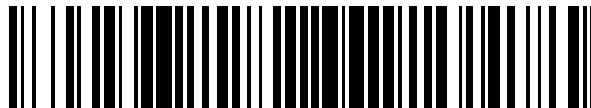


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 744 898**

51 Int. Cl.:

C12N 9/50 (2006.01)

A23L 2/84 (2006.01)

A23J 3/34 (2006.01)

C12N 9/62 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.05.2015 PCT/EP2015/061026**

87 Fecha y número de publicación internacional: **26.11.2015 WO15177171**

96 Fecha de presentación y número de la solicitud europea: **19.05.2015 E 15724252 (0)**

97 Fecha y número de publicación de la concesión europea: **26.06.2019 EP 3146043**

54 Título: **Endoproteasa específica de prolina**

30 Prioridad:

19.05.2014 EP 14168866

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

26.02.2020

73 Titular/es:

**DSM IP ASSETS B.V. (100.0%)
Het Overloon 1
6411 TE Heerlen, NL**

72 Inventor/es:

**LAAN, VAN DER, JAN METSKE;
VONDERVOORT, VAN DE, PETER JOZEF IDA;
CHRISTIS, CHANTAL;
SPAANS, MARTINE;
BRUINE-PAULUS, DE, ANGELA y
MUTSAERS, JOHANNA HENRICA GERDINA
MARIA**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 744 898 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

Endoproteasa específica de prolina

5 La presente invención se refiere a un polipéptido que tiene actividad de endoproteasa específica de prolina, a una composición que comprende el polipéptido, a un ácido nucleico que codifica una endoproteasa específica de prolina, a un vector de expresión que comprende el ácido nucleico que codifica una endoproteasa específica de prolina, a una célula hospedante recombinante, a un método para preparar endoproteasa específica de prolina, y a un procedimiento para preparar un producto alimentario o de pienso en el que se usa la endoproteasa específica de prolina.

Antecedentes

10 Las endoproteasas específicas de prolina son enzimas que hidrolizan una proteína o péptido en una posición en la que hay una prolina en la proteína o péptido.

Una endoproteasa específica de prolina puede derivar, por ejemplo, de *Aspergillus niger* o *Penicillium chrysogenum*, tal como se describe en los documentos WO2002/046381 y WO2009/144269, respectivamente.

15 Se conocen otras endoproteasas específicas de prolina desde el documento WO2012/174127. El documento WO2012/174127 describe proteasa específica de prolina procedente de *Botryotinia fuckeliana*, *Aspergillus clavatus*, *Sclerotinia sclerotium*, *Mycosphaerella graminicola*, *Neurospora crasse*, *Talaromyces stipitatus* y *Gibberella zeae*.

20 La endoproteasa específica de prolina se puede usar en varias aplicaciones, por ejemplo en la degradación de gluten (véase, por ejemplo, el documento WO2005/027953 o WO2003/068170). El gluten es la fracción proteica insoluble de cereales como el trigo, centeno, avena o cebada. El gluten es una mezcla compleja de moléculas de glutenina y prolamina que se piensa que provocan efectos tóxicos, por ejemplo en pacientes que sufren celiaquía. Se considera que la celiaquía o enfermedad celíaca es una enfermedad autoinmune. Los pacientes que sufren celiaquía necesitan seguir una dieta estricta libre de gluten, que es muy difícil de seguir debido a que el gluten se usa tan ampliamente. El uso de endoproteasa específica de prolina como medicamento o suplemento dietético puede aliviar la necesidad de una dieta estricta libre de gluten (documento WO2003/068170).

25 Las endoproteasas específicas de prolina también se usan para reducir la turbidez en la cerveza, en la que se puede añadir la proteasa específica de prolina durante varias etapas de un procedimiento de producción de la cerveza, por ejemplo como se describe en el documento WO 2002/046381, o WO2007101888A2.

El objetivo de la presente invención es una endoproteasa alternativa específica de prolina con características mejoradas.

30 Sumario

En un aspecto, la presente descripción se refiere a un polipéptido aislado que tiene actividad de endoproteasa específica de prolina seleccionado del grupo que consiste en

i. un polipéptido que comprende una secuencia de polipéptido maduro de SEQ ID NO: 2;

35 ii. un polipéptido que tiene al menos 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad con una secuencia de polipéptido maduro de SEQ ID NO: 2;

iii. un polipéptido codificado por un ácido nucleico que se hibrida bajo restricción media, preferiblemente en condiciones de restricción elevadas, a la hebra complementaria de una secuencia codificante de polipéptido maduro de SEQ ID NO: 1;

40 iv. un polipéptido codificado por un ácido nucleico que tiene al menos 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con una secuencia codificante de polipéptido maduro de SEQ ID NO: 1;

En otro aspecto, la presente invención se refiere a una composición que comprende un polipéptido como se describe aquí.

45 En un aspecto, la presente descripción se refiere a un ácido nucleico que tiene al menos 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con una secuencia codificante de polipéptido maduro de SEQ ID NO: 1.

La presente invención también se refiere a un vector de expresión que comprende un polinucleótido que codifica un polipéptido según la presente invención.

50 En otro aspecto, la presente invención se refiere a una célula hospedante que comprende una secuencia polinucleotídica, o un vector de expresión como se describe aquí.

En todavía otro aspecto, la presente invención se refiere a un método para la preparación de un polipéptido, que comprende cultivar una célula hospedante como se describe aquí en condiciones que permiten la expresión del polipéptido, y preparar el polipéptido.

- 5 En otro aspecto, la presente invención se refiere a un procedimiento para la preparación de un producto alimentario o de pienso, que comprende incubar una forma intermedia del producto alimentario o de pienso con un polipéptido como se describe aquí, y preparar el producto alimentario.

La presente descripción también se refiere a un producto alimentario o de pienso obtenible mediante un procedimiento como se describe aquí.

Definiciones

- 10 La expresión “productos horneados” se define aquí como cualquier producto preparado a partir de una masa espesa o de una masa licuada. El producto puede tener un carácter blando o un carácter crujiente, y puede ser de un tipo blanco, claro u oscuro. Los productos horneados incluyen, pero no se limitan a, pan, tal como, por ejemplo, pan blanco, integral, o de centeno, pan de tipo baguette francesa, productos de masa laminados, tales como pasteles (daneses), croissants u hojaldre, pan de pita, tortillas, tacos, bizcochos, tortitas, galletitas, galletas, rosquillas, roscas
15 de pan, tapas de masa, panecillos, pan al vapor, y pan crujiente. Los tipos de productos horneados, los métodos para caracterizarlos y producirlos son conocidos por los expertos en la técnica; véase, por ejemplo, “Baking Science and Technology”, por E.J. Pyle, L.A. Gorton, 2008, (2 volúmenes) Sosland Publishing Company, Kansas, USA, o “Baked Products: Science, Technology and Practice” por S.P. Cauvain, L.S. Young, 2006, Blackwell Publishing Ltd, Oxford, UK.

- 20 La expresión “hebra complementaria” se puede usar de forma intercambiable con el término “complemento”. El complemento de una hebra de ácido nucleico puede ser el complemento de una hebra codificante o el complemento de una hebra no codificante. Cuando se hace referencia a ácidos nucleicos bicatenarios, el complemento de un ácido nucleico que codifica un polipéptido se refiere a la hebra complementaria de la hebra que codifica la secuencia de aminoácidos o a cualquier molécula de ácido nucleico que la contenga.

- 25 La expresión “secuencia de control” se puede usar de forma intercambiable con la expresión “secuencia de ácido nucleico que regula la expresión”. La expresión, como se usa aquí, se refiere a secuencias de ácidos nucleicos necesarias para y/o que afectan a la expresión de una secuencia codificante enlazada operablemente en un organismo hospedante particular o *in vitro*. Cuando dos secuencias de ácidos nucleicos están enlazadas operablemente, habitualmente estarán en la misma orientación y también en el mismo marco de lectura.
30 Habitualmente serán esencialmente contiguas, aunque esto puede no ser necesario. Las secuencias de ácidos nucleicos que regulan la expresión, tales como, entre otras, las secuencias apropiadas de iniciación de la transcripción, de terminación, promotoras, líder, de péptido señal, de propéptido, de prepropéptido, o potenciadoras, la secuencia de Shine-Dalgarno, secuencias represoras o activadoras, las señales de procesamiento del ARN eficientes, tales como señales de ajuste y de poliadenilación, las secuencias que estabilizan ARNm citoplásmico, las
35 secuencias que potencian la eficiencia de la traducción (por ejemplo, sitios de unión al ribosoma), las secuencias que potencian la estabilidad de la proteína, y, cuando se desee, las secuencias que potencian la secreción de la proteína, pueden ser cualquier secuencia de ácido nucleico que muestre actividad en el organismo hospedante de elección, y pueden derivar de genes que codifican proteínas, que bien son endógenas o heterólogas a una célula hospedante. Cada secuencia de control puede ser nativa o extraña a la secuencia de ácido nucleico que codifica el polipéptido. Cuando se desea, la secuencia de control se puede proporcionar con enlazadores con el fin de
40 introducir sitios de restricción específicos que facilitan la ligación de las secuencias de control con la región codificante de la secuencia de ácido nucleico que codifica un polipéptido. Las secuencias de control se pueden optimizar para sus fines específicos.

- 45 Un “producto lácteo” se refiere a cualquier tipo de producto a base de leche destinado a ser usado como alimento, pienso, o bebida, incluyendo, pero sin limitarse a, queso, leche, leche desnatada, leche ácida, leche de mantequilla, leche condensada, untas, margarinas, yogur, helado, polvo de leche, mantequilla, EMC (Queso Modificado con Enzimas), dulce de leche, blanqueante del café, leche en polvo para el café, nata, suero de mantequilla, análogo lácteo, etc. El queso puede ser cualquier tipo de queso, por ejemplo queso fresco, queso duro, requesón, queso crema, queso de moho blanco, queso de moho azul, y queso fundido. Los ejemplos de queso fresco son Ricotta, queso crema, Neufchatel or queso Cottage. Los ejemplos de queso duro son Chester, Danbo, Manchego, Saint
50 Paulin, Cheddar, Monterey, Colby, Edam, Gouda, Muenster, tipo suizo, Gruyere, Emmenthaler, Parmigiano Reggiano, Grana Padano, Parmesano, Pecorino, Provolone, y Romano. Los ejemplos de requesón tales como queso feta, queso Quotija, queso hilado, tal como Mozzarella, y queso fresco. Los ejemplos de queso crema son queso Philadelphia. Los ejemplos de queso de moho blanco son queso Brie y Camembert. Los ejemplos de queso
55 de moho azul son Gorgonzola y queso azul danés.

Como se usa aquí, el término “endógena” se refiere a una secuencia de ácido nucleico o de aminoácidos que aparece de forma natural en un hospedante.

