(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 772 032**

(51) Int. Cl.:

C12N 9/26 (2006.01)
C12N 15/82 (2006.01)
C12P 7/16 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **13.12.2013** **PCT/US2013/074957**

(87) Fecha y número de publicación internacional: **26.06.2014** **WO14099653**

(96) Fecha de presentación y número de la solicitud europea: **13.12.2013** **E 13864154 (3)**

(97) Fecha y número de publicación de la concesión europea: **13.11.2019** **EP 2935574**

(54) Título: **Alfa-amilasas y polinucleótidos que las codifican**

(30) Prioridad:

17.12.2012 US 201261738145 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
07.07.2020

(73) Titular/es:

NOVOZYMES A/S (100.0%)
Krogshoejvej 36
2880 Bagsvaerd, DK

(72) Inventor/es:

COWARD-KELLY, GUILLERMO;
FUKUYAMA, SHIRO;
TSUTSUMI, NORIKO y
AYABE, KEIICHI

(74) Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 772 032 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

Alfa-amilasas y polinucleótidos que las codifican

ANTECEDCENTES DE LA INVENCIÓN

Campo de la invención

[0001] La presente invención se refiere a alfa-amilasas, polinucleótidos que codifican las alfa-amilasas, construcciones de ácido nucleico o vectores de expresión que comprenden los polinucleótidos, células hospedadoras recombinantes que comprenden los polinucleótidos, métodos para producir las alfa-amilasas y métodos para usar las alfa-amilasas. En formas de realización de la presente descripción, se mejora la actividad de degradación del almidón crudo.

Descripción de la técnica relacionada

[0002] La degradación enzimática del almidón forma parte de muchos procesos industriales que incluyen la elaboración de cerveza, la producción de jarabes con alto contenido de glucosa o fructosa y la producción de etanol para bebidas o como combustible. En su estado natural, el almidón es bastante resistente a la degradación por muchas enzimas y, por lo tanto, la degradación enzimática industrial del almidón se inicia tradicionalmente mediante un paso de calentamiento en el que el almidón se gelatiniza, lo que hace que el almidón sea más sensible a muchas enzimas. Algunas enzimas pueden actuar sobre el almidón no gelatinizado, y comúnmente se las conoce como poseedoras de actividad degradante del almidón crudo. El uso de estas enzimas permite procesos mejorados, que incluyen, por ejemplo, la reducción del paso de calentamiento en el procesamiento del almidón.

[0003] Las alfa-amilasas (alfa-1,4-glucano 4 glucanohidrolasas, EC. 3.2.1.1) constituyen un grupo de enzimas que cataliza la hidrólisis del almidón y otros oligo y polisacáridos 1,4 glucosídicos lineales y ramificados. Las enzimas alfa-amilasas se han utilizado para una variedad de propósitos industriales diferentes, incluyendo la licuefacción de almidón, la producción de etanol, el desaprestado de textiles, el lavado de textiles, la modificación del almidón en la industria del papel y la pulpa, la elaboración de cerveza y la panadería.

[0004] WO 2010/091221 describe un polipéptido que tiene actividad de alfa-amilasa de *Aspergillus terreus.* La base de datos UniProt XP002576027 revela la secuencia de ácido nucleico del genoma de Q0C881 (*Aspergillus terreus*) para una alfa-amilasa.

[0005] WO 2008/080093 divulga alfa-amilasas y glucoamilasas y su uso en la fabricación de biocombustible.

[0006] Sigue existiendo una necesidad en la técnica de alfa-amilasas mejoradas, incluidas las alfa-amilasas que tienen actividad degradante del almidón crudo.

Resumen de la invención

[0007] La presente invención se refiere a polipéptidos aislados que tienen actividad de alfa-amilasa seleccionados del grupo que consiste en:

(a) un polipéptido que incluye una secuencia de aminoácidos del polipéptido maduro de la secuencia de aminoácidos de la SEQ ID N.º: 2, SEQ ID N.º:4, o SEQ ID N.º:6;
(b) un polipéptido que incluye una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el polipéptido maduro de la secuencia de aminoácidos de la SEQ ID N.º: 2, SEQ ID N.º:4, o SEQ ID N.º:6;
(c) un polipéptido codificado por un polinucleótido con al menos 90% de identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º:3, o SEQ ID N.º:5; y
(d) un polipéptido codificado por un polinucleótido que se hibrida bajo condiciones de astringencia media con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º:3, o SEQ ID N.º:5.

[0008] La presente invención también se refiere a polinucleótidos que codifican los polipéptidos de la presente invención; construcciones de ácido nucleico; vectores de expresión recombinantes; células hospedadoras recombinantes que comprenden los polinucleótidos; y métodos para producir los polipéptidos.

[0009] La presente invención también se refiere a un método para la producción de etanol usando un polipéptido de la presente invención. La presente invención también se refiere a un método para la producción de etanol a partir de almidón no gelatinizado usando un polipéptido de la presente invención.

Definiciones

[0010] **Actividad de alfa-amilasa**: El término "actividad de alfa-amilasa" se define en este documento como una actividad que cataliza la endohidrólisis de enlaces (1,4)-alfa-D-glucosídicos en polisacáridos que contienen tres o más unidades de D-glucosa unidas por enlaces alfa-(1,4). El término "actividad de alfa-amilasa" corresponde a las enzimas agrupadas en E.C. 3.2.1.1. Para los fines de la presente invención, la actividad de alfa-amilasa se determina de acuerdo con el procedimiento descrito en la sección "Ejemplo".

[0011] **Variante alélica**: el término "variante alélica" significa cualquiera de dos o más formas alternativas de un gen que ocupan el mismo locus cromosómico. La variación alélica surge naturalmente a través de la mutación y puede resultar en polimorfismos dentro de las poblaciones. Las mutaciones génicas pueden ser silenciosas (sin cambio en el polipéptido codificado) o pueden codificar polipéptidos con secuencias de aminoácidos modificadas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

[0012] **ADNc**: el término "ADNc" significa una molécula de ADN que se puede preparar por transcripción inversa a partir de una molécula de ARNm empalmado maduro obtenida de una célula eucariota. El ADNc carece de secuencias de intrones que pueden estar presentes en el ADN genómico correspondiente. El transcrito de ARN primario inicial es un precursor del ARNm que se procesa a través de una serie de pasos, incluido el empalme, antes de aparecer como ARNm empalmado maduro.

[0013] **Secuencia codificante**: el término "secuencia codificante" significa un polinucleótido, que especifica directamente la secuencia de aminoácidos de un polipéptido. Los límites de la secuencia codificante están determinados generalmente por un marco de lectura abierto, que empieza normalmente con un codón de inicio como ATG, GTG o TTG y acaba con un codón de terminación tal como TAA, TAG o TGA. La secuencia codificante puede ser un ADN genómico, ADNc, ADN sintético o una combinación de estos.

[0014] **Secuencias de control**: el término "secuencias de control" significa las secuencias de ácido nucleico necesarias para la expresión de un polinucleótido que codifica un polipéptido maduro de la presente invención. Cada secuencia de control puede ser nativa (es decir, del mismo gen) o exógena (de un gen diferente) respecto al polinucleótido que codifica el polipéptido o nativa o exógena entre sí. Tales secuencias de control incluyen, pero de forma no limitativa, un líder, una secuencia de poliadenilación, una secuencia de propéptido, un promotor, una secuencia de péptido señal y un terminador de la transcripción. Como mínimo, las secuencias de control incluyen un promotor y señales de parada transcripcional y traduccional. Las secuencias de control pueden estar provistas de conectores con el fin de introducir sitios de restricción específicos que faciliten el ligamiento de las secuencias de control con la región codificante del polinucleótido que codifica una variante.

[0015] **Expresión**: el término "expresión" incluye cualquier paso implicado en la producción de un polipéptido incluyendo, pero de forma no limitativa, la transcripción, la modificación postranscripcional, la traducción, la modificación postraduccional y la secreción.

[0016] **Vector de expresión**: el término "vector de expresión" significa una molécula lineal o circular de ADN que comprende un polinucleótido que codifica un polipéptido y que está unido operativamente a secuencias de control que permiten su expresión.

[0017] **Fragmento**: el término "fragmento" significa un polipéptido de la presente invención que tiene uno o más (por ejemplo, varios) aminoácidos ausentes del extremo amino y/o carboxilo terminal de un dominio o polipéptido maduro, en donde el fragmento tiene actividad de alfa-amilasa. En un aspecto, un fragmento contiene al menos 497 residuos de aminoácidos, al menos 526 residuos de aminoácidos, o al menos 555 residuos de aminoácidos de las SEQ ID N.º: 2 o 6. En otro aspecto, un fragmento contiene al menos 528 residuos de aminoácidos, al menos 559 residuos de aminoácidos, o al menos 590 residuos de aminoácidos de la SEQ ID N.º: 4.

[0018] **Condiciones de astringencia alta**: el término "condiciones de astringencia alta" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42 °C en SSPE 5X, SDS al 0,3%, 200 microgramos/ml de ADN de esperma de salmón cortado y desnaturalizado, y formamida al 50%, después de procedimientos estándar de transferencia Southern durante 12 a 24 horas. El material de soporte finalmente se lava tres veces, cada una durante 15 minutos, usando SSC 2X, SDS al 0,2% a 65 °C.

[0019**] Célula hospedadora**: el término "célula hospedadora" significa cualquier tipo de célula que es susceptible a la transformación, la transfección, la transducción y similares con una construcción de ácido nucleico o un vector de expresión que comprende un polinucleótido de la presente invención. El término "célula hospedadora" abarca cualquier descendiente de una célula progenitora que no sea idéntica a la célula progenitora debido a las mutaciones que ocurren durante la replicación.

[0020] **Aislado**: El término "aislado" significa una sustancia en una forma o ambiente que no se da en la naturaleza. Los ejemplos no limitativos de sustancias aisladas incluyen (1) cualquier sustancia que no se produce de forma natural, (2) cualquier sustancia incluyendo, pero sin limitarse a, cualquier enzima, variante, ácido nucleico, proteína, péptido o cofactor, que se extraiga al menos parcialmente de uno o más o todos los constituyentes naturales con los que está asociado/a en la naturaleza; (3) cualquier sustancia modificada por la mano del hombre en relación

con esa sustancia encontrada en la naturaleza; o (4) cualquier sustancia modificada al aumentar la cantidad de la sustancia con respecto a otros componentes con los que está asociada naturalmente (por ejemplo, copias múltiples de un gen que codifica la sustancia; uso de un promotor más fuerte que el promotor asociado naturalmente con el gen que codifica la sustancia). Una sustancia aislada puede estar presente en una muestra de caldo de fermentación. Por ejemplo, el polipéptido de la presente invención puede usarse en aplicaciones industriales en forma de un producto de caldo de fermentación, es decir, el polipéptido de la presente invención es un componente de un caldo de fermentación usado como producto en aplicaciones industriales (por ejemplo, producción de etanol). El producto de caldo de fermentación, además del polipéptido de la presente invención, comprenderá ingredientes adicionales utilizados en el proceso de fermentación, tales como, por ejemplo, células (incluyendo las células hospedadoras que contienen el gen que codifica el polipéptido de la presente invención que se utilizan para producir el polipéptido de interés), restos celulares, biomasa, medios de fermentación y/o productos de fermentación. El caldo de fermentación puede someterse de manera opcional a uno o más pasos de purificación (incluyendo filtración) para eliminar o reducir uno o más componentes de un proceso de fermentación. Por consiguiente, un polipéptido "aislado" de la presente invención puede estar presente en dicho producto de caldo de fermentación.

[0021] **Polipéptido maduro**: el término "polipéptido maduro" significa un polipéptido en su forma final tras la traducción y cualquier modificación postraduccional, tal como el procesamiento N-terminal, el truncamiento C-terminal, la glicosilación, la fosforilación, etc. Se sabe en la técnica que una célula hospedadora puede producir una mezcla de dos o más polipéptidos maduros diferentes (es decir, con un aminoácido C-terminal y/o N-terminal diferente) expresado por el mismo polinucleótido. También se sabe en la técnica que diferentes células hospedadoras procesan los polipéptidos de manera diferente y, por lo tanto, una célula hospedadora que expresa un polinucleótido puede producir un polipéptido maduro diferente (por ejemplo, con un amino C-terminal y/o N-terminal diferente) en comparación con otra célula hospedadora que expresa el mismo polinucleótido.

[0022] En un aspecto, el polipéptido maduro son los aminoácidos 1 a 585 de la SEQ ID N.º: 2 basado en SignalP (Nielsen et al., 1997, Protein Engineering 10: 1-6) que predice que los aminoácidos -1 a -21 de la SEQ ID N.º: 2 son un péptido señal.

[0023] En otro aspecto, el polipéptido maduro son los aminoácidos 1 a 622 de la SEQ ID N.º: 4 basado en SignalP (Nielsen et al., 1997, Protein Engineering 10: 1-6) que predice que los aminoácidos -1 a -21 de la SEQ ID N.º: 4 son un péptido señal.

[0024] En otro aspecto, el polipéptido maduro son los aminoácidos 1 a 607 de la SEQ ID N.º: 6 basado en SignalP (Nielsen et al., 1997, Protein Engineering 10: 1-6) que predice que los aminoácidos -1 a -21 de la SEQ ID N.º: 6 son un péptido señal.

[0025] **Secuencia codificante del polipéptido maduro**: El término "secuencia codificante del polipéptido maduro" significa un polinucleótido que codifica un polipéptido maduro que tiene actividad de alfa-amilasa.

[0026] **Condiciones de astringencia media**: el término "condiciones de astringencia media" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42 °C en SSPE 5X, SDS al 0,3%, 200 microgramos/ml de ADN de esperma de salmón cortado y desnaturalizado, y formamida al 35%, después de procedimientos estándar de transferencia Southern durante 12 a 24 horas. El material de soporte finalmente se lava tres veces, cada una durante 15 minutos, usando SSC 2X, SDS al 0,2% a 55 °C.

[0027] **Condiciones de astringencia media-alta**: el término "condiciones de astringencia media-alta" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42°C en SSPE 5X, SDS al 0,3%, 200 microgramos/ml de ADN de esperma de salmón cortado y desnaturalizado, y formamida al 35%, después de procedimientos estándar de transferencia Southern durante 12 a 24 horas. El material de soporte finalmente se lava tres veces, cada una durante 15 minutos, usando SSC 2X, SDS al 0,2% a 60°C.

[0028] **Construcción de ácido nucleico**: El término "construcción de ácido nucleico" significa una molécula de ácido nucleico, monocatenaria o bicatenaria, que se aísla de un gen natural o que se modifica para que contenga segmentos de ácidos nucleicos de una manera que de otra manera no existiría en la naturaleza o que es sintética, que comprende una o más secuencias de control.

