO. 20	ATATGAATTCAGCAGCGTCTGCCGCGGGTAAACC	EcoRI
SEQ ID NO. 21	ATATAAGCTTAAGGAGGATATTATGACATTGACTAAACTG AAA	HindIII
SEQ ID NO. 22	ATATGAATTCAGCGGCAAGTGCCTGACTGGAAAA	EcoRI

Ejemplo 2: Expresión de las proteínas

Una cepa de *Bacillus licheniformis* fue transformada con los plásmidos 1 a 5 para obtener las diversas cepas de producción de proteasa. Para la inoculación de cultivos se usaron colonias individuales de placas de agar que fueron incubadas durante una noche. Para la determinación cuantitativa de la eficiencia de secreción, las colonias individuales fueron transferidas directamente de las placas de agar a Deepwell-MTP (placa de microtitulación; 96 cavidades cada una de 1 ml de medio LB selectivo). En tal caso, cada colonia individual fue transferida a por lo menos dos cavidades en paralelo para obtener determinación por duplicado o triplicado mediante el cultivo múltiple del clon respectivo. Para la inoculación de las Deepwell-MTP se usaron exclusivamente clones que fueron incubados durante una noche a 37 °C en el agitador de placas de microtitulación (Timix 5, compañía Edmy-Bühler, Hechingen), todos los clones fueron replicados en placas de agar LB y, a continuación, las células fueron sedimentadas por medio de centrifugación (4000 rpm, 20 min, 4 °C). Todas las etapas de aplicación con pipeta siguientes fueron realizadas con ayuda de pipetas de canales múltiples (Eppendorf, Hamburg), en cuyo caso fue usado el modo "reverse pipetting" (pipeteo inverso) y no fue pipeteado ningún volumen menor de 15 µl. El volumen más pequeño se cargó respectivamente en la MTP y los

volúmenes más grandes se agregaron a este y la MTP fue mezclada en cada etapa de dilución durante 10 segundos en el espectrofotómetro "Spectramax 250" (Molecular Devices, Sunnyvale, EUA). Para la realización de las diluciones correspondientes se retiró el sobrenadante de cultivo usando la pipeta de canales múltiples y se transfirió a placas de microtitulación (96 cavidades, fondo de F, transparente, compañía Greiner Bio-One, Frickenhausen).

5 A continuación, la actividad proteolíticamente en los sobrenadantes de divo o diluciones fue determinada mediante la liberación del cromóforo para-nitroanilina (pNA) del sustrato suc-L-Ala-L-Ala-L-Pro-L-Phe-p-Nitroanilid (suc-AAPF-pNA). La proteasa disocia el sustrato y libera pNA. La liberación de pNA causa un incremento en la absorbancia a 410 nm, yo paso en el tiempo es una medida de la actividad enzimática (cf. Del Mar et al., Anal.Biochem., 99: 316-320, 1979).

10 Para la determinación de la eficiencia de secreción de las diversas cepas, se cultivó un constructo de control interno (plásmido 1 o plásmido 2) en cada cultivo de MTP. La actividad proteolítica de la cepa terminada en el sobrenadante de cultivo con el constructo de control se define como 100 %.

Las cepas con los plásmidos según la invención 3 y 4, en comparación con el control que contenía el plásmido 1, lograron una actividad de proteasa incrementada en 194 % +/- 48 y 230 % +/- 38 (cf. Figura 2).

15 La cepa con el plásmido 5 según la invención, en comparación con el control que contenía el plásmido 2, logró una actividad de proteasa incrementada en 44 % +/- 10 (cf. Figura 3).

Descripción de las figuras

Figura 1: Esquema de la estrategia de clonación en el vector de expresión de *Bacillus* pBSMu5 (modificados según Brockmeier et al., 2006). (A) los fragmentos de ADN de los péptidos señal fueron amplificados en la posición N terminal con un sitio de restricción HindIII, un sitio de unión a ribosoma estandarizado (RBS), seguido por una región espaciadora y el codón de inicio estandarizado para metionina. Entre el péptido señal y la posición N terminal de la proteasa que ha de secretarse se unió un acoplamiento con una alanina en la posición "+1" y el sitio de restricción EcoRI. (B) el vector de *Bacillus* pBSMu5 con el promotor Hpall, el respectivo objetivo de secreción (clonado mediante EcoRI y BamHI), así como el casete de resistencia a la kanamicina y la proteína de replicación repB para *Bacillus*.

