

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 606 553**

51 Int. Cl.:

C12N 15/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.05.2012 PCT/EP2012/059901**

87 Fecha y número de publicación internacional: **06.12.2012 WO12163855**

96 Fecha de presentación y número de la solicitud europea: **25.05.2012 E 12723512 (5)**

97 Fecha y número de publicación de la concesión europea: **14.09.2016 EP 2714902**

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

30 Prioridad:

31.05.2011 DE 102011118032

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

24.03.2017

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 2 606 553 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 una secreción de proteína mejorada

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

10 Para la preparación de proteínas, pueden usarse células hospederas, 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 hospederas de tal manera que es expresado por éstas. Frecuentemente, éste se encuentra en un denominado vector de expresión junto con una o más secuencias promotoras (promotores), por lo cual se hace posible la expresión génica.

15 Para la producción biotecnológica a escala industrial, las células hospederas relevantes son cultivadas en fermentadores que están correspondientemente adaptados a las propiedades metabólicas de las células. Durante el cultivo, las células hospederas metabolizan el sustrato suministrado y forman el producto deseado, que después de que termina la fermentación, es usualmente separado de los organismos de producción y es purificado y/o concentrado de la suspensión del fermentador y/o medio de fermentación.

20 Fundamentalmente es deseable obtener un rendimiento alto del producto en la fermentación. El rendimiento de producto depende de múltiples factores; por ejemplo, las células hospederas usualmente forman, además del producto propiamente deseado, una gran cantidad de sustancias adicionales que generalmente no son de interés. Además, la expresión de un transgén y por lo tanto el rendimiento del producto depende sustancialmente del sistema de expresión usado. Por ejemplo, la solicitud de patente internacional WO 91/02792 describe la producción fermentativa mejorada de una proteasa alcalina de *Bacillus lentus* en una cepa optimizada de *Bacillus licheniformis* bajo el control de las secuencias reguladoras de gen de *Bacillus licheniformis*, principalmente el promotor de *Bacillus licheniformis*.

25 Para la producción industrial de proteínas, por ejemplo enzimas hidrolíticas, se da preferencia al uso de células hospederas capaces de secretar grandes cantidades de la proteína en el sobrenadante de cultivo, lo que hace redundante la compleja disrupción de células, que es necesaria en la producción intracelular. Para este propósito, se da preferencia al uso de células hospederas, por ejemplo especies de *Bacillus*, que pueden cultivarse usando medios económicos de cultivo en fermentaciones eficientes de alta densidad celular y son capaces de secretar más
30 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 hospedera y codifican la proteína que ha de ser secretada. La proteína expresada usualmente comprende un péptido de señal (secuencia de señal) que produce la exportación del mismo desde las células hospederas. El péptido de señal es habitualmente parte de la cadena de polipéptido traducida en la célula hospedera, pero además puede ser disociada después de la traducción por la proteína, dentro o fuera de la célula hospedera.
35

40 Precisamente para esta producción extracelular de proteínas heterólogas, no obstante, numerosas imitaciones y por consiguiente una demanda alta para optimizar los fenómenos de secreción. Una de estas imitaciones es la selección de un péptido de señal que permita la exportación eficiente de la proteína objetivo desde la célula hospedera. Los péptidos de señal pueden ser básicamente recién combinados 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 de cribado de una biblioteca de péptidos de señal en el ejemplo de una cutinasa. Sin embargo, no todo péptido de señal provoca una exportación suficiente desde la proteína en condiciones de fermentación, principalmente en condiciones industriales de fermentación o de escala industrial.

45 El documento WO 2011/015327 A1 divulga un procedimiento para la producción de nucleasas. El documento WO 03/054185 A1 divulga nuevas proteasas de tipo silvestre del tipo Subtilisin. El documento DE 10 2007 049 830 A1 divulgan la preparación de proteasas con ayuda de permutación circular.

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

50 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 una proteína y la proteína comprende un péptido de señal y otra secuencia de aminoácidos; el péptido de señal comprende una secuencia de aminoácidos que es idéntica en al

menos 80 % a la secuencia de aminoácidos indicada en SEQ ID NO. 2, y la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa.

5 De manera sorprendente se ha encontrado que por medio de un vector de expresión que codifica una proteína con un péptido de señal de este tipo, se logra una secreción mejorada de la proteína desde una célula hospedera que incluye el vector de expresión y expresa la secuencia de ácido nucleico b). Como resultado, en realizaciones preferidas de la invención es posible incrementar el rendimiento de producto de proteína en una fermentación.

Un vector de expresión es una secuencia de ácido nucleico que ocasiona que puede expresarse la proteína en una célula hospedera, principalmente un microorganismo. Éste comprende la información genética, es decir aquella secuencia de ácido nucleico (gen) b) que codifica la proteína.

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

20 Los vectores son elementos genéticos que se componen de ácidos nucleicos, de preferencia ácido desoxirribonucleico (ADN) y son conocidos por el experto en el campo de la biotecnología. Principalmente al usar en bacterias son plásmidos específicos, 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 preponderantemente vectores o plásmidos sintéticos con elementos del origen más diverso. Con los otros elementos genéticos respectivamente presentes, los vectores son capaces de establecerse en células hospederas en las cuales se han introducido preferiblemente mediante transformación, durante múltiples generaciones como unidades estables. En el contexto de la invención en este caso no es considerable si se establecen de modo extracromosómico como unidades propias o se integran a cromosoma o a ADN cromosómico. Cuál de los numerosos sistemas se selecciona depende del caso individual. Decisivos pueden ser, por ejemplo, la cantidad de copias que puede lograrse, los sistemas disponibles de selección, entre éstos ante todo la resistencia a antibióticos o la capacidad de cultivarse de las células hospederas capaces de alojar los vectores.

Los vectores de expresión pueden además regularse a través de cambios en las condiciones de cultivo como, por ejemplo, la densidad celular o la adición de determinados compuestos. Un ejemplo de un compuesto tal es el derivado de galactosa isopropil- β -D-tiogalactopiranosido (IPTG), que se usa como un activador del operón de lactosa bacteriano (operón lac).

35 Un vector de expresión comprende además al menos una secuencia de ácido nucleico, preferiblemente ADN, con una función de control para la expresión de la secuencia de ácido nucleico b) que codifica la proteína (una secuencia llamada reguladora del gen). Una secuencia reguladora del gen en este caso es cualquier secuencia de ácido nucleico cuya presencia en la célula hospedera respectiva influya, preferiblemente incremente, la frecuencia de transcripción de la secuencia de ácido nucleico b) que codifica la proteína. De preferencia se trata de una secuencia de promotor de la que una secuencia de este tipo es esencial para la expresión de la secuencia de ácido nucleico b). Un vector de expresión según la invención también puede comprender, no obstante, todavía otras secuencias reguladoras de gen, por ejemplo una o varias secuencias promotoras. Un vector de expresión en el marco de la invención comprende, por consiguiente, al menos una unidad funcional de la secuencia de ácido nucleico b y promotor (casete de expresión). Esta puede, pero no tiene que obligatoriamente existir como unidad física. El promotor causa la expresión de la secuencia de ácido nucleico b) en la célula hospedera. En el contexto de la presente invención, un vector de expresión también puede limitarse a los casetes de expresión puros de promotor y la secuencia de ácido nucleico b) que va expresarse, y este casete de expresión puede presentarse de modo integrado a nivel extracromosómico o también cromosómico. Configuraciones de este tipo de los vectores de expresión según la invención representan respectivamente una forma de realización separada de la invención

50 La presencia de por lo menos un promotor es consecuentemente esencial para un vector de expresión de conformidad con la invención. Por lo tanto, por un promotor se entiende una secuencia de ADN que permite la expresión regulada de un gen. Una secuencia de promotor es naturalmente un componente de un gen y a menudo está situada en su extremo 5' y por lo tanto antes de la región codificante de ARN. Preferiblemente, la secuencia de promotor en un vector de expresión de conformidad con la invención está situada hacia el extremo 5' de la secuencia de ácido nucleico b) que codifica la proteína. La propiedad más importante de un promotor es la interacción específica con al menos una proteína o polipéptido de unión a ADN, que es mediador del inicio de la transcripción del gen por medio de una ARN polimerasa y se denomina factor de transcripción. Varios factores de transcripción y/o proteínas adicionales frecuentemente participan en el inicio de la transcripción por medio de una ARN polimerasa. Un promotor por lo tanto es preferiblemente una secuencia de ADN que tiene actividad de promotor, es decir, una

secuencia de ADN a la cual se une al menos un factor de transcripción por lo menos transitoriamente a fin de iniciar la transcripción de un gen. La potencia de un promotor es medible por la frecuencia de transcripción del gen expresado, es decir, a través de la cantidad de moléculas de ARN, muy particularmente moléculas de ARNm, generadas por unidad de tiempo.