[0029] **Operativamente unido**: el término "operativamente unido/a" significa una configuración en la que una secuencia de control se coloca en una posición apropiada respecto a la secuencia codificante de un polinucleótido de manera que la secuencia de control dirija la expresión de la secuencia codificante.

[0030] **Identidad de secuencia**: la relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos se describe por el parámetro "identidad de secuencia".

[0031] Para los fines de la presente invención, el grado de identidad de secuencia entre dos secuencias de aminoácidos se determina usando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol.

48: 443-453) como se implementa en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferiblemente la versión 5.0.0 o posterior. Los parámetros utilizados son una penalización por apertura de espacio de 10, una penalización por extensión de espacio de 0,5 y la matriz de sustitución EBLOSUM62 (versión de EMBOSS de BLOSUM62). El resultado de Needle etiquetado "identidad más larga" (obtenido usando la opción -nobrief) se usa como el porcentaje de identidad y se calcula de la siguiente manera:

(Residuos idénticos x 100)/(Longitud del alineamiento – Número total de espacios en el alineamiento)

[0032] Para los fines de la presente invención, el grado de identidad de secuencia entre dos secuencias de desoxirribonucleótidos se determina usando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, supra) como se implementa en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000,supra), preferiblemente la versión 5.0.0 o posterior. Los parámetros opcionales usados son una penalización por apertura de espacio de 10, una penalización por extensión de espacio de 0,5 y la matriz de sustitución EDNAFULL (versión de EMBOSS de NCBI NUC4.4). El resultado de Needle etiquetado "identidad más larga" (obtenido usando la opción -nobrief) se usa como el porcentaje de identidad y se calcula de la siguiente manera:

(Desoxirribonucleótidos idénticos x 100)/(Longitud del alineamiento – Número total de espacios en el alineamiento)

[0033] **Subsecuencia**: el término "subsecuencia" significa un polinucleótido que tiene uno o más (por ejemplo, varios) nucleótidos ausentes del extremo 5' y/o 3' terminal de una secuencia codificante de polipéptido maduro; en donde la subsecuencia codifica un fragmento que tiene actividad alfa-amilasa.

[0034] **Variante**: el término "variante" significa un polipéptido con actividad de alfa-amilasa que comprende una modificación, es decir, una sustitución, inserción y/o deleción, en una o más (varias) posiciones. Una sustitución significa una sustitución de un aminoácido que ocupa una posición por un aminoácido diferente; una deleción significa la eliminación de un aminoácido que ocupa una posición; y una inserción significa la adición de uno o más, por ejemplo 1-3, aminoácidos adyacentes a un aminoácido que ocupa una posición.

[0035] **Condiciones de astringencia muy alta**: el término "condiciones de astringencia muy alta" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42 °C en SSPE 5X, SDS al 0,3%, 200 microgramos/ml de ADN de esperma de salmón cortado y desnaturalizado, y formamida al 50%, siguiendo procedimientos estándar de transferencia Southern durante 12 a 24 horas. El material de soporte finalmente se lava tres veces, cada una durante 15 minutos, usando SSC 2X, SDS al 0,2% a 70 °C.

Descripción detallada de la invención

Polipéptidos que tienen actividad de alfa-amilasa

[0036] La presente divulgación se refiere a nuevas secuencias de alfa-amilasa.

[0037] En el primer aspecto, la presente invención se refiere a un polipéptido aislado que tiene actividad de alfa-amilasa, seleccionado del grupo que consiste en:

(a) un polipéptido que incluye una secuencia de aminoácidos del polipéptido maduro de la secuencia de aminoácidos de la SEQ ID N.º: 2, SEQ ID N.º:4, o SEQ ID N.º:6;
(b) un polipéptido que incluye una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el polipéptido maduro de la secuencia de aminoácidos de la SEQ ID N.º: 2, SEQ ID N.º:4, o SEQ ID N.º:6;
(c) un polipéptido codificado por un polinucleótido con al menos 90% de identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º:3, o SEQ ID N.º:5; y
(d) un polipéptido codificado por un polinucleótido que se hibrida bajo condiciones de astringencia media con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º:3, o SEQ ID N.º:5.

[0038] Las nuevas secuencias de alfa-amilasa incluyen el polipéptido maduro de la SEQ ID N.º: 2, el polipéptido maduro de la SEQ ID N.º:4 y el polipéptido maduro de la SEQ ID N.º:6. Un polipéptido maduro de la SEQ ID N.º:2 se muestra también como la secuencia de aminoácidos de la SEQ ID N.º: 1 (residuos 1-585). Un polipéptido maduro de la SEQ ID N.º:4 se muestra también como la secuencia de aminoácidos de la SEQ ID N.º: 3 (residuos 1-622). Un polipéptido maduro de la SEQ ID N.º:6 se muestra como la secuencia de aminoácidos de la SEQ ID N.º: 5 (residuos 1-607).

[0039] En una forma de realización, la presente invención se refiere a polipéptidos aislados que tienen una identidad de secuencia con el polipéptido maduro de la SEQ ID N.º: 2 de al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%,

o 100%, que tienen actividad de alfa-amilasa. En un aspecto, los polipéptidos difieren del polipéptido maduro de la SEQ ID N.º: 2 en no más de 10 aminoácidos, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, o 9 aminoácidos.

[0040] En otro aspecto, el polipéptido comprende o consiste en el polipéptido maduro de la SEQ ID N.º: 2. Un polipéptido de la presente invención comprende preferiblemente o consiste en la secuencia de aminoácidos de la SEQ ID N.º: 2 o una variante alélica de esta.

[0041] En una forma de realización, la presente invención se refiere a polipéptidos aislados que tienen una identidad de secuencia con el polipéptido maduro de la SEQ ID N.º:4 de al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100%, que tienen actividad de alfa-amilasa. En un aspecto, los polipéptidos difieren del polipéptido maduro de la SEQ ID N.º:4 en no más de 10 aminoácidos, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, o 9 aminoácidos.

[0042] En otro aspecto, el polipéptido comprende o consiste en el polipéptido maduro de la SEQ ID N.º: 4.

[0043] En una forma de realización, la presente invención se refiere a polipéptidos aislados que tienen una identidad de secuencia con el polipéptido de la SEQ ID N.º:6 de al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100%, que tienen actividad de alfa-amilasa. En un aspecto, los polipéptidos difieren del polipéptido maduro de la SEQ ID N.º:6 en no más de 10 aminoácidos, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, o 9 aminoácidos.

[0044] En otro aspecto, el polipéptido comprende o consiste en el polipéptido maduro de la SEQ ID N.º: 6.

[0045] La presente invención se refiere a un polipéptido aislado que tiene actividad de alfa-amilasa codificado por un polinucleótido que se hibrida en condiciones de astringencia media, condiciones de astringencia media-alta, condiciones de astringencia alta, o condiciones de astringencia muy alta con una secuencia de ácidos nucleicos que codifica de la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º:3, o SEQ ID N.º:5, o (ii) el complemento en toda su longitud de la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º:3, o SEQ ID N.º:5.

[0046] En otra forma de realización, la presente invención se refiere a un polipéptido aislado que tiene actividad de alfa-amilasa codificado por un polinucleótido con una identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º:3 o SEQ ID N.º:5 de al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100%.

[0047] Los cambios de aminoácidos pueden ser de naturaleza menor, es decir, sustituciones o inserciones conservadoras de aminoácidos que no afectan significativamente al plegamiento y/o la actividad de la proteína. Los ejemplos de sustituciones conservadoras están dentro de los grupos de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina) y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Las sustituciones de aminoácidos que generalmente no alteran la actividad específica son conocidas en la técnica y se describen, por ejemplo, por H. Neurath y R.L. Hill, 1979, en The Proteins, Academic Press, Nueva York. Las sustituciones comunes son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu y Asp/Gly.

[0048] Alternativamente, los cambios de aminoácidos son de tal naturaleza que las propiedades físicoquímicas de los polipéptidos se ven alteradas. Por ejemplo, los cambios de aminoácidos pueden mejorar la termoestabilidad del polipéptido, alterar la especificidad de sustrato, cambiar el pH óptimo, y similares.

[0049] Las sustituciones, deleciones y/o inserciones de aminoácidos simples o múltiples se pueden realizar y probar utilizando métodos conocidos de mutagénesis, recombinación y/o barajado, seguidos de un procedimiento de cribado relevante, como los descritos por Reidhaar-Olson y Sauer, 1988, Science 241: 53-57; Bowie y Sauer, 1989, Proc. Natl. Acad. Sci. Estados Unidos 86: 2152-2156; WO 95/17413; o WO 95/22625. Otros métodos que pueden usarse incluyen PCR propensa a error, visualización de fagos (por ejemplo, Lowman et al., 1991, Biochemistry 30: 10832-10837; patente de EE. UU. 5,223,409; WO 92/06204) y mutagénesis dirigida (Derbyshire et al., 1986, Gene 46: 145; Ner et al., 1988, DNA 7: 127).

[0050] Los métodos de mutagénesis/barajado se pueden combinar con métodos de cribado automatizados de alto rendimiento para detectar la actividad de polipéptidos mutagenizados clonados expresados por células hospedadoras (Ness et al., 1999, Nature Biotechnology 17: 893-896). Las moléculas de ADN mutagenizadas que codifican polipéptidos activos pueden recuperarse de las células hospedadoras y secuenciarse rápidamente usando métodos estándar en la técnica. Estos métodos permiten la determinación rápida de la importancia de los residuos de aminoácidos individuales en un polipéptido.

[0051] Las variantes de los polipéptidos pueden construirse sobre la base del polinucleótido presentado como la secuencia codificante del polipéptido maduro, por ejemplo, una subsecuencia de este, y/o mediante la introducción de sustituciones de nucleótidos que no dan como resultado un cambio en la secuencia de aminoácidos del polipéptido, pero que corresponden al uso de codones del organismo hospedador destinados a la producción de la enzima, o mediante la introducción de sustituciones de nucleótidos que pueden dar lugar a una secuencia de aminoácidos diferente. Para una descripción general de la sustitución de nucleótidos, véase, por ejemplo, Ford et al., 1991, Protein Expression and Purification 2: 95-107.

[0052] Los aminoácidos esenciales de un polipéptido pueden identificarse de acuerdo con procedimientos conocidos en la técnica, tales como mutagénesis dirigida o mutagénesis de barrido de alanina (Cunningham y Wells, 1989, Science 244: 1081-1085). En la última técnica, se introducen mutaciones de alanina individuales en cada residuo de la molécula, y se somete a las moléculas mutantes resultantes a una prueba para determinar la actividad de alfa-amilasa para identificar residuos de aminoácidos que son fundamentales para la actividad de la molécula. Véase también Hilton et al., 1996, J. Biol. Chem 271: 4699-4708. El sitio activo de la enzima u otra interacción biológica también se puede determinar mediante el análisis físico de la estructura, según se determina mediante técnicas tales como resonancia magnética nuclear, cristalografía, difracción de electrones o etiquetado de fotoafinidad, junto con la mutación de aminoácidos de sitios de contacto putativos. Véase, por ejemplo, de Vos et al., 1992, Science 255: 306-312; Smith et al., 1992, J. Mol. Biol. 224: 899-904; Wlodaver et al., 1992, FEBS Lett. 309: 59-64. La identidad de los aminoácidos esenciales también se puede inferir de una alineación con un polipéptido relacionado.

Polinucleótidos

[0053] La presente invención también se refiere a polinucleótidos aislados que codifican un polipéptido de la presente invención, como se describe en este caso.

[0054] En una forma de realización, la presente invención se refiere a un polinucleótido con una identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º: 3 o SEQ ID N.º: 5, de al menos 90 %, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o al menos 99%.

Construcciones de ácido nucleico

[0055] La presente invención también se refiere a construcciones de ácido nucleico que comprenden un polinucleótido de la presente invención operativamente unido a una o más secuencias de control que dirigen la expresión de la secuencia de codificación en una célula hospedadora adecuada en condiciones compatibles con las secuencias de control.

[0056] Un polinucleótido puede manipularse de varias maneras para lograr la expresión del polipéptido. La manipulación del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar polinucleótidos que utilizan métodos de ADN recombinante son ampliamente conocidas en la técnica.

[0057] La secuencia de control puede ser un promotor, un polinucleótido que es reconocido por una célula hospedadora para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. El promotor contiene secuencias de control transcripcional que median en la expresión del polipéptido. El promotor puede ser cualquier polinucleótido que muestre actividad transcripcional en la célula hospedadora, incluyendo los promotores mutantes, truncados e híbridos, y puede obtenerse de genes que codifican polipéptidos extracelulares o intracelulares, ya sean homólogos o heterólogos a la célula hospedadora.

[0058] Algunos ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácido nucleico de la presente invención en una célula hospedadora bacteriana son los promotores obtenidos de gen de la alfa-amilasa de *Bacillus amyloliquefaciens* (*amyQ*), del gen de la alfa-amilasa de *Bacillus licheniformis* (*amyL*), del gen de la penicilinasa de *Bacillus licheniformis* (*penP*), del gen de la amilasa maltogénica de *Bacillus stearothermophilus* (*amyM*), del gen de la levansacarasa de *Bacillus subtilis* (*sacB*), los genes *xylA* y *xylB* de *Bacillus subtilis*, el gen *cryIIIA* de *Bacillus Thuuringiensis* (Agaisse y Lereclus, 1994, Molecular Microbiology 13: 97-107), el operón *lac* de *E. coli,* el promotor *trc* de *E. coli* (Egon et al., 1988, Gene 69: 301-315), el gen de agarasa de *Streptomyces coelicolor* (*dagA*), y el gen de beta-lactamasa procariota (Villa-Kamaroff et al., 1978, Proc. Natl. Acad. Sci. Estados Unidos 75: 3727-3731), así como el promotor tac (DeBoer et al., 1983, Proc. Natl. Acad. Sci. Estados Unidos 80: 21-25) Otros promotores se describen en "Useful proteins from recombinant bacteria" en Gilbert et al., 1980, Scientific American 242: 74-94; y en Sambrook *et al.* 1989, *supra.* Algunos ejemplos de promotores en tándem se describen en WO 99/43835.