Las endopeptidasas, endoproteinasas o endoproteasas son enzimas que son capaces de romper enlaces peptídicos

- de aminoácidos no terminales (es decir, dentro de la proteína), en contraste con las exopeptidasas, que rompen enlaces peptídicos del término amino o del término carboxilo. Las endopeptidasas no tienden a romper péptidos en monómeros, sino que dan como resultado fragmentos peptídicos relativamente grandes. La generación específica de fragmentos relativamente grandes es muy preferida en muchas aplicaciones relacionadas con los alimentos y los piensos. Un caso particular de endopeptidasa es la oligopeptidasa, cuyos sustratos son oligopéptidos en lugar de proteínas.
- El término “expresión” incluye cualquier etapa implicada en la producción del polipéptido, incluyendo, pero sin limitarse a, transcripción, modificación post-transcripcional, traducción, modificación post-traducciona, y secreción.
- Un vector de expresión comprende un polinucleótido que codifica un polipéptido, enlazado operablemente a las secuencias de control apropiadas (tales como un promotor, y señales de parada transcripcionales y traduccionales) para la expresión y/o traducción *in vitro*, o en la célula hospedante, del polinucleótido.
- El vector de expresión puede ser cualquier vector (por ejemplo, un plásmido o virus), que se puede someter convenientemente a procedimientos de ADN recombinante y pueden provocar la expresión del polinucleótido. La elección del vector dependerá típicamente de la compatibilidad del vector con la célula en la que se va a introducir el vector. Los vectores pueden ser plásmidos circulares lineales o cerrados. El vector puede ser un vector que se replica de forma autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo un plásmido, un elemento extracromosómico, un minicromosoma, o un cromosoma artificial. Como alternativa, el vector puede ser aquel que, cuando se introduce en la célula hospedante, se integra en el genoma y se replica junto con el cromosoma o cromosomas en los que se ha integrado. El vector de clonación integrante se puede integrar en un locus diana al azar o en uno predeterminado en los cromosomas de la célula hospedante. El sistema de vector puede ser un único vector o plásmido, o dos o más vectores o plásmidos, que contienen juntos el ADN total a introducir en el genoma de la célula hospedante, o un transposón.
- Una célula hospedante como se define aquí es un organismo adecuado para la manipulación genética, y aquel que se puede cultivar a densidades celulares útiles para la producción industrial de un producto diana, tal como un polipéptido según la presente invención. Una célula hospedante puede ser una célula hospedante encontrada en la naturaleza, o una célula hospedante derivada de una célula hospedante progenitora tras la manipulación genética o mutagénesis clásica. Ventajosamente, una célula hospedante es una célula hospedante recombinante.
- Una célula hospedante puede ser una célula hospedante procariota, arqueabacteriana, o eucariota. Una célula hospedante procariota puede ser, pero no se limita a, una célula hospedante bacteriana. Una célula hospedante eucariota puede ser, pero no se limita a, una célula hospedante de una levadura, un hongo, una ameba, una alga, una planta, un animal, o un insecto.
- El término “heteróloga”, como se usa aquí, se refiere a secuencias de ácidos nucleicos o de aminoácidos que no aparecen de forma natural en una célula hospedante. En otras palabras, la secuencia de ácido nucleico o de aminoácidos no es idéntica a la encontrada de forma natural en la célula hospedante.
- El término “hibridación” significa el apareamiento de hebras sustancialmente complementarias de compuestos oligoméricos, tales como compuestos de ácido nucleico.
- La hibridación se puede llevar a cabo en condiciones de restricción baja, media o alta. Las condiciones de hibridación de baja restricción comprenden hibridar en 6X cloruro de sodio/citrato de sodio (SSC) a alrededor de 45°C, seguido de dos lavados en 0,2X SSC, 0,1% de SDS al menos a 50°C (la temperatura de los lavados se puede incrementar hasta 55°C para condiciones de baja restricción). Las condiciones de hibridación de restricción media comprenden hibridar en 6X SSC a alrededor de 45°C, seguido de uno o más lavados en 0,2X SSC, 0,1% de SDS a 60°C, y las condiciones de hibridación de alta restricción comprenden hibridar en 6X SSC a alrededor de 45°C, seguido de uno o más lavados en 0,2X SSC, 0,1% de SDS a 65°C.
- Un ácido nucleico o secuencia polinucleotídica se define aquí como un polímero de nucleótido que comprende al menos 5 unidades nucleotídicas o de ácido nucleico. Un nucleótido o ácido nucleico se refiere a ARN y ADN. Las expresiones “ácido nucleico” y “secuencia polinucleotídica” se usan aquí de forma intercambiable.
- Un “péptido” se refiere a una cadena corta de restos de aminoácidos enlazados mediante un enlace peptídico (amídico). El péptido más corto, un dipéptido, consiste en 2 aminoácidos unidos mediante un único enlace peptídico.
- El término “polipéptido” se refiere a una molécula que comprende restos de aminoácidos enlazados mediante enlaces peptídicos, y que contiene más de cinco restos de aminoácidos. El término “proteína”, como se usa aquí, es sinónimo del término “polipéptido”, y también se puede referir a dos o más polipéptidos. De este modo, los términos “proteína” y “polipéptido” se pueden usar de forma intercambiable. Los polipéptidos se pueden modificar opcionalmente (por ejemplo, glicosilar, fosforilar, acilar, farnesilar, prenilar, sulfonar, y similar) para añadir funcionalidad. Los polipéptidos que muestran actividad en presencia de un sustrato específico en ciertas condiciones se pueden denominar como enzimas. Se entenderá que, como resultado de la degeneración del código genético, se puede producir una multitud de secuencias nucleotídicas que codifican un polipéptido dado.

Un "fragmento de ácido nucleico aislado" es un fragmento de ácido nucleico que no aparece de forma natural como un fragmento y no se encontraría en el estado natural.

La expresión "polipéptido aislado", como se usa aquí, significa un polipéptido que se elimina de al menos un componente, por ejemplo otro material polipeptídico, con el que está asociado de forma natural. El polipéptido aislado puede estar libre de cualesquiera otras impurezas. El polipéptido aislado puede ser al menos 50% puro, por ejemplo al menos 60% puro, al menos 70% puro, al menos 75% puro, al menos 80% puro, al menos 85% puro, al menos 80% puro, al menos 90% puro, o al menos 95% puro, 96%, 97%, 98%, 99%, 99,5%, 99,9%, según se determina mediante SDS-PAGE o cualquier otro método analítico adecuado para este fin y conocido por la persona experta en la técnica. Un polipéptido aislado puede ser producido por una célula hospedante recombinante.

Un "polipéptido maduro" se define aquí como un polipéptido en su forma final, y se obtiene tras la traducción de un ARNm en polipéptido y las modificaciones post-traduccionales de dicho polipéptido. Las modificaciones post-traduccionales incluyen el procesamiento N-terminal, el truncamiento C-terminal, glicosilación, fosforilación, y la eliminación de secuencias líder tales como péptidos señal, propéptidos y/o prepropéptidos mediante escisión.

Una "secuencia codificante de polipéptido maduro" significa un polinucleótido que codifica un polipéptido maduro.

La expresión "constructo de ácido nucleico" se denomina aquí como una molécula de ácido nucleico, ya sea mono- o bicatenaria, que se aísla de un gen de origen natural, o que se ha modificado para que contenga segmentos de ácido nucleico que se combinan yuxtaponen de una manera que de otro modo no existiría en la naturaleza. La expresión constructo de ácido nucleico es sinónima del término "casete de expresión" cuando el constructo de ácido nucleico contiene todas las secuencias de control requeridas para la expresión de una secuencia codificante, en el que dichas secuencias de control están enlazadas operablemente a dicha secuencia codificante.

Una "endoproteasa específica de prolina" es una proteasa que hidroliza una proteína o péptido en una posición en la que la proteína o péptido contiene un resto de prolina. Una endoproteasa específica de prolina puede tener actividad de endoproteasa específica de prolina y/o de oligopeptidasa específica de prolina (EC3.4.21.26). Una endoproteasa específica de prolina es preferiblemente una enzima que hidroliza un enlace peptídico en el extremo carboxi terminal de restos de prolina, que da como resultado un fragmento peptídico y/o polipeptídico con una prolina C-terminal.

El término "promotor" se define aquí como una secuencia de ADN que se une a ARN polimerasa y dirige la polimerasa al sitio correcto de comienzo transcripcional en dirección 3' de una secuencia de ácido nucleico para iniciar la transcripción.

El término "recombinante", cuando se usa en referencia a una célula, ácido nucleico, proteína o vector, indica que la célula, ácido nucleico, proteína o vector se ha modificado mediante la introducción de un ácido nucleico o proteína heterólogos, o mediante la alteración de un ácido nucleico o proteína nativos, o que la célula deriva de una célula así modificada. De este modo, por ejemplo, las células recombinantes expresan genes que no se encuentran en la forma nativa (no recombinante) de la célula, o expresan genes nativos que de otro modo se expresan anormalmente, están subexpresados, o no se expresan en absoluto. El término "recombinante" es sinónimo de "genéticamente modificado" y "transgénico".

Identidad de secuencia. Identidad de secuencia u homología de secuencia se usan aquí de forma intercambiable. Para los fines de esta invención, se define aquí que, a fin de determinar el porcentaje de homología de secuencia o identidad de secuencia de dos secuencias de aminoácidos o de dos secuencias de ácidos nucleicos, las secuencias se alinean con fines de comparación óptima. A fin de optimizar el alineamiento entre las dos secuencias, se pueden introducir espacios vacíos en cualquiera de las dos secuencias que se comparan. Tal alineamiento se puede llevar a cabo a lo largo de toda la longitud de las dos secuencias que se comparan. Como alternativa, el alineamiento se puede llevar a cabo a lo largo de una longitud más corta, por ejemplo a lo largo de alrededor de 20, alrededor de 50, alrededor de 100 o más ácidos nucleicos/bases o aminoácidos. La identidad de secuencia es el porcentaje de coincidencias idénticas entre las dos secuencias a lo largo de la región alineada dada a conocer. El porcentaje de identidad de secuencia entre dos secuencias de aminoácidos o entre dos secuencias nucleotídicas se puede determinar usando el algoritmo de Needleman y Wunsch para el alineamiento de dos secuencias. (Needleman, S. B. y Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). Tanto las secuencias de aminoácidos como las secuencias nucleotídicas se pueden alinear mediante el algoritmo. El algoritmo de Needleman-Wunsch se ha implementado en el programa de ordenador NEEDLE. Para los fines de esta invención, se usó el programa NEEDLE del paquete EMBOSS (versión 2.8.0 o superior, EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice, P. Longden, I. y Bleasby, A. Trends in Genetics 16, (6) p. 276-277, <http://emboss.bioinformatics.nl/>). Para las secuencias de proteínas, se usa EBLOSUM62 para la matriz de sustitución. Para la secuencia nucleotídica, se usa EDNAFULL. Los parámetros opcionales usados son una penalización de espacio vacío de 10, y una penalización de extensión del espacio vacío de 0,5. La persona experta apreciará que todos estos parámetros diferentes producirán resultados ligeramente diferentes, pero que el porcentaje de identidad global de dos secuencias no se altera significativamente cuando se usan algoritmos diferentes.

Tras el alineamiento mediante el programa NEEDLE como se describe anteriormente, el porcentaje de identidad de secuencia entre una secuencia de consulta y una secuencia de la invención se calcula como sigue: número de

posiciones correspondientes en el alineamiento que muestran un aminoácido idéntico o nucleótido idéntico en ambas secuencias, dividido entre la longitud total del alineamiento tras restar el número total de espacios vacíos en el alineamiento. La identidad como se define aquí se puede obtener a partir de NEEDLE usando la opción NOBRIEF, y está marcada en el resultado del programa como "identidad más larga".

5 El ácido nucleico y secuencias proteicas de la presente invención se pueden usar además como una "secuencia de consulta" para llevar a cabo una búsqueda frente a bases de datos públicas para, por ejemplo, identificar otros miembros de la familia o secuencias relacionadas. Tales búsquedas se pueden llevar a cabo usando los programas NBLAST y XBLAST (versión 2.0) de Altschul, et al. (1990) J. Mol. Biol. 215:403-10. las búsquedas nucleotídicas de BLAST se pueden llevar a cabo con el programa NBLAST, puntuación = 100, longitud de palabra = 12, para obtener
10 secuencias nucleotídicas homólogas a moléculas de ácidos nucleicos de la invención. Las búsquedas de proteínas de BLAST se pueden llevar a cabo con el programa XBLAST, puntuación = 50, longitud de palabra = 3, para obtener secuencias de aminoácidos homólogas a moléculas proteicas de la invención. Para obtener alineamientos con espacios vacíos, con fines comparativos, se puede utilizar BLAST con espacios vacíos, como se describe en Altschul et al., (1997) Nucleic Acids Res. 25(17): 3389-3402. Cuando se utilizan los programas BLAST y BLAST con espacios vacíos, se pueden usar los parámetros por defecto de los programas respectivos (por ejemplo, XBLAST y NBLAST). Véase la página de inicio del National Center for Biotechnology Information en
15 <http://www.ncbi.nlm.nih.gov/>.