5 La secuencia de un promotor (a) y la secuencia de ácido nucleico (b) se encuentran una detrás de otra en el vector de expresión. La secuencia del promotor (a) se encuentra más preferiblemente delante de la secuencia de ácido nucleico (d) en la molécula de ácido nucleico (en la orientación 5' → 3'). De modo igualmente preferido, entre ambas secuencias de ácido nucleico (a) y (b) no se encuentran secuencias de ácido nucleico que reduzcan la frecuencia de transcripción de la secuencia de ácido nucleico (b) que codifica la proteína. Todos los datos previos se refieren a aquella hebra de ADN que contiene la secuencia de ácido nucleico (b) que codifica la proteína (la hebra codificante) y no la hebra codogénica correspondiente de ADN. A partir de la secuencia de ácido nucleico (d) que codifica la proteína se encuentra consecuentemente la secuencia de promotor (a) de preferencia en dirección ascendente, es decir en dirección 5', sobre esta hebra de ADN.

10 La secuencia de ácido nucleico b) codifica la proteína que ha de ser secretada. En este caso se trata de aquella proteína que debe prepararse con ayuda de un vector de expresión según la invención (proteína objetivo).

15 La proteína codificada por la secuencia de ácido nucleico b) comprende un péptido de señal con una secuencia de aminoácidos que es por lo menos 80 % idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 2 o es por lo menos 80 % idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 4 o es por lo menos 80 % idéntica a la secuencia de aminoácidos indicada en la SEQ ID NO. 6. Se estableció que dichos péptidos de señal efectúan una secreción eficiente de la proteína que los contiene, principalmente proteína recombinante. Con preferencia cada vez mayor, el péptido de señal comprende una secuencia de aminoácidos que es por lo 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 100 % idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 2, o es por lo 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 100 % idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 4 o es 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 100 % idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 6. El péptido de señal presenta de manera particularmente preferida una secuencia de aminoácidos que es 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 100 % idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 2, o es por lo 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 100 % idéntica a la secuencia de aminoácidos indicada en SEQ ID NO. 4, o es 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 100 % idéntica a la secuencia de aminoácidos picada en SEQ ID NO. 6.

20 De manera muy particular respectivamente se prefieren secuencias 100 % idénticas de modo que un vector de expresión preferido de manera correspondiente se caracteriza porque el péptido de señal codificado por la secuencia de ácido nucleico b) tiene una secuencia de aminoácidos de acuerdo con SEQ ID NO. 2, SEQ ID NO. 4 o SEQ ID NO. 6. Para péptidos de señal de este tipo se indican secuencias de ácido nucleico codificantes, particularmente preferidas, en SEQ ID NO. 1, SEQ ID NO. 3 y SEQ ID NO. 5.

25 Además es posible que en lugar de los péptidos de señal mencionados que hacen posible una secreción de la proteína se usen secuencias homólogas estructuralmente a estas secuencias. Por una secuencia estructuralmente homóloga se entiende una secuencia de aminoácidos cuya sucesión de aminoácidos presenta un doblez espacial comparable con un péptido de señal con una secuencia de aminoácidos de acuerdo con SEQ ID NO. 2, SEQ ID NO. 4 o SEQ ID NO. 6. Este doblez espacial causa que sea reconocida por la célula hospedera como secuencia secretora de señal y en consecuencia la proteína que contiene la secuencia de señal estructuralmente homóloga se descargue de la célula hospedera. De preferencia se efectúa una interacción con el sistema de translocación usado por la célula hospedera. La secuencia de aminoácidos estructuralmente homóloga se enlaza, por lo tanto, de preferencia directa o indirectamente con al menos un componente del sistema de translocación de la célula hospedera. Por enlace directo se entiende una interacción directa; por enlace indirecto se entiende que la interacción puede efectuarse mediante uno o varios componentes adicionales, principalmente proteínas u otras moléculas que funcionan como adaptadores 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 hospedera.

30 La determinación de la entidad de secuencias de ácido nucleico y aminoácidos se efectúa mediante una comparación de secuencias. Una comparación de este tipo se efectúa asignando entre sí sucesiones similares en la secuencia de nucleótidos o secuencia de aminoácidos. Esta comparación de secuencias se efectúa de preferencia con base en el algoritmo BLAST establecido en el estado de la técnica y habitualmente usado (compárese, por ejemplo, Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410, y Altschul, Stephan F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Hheng

Zhang, Webb Miller, and David J. Lipman (1997): "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs"; *Nucleic Acids Res.*, 25, páginas 3389-3402) y ocurre en principio asignando entre sí sucesiones similares de nucleótidos o aminoácidos en las secuencias de ácido nucleico y de aminoácidos. Una asignación tabular de las posiciones relevantes se denomina alineación. Otro algoritmo disponible en el estado de la técnica es el algoritmo FASTA. Las comparaciones de secuencias (alineaciones), principalmente comparaciones múltiples de secuencias, se producen habitualmente con programas de ordenador. Con frecuencia 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 basados en estos programas o algoritmos. En el contexto de la presente invención se producen comparaciones de secuencias y alineaciones preferiblemente con el programa de ordenador Vector NTI® Suite 10.3 (Invitrogen Corporation, 1600 Faraday Avenue, Carlsbad, California, Estados Unidos) con los parámetros estándares (por defecto) predefinidos.

Una comparación de este tipo permite una declaración sobre la similitud de las secuencias comparadas entre sí. Habitualmente se indica en porcentaje de identidad, es decir la fracción de los nucleótidos o residuos de aminoácidos idénticos en las mismas posiciones o posiciones correspondientes entre sí en una alineación. El concepto amplio de homología toma en consideración, en el caso de secuencias de aminoácidos, a sustituciones de aminoácidos conservadas, es decir aminoácidos que tienen propiedades similares ya que estos la mayoría de las veces ejercen actividades o funciones similares dentro de la proteína. Por lo tanto, la similitud de las secuencias comparadas también puede indicarse en porcentaje de homología o en porcentaje de similitud. Los datos de identidad y/o de homología pueden reportarse por polipéptidos o genes completos o solamente por regiones individuales. Las regiones homologadas o idénticas de diferentes secuencias de ácido nucleico o de aminoácidos son, por lo tanto, definidos mediante coincidencias en las secuencias. Frecuentemente tienen funciones iguales o similares. Pueden ser pequeñas y comprender solamente unos pocos nucleótidos o aminoácidos. Con frecuencia tales regiones pequeñas ejercen funciones esenciales para la actividad total de la proteína. Por lo tanto, puede ser práctico referir las coincidencias de secuencias solamente a regiones individuales, opcionalmente pequeñas. En tanto no se indique algo diferente, sin embargo, los datos de identidad y de homología en la presente invención se refieren a la longitud total de la secuencia de ácido nucleico o de aminoácidos respectivamente indicada.