Figura 2: Actividad de proteasa relativa en el sobrenadante de cultivo de *Bacillus licheniformis* con la proteasa según SEQ ID NO. 8 y tres péptidos señal diferentes en pBSMu5. La actividad proteolítica de plásmido 1 fue definida como 100 % (control). Los valores fueron determinados en al menos dos cultivos independientes entre sí. Las barras de error indican la desviación estándar.

Figura 3: Actividad de proteasa relativa en el sobrenadante de cultivo de *Bacillus licheniformis* con la proteasa según SEQ ID NO. 9 y dos péptidos señal diferentes en pBSMu5. La actividad autolítica del plásmido de constructo 2 fue definida como 100 % (control). Los valores fueron determinados en al menos cultivos independientes entre sí. Las barras de error indican la desviación estándar.

Listado de secuencias

<110> BASF SE
 <120> Vectores de expresión para secreción mejorada de proteína
 <130> 0000075519EP02
 <150> DE 102011118032.3
 <151> 2011-05-31
 <160> 22
 <170> PatentIn version 3.5
 <210> 1
 <211> 126
 <212> DNA
 <213> *Bacillus subtilis*
 <220>
 <221> CDS

ES 2 763 577 T3

<222> (1)..(126)

<400> 1

```

atg gcg aaa cca cta tca aaa ggg gga att ttg gtg aaa aaa gta ttg      48
Met Ala Lys Pro Leu Ser Lys Gly Gly Ile Leu Val Lys Lys Val Leu
1           5           10           15

att gca ggt gca gta gga aca gca gtt ctt ttc gga acc ctt tca tca      96
Ile Ala Gly Ala Val Gly Thr Ala Val Leu Phe Gly Thr Leu Ser Ser
           20           25           30

ggt ata cca ggt tta ccc gcg gca gac gct      126
Gly Ile Pro Gly Leu Pro Ala Ala Asp Ala
           35           40
    
```

<210> 2

5 <211> 42

<212> PRT

<213> Bacillus subtilis

<400> 2

```

Met Ala Lys Pro Leu Ser Lys Gly Gly Ile Leu Val Lys Lys Val Leu
1           5           10           15

Ile Ala Gly Ala Val Gly Thr Ala Val Leu Phe Gly Thr Leu Ser Ser
           20           25           30

Gly Ile Pro Gly Leu Pro Ala Ala Asp Ala
           35           40
    
```

10 <210> 3

<211> 81

<212> DNA

<213> Bacillus licheniformis

<220>

15 <221> CDS

<222> (1)..(81)

<400> 3

```

atg aaa aaa cta ttc aaa acc att gtt aca ctg tca ctc ttg att tct      48
Met Lys Lys Leu Phe Lys Thr Ile Val Thr Leu Ser Leu Leu Ile Ser
1           5           10           15

gga acg ctt tta ttc tca caa tct gcg gcg gct      81
Gly Thr Leu Leu Phe Ser Gln Ser Ala Ala Ala
           20           25
    
```

<210> 4

20 <211> 27

ES 2 763 577 T3

<212> PRT

<213> Bacillus licheniformis

<400> 4

Met Lys Lys Leu Phe Lys Thr Ile Val Thr Leu Ser Leu Leu Ile Ser
1 5 10 15

Gly Thr Leu Leu Phe Ser Gln Ser Ala Ala Ala
20 25

5 <210> 5

<211> 84

<212> DNA

<213> Bacillus subtilis

<220>

10 <221> CDS

<222> (1)..(84)