La expresión "sustancialmente puro" con respecto a polipéptidos se refiere a una preparación polipeptídica que contiene como máximo 50% en peso de otro material polipeptídico. Los polipéptidos descritos aquí están preferiblemente en una forma sustancialmente pura. En particular, se prefiere que los polipéptidos descritos aquí estén en "forma esencialmente pura", es decir, que la preparación polipeptídica esté esencialmente libre de otro material polipeptídico. Opcionalmente, el polipéptido puede estar también esencialmente libre de material no polipeptídico, tal como ácidos nucleicos, lípidos, componentes de los medios, y similares. Aquí, la expresión "polipéptido sustancialmente puro" es sinónimo de las expresiones "polipéptido aislado" y "polipéptido en forma aislada".
20 La expresión "sustancialmente puro" con respecto a polinucleótidos se refiere a una preparación polinucleotídica que contiene como máximo 50% en peso de otro material polinucleotídico. Los polinucleótidos descritos aquí están preferiblemente en una forma sustancialmente pura. En particular, se prefiere que el polinucleótido descrito aquí esté en "forma esencialmente pura", es decir, que la preparación polinucleotídica esté esencialmente libre de otro material polinucleotídico. Opcionalmente, el polinucleótido también puede estar esencialmente libre de otro material no polinucleotídico, tales como polipéptidos, lípidos, componentes de los medios, y similares. Aquí, la expresión "polinucleótido sustancialmente puro" es sinónima de las expresiones "polinucleótido aislado" y "polinucleótido en forma aislada".
25

Una "molécula sintética", tal como un ácido nucleico sintético o un polipéptido sintético, se produce mediante una síntesis química o enzimática in vitro. Incluye, pero no se limita a, ácidos nucleicos variantes obtenidos con el uso óptimo de codones para organismos hospedantes de elección. Un ácido nucleico sintético se puede optimizar para el uso de codones, preferiblemente según los métodos descritos en los documentos WO2006/077258 y/o WO2008/000632; el documento WO2008/000632 aborda la optimización de pares de codones. La optimización de pares de codones es un método en el que las secuencias nucleotídicas que codifican un polipéptido que se ha modificado con respecto a su uso de codones, en particular los pares de codones que se usan, se optimizan para obtener una expresión mejorada de la secuencia nucleotídica que codifica el polipéptido, y/o para la producción mejorada del polipéptido codificado. Los pares de codones se definen como un conjunto de dos tripletes subsiguientes (codones) en una secuencia codificante. Los expertos en la técnica sabrán que es necesario adaptar el uso de codones dependiendo de la especie hospedante, dando posiblemente como resultado variantes con una desviación significativa de homología a partir de SEQ ID NO: 1, pero codificando todavía el polipéptido según la invención.
30
35
40
45

Como se usa aquí, los términos "variante", "derivado", "mutante" o "homólogo" se pueden usar de forma intercambiable. Se pueden referir a polipéptidos o a ácidos nucleicos. Las variantes incluyen sustituciones, inserciones, supresiones, truncamientos, transversiones, y/o inversiones, en una o más localizaciones con respecto a una secuencia de referencia. Las variantes se pueden obtener, por ejemplo, mediante mutagénesis por saturación de sitio, mutagénesis de barrido, mutagénesis de inserción, mutagénesis aleatoria, mutagénesis dirigida al sitio, y evolución dirigida, así como otros diversos enfoques de recombinación conocidos por una persona experta en la técnica. Los genes variantes de ácidos nucleicos se pueden sintetizar artificialmente mediante técnicas conocidas en la técnica.
50

FIGURAS

55 Figura 1: Vector pGBTOP-16 usado para la clonación del gen GLA. El vector pGBTOP-16 deriva del vector pGBTOP-12 descrito en el documento WO 2011/009700. Además de pGBTOP-12, contiene el gen ccdB de E. coli para la selección positiva para la presencia de un inserto entre los sitios de clonación EcoRI y Pacl. El sitio de restricción Pacl sustituye al sitio de restricción SnaBI presente en pGBTOP-12. Este vector se linealiza mediante digestión con NotI antes de la transformación.

60 Descripción detallada

En un aspecto, la presente invención se refiere a un polipéptido que tiene actividad de endoproteasa específica de prolina, seleccionado del grupo que consiste en

- i. un polipéptido que comprende una secuencia de polipéptido maduro de SEQ ID NO: 2;
- ii. un polipéptido que tiene al menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad con una secuencia de polipéptido maduro de SEQ ID NO: 2;
- iii. un polipéptido codificado por un ácido nucleico que tiene al menos 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con una secuencia codificante de polipéptido maduro de SEQ ID NO: 1;

en el que la secuencia de polipéptido maduro de SEQ ID NO: 2 comprende los aminoácidos 36 a 526, aminoácidos 37 a 526, aminoácidos 39 a 526, aminoácidos 40 a 526, o aminoácidos 41 a 526 de SEQ ID NO: 2, en el que la metionina en la posición 1 de SEQ ID NO: 2 se cuenta como 1, y en el que el polipéptido tiene al menos 50% de actividad residual después de que el polipéptido se ha mantenido a una temperatura de al menos 70°C durante 15 min., en el que la actividad se mide con acetil AlaAlaPro-para-nitroaloinina (Ac-AAP-pNA) como sustrato.

En una realización, la presente invención se refiere a un polipéptido que es un polipéptido aislado, sustancialmente puro, puro, recombinante, sintético, o variante, del polipéptido como se describe aquí.

En otro aspecto, la presente invención se refiere a un polipéptido aislado que tiene actividad de endoproteasa específica de prolina, que comprende la secuencia de aminoácidos de SEQ ID NO: 2. Una endoproteasa específica de prolina, como se describe aquí, puede derivar de *Rasamsonia emersonii*. El término "derivado" o "derivable", con respecto al origen de un polipéptido como se describe aquí, significa que, cuando se lleva a cabo una búsqueda de BLAST con un polipéptido según la presente invención, el polipéptido de la presente invención puede derivar de una fuente natural, tal como una célula microbiana, de la cual un polipéptido endógeno muestra el porcentaje más elevado de homología o identidad con el polipéptido como se describe aquí.

Ventajosamente, un polipéptido que tiene actividad de endoproteasa específica de prolina proporcionado por la presente invención es relativamente termoestable. Sorprendentemente, se encontró que un polipéptido según la presente invención es más termoestable que una endoproteasa específica de prolina procedente de *Aspergillus niger*. Se encontró que un polipéptido según la presente descripción tuvo al menos 50%, o al menos 55%, o al menos 60% de actividad residual después de que el polipéptido se hubo conservado a una temperatura de al menos 70°C durante 15 min., tal como a 70°C durante 15 min., o 71°C durante 15 min., en el que la actividad se mide con acetil AlaAlaPro-para-nitroaloinina (Ac-AAP-pNA) como sustrato. En consecuencia, una endoproteasa específica de prolina, como se describe aquí, se puede usar fácilmente en aplicaciones alimentarias, en las que la industria alimentaria desea aplicar la enzima a mayores temperaturas, por ejemplo durante la maceración en un procedimiento para la preparación de cerveza.

Preferiblemente, un polipéptido proporcionado por la invención puede ser un polipéptido que tiene al menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad con respecto a la secuencia de polipéptido maduro de SEQ ID NO: 2. Un polipéptido según SEQ ID NO: 2 comprende una pre-pro-secuencia. Al segregar el polipéptido a través de una membrana de la célula hospedante, se elimina la pre-pro-secuencia, tal como los aminoácidos 1 a 35, 1 a 36, 1 a 37, 1 a 38, 1 a 39, 1 a 40, 1 a 41, 1 a 42, 1 a 43, 1 a 44, 1 a 45, 1 a 46, 1 a 47, 1 a 48, 1 a 49, 1 a 50, o 1 a 51 de SEQ ID NO: 2. La secuencia de polipéptido maduro que tiene actividad de endoproteasa específica de prolina de SEQ ID NO: 2 puede comprender los aminoácidos 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o 51 a los aminoácidos 512, 513, 514, 515, 516, 517, 518, 519, 520 o 526 de SEQ ID NO:2, en la que el aminoácido metionina en la posición 1 en SEQ ID NO: 2 se cuenta como el número 1. La secuencia de polipéptido maduro de SEQ ID NO: 2 puede comprender o contener los aminoácidos 36 a 526 de SEQ ID NO: 2, en la que el aminoácido metionina en la posición 1 en SEQ ID NO: 2 se cuenta como el número 1.

Un polipéptido según la presente invención puede ser codificado por cualquier secuencia polinucleotídica adecuada. Típicamente, una secuencia polinucleotídica está optimizada en los codones, o es una secuencia optimizada para el par de codones para la expresión de un polipéptido como se describe aquí en una célula hospedante particular.

Un polipéptido como se describe aquí puede ser codificado por un polinucleótido que se hibrida bajo restricción media, preferiblemente en condiciones de alta restricción, a la hebra complementaria de la secuencia codificante del polipéptido maduro de SEQ ID NO: 1. SEQ ID NO: 1 es una secuencia polinucleotídica con codones optimizados para la expresión de un polipéptido como se describe aquí en una célula hospedante de *Aspergillus niger*.

Un polipéptido como se describe aquí también puede ser codificado por un ácido nucleico que tiene al menos 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con respecto a una secuencia codificante de polipéptido maduro de SEQ ID NO: 1.

Un polipéptido que tiene actividad de endoproteasa específica de prolina, como se describe aquí, también puede ser una variante de un polipéptido maduro de SEQ ID NO: 2, que comprende una sustitución, supresión y/o inserción en

una o más posiciones del polipéptido maduro de SEQ ID NO: 2. Una variante del polipéptido maduro de SEQ ID NO: 2 puede ser una secuencia de aminoácidos que difiera en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 o 12 aminoácidos de los aminoácidos del polipéptido maduro de SEQ ID NO: 2.

5 En una realización, la presente descripción presenta un fragmento biológicamente activo de un polipéptido como se describe aquí.

Los fragmentos biológicamente activos de un polipéptido de la invención incluyen polipéptidos que comprenden secuencias de aminoácidos suficientemente idénticas a o derivadas de la secuencia de aminoácidos de la proteína de endoproteasa específica de prolina (por ejemplo, la secuencia de aminoácidos madura de SEQ ID NO: 2), que incluye menos aminoácidos que la proteína de longitud completa, pero que exhibe al menos una actividad biológica de la proteína de longitud completa correspondiente. Típicamente, los fragmentos biológicamente activos comprenden un dominio o motivo con al menos una actividad de la proteína de endoproteasa específica de prolina. Un fragmento biológicamente activo puede comprender, por ejemplo, un dominio catalítico. Un fragmento biológicamente activo de la proteína de la invención puede ser un polipéptido que tiene, por ejemplo, 10, 25, 50, 100 o más aminoácidos de longitud. Además, otras porciones biológicamente activas, en las que se suprimen otras regiones de la proteína, se pueden preparar mediante técnicas recombinantes, y se pueden evaluar en busca de una o más de las actividades biológicas de la forma nativa de un polipéptido de la invención.

La descripción también presenta fragmentos de ácidos nucleicos que codifican los fragmentos biológicamente activos anteriores de la proteína de endoproteasa específica de prolina.

20 Un polipéptido según la presente invención puede ser una proteína de fusión. Las técnicas para producir polipéptidos de fusión son conocidas en la técnica, e incluyen ligar las secuencias codificantes que codifican los polipéptidos de manera que estén en el marco. La expresión del polipéptido fusionado está bajo el control del mismo promotor o promotores y terminador. Los polipéptidos híbridos pueden comprender una combinación de secuencias polipeptídicas parciales o completas obtenidas a partir de al menos dos polipéptidos diferentes, en el que uno o más pueden ser heterólogos a una célula hospedante. Tales polipéptidos de fusión procedentes de al menos dos polipéptidos diferentes pueden comprender un dominio de unión procedente de un polipéptido, enlazado operablemente a un dominio catalítico procedente de un segundo polipéptido. Los ejemplos de polipéptidos de fusión y fusiones de secuencias señal son, por ejemplo, como se describen en los documentos WO2010/121933, WO2013/007820 y WO2013/007821.

30 Un polipéptido según la presente invención puede derivar de cualquier eucariota adecuada. Una célula eucariota puede ser una célula de mamífero, de insecto, de planta, fúngica, o de alga.