La proteína codificada por la secuencia de ácido nucleico d) comprende además otra secuencia de aminoácidos. Esta secuencia de aminoácidos es consecuentemente la secuencia de aminoácidos propia de la proteína sin péptido de señal. De preferencia se trata de la secuencia de aminoácidos de una proteína madura. Por proteína madura se entiende su forma tratada terminada, puesto que es posible que una forma inmadura se codifique por un gen correspondiente, la cual es tratada después de la traducción para dar la forma madura. A manera de ejemplo, las formas inmaduras de la proteína pueden comprender péptido de señal y/o propéptidos o elongaciones en el terminal N y/o terminal C que ya no están presentes en la forma madura. A manera de ejemplo, las formas inmaduras de proteasas, principalmente subtilasas y entre éstas ante todo subtilisinas, comprenden un péptido de señal así como un péptido que ya no están presentes en la forma madura de la proteasa. De modo alternativo, la otra secuencia de aminoácidos es una secuencia de aminoácidos de una proteína inmadura que comprende un propéptido. Una configuración de este tipo se toman en consideración principalmente incluso para proteasas, principalmente subtilasas y entre éstas ante todo subtilisinas. En configuraciones principalmente preferidas la otra secuencia de aminoácidos no comprende otro péptido de señal. En configuraciones de este tipo según la invención, por lo tanto, solamente el péptido de señal de la invención produce la secreción de la proteína desde una célula hospedera.

La otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa. También se divulgan secuencias de aminoácidos de la proteína en las cuales la secuencia de aminoácidos comprende amilasa, celulasa, hemicelulasa, mananasa, tanasa, xilanasa, xantanasa, xiloglucanasa, β -glucosidasa, un enzima disociadores de pectina, carragenasa, perhidrolasa, oxidasa, oxidoreductasa o una lipasa. La otra secuencia de aminoácidos de la proteína comprende de manera muy particularmente preferida la secuencia de aminoácidos de una subtilisina.

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

Entre las proteasas se prefieren subtilisinas. Ejemplos de estas son las subtilisinas BPN' y Carlsberg, la proteasa PB92, la subtilisina 147 y 309, la proteasa alcalina de *Bacillus lentus*, subtilisina DY y las enzimas termitasa, proteinasa K las proteasa TW3 y TW7 que deben ser asignadas a las subtilasas, aunque ya no a las subtilisinas en un sentido estrecho. La subtilisina Carlsberg se encuentra disponible en forma desarrollada adicionalmente bajo el nombre comercial Alcalase® de la compañía Novozymes A/S, Bagsvaerd, Dinamarca. La subtilisina 147 y 309 son vendidas bajo los nombres comerciales Esperase®, y Savinase® de la compañía Novozymes. De la proteasa de *Bacillus lentus* DSM 5483 se derivan las variantes de proteasa conocidas bajo la denominación BLAP®. Otras proteasas preferidas son también, por ejemplo, las enzimas conocidas bajo la denominación PUR. Otras proteasas también se encuentran 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 la denominación Proteinase K-16 de la compañía Kao Corp., Tokio, Japón. También se prefieren las proteasas de *Bacillus gibsonii* y *Bacillus pumilus* que 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 amyloliquefaciens* o de *Bacillus stearothermophilus* así como principalmente de sus desarrollos adicionales mejorados para uso en productos detergentes o de limpieza. La enzima de *Bacillus licheniformis* se encuentra disponible en la compañía Novozymes bajo el nombre Termamyl® y en la compañía Danisco/Genencor bajo el nombre Purastar®ST. Los productos de desarrollo adicional de esta α -amilasa se encuentran disponibles en la compañía Novozymes bajo los nombres comerciales Duramyl® y Termamyl®Ultra, en la compañía Danisco/Genencor bajo el nombre Purastar®OxAm y en la compañía Daiwa Seiko Inc., Tokio, Japón, como Keistase®. La α -amilasa de *Bacillus amyloliquefaciens* es vendida en la compañía Novozymes bajo el nombre BAN®, y las variantes derivadas de la α -amilasa de *Bacillus stearothermophilus* bajo los nombres BSG® y Novamyl®, también en la compañía Novozymes. Además, pueden mencionarse la α -amilasa de *Bacillus sp. A 7-7* (DSM 12368) y las ciclodextrina-glucanotransferasa (CGTasa) de *Bacillus agaradherens* (DSM 9948). También pueden emplearse productos de fusión de todas las moléculas mencionadas. Además, son adecuados los desarrollos adicionales de la α -amilasa de *Aspergillus niger* y *A. oryzae* disponibles bajo los nombres comerciales Fungamyl® en la compañía Novozymes. Otros productos comerciales ventajosos son, por ejemplo, la amilasa Powerase® de la compañía Danisco/Genencor y las amilasas Amylase-LT®, Stainzyme® y Stainzyme plus®, estas últimas de la compañía Novozymes. También pueden producirse variantes de estas enzimas disponibles mediante mutaciones puntuales. Otras amilasas preferidas se divulgan en las publicaciones internacionales de solicitud de patente WO 00/60060, WO 03/002711, WO 03/054177 y WO07/079938, a cuya divulgación se hace referencia expresamente y su contenido de divulgación correspondiente se incorpora a la presente solicitud de patente de manera expresa. Las amilasas también son de preferencia α -amilasas.

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

Ejemplos de celulasas (endoglucanasas, EG) comprenden secuencias de preparación de celulasa rica en endoglucanasa(EG) de hongos, o los desarrollos adicionales de la misma, que son suministrados por Novozymes bajo el nombre comercial Celluzyme®. Los productos Endolase® y Carezyme®, igualmente disponibles en Novozymes, se basan en 50 kD-EG y 43 kD-EG, respectivamente, de *Humicola insolens* DSM 1800. Otros productos comerciales de esta compañía que pueden prepararse son Cellusoft®, Renozyme® y Celluclean®. Es posible además preparar, por ejemplo, celulasas que están disponibles en AB Enzymes, Finlandia, bajo los nombres comerciales Ecostone® y Biotouch® y que se basan al menos parcialmente en 20 kD-EG de *Melanocarpus*. Celulasas adicionales de AB Enzymes son Econase® y Ecopulp®. Otras celulasas adecuadas pueden obtenerse de *Bacillus sp. CBS 670.93* y *CBS 669.93*, y aquella del *Bacillus sp. CSS 670.93* se encuentra disponible en Danisco/Genencor bajo el nombre comercial de Puradax®. Otros productos comerciales de Danisco/Genencor que se pueden preparar son "celulasa L de detergente de Genencor" e IndiAge® Neutra:

Variantes de estas enzimas obtenibles por mutaciones puntuales también pueden prepararse de acuerdo con la invención. Las celulasas particularmente preferidas son variantes de celulasa de *Thielavia terrestris* que se describen en la publicación de solicitud internacional de patente WO 98/12307, celulasas de *Melanocarpus*, muy particularmente *Melanocarpus albomyces*, que se describen en la publicación de solicitud internacional de patente WO 97/14804, celulasas de tipo EGIII de *Trichoderma reesei* que se divulgan en la solicitud de patente europea EP 1 305 432 o variantes obtenibles de las mismas, principalmente aquellas que se divulgan en las solicitudes de patentes europeas EP 1240525 y EP 1305432, así como celulasas que se divulgan en las publicaciones de solicitudes internacionales de patente WO 1992006165, WO 96/29397 y WO 02/099091, a cuya respectiva divulgación, por lo tanto, se hace referencia expresa y cuyo contenido divulgado correspondiente se incorpora, por lo tanto, de manera expresa a la presente solicitud de patente.