[0059] Algunos ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácido nucleico de la presente invención en una célula hospedadora fúngica filamentosa son promotores obtenidos de los genes para acetamidasa de *Aspergillus nidulans*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en

ácido de *Aspergillus niger*, glucoamilasa de *Aspergillus niger* o *Aspergillus awamori* (glaA), TAKA amilasa de *Aspergillus oryzae*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), amiloglucosidasa de *Fusarium venenatum* (WO 00/56900), Daria de *Fusarium venenatum* (WO 00/56900), Quinn de *Fusarium venenatum* (WO 00/56900), lipasa de *Rhizomucor miehei*, proteinasa aspártica de *Rhizomucor miehei*, 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*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, al igual que el promotor NA2-tpi (un promotor modificado a partir de un gen de alfa-amilasa neutra de *Aspergillus* donde el líder no traducido se ha sustituido por un líder no traducido a partir de un gen de triosa fosfato isomerasa de *Aspergillus*; los ejemplos no limitativos incluyen promotores modificados a partir de un gen de alfa-amilasa neutra de *Aspergillus niger* en el que el líder no traducido se ha sustituido por un líder no traducido a partir de un gen de triosa fosfato isomerasa de *Aspergillus nidulans* o *Aspergillus oryzae*); y promotores mutantes, truncados e híbridos de estos.

[0060] En una levadura hospedadora, los promotores útiles se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), galactoquinasa de *Saccharomyces cerevisiae* (GAL1), alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH1, ADH2/GAP), triosa fosfato isomerasa de *Saccharomyces cerevisiae* (TPI), metalotioneína de *Saccharomyces cerevisiae* (CUP1) y 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para células de levadura hospedadoras están descritos por Romanos et al., 1992, Yeast 8: 423-488.

[0061] La secuencia de control también puede ser un terminador de la transcripción, que es reconocido por una célula hospedadora para terminar la transcripción. El terminador está unido operativamente al extremo 3' terminal del polinucleótido que codifica el polipéptido. Cualquier terminador que sea funcional en la célula hospedadora puede usarse en la presente invención.

[0062] Los terminadores preferidos para las células hospedadoras bacterianas se obtienen de los genes para proteasa alcalina de *Bacillus clausii* (aprH), alfa-amilasa de *Bacillus licheniformis* (amyL), y ARN ribosómico de *Escherichia coli* (rrnB).

[0063] Los terminadores preferidos para células hospedadoras de hongos filamentosos se obtienen de los genes para antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, alfa-glucosidasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae*, y proteasa de tipo tripsina de *Fusarium oxysporum*.

[0064] Los terminadores preferidos para células hospedadoras de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo C de *Saccharomyces cerevisiae* (CYC1), y gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae*. Romanos et al. 1992, supra describen otros terminadores útiles para células hospedadoras de levadura.

[0065] La secuencia de control también puede ser una región estabilizadora de ARNm en dirección hacia 5' de un promotor y en dirección hacia 3' de la secuencia de codificación de un gen que aumenta la expresión del gen.

[0066] Se obtienen ejemplos de regiones estabilizadoras de ARNm adecuadas a partir de un gen *cryIIIA* de *Bacillus thuringiensis* (WO 94/25612) y Gen SP82 de *Bacillus subtilis* (Hue et al., 1995, Journal of Bacteriology 177: 3465-3471).

[0067] La secuencia de control también puede ser un líder, una región no traducida de un ARNm que es importante para la traducción por la célula hospedadora. El líder está operativamente unido al extremo 5' del polinucleótido que codifica el polipéptido. Se puede usar cualquier líder que sea funcional en la célula hospedadora.

[0068] Los líderes preferidos para células hospedadoras de hongos filamentosos se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

[0069] Los líderes adecuados para las células hospedadoras de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*, factor alfa de *Saccharomyces cerevisiae* y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP).

[0070] La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia unida operativamente al extremo 3' del polinucleótido y, cuando se transcribe, es reconocida por la célula hospedadora como una señal para agregar residuos de poliadenosina al ARNm transcrito. Se puede usar cualquier secuencia de poliadenilación que sea funcional en la célula hospedadora.

[0071] Las secuencias de poliadenilación preferidas para células hospedadoras fúngicas filamentosas se obtienen de los genes para antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, alfa-glucosidasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae*, y proteasa de tipo tripsina de *Fusarium oxysporum*.

[0072] Las secuencias de poliadenilación útiles para células hospedadoras de levadura son descritas por Guo y Sherman, 1995, Mol. Cellular Biol. 15: 5983-5990.

[0073] La secuencia de control también puede ser una región de codificación de péptido señal que codifica un péptido señal unido al extremo N terminal de un polipéptido y dirige el polipéptido hacia la vía secretora de la célula. El extremo 5' de la secuencia de codificación del polinucleótido puede contener inherentemente una secuencia de codificación de péptido señal unida de manera natural en el marco de lectura de traducción con el segmento de la secuencia de codificación que codifica el polipéptido. Alternativamente, el extremo 5' de la secuencia de codificación puede contener una secuencia de codificación de péptido señal que es exógena a la secuencia de codificación. Puede requerirse una secuencia codificante de péptido señal exógena cuando la secuencia codificante no contiene de manera natural una secuencia codificante de péptido señal. Alternativamente, una secuencia codificante de péptido señal exógena puede reemplazar simplemente la secuencia codificante de péptido señal natural para mejorar la secreción del polipéptido. Sin embargo, puede usarse cualquier secuencia de codificación de péptido señal que dirija el polipéptido expresado hacia la vía secretora de una célula hospedadora.

[0074] Las secuencias de codificación de péptidos señal eficaces para células hospedadoras bacterianas son las secuencias de codificación de péptidos señal obtenidas de los genes para amilasa maltogénica de *Bacillus* NCIB 11837, subtilisina de *Bacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, alfa-amilasa de *Bacillus stearothermophilus*, proteasas neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM), y prsA de *Bacillus subtilis*. Otros péptidos señal adicionales son descritos por Simonen y Palva, 1993, Microbiological Reviews 57: 109-137.

[0075] Las secuencias de codificación de péptidos señal efectivas para células hospedadoras fúngicas filamentosas son las secuencias de codificación de péptidos señal obtenidas de los genes para amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae*, celulasa de *Humicola insolens*, endoglucanasa V de *Humicola insolens*, lipasa de *Humicola lanuginosa*, y proteinasa aspártica de *Rhizomucor miehei*.

[0076] Los péptidos señal útiles para células hospedadoras de levadura se obtienen de los genes para factor alfa de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras secuencias codificantes de péptido señal útiles son descritas por Romanos et al., 1992, supra.

[0077] La secuencia de control también puede ser una secuencia codificante de propéptido que codifica un propéptido posicionado en el extremo N-terminal de un polipéptido. El polipéptido resultante se conoce como una proenzima o propolipéptido (o un zimógeno en algunos casos). Un propolipéptido generalmente es inactivo y se puede convertir en un polipéptido activo por escisión catalítica o autocatalítica del propéptido a partir del propolipéptido. La secuencia codificante del propéptido se puede obtener de los genes para proteasa alcalina de *Bacillus subtilis* (aprE), proteasa neutra de *Bacillus subtilis* (nprT), lacasa de *Myceliophthora thermophila* (WO 95/33836), proteinasa aspártica de *Rhizomucor miehei*, y factor alfa de *Saccharomyces cerevisiae*.

[0078] Cuando están presentes las secuencias tanto del péptido señal como del propéptido, la secuencia del propéptido se coloca junto al extremo N-terminal de un polipéptido y la secuencia del péptido señal se coloca junto al extremo N-terminal de la secuencia del propéptido.

[0079] También puede ser deseable agregar secuencias reguladoras que regulen la expresión del polipéptido en relación con el crecimiento de la célula hospedadora. Algunos ejemplos de sistemas reguladores son aquellos que hacen que la expresión del gen se active o se desactive en respuesta a un estímulo químico o físico, incluyendo la presencia de un compuesto regulador. Los sistemas reguladores en los sistemas procariotas incluyen los sistemas del operón *lac, tac*, y *trp*. En levaduras, se puede usar el sistema ADH2 o el sistema GAL1. En hongos filamentosos, se pueden usar el promotor de glucoamilasa de *Aspergillus niger,* el promotor de TAKA alfa-amilasa de *Aspergillus oryzae*, y promotor de glucoamilasa de *Aspergillus oryzae*. Otros ejemplos de secuencias reguladoras son las que permiten la amplificación génica. En sistemas eucarióticos, estas secuencias reguladoras incluyen el gen de dihidrofolato reductasa que se amplifica en presencia de metotrexato, y los genes de metalotioneína que se amplifican con metales pesados. En estos casos, el polinucleótido que codifica el polipéptido estaría operativamente enlazado con la secuencia reguladora.

Vectores de expresión

[0080] La presente invención también se refiere a vectores de expresión recombinantes que comprenden un polinucleótido de la presente invención, un promotor y señales de parada transcripcionales y traduccionales. Las diversas secuencias de nucleótidos y de control pueden unirse para producir un vector de expresión recombinante

que puede incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución del polinucleótido que codifica el polipéptido en dichos sitios. Alternativamente, el polinucleótido puede expresarse insertando el polinucleótido o una construcción de ácido nucleico que comprende el polinucleótido en un vector apropiado para la expresión. Al crear el vector de expresión, la secuencia de codificación se ubica en el vector de modo que la secuencia de codificación se une operativamente con las secuencias de control apropiadas para la expresión.

[0081] El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus) que se puede someter convenientemente a procedimientos de ADN recombinante y puede provocar la expresión del polinucleótido. La elección del vector dependerá típicamente de la compatibilidad del vector con la célula hospedadora en la que el vector se va a introducir. El vector puede ser un plásmido lineal o circular cerrado.

[0082] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para asegurar la autorreplicación. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula hospedadora, se integra en el genoma y se replica junto con el/los cromosoma(s) en el/los que se ha integrado. Además, se puede usar un único vector o plásmido o dos o más vectores o plásmidos que, juntos, contienen el ADN total que se desea introducir en el genoma de la célula hospedadora, o un transposón.

[0083] El vector contiene preferiblemente uno o más marcadores seleccionables que permiten una selección fácil de células transformadas, transfectadas, transducidas o similares. Un marcador seleccionable es un gen cuyo producto proporciona resistencia a biocidas o resistencia viral, resistencia a metales pesados, prototrofia a auxótrofos, y similares.

[0084] Los marcadores seleccionables para su uso en una célula hospedadora fúngica filamentosa incluyen, pero de forma no limitativa, *amdS* (acetamidasa), *argB* (ornitina-carbamoiltransferasa), *bar* (fosfonitricina acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrato-reductasa), *pyrG* (orotidina-5'-fosfato-descarboxilasa), *sC* (sulfato adeniltransferasa), y *trpC* (antranilato sintasa), así como equivalentes de los mismos. Para el uso en una célula de *Aspergillus* se prefieren los genes *pyrG* y *amdS de Aspergillus nidulans* o *Aspergillus oryzae* y un gen *bar* de *Streptomyces hygroscopicus*.

[0085] El vector contiene preferiblemente un elemento(s) que permite(n) la integración del vector en el genoma de la célula hospedadora o la replicación autónoma del vector en la célula independiente del genoma.

[0086] Para la integración en el genoma de la célula hospedadora, el vector puede basarse en la secuencia del polinucleótido que codifica el polipéptido o cualquier otro elemento del vector para la integración en el genoma mediante recombinación homóloga o no homóloga. Alternativamente, el vector puede contener polinucleótidos adicionales para dirigir la integración por recombinación homóloga en el genoma de la célula hospedadora en una ubicación precisa en el/los cromosoma(s). Para aumentar la probabilidad de integración en una ubicación precisa, los elementos de integración deberían contener un número suficiente de ácidos nucleicos, como 100 a 10 000 pares de bases, 400 a 10 000 pares de bases y 800 a 10 000 pares de bases, que tengan un alto grado de identidad de secuencia con la secuencia de destino correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que sea homóloga con la secuencia de destino en el genoma de la célula hospedadora. Además, los elementos integracionales pueden ser polinucleótidos no codificantes o codificantes. Por otro lado, el vector puede integrarse en el genoma de la célula hospedadora por recombinación no homóloga.

[0087] Para la replicación autónoma, el vector puede comprender además un origen de replicación que permita que el vector se replique de forma autónoma en la célula hospedadora en cuestión. El origen de replicación puede ser cualquier replicador de plásmidos que medie la replicación autónoma que funciona en una célula. El término "origen de replicación" o "replicador de plásmidos" significa un polinucleótido que permite que un plásmido o vector se replique *in vivo*.

[0088] Algunos ejemplos de orígenes de replicación útiles en una célula fúngica filamentosa son AMA1 y ANS1 (Gems et al., 1991, Gene 98: 61-67; Cullen et al., 1987, Nucleic Acids Res. 15: 9163-9175; WO 00/24883). El aislamiento del gen AMA1 y la construcción de plásmidos o vectores que comprenden el gen pueden realizarse de acuerdo con los métodos descritos en WO 00/24883.

[0089] Algunos ejemplos de orígenes de replicación para usar en una célula hospedadora de levadura son el origen de replicación de 2μ, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6.

[0090] Se puede insertar más de una copia de un polinucleótido de la presente invención en una célula hospedadora para aumentar la producción de un polipéptido. Se puede obtener un aumento en el número de copias del polinucleótido integrando al menos una copia adicional de la secuencia en el genoma de la célula hospedadora o incluyendo un gen marcador amplificable seleccionable con el polinucleótido donde las células que

contienen copias amplificadas del gen marcador seleccionable, y de este modo, las copias adicionales del polinucleótido, se pueden seleccionar cultivando las células en presencia del agente seleccionable apropiado.

[0091] Los procedimientos utilizados para ligar los elementos descritos anteriormente para construir los vectores de expresión recombinantes de la presente invención son ampliamente conocidos por un experto en la materia (véase, por ejemplo, Sambrook *et al.* 1989, *supra*).

Células hospedadoras

[0092] La presente invención también se refiere a células hospedadoras recombinantes, que comprenden un polinucleótido de la presente invención operativamente unido a una o más secuencias de control que dirigen la producción de un polipéptido de la presente invención. Una construcción o vector que comprende un polinucleótido se introduce en una célula hospedadora de modo que la construcción o el vector se mantenga como un integrante cromosómico o como un vector extracromosómico autorreplicante como se ha descrito anteriormente. La célula hospedadora puede ser cualquier célula útil en la producción recombinante de un polipéptido de la presente invención, por ejemplo una procariota (célula bacteriana) o una eucariota (como una célula de mamífero, de insecto, vegetal o fúngica).

[0093] En un aspecto preferido, la célula hospedadora es una célula fúngica. "Hongos", como se usa en el presente documento, incluye el filo Ascomycota, Basidiomycota, Chytridiomycota y Zygomycota, así como los Oomycota y todos los hongos mitospóricos (según lo definido por Hawksworth et al., En, Ainsworth and Bisby's Dictionary of The Fungi, octava edición, 1995, CAB International, University Press, Cambridge, Reino Unido.).