Un polipéptido según la presente invención también puede derivar de una célula fúngica filamentosa o de una célula fúngica filamentosa termófila. Las células fúngicas filamentosas preferidas pertenecen a una especie de un *Acremonium*, *Aspergillus*, *Chrysosporium*, *Myceliophthora*, *Penicillium*, *Talaromyces*, *Rasamsonia*, *Thielavia*, *Fusarium* o *Trichoderma*, *Amorphotheca*, *Pseudocercospora*, *Trametes*, *Rhizomucor*, *Calcarisporiella*, *Thermomyces*, *Thermoascus*, *Coryascus*, *Myricoccum*, *Scytalidium*, *Chaetomium*, *Paecilomyces*, *Corynascus*, *Malbranchea*, *Stilbella*, *Thermomyces*, *Dactylomyces*, *Humicola*, *Chaetomium*, *Melanocarpus*, *Rhizomucor*, *Lentinula*, género *Anaeromyces*, y lo más preferible, pertenece a una especie de *Aspergillus niger*, *Acremonium alabamense*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus sojae*, *Aspergillus fumigatus*, *Talaromyces emersonii*, *Rasamsonia emersonii*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium oxysporum*, *Myceliophthora thermophila*, *Trichoderma reesei*, *Thielavia terrestris*, *Penicillium chrysogenum*, *Amorphotheca resiniae*, *Aureobasidium pullulans*, *Pseudocercospora herpotrichoides*, *Trametes versicolor* 52J, *Rhizomucor pusillus*, *Calcarisporiella thermophila*, *Talaromyces thermophilus*, *Thermomyces lanuginosus*, *Thermoascus auratiacus*, *Corynascus thermophilus*, *Myricoccum thermophilum*, *Scytalidium thermophilum*, *Myceliophthora hinnulea*, *Chaetomium thermophilum*, *Paecilomyces byssochlamydoides*, *Corynascus sepedonium*, *Malbranchea cinnamomea*, *Thielavia australiensis*, *Stilbella thermophila*, *Thermomyces stellatus*, *Talaromyces emersonii*, *Dactylomyces thermophilus*, *Humicola hyalothermophila*, *Acremonium thermophilum*, *Chaetomium olivicolor*, *Melanocarpus albomyces*, *Rhizomucor miehei*, *Lentinula edodes* o *Anaeromyces mucronatus*. Un polipéptido es derivable preferiblemente de *Rasamsonia emersonii*.

50 Un polipéptido según la presente invención puede ser un polipéptido de origen natural, o un polipéptido genéticamente modificado o recombinante.

Un polipéptido como se describe aquí puede estar purificado. La purificación de una proteína es conocida por una persona experta en la técnica.

En un aspecto, la presente invención se refiere a una composición que comprende un polipéptido como se describe aquí.

55 Una composición como se describe aquí puede comprender un soporte, un excipiente, una enzima auxiliar, u otros compuestos. Típicamente, una composición, o una formulación, comprende un compuesto con el que se puede formular una endoproteasa específica de prolina. Como se usa aquí, un excipiente es una sustancia inactiva formulada junto con un polipéptido como se describe aquí, por ejemplo sacarosa o lactosa, glicerol, sorbitol, o

cloruro de sodio. Una composición que comprende un polipéptido como se describe aquí puede ser una composición líquida o una composición sólida. Una composición líquida comprende habitualmente agua. Cuando se formula como una composición líquida, la composición comprende habitualmente componentes que disminuyen la actividad del agua, tales como glicerol, sorbitol o cloruro de sodio (NaCl). Una composición sólida que comprende un polipéptido como se describe aquí puede comprender un granulado que comprende la enzima, o la composición comprende un polipéptido encapsulado en matrices líquidas como liposomas o en geles como alginato o carrageenanos. Hay muchas técnicas conocidas en la técnica para encapsular o granular un polipéptido o enzima (véase, por ejemplo, G.M.H. Meesters, "Encapsulation of Enzymes and Peptides", Capítulo 9, en N.J. Zuidam and V.A. Nedović (eds.) "Encapsulation Technologies for Active Food Ingredients and food processing" 2010).

Una composición como se describe aquí también puede comprender un soporte que comprende un polipéptido como se describe aquí. Un polipéptido como se describe aquí se puede unir o inmovilizar a un soporte mediante tecnologías conocidas en la técnica.

La presente descripción también se refiere a un procedimiento para preparar una composición que comprende un polipéptido como se describe aquí, que puede comprender secar por pulverización un medio de fermentación que comprende el polipéptido, o granular o encapsular un polipéptido como se describe aquí, y preparar la composición.

En otro aspecto, la presente invención se refiere a una secuencia polinucleotídica que codifica un polipéptido como se describe aquí, que tiene menos 85%, 90%, 95%, 96%, 97%, 98%, o 99% de identidad de secuencia con SEQ ID NO: 1, o a la secuencia codificante de polipéptido maduro de SEQ ID NO: 1. Una secuencia polinucleotídica como se describe aquí puede comprender SEQ ID NO: 1, o puede comprender la secuencia codificante de polipéptido maduro de SEQ ID NO: 1.

En otra realización de la presente invención, se describe un ácido nucleico que es un ácido nucleico aislado, sustancialmente puro, puro, recombinante, sintético o variante del ácido nucleico de SEQ ID NO: 1. Una secuencia de ácido nucleico variante puede tener, por ejemplo, al menos 70% de identidad de secuencia con SEQ ID NO: 1.

En otro aspecto, la presente invención se refiere a un vector de expresión que comprende un polinucleótido como se describe aquí enlazado operablemente a una o más secuencias de control que dirigen la expresión del polipéptido en una célula hospedante de expresión.

Hay varias maneras de insertar un ácido nucleico en un constructo de ácido nucleico o en un vector, que son conocidas por una persona experta en la técnica; véase, por ejemplo, Sambrook y Russell, *Molecular Cloning: A Laboratory Manual*, 3ª Ed., CSHL Press, Cold Spring Harbor, NY, 2001. Puede ser deseable manipular un ácido nucleico que codifica un polipéptido de la presente invención con secuencias de control, tales como secuencias promotoras y terminadoras.

Un promotor puede ser cualquier secuencia promotora apropiada adecuada para una célula hospedante eucariota o procariota, que muestra actividad transcripcional, incluyendo promotores mutantes, truncados, e híbridos, y se puede obtener a partir de polinucleótidos que codifican polipéptidos extracelulares o intracelulares ya sea endógenos (nativos) o heterólogos (extraños) a la célula. El promotor puede ser un promotor constitutivo o uno inducible. Preferiblemente, el promotor es un promotor inducible, por ejemplo un promotor inducible por almidón. Los promotores adecuados en hongos filamentosos son promotores que se pueden seleccionar del grupo que incluye, pero no se limita a, promotores obtenidos de los polinucleótidos que codifican TAKA amilasa de *A. oryzae*, proteinasa aspártica de *Rhizomucor miehei*, promotor *gpdA* de *Aspergillus*, alfa-amilasa neutra de *A. niger*, alfa-amilasa estable a ácidos de *A. niger*, glucoamilasa (*glaA*) de *A. niger* o *A. awamori*, endoxilanasas (*xlnA*) o beta-xilosidasas (*xlnD*) de *A. niger* o *A. awamori*, celobiohidrolasa I (CBHI) de *T. reesei*, lipasa de *R. miehei*, proteasa alcalina de *A. oryzae*, triosa fosfato isomerasa de *A. oryzae*, acetamidasa de *A. nidulans*, amiloglucosidasa de *Fusarium venenatum* (documento WO 00/56900), Dania de *Fusarium venenatum* (documento WO 00/56900), Quinn de *Fusarium venenatum* (documento WO 00/56900), proteasa similar a tripsina de *Fusarium oxysporum* (documento WO 96/00787), beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasas I de *Trichoderma reesei*, xilanasas II de *Trichoderma reesei*, beta-xilosidasas de *Trichoderma reesei*, así como el promotor NA2-tpi (un híbrido de los promotores procedentes de los polinucleótidos que codifican alfa-amilasa neutra de *A. niger* y triosa fosfato isomerasa de *A. oryzae*), y promotores, mutantes, truncados e híbridos de los mismos.

Se puede usar cualquier terminador que sea funcional en una célula como se describe aquí, que sean conocidos por una persona experta en la técnica. Los ejemplos de secuencias terminadoras adecuadas en hongos filamentosos incluyen secuencias terminadoras de un gen fúngico filamentosos, tal como de genes de *Aspergillus*, por ejemplo del gen que codifica TAKA amilasa de *A. oryzae*, los genes que codifican glucoamilasa (*glaA*) de *A. niger*, antranilato sintasa de *A. nidulans*, alfa-glucosidasa de *A. niger*, *trpC*, y/o proteasa similar a tripsina de *Fusarium oxysporum*.

En otro aspecto, la presente invención se refiere a una célula hospedante que comprende un constructo de ácido nucleico o un vector de expresión como se describe aquí. Una célula hospedante adecuada puede ser una célula de

mamífero, de insecto, de planta, fúngica, o de alga, o una célula bacteriana. Una célula hospedante adecuada puede ser una célula fúngica, por ejemplo del género *Acremonium*, *Aspergillus*, *Chrysosporium*, *Fusarium*, *Myceliophthora*, *Penicillium*, *Rasamsonia*, *Talaromyces*, *Thielavia*, *Trichoderma*, *Saccaromyces*, *Kluyveromyces*, *Pichia*, por ejemplo *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus foetidus*, *A. oryzae*, *A. sojae*, *Talaromyces emersonii*, *Rasamsonia emersonii*, *Chrysosporium lucknowense*, *Fusarium oxysporum*, *Myceliophthora thermophila*, *Thielavia terrestris* o *Trichoderma reesei* o *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Pichia pastoris*.

Una célula hospedante puede ser una célula hospedante recombinante o transgénica. La célula hospedante se puede modificar genéticamente con un constructo de ácido nucleico o con un vector de expresión como se describe aquí con técnicas estándar conocidas en la técnica, tal como electroporación, transformación por protoplastos, o conjugación, por ejemplo como se describe en Sambrook y Russell, *Molecular Cloning: A Laboratory Manual*, 3ª Ed., CSHL Press, Cold Spring Harbor, NY, 2001.

En otro aspecto, la presente invención se refiere a un procedimiento para la producción de un polipéptido como se describe aquí, que comprende cultivar una célula hospedante en un medio de fermentación adecuado en condiciones que conduzcan a la producción del polipéptido, y producir el polipéptido. Una persona experta en la técnica sabe cómo llevar a cabo un procedimiento para la producción de un polipéptido como se describe aquí dependiendo de la célula hospedante usada, tal como el pH, la temperatura y composición de un medio de fermentación. Las células hospedantes se pueden cultivar en matraces de agitación, o en fermentadores que tienen un volumen de 0,5 o 1 litro o más a 10 a 100 o más metros cúbicos. El cultivo se puede llevar a cabo aerobia o anaerobiamente, dependiendo de los requisitos de la célula hospedante.

Ventajosamente, un polipéptido como se describe aquí se recupera o aísla del medio de fermentación.

Un polipéptido que tiene actividad de endoproteasa específica de prolina, o una composición que comprende un polipéptido como se describe aquí, se puede usar en una gran variedad de aplicaciones, por ejemplo en la producción de un producto alimentario o de pienso, tal como en la producción de un hidrolizado proteico. Varias proteínas alimentarias contienen subfracciones muy alérgicas que pueden ser incluso tóxicas a individuos específicos, tal como el gluten, que contiene prolaminas con secuencias peptídicas ricas en prolina. Estas proteínas se pueden someter a la nueva enzima para aliviar su antigenicidad o su toxicidad.

Un grupo de personas para las que el gluten es tóxico son los individuos que sufren celiaquía. La celiaquía, también conocida como enfermedad celíaca, es una enfermedad autoinmune del intestino delgado causada por la ingestión de proteínas del gluten procedentes de cereales, tales como alfa-gliadina del trigo, hordeína de la cebada, secalina del centeno, y avenina de avenas.

En consecuencia, un polipéptido que tiene actividad de endoproteasa específica de prolina, o una composición que comprende un polipéptido como se describe aquí, se puede usar en la preparación de un suplemento dietético o como un medicamento en el tratamiento de un paciente que sufre celiaquía, y/o en el tratamiento de personas intolerantes al gluten.

Un polipéptido como se describe aquí también se puede usar como un auxiliar del procesamiento para hidrolizar gluten en un producto alimentario.