Además, pueden prepararse otras enzimas que se recopilan bajo el término hemicelulasas. Estas incluyen, por ejemplo, mananasas, xantanoliasas, xantanasas, xiloglucanasas, xilanasas, pululanasas, enzimas disociadoras de

- pectina y β -glucanasas. La β -glucanasa obtenida de *Bacillus subtilis* está disponible bajo el nombre de Cereflo® de Novozymes. Las hemicelulasas particularmente preferidas son mananasas, que son vendidas, por ejemplo, bajo los nombres comerciales Mannaway® de Novozymes o Purabrite® de Genencor. Las enzimas disociadoras de pectina asimismo incluyen enzimas con las denominaciones pectinasa, pectatoliasa, pectinesterasa, pectinadesmetoxilasa, pectinametoxilasa, pectinametilsterasa, pectasa, pectinametilsterasa, pectinoesterasa, pectinapectilhidrolasa, pectinadespolimerasa, endopoligalacturonasa, pectolasa, pectina hidrolasa, pectina poligalacturonasa, endopoligalacturonasa, pectolasa, pectinahidrolasa, pectina-poligalacturonasa, poli- α -1,4-galacturonido glicanohidrolasa, endogalacturonasa, endo-D-galacturonasa, galacturano 1,4- α -galacturonidasa, exopoligalacturonasa, poli(galacturonato) hidrolasa, exo-D-galacturonasa, exo-D-galacturonanasa, exopoli-D-galacturonasa, exo-poli- α -galacturonosidasa, exo-poli-galacturonosidasa. Ejemplos de enzimas adecuadas correspondientes están disponibles, por ejemplo, bajo los nombres Gamanase®, Pektinex AR®, X-Pect® o Pectaway® de Novozymes, bajo el nombre Rohapect UF®, Rohapect TPL®, Rohapect PTE100®, Rohapect MPE®, Rohapect MA plus HC, Rohapect DA12L®, Rohapect 10L®, Rohapect B1L® de la compañía AB Enzymes y bajo el nombre Pyrolase® de la compañía Diversa Corp., San Diego, CA, Estados Unidos de América.
- Además, también es posible preparar oxidorreductasas, por ejemplo, oxidasas, oxigenasas, catalasas, peroxidasas, tales como haloperoxidasas, cloroperoxidasas, bromoperoxidasas, lignina-peroxidasas, glucosa-peroxidasas o manganeso-peroxidasas, dioxigenasas o lacasas (fenol-oxidasas, polifenol-oxidasas). Productos comerciales adecuados que pueden mencionarse son Denilite® 1 y 2 de Novozymes. Otras enzimas se divulgan en las solicitudes internacionales de patente WO 98/45398, WO 2005/056782, WO 2004/058961 y WO 2005/124012.
- En otra forma de realización de la invención, la otra secuencia de aminoácidos no está presente de modo natural junto con el péptido de 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 de manera natural significa por lo tanto que las dos secuencias de aminoácidos no son componentes de una proteína propia del microorganismo. Una proteína que comprende el péptido de señal y la otra secuencia de aminoácidos consecuentemente no puede ser expresada en el microorganismo por una secuencia de ácido nucleico que es, parte del ADN cromosómico del microorganismo en su forma de tipo silvestre. Dicha proteína y/o la secuencia de ácido nucleico que la codifica en cada caso no está consecuentemente presente en forma de tipo silvestre del microorganismo y/o no puede ser aislada de la forma de tipo silvestre del microorganismo. Ambas secuencias – el péptido de señal y la otra secuencia de aminoácidos – tienen por lo tanto que ser asignadas a dos diferentes cadenas de polipéptido en una forma de tipo silvestre de un microorganismo, en tanto ambas están, o pueden estar, presentes en su totalidad en forma de tipo silvestre de un microorganismo. Por lo tanto, en el contexto de esta configuración de la invención, el péptido de señal y la otra secuencia de aminoácido, o los ácidos nucleicos que los codifican, fueron combinados recientemente con ayuda de procedimientos de ingeniería genética y esta combinación del péptido de 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 del péptido de señal con la otra secuencia de aminoácidos, y específicamente ni al nivel de ADN ni al de proteína. Sin embargo, el péptido de señal y la otra secuencia de aminoácidos, o las secuencias de ácido nucleico que los codifican, pueden ser respectivamente de origen natural, pero su combinación no existe en la naturaleza. Pero, el péptido de señal y la otra secuencia de aminoácidos pueden provenir por sí mismos del mismo microorganismo o incluso de diferentes microorganismos.
- En una forma de realización preferida, un ácido nucleico de la invención se caracteriza porque es un ácido nucleico no natural. No natural significa que un ácido nucleico de la invención no puede aislarse de un organismo en su forma de tipo silvestre existente en la naturaleza. Por lo tanto, principalmente y con respecto a las bacterias de tipo silvestre, un ácido nucleico de la invención no es un ácido nucleico propio de bacterias. Las secuencias (a) y (b) no provienen preferiblemente del mismo o de los mismos organismos, principalmente bacterias, sino que provienen de diferentes organismos, principalmente bacterias. Las diferentes bacterias son, por ejemplo, bacterias que pertenecen a diferentes cepas o especies o géneros.
- En otra forma de realización de la invención el vector de expresión se caracteriza porque el péptido de señal está dispuesto N-terminal 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: N-terminal -péptido de señal - (secuencia de aminoácidos adicional opcional) -otra secuencia de aminoácidos- C-terminal. Una estructura de este tipo de la proteína que va a ser expresada ha demostrado 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 enlazamiento que está dispuesta entre el péptido señal de la otra secuencia de aminoácidos de la proteína. La proteína codificada por la secuencia de ácido nucleico b) presenta por lo tanto la siguiente estructura: N-terminal -péptido de señal -secuencia de enlazamiento (también "acoplamiento" o "espaciador") -otra secuencia de aminoácidos - C-terminal. Una estructura de este tipo de la proteína que va a expresarse también ha demostrado ser particularmente ventajoso. La secuencia de enlazamiento tiene preferiblemente una longitud entre 1 de 50 aminoácidos, 2 y 25 aminoácidos, entre 2 y 15 aminoácidos, entre 3 y 10 aminoácido y de modo particularmente preferido entre 3 y 5 aminoácidos. Un ejemplo de

una secuencia de enlazamiento particularmente preferida es la sucesión de aminoácidos alanina, ácido glutámico y fenilalanina (desde el N-terminal al C-terminal).

5 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 y la secuencia de aminoácidos de la proteasa es idéntica en al menos 80 % a la SEQ ID NO. 7. La secuencia de aminoácidos de la proteasa 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 en 100 % a la SEQ ID NO. 7.

10 De modo alternativo, la otra secuencia de aminoácidos de la proteína es la secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a la 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 la SEQ ID NO. 8.

15 De modo alternativo, la otra secuencia de aminoácidos de la proteína es la secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a la 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 % a la SEQ ID NO. 9.

20 De modo alternativo, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a la SEQ ID NO. 10 y en la posición 99 en el conteo según SEQ ID NO. 10 tiene el aminoácido ácido glutámico (E) o ácido aspártico (D). 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 a la SEQ ID NO. 10 en las posiciones 1 a 98 y 100 a 269 en el conteo según la SEQ ID NO. 10.

25 De modo alternativo, la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa que es idéntica en al menos 80 % a la SEQ ID NO. 10 y en la posición 99 en el conteo según SEQ ID NO. 10 presenta el aminoácido ácido glutámico (E) o ácido aspártico (D) y también en el conteo 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),
- 30 (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).