<400> 5

atg aca ttg act aaa ctg aaa atg ctg agt atg tta acc gtg atg att 48
Met Thr Leu Thr Lys Leu Lys Met Leu Ser Met Leu Thr Val Met Ile
1 5 10 15

gca tct tta ttc att ttt tcc agt cag gca ctt gcc 84
Ala Ser Leu Phe Ile Phe Ser Ser Gln Ala Leu Ala
20 25

<210> 6

15 <211> 28

<212> PRT

<213> Bacillus subtilis

<400> 6

Met Thr Leu Thr Lys Leu Lys Met Leu Ser Met Leu Thr Val Met Ile
1 5 10 15

Ala Ser Leu Phe Ile Phe Ser Ser Gln Ala Leu Ala

20

25

20 <210> 7

<211> 269

<212> PRT

<213> Bacillus lentus

<400> 7

ES 2 763 577 T3

Ala Gln Thr Ile Pro Trp Gly Ile Ser Arg Val Gln Ala Pro Ala Ala
1 5 10 15

His Asn Arg Gly Leu Thr Gly Ser Gly Val Lys Val Ala Val Leu Asp
20 25 30

Thr Gly Ile Ser Thr His Pro Asp Leu Asn Ile Arg Gly Gly Ala Ser
35 40 45

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

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

Gly Val Ala Pro Ser Ala Glu Leu Tyr Ala Val Lys Val Leu Gly Ala
85 90 95

Asp Gly Arg Gly Ala Ile Ser Ser Ile Ala Gln Gly Leu Glu Trp Ala
100 105 110

Gly Asn Asn Gly Met His Val Ala Asn Leu Ser Leu Gly Ser Pro Ser
115 120 125

Pro Ser Ala Thr Leu Glu Gln Ala Val Asn Ser Ala Thr Ser Arg Gly
130 135 140

Val Leu Val Val Ala Ala Ser Gly Asn Ser Gly Ala Ser Ser Ile Ser
145 150 155 160

Tyr Pro Ala Arg Tyr Ala Asn Ala Met Ala Val Gly Ala Thr Asp Gln
165 170 175

Asn Asn Asn Arg Ala Ser Phe Ser Gln Tyr Gly Ala Gly Leu Asp Ile
180 185 190

Val Ala Pro Gly Val Asn Ile Gln Ser Thr Tyr Pro Gly Ser Thr Tyr
195 200 205

Ala Ser Leu Asn Gly Thr Ser Met Ala Thr Pro His Val Ala Gly Ala

ES 2 763 577 T3

210						215									220
Ala	Ala	Leu	Val	Lys	Gln	Lys	Asn	Pro	Ser	Trp	Ser	Asn	Val	Gln	Ile
225					230					235					240
Arg	Asn	His	Leu	Lys	Asn	Thr	Ala	Thr	Ser	Leu	Gly	Ser	Thr	Asn	Leu
				245					250					255	
Tyr	Gly	Ser	Gly	Leu	Val	Asn	Ala	Glu	Ala	Ala	Thr	Arg			
			260					265							