En consecuencia, la presente invención se refiere a un procedimiento para la preparación de un producto alimentario o de pienso, que comprende incubar una forma intermedia de un producto alimentario o de pienso con un polipéptido o composición que comprende el polipéptido como se describe aquí, y preparar el producto alimentario o de pienso. Un producto alimentario en un procedimiento como se describe aquí incluye una bebida, tal como cerveza, vino, o zumo de frutas, o un producto horneado, o un producto lácteo, pero no se limita a ellos.

Incubar una forma intermedia de un producto alimentario o de pienso con un polipéptido o composición que comprende el polipéptido como se describe aquí puede comprender añadir el polipéptido o composición que comprende el polipéptido a la forma intermedia del producto alimentario o de pienso.

Un procedimiento para la preparación de un producto alimentario puede ser un procedimiento para la preparación de cerveza. Habitualmente, un procedimiento para la preparación de cerveza comprende macerar la malta para obtener una pulpa, y filtrar la pulpa para obtener un mosto, hervir el mosto, por ejemplo con lúpulo, y fermentar el mosto inoculándole al mosto la levadura. Tras la fermentación, un procedimiento para la preparación de cerveza comprende típicamente una fase de maduración y de estabilización, y habitualmente también una fase de filtración, dependiendo del tipo de cerveza que se produce.

Una forma intermedia de un producto alimentario puede ser cualquier forma de un producto alimentario durante la preparación del producto alimentario. Una forma intermedia de cerveza puede ser, por ejemplo, una pulpa, un mosto, un caldo de fermentación, o una cerveza verde. Como se usa aquí, cerveza verde es una cerveza que resulta de una fermentación primaria, y puede contener típicamente algo de levadura sin sedimentar y componentes de sabor indeseable. Una forma intermedia de pan puede ser por ejemplo una masa espesa o una masa licuada. Ventajosamente, un polipéptido que tiene actividad de endoproteasa específica de prolina, o una composición que comprende el polipéptido según la presente invención, se incuba con una pulpa durante la maceración en un

procedimiento para la preparación de cerveza. Sorprendentemente, una cantidad significativa de proteínas activamente turbias se reduce mediante un polipéptido según la presente invención durante la maceración. En consecuencia, se puede evitar la formación de turbidez en la cerveza final. Es ventajoso que un polipéptido que tiene actividad de endoproteasa específica de prolina según la presente invención sea activo durante la maceración, puesto que, tras la maceración, el mosto separado se hierve y la enzima se inactiva. Esto es ventajoso puesto que es indeseable que las enzimas estén activas en un producto alimentario final, tal como una cerveza.

La maceración es un procedimiento durante el cual el almidón se rompe en azúcares mediante enzimas naturales presentes en el grano, conocido por una persona experta en la técnica. Típicamente, la maceración comprende llevar la malta, es decir, una mezcla de un grano molido tal como cebada malteada y agua, poco a poco hasta una temperatura de alrededor de 45-52°C hasta una temperatura de 71 a 76 durante un cierto período de tiempo, lo que es conocido por una persona experta en la técnica. Habitualmente, la maceración comprende llevar la malta hasta una temperatura de 45-52°C, después hasta una temperatura de 60 a 65°C, después a una temperatura de alrededor de 71 a 76°C, y opcionalmente hasta una temperatura final de 76 a 79°C.

Tras la maceración, la fracción líquida, también denominada el "mosto", se separa de la fracción sólida. Subsiguientemente, el mosto se hierve típicamente con ingredientes adicionales, tal como lúpulo. Durante la ebullición del mosto, se inactivan las enzimas presentes en el mosto.

En una realización, un procedimiento para la preparación de cerveza comprende incubar una pulpa con un polipéptido o una composición que comprende un polipéptido como se describe aquí, y preparar la cerveza.

Las proteínas activamente turbias son proteínas ricas en prolina que pueden interactuar con polifenoles para formar agregados de proteína-polifenol. Estos agregados de proteína-polifenol provocan la formación de turbidez, también denominada "turbidez fría" en la cerveza. En consecuencia, un aspecto de la presente invención se refiere al uso de un polipéptido como se describe aquí, o una composición como se describe, en la reducción de la turbidez en una bebida. Una bebida como se usa aquí puede ser una cerveza.

Un producto alimentario y/o una forma intermedia de un producto alimentario puede comprender gluten.

Se encontró que un polipéptido que tiene actividad de endoproteasa específica de prolina como se describe aquí fue capaz de hidrolizar los epítomos tóxicos en el gluten en fragmentos no tóxicos. En consecuencia, la descripción se refiere al uso de un polipéptido o de una composición que comprende el polipéptido como se describe aquí en la reducción de gluten en un producto alimentario.

Un procedimiento para la preparación de un producto alimentario según la presente invención puede comprender una etapa de pasteurizar el producto alimentario. La pasteurización comprende habitualmente calentar un producto alimentario, o una forma intermedia de un producto alimentario, por ejemplo llevando el producto alimentario o la forma intermedia de un producto alimentario hasta una temperatura de entre 60 y 68°C, entre 10 y 20 min., o entre 12 y 18 min., o a una temperatura de entre 70-74°C, tal como alrededor de 72°C, durante al menos 5, 10 o 15 segundos.

Un producto alimentario, en un procedimiento como se describe aquí, también puede ser un hidrolizado de proteína. En consecuencia, la presente descripción se refiere a un procedimiento para la preparación de un hidrolizado de proteína, que comprende poner en contacto un sustrato proteico con un polipéptido o una composición como se describe aquí, y producir el hidrolizado de proteína. Un hidrolizado de proteína se puede preparar a partir de cualquier sustrato proteico adecuado, por ejemplo un sustrato proteico que sea rico en restos de prolina, tal como gluten en cereales o caseínas en leche bobina.

En un aspecto, la presente descripción se refiere a un producto alimentario obtenible mediante un procedimiento para la preparación de un producto alimentario como se describe aquí.

EJEMPLOS

MATERIALES Y MÉTODOS

45 **Ejemplo 1. Clonación, expresión y recuperación de endoproteasa específica de prolina (PEP) BC2G079**

Ejemplo 1.1. Clonación y expresión

La secuencia proteica de la endoproteasa específica de prolina (PEP) de *Rasamsonia emersonii* se muestra en SEQ ID NO: 2, y se denomina PEP BC2G079.

Para la expresión de esta proteína en *Aspergillus niger*, se diseñó una secuencia de ADN de codones adaptados que contiene sitios de restricción adicionales para la subclonación en un vector de expresión de *Aspergillus*. La adaptación de los codones se llevó a cabo como se describe en el documento WO 2008/000632. La secuencia de ADN con codones optimizados para *A. niger* del gen que codifica la proteína PEP de SEQ ID NO: 2 se muestra en SEQ ID NO: 1.

La secuencia de iniciación traduccional del promotor de glucoamilasa *gluA* se modificó en 5'-CACCGTCAAA ATG-3', y se usó una secuencia de terminación traduccional óptima 5'-TAAA-3' en la generación del constructo de expresión (como se detalla también en el documento WO2006/077258). Un fragmento de ADN, que contiene, entre otros, parte del promotor de glucoamilasa y el gen que codifica PEP, se sintetizó completamente, se purificó y se digirió con EcoRI y Pacl. El vector pGBTOP-16 (Figura 1) se linealizó mediante digestión con EcoRI/Pacl, y el fragmento del vector linealizado se purificó subsiguientemente mediante extracción en gel. El fragmento de ADN que contiene la región codificante de PEP se clonó en el vector pGBTOP-16, dando como resultado pGBTOP-PEP. Subsiguientemente, GBA 306 de *A. niger* ($\Delta gluA$, $\Delta pepA$, $\Delta hdfa$, amplicón de BamHI adaptado, $\Delta amyBII$, $\Delta amyBI$, $\Delta amyA$, cepa negativa a alfa-amilasa y glucoamilasa) se transformó con el vector pGBTOP-PEP linealizado mediante digestión con *NotI*, en un protocolo de cotransformación con pGBAAS-4 linealizado, con la cepa y métodos como se describen en el documento WO 2011/009700 y referencias allí, y se seleccionó en medio que contiene acetamida, y la colonia se purificó según los procedimientos estándar. La transformación y selección se llevaron a cabo como se describió en los documentos WO 98/46772 y WO 99/32617. Las cepas que contienen el gen PEP que codifica PEP BC2G079 se seleccionaron vía PCR con cebadores específicos para el gen PEP para verificar la presencia del casete de expresión pGBTOP-PEP. Se seleccionó un único transformante, denominado PEP1, y posteriormente se replicó en placa para obtener un inóculo de una cepa individual.

Ejemplo 1.2. Producción de PEP BC2G079 en la cepa PEP1 de *A. niger*

Se prepararon esporas recientes de PEP-1 de *A. niger*. 4 matraces de agitación con 100 ml de medio 1 de fermentación (10% p/v de Sólidos de Licor de Maíz, 1% p/v de glucosa.H₂O, 0,1% p/v de NaH₂PO₄.H₂O, 0,05% p/v de MgSO₄.7H₂O, 0,025% p/v de Basildon, pH 5,8), en matraces de agitación de 500 ml con matraces con deflectores, se inocularon con 10⁷ esporas. Estos precultivos se incubaron a 34°C y 170 rpm durante 16-24 horas. De los precultivos, se usaron 50 ml para la inoculación de 1 matraz de agitación con 1 litro de medio 2 de fermentación (15% p/v de maltosa, 6% p/v de bacto-soytone, 1,5% p/v de (NH₄)₂SO₄, 0,1% p/v de NaH₂PO₄.H₂O, 0,1% p/v de MgSO₄.7H₂O, 0,1% p/v de L-arginina, 8‰ p/v de Tween-80, 2‰ p/v de Basildon, 2% p/v de MES pH 5,1) en un tamaño de matraz de agitación de 5 litros, y se agitaron a 34°C y 170 rpm. Después de 3, 4, 5, y 6 días de incubación, el pH del cultivo se redujo hasta pH 5,0 usando HCl 2N, y las muestras procedentes de estos puntos de tiempo se analizaron para determinar la actividad de PEP. Se tomaron muestras de 50 ml, y el sobrenadante se separó de la biomasa mediante centrifugación y posterior filtración. La muestra con la mayor actividad se usó para caracterizar la PEP producida.

Ejemplo 1.3. Producción de la endoproteasa específica de prolina de referencia a partir de *A. niger*

La endoproteasa específica de prolina de *A. niger* es conocida desde el documento WO2002/046381. La secuencia proteica de la endoproteasa específica de prolina (PEP) de *A. niger* se muestra en SEQ ID NO: 5, en la que los primeros 17 aminoácidos son una secuencia señal de pectin-metil esterasa de *A. niger* (PMeA ss; SEQ ID NO: 3), y la siguiente parte comprende 19 aminoácidos de la prosequencia de endoproteasa específica de prolina de *A. niger* (SEQ ID NO: 4).

La expresión y producción de PEP de *A. niger* se llevó a cabo de la misma manera como se describe en los ejemplos 1.1. y 1.2. La PEP de *A. niger* se usa como la enzima de referencia en el Ejemplo 2.

Ejemplo 2. Medidas de la actividad de endoproteasa específica de prolina (PEP)

100 µl de sobrenadante de cultivo como se produce en el Ejemplo 1, diluido en amortiguador de acetato de sodio 0,1 M a pH 4,5 con NaCl 50 mM, se incubaron con 100 µl de Ac-AAP-pNA 6 mM (acetil-AlaAlaPro-paranitroanilina, de Selleckchem o CPC Scientific; pureza >95,0%, basado en el análisis de HPLC) en amortiguador de NaAc 0,1 M a pH 4,5 con NaCl 50 mM, en MTP (placa de microtitulación) de fondo plano de 96 pocillos Nunc. Después de 60 minutos a 20°C, la reacción se detuvo añadiendo 40 µl de HCl 1 M. La pNA que se hubo liberado mediante PEP se midió en un espectrofotómetro de MTP Tecan a 405 nm (A405) (www.tecan.com). El blanco se preparó mezclando el sobrenadante de cultivo diluido con la disolución de sustrato que se había mezclado previamente con la disolución de HCl. La actividad se expresa como pNASU. 1 pNASU es la cantidad de enzima que libera a partir de Ac-AAP-pNA en 1 hora la cantidad de pNA que corresponde a un incremento en la absorción a 405 nm de 1 OD, usando las condiciones como se describen anteriormente. La A405 no debería estar por debajo del valor del blanco al comienzo de la reacción, o por encima de 2,5 al final de la reacción, ni la A405 puede exceder el intervalo lineal del espectrofotómetro que se use.

Tabla 1. Actividad relativa de PEP BC2G079, en comparación con la referencia de PEP.

Descripción	Actividad (pNASU/ml)	Actividad específica (pNASU/mg)
Referencia de PEP	100%	100%
PEP BC2G079	171%	150%

5 A fin de comparar la actividad específica de PEP BC2G079 con la referencia PEP, se determinaron las concentraciones de proteína de respectivamente la PEP BC2G079 y la referencia PEP midiendo la absorbancia a 280 nm tras la filtración en gel sobre una columna PD10 para eliminar compuestos de bajo peso molecular que puede interferir con la medida a 280 nm. La SDS-PAGE cuantitativa, llevada a cabo como se describió en el documento WO2013160316, mostró que más del 80% de la proteína presente es la PEP de interés. El coeficiente de extinción se calculó usando la herramienta Prot Param, según se puede acceder a través del sitio web <http://web.expasy.org/protparam/>. La secuencia madura de SEQ ID NO: 2 (aminoácidos 36-526) se usó para calcular el coeficiente de extinción para PEP BC2G079 ($A_{280}^{1\text{cm},1\text{mg/ml}} = 2,81$). La secuencia madura de SEQ ID NO 5 (aminoácidos 37-521) se usó para la referencia PEP ($A_{280}^{1\text{cm},1\text{mg/ml}} = 3,05$).

10 **Ejemplo 3. Estabilidad térmica de la endoproteasa específica de prolina (PEP) BC2G079**

15 Para evaluar la estabilidad térmica de PEP BC2G079, el ensayo de actividad estuvo precedido por una incubación de alícuotas de 100 µl de una disolución de diez veces del sobrenadante de cultivo producido en el Ejemplo 1 en amortiguador (NaAc 0,1 M pH 4,5, con NaCl 50 mM) a 55°C y 65°C durante 15 min. en una placa de PCR en una máquina de PCR. Tras la incubación durante 15 min., las muestras se enfriaron rápidamente hasta 25°C en la máquina de PCR. Se midió la pNASU/ml de cada muestra. La actividad inicial medida antes de la incubación a temperatura elevada (0 minutos) se usa como referencia (100%) para determinar la actividad residual.

Tabla 2. Actividad de endoproteasa específica de prolina de PEP BC2G079 tras la incubación durante 15 minutos a 55°C y 65°C

Descripción	Actividad residual tras 15' a 55°C (pNASU)	Actividad residual tras 15' a 65°C (pNASU)
PEP BC2G079	105%	84%

20 Los resultados en la Tabla 1 muestran que PEP BCG079 derivada de *Rasamsonia emersonii* es relativamente estable a una temperatura de 65 grados Celsius.

Ejemplo 5. Perfil de estabilidad térmica de PEP BC2G079 en comparación con la referencia PEP de *A. niger*

25 Para evaluar el perfil de estabilidad térmica de PEP BC2G079 en condiciones más extremas, el ensayo de actividad estuvo precedido por una etapa de calentamiento. Alícuotas de 100 µl de un sobrenadante de cultivo como se produce en el Ejemplo 1 se diluyeron diez veces en amortiguador (NaAc 0,1 M pH 4,5, con NaCl 50 mM), y se calentaron a lo largo de un intervalo de temperaturas en una placa de PCR en una máquina de PCR. Después del período de incubación de 15 min., las muestras se enfriaron rápidamente hasta 25°C en la máquina de PCR. Se midió la pNASU/ml de cada muestra. La actividad inicial medida antes de la incubación a temperatura elevada (0 minutos) se usó como referencia (100%) para determinar la actividad residual. Los resultados en la tabla 3 muestran que una endoproteasa específica de prolina (BCG079) procedente de *Rasamsonia emersonii* es sustancialmente más termoestable en comparación con la endoproteasa específica de prolina de referencia procedente de *A. niger*.

Tabla 3. Actividad residual de endoproteasa específica de prolina de PEP BC2G079 y de la referencia PEP de *A. niger* tras la incubación de 15 minutos a diversas temperaturas.

Temperatura de incubación (°C)	Actividad residual de la referencia PEP (pNASU/ml)	Actividad residual de BC2G079 (pNASU/ml)
51,4	100%	
54,4	97%	
59,0	93%	
63,5		100%
64,5	76%	
66,6		91%
69,3	24%	
71,2		65%
72,4	1%	

Temperatura de incubación (°C)	Actividad residual de la referencia PEP (pNASU/ml)	Actividad residual de BC2G079 (pNASU/ml)
74,0	0%	
76,7		11%
81,4		0%

EJEMPLO 6. Perfil de actividad de pH de PEP BC2G079

Para evaluar el intervalo operacional de pH de PEP BC2G079, el sobrenadante de cultivo y el sustrato Ac-AAP-pNA se diluyeron ambos en amortiguadores que oscilan desde pH 3,5 hasta pH 7 (ácido cítrico 0,1 M/Na₂HPO₄, con NaCl 50 mM). Subsiguientemente, se mezclaron 100 µl del sobrenadante diluido con 100 µl de la disolución de sustrato Ac-AAP-pNA 6 mM en una MTP de fondo plano de 96 pocillos. La reacción se detuvo tras 60 min. a 20°C con la adición de 40 µl de NaOH 1 M. La actividad de la PEP se midió mediante el incremento en la absorbancia a 405 nm. Los resultados en la Tabla 4 muestran que la mayor actividad para BC2G079 se observó a pH = 5,0, que se ajustó a 100%, mientras que la referencia PEP de *A. niger* muestra la mayor actividad (100%) a pH = 5,5. Globalmente, el perfil de actividad de pH de BC2G079 se ha desplazado 0,5 unidades de pH hacia condiciones de operación de pH más ácido. Como consecuencia, por debajo de pH 5,5, PEP BC2G079 exhibe mayor actividad relativa en comparación con la referencia de PEP.

Tabla 4. Perfil de actividad de pH de PEP BC2G079

pH	PEP BC2G079 (pNASU/ml)	PEP de referencia (pNASU/ml)
3,5	59%	54%
4,0	81%	68%
4,5	97%	78%
5,0	100%	97%
5,5	86%	100%
6,0	63%	92%
6,5	41%	60%
7,0	24%	34%

15 EJEMPLO 7. Ensayo de maceración

7.1. Métodos

Para determinar las PNACU, se mezcla una alícuota de 100 µl en una MTP de fondo plano de 96 pocillos Nunc con 100 µl de Ac-AAP-pNA 6 mM en amortiguador de NaAc 0,1 M, pH 4,5, que contiene NaCl 50 mM. El incremento en la OD a 405 nm se registra como función del tiempo en un espectrofotómetro de MTP Tecan. Las PNACU se deberían de calcular desde la parte lineal de la curva, preferiblemente la pendiente inicial. El ensayo se realiza a 20°C. A fin de medir el incremento en la OD a lo largo de un período de tiempo suficientemente largo mientras se permanece en el intervalo lineal del detector, la muestra enzimática se diluye en consecuencia con amortiguador de NaAc 0,1 M, pH 4,5, con NaCl 50 mM. Una PNACU es la cantidad de enzima que libera en una hora a partir de Ac-AAP-pNA la cantidad de pNA que corresponde a un incremento de la absorción a 405 nm de una unidad de OD.

25 Las proteínas activamente turbias se midieron con un tannómetro, usando para este método las instrucciones operativas de Pfeuffer. Se añadió ácido tánico a las muestras, y la turbidez se midió bajo un ángulo de dispersión de 90 grados expresada en unidades EBC y dada a conocer para la adición de 2,5, 5 y 10 mg/l de ácido tánico.

La gliadina en el mosto se determinó con un Elisa competitivo (RIDASCREEN® Gliadin competitive (R-Biopharm), como una medida para el contenido de gluten.

30 7.2. Degradación de las proteínas activamente turbias y del gluten durante la maceración

La degradación de las proteínas activamente turbias y del gluten mediante una enzima PEP BC2G079 termoestable

5 en un proceso de maceración se determinó y se comparó con el comportamiento de la enzima de referencia PEP menos estable procedente de *A. niger*. En un experimento de maceración, se añadieron 5 ml de PEP BC2G079 termoestable, que contiene 460 PNACU/ml, a 200 ml de pulpa. En otro experimento de maceración, se añadieron 5 ml de la PEP de referencia, que contiene 460 PNACU/ml. Las enzimas se añadieron al comienzo del proceso de maceración. Se llevó a cabo una tercera maceración, en la que se añadió el mismo volumen de agua corriente en lugar de la disolución enzimática, para que sirva como una comparación con el blanco.

10 Los ensayos de maceración a escala de laboratorio se llevaron a cabo en un baño de maceración (Lochner Labor technik, Alemania). Se usaron 80 gramos de malta EBC estándar molida en 200 ml de agua. La pulpa se calentó por etapas como se muestra en la Tabla 5. Las pulpas se agitaron continuamente a 100 rpm durante el proceso de maceración. Al final de la maceración, la pulpa se filtró sobre un papel de filtro (Macherey-Nagel, MN614 ¼, diámetro 320 mm). El filtrado se denomina el mosto. Se tomó una muestra del mosto y se analizó para determinar la presencia de proteínas sensibles a la turbidez y de gluten, y para la determinación de la actividad de endoproteasa específica de prolina (pNASU).

Tabla 5. Esquema de maceración

Tiempo desde el comienzo (min)	Temperatura (°C)
0	50
15	50
28	63
43	63
55	75
70	75
73	78
78	78

15 Los resultados de la degradación de proteínas sensibles a la turbidez y de gliadina se muestran en la Tabla 6 y Tabla 7.

20 Tabla 6. Cantidad de proteínas activamente turbias (medido en unidades EBC) en muestras de mosto procedentes de pulpas en las que no había PEP, y de pulpas que se habían tratado con una PEP BC2G079 termoestable y una PEP de referencia

Ácido tánico	Turbidez EBC tras la adición de ácido tánico		
	Muestras de mosto		
	Sin PEP	PEP de referencia	PEP termoestable
2,5 mg/l	1,2	0,7	0,3
5 mg/l	4,0	2,7	0,8
10mg/l	9,3	6,2	1,5

Tabla 7. Contenido de gliadina en mosto procedente de pulpas en las que no había PEP, y de pulpas que se habían tratado con una PEP BC2G079 termoestable y una PEP de referencia

Muestras de mosto	Gliadina, en ppm
Sin PEP	362
PEP de referencia	135
PEP termoestable	82

Los resultados en la Tabla 6 y 7 muestran que una PEP termoestable, como BC2G079, es mucho más eficiente degradando proteínas activamente turbias y gliadina durante el proceso de maceración en comparación con la PEP de referencia menos estable. También se determinó la actividad residual de PEP al final del proceso de maceración descrito anteriormente (Tabla 8). Solamente la PEP BC2G079 termoestable es todavía activa, mientras que la PEP de referencia está totalmente inactivada. La mayor termoestabilidad se correlaciona con la degradación de proteína sensible a la turbidez y la eliminación de gliadina. Los datos de maceración muestran que una PEP termoestable es ventajosa en la maceración, puesto que la termoestabilidad prolonga significativamente el tiempo durante el cual la PEP es activa en la maceración.

10 Tabla 8. Actividad residual de PEP al final de la maceración

	Actividad residual al final de la maceración (pNASU/ml)
PEP de referencia	0%
PEP termoestable	41%

EJEMPLO 8. Determinación de la masa molecular de la PEP BC2G079 madura con LC-MS

Para el análisis mediante LC-MS de la PEP BC2G079 madura, se mezcló una muestra de 100 µl del sobrenadante de un medio de fermentación producido en el Ejemplo 1.2. con 100 µl de TCA al 20% (Chem Lab NV, Bélgica). Esta disolución se colocó en hielo durante 1 hora. Tras precipitar las proteínas, la muestra se centrifugó a 14000 rpm y 4°C, durante 15 min. Tras la centrifugación, el sobrenadante se eliminó, y el pelete se lavó una vez con 500 µl de acetona (-20°C, Sigma-Aldrich, Países Bajos) y se centrifugó a 14000 rpm y 4°C durante 10 min. Se eliminó el sobrenadante, y el pelete se disolvió en 50 µl de NaOH 50 mM (Sigma-Aldrich, Países Bajos), y después se añadieron 350 µl de NaHCO₃ 100 mM (Sigma-Aldrich, Países Bajos). Después, 200 µl de la muestra se redujeron y se alquilaron. Para la reducción, se añadieron 5 µl de TCEP (Sigma-Aldrich, Países Bajos) a la disolución, y se incubó a temperatura ambiente en una termomezcladora a 1000 rpm durante 30 min. Para la alquilación, se añadieron 5 µl de IAA 550 mM (yodoacetamida, Sigma-Aldrich, Países Bajos), y se incubó en una termomezcladora a 1000 rpm en la oscuridad durante 30 min. A fin de desglicosilar la muestra, se añadieron 20 µl de PNGasa F (Promega, USA). La mezcla se incubó en la termomezcladora a 1000 rpm a 37°C toda la noche. Al día siguiente, se añadió ácido fórmico al 1% (Merk, Alemania), a fin de diluir la muestra hasta una concentración de 50 µg/ml. La concentración de proteína se midió en el ensayo de proteína cuantitativo Qubit (Life technologies). La muestra se analizó en el Acquity I-class - Synapt G2-S (Waters, UK), con los siguientes parámetros: columna: Waters Acquity UPLC BEH300 C4 1,7 µm tamaño de poros 300Å 2,1x50 mm; temperatura de la columna: 75°C; volumen de inyección: 1 µl; fase móvil A: ácido fórmico 0,1% en Mili-Q (Biosolve, Países Bajos); fase móvil B: ácido fórmico 0,1% en acetonitrilo (Biosolve, Países Bajos). Se aplicó un gradiente de 15 minutos a la columna variando la fase A y B (de 20 a 50% de B). Los ajustes del detector de MS fueron: intervalo de masas de adquisición 500-3500 m/z, tiempo de barrido 1 s, ESI positiva, modo de resolución de MS TOF, con corrección de datos usando Leu-Enk (Sigma-Aldrich, Países Bajos) como masa conocida en el vuelo durante el experimento. La deconvolución espectral de los datos, extracción del estado de la carga, se realizó con una herramienta de software Waters MassLynx MaxEnt1: resolución de masa de salida = 1 Da/canal; modelo de daño: gaussiano (FWHH = 0,750 Da; relaciones de intensidad minim = 33% izquierda y derecha); repítase hasta converger.

Tras la corrección para la alquilación y desglicosilación, los espectros de masas muestran una especie principal con una masa molecular que corresponde a los aminoácidos 36-526 de SEQ ID NO 2. Además, hay una especie minoritaria con una masa molecular que corresponde a los aminoácidos 41-526 de SEQ ID NO 2. Como consecuencia, el término N de la especie principal en la muestra de PEP BC2G079 madura corresponde a RDPLHGPT, y el término N de la especie minoritaria comienza en GPTNAS. Ambas especies muestran múltiples picos en el espectro de masas. Cada pico se desplazó en 57 Da, o una multitud de 57 Da, indicando una diferencia en la alquilación. Una única alquilación añade 57 Da por cisteína. En los espectros, se observó la adición de 1, 3, 5 y 7 IAA, de acuerdo con la presencia de 7 cisteínas en la secuencia de aminoácidos. Las masas observadas para las especies principal y minoritaria completamente alquiladas son respectivamente 55218 Da y 54598 Da. Puesto que la desglicosilación añade un 1 Da adicional por sitio desglicosilado, los datos sugieren que 6-7 de los 8 sitios de glicosilación teóricos están realmente glicosilados.

LISTADO DE SECUENCIAS

<110> DSM IP Assets B.V.
 50 <120> Endoproteasa específica de prolina
 <130> 30254-WO-PCT

ES 2 744 898 T3

<160> 6

<170> BiSSAP 1.2

<210> 1

<211> 1580

5 <212> ADN

<213> Secuencia artificial

<220>

<221> fuente

<222> 1..1580

10 <223> /organismo="Secuencia artificial" /nota="Ácido nucleico con codones optimizados que codifica endoproteasa específica de prolina de Rasamsonia emersonii BC2G079 para la expresión en A. niger" /tipo_mol="ADN sin asignar"

<400> 1

atgccctccc	tctcctccct	cgttgccctg	actgcttctc	ttgtctctct	ggctgctgcc	60
gctgctcctc	gtctccctct	tctcctcgc	cctcccttgc	ctccccgtga	ccccttgcac	120
ggacctacca	acgcctccgc	cactttccag	cagctcatcg	accacaacca	ccccgagctt	180
ggcaccttct	cccagcgccta	ctggtggaat	gatgagtctt	ggaagggtcc	cggctctccc	240
gttgtccttt	tcacccccgg	tgaagaagat	gccagcggtt	acgtgggcta	cctgaagaac	300
accaccatca	cgggtctgat	cgctcagacc	atcggtggtg	ccgtcatcgt	cctcgaacac	360
cgctactggg	gccagtcctc	cccctacgac	tctctgacca	ccaagaacct	gcagtacctg	420
accctcaagc	agtccattgc	cgacctcacc	tacttcgcca	agaccgtcaa	gctccccctc	480
gaccgcaacg	gcagctccaa	cgccgacaag	gctccctggg	ttctcagcgg	tggaagctac	540
tctggtgctc	tctccgcctg	gactgccagc	acctcccccg	gtactttctg	ggcctaccac	600
gccagctctg	ctcctgttga	ggccatctac	gattactggc	agtacttcgc	tcccgtgcag	660
gatggattgc	ctgccaaactg	ctcgaaggac	ctctcccgtg	tcgtcgacta	catcgactcc	720
gttctccagt	cgggcaacgc	cactgccaaag	caacagctca	aggacctttt	cggctctgggt	780
gctctggagc	acgacgatga	cttcgcctcc	gctcttgaga	acggcccttg	gctctggcag	840
tcgaactcgt	tctacgacc	ctaccctcct	gtctacgagt	tctgcgacta	cgttgagaac	900
gcctacgcca	gccctcccgt	tgctgctggt	cccgatgggtg	ttggtctgga	gaaggetctg	960
tctggctacg	ccacctgggtg	gaacaaggtc	ttcttccccg	gctactgcgc	tacctacggc	1020
tactggtcct	ccaacgactc	cattgcctgc	ttcgacacct	acaaccagtc	gtcggccatg	1080

ES 2 744 898 T3

ttcaccgacc tttccgtctc caacactatc aaccgccagt ggaactggtt cctctgcaac 1140
gagcccttct tctactggca ggatgggtgct cccaagaacg tccccagcat tgtctctcgt 1200
ctggtcactg ctgagtactg gcagcgccag tgccccttgt tcttcctga agaggatggc 1260
tacacctacg gaagcgccaa gggcaagact gctgccgatg tcaacgcctg gaccaagggc 1320
tggttcttga ctaacaccac ccgtctgatc tggaccaacg gcgagcttga cccctggcgc 1380
tctgtgggtg tcagcagcaa gttccgtccc ggtggtcccc tccagtccac tccccaggct 1440
cctctgcagc tcattcccga ggggtgtccac tgctacgatc tgatcctcaa gaacgccgag 1500
gccaacgccg gtgtgcagcg tgttgcacc aacgaggttg ctcagatcaa ggcctgggtg 1560
aacgaatact accgtaagta 1580

<210> 2

<211> 526

<212> PRT

5 <213> Secuencia artificial

<220>

<223> *Rasamsonia emersonii*

<400> 2

Met Pro Ser Leu Ser Ser Leu Val Ala Leu Thr Ala Ser Leu Val Ser
1 5 10 15
Leu Ala Ala Ala Ala Pro Arg Leu Pro Leu Pro Pro Arg Pro Pro
20 25 30
Leu Pro Pro Arg Asp Pro Leu His Gly Pro Thr Asn Ala Ser Ala Thr
35 40 45
Phe Gln Gln Leu Ile Asp His Asn His Pro Glu Leu Gly Thr Phe Ser
50 55 60
Gln Arg Tyr Trp Trp Asn Asp Glu Phe Trp Lys Gly Pro Gly Ser Pro
65 70 75 80
Val Val Leu Phe Thr Pro Gly Glu Glu Asp Ala Ser Gly Tyr Val Gly
85 90 95
Tyr Leu Lys Asn Thr Thr Ile Thr Gly Leu Ile Ala Gln Thr Ile Gly
100 105 110
Gly Ala Val Ile Val Leu Glu His Arg Tyr Trp Gly Gln Ser Ser Pro
115 120 125
Tyr Asp Ser Leu Thr Thr Lys Asn Leu Gln Tyr Leu Thr Leu Lys Gln
130 135 140
Ser Ile Ala Asp Leu Thr Tyr Phe Ala Lys Thr Val Lys Leu Pro Phe
145 150 155 160
Asp Arg Asn Gly Ser Ser Asn Ala Asp Lys Ala Pro Trp Val Leu Ser
165 170 175
Gly Gly Ser Tyr Ser Gly Ala Leu Ser Ala Trp Thr Ala Ser Thr Ser
180 185 190
Pro Gly Thr Phe Trp Ala Tyr His Ala Ser Ser Ala Pro Val Glu Ala
195 200 205
Ile Tyr Asp Tyr Trp Gln Tyr Phe Ala Pro Val Gln Asp Gly Leu Pro
210 215 220
Ala Asn Cys Ser Lys Asp Leu Ser Arg Val Val Asp Tyr Ile Asp Ser
225 230 235 240
Val Leu Gln Ser Gly Asn Ala Thr Ala Lys Gln Gln Leu Lys Asp Leu
245 250 255
Phe Gly Leu Gly Ala Leu Glu His Asp Asp Asp Phe Ala Ser Ala Leu
260 265 270
Glu Asn Gly Pro Trp Leu Trp Gln Ser Asn Ser Phe Tyr Asp Pro Tyr

<213> Secuencia artificial

<220>

<223> Endoproteasa especifica de prolina de Aspergillus niger que comprende una secuencia señal de PMeA de A. niger

5 <400> 5

```

Met Val Lys Ser Ile Leu Ala Ser Val Phe Phe Ala Ala Thr Ala Leu
1          5          10          15
Ala Ala Arg Pro Arg Leu Val Pro Lys Pro Val Ser Arg Pro Ala Ser
20          25          30
Ser Lys Ser Ala Ala Thr Thr Gly Glu Ala Tyr Phe Glu Gln Leu Leu
35          40          45
Asp His His Asn Pro Glu Lys Gly Thr Phe Ser Gln Arg Tyr Trp Trp
50          55          60
Ser Thr Glu Tyr Trp Gly Gly Pro Gly Ser Pro Val Val Leu Phe Thr
65          70          75          80
Pro Gly Glu Val Ser Ala Asp Gly Tyr Glu Gly Tyr Leu Thr Asn Glu
85          90          95
Thr Leu Thr Gly Val Tyr Ala Gln Glu Ile Gln Gly Ala Val Ile Leu
100         105
Ile Glu His Arg Tyr Trp Gly Asp Ser Ser Pro Tyr Glu Val Leu Asn
115         120         125
Ala Glu Thr Leu Gln Tyr Leu Thr Leu Asp Gln Ala Ile Leu Asp Met
130         135         140
Thr Tyr Phe Ala Glu Thr Val Lys Leu Gln Phe Asp Asn Ser Thr Arg
145         150         155         160
Ser Asn Ala Gln Asn Ala Pro Trp Val Met Val Gly Gly Ser Tyr Ser
165         170         175
Gly Ala Leu Thr Ala Trp Thr Glu Ser Val Ala Pro Gly Thr Phe Trp
180         185         190
Ala Tyr His Ala Thr Ser Ala Pro Val Glu Ala Ile Tyr Asp Tyr Trp
195         200         205
Gln Tyr Phe Tyr Pro Ile Gln Gln Gly Met Ala Gln Asn Cys Ser Lys
210         215         220
Asp Val Ser Leu Val Ala Glu Tyr Val Asp Lys Ile Gly Lys Asn Gly
225         230         235         240
Thr Ala Lys Glu Gln Gln Ala Leu Lys Glu Leu Phe Gly Leu Gly Ala
245         250         255
Val Glu His Phe Asp Asp Phe Ala Ala Val Leu Pro Asn Gly Pro Tyr
260         265         270
Leu Trp Gln Asp Asn Asp Phe Ala Thr Gly Tyr Ser Ser Phe Phe Gln
275         280         285
Phe Cys Asp Ala Val Glu Gly Val Glu Ala Gly Ala Ala Val Thr Pro
290         295         300
Gly Pro Glu Gly Val Gly Leu Glu Lys Ala Leu Ala Asn Tyr Ala Asn
305         310         315         320
Trp Phe Asn Ser Thr Ile Leu Pro Asp Tyr Cys Ala Ser Tyr Gly Tyr
325         330         335
Trp Thr Asp Glu Trp Ser Val Ala Cys Phe Asp Ser Tyr Asn Ala Ser
340         345         350
Ser Pro Ile Tyr Thr Asp Thr Ser Val Gly Asn Ala Val Asp Arg Gln
355         360         365
Trp Glu Trp Phe Leu Cys Asn Glu Pro Phe Phe Tyr Trp Gln Asp Gly
370         375         380
Ala Pro Glu Gly Thr Ser Thr Ile Val Pro Arg Leu Val Ser Ala Ser
385         390         395         400
Tyr Trp Gln Arg Gln Cys Pro Leu Tyr Phe Pro Glu Thr Asn Gly Tyr
405         410         415
Thr Tyr Gly Ser Ala Lys Gly Lys Asn Ala Ala Thr Val Asn Ser Trp
420         425         430
Thr Gly Gly Trp Asp Met Thr Arg Asn Thr Thr Arg Leu Ile Trp Thr
435         440         445

```

ES 2 744 898 T3

Asn Gly Gln Tyr Asp Pro Trp Arg Asp Ser Gly Val Ser Ser Thr Phe
 450 455 460
 Arg Pro Gly Gly Pro Leu Ala Ser Thr Ala Asn Glu Pro Val Gln Ile
 465 470 475
 Ile Pro Gly Gly Phe His Cys Ser Asp Leu Tyr Met Ala Asp Tyr Tyr
 485 490 495
 Ala Asn Glu Gly Val Lys Lys Val Val Asp Asn Glu Val Lys Gln Ile
 500 505 510
 Lys Glu Trp Val Glu Glu Tyr Tyr Ala
 515 520

<210> 6

<211> 1566

<212> ADN

5 <213> Secuencia artificial

<220>

<221> fuente

<222> 1..1566

10 <223> /organismo="Secuencia artificial" /nota=" Endoproteasa especifica de prolina de Aspergillus niger que comprende una secuencia señal de PMeA de A. niger" /tipo_mol="AND sin asignar"

<400> 6

atggtcaagt ccatcctggc ctccgtcttc ttcgctgcc ctgctcttgc tgcaaggcct 60
 cgtctcgttc ccaagcccgt ttctcgtccc gccagctcca agtccgctgc tactactggt 120
 gaggcctact ttgaacagct gttggaccac cacaaccctg agaagggtac tttctcgsaa 180
 agatactggt ggagcaccga gtactgggggt ggtcccggat cccccgttgt cctgttcaact 240
 cccggtgagg tcagcgtga tggctacgag ggttatctga ccaacgagac tctcaccggt 300
 gtctacgccc aggagattca ggggtcgtgc atcctgatcg aacaccgata ctggggtgac 360
 tcgtctccct acgaggtgct gaacgccgag actctccagt acttgaccct cgaccaggct 420
 atccttgata tgacctactt cgccgaaacc gtcaagctcc agtttgacaa ctccaccgcg 480
 tccaacgctc agaacgctcc ttgggttatg gtcggcggca gctacagcgg tgctctgact 540
 gcttgaccg agtccgttgc tcccggcacc ttctgggctt accacgccac ctctgctcct 600
 gttgaggcca tctacgacta ctggcaatac ttctacccca ttcagcaggg tatggctcag 660
 aactgctcca aagatgtctc tctttagca gaatacgtcg acaagatcgg caagaacggc 720
 actgccaagg agcaacaggc tctgaaggag cttttcggcc taggagcagt ggagcacttc 780
 gacgacttcg ccgctgttct gcccaacggt ccttacctct ggcaagacaa cgactttgcc 840
 accggttact cttctttctt ccagttctgt gatgccgtcg aggggtgtcga ggctggtgct 900
 gccgtcacc cgggtcctga aggtgttggt ctggaaaagg cccttgctaa ctacgcgaac 960
 tggttcaact ctaccatcct ccccgattac tgcgccagct acggctactg gactgacgag 1020
 tggtcctcgt cctgcttcga ctctacaac gcctcctctc ctatatacac cgacaccagc 1080
 gttgtaacg ccgtcgaccg tcagtgggag tggttcctct gcaatgagcc cttcttttac 1140

ES 2 744 898 T3

tggcaggacg	gtgccccga	gggtacttca	acgatagtac	cccgcttagt	gtccgcctcc	1200
tactggcagc	gtcaatgtcc	gttgtacttc	cccgagacta	acggttacac	ctacggctcc	1260
gccaagggaa	agaacgccgc	caccgtcaac	agctggaccg	gtggctggga	catgaccctg	1320
aacaccaccc	gtctgatctg	gacgaacggc	caatacgacc	cctggcgtga	ctccggtgtc	1380
tcttccacct	tccgccccgg	tggtcccctc	gcttcgaccg	ccaacgagcc	cgtccagata	1440
ataccgggtg	gtttcattg	ctccgacctc	tacatggcag	actactacgc	caacgagggc	1500
gtcaagaagg	ttgtcgacaa	cgaagtcaaa	caaatcaagg	agtggggtga	ggaatactac	1560
gcgtaa						1566

REIVINDICACIONES

1. Un polipéptido que tiene actividad de endoproteasa específica de prolina, seleccionado del grupo que consiste en
 - i. un polipéptido que comprende una secuencia de polipéptido maduro de SEQ ID NO: 2;
 - 5 ii. un polipéptido que tiene al menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad con una secuencia de polipéptido maduro de SEQ ID NO: 2;
 - iii. un polipéptido codificado por un ácido nucleico que tiene al menos 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad con una secuencia codificante de polipéptido maduro de SEQ ID NO: 1;

10 en el que la secuencia de polipéptido maduro de SEQ ID NO: 2 comprende los aminoácidos 36 a 526, aminoácidos 37 a 526, aminoácidos 39 a 526, aminoácidos 40 a 526, o aminoácidos 41 a 526 de SEQ ID NO: 2, en el que la metionina en la posición 1 de SEQ ID NO: 2 se cuenta como 1, y en el que el polipéptido tiene al menos 50% de actividad residual después de que el polipéptido se ha mantenido a una temperatura de al menos 70°C durante 15 min., en el que la actividad se mide con acetil AlaAlaPro-para-nitroaloinina (Ac-AAP-pNA) como sustrato.
- 15 2. Una composición que comprende un polipéptido según la reivindicación 1.
3. Una composición según la reivindicación 2, que comprende un soporte, o un excipiente.
4. Un ácido nucleico que codifica una endoproteasa específica de prolina, que tiene al menos 85%, 90%, 95%, 96%, 97%, 98%, o 99% de identidad de secuencia con una secuencia codificante de polipéptido maduro de SEQ ID NO: 1.
5. Un vector de expresión que comprende un ácido nucleico según la reivindicación 4 enlazado operablemente a una o más secuencias de control que dirigen la expresión del polipéptido en una célula hospedante.
- 20 6. Una célula hospedante recombinante que comprende un ácido nucleico según la reivindicación 4, o un vector de expresión según la reivindicación 5, en los que el ácido nucleico es heterólogo a la célula hospedante.
7. Un método para la preparación de un polipéptido según la reivindicación 1, que comprende cultivar una célula hospedante según la reivindicación 6 en un medio de fermentación adecuado, en condiciones que permitan la expresión del polipéptido, y preparar el polipéptido, y opcionalmente además, que comprende recuperar el polipéptido.
- 25 8. Un procedimiento para la preparación de un producto alimentario, que comprende poner en contacto un producto alimentario con un polipéptido según la reivindicación 1, o con una composición según la reivindicación 2 o 3, durante la preparación del producto alimentario, y preparar el producto alimentario o de pienso.
9. Un procedimiento según la reivindicación 8, en el que el producto alimentario es una bebida, preferiblemente una cerveza.
- 30 10. Un procedimiento según la reivindicación 9, en el que la forma intermedia del producto alimentario es una pulpa.
11. Un procedimiento según una cualquiera de las reivindicaciones 8 a 10, en el que el producto alimentario comprende gluten.
12. Uso de un polipéptido según la reivindicación 1, o una composición según la reivindicación 2 o 3, para reducir la turbidez en una bebida.
- 35 13. Uso según la reivindicación 12, en el que la bebida es una cerveza.

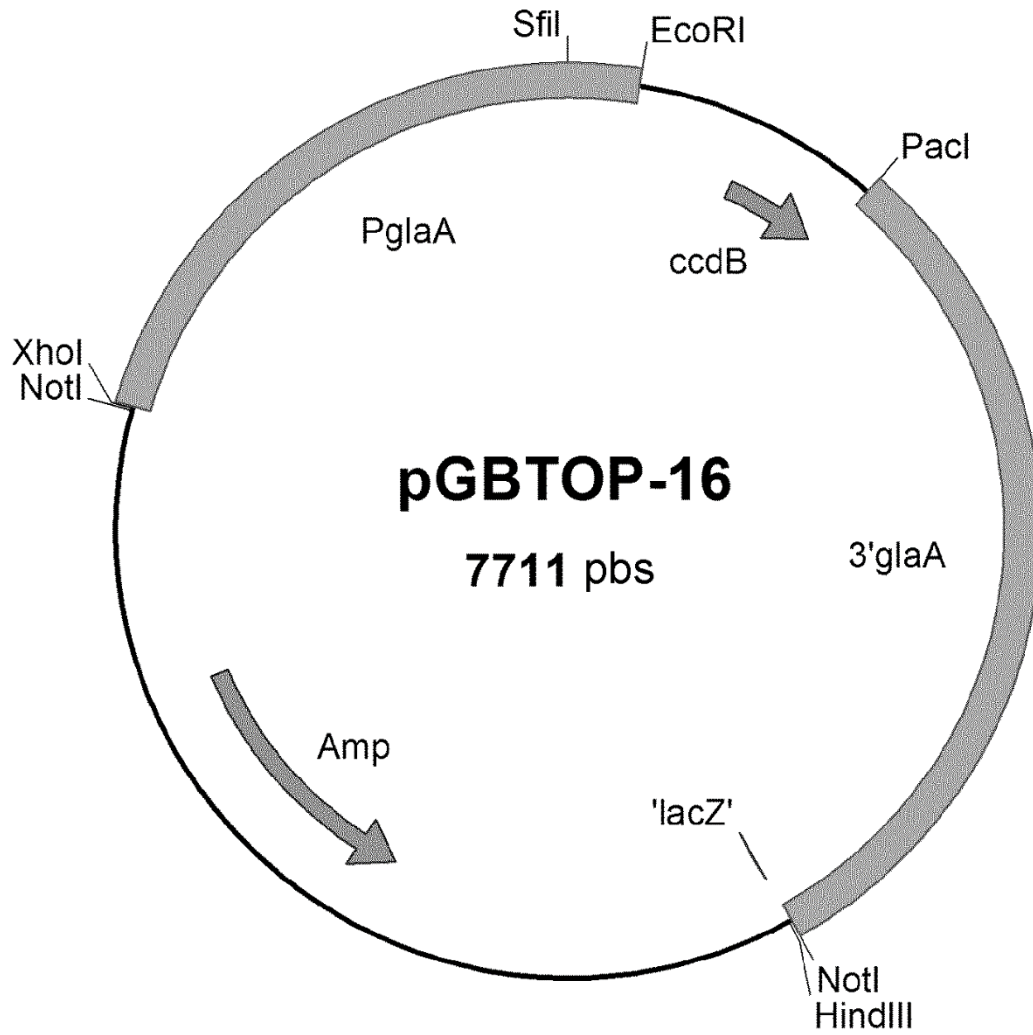


Fig 1