35 La secuencia de aminoácidos de esta proteasa es idéntica preferiblemente en 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 a la 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 en al menos dos

40 posiciones en comparación con la SEQ ID NO. 10, y la primera modificación en el conteo según SEQ ID NO. 10 es ácido glutámico en la posición 99 y la segunda modificación en el conteo según SEQ ID NO. 10 se selecciona del grupo que se compone de:

- (a) treonina en la posición 3 (3T),
- (b) isoleucina en la posición 4 (4I),
- 45 (c) alanina, treonina o arginina en la posición 61 (61A, 61T o 61 R),
- (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),

5 (i) combinaciones de los aminoácidos (a) a (h).

También de manera muy particularmente preferida la otra secuencia de aminoácidos de la proteína comprende la secuencia de aminoácidos de una proteasa que en comparación con la SEQ ID NO. 10 tiene una secuencia de aminoácidos modificada en al menos dos posiciones y la primera modificación en el conteo según SEQ ID NO. 10 es ácido aspártico en la posición 99 y la segunda modificación en el conteo según SEQ ID NO. 10 se selecciona del grupo que se compone de:

10

- (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 61 R),
- (d) ácido aspártico o ácido glutámico en la posición 154 (154D o 154E),

15

- (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).

20

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

25

30

En términos generales es válido que la otra secuencia de aminoácidos de la proteína no tenga que comprender solamente la secuencia de aminoácidos de una proteína madura; más bien, pueden estar comprendidas otras secuencias de aminoácidos, como por ejemplo los propéptidos de esta secuencia de aminoácidos. Esto aplica no solamente para proteasas sino para todas las proteínas, principalmente también todas las especies de enzimas.

35

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

40

45

Con todos los objetos de la invención y formas de realización mencionados, como otros objetos de la invención están asociadas las células hospederas no humanas que contienen vectores según la invención, los procedimientos de preparación en los cuales se usan células hospederas correspondientes y los usos de vectores o células hospederas correspondientes. Por lo tanto, las realizaciones precedentes se refieren a estos objetos de la invención de manera correspondiente.

Otro objeto de la invención es una célula hospedera no humana que contiene un vector de expresión según la invención. Un vector de expresión según la invención se introduce preferiblemente a la célula hospedera mediante

5 su transformación. Esto se efectúa según la invención de manera preferida transformando un vector de la invención en un microorganismo el cual luego representa una célula hospedera según la invención. De modo alternativo, a una célula hospedera también pueden introducirse componentes individuales, es decir partes o fragmentos de ácido nucleico, por ejemplo los componentes (a) y/o (b) de un vector según la invención de tal modo que la célula hospedera resultante luego contenga un vector según la invención. Este procedimiento es adecuado particularmente cuando la célula hospedera contiene ya uno 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 los procedimientos para la transformación de células y y son suficientemente conocidos por el especialista. Como células hospederas son adecuadas en teoría todas las células, es decir células procariotas o eucariotas. Se prefieren aquellas células hospederas que se dejan manejar genéticamente de manera ventajosa, lo cual afecta por ejemplo a la transformación con el vector y a su establecimiento estable, por ejemplo hongos o bacterias unicelulares. Las células hospederas preferidas se caracterizan además por una buena capacidad de manipulación microbiológica y biotecnológica. Esto se refiere por ejemplo a una fácil capacidad de cultivo, altas tasas de crecimiento, demandas bajas a los medios de fermentación y buenas tasas de producción y de secreción para proteínas extrañas. Con frecuencia, para el caso individual, tienen que determinarse experimentalmente los sistemas de expresión óptimos entre la abundancia de los diferentes sistemas disponibles en el estado de la técnica.

20 Otras formas de realización preferidas representan aquellas células hospederas que son regulables en su actividad debido a elementos de regulación genética que, por ejemplo, se hacen disponibles en el vector, pero también pueden estar presentes desde el principio en estas células. Por ejemplo, pueden estimularse para expresar mediante adición controlada de compuestos químicos que sirven como activadores, cambiando las condiciones de cultivo o al alcanzar una densidad particular de células. Esto permite una producción económica de las proteínas.

25 Las células hospederas preferidas son células procariotas o bacterianas. Las bacterias se caracterizan por tiempos breves de generación y bajas demandas a las condiciones de cultivo. De esta manera pueden establecerse procedimientos económicos. Además, el especialista en bacterias en la tecnología de fermentación dispone de una extensa experiencia. Para una producción especial pueden ser adecuadas bacterias gram-negativas o gram-positivas por las más diversas razones, que particularmente han de determinarse experimentalmente, tal como fuentes de nutrientes, velocidad de formación del producto, demandas de tiempo, etcétera.

30 En el caso de bacterias gram-negativas, como por ejemplo *Escherichia coli*, se secreta una gran cantidad de polipéptidos al espacio periplasmático, es decir al compartimiento entre las dos membranas que encierran las células. Esto puede ser ventajoso para aplicaciones especiales. Además, amén es posible configurar bacterias gram- negativas de tal manera que descarguen los polipéptidos expresados no solamente al espacio periplasmático, sino al medio que rodea la bacteria. Por el contrario, las bacterias gram- positivas, por ejemplo, bacilos o Actinomycetaceae u otros representantes de los Actinomycetales no tienen membrana exterior de modo que las proteínas secretadas se liberan inmediatamente al medio que rodea la bacteria, por lo regular el medio nutriente, del cual pueden purificarse los polipéptidos expresados. Pueden aislarse directamente del medio o tratarse adicionalmente. Además, las bacterias gram-positivas están relacionadas o son idénticas a la mayoría de los organismos de origen para enzimas industrialmente importantes y casi siempre forman enzimas comparables de modo que disponen de un uso de codón similar y su aparato de síntesis de proteínas está adaptado correspondientemente de manera natural.

40 Por uso de codón se entiende la traducción del código genético en aminoácidos, es decir cuál sucesión de nucleótidos (triplete o triplete base) codifica cuál aminoácido o cuál función, por ejemplo el inicio y el fin de la región que va a traducirse, los sitios de unión para diferentes proteínas, etcétera. De esta manera, cada organismo, principalmente cada cepa de producción, poseen un determinado uso de codón. Pueden ocurrir embotellamientos en la biosíntesis de proteína si los codones que se encuentran en el ácido nucleico transgénico en la célula hospedera se enfrentan a un número comparativamente bajo de ARNt cargados. Por el contrario, los codones sinónimos que codifican los mismos aminoácidos y pueden traducirse de manera más eficiente dependiendo del hospedero. Esta transcripción opcionalmente necesaria depende de esta manera de la elección del sistema de expresión. Principalmente en el caso de muestras compuestas por organismos desconocidos, eventualmente no cultivables puede ser necesaria una adaptación correspondiente.

50 En teoría, la presente invención es aplicable a todos los microorganismos, principalmente a todos los microorganismos fermentables, de modo particularmente preferido a aquellos del género *Bacillus*, y conduce a que empleando tales microorganismos como organismos de producción puede realizarse un rendimiento de producto incrementado en una fermentación. Tales microorganismos representan células hospederas preferidas en el sentido de la invención.

55 En otra forma de realización de la invención, la célula hospedera se caracteriza por lo tanto porque es una bacteria, preferiblemente una seleccionada del grupo de los géneros de *Escherichia*, *Klebsiella*, *Bacillus*, *Staphylococcus*, *Corynebacterium*, *Arthrobacter*, *Streptomyces*, *Stenotrophomonas* y *Pseudomonas*; más preferiblemente una seleccionada del grupo de *Escherichia coli*, *Klebsiella planticola*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus amyloliquefaciens*, *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. Se prefieren muy particularmente Bacillus licheniformis.

5 La célula hospedera también puede ser una célula eucariota que se caracteriza porque tiene un núcleo celular. Otro objeto de la invención representa por lo tanto una célula hospedera que se caracteriza porque tiene un núcleo celular.