<210> 8

<211> 269

<212> PRT

5 <213> Bacillus lentus

<400> 8

ES 2 763 577 T3

Ala Gln Thr Ile Pro Trp Gly Ile Ser Arg Val Gln Ala Pro Ala Ala
 1 5 10 15

His Asn Arg Gly Leu Thr Gly Ser Gly Val Lys Val Ala Val Leu Asp
 20 25 30

Thr Gly Ile Ser Thr His Pro Asp Leu Asn Ile Arg Gly Gly Ala Ser
 35 40 45

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

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

Gly Val Ala Pro Ser Ala Glu Leu Tyr Ala Val Lys Val Leu Gly Ala
 85 90 95

Asp Gly Arg Gly Ala Ile Ser Ser Ile Ala Gln Gly Leu Glu Trp Ala
 100 105 110

Gly Asn Asn Gly Met His Val Ala Asn Leu Ser Leu Gly Ser Pro Ser
 115 120 125

Pro Ser Ala Thr Leu Glu Gln Ala Val Asn Ser Ala Thr Ser Arg Gly
 130 135 140

Val Leu Val Val Ala Ala Ser Gly Asn Ser Gly Ala Ser Ser Ile Ser
 145 150 155 160

Tyr Pro Ala Arg Tyr Ala Asn Ala Met Ala Val Gly Ala Thr Asp Gln

ES 2 763 577 T3

				165					170					175			
Asn	Asn	Asn	Arg	Ala	Ser	Phe	Ser	Gln	Tyr	Gly	Ala	Gly	Leu	Asp	Ile		
			180					185					190				
Met	Ala	Pro	Gly	Val	Asn	Ile	Gln	Ser	Thr	Tyr	Pro	Gly	Ser	Thr	Tyr		
		195					200					205					
Ala	Ser	Asp	Asn	Gly	Thr	Ser	Met	Ala	Thr	Pro	His	Val	Ala	Gly	Ala		
	210					215					220						
Ala	Ala	Leu	Val	Lys	Gln	Lys	Asn	Pro	Ser	Trp	Ser	Asn	Val	Gln	Ile		
225					230					235					240		
Arg	Asn	His	Leu	Lys	Asn	Thr	Ala	Thr	Ser	Leu	Gly	Ser	Thr	Asn	Leu		
				245					250					255			
Tyr	Gly	Ser	Gly	Leu	Val	Asn	Ala	Glu	Ala	Ala	Thr	Arg					
			260					265									

<210> 9

<211> 269

<212> PRT

5 <213> Bacillus gibsonii

<400> 9

ES 2 763 577 T3

Gln Gln Thr Val Pro Trp Gly Ile Thr Arg Val Gln Ala Pro Thr Val
 1 5 10 15

His Asn Arg Gly Ile Thr Gly Ser Gly Val Lys Val Ala Ile Leu Asp
 20 25 30

Thr Gly Ile Ala Gln His Ser Asp Leu Thr Ile Arg Gly Gly Ala Ser
 35 40 45

Phe Val Pro Gly Glu Ser Thr Thr Ala Asp Leu Asn Gly His Gly Thr
 50 55 60

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

Gly Val Ala Pro Ser Ala Asp Leu Tyr Ala Val Lys Val Leu Gly Ala
 85 90 95

Asn Gly Arg Gly Ser Val Ser Gly Ile Ala Gln Gly Leu Glu Trp Ala
 100 105 110

Ala Thr Asn Asn Met His Ile Ala Asn Met Ser Leu Gly Ser Asp Ala

ES 2 763 577 T3

	115						120						125			
Pro	Ser	Thr	Thr	Leu	Glu	Arg	Ala	Val	Asn	Tyr	Ala	Thr	Ser	Arg	Gly	
	130					135					140					
Val	Leu	Val	Ile	Ala	Ala	Thr	Gly	Asn	Asn	Gly	Thr	Gly	Ser	Ile	Gly	
145					150					155					160	
Tyr	Pro	Ala	Arg	Tyr	Ala	Asn	Ala	Met	Ala	Val	Gly	Ala	Thr	Asp	Gln	
				165					170					175		
Asn	Asn	Arg	Arg	Ala	Ser	Phe	Ser	Gln	Tyr	Gly	Thr	Gly	Ile	Asp	Ile	
			180					185					190			
Val	Ala	Pro	Gly	Val	Gly	Ile	Gln	Ser	Thr	Tyr	Leu	Asn	Asn	Ser	Tyr	
		195					200					205				
Ala	Ser	Met	Pro	Gly	Thr	Ser	Met	Ala	Thr	Pro	His	Val	Ala	Gly	Val	
	210					215					220					
Ala	Ala	Leu	Val	Lys	Gln	Lys	Asn	Pro	Ser	Trp	Asn	Ala	Thr	Gln	Ile	
225					230					235					240	
Arg	Asn	His	Leu	Lys	Asn	Thr	Ala	Thr	Asn	Leu	Gly	Asn	Ser	Ser	Gln	
				245					250					255		
Phe	Gly	Ser	Gly	Leu	Val	Asn	Ala	Asp	Ala	Ala	Thr	Arg				
			260					265								

<210> 10

<211> 269

<212> PRT

5 <213> Bacillus lentus

<400> 10

ES 2 763 577 T3

65					70						75				80
Gly	Val	Ala	Pro	Ser	Ala	Glu	Leu	Tyr	Ala	Val	Lys	Val	Leu	Gly	Ala
				85					90					95	
Asp	Gly	Arg	Gly	Ala	Ile	Ser	Ser	Ile	Ala	Gln	Gly	Leu	Glu	Trp	Ala
			100					105					110		
Gly	Asn	Asn	Gly	Met	His	Val	Ala	Asn	Leu	Ser	Leu	Gly	Ser	Pro	Ser
		115					120					125			
Pro	Ser	Ala	Thr	Leu	Glu	Gln	Ala	Val	Asn	Ser	Ala	Thr	Ser	Arg	Gly
	130					135					140				
Val	Leu	Val	Val	Ala	Ala	Ser	Gly	Asn	Ser	Gly	Ala	Ser	Ser	Ile	Ser
145					150					155					160
Tyr	Pro	Ala	Arg	Tyr	Ala	Asn	Ala	Met	Ala	Val	Gly	Ala	Thr	Asp	Gln
				165					170					175	
Asn	Asn	Asn	Arg	Ala	Ser	Phe	Ser	Gln	Tyr	Gly	Ala	Gly	Leu	Asp	Ile
			180					185					190		
Val	Ala	Pro	Gly	Val	Asn	Val	Gln	Ser	Thr	Tyr	Pro	Gly	Ser	Thr	Tyr
		195					200					205			
Ala	Ser	Leu	Asn	Gly	Thr	Ser	Met	Ala	Thr	Pro	His	Val	Ala	Gly	Ala
	210					215					220				
Ala	Ala	Leu	Val	Lys	Gln	Lys	Asn	Pro	Ser	Trp	Ser	Asn	Val	Gln	Ile
225					230					235					240
Arg	Asn	His	Leu	Lys	Asn	Thr	Ala	Thr	Ser	Leu	Gly	Ser	Thr	Asn	Leu
				245					250					255	
Tyr	Gly	Ser	Gly	Leu	Val	Asn	Ala	Glu	Ala	Ala	Thr	Arg			
			260					265							

<210> 11

<211> 30

<212> DNA

<213> Artificial
<220>
<223> cebador en dirección ascendente
<400> 11
5 atatgaattc gctgaggaag caaaagaaaa 30
 <210> 12
 <211> 31
 <212> DNA
 <213> Artificial
10 <220>
 <223> cebador en dirección descendente
 <400> 12
 atatggatcc ttagcgtgtt gccgcttctg c 31
 <210> 13
15 <211> 30
 <212> DNA
 <213> Artificial
 <220>
 <223> cebador en dirección ascendente
20 <400> 13
 atatgaattc gctgaggaag caaaagaaaa 30
 <210> 14
 <211> 31
 <212> DNA
25 <213> Artificial
 <220>
 <223> cebador en dirección descendente
 <400> 14
 atatggatcc ttagcgcgtt gctgcatctg c 31
30 <210> 15
 <211> 43
 <212> DNA
 <213> Artificial
 <220>
35 <223> cebador en dirección ascendente
 <400> 15
 atataagctt aaggaggata ttatgatgag gaaaaagagt ttt 43

<210> 16
<211> 34
<212> DNA
<213> Artificial
5 <220>
<223> cebador en dirección descendente
<400> 16
atatgaattc agctgcagaa gcggaatcgc tga 34
<210> 17
10 <211> 43
<212> DNA
<213> Artificial
<220>
<223> cebador en dirección ascendente
15 <400> 17
atataagctt aaggaggata ttatgaaaaa actattcaaa acc 43
<210> 18
<211> 34
<212> DNA
20 <213> Artificial
<220>
<223> cebador en dirección descendente
<400> 18
atatgaattc agcagccgcc gcagattgtg agaa 34
25 <210> 19
<211> 43
<212> DNA
<213> Artificial
<220>
30 <223> cebador en dirección ascendente
<400> 19
atataagctt aaggaggata ttatggcgaa accactatca aaa 43
<210> 20
<211> 34
35 <212> DNA
<213> Artificial
<220>

<223> cebador en dirección descendente

<400> 20

atatgaattc agcagcgtct gccgcgggta aacc 34

<210> 21

5 <211> 43

<212> DNA

<213> Artificial

<220>

<223> cebador en dirección ascendente

10 <400> 21

atataagctt aaggaggata ttatgacatt gactaaactg aaa 43

<210> 22

<211> 34

<212> DNA

15 <213> Artificial

<220>

<223> cebador en dirección descendente

<400> 22

atatgaattc agcggcaagt gcctgactgg aaaa 34

20

REIVINDICACIONES

1. Vector de expresión que comprende
 - a) una secuencia de promotor y
 - b) una secuencia de ácido nucleico que codifica para una proteína, en donde la proteína comprende un péptido señal y otra secuencia de aminoácidos y el péptido señal comprende una secuencia de aminoácidos que es idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 6 en al menos el 80 %, y la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa.
2. Vector de expresión según la reivindicación 1, caracterizado porque el péptido señal codificado por la secuencia de ácido nucleico b) presenta una secuencia de aminoácidos según SEQ ID NO. 6.
3. Vector de expresión según las reivindicaciones 1 o 2, caracterizado porque el péptido señal se encuentra dispuesto en posición N terminal de la otra secuencia de aminoácidos en la proteína codificada por la secuencia de ácido nucleico b).
4. Vector de expresión según una de las reivindicaciones 1 a 3, caracterizado porque la proteína codificada por la secuencia de ácido nucleico b) comprende además una secuencia de conexión que se encuentra dispuesta entre el péptido señal y la otra secuencia de aminoácidos de la proteína, principalmente en donde la secuencia de conexión presenta una longitud de entre 1 y 50 aminoácidos.
5. Vector de expresión según una de las reivindicaciones 1 a 4, caracterizado porque la secuencia de aminoácidos de la proteasa
 - es idéntica en al menos el 80 % a SEQ ID NO. 7, o
 - es idéntica en al menos el 80 % a SEQ ID NO. 8, o
 - es idéntica en al menos el 80 % a SEQ ID NO. 9, o
 - es idéntica en al menos el 80 % a SEQ ID NO. 10 y en la posición 99 en el listado según SEQ ID NO. 10 presenta el aminoácido ácido glutámico (E) o ácido aspártico (D), o es idéntica en al menos el 80 % a SEQ ID NO. 10 y en la posición 99 en el listado según SEQ ID NO. 10 presenta el aminoácido ácido glutámico (E) o ácido aspártico (D) y además en el listado según SEQ ID NO. 10 presenta al menos uno de los siguientes aminoácidos:
 - (a) treonina en la posición 3 (3T),
 - (b) isoleucina en la posición 4 (4I),
 - (c) alanina, treonina o arginina en la posición 61 (61A, 61T o 61R),
 - (d) ácido aspártico o ácido glutámico en la posición 154 (154D o 154E),
 - (e) prolina en la posición 188 (188P),
 - (f) metionina en la posición 193 (193M),
 - (g) isoleucina en la posición 199 (199I),
 - (h) ácido aspártico, ácido glutámico o glicina en la posición 211 (211D, 211E o 211G),
 - (i) combinaciones de los aminoácidos (a) a (h).
6. Célula anfitriona no humana que contiene un vector de expresión según una de las reivindicaciones 1 a 5.
7. Célula anfitriona según la reivindicación 6, caracterizada porque es una bacteria, preferiblemente una que se selecciona del grupo de los géneros de *Escherichia*, *Klebsiella*, *Bacillus*, *Staphylococcus*, *Corynebacterium*, *Arthrobacter*, *Streptomyces*, *Stenotrophomonas* y *Pseudomonas*, más preferiblemente una que se selecciona del grupo de *Escherichia coli*, *Klebsiella planticola*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus amiloliquefaciens*, *Bacillus subtilis*, *Bacillus alcalophilus*, *Bacillus globigii*, *Bacillus gibsonii*, *Bacillus clausii*, *Bacillus halodurans*, *Bacillus pumilus*, *Staphylococcus carnosus*, *Corynebacterium glutamicum*, *Arthrobacter oxidans*, *Streptomyces lividans*, *Streptomyces coelicolor* y *Stenotrophomonas maltophilia*, principalmente *Bacillus licheniformis*.
8. Procedimiento para la preparación de una proteasa que comprende los pasos procedimentales de
 - (a) cultivar una célula anfitriona según una de las reivindicaciones 6 o 7
 - (b) aislar la proteasa del medio de cultivo o de la célula anfitriona.

9. Uso de un vector de expresión según una de las reivindicaciones 1 a 5 o de una célula anfitriona según una de las reivindicaciones 6 o 7 para la preparación de una proteasa.

Figura 1

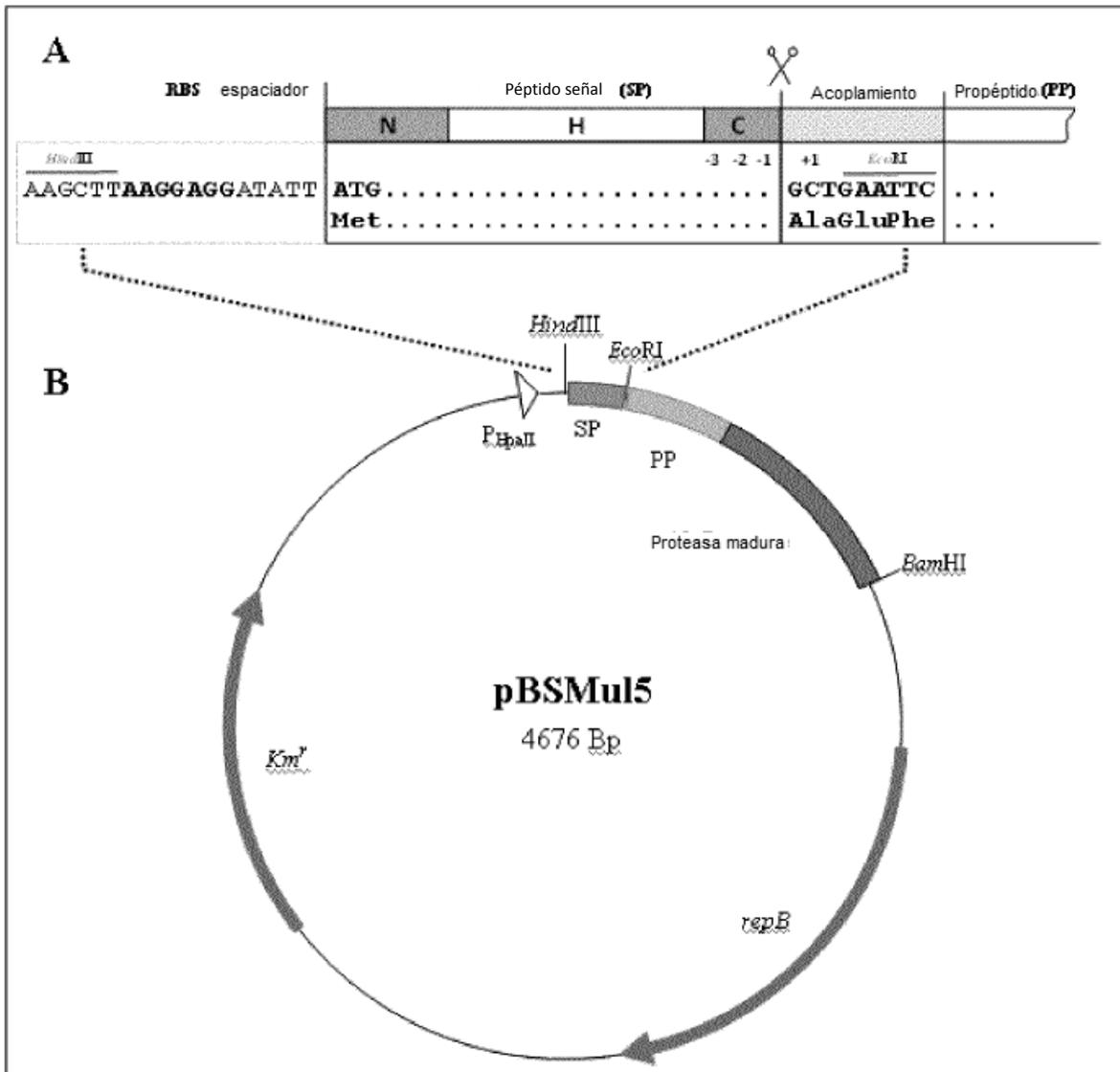


Figura 2

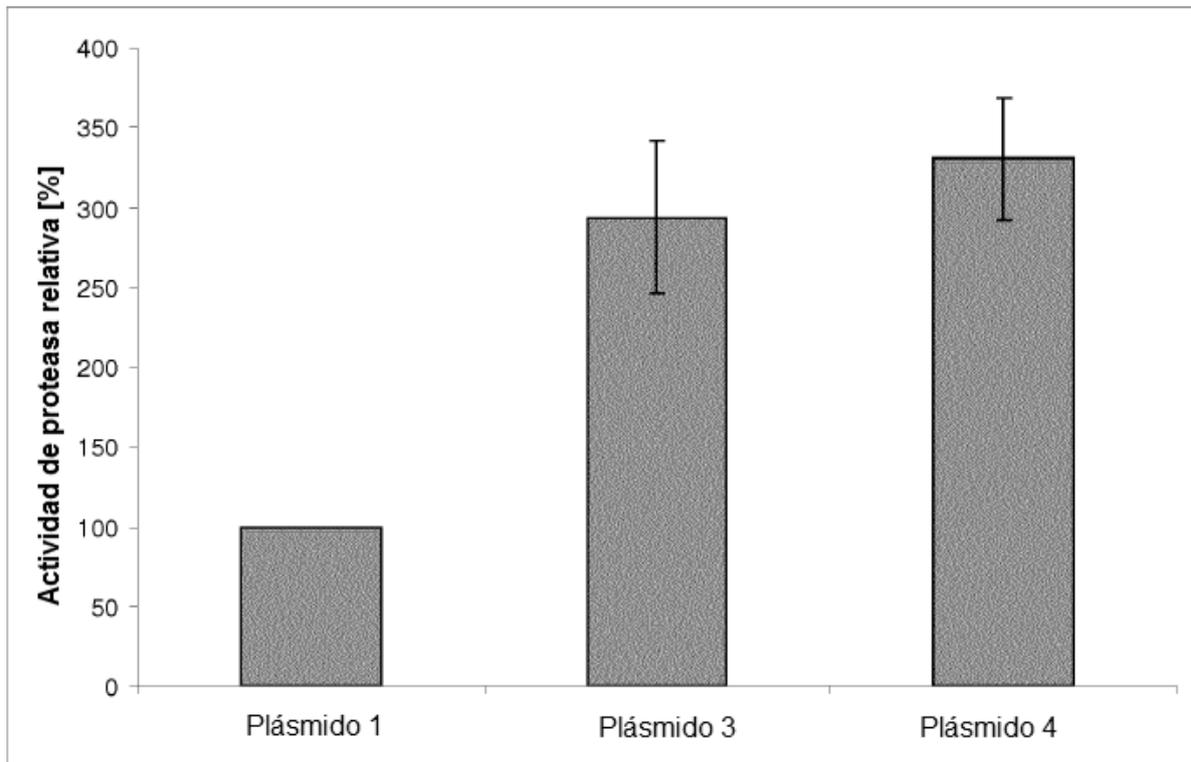


Figura 3