[0094] La célula hospedadora fúngica puede ser una célula de levadura. "Levadura", como se usa en el presente documento, incluye levaduras ascosporógenas (Endomycetales), levaduras basidioesporógenas y levaduras que pertenece a los hongos imperfectos (Blastomicetos). Dado que la clasificación de la levadura puede cambiar en el futuro, para los fines de esta invención, la levadura se definirá como se describe en Biology and Activities of Yeast (Skinner, Passmore y Davenport, editors, Soc. App. Bacteriol. Symposium Series No. 9, 1980).

[0095] La célula hospedadora de levadura puede ser una célula de *Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces,* o *Yarrowia*, como una célula de *Kluyveromyces lactis, Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis, Saccharomyces oviformis,* o *Yarrowia lipolytica*.

[0096] La célula hospedadora fúngica puede ser una célula fúngica filamentosa. Los "hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota (según lo definido por Hawksworth *et al.* 1995, *supra*). Los hongos filamentosos se caracterizan generalmente por una pared micelial compuesta de quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. El crecimiento vegetativo es por alargamiento hifal y el catabolismo del carbono es obligatoriamente aeróbico. En contraste, el crecimiento vegetativo de levaduras como *Saccharomyces cerevisiae* es por gemación de un tallo unicelular y el catabolismo del carbono puede ser fermentativo.

[0097] La célula hospedadora fúngica filamentosa puede ser una célula de *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysosporium, Coprinus, Coriolus, Cryptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes,* o *Trichoderma.*

[0098] Por ejemplo, la célula hospedadora filamentosa fúngica puede ser una célula de *Aspergillus awamori, Aspergillus foetidus, Aspergillus fumigatus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Chrysosporium inops, Chrysosporium keratinophilum, Chrysosporium lucknowense, Chrysosporium merdarium, Chrysosporium pannicola, Chrysosporium queenslandicum, Chrysosporium tropicum, Chrysosporium zonatum, Coprinus cinereus, Coriolus hirsutus, Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Humicola insolens, Humicola lanuginosa, Mucormiehei, Myceliophthora thermophila, Neurospora crassa, Penicillium purpurogenum, Phanerochaete chrysosporium, Phlebia radiata, Pleurotus eryngii, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei,* o *Trichoderma viride.*

[0099] Las células fúngicas se pueden transformar por un proceso que implica la formación de protoplastos, la transformación de los protoplastos y la regeneración de la pared celular de una manera conocida per se. Los procedimientos adecuados para la transformación de células hospedadoras de *Aspergillus* y *Trichoderma* se

describen en EP 238023, Yelton et al., 1984, Proc. Natl. Acad. Sci. EE. UU 81: 1470-1474, y Christensen et al., 1988, Bio/Technology 6: 1419-1422. Métodos adecuados para transformar especies de *Fusarium* son descritos por Malardier et al., 1989, gen 78: 147-156, y WO 96/00787. Las levaduras se pueden transformar utilizando los procedimientos descritos por Becker y Guarente, en Abelson, J.N. Y Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., Nueva York; Ito et al., 1983, J. Bacteriol. 153: 163; e Hinnen et al., 1978, Proc. Natl. Acad. Sci. EE. UU. 75: 1920.

Métodos de producción

[0100] La presente invención también se refiere a métodos de producción de un polipéptido de la presente invención, que comprende:

a) el cultivo de una célula hospedadora que comprende el polinucleótido que codifica el polipéptido de la presente invención operativamente enlazado a una o más secuencias de control que dirigen la producción del polipéptido en condiciones propicias para la producción del polipéptido; y
(b) la recuperación del polipéptido.

[0101] En un aspecto preferido, la célula es una célula de *Aspergillus*, como, *Aspergillus foetidus, Aspergillus fumigatus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Aspergillus terreus.*

[0102] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención, que comprende (a) cultivar una célula hospedadora recombinante de la presente invención en condiciones propicias para la producción del polipéptido; y (b) recuperar el polipéptido.

[0103] Las células hospedadoras se cultivan en un medio nutritivo adecuado para la producción del polipéptido usando métodos conocidos en la técnica. Por ejemplo, la célula puede cultivarse mediante cultivo en matraz con agitación, o fermentación a pequeña o gran escala (incluyendo fermentaciones continuas, discontinuas, discontinuas o en estado sólido) en fermentadores de laboratorio o industriales en un medio adecuado y en condiciones que permitan expresar y/o aislar el polipéptido. El cultivo se lleva a cabo en un medio nutritivo adecuado que comprende fuentes de carbono y nitrógeno y sales inorgánicas, utilizando procedimientos conocidos en la técnica. Los medios adecuados están disponibles de proveedores comerciales o pueden prepararse de acuerdo con las composiciones publicadas (por ejemplo, en catálogos de la American Type Culture Collection). Si el polipéptido se secreta en el medio nutritivo, el polipéptido se puede recuperar directamente del medio. Si el polipéptido no se secreta, se puede recuperar a partir de lisados celulares.

[0104] El polipéptido puede detectarse usando métodos conocidos en la técnica que son específicos para los polipéptidos. Estos métodos de detección incluyen, pero no se limitan a, el uso de anticuerpos específicos, la formación de un producto enzimático o la desaparición de un sustrato enzimático. Por ejemplo, se puede usar un ensayo enzimático para determinar la actividad del polipéptido.

[0105] El polipéptido puede recuperarse usando métodos conocidos en la técnica. Por ejemplo, el polipéptido puede recuperarse del medio nutritivo mediante procedimientos convencionales que incluyen, pero no se limitan a, recolección, centrifugación, filtración, extracción, secado por pulverización, evaporación o precipitación.

[0106] El polipéptido se puede purificar mediante una variedad de procedimientos conocidos en la técnica incluyendo, pero no limitados a, cromatografía (por ejemplo, de intercambio iónico, de afinidad, hidrofóbica, cromatoenfoque, y de exclusión de tamaño), procedimientos electroforéticos (por ejemplo, isoelectroenfoque preparativo), solubilidad diferencial (por ejemplo, precipitación de sulfato amónico), SDS-PAGE, o extracción (véase, por ejemplo, Protein Purification, Janson and Ryden, editors, VCH Publishers, Nueva York, 1989) para obtener polipéptidos sustancialmente puros.

[0107] En un aspecto alternativo, el polipéptido no se recupera, sino que una célula hospedadora de la presente invención que expresa el polipéptido se usa como fuente del polipéptido.

Composiciones

[0108] Se describen composiciones que comprenden un polipéptido de la presente invención. Dichas composiciones con polipéptido pueden prepararse de acuerdo con métodos conocidos en la técnica y pueden estar en forma de una composición líquida o seca. El polipéptido para incluir en la composición puede estabilizarse de acuerdo con métodos conocidos en la técnica.

[0109] La composición con polipéptido puede estar en forma de granulado, un microgranulado o un polvo. Los métodos para preparar tales composiciones son ampliamente conocidos en la técnica.

[0110] La composición con polipéptido puede estar en forma de un producto de caldo de fermentación. El producto de caldo de fermentación, además del polipéptido de la presente invención, comprenderá ingredientes adicionales

utilizados en el proceso de fermentación, tales como, por ejemplo, células (incluyendo las células hospedadoras que contienen el gen que codifica el polipéptido de la presente invención que se utilizan para producir el polipéptido de interés), restos celulares, biomasa, medios de fermentación y/o productos de fermentación. El caldo de fermentación se puede someter opcionalmente a uno o más pasos de purificación (incluyendo filtración) para eliminar o reducir uno más componentes de un proceso de fermentación.

[0111] La composición con polipéptido puede comprender además una enzima seleccionada del grupo que comprende otra alfa-amilasa (EC 3.2.1.1), una beta-amilasa (EC 3.2.1.2), una glucoamilasa (EC3.2.1.3), una pululanasa (EC 3.2.1.41), una fitasa (EC 3.1.2.28) y una proteasa (EC 3.4.), y sus combinaciones (por ejemplo, el polipéptido de la presente invención y una glucoamilasa, o el polipéptido de la presente invención y una glucoamilasa y una proteasa)

[0112] En un aspecto particular, la composición con polipéptido comprende además una glucoamilasa. El polipéptido se puede combinar con glucoamilasa comercial, tal como la preparación de glucoamilasa suministrada por Novozymes A/S como SPIRIZYME FUEL. La glucoamilasa también se puede derivar a partir de una cepa de *Aspergillus* sp., tal como *Aspergillus niger*, o a partir de una cepa de *Talaromyces* sp. y derivarse en particular de *Talaromyces leycettanus* como la glucoamilasa descrita en la patente de EE. UU. nº Reg. 32,153, *Talaromyces duponti* y/o *Talaromyces thermopiles* tal como las glucoamilasas descritas en la patente de EE. UU. nº 4,587,215 y derivarse más preferiblemente de *Talaromyces emersonii*. En un aspecto, la glucoamilasa se deriva de la cepa de *Talaromyces emersonii* CBS 793.97 y/o con la secuencia descrita como la SEQ ID N.º: 7 en WO 99/28448. En otro aspecto, la actividad de glucoamilasa se deriva a partir de una cepa del género *Trametes*, preferiblemente *Trametes cingulata*. Otras glucoamilasas incluyen la glucoamilasa con la secuencia de aminoácidos del polipéptido maduro de la SEQ ID N.º: 2 de WO 2006/069289. Glucoamilasa también puede incluir las glucoamilasas del género *Pachykytospora*, preferiblemente *Pachykytospora papyracea* o la cepa de *E. Coli* depositada en DSMZ con el nº DSM 17105, e incluye la glucoamilasa con la secuencia de aminoácidos del polipéptido maduro del polipéptido maduro de la SEQ ID N.º: 5 de WO 2006/069289. Otras glucoamilasas incluyen las que tienen una secuencia de aminoácidos que tiene al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o incluso al menos 99% de homología con la secuencia de aminoácidos anteriormente mencionada.

[0113] La dosificación de la composición con polipéptido de la invención y otras condiciones en las cuales se usa la composición pueden determinarse en función de métodos conocidos en la técnica.

Usos

[0114] La presente invención también se refiere a métodos para usar los polipéptidos que tienen actividad de alfa-amilasa, o composiciones de estos.

[0115] El polipéptido de la presente invención o la composición que comprende un polipéptido de la presente invención se puede usar en la conversión de almidón, la conversión de almidón a azúcar y la producción de etanol, etc., por ejemplo, en la licuefacción y/o sacarificación de un almidón gelatinizado, un almidón granular o un almidón parcialmente gelatinizado. Un almidón parcialmente gelatinizado es un almidón que en cierta medida está gelatinizado, es decir, en el que parte del almidón se ha hinchado y gelatinizado irreversiblemente y parte del almidón todavía está presente en un estado granular.

[0116] El polipéptido de la invención o la composición que comprende un polipéptido de la presente invención puede usarse en un proceso para licuar un almidón gelatinizado, un almidón granular o un sustrato de almidón parcialmente gelatinizado en medio acuoso con el polipéptido de la presente invención.

[0117] Un uso preferido de un polipéptido de la presente invención es en un proceso de fermentación para producir glucosa y/o maltosa adecuada para la conversión en un producto de fermentación por un organismo de fermentación, preferiblemente una levadura. Dichos procesos de fermentación incluyen un proceso para producir etanol como combustible o para bebidas (alcohol potable), un proceso para producir una bebida, un proceso para producir compuestos orgánicos deseados, como ácido cítrico, ácido itacónico, ácido láctico, ácido glucónico, gluconato de sodio, gluconato de calcio, gluconato de potasio, glucono delta lactona o eritorbato de sodio; cetonas; aminoácidos, como el ácido glutámico (monoglutaminato de sodio), pero también compuestos más complejos como antibióticos, como la penicilina, la tetraciclina; enzimas; vitaminas, tales como riboflavina, B12, betacaroteno; hormonas, que son difíciles de producir sintéticamente.

[0118] En una forma de realización preferida, el polipéptido de la presente invención se usa en un proceso que comprende fermentación para producir un producto de fermentación (por ejemplo, etanol), a partir de un almidón gelatinizado. Tal proceso para producir etanol a partir de almidón gelatinizado por fermentación comprende: (i) licuar el almidón gelatinizado con un polipéptido con actividad alfa-amilasa de la presente invención; (ii) sacarificación de la pasta licuada obtenida; (iii) fermentar el material obtenido en el paso (ii) en presencia de un organismo de fermentación. Opcionalmente, el proceso comprende además la recuperación del etanol. La

sacarificación y la fermentación pueden llevarse a cabo como un proceso de sacarificación y fermentación simultáneas (proceso SSF).

[0119] En otra forma de realización preferida, el polipéptido de la presente invención se usa en un proceso que comprende fermentación para producir un producto de fermentación, por ejemplo, etanol, a partir de un almidón no gelatinizado ("crudo"). Tal proceso para producir etanol a partir de material que contiene almidón no gelatinizado por fermentación comprende: (i) poner en contacto el almidón no gelatinizado con un polipéptido con actividad alfa-amilasa de la presente invención para degradar el almidón no gelatinizado; (ii) sacarificar la pasta obtenida; (iii) fermentar el material obtenido en el paso (ii) en presencia de un organismo de fermentación. Opcionalmente, el proceso comprende además la recuperación del etanol. La sacarificación y la fermentación pueden llevarse a cabo como un proceso de sacarificación y fermentación simultáneas (proceso SSF).

[0120] El material que contiene almidón utilizado en los métodos de la presente invención puede ser cualquier material vegetal que contenga almidón. Los materiales que contienen almidón preferidos se seleccionan del grupo que consiste en: tubérculos, raíces y granos enteros; y cualquier combinación de estos. En una forma de realización, el material que contiene almidón se obtiene a partir de cereales. El material que contiene almidón puede, por ejemplo, seleccionarse de los grupos que consisten en maíz, raquis de maíz, trigo, cebada, yuca, sorgo, centeno, zahína y patata; o cualquier combinación de estos. Cuando el producto de fermentación es etanol, el material que contiene almidón es preferiblemente granos enteros o al menos principalmente granos enteros. La materia prima también puede consistir o comprender un producto secundario del procesamiento de almidón.

[0121] En formas de realización adicionales, el polipéptido de la presente invención también puede ser útil para el desaprestado de textiles, tejidos o prendas mediante el tratamiento de un textil, tejido o prenda con un polipéptido de la presente invención, para la producción de un producto o masa de panadería, mediante el tratamiento de una masa con un polipéptido de la presente invención, y opcionalmente su horneado, como ingrediente en un proceso de producción de detergente y de pulpa y papel mediante el tratamiento de una pulpa de fabricación de papel con un polipéptido de la presente invención.

[0122] La presente invención se describe a continuación mediante los ejemplos siguientes, que no deberían ser interpretados como una limitación del alcance de la invención.

Ejemplos

Ensayos sobre la actividad de alfa-amilasa

1. Ensayo de Phadebas

[0123] La actividad de la alfa-amilasa se determina mediante un método que emplea comprimidos Phadebas® como sustrato. Los comprimidos de Phadebas (Phadebas® Amylase Test, suministrado por Pharmacia Diagnostic) contienen un polímero de almidón de color azul insoluble reticulado, que se ha mezclado con albúmina de suero bovino y una sustancia tampón y se ha transformado en comprimidos.

[0124] Para cada medición individual se suspende un comprimido en un tubo que contiene 5 ml de tampón Britton-Robinson 50 mM (ácido acético 50 mM, ácido fosfórico 50 mM, ácido bórico 50 mM, CaCl 0,1 mM_2, pH ajustado al valor de interés con NaOH). La prueba se realiza en un baño de agua a la temperatura de interés. La alfa-amilasa que se desea analizar se diluye en x ml de tampón Britton-Robinson 50 mM. Se agrega 1 ml de esta solución de alfa-amilasa a los 5 ml de tampón Britton-Robinson 50 mM. El almidón es hidrolizado por la alfa-amilasa, lo que proporciona fragmentos azules solubles. La absorbancia de la solución azul resultante, medida espectrofotométricamente a 620 nm, es una función de la actividad alfa-amilasa.

[0125] Es importante que la absorbancia medida de 620 nm después de 10 o 15 minutos de incubación (tiempo de prueba) esté en el rango de 0,2 a 2,0 unidades de absorbancia a 620 nm. En este rango de absorbancia hay linealidad entre actividad y absorbancia (ley de Lambert-Beer). Por lo tanto, la dilución de la enzima debe ajustarse para cumplir este criterio. En un conjunto específico de condiciones (temperatura, pH, tiempo de reacción, condiciones de tampón), 1 mg de una alfa-amilasa dada hidrolizará una cierta cantidad de sustrato y se producirá un color azul. La intensidad del color se mide a 620 nm. La absorbancia medida es directamente proporcional a la actividad específica (actividad/mg de proteína alfa-amilasa pura) de la alfa-amilasa en cuestión en el conjunto de condiciones dadas.

2. Método alternativo

[0126] La actividad de alfa-amilasa se determina mediante un método que emplea el sustrato PNP-G_7. PNP-G_7, que es una abreviatura de p-nitrofenil-alfa,D-maltoheptaósido, es un oligosacárido bloqueado que puede ser escindido por una endo-amilasa. Después de la escisión, la alfa-glucosidasa incluida en un kit disponible en el mercado digiere el sustrato para liberar una molécula de PNP libre que tiene un color amarillo y, por lo tanto, puede

medirse por espectofometría visible a λ = 405 nm (400-420 nm). Kits que contienen PNP-G_7 el sustrato y la alfa-glucosidasa están disponibles comercialmente en Roche y en otros.

[0127] Para preparar la solución de reactivo, se añaden 10 ml de solución de sustrato/tampón a 50 ml de solución de enzima/tampón según lo recomendado por el fabricante. El ensayo se realiza transfiriendo muestras de 20 micro l a una placa de microtitulación de 96 pocillos e incubándolas a 25 °C. Se agregan 200 micro l de una solución de reactivo preequilibrada a 25 °C. La solución se mezcla y se incuba previamente 1 minuto y la absorción se mide cada 30 segundos durante 4 minutos a una DO de 405 nm en un lector de ELISA.

[0128] La pendiente de la curva de absorción dependiente del tiempo es directamente proporcional a la actividad de la alfa-amilasa en cuestión en el conjunto de condiciones dado.

Ejemplo 1

[0129] Se evaluó una alfa-amilasa de la presente invención (SEQ ID N.º: 2) en un ensayo de fermentación de almidón crudo y se comparó tanto con una alfa-amilasa híbrida (descrita en WO 2006/069290 como poseedora del dominio catalítico de *Rhizomucor pusillus* (SEQ ID N.º: 20), el enlazador de *Aspergillus niger* (SEQ ID N.º: 72) y el dominio de unión a carbohidratos de *Aspergillus niger* (SEQ ID N.º: 96)) como con la alfa-amilasa de *Aspergillus terreus* (se muestra como la SEQ ID N.º: 2 en WO 2010/091221).

Materiales y métodos

[0130] Se agregaron aproximadamente 405 g de maíz amarillo "diente de caballo" (obtenido de Hawkeye Renewables de Shell Rock, Iowa; molido internamente) a 595 g de agua del grifo y se determinó que el nivel de sólidos secos (DS) era de 34,42% de DS. Esta mezcla se complementó con 3 ppm de penicilina y 1000 ppm de urea. La suspensión se ajustó a pH 4,5 con H_2SO_4 al 40%. Se añadieron aproximadamente 5 g de esta suspensión a tubos de 15 ml. A cada tubo se dosificó DK193 AMG purificada (AMG de *Trametes cingulata* divulgada en WO 2006/069289 como la SEQ ID N.º: 2) a 0,0801 mg de EP/g de DS y las alfa-amilasas se dosificaron a 0,0225 mg de EP/g de DS. Las dosis enzimáticas reales se basaron en el peso exacto de la suspensión de maíz en cada tubo de acuerdo con la siguiente fórmula:

$$\text{Dosis enz. (µL)} = \frac{\text{Dosis enz. final (mg/g } de\ DS) \ x\ Peso\ de\ la\ masa\ (g)\ x\ Contenido\ de\ sólidos\ secos\ (\%\ de\ DS)}{\text{Conc. solución madre (mg/mL) x1000}}$$

[0131] Se añadió agua a cada tubo para llevar el volumen agregado total (enzima + agua) al 2% del peso inicial de la pasta. Esta corrección de volumen lleva todos los tubos del experimento al mismo porcentaje total de sólidos, haciendo que las concentraciones de etanol sean directamente comparables entre los tratamientos. Después de la adición de enzima y agua, se agregaron 200 µL de propagación de levadura (0,024 g de levadura Fermentis Ethanol Red, incubada durante la noche a 32 °C en 50 ml de pasta de maíz licuada filtrada y 5,1 µL de Spirizyme Plus AMG) a cada tubo.

[0132] Los tubos se incubaron en una sala con temperatura controlada a 32 °C y se realizaron seis fermentaciones replicadas de cada tratamiento. Todos los tubos se agitaron a las 24 y 48 horas. Una muestra se sacrificó para el análisis por HPLC a las 24 horas, dos a las 48 horas y tres a las 70 horas. La preparación de HPLC consistió en detener la reacción mediante la adición de 50 µl de 40% de H_2SO_4, centrifugado durante 10 minutos a 1462xg, y filtrado a través de un filtro de 0,45 µm. Las muestras se almacenaron a 4 °C. Se usó un sistema 1100 HPLC de Agilent™ junto con un refractómetro diferencial para determinar las concentraciones de etanol y oligosacáridos. La columna de separación era una columna de exclusión de iones BioRad™ Aminex HPX-87H (300 mm x 7,8 mm).

[0133] Los datos se analizaron en JMP (SAS, Cary, NC). Los valores atípicos se eliminaron en función de la prueba F ($p < 0,05$). Los tratamientos se compararon con el control con el test HSD de Tukey-Kramer ($p < 0,05$).

Resultados y discusión

[0134] Como se muestra en la Tabla 1, en estas condiciones experimentales, una alfa-amilasa de la presente invención (SEQ ID N.º: 2) funcionó mejor que la alfa-amilasa híbrida (WO 2006/069290) que mostraba una mejora del 2,2 % en el punto de tiempo de 70 horas en comparación con la alfa-amilasa híbrida (WO 2006/069290) y también mejor que la alfa-amilasa de *Aspergillus terreus.*

Tabla 1

Tratamiento (70 h)	*Rendimiento de etanol*
Invención (SEQ ID N.º: 2)	102,21%
Alfa-amilasa híbrida de WO 2006/069290	100,00%
A. terreus	96,68%

Ejemplo 2

Materiales y métodos

[0135] Se agregaron aproximadamente 405 g de maíz amarillo "diente de caballo" (obtenido de Hawkeye Renewables de Shell Rock, Iowa; molido internamente) a 595 g de agua del grifo y se determinó que el nivel de sólidos secos (DS) era de 34,42% de DS. Esta mezcla se complementó con 3 ppm de penicilina y 1000 ppm de urea. La suspensión se ajustó a pH 4,5 con 40% de H_2SO_4. Se añadieron aproximadamente 5 g de esta suspensión a tubos de 15 ml. Cada tubo se dosificó con DK193 AMG purificada a 0,0623 mg de EP/g de DS y las alfa-amilasas se dosificaron a 0,0175 mg de EP/g de DS. Las dosis enzimáticas reales se basaron en el peso exacto de la suspensión de maíz en cada tubo de acuerdo con la siguiente fórmula:

$$\text{Dosis enz. (µL)} = \frac{\text{Dosis enz. final (mg/g } de\ DS) \ x\ Peso\ de\ la\ masa\ (g)\ x\ Contenido\ sólido\ seco\ (\%\ de\ DS)}{\text{Conc. solución madre (mg/mL) x1000}}$$

[0136] Se añadió agua a cada tubo para llevar el volumen agregado total (enzima + agua) al 2% del peso inicial de la pasta. Esta corrección de volumen lleva todos los tubos del experimento al mismo porcentaje total de sólidos, haciendo que las concentraciones de etanol sean directamente comparables entre los tratamientos. Después de la adición de enzima y agua, se añadieron a cada tubo 200 µL de propagación de levadura (0,024 g de levadura Fermentis Ethanol Red, incubada durante la noche a 32 °C en 50 ml de la pasta de maíz licuada filtrada y 5,1 µL de Spirizyme Plus AMG).

[0137] Los tubos se incubaron en una habitación con temperatura controlada a 32 °C y se realizaron seis fermentaciones replicadas de cada tratamiento. Todos los tubos se sometieron a agitación a las 24 y 48 horas. Una muestra se sacrificó para el análisis por HPLC a las 24 horas, dos a las 48 horas y tres a las 70 horas. La preparación de HPLC consistió en detener la reacción mediante la adición de 50 µl de H_2SO_4 al 40%, centrifugación durante 10 minutos a 1462xg y filtrar a través de un filtro de 0,45 µm. Las muestras se almacenaron a 4 °C. Se usó un sistema Agilent™ HPLC 1100 junto con un refractómetro diferencial para determinar las concentraciones de etanol y oligosacáridos. La columna de separación era una columna de exclusión de iones BioRad ™ Aminex HPX-87H (300 mm x 7,8 mm).

[0138] Los datos se analizaron en JMP (SAS, Cary, NC). Los valores atípicos se eliminaron en función de la prueba F ($p < 0,05$). Los tratamientos se compararon con el control con el test HSD de Tukey-Kramer ($p < 0,05$).

Resultados y discusión

[0139] Como se muestra en la Tabla 2, en estas condiciones experimentales, una alfa-amilasa de la presente invención (SEQ ID N.º: 2) tuvo mejores resultados que la alfa-amilasa de *Aspergillus terreus* que mostró una mejora del 1,5% en el punto de tiempo de 70 horas. Las otras alfa-amilasas (SEQ ID N.º: 4 y SEQ ID N.º: 6) también obtuvieron mejores resultados que la alfa-amilasa de *Aspergillus terreus.*

Tabla 2.

Tratamiento (70 h)	*Rendimiento de etanol*
Invención (SEQ ID N.º: 2)	101,5%
Invención (SEQ ID N.º: 4)	101,1%
Invención (SEQ ID N.º: 6)	100,3%
A. terreus	100,0%

LISTADO DE SECUENCIAS

[0140]

<110> Novozymes A/S Fukuyama, Shiro Tsutsumi, Noriko Coward-Kelly, Guillermo Ayabe, Keiichi

<120> Híbridos de amilasa de Aspergillus terreus

<130> 12322-WO-PCT

<160> 6

<170> versión de PatentIn 3.5

<210> 1
<211> 1821
<212> ADN
<213> Secuencia artificial

<220>
<223> híbrido

<220>
<221> sig_peptide
<222> (1)..(63)

<220>
<221> CDS
<222> (1)..(1818)

<220>
<221> mat_peptide
<222> (64)..(1818)

<400> 1
```
atg aga tta tcg act tcg agt ctc ttc ctt tcc gtg tct ctg ctg ggg         48
Met Arg Leu Ser Thr Ser Ser Leu Phe Leu Ser Val Ser Leu Leu Gly
    -20                 -15                 -10

aag ctg gcc ctc ggg ctg acc cca gca gaa tgg cgc agc cag tca atc         96
Lys Leu Ala Leu Gly Leu Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile
-5              -1  1               5                   10

tac ttc ctg ttg acc gat cgc ttt ggt cga acg gac aat tcc aca act        144
Tyr Phe Leu Leu Thr Asp Arg Phe Gly Arg Thr Asp Asn Ser Thr Thr
            15                  20                  25

gcc gct tgc gat acc tct gac aga gta tac tgc ggt ggt agc tgg cag        192
Ala Ala Cys Asp Thr Ser Asp Arg Val Tyr Cys Gly Gly Ser Trp Gln
        30                  35                  40

gga atc atc aac caa ctc gat tac atc caa ggg atg gga ttc act gcc        240
Gly Ile Ile Asn Gln Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala
    45                  50                  55

atc tgg atc acc ccg gtc act gga cag ttc tac gaa aac acg ggc gac        288
Ile Trp Ile Thr Pro Val Thr Gly Gln Phe Tyr Glu Asn Thr Gly Asp
60                  65                  70                  75

ggc acc tct tac cat gga tac tgg cag cag gac atc tac gat ctc aac        336
Gly Thr Ser Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Asp Leu Asn
                80                  85                  90

tac aac tac gga acg gcc caa gac ctc aag aac cta gcc aat gct ttg        384
Tyr Asn Tyr Gly Thr Ala Gln Asp Leu Lys Asn Leu Ala Asn Ala Leu
            95                  100                 105

cac gag cgc ggc atg tat ttg atg gtt gat gtg gtt gcc aac cac atg        432
His Glu Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met
        110                 115                 120

ggc tat gac gga gcg gga aac acc gtg gac tac agt gtt ttc aac ccc        480
Gly Tyr Asp Gly Ala Gly Asn Thr Val Asp Tyr Ser Val Phe Asn Pro
```

```
    125                 130                 135

ttc tcc tcc tcc tcc tac ttt cac cca tac tgc ctc atc tcc aat tac        528
Phe Ser Ser Ser Ser Tyr Phe His Pro Tyr Cys Leu Ile Ser Asn Tyr
140                 145                 150                 155

gac aac cag acc aat gtt gaa gac tgc tgg ctg ggc gat acc acc gtt        576
Asp Asn Gln Thr Asn Val Glu Asp Cys Trp Leu Gly Asp Thr Thr Val
                160                 165                 170

tcg ctg cca gat ctc gac acg aca agc aca gcc gtg cgg aac atc tgg        624
Ser Leu Pro Asp Leu Asp Thr Thr Ser Thr Ala Val Arg Asn Ile Trp
            175                 180                 185

tac gac tgg gtg gca gat ttg gtc gcc aac tat tct atc gac ggt ctg        672
Tyr Asp Trp Val Ala Asp Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu
        190                 195                 200

cgt gtc gac act gta aaa cac gtc gaa aaa gac ttt tgg ccc ggc tat        720
Arg Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Tyr
    205                 210                 215

aac agc gca gca ggc gtc tac tgt gtc ggc gag gtc tac tca ggc gat        768
Asn Ser Ala Ala Gly Val Tyr Cys Val Gly Glu Val Tyr Ser Gly Asp
220                 225                 230                 235

cct gca tac aca tgc ccg tac cag aac tac atg gac ggt gtg ctc aac        816
Pro Ala Tyr Thr Cys Pro Tyr Gln Asn Tyr Met Asp Gly Val Leu Asn
                240                 245                 250

tat cca att tac tac cag ctt ctc tat gcg ttt gag tcg tcc agc ggc        864
Tyr Pro Ile Tyr Tyr Gln Leu Leu Tyr Ala Phe Glu Ser Ser Ser Gly
            255                 260                 265

agc atc agc gat ctc tac aac atg atc agc tcc gtt gcc tcc agc tgc        912
Ser Ile Ser Asp Leu Tyr Asn Met Ile Ser Ser Val Ala Ser Ser Cys
        270                 275                 280

aag gat ccc aca ctc ctg ggg aac ttc atc gag aac cac gat aac ccc        960
Lys Asp Pro Thr Leu Leu Gly Asn Phe Ile Glu Asn His Asp Asn Pro
    285                 290                 295

cgc ttc gct tcc tac acg agc gac tac tcg cag gct aag aac gtg atc       1008
Arg Phe Ala Ser Tyr Thr Ser Asp Tyr Ser Gln Ala Lys Asn Val Ile
300                 305                 310                 315

acc ttc atc ttc ctg agc gac ggt atc ccc atc gtc tac gcc gga cag       1056
Thr Phe Ile Phe Leu Ser Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln
                320                 325                 330

gaa cag cac tac agc gga ggc agc gac cca gct aac cgc gag gcc acc       1104
Glu Gln His Tyr Ser Gly Gly Ser Asp Pro Ala Asn Arg Glu Ala Thr
            335                 340                 345

tgg ctg tcc gga tac tcc acc agc gct acg ctg tac acc tgg atc gcc       1152
Trp Leu Ser Gly Tyr Ser Thr Ser Ala Thr Leu Tyr Thr Trp Ile Ala
        350                 355                 360

acc aca aac cag atc cgc agc ctg gcg atc tcc aag gac gcg gga tac       1200
Thr Thr Asn Gln Ile Arg Ser Leu Ala Ile Ser Lys Asp Ala Gly Tyr
```

365 370 375

gtg cag gcc aag aac aac ccc ttc tac tcc gac tcc aac acc atc gcc 1248
Val Gln Ala Lys Asn Asn Pro Phe Tyr Ser Asp Ser Asn Thr Ile Ala
380 385 390 395

atg cgc aag ggc acg aca gcc ggc gcg caa gtc atc acc gtc ctc agc 1296
Met Arg Lys Gly Thr Thr Ala Gly Ala Gln Val Ile Thr Val Leu Ser
400 405 410

aac aag ggc gcc tcc ggc agc tcc tac act ctc tct ttg agt ggt acc 1344
Asn Lys Gly Ala Ser Gly Ser Ser Tyr Thr Leu Ser Leu Ser Gly Thr
415 420 425

ggc tac tcc gcc ggc gtg acc ctg gtc gag acg tac acc tgc acc acg 1392
Gly Tyr Ser Ala Gly Val Thr Leu Val Glu Thr Tyr Thr Cys Thr Thr
430 435 440

gtg act gta gac tcg agc ggc aac ctg ccc gtc cca atg aca tcc ggc 1440
Val Thr Val Asp Ser Ser Gly Asn Leu Pro Val Pro Met Thr Ser Gly
445 450 455

ttg ccg cga gtg ttt gtc ccg tcg tcc tgg gtg aat ggg agc gcg ctt 1488
Leu Pro Arg Val Phe Val Pro Ser Ser Trp Val Asn Gly Ser Ala Leu
460 465 470 475

tgc aac ggt gct aca agc ccg ggt ggc tcc tcg ggt agt gtc gag gtc 1536
Cys Asn Gly Ala Thr Ser Pro Gly Gly Ser Ser Gly Ser Val Glu Val
480 485 490

act ttc gac gtt tac gct acc aca gta tat ggc cag aac atc tat atc 1584
Thr Phe Asp Val Tyr Ala Thr Thr Val Tyr Gly Gln Asn Ile Tyr Ile
495 500 505

acc ggt gat gtg agt gag ctc ggc aac tgg aca ccc gcc aat ggt gtt 1632
Thr Gly Asp Val Ser Glu Leu Gly Asn Trp Thr Pro Ala Asn Gly Val
510 515 520

gca ctc tct tct gct aac tac ccc acc tgg agt gcc acg atc gct ctc 1680
Ala Leu Ser Ser Ala Asn Tyr Pro Thr Trp Ser Ala Thr Ile Ala Leu
525 530 535

ccc gct gac acg aca atc cag tac aag tat gtc aac att gac ggc agc 1728
Pro Ala Asp Thr Thr Ile Gln Tyr Lys Tyr Val Asn Ile Asp Gly Ser
540 545 550 555

acc gtc atc tgg gag gat gct atc agc aat cgc gag atc acg acg ccc 1776
Thr Val Ile Trp Glu Asp Ala Ile Ser Asn Arg Glu Ile Thr Thr Pro
560 565 570

gcc agc ggc aca tac acc gaa aaa gac act tgg gat gaa tct tag 1821
Ala Ser Gly Thr Tyr Thr Glu Lys Asp Thr Trp Asp Glu Ser
575 580 585

<210> 2
<211> 606
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 2

Met Arg Leu Ser Thr Ser Ser Leu Phe Leu Ser Val Ser Leu Leu Gly
-20 -15 -10

Lys Leu Ala Leu Gly Leu Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile
-5 -1 1 5 10

Tyr Phe Leu Leu Thr Asp Arg Phe Gly Arg Thr Asp Asn Ser Thr Thr
15 20 25

Ala Ala Cys Asp Thr Ser Asp Arg Val Tyr Cys Gly Gly Ser Trp Gln
30 35 40

Gly Ile Ile Asn Gln Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala
45 50 55

Ile Trp Ile Thr Pro Val Thr Gly Gln Phe Tyr Glu Asn Thr Gly Asp
60 65 70 75

Gly Thr Ser Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Asp Leu Asn
80 85 90

Tyr Asn Tyr Gly Thr Ala Gln Asp Leu Lys Asn Leu Ala Asn Ala Leu
95 100 105

His Glu Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met
110 115 120

Gly Tyr Asp Gly Ala Gly Asn Thr Val Asp Tyr Ser Val Phe Asn Pro
125 130 135

Phe Ser Ser Ser Ser Tyr Phe His Pro Tyr Cys Leu Ile Ser Asn Tyr
140 145 150 155

Asp Asn Gln Thr Asn Val Glu Asp Cys Trp Leu Gly Asp Thr Thr Val
160 165 170

Ser Leu Pro Asp Leu Asp Thr Thr Ser Thr Ala Val Arg Asn Ile Trp
175 180 185

Tyr Asp Trp Val Ala Asp Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu
190 195 200

Arg Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Tyr
205 210 215

Asn Ser Ala Ala Gly Val Tyr Cys Val Gly Glu Val Tyr Ser Gly Asp
220 225 230 235

Pro Ala Tyr Thr Cys Pro Tyr Gln Asn Tyr Met Asp Gly Val Leu Asn
240 245 250

Tyr Pro Ile Tyr Tyr Gln Leu Leu Tyr Ala Phe Glu Ser Ser Ser Gly
255 260 265

Ser Ile Ser Asp Leu Tyr Asn Met Ile Ser Ser Val Ala Ser Ser Cys
270 275 280

Lys Asp Pro Thr Leu Leu Gly Asn Phe Ile Glu Asn His Asp Asn Pro
285 290 295

Arg Phe Ala Ser Tyr Thr Ser Asp Tyr Ser Gln Ala Lys Asn Val Ile
300 305 310 315

Thr Phe Ile Phe Leu Ser Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln
320 325 330

Glu Gln His Tyr Ser Gly Gly Ser Asp Pro Ala Asn Arg Glu Ala Thr
335 340 345

Trp Leu Ser Gly Tyr Ser Thr Ser Ala Thr Leu Tyr Thr Trp Ile Ala
350 355 360

Thr Thr Asn Gln Ile Arg Ser Leu Ala Ile Ser Lys Asp Ala Gly Tyr
365 370 375

Val Gln Ala Lys Asn Asn Pro Phe Tyr Ser Asp Ser Asn Thr Ile Ala
380 385 390 395

Met Arg Lys Gly Thr Thr Ala Gly Ala Gln Val Ile Thr Val Leu Ser
400 405 410

Asn Lys Gly Ala Ser Gly Ser Ser Tyr Thr Leu Ser Leu Ser Gly Thr
415 420 425

Gly Tyr Ser Ala Gly Val Thr Leu Val Glu Thr Tyr Thr Cys Thr Thr
430 435 440

Val Thr Val Asp Ser Ser Gly Asn Leu Pro Val Pro Met Thr Ser Gly
445 450 455

Leu Pro Arg Val Phe Val Pro Ser Ser Trp Val Asn Gly Ser Ala Leu
460 465 470 475

Cys Asn Gly Ala Thr Ser Pro Gly Gly Ser Ser Gly Ser Val Glu Val
480 485 490

Thr Phe Asp Val Tyr Ala Thr Thr Val Tyr Gly Gln Asn Ile Tyr Ile
495 500 505

Thr Gly Asp Val Ser Glu Leu Gly Asn Trp Thr Pro Ala Asn Gly Val
510 515 520

Ala Leu Ser Ser Ala Asn Tyr Pro Thr Trp Ser Ala Thr Ile Ala Leu
525 530 535

Pro Ala Asp Thr Thr Ile Gln Tyr Lys Tyr Val Asn Ile Asp Gly Ser
540 545 550 555

Thr Val Ile Trp Glu Asp Ala Ile Ser Asn Arg Glu Ile Thr Thr Pro
560 565 570

Ala Ser Gly Thr Tyr Thr Glu Lys Asp Thr Trp Asp Glu Ser
575 580 585

<210> 3
<211> 1932
<212> ADN
<213> Secuencia artificial

<220>
<223> Híbrido

<220>
<221> sig_peptide
<222> (1)..(63)

<220>
<221> CDS
<222> (1)..(1929)

<220>
<221> mat_peptide
<222> (64)..(1929)

<400> 3
atg aga tta tcg act tcg agt ctc ttc ctt tcc gtg tct ctg ctg ggg 48

```
Met Arg Leu Ser Thr Ser Ser Leu Phe Leu Ser Val Ser Leu Leu Gly
    -20                 -15                 -10

aag ctg gcc ctc ggg ctg acc cca gca gaa tgg cgc agc cag tca atc      96
Lys Leu Ala Leu Gly Leu Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile
-5              -1  1               5                   10

tac ttc ctg ttg acc gat cgc ttt ggt cga acg gac aat tcc aca act     144
Tyr Phe Leu Leu Thr Asp Arg Phe Gly Arg Thr Asp Asn Ser Thr Thr
            15                  20                  25

gcc gct tgc gat acc tct gac aga gta tac tgc ggt ggt agc tgg cag     192
Ala Ala Cys Asp Thr Ser Asp Arg Val Tyr Cys Gly Gly Ser Trp Gln
        30                  35                  40

gga atc atc aac caa ctc gat tac atc caa ggg atg gga ttc act gcc     240
Gly Ile Ile Asn Gln Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala
    45                  50                  55

atc tgg atc acc ccg gtc act gga cag ttc tac gaa aac acg ggc gac     288
Ile Trp Ile Thr Pro Val Thr Gly Gln Phe Tyr Glu Asn Thr Gly Asp
60                  65                  70                  75

ggc acc tct tac cat gga tac tgg cag cag gac atc tac gat ctc aac     336
Gly Thr Ser Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Asp Leu Asn
                80                  85                  90

tac aac tac gga acg gcc caa gac ctc aag aac cta gcc aat gct ttg     384
Tyr Asn Tyr Gly Thr Ala Gln Asp Leu Lys Asn Leu Ala Asn Ala Leu
            95                  100                 105

cac gag cgc ggc atg tat ttg atg gtt gat gtg gtt gcc aac cac atg     432
His Glu Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met
        110                 115                 120

ggc tat gac gga gcg gga aac acc gtg gac tac agt gtt ttc aac ccc     480
Gly Tyr Asp Gly Ala Gly Asn Thr Val Asp Tyr Ser Val Phe Asn Pro
    125                 130                 135

ttc tcc tcc tcc tcc tac ttt cac cca tac tgc ctc atc tcc aat tac     528
Phe Ser Ser Ser Ser Tyr Phe His Pro Tyr Cys Leu Ile Ser Asn Tyr
140                 145                 150                 155

gac aac cag acc aat gtt gaa gac tgc tgg ctg ggc gat acc acc gtt     576
Asp Asn Gln Thr Asn Val Glu Asp Cys Trp Leu Gly Asp Thr Thr Val
                160                 165                 170

tcg ctg cca gat ctc gac acg aca agc aca gcc gtg cgg aac atc tgg     624
Ser Leu Pro Asp Leu Asp Thr Thr Ser Thr Ala Val Arg Asn Ile Trp
            175                 180                 185

tac gac tgg gtg gca gat ttg gtc gcc aac tat tct atc gac ggt ctg     672
Tyr Asp Trp Val Ala Asp Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu
        190                 195                 200

cgt gtc gac act gta aaa cac gtc gaa aaa gac ttt tgg ccc ggc tat     720
Arg Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Tyr
    205                 210                 215

aac agc gca gca ggc gtc tac tgt gtc ggc gag gtc tac tca ggc gat     768
```

Asn Ser Ala Ala Gly Val Tyr Cys Val Gly Glu Val Tyr Ser Gly Asp
220 225 230 235

cct gca tac aca tgc ccg tac cag aac tac atg gac ggt gtg ctc aac 816
Pro Ala Tyr Thr Cys Pro Tyr Gln Asn Tyr Met Asp Gly Val Leu Asn
240 245 250

tat cca att tac tac cag ctt ctc tat gcg ttt gag tcg tcc agc ggc 864
Tyr Pro Ile Tyr Tyr Gln Leu Leu Tyr Ala Phe Glu Ser Ser Ser Gly
255 260 265

agc atc agc gat ctc tac aac atg atc agc tcc gtt gcc tcc agc tgc 912
Ser Ile Ser Asp Leu Tyr Asn Met Ile Ser Ser Val Ala Ser Ser Cys
270 275 280

aag gat ccc aca ctc ctg ggg aac ttc atc gag aac cac gat aac ccc 960
Lys Asp Pro Thr Leu Leu Gly Asn Phe Ile Glu Asn His Asp Asn Pro
285 290 295

cgc ttc gct tcc tac acg agc gac tac tcg cag gct aag aac gtg atc 1008
Arg Phe Ala Ser Tyr Thr Ser Asp Tyr Ser Gln Ala Lys Asn Val Ile
300 305 310 315

acc ttc atc ttc ctg agc gac ggt atc ccc atc gtc tac gcc gga cag 1056
Thr Phe Ile Phe Leu Ser Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln
320 325 330

gaa cag cac tac agc gga ggc agc gac cca gct aac cgc gag gcc acc 1104
Glu Gln His Tyr Ser Gly Gly Ser Asp Pro Ala Asn Arg Glu Ala Thr
335 340 345

tgg ctg tcc gga tac tcc acc agc gct acg ctg tac acc tgg atc gcc 1152
Trp Leu Ser Gly Tyr Ser Thr Ser Ala Thr Leu Tyr Thr Trp Ile Ala
350 355 360

acc aca aac cag atc cgc agc ctg gcg atc tcc aag gac gcg gga tac 1200
Thr Thr Asn Gln Ile Arg Ser Leu Ala Ile Ser Lys Asp Ala Gly Tyr
365 370 375

gtg cag gcc aag aac aac ccc ttc tac tcc gac tcc aac acc atc gcc 1248
Val Gln Ala Lys Asn Asn Pro Phe Tyr Ser Asp Ser Asn Thr Ile Ala
380 385 390 395

atg cgc aag ggc acg aca gcc ggc gcg caa gtc atc acc gtc ctc agc 1296
Met Arg Lys Gly Thr Thr Ala Gly Ala Gln Val Ile Thr Val Leu Ser
400 405 410

aac aag ggc gcc tcc ggc agc tcc tac act ctc tct ttg agt ggt acc 1344
Asn Lys Gly Ala Ser Gly Ser Ser Tyr Thr Leu Ser Leu Ser Gly Thr
415 420 425

ggc tac tcc gcc ggc gtg acc ctg gtc gag acg tac acc tgc acc acg 1392
Gly Tyr Ser Ala Gly Val Thr Leu Val Glu Thr Tyr Thr Cys Thr Thr
430 435 440

gtg act gta gac tcg agc ggc aac ctg ccc gtc cca atg aca tcc ggc 1440
Val Thr Val Asp Ser Ser Gly Asn Leu Pro Val Pro Met Thr Ser Gly
445 450 455

ttg ccg cga gtg ttt gtc ccg tcg tcc tgg gtg aat ggg agc gcg ctt 1488

Leu Pro Arg Val Phe Val Pro Ser Ser Trp Val Asn Gly Ser Ala Leu
460 465 470 475

tgc aac act ggc ggc acc act acg acg gct acc ccc act ggc tcc ggc 1536
Cys Asn Thr Gly Gly Thr Thr Thr Thr Ala Thr Pro Thr Gly Ser Gly
480 485 490

agc gtg acc tcg acc agc aag acc acc gcg act gcc agc aag acc agc 1584
Ser Val Thr Ser Thr Ser Lys Thr Thr Ala Thr Ala Ser Lys Thr Ser
495 500 505

acc agt acg tca tca acc tcc tgt acc act ccc acc gcc gtg gct gtg 1632
Thr Ser Thr Ser Ser Thr Ser Cys Thr Thr Pro Thr Ala Val Ala Val
510 515 520

act ttc gat ctg aca gct acc acc acc tac ggc gag aac atc tac ctg 1680
Thr Phe Asp Leu Thr Ala Thr Thr Thr Tyr Gly Glu Asn Ile Tyr Leu
525 530 535

gtc gga tcg atc tct cag ctg ggt gac tgg gaa acc agc gac ggc ata 1728
Val Gly Ser Ile Ser Gln Leu Gly Asp Trp Glu Thr Ser Asp Gly Ile
540 545 550 555

gct ctg agt gct gac aag tac act tcc agc gac ccg ctc tgg tat gtc 1776
Ala Leu Ser Ala Asp Lys Tyr Thr Ser Ser Asp Pro Leu Trp Tyr Val
560 565 570

act gtg act ctg ccg gct ggt gag tcg ttt gag tac aag ttt atc cgc 1824
Thr Val Thr Leu Pro Ala Gly Glu Ser Phe Glu Tyr Lys Phe Ile Arg
575 580 585

att gag agc gat gac tcc gtg gag tgg gag agt gat ccc aac cga gaa 1872
Ile Glu Ser Asp Asp Ser Val Glu Trp Glu Ser Asp Pro Asn Arg Glu
590 595 600

tac acc gtt cct cag gcg tgc gga acg tcg acc gcg acg gtg act gac 1920
Tyr Thr Val Pro Gln Ala Cys Gly Thr Ser Thr Ala Thr Val Thr Asp
605 610 615

acc tgg cgg tag 1932
Thr Trp Arg
620

<210> 4
<211> 643
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 4

Met Arg Leu Ser Thr Ser Ser Leu Phe Leu Ser Val Ser Leu Leu Gly
-20 -15 -10

Lys Leu Ala Leu Gly Leu Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile
-5 -1 1 5 10

Tyr Phe Leu Leu Thr Asp Arg Phe Gly Arg Thr Asp Asn Ser Thr Thr
15 20 25

Ala Ala Cys Asp Thr Ser Asp Arg Val Tyr Cys Gly Gly Ser Trp Gln
30 35 40

Gly Ile Ile Asn Gln Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala
45 50 55

Ile Trp Ile Thr Pro Val Thr Gly Gln Phe Tyr Glu Asn Thr Gly Asp
60 65 70 75

Gly Thr Ser Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Asp Leu Asn
80 85 90

Tyr Asn Tyr Gly Thr Ala Gln Asp Leu Lys Asn Leu Ala Asn Ala Leu
95 100 105

His Glu Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met
110 115 120

Gly Tyr Asp Gly Ala Gly Asn Thr Val Asp Tyr Ser Val Phe Asn Pro
125 130 135

Phe Ser Ser Ser Ser Tyr Phe His Pro Tyr Cys Leu Ile Ser Asn Tyr
140 145 150 155

Asp Asn Gln Thr Asn Val Glu Asp Cys Trp Leu Gly Asp Thr Thr Val
160 165 170

Ser Leu Pro Asp Leu Asp Thr Thr Ser Thr Ala Val Arg Asn Ile Trp
175 180 185

Tyr Asp Trp Val Ala Asp Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu
190 195 200

Arg Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Tyr
205 210 215

Asn Ser Ala Ala Gly Val Tyr Cys Val Gly Glu Val Tyr Ser Gly Asp
220 225 230 235

Pro Ala Tyr Thr Cys Pro Tyr Gln Asn Tyr Met Asp Gly Val Leu Asn
240 245 250

Tyr Pro Ile Tyr Tyr Gln Leu Leu Tyr Ala Phe Glu Ser Ser Ser Gly
255 260 265

Ser Ile Ser Asp Leu Tyr Asn Met Ile Ser Ser Val Ala Ser Ser Cys
270 275 280

Lys Asp Pro Thr Leu Leu Gly Asn Phe Ile Glu Asn His Asp Asn Pro
285 290 295

Arg Phe Ala Ser Tyr Thr Ser Asp Tyr Ser Gln Ala Lys Asn Val Ile
300 305 310 315

Thr Phe Ile Phe Leu Ser Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln
320 325 330

Glu Gln His Tyr Ser Gly Gly Ser Asp Pro Ala Asn Arg Glu Ala Thr
335 340 345

Trp Leu Ser Gly Tyr Ser Thr Ser Ala Thr Leu Tyr Thr Trp Ile Ala
350 355 360

Thr Thr Asn Gln Ile Arg Ser Leu Ala Ile Ser Lys Asp Ala Gly Tyr
365 370 375

Val Gln Ala Lys Asn Asn Pro Phe Tyr Ser Asp Ser Asn Thr Ile Ala
380 385 390 395

Met Arg Lys Gly Thr Thr Ala Gly Ala Gln Val Ile Thr Val Leu Ser
400 405 410

Asn Lys Gly Ala Ser Gly Ser Ser Tyr Thr Leu Ser Leu Ser Gly Thr
415 420 425

Gly Tyr Ser Ala Gly Val Thr Leu Val Glu Thr Tyr Thr Cys Thr Thr
430 435 440

Val Thr Val Asp Ser Ser Gly Asn Leu Pro Val Pro Met Thr Ser Gly
445 450 455

Leu Pro Arg Val Phe Val Pro Ser Ser Trp Val Asn Gly Ser Ala Leu
460 465 470 475

Cys Asn Thr Gly Gly Thr Thr Thr Thr Ala Thr Pro Thr Gly Ser Gly
480 485 490

Ser Val Thr Ser Thr Ser Lys Thr Thr Ala Thr Ala Ser Lys Thr Ser
495 500 505

Thr Ser Thr Ser Ser Thr Ser Cys Thr Thr Pro Thr Ala Val Ala Val
510 515 520

Thr Phe Asp Leu Thr Ala Thr Thr Thr Tyr Gly Glu Asn Ile Tyr Leu
525 530 535

Val Gly Ser Ile Ser Gln Leu Gly Asp Trp Glu Thr Ser Asp Gly Ile
540 545 550 555

Ala Leu Ser Ala Asp Lys Tyr Thr Ser Ser Asp Pro Leu Trp Tyr Val
560 565 570

Thr Val Thr Leu Pro Ala Gly Glu Ser Phe Glu Tyr Lys Phe Ile Arg
575 580 585

Ile Glu Ser Asp Asp Ser Val Glu Trp Glu Ser Asp Pro Asn Arg Glu
590 595 600

Tyr Thr Val Pro Gln Ala Cys Gly Thr Ser Thr Ala Thr Val Thr Asp
605 610 615

Thr Trp Arg
620

<210> 5
<211> 1887
<212> ADN
<213> Secuencia artificial

<220>
<223> Híbrido

<220>
<221> sig_peptide
<222> (1)..(63)

<220>
<221> CDS
<222> (1)..(1884)

<220>
<221> mat_peptide
<222> (64)..(1884)

<400> 5

```
atg aga tta tcg act tcg agt ctc ttc ctt tcc gtg tct ctg ctg ggg      48
Met Arg Leu Ser Thr Ser Ser Leu Phe Leu Ser Val Ser Leu Leu Gly
    -20                 -15                 -10

aag ctg gcc ctc ggg ctg acc cca gca gaa tgg cgc agc cag tca atc      96
Lys Leu Ala Leu Gly Leu Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile
-5              -1  1               5                   10

tac ttc ctg ttg acc gat cgc ttt ggt cga acg gac aat tcc aca act     144
Tyr Phe Leu Leu Thr Asp Arg Phe Gly Arg Thr Asp Asn Ser Thr Thr
            15                  20                  25

gcc gct tgc gat acc tct gac aga gta tac tgc ggt ggt agc tgg cag     192
Ala Ala Cys Asp Thr Ser Asp Arg Val Tyr Cys Gly Gly Ser Trp Gln
        30                  35                  40

gga atc atc aac caa ctc gat tac atc caa ggg atg gga ttc act gcc     240
Gly Ile Ile Asn Gln Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala
    45                  50                  55

atc tgg atc acc ccg gtc act gga cag ttc tac gaa aac acg ggc gac     288
Ile Trp Ile Thr Pro Val Thr Gly Gln Phe Tyr Glu Asn Thr Gly Asp
60                  65                  70                  75

ggc acc tct tac cat gga tac tgg cag cag gac atc tac gat ctc aac     336
Gly Thr Ser Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Asp Leu Asn
                80                  85                  90

tac aac tac gga acg gcc caa gac ctc aag aac cta gcc aat gct ttg     384
Tyr Asn Tyr Gly Thr Ala Gln Asp Leu Lys Asn Leu Ala Asn Ala Leu
            95                  100                 105

cac gag cgc ggc atg tat ttg atg gtt gat gtg gtt gcc aac cac atg     432
His Glu Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met
        110                 115                 120

ggc tat gac gga gcg gga aac acc gtg gac tac agt gtt ttc aac ccc     480
Gly Tyr Asp Gly Ala Gly Asn Thr Val Asp Tyr Ser Val Phe Asn Pro
    125                 130                 135

ttc tcc tcc tcc tcc tac ttt cac cca tac tgc ctc atc tcc aat tac     528
Phe Ser Ser Ser Ser Tyr Phe His Pro Tyr Cys Leu Ile Ser Asn Tyr
140                 145                 150                 155

gac aac cag acc aat gtt gaa gac tgc tgg ctg ggc gat acc acc gtt     576
Asp Asn Gln Thr Asn Val Glu Asp Cys Trp Leu Gly Asp Thr Thr Val
                160                 165                 170

tcg ctg cca gat ctc gac acg aca agc aca gcc gtg cgg aac atc tgg     624
Ser Leu Pro Asp Leu Asp Thr Thr Ser Thr Ala Val Arg Asn Ile Trp
            175                 180                 185

tac gac tgg gtg gca gat ttg gtc gcc aac tat tct atc gac ggt ctg     672
Tyr Asp Trp Val Ala Asp Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu
        190                 195                 200

cgt gtc gac act gta aaa cac gtc gaa aaa gac ttt tgg ccc ggc tat     720
Arg Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Tyr
    205                 210                 215
```

aac agc gca gca ggc gtc tac tgt gtc ggc gag gtc tac tca ggc gat 768
Asn Ser Ala Ala Gly Val Tyr Cys Val Gly Glu Val Tyr Ser Gly Asp
220 225 230 235

cct gca tac aca tgc ccg tac cag aac tac atg gac ggt gtg ctc aac 816
Pro Ala Tyr Thr Cys Pro Tyr Gln Asn Tyr Met Asp Gly Val Leu Asn
240 245 250

tat cca att tac tac cag ctt ctc tat gcg ttt gag tcg tcc agc ggc 864
Tyr Pro Ile Tyr Tyr Gln Leu Leu Tyr Ala Phe Glu Ser Ser Ser Gly
255 260 265

agc atc agc gat ctc tac aac atg atc agc tcc gtt gcc tcc agc tgc 912
Ser Ile Ser Asp Leu Tyr Asn Met Ile Ser Ser Val Ala Ser Ser Cys
270 275 280