10 A diferencia de las células procariotas, las células eucariotas están en capacidad de modificar la proteína formada después de la traducción. Ejemplos de estas son hongos tales como Saccharomyces o Kluyveromyces. Esto puede ser particularmente ventajoso, por ejemplo, si las proteínas deben experimentar modificaciones específicas en conexión con su síntesis, las cuales hacen posible sistemas de este tipo. Las modificaciones que los sistemas eucariotas realizan particularmente en conexión con la síntesis de proteínas incluyen por ejemplo el enlazamiento de compuestos de bajo peso molecular, como anclas de membrana u oligosacáridos. Dichas modificaciones de oligosacáridos pueden ser deseables, por ejemplo, para reducir la alergenicidad de una proteína expresada. Puede ser ventajosa una co-expresión con las enzimas formadas de manera natural por las células de este tipo, como por ejemplo celulasas. Además, los sistemas de expresión de hongos termofílicos pueden ser particularmente adecuados por ejemplo para la expresión de variantes resistentes a temperatura.

15 En el contexto de la invención, como productos que se forman durante la fermentación se consideran las proteínas que se codifican por la secuencia de ácido nucleico (b), principalmente aquellas como se han descrito previamente. En este caso se trata de proteasas y muy particularmente preferible de subtilisinas.

20 Las células hospederas pueden modificarse además con respecto a sus requerimientos a las condiciones de cultivo, pueden tener otros marcadores de selección o adicionales o expresar otras proteínas o adicionales. Principalmente también puede tratarse de tales células hospederas que expresan varias proteínas o enzimas. Son secretadas preferiblemente en el medio que rodea las células hospederas.

25 Las células hospederas de la invención se cultivan y se fermentan de una manera conocida como tal, por ejemplo en sistemas discontinuos o sistemas continuos. En el primer caso se inocula un medio de cultivo adecuado con las células hospederas y después de un período a determinarse experimentalmente se cosecha el producto del medio. Las fermentaciones continuas se caracterizan por lograr un estado de equilibrio en el cual durante un período comparativamente largo las células mueren parcialmente pero también crecen nuevamente el producto puede ser retirado al mismo tiempo del medio.

30 De acuerdo con la invención las células hospederas se usan preferiblemente para preparar proteínas codificadas por la secuencia de ácido nucleico (b). Otro objeto de la invención es, por lo tanto, un procedimiento para producir una proteasa, el cual comprende

a) cultivar una célula hospedera según la invención

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

35 Este objeto de la invención comprende preferiblemente procedimientos de fermentación. Los procedimientos de fermentación son conocidos como tales del estado de la técnica y representan la propia etapa de producción industrial, seguida por lo regular por un procedimiento de purificación adecuado del producto preparado, por ejemplo la proteína. Todos los procedimientos de fermentación que se basan en un procedimiento correspondiente para la producción de una proteasa representan formas de realización de este objeto de la invención.

40 A este respecto, las respectivas condiciones óptimas para los procedimientos de preparación, principalmente las condiciones de cultivo óptimas para las células hospederas usadas, deben determinarse experimentalmente de acuerdo con el conocimiento de un especialista, por ejemplo respecto del volumen de fermentación y/o composición del medio y/o suministro de oxígeno y/o velocidad de agitación.

45 Los procedimientos de fermentación caracterizados porque la fermentación se lleva a cabo mediante una estrategia de suministro continua se toman principalmente en consideración. En este caso, los componentes del medio que son consumidos por el cultivo en curso se alimentan continuamente; esto se conoce como una estrategia de alimentación continua. Como resultado, pueden lograrse incrementos considerables tanto en la densidad de células como también en la masa celular o en masa seca y/o ante todo en la actividad de la proteína de interés, preferiblemente un enzima.

50 Además, la fermentación también puede configurarse de tal manera que los productos metabólicos no deseados sean filtrados o neutralizados por la adición de un regulador de pH o los contraiones apropiados respectivamente.

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 hospedera, es decir un tratamiento del producto a

partir de la masa celular (masa seca). De acuerdo con la invención, los marcadores de secreción adecuados a este respecto se proporcionan con los péptidos de señal.

5 Todos los hechos explicados previamente pueden combinarse para formar procedimientos de preparación de proteínas. A este respecto es concebible una gran cantidad de posibilidades de combinaciones de etapas del procedimiento. El procedimiento óptimo tiene que determinarse para cada caso individual específico.

10 Otro objeto de la invención es el uso de un vector de expresión según la invención o de una célula hospedera según la invención para producción de una proteasa todos los hechos, objetos y formas de realización que ya se han descrito previamente también son aplicables a este objeto de la invención. Por lo tanto, se hace referencia expresa en este sitio a la divulgación del sitio correspondiente con la indicación de que dicha divulgación también se aplica a los usos según la invención (uso del vector o de la célula hospedera).

Ejemplos:

15 Todas las etapas de trabajo de biología molecular siguen procedimientos estándares tal como se indica, por ejemplo, en el manual de Fritsch, Sambrook y Maniatis "Molecular cloning: a laboratory manual", Cold Spring Harbour Laboratory Press, New York, 1989, o en obras especializadas comparables. Se usaron enzimas y kits de acuerdo con las indicaciones de los respectivos fabricantes.

Ejemplo 1: Preparación de lectores de expresión según la invención

20 El plásmido pBSMuL3 (Brockmeier et al., 2006) fue acortado mediante digestión de restricción de SacI y religación subsiguiente alrededor de la porción de E.coli. El plásmido resultante, pBSMuL5 (compárese la figura 1) sirvió como vector para la clonación de las proteasas incluyendo propéptido 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 plásmido resultante sirvieron como vectores para clonar péptidos de señales en los sitios de restricción HindIII y EcoRI. El fragmento de ADN del péptido de señal de control SubC (B. licheniformis, NCBI ("National Center for Biotechnology Information") número de acceso ("Accession Number"): X91260.1) como sello distintivo fue amplificado con ayuda de los cebadores según SEQ ID NO. 15 y SEQ ID NO. 16 y se tronaron 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 una proteína con el péptido de 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 sello distintivo. El fragmento de ADN del péptido de señal según SEQ ID NO. 2 fue amplificado con ayuda de los cebadores según SEQ ID NO. 19 y SEQ ID NO. 20, el fragmento de ADN del péptido de 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 de señal según SEQ ID NO. 6 con los cebadores según SEQ ID NO. 21 y SEQ ID NO. 22. Mientras tanto los fragmentos de ADN de los péptidos de señal según SEQ ID NO. 2 y 4 fueron coronados respectivamente es en el vector que codifica una proteasa según SEQ ID NO. 8 (plásmidos 3 y 4), el fragmento de ADN del péptido de señal según SEQ ID NO. 6 fue insertado en el vector que codifica una proteasa según SEQ ID NO. 9 (plásmido 5). Asociada con la clonación, una secuencia de 9 nucleótidos que codifican la sucesión de los aminoácidos AEF (compárese la figura 1) fue introducida entre la secuencia de ADN del péptido de señal respectivo y la secuencia de ADN del propéptido de la proteasa respectiva. Esta, así llamada, secuencia de conexión contiene la secuencia de reconocimiento de la endonucleasa de restricción EcoRI.

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

Tabla 1:

Denominación	Secuencia de nucleótidos (en orientación 5'→ 3'; el subrayado indica los sitios de restricción)	Sitio de restricción
SEQ ID NO. 11	ATATGAATTCGCTGAGGAAGCAAAGAAAA	EcoRI
SEQ ID NO. 12	ATATGGATCCTTAGCGTGTGGCCGCTTCTGC	BamHI
SEQ ID NO. 13	ATATGAATTCGCTGAGGAAGCAAAGAAAA	EcoRI
SEQ ID NO. 14	ATATGGATCCTTAGCGCGTTGCTGCATCTGC	BamHI
SEQ ID NO. 15	ATATAAGCTTAAGGAGGATATTATGATGAGGAAAAGAGT 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

Denominación	Secuencia de nucleótidos (en orientación 5'→ 3'; el subrayado indica los sitios de restricción)	Sitio de restricción
SEQ ID NO. 21	ATATA <u>AAGCTT</u> AAGGAGGATATTATGACATTGACTAACTG AAA	HindIII
SEQ ID NO. 22	ATATGAATTCAGCGGCAAGTGCCTGACTGGAAAA	EcoRIII

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 toda la noche (16 h). Para la determinación cuantitativa de la eficiencia de secreción, fueron transferidas directamente las colonias individuales desde las placas de agar a Deepwell-MTP (placas de microtitulación; 96 cavidades cada una de 1 ml de medio LB selectivo). En tal caso cada colonia individual fue transferida en paralelo en al menos dos cavidades con el fin de obtener una determinación doble o triple mediante el cultivo múltiple del clon respectivo. Para la inoculación de las Deepwell-MTP se usaron clones exclusivamente que hubieran sido incubados a 37 °C por toda la noche. Después del cultivo durante 20 h a 37 °C en el agitador de placas de microtitulación (Timix 5, de la compañía Edmy-Bühler, Hechingen) todos los clones fueron replicados en placas de agar de LB y a continuación las células fueron sedimentadas por medio de centrifugación (4000 rpm, 20 min, 4°C). Toda las siguientes etapas de aplicación con pipeta se llevaron a cabo usando pipetas de canales múltiples (Eppendorf, Hamburgo) y se usó el modo "reverse pipetting"; no se dosificó con pipeta ningún volumen inferior a los 15 µl. En cada caso el volumen más pequeño fue cargado inicialmente en la MTP y los volúmenes más grandes se adicionaron a la misma; la MTP en cada etapa de dilución fue mezclada por 10 segundos en el espectrofotómetro "Spektramax 250" (Molecular Devices, Sunnyvale, Estados Unidos). Para la generación 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 en F, transparente, de la compañía Greiner Bio-One, Frickenhausen).

Continuación la actividad proteolítica en los sobrenadantes de cultivo o diluciones fue determinada mediante la liberación del cromóforo para-nitroanilina (pNA) del sustrato suc-L-Ala-L-Ala-L-Pro-L-Phe-p-nitroanilida (suc-AAPF-pNA). La proteasa disocia el sustrato y libera pNA. La liberación de pNA provoca un incremento en la absorbancia a 410 nm y su cambio en el tiempo es una medida de la actividad enzimática (compárese Del Mar et al., Anal.Biochem., 99: 316-320, 1979).

Para la determinación de la eficiencia de secreción de las diversas cepas, se cultivó conjuntamente un constructo de control interno (plásmido 1 o plásmidos 2) en cada cultivo de MTP. La actividad proteolítica de la cepa con el constructo de control, determinada en el sobrenadante de cultivo, fue definida 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 (compárese la figura 2).

La cepa con el plásmido 5 según la invención, en comparación con el control que contenía el plásmidos 2, alcanzó una actividad de proteasa incrementada en 44 % +/- 10 (compárese la figura 3).

Descripción de las figuras

Figura 1: Esquema de la estrategia de clonación en el vector de expresión de *Bacillus* pBSMu15 (modificado según Brockmeier et al., 2006). (A) los fragmentos de ADN de los péptidos de señal fueron amplificados en el N-terminal con un sitio de restricción HindIII, un sitio estandarizado de enlace a ribosomas (RBS), seguido por una región espaciadora y el cordón de inicio estandarizado para metionina. Un acoplador que tenía una alanina en la posición "+1" y el sitio de restricción EcoRI fue insertado entre el partido de señal y el N-terminal de la proteasa que va a ser secretada. (B) El vector de *Bacillus* pBSMu15 con el promotor Hpall, el respectivo objetivo de secreción (coronado mediante EcoRI y BamHI) y 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 de señales diferentes en pBSMu15. La actividad proteolítico del plásmido de constructo 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 distintos péptidos de señales en pBSMu15. La actividad proteolítico del plásmido de constructo 2 fue definida como 100 % (control) los valores se determinaron al menos en dos cultivos independientes uno de otro. Las barras de error indican la desviación estándar.

Listado de secuencias

<110> Henkel AG & Co. KGaA

<120> Vectores de expresión para secreción mejorada de proteína

<130> H 09309 PCT

<150> DE 102011118032.3

5 <151> 2011-05-31

<160> 22

<170> PatentIn version 3.5

<210> 1

<211> 126

10 <212> ADN

<213> Bacillus subtilis

<220>

<221> CDS

<222> (1)..(126)

15 <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

<211> 42

20 <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
    
```

<210> 3

ES 2 606 553 T3

<211> 81

<212> ADN

<213> Bacillus licheniformis

<220>

5 <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

10 <211> 27

<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
    
```

15 <210> 5

<211> 84

<212> ADN

<213> Bacillus subtilis

<220>

20 <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

ES 2 606 553 T3

<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

5

<210> 7

<211> 269

<212> PRT

<213> Bacillus lentus

10 <400> 7

ES 2 606 553 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
 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

5 <212> PRT

ES 2 606 553 T3

<213> Bacillus lentus

<400> 8

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

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

5 <210> 9

ES 2 606 553 T3

<211> 269

<212> PRT

<213> *Bacillus gibsonii*

<400> 9

ES 2 606 553 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
 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

ES 2 606 553 T3

<212> PRT

<213> *Bacillus lentus*

<400> 10

ES 2 606 553 T3

Ala Gln Ser Val 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 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> ADN
<213> Artificial
<220>
<223> cebador corriente arriba
5 <400> 11
atatgaattc gctgaggaag caaaagaaaa 30
<210> 12
<211> 31
<212> ADN
10 <213> Artificial
<220>
<223> cebador corriente abajo
<400> 12
atatggatcc ttagcgtgtt gccgcttctg c 31
15 <210> 13
<211> 30
<212> ADN
<213> Artificial
<220>
20 <223> cebador corriente arriba
<400> 13
atatgaattc gctgaggaag caaaagaaaa 30
<210> 14
<211> 31
25 <212> ADN
<213> Artificial
<220>
<223> cebador corriente abajo
<400> 14
30 atatggatcc ttagcgcgtt gctgcatctg c 31
<210> 15
<211> 43
<212> ADN
<213> Artificial