aag gat ccc aca ctc ctg ggg aac ttc atc gag aac cac gat aac ccc 960
Lys Asp Pro Thr Leu Leu Gly Asn Phe Ile Glu Asn His Asp Asn Pro
285 290 295

cgc ttc gct tcc tac acg agc gac tac tcg cag gct aag aac gtg atc 1008
Arg Phe Ala Ser Tyr Thr Ser Asp Tyr Ser Gln Ala Lys Asn Val Ile
300 305 310 315

acc ttc atc ttc ctg agc gac ggt atc ccc atc gtc tac gcc gga cag 1056
Thr Phe Ile Phe Leu Ser Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln
320 325 330

gaa cag cac tac agc gga ggc agc gac cca gct aac cgc gag gcc acc 1104
Glu Gln His Tyr Ser Gly Gly Ser Asp Pro Ala Asn Arg Glu Ala Thr
335 340 345

tgg ctg tcc gga tac tcc acc agc gct acg ctg tac acc tgg atc gcc 1152
Trp Leu Ser Gly Tyr Ser Thr Ser Ala Thr Leu Tyr Thr Trp Ile Ala
350 355 360

acc aca aac cag atc cgc agc ctg gcg atc tcc aag gac gcg gga tac 1200
Thr Thr Asn Gln Ile Arg Ser Leu Ala Ile Ser Lys Asp Ala Gly Tyr
365 370 375

gtg cag gcc aag aac aac ccc ttc tac tcc gac tcc aac acc atc gcc 1248
Val Gln Ala Lys Asn Asn Pro Phe Tyr Ser Asp Ser Asn Thr Ile Ala
380 385 390 395

atg cgc aag ggc acg aca gcc ggc gcg caa gtc atc acc gtc ctc agc 1296
Met Arg Lys Gly Thr Thr Ala Gly Ala Gln Val Ile Thr Val Leu Ser
400 405 410

aac aag ggc gcc tcc ggc agc tcc tac act ctc tct ttg agt ggt acc 1344
Asn Lys Gly Ala Ser Gly Ser Ser Tyr Thr Leu Ser Leu Ser Gly Thr
415 420 425

ggc tac tcc gcc ggc gtg acc ctg gtc gag acg tac acc tgc acc acg 1392
Gly Tyr Ser Ala Gly Val Thr Leu Val Glu Thr Tyr Thr Cys Thr Thr
430 435 440

gtg act gta gac tcg agc ggc aac ctg ccc gtc cca atg aca tcc ggc 1440
Val Thr Val Asp Ser Ser Gly Asn Leu Pro Val Pro Met Thr Ser Gly
445 450 455

ttg ccg cga gtg ttt gtc ccg tcg tcc tgg gtg aat ggg agc gcg ctt 1488
Leu Pro Arg Val Phe Val Pro Ser Ser Trp Val Asn Gly Ser Ala Leu
460 465 470 475

tgc aac acg acg ggg acg ggg ggc agc act ggc act act acc gct tcc 1536
Cys Asn Thr Thr Gly Thr Gly Gly Ser Thr Gly Thr Thr Thr Ala Ser
480 485 490

gag acc gga gga tca tcg cca acc tcg aca gca tgc gcg tcc gtc ccc 1584
Glu Thr Gly Gly Ser Ser Pro Thr Ser Thr Ala Cys Ala Ser Val Pro
495 500 505

gta act ttt aac gag aag gtt acg acc gtc gtc ggg gag acg atc aag 1632
Val Thr Phe Asn Glu Lys Val Thr Thr Val Val Gly Glu Thr Ile Lys
510 515 520

ata tcc ggc agc gtg gcc gcc ctc ggc gac tgg gcc acg ggc agc gcg 1680
Ile Ser Gly Ser Val Ala Ala Leu Gly Asp Trp Ala Thr Gly Ser Ala
525 530 535

gtg gcc ctg agc gcc gcg agc tac acg tcc agc aac ccg cag tgg gac 1728
Val Ala Leu Ser Ala Ala Ser Tyr Thr Ser Ser Asn Pro Gln Trp Asp
540 545 550 555

gtg acc atc agc ttt gcg ccg ggc acc gtg atc gag tac aag tac atc 1776
Val Thr Ile Ser Phe Ala Pro Gly Thr Val Ile Glu Tyr Lys Tyr Ile
560 565 570

aac gtg gcg agc agc ggg gcc gtg acc tgg gag gcc gac ccg aac cac 1824
Asn Val Ala Ser Ser Gly Ala Val Thr Trp Glu Ala Asp Pro Asn His
575 580 585

acg tac acg gtc ccg gcg tcc tgc gcc acg gcc gcc gtc gtg tct gac 1872
Thr Tyr Thr Val Pro Ala Ser Cys Ala Thr Ala Ala Val Val Ser Asp
590 595 600

acg tgg cag acg tga 1887
Thr Trp Gln Thr
605

<210> 6
<211> 628
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción sintética

<400> 6

Met Arg Leu Ser Thr Ser Ser Leu Phe Leu Ser Val Ser Leu Leu Gly
-20 -15 -10

Lys Leu Ala Leu Gly Leu Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile
-5 -1 1 5 10

Tyr Phe Leu Leu Thr Asp Arg Phe Gly Arg Thr Asp Asn Ser Thr Thr

15 20 25

Ala Ala Cys Asp Thr Ser Asp Arg Val Tyr Cys Gly Gly Ser Trp Gln
30 35 40

Gly Ile Ile Asn Gln Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala
45 50 55

Ile Trp Ile Thr Pro Val Thr Gly Gln Phe Tyr Glu Asn Thr Gly Asp
60 65 70 75

Gly Thr Ser Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Asp Leu Asn
80 85 90

Tyr Asn Tyr Gly Thr Ala Gln Asp Leu Lys Asn Leu Ala Asn Ala Leu
95 100 105

His Glu Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met
110 115 120

Gly Tyr Asp Gly Ala Gly Asn Thr Val Asp Tyr Ser Val Phe Asn Pro
125 130 135

Phe Ser Ser Ser Ser Tyr Phe His Pro Tyr Cys Leu Ile Ser Asn Tyr
140 145 150 155

Asp Asn Gln Thr Asn Val Glu Asp Cys Trp Leu Gly Asp Thr Thr Val
160 165 170

Ser Leu Pro Asp Leu Asp Thr Thr Ser Thr Ala Val Arg Asn Ile Trp
175 180 185

Tyr Asp Trp Val Ala Asp Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu
190 195 200

Arg Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Tyr
205 210 215

Asn Ser Ala Ala Gly Val Tyr Cys Val Gly Glu Val Tyr Ser Gly Asp
220 225 230 235

Pro Ala Tyr Thr Cys Pro Tyr Gln Asn Tyr Met Asp Gly Val Leu Asn
240 245 250

Tyr Pro Ile Tyr Tyr Gln Leu Leu Tyr Ala Phe Glu Ser Ser Ser Gly

255 260 265

Ser Ile Ser Asp Leu Tyr Asn Met Ile Ser Ser Val Ala Ser Ser Cys
270 275 280

Lys Asp Pro Thr Leu Leu Gly Asn Phe Ile Glu Asn His Asp Asn Pro
285 290 295

Arg Phe Ala Ser Tyr Thr Ser Asp Tyr Ser Gln Ala Lys Asn Val Ile
300 305 310 315

Thr Phe Ile Phe Leu Ser Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln
320 325 330

Glu Gln His Tyr Ser Gly Gly Ser Asp Pro Ala Asn Arg Glu Ala Thr
335 340 345

Trp Leu Ser Gly Tyr Ser Thr Ser Ala Thr Leu Tyr Thr Trp Ile Ala
350 355 360

Thr Thr Asn Gln Ile Arg Ser Leu Ala Ile Ser Lys Asp Ala Gly Tyr
365 370 375

Val Gln Ala Lys Asn Asn Pro Phe Tyr Ser Asp Ser Asn Thr Ile Ala
380 385 390 395

Met Arg Lys Gly Thr Thr Ala Gly Ala Gln Val Ile Thr Val Leu Ser
400 405 410

Asn Lys Gly Ala Ser Gly Ser Ser Tyr Thr Leu Ser Leu Ser Gly Thr
415 420 425

Gly Tyr Ser Ala Gly Val Thr Leu Val Glu Thr Tyr Thr Cys Thr Thr
430 435 440

Val Thr Val Asp Ser Ser Gly Asn Leu Pro Val Pro Met Thr Ser Gly
445 450 455

Leu Pro Arg Val Phe Val Pro Ser Ser Trp Val Asn Gly Ser Ala Leu
460 465 470 475

Cys Asn Thr Thr Gly Thr Gly Gly Ser Thr Gly Thr Thr Thr Ala Ser
480 485 490

Glu Thr Gly Gly Ser Ser Pro Thr Ser Thr Ala Cys Ala Ser Val Pro

495 500 505

Val Thr Phe Asn Glu Lys Val Thr Thr Val Val Gly Glu Thr Ile Lys
510 515 520

Ile Ser Gly Ser Val Ala Ala Leu Gly Asp Trp Ala Thr Gly Ser Ala
525 530 535

Val Ala Leu Ser Ala Ala Ser Tyr Thr Ser Ser Asn Pro Gln Trp Asp
540 545 550 555

Val Thr Ile Ser Phe Ala Pro Gly Thr Val Ile Glu Tyr Lys Tyr Ile
560 565 570

Asn Val Ala Ser Ser Gly Ala Val Thr Trp Glu Ala Asp Pro Asn His
575 580 585

Thr Tyr Thr Val Pro Ala Ser Cys Ala Thr Ala Ala Val Val Ser Asp
590 595 600

Thr Trp Gln Thr
605

REIVINDICACIONES

1. Polipéptido aislado que tiene actividad alfa-amilasa, seleccionado del grupo que consiste en:

(a) un polipéptido que comprende una secuencia de aminoácidos del polipéptido maduro de la secuencia de aminoácidos de la SEQ ID N.º: 2, SEQ ID N.º: 4 o SEQ ID N.º: 6;
(b) un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un 90% de identidad de secuencia con el polipéptido maduro de la secuencia de aminoácidos de la SEQ ID N.º: 2, SEQ ID N.º: 4 o SEQ ID N.º: 6;
(c) un polipéptido codificado por un polinucleótido que tiene al menos un 90% de identidad de secuencia con la secuencia de codificación de polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º: 3 o SEQ ID N.º: 5; y
(d) un polipéptido codificado por un polinucleótido que se hibrida en condiciones de astringencia media con la secuencia de codificación de polipéptidos maduros de la SEQ ID N.º: 1, SEQ ID N.º: 3 o SEQ ID N.º: 5.

2. Polipéptido de la reivindicación 1, que es un polipéptido que comprende una secuencia de aminoácidos que tiene al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% de identidad de secuencia con el polipéptido maduro de la secuencia de aminoácidos de la SEQ ID N.º: 2, SEQ ID N.º: 4 o SEQ ID N.º: 6.

3. Polipéptido de la reivindicación 1, que es un polipéptido codificado por un polinucleótido que tiene al menos un 95%, o al menos un 99% de identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1, SEQ ID N.º: 3 o SEQ ID N.º: 5.

4. Polipéptido de la reivindicación 1, un polipéptido codificado por un polinucleótido que se hibrida en condiciones de astringencia media con la secuencia codificante del polipéptido maduro de la SEQ ID N.º: 1 o SEQ ID N.º: 3 o SEQ ID N.º: 5.

5. Polipéptido de la reivindicación 1 que comprende o que consiste en los residuos 1-585 de la SEQ ID Nº: 2 o el polipéptido maduro codificado por la SEQ ID Nº: 1.

6. Polipéptido de la reivindicación 1 que comprende o que consiste en los residuos 1-622 de la SEQ ID Nº: 4 o el polipéptido maduro codificado por la SEQ ID Nº: 3.

7. Polipéptido de la reivindicación 1 que comprende o que consiste en los residuos 1-607 de la SEQ ID Nº: 6 o el polipéptido maduro codificado por la SEQ ID Nº: 5.

8. Polinucleótido aislado que codifica el polipéptido de cualquiera de las reivindicaciones 1-7.

9. Construcción de ácido nucleico o vector de expresión que comprende el polinucleótido de la reivindicación 8 unido operativamente a una o más secuencias de control que dirigen la producción del polipéptido en un huésped de expresión.

10. Célula hospedadora recombinante que comprende el polinucleótido de la reivindicación 8 operativamente unido a una o más secuencias de control que dirigen la producción del polipéptido.

11. Célula hospedadora recombinante de la reivindicación 10, en donde la célula hospedadora es una célula fúngica, tal como una levadura o una célula fúngica filamentosa.

12. Célula hospedadora recombinante de la reivindicación 10 u 11, en la que la célula hospedadora es una célula de *Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces,* o *Yarrowia*, como una célula de *Kluyveromyces lactis, Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis, Saccharomyces oviformis,* o *Yarrowia Lipolytica.*

13. Célula hospedadora recombinante de la reivindicación 10 u 11, en la que la célula hospedadora es una célula de *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysosporium, Coprinus, Coriolus, Cryptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes,* o *Trichoderma.*

14. Método para producir un polipéptido que tiene actividad de alfa-amilasa, que comprende:

(a) cultivar la célula hospedadora de la reivindicación 10 en condiciones propicias para la producción del polipéptido; y
(b) recuperar el polipéptido.

15. Proceso para producir un producto de fermentación, que comprende:

(a) tratar un material que contiene almidón con el polipéptido de cualquiera de las reivindicaciones 1-7;
(b) fermentar el material de (a) usando un organismo de fermentación para producir un producto de fermentación.

16. Proceso de la reivindicación 15, en donde (a) comprende (i) licuar el almidón gelatinizado con un polipéptido de cualquiera de las reivindicaciones 1-7 y (ii) sacarificar la pasta licuada obtenido usando una glucoamilasa.

17. Proceso de la reivindicación 15, en donde el producto de fermentación se selecciona del grupo que consiste en etanol como combustible, alcohol potable, una bebida o compuestos orgánicos.

18. Proceso de la reivindicación 15, que comprende recuperar el producto de fermentación.

19. Proceso para producir un producto de fermentación que comprende:

(a) tratar un material que contiene almidón con una alfa-amilasa de cualquiera de las reivindicaciones 1-7 a una temperatura inferior a la temperatura de gelatinización inicial de dicho material que contiene almidón; y
(b) fermentar el material de almidón tratado usando un organismo de fermentación para producir un producto de fermentación.

20. Proceso de la reivindicación 19, en donde los pasos (a) y (b) se llevan a cabo secuencial o simultáneamente.

21. Proceso de la reivindicación 19, en donde el producto de fermentación es etanol como combustible.