<220>

<223> cebador corriente arriba

<400> 15

atataagctt aaggaggata ttatgatgag gaaaaagagt ttt 43

5 <210> 16

<211> 34

<212> ADN

<213> Artificial

<220>

10 <223> cebador corriente abajo

<400> 16

atatgaattc agctgcagaa gcggaatcgc tga 34

<210> 17

<211> 43

15 <212> ADN

<213> Artificial

<220>

<223> cebador corriente arriba

<400> 17

20 atataagctt aaggaggata ttatgaaaaa actattcaaa acc 43

<210> 18

<211> 34

<212> ADN

<213> Artificial

25 <220>

<223> cebador corriente abajo

<400> 18

atatgaattc agcagccgcc gcagattgtg agaa 34

<210> 19

30 <211> 43

<212> ADN

<213> Artificial

<220>

<223> cebador corriente arriba

ES 2 606 553 T3

<400> 19

atataagctt aaggaggata ttatggcgaa accactatca aaa 43

<210> 20

<211> 34

5 <212> ADN

<213> Artificial

<220>

<223> cebador corriente abajo

<400> 20

10 atatgaattc agcagcgtct gccgcgggta aacc 34

<210> 21

<211> 43

<212> ADN

<213> Artificial

15 <220>

<223> cebador corriente arriba

<400> 21

atataagctt aaggaggata ttatgacatt gactaaactg aaa 43

<210> 22

20 <211> 34

<212> ADN

<213> Artificial

<220>

<223> cebador corriente abajo

25 <400> 22

atatgaattc agcggcaagt gcctgactgg aaaa 34

REIVINDICACIONES

1. Vector de expresión que comprende
 - a) una secuencia de promotor y
 - b) una secuencia de ácido nucleico que codifica una proteína, comprendiendo la proteína un péptido de señal y otra secuencia de aminoácidos y el péptido de señal comprende una secuencia de aminoácidos que es idéntica en al menos el 80 % a la secuencia de aminoácidos indicada en la SEQ ID NO. 2, 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 de señal codificado por la secuencia de ácido nucleico b) presenta una secuencia de aminoácidos según la SEQ ID NO. 2.
3. Vector de expresión según una de las reivindicaciones 1 o 2, caracterizado porque el péptido de señal se encuentra dispuesto en la 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 bis 3, caracterizado porque la proteína codificada por la secuencia de ácido nucleico b) comprende además una secuencia de enlazamiento que se encuentra dispuesta entre el péptido de señal y la otra secuencia de aminoácidos de la proteína, presentando la secuencia de enlazamiento principalmente 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 80 % a la SEQ ID NO. 7, o
 - es idéntica en al menos 80 % a la SEQ ID NO. 8, o
 - es idéntica en al menos 80 % a la SEQ ID NO. 9, o
 - es idéntica en al menos 80 % a la SEQ ID NO. 10 y en la posición 99 en el conteo según la SEQ ID NO. 10 presenta los aminoácidos ácido glutámico (E) o ácido aspártico (D), o
 - es idéntica en al menos 80 % a la SEQ ID NO. 10 y en la posición 99 en el conteo según la SEQ ID NO. 10 presenta los aminoácidos ácido glutámico (E) o ácido aspártico (D) y además en el conteo según la 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 (211 D, 211 E o 211 G),
 - (i) combinaciones de los aminoácidos (a) a (h).
6. Célula hospedera no humana que contiene un vector de expresión según una de las reivindicaciones 1 a 5.
7. Célula hospedera según la reivindicación 6, caracterizada porque es una bacteria, preferiblemente una que se selecciona del grupo de los géneros *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 amyloliquefaciens*, *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 producción de una proteasa que comprende las etapas del procedimiento

(a) cultivar una célula hospedera según una de las reivindicaciones 6 o 7

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

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

Figura 1

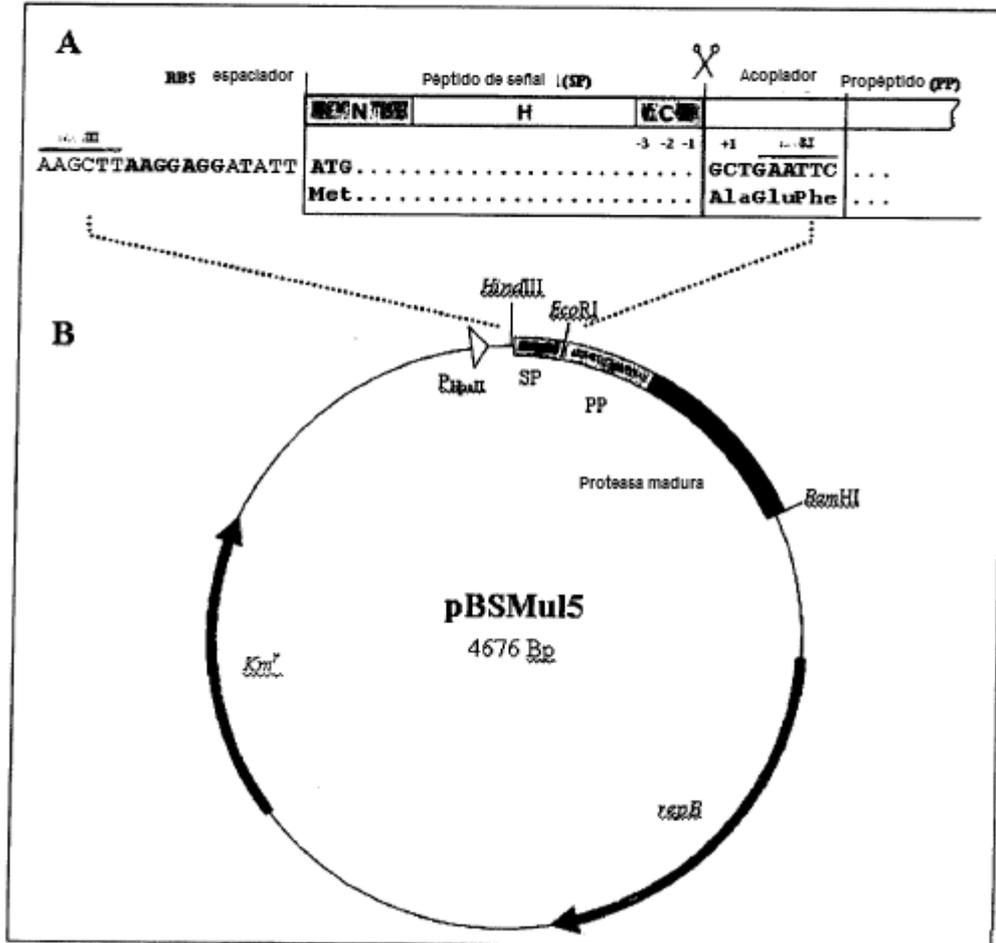


Figura 2:

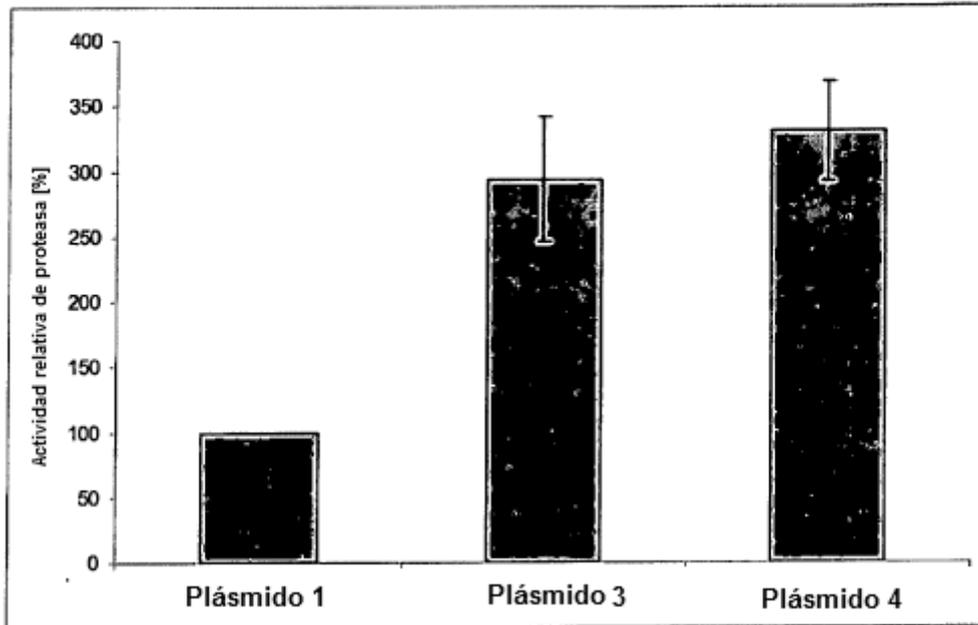


Figura 3:

