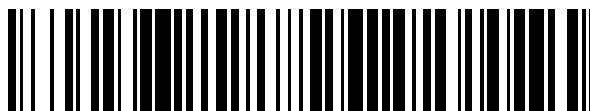


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 711 165**

51 Int. Cl.:

C12P 21/00 (2006.01)

C12N 9/10 (2006.01)

C07K 16/00 (2006.01)

C12N 15/80 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.07.2014 PCT/EP2014/064818**

87 Fecha y número de publicación internacional: **15.01.2015 WO15004239**

96 Fecha de presentación y número de la solicitud europea: **10.07.2014 E 14738496 (0)**

97 Fecha y número de publicación de la concesión europea: **28.11.2018 EP 3019621**

54 Título: **Producción de glucoproteínas que tienen una ocupación del sitio de N-glucosilación aumentada**

30 Prioridad:

10.07.2013 EP 13175997

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

30.04.2019

73 Titular/es:

**GLYKOS FINLAND OY (100.0%)
Viikinkaari 6
00790 Helsinki, FI**

72 Inventor/es:

**NATUNEN, JARI;
LANDOWSKI, CHRISTOPHER;
SALOHEIMO, MARKKU;
OSTERMEIER, CHRISTIAN;
SOMMER, BENJAMIN PATRICK y
WAHL, RAMON**

74 Agente/Representante:

CAMPello ESTEBARANZ, Reyes

ES 2 711 165 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

Producción de glucoproteínas que tienen una ocupación del sitio de N-glucosilación aumentada

5 CAMPO TÉCNICO

La presente divulgación se refiere a composiciones y métodos útiles para la producción de proteínas heterólogas, por ejemplo, anticuerpos recombinantes, en células fúngicas filamentosas.

10 ANTECEDENTES

La modificación postraduccional de proteínas eucariotas, particularmente proteínas terapéuticas tales como inmunoglobulinas, a menudo es necesaria para el plegamiento y la función apropiados de las proteínas. Debido a que los sistemas de expresión procariotas estándar carecen de la maquinaria adecuada necesaria para tales modificaciones, deben usarse sistemas de expresión alternativos en la producción de estas proteínas terapéuticas.

Incluso cuando las proteínas eucariotas no tienen modificaciones postraduccionales, los sistemas de expresión procariotas a menudo carecen de las proteínas chaperonas necesarias para un plegamiento adecuado. La levadura y los hongos son opciones atractivas para expresar proteínas, ya que se pueden cultivar fácilmente a gran escala en medios sencillos, lo que permite bajos costes de producción, y la levadura y los hongos tienen maquinaria postraduccional y chaperonas que realizan funciones similares a las de las células de mamíferos. Además, hay herramientas disponibles para manipular la composición genética relativamente simple de las células de levadura y fúngicas, así como las células eucariotas más complejas, tales como las células de mamíferos o de insecto (De Pourcq et al., Appl Microbiol Biotechnol, 87(5):1617-31).

Sin embargo, las modificaciones postraduccionales que se producen en levaduras y hongos pueden seguir siendo una preocupación para la producción de una proteína terapéutica recombinante. En particular, la insuficiente N-glucosilación es uno de los mayores obstáculos a superar en la producción de productos biofarmacéuticos para aplicaciones humanas en hongos.

Se ha demostrado que la N-glucosilación, que se refiere a la unión de la molécula de azúcar a un átomo de nitrógeno de una cadena lateral de asparagina, modula la farmacocinética y la farmacodinámica de las proteínas terapéuticas.

Cuando las proteínas recombinantes se expresan en células fúngicas filamentosas, tales como las células del hongo *Trichoderma*, la proporción de sitios de N-glucosilación que de hecho están glucosilados, es generalmente menor que para la misma proteína expresada en un sistema de mamífero, tal como las células CHO.

El documento WO2011/106389, titulado "Methods for increasing N-glycosylation site occupancy on therapeutic glycoproteins produced in *Pichia pastoris*", describe células de *Pichia pastoris* que sobreexpresan oligotransferasa heteróloga de subunidad única, y son capaces de producir glucoproteínas con N-glucosilación mejorada.

De manera similar, Choi et al. describen la N-glucosilación mejorada de proteínas recombinantes mediante la expresión heteróloga de oligotransferasa heteróloga de subunidad única (Choi et al., Appl Microbiol Biotechnol, 95(3): 671-82).

Los mismos autores también han descrito, en el documento WO2013062939, métodos para aumentar la ocupación de N-glicanos y reducir la producción de N-glicanos híbridos en cepas de *Pichia pastoris* que carecen de actividad de alfa-1,3-manosiltransferasa (alteración de Alg3p).

Faltan informes de sistemas de expresión de células fúngicas que expresen N-glicanos fucosilados de tipo humano. De hecho, debido al enfoque de la industria en la tecnología de cultivo de células de mamífero durante tanto tiempo, los sistemas de expresión de células fúngicas tal como *Trichoderma*, no están tan bien establecidos para la producción de proteínas terapéuticas como el cultivo de células de mamífero y, por lo tanto, padecen inconvenientes cuando se expresan proteínas de mamífero. En particular, sigue existiendo la necesidad en la técnica de células fúngicas filamentosas mejoradas, tales como células del hongo *Trichoderma*, que puedan producir de manera estable proteínas heterólogas con una mayor ocupación del sitio de N-glucosilación, preferiblemente a altos niveles de expresión.

RESUMEN

La invención se define en las reivindicaciones.

La presente divulgación se refiere a métodos mejorados para producir glucoproteínas con un aumento de la ocupación del sitio de N-glucosilación en sistemas de expresión fúngicos filamentosos, y más específicamente, glucoproteínas, tales como anticuerpos o inmunoglobulinas relacionadas o proteínas de fusión.

La presente divulgación se basa, en parte, en el descubrimiento sorprendente de que las células fúngicas filamentosas, tales como las células de *Trichoderma*, pueden modificarse genéticamente para expresar la actividad de oligosacaril transferasa, sin afectar adversamente al rendimiento de las glucoproteínas producidas.

Por consiguiente, en un primer aspecto, la divulgación se refiere a una célula fúngica filamentosa que comprende

- i. una o más mutaciones que reducen o eliminan una o más actividades de proteasa endógena en comparación con una célula fúngica filamentosa parental que no tiene dicha mutación o mutaciones,
- ii. un polinucleótido que codifica una subunidad catalítica heteróloga de la oligosacaril transferasa, y
- iii. un polinucleótido que codifica una glucoproteína heteróloga,

en la que dicha subunidad catalítica de oligosacaril transferasa se selecciona de subunidades catalíticas de oligosacaril transferasa de *Leishmania*.

En una forma de realización, dicha célula fúngica filamentosa tiene al menos una reducción de dos veces, preferiblemente al menos una reducción de tres veces, incluso más preferiblemente al menos una reducción de cuatro veces, al menos una reducción de cinco veces, en la actividad de proteasa total en comparación con una célula fúngica filamentosa parental que no tiene la mutación o mutaciones deficientes en proteasa.

En una forma de realización de la divulgación, dicha célula fúngica filamentosa es una célula de *Trichoderma*, *Neurospora*, *Myceliophthora*, *Chrysosporium*, *Aspergillus*, o *Fusarium*.

En una forma de realización de la invención, el polinucleótido que codifica la subunidad catalítica heteróloga de la oligosacaril transferasa comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en la SEQ ID NO: 2, SEQ ID NO: 9, SEQ ID NO: 88 y SEQ ID NO: 90 o un polinucleótido que codifica un polipéptido variante funcional que tiene al menos un 50%, al menos un 60%, al menos un 70% de identidad, al menos un 80% de identidad, al menos un 90% de identidad, o al menos un 95% de identidad con la SEQ ID NO: 1, SEQ ID NO: 8, SEQ ID NO: 89 o SEQ ID NO: 91, teniendo dicho polipéptido variante funcional actividad de oligosacariltransferasa.

En otra forma de realización, dicho polinucleótido que codifica la subunidad catalítica heteróloga de oligosacaril transferasa está bajo el control de un promotor para la expresión constitutiva de dicha oligosacaril transferasa en dicha célula.

En una forma de realización de la divulgación, la ocupación del sitio de N-glucosilación de la glucoproteína heteróloga expresada en una célula fúngica filamentosa es de al menos el 80%, al menos el 90%, al menos el 95%, al menos el 99%, o el 100%.

En una forma de realización específica, la ocupación del sitio de N-glucosilación de la glucoproteína heteróloga es de al menos el 95% y las glucoformas Man3, Man5, G0, G1 y/o G2 representan al menos el 50% de los N-glicanos neutros totales de la glucoproteína heteróloga.

En una forma de realización de la divulgación, la célula fúngica filamentosa es una célula de *Trichoderma*, por ejemplo, *Trichoderma reesei*, y dicha célula comprende mutaciones que reducen o eliminan la actividad de

- las tres proteasas endógenas pep1, tsp1, y slp1;
- las tres proteasas endógenas gap1, slp1, y pep1;
- las tres proteasas endógenas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, pep8, pep9, pep11, pep12, tsp1, slp1, slp2, slp3, slp7, gap1 y gap2;
- de tres a seis proteasas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, tsp1, slp1, slp2, slp3, gap1 y gap2;
- de siete a diez proteasas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, pep7, pep8, pep9, tsp1, slp1, slp2, slp3, slp5, slp6, slp7, slp8, tpp1, gap1 y gap2.

En una forma de realización, la célula fúngica comprende además una mutación en el gen que codifica ALG3 que reduce o elimina la expresión de ALG3 correspondiente en comparación con el nivel de expresión del gen ALG3 en una célula parental que no tiene tal mutación.

5

En una forma de realización, la célula fúngica comprende además un polinucleótido que codifica un dominio catalítico de N-acetilglucosaminiltransferasa I y un dominio catalítico de N-acetilglucosaminiltransferasa II.

En una forma de realización, la célula fúngica comprende además uno o más polinucleótidos que codifican un polipéptido seleccionado del grupo que consiste en:

10

- i. α 1, 2 manosidasa;
- ii. dominio catalítico de N-acetilglucosaminiltransferasa I;
- iii. α -manosidasa II;
- 15 iv. dominio catalítico de N-acetilglucosaminiltransferasa II;
- v. β 1,4 galactosiltransferasa; y,
- vi. fucosiltransferasa.

15

En una forma de realización de la divulgación, la glucoproteína heteróloga es una glucoproteína de mamífero.

20

En una forma de realización específica, dicha glucoproteína de mamífero se selecciona del grupo que consiste en un anticuerpo, una inmunoglobulina o una fusión de proteínas que comprende el fragmento Fc de una inmunoglobulina.

En otra forma de realización específica, dicha glucoproteína de mamífero es un anticuerpo terapéutico.

25

En otro aspecto, la divulgación también se refiere a un método para aumentar la ocupación del sitio de N-glucosilación de la glucoproteína heteróloga producida en una célula huésped fúngica filamentosa, que comprende:

30

- a) proporcionar una célula huésped fúngica filamentosa, por ejemplo, una célula de *Trichoderma*, que tiene un gen STT3D de *Leishmania* que codifica una subunidad catalítica de oligosacaril transferasa, o una variante funcional de la misma, y un polinucleótido que codifica una glucoproteína heteróloga,
- b) cultivar la célula huésped en condiciones apropiadas para la expresión del gen STT3D o su variante funcional, o dicha variante funcional, y la producción de la glucoproteína heteróloga;

35 en el que las glucoproteínas heterólogas expresadas presentan una mayor ocupación del sitio de N-glucosilación en comparación con las glucoproteínas heterólogas expresadas en una célula fúngica filamentosa parental correspondiente que no expresa dicha subunidad catalítica de oligosacaril transferasa.

La divulgación también se refiere a un método para producir una composición de glucoproteína heteróloga, con un aumento de la ocupación del sitio de N-glucosilación, que comprende:

40

- a) proporcionar una célula fúngica filamentosa, por ejemplo, una célula de *Trichoderma*, que tiene un gen STT3D de *Leishmania* que codifica una subunidad catalítica de oligosacaril transferasa, o una variante funcional de la misma, y un polinucleótido que codifica una glucoproteína heteróloga,
- 45 b) cultivar la célula en condiciones apropiadas para la expresión del gen STT3D o su variante funcional, y la producción de la composición de glucoproteína heteróloga; y,
- c) recuperar y, opcionalmente, purificar la composición de glucoproteína heteróloga.

En ciertas formas de realización del método de la divulgación, dicho gen STT3D de *Leishmania* que codifica una subunidad catalítica de oligosacaril transferasa comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en la SEQ ID NO: 2, SEQ ID NO: 9, SEQ ID NO: 88 y SEQ ID NO: 90, o un polinucleótido que codifica un polipéptido variante funcional que tiene al menos un 50%, al menos un 60%, al menos un 70% de identidad, al menos un 80% de identidad, al menos un 90% de identidad, o al menos un 95% de identidad con la SEQ ID NO: 1, SEQ ID NO: 8, SEQ ID NO: 89 o SEQ ID NO: 91, teniendo dicho polipéptido variante funcional actividad de oligosacariltransferasa.

55

En una forma de realización, dicho polinucleótido que codifica dicha glucoproteína heteróloga comprende además un polinucleótido que codifica el dominio catalítico de CBH1 y el enlazador como una proteína portadora y/o el promotor cbh1.

En una forma de realización de la divulgación, el cultivo es en un medio comprende un inhibidor de proteasa.

En una forma de realización específica, el cultivo es en un medio que comprende uno o dos inhibidores de proteasa seleccionados de SBTI y quimiostatina.

En una forma de realización del método de la divulgación, la ocupación del sitio de N-glucosilación de la composición de glucoproteína producida es de al menos el 80%, al menos el 90%, al menos el 95%, al menos el 99% o el 100%.

10

En un aspecto, la divulgación también se refiere a una composición de glucoproteína obtenible mediante el método descrito anteriormente.

En un aspecto, la divulgación se refiere a una composición de anticuerpo obtenible por el método descrito anteriormente.

15

En una forma de realización, la divulgación se refiere a la composición de anticuerpo descrita anteriormente, en la que la ocupación del sitio de N-glucosilación es de al menos el 80%, al menos el 90%, al menos el 95%, al menos el 99% o el 100%.

20

En una forma de realización, la divulgación se refiere a la composición de anticuerpo descrita anteriormente, en la que dicha composición de anticuerpo comprende, además, como una glucoforma principal, ya sea:

- i. $\text{Man}\alpha 3[\text{Man}\alpha 6(\text{Man}\alpha 3)\text{Man}\alpha 6]\text{Man}\beta 4\text{GlcNA}\beta 4\text{GlcNAc}$ (glucoforma Man5);
- ii. $\text{GlcNAc}\beta 2\text{Man}\alpha 3[\text{Man}\alpha 6(\text{Man}\alpha 3)\text{Man}\alpha 6]\text{Man}\beta 4\text{GlcNA}\beta 4\text{GlcNAc}$ (glucoforma GlcNAcMan5);
- iii. $\text{Man}\alpha 6(\text{Man}\alpha 3)\text{Man}\beta 4\text{GlcNA}\beta 4\text{GlcNAc}$ (glucoforma Man3);
- iv. $\text{Man}\alpha 6(\text{GlcNAc}\beta 2\text{Man}\alpha 3)\text{Man}\beta 4\text{GlcNA}\beta 4\text{GlcNAc}$ (glucoforma GlcNAcMan3); o,
- v. N-glicanos de tipo complejo seleccionados de la glucoforma G0, G1, o G2.

25

30 DESCRIPCIÓN DE LAS FIGURAS

Figura 1. Diseño de casete de expresión esquemático para STT3 de Leishmania major dirigido al locus de xilanas 1.

Figura 2. Espectros ejemplares de la cepa parental M317 (pyr4- de M304) y clon STT3 de L. major 26B-a (M421). K significa lisina.

35

Figura 3. Mapa esquemático de los casetes de expresión de STT3.

Figura 4. Estructuras de glicano producidas en cepas Δalg3 .

Figura 5. Datos de actividad de proteasa normalizada de sobrenadantes de cultivo de los sobrenadantes de delección de proteasa y la cepa parental. La actividad de la proteasa se midió a pH 5,5 en las primeras 5 cepas y a pH 4,5 en las últimas tres cepas de delección. La actividad de proteasa es contra la caseína verde fluorescente. La cepa de delección de seis proteasas tiene solo el 6% de la cepa parental de tipo silvestre y la actividad de la proteasa de la cepa de delección de 7 proteasas fue aproximadamente un 40% menor que la actividad de la cepa de delección de 6 proteasas.

40

45 DESCRIPCIÓN DETALLADA

Definiciones

Como se usa en el presente documento, un "sistema de expresión" o una "célula huésped" se refiere a la célula que está modificada genéticamente para permitir la transcripción, traducción y plegamiento apropiado de un polipéptido o una proteína de interés, típicamente de proteína de mamífero.

50

El término "polinucleótido" u "oligonucleótido" o "ácido nucleico", como se usa en el presente documento, se refiere típicamente a un polímero de al menos dos nucleótidos unidos por un enlace fosfodiéster y puede consistir en ribonucleótidos o desoxinucleótidos o sus derivados que pueden introducirse en una célula huésped para la modificación genética de dicha célula huésped. Por ejemplo, un polinucleótido puede codificar una secuencia codificante de una proteína, y/o comprender secuencias de control o reguladoras de una secuencia codificante de una proteína, tales como secuencias potenciadoras o promotoras o terminador. Un polinucleótido puede comprender, por ejemplo, la secuencia codificante nativa de un gen o sus fragmentos, o secuencias variantes que se

55

han optimizado para la expresión génica óptima en una célula huésped específica (por ejemplo, para tener en cuenta el sesgo del codón).

Como se usa en el presente documento, el término "optimizado" con referencia a un polinucleótido significa que un polinucleótido se ha alterado para codificar una secuencia de aminoácidos usando codones que se prefieren en la célula u organismo de producción, por ejemplo, una célula fúngica filamentosa tal como una célula de *Trichoderma*. Las secuencias de nucleótidos heterólogas que se transfectan en una célula huésped típicamente se optimizan para retener completamente, o lo máximo posible, la secuencia de aminoácidos codificada originalmente por la secuencia de nucleótidos original (no optimizada). Las secuencias optimizadas en el presente documento se han diseñado para tener codones que se prefieren en la célula u organismo de producción correspondiente, por ejemplo, la célula fúngica filamentosa. Las secuencias de aminoácidos codificadas por secuencias de nucleótidos optimizadas también pueden denominarse optimizadas.

Como se usa en el presente documento, un "péptido" o un "polipéptido" es una secuencia de aminoácidos que incluye una pluralidad de residuos de aminoácidos polimerizados consecutivos. El péptido o polipéptido puede incluir residuos de aminoácidos modificados, residuos de aminoácidos de origen natural no codificados por un codón, y residuos de aminoácidos de origen no natural. Como se usa en el presente documento, una "proteína" puede referirse a un péptido o un polipéptido o una combinación de más de un péptido o polipéptido ensamblados juntos por enlaces covalentes o no covalentes. A menos que se especifique, el término "proteína" puede incluir una o más secuencias de aminoácidos con sus modificaciones postraduccionales, y en particular con modificaciones de O-manosilación o N-glicano.

Como se usa en el presente documento, el término "glucoproteína" se refiere a una proteína que comprende al menos un glicano ligado a N unido al menos a un residuo de asparagina de una proteína, o al menos una manosa unida al menos a una serina o treonina que da como resultado la O-manosilación. Dado que las glucoproteínas según se producen en un sistema de expresión de célula huésped se producen normalmente como una mezcla de diferentes patrones de glucosilación, los términos "glucoproteína" o "composición de glucoproteína" incluyen las mezclas de glucoproteínas según se producen por una célula huésped, con diferentes patrones de glucosilación, a menos que se definan específicamente.

Los términos "N-glucosilación" u "actividad de oligosacaril transferasa" se usan en el presente documento para referirse al enlace covalente de al menos una cadena de oligosacáridos al nitrógeno de amida de cadena lateral del residuo de asparagina (Asn) de un polipéptido.

Como se usa en el presente documento, "glicano" se refiere a una cadena de oligosacáridos que puede unirse a un vehículo tal como un aminoácido, péptido, polipéptido, lípido o un conjugado de extremo reductor. En ciertas formas de realización, la divulgación se refiere a glicanos unidos a N ("N-glicano") conjugados a un sitio de N-glucosilación polipeptídica tal como -Asn-Xaa-Ser/Thr- mediante un enlace N a nitrógeno de amida de cadena lateral de un residuo de asparagina (Asn), donde Xaa es cualquier residuo de aminoácido excepto Pro. La divulgación puede relacionarse adicionalmente con los glicanos como parte de las estructuras lipídicas precursoras de dolicol-fosfoligosacáridos (DoI-P-P-OS), que son precursores de los glicanos unidos a N en el retículo endoplásmico de células eucariotas. Los oligosacáridos precursores están unidos desde su extremo reductor a dos residuos de fosfato en el lípido dolicol. Por ejemplo, la α 3-manosiltransferasa Alg3 modifica el precursor de DoI-P-P-oligosacáridos de los N-glicanos. Generalmente, las estructuras de glicano descritas en el presente documento son estructuras terminales de glicano, donde los residuos no reductores no están modificados por otro residuo o residuos de monosacáridos.

Como se usa a lo largo de la presente divulgación, la nomenclatura de glucolípidos y carbohidratos es esencialmente de acuerdo con las recomendaciones de la IUPAC-IUB Commission on Biochemical Nomenclature (por ejemplo, Carbohydrate Res. 1998, 312, 167; Carbohydrate Res. 1997, 297, 1; Eur. J. Biochem. 1998, 257, 29). Se supone que Gal (galactosa), Glc (glucosa), GlcNAc (N-acetilglucosamina), GalNAc (N-acetilgalactosamina), Man (manosa) y Neu5Ac son de la configuración D, Fuc de la configuración L, y todas las unidades de monosacáridos en forma de piranosa (D-Galp, D-Glcp, D-GlcpNAc, D-GalpNAc, D-Manp, L-Fucp, D-Neup5Ac). El grupo amina es como se define para la galactosa natural y las glucosaminas en la posición 2 de GalNAc o GlcNAc. Los enlaces glucosídicos se muestran parcialmente en la nomenclatura más corta y en parte en la más larga, los enlaces de los residuos de SA/Neu5X de ácido siálico α 3 y α 6 significan lo mismo que α 2-3 y α 2-6, respectivamente, y para los residuos de monosacáridos de hexosa α 1-3, α 1-6, β 1-2, β 1-3, β 1-4 y β 1-6 se pueden acortar como α 3, α 6, β 2, β 3, β 4 y β 6, respectivamente. Lactosamina se refiere a N-acetil-lactosamina de tipo II, Gal β 4GlcNAc, y/o N-acetil-lactosamina de tipo I. Gal β 3GlcNAc y ácido siálico (SA) se refieren a ácido N-acetilneuramínico (Neu5Ac), ácido N-glicolilneuramínico (Neu5Gc), o cualquier otro ácido siálico natural que incluya derivados de Neu5X. El ácido siálico

se conoce como NeuNX o Neu5X, donde preferiblemente X es Ac o Gc. Ocasionalmente, Neu5Ac/Gc/X puede denominarse NeuNAc/NeuNGc/NeuNX.

Los azúcares que constituyen típicamente los N-glicanos encontrados en la glucoproteína de mamífero, incluyen, sin limitación, N-acetilglucosamina (abreviada en lo sucesivo aquí como "GlcNAc"), manosa (abreviada en lo sucesivo aquí como "Man"), glucosa (abreviada en lo sucesivo aquí como "Glc"), galactosa (abreviada en lo sucesivo aquí como "Gal"), y ácido siálico (abreviada en lo sucesivo aquí como "Neu5Ac"). Los N-glicanos comparten un pentasacárido común conocido como la estructura "central" $\text{Man}_3\text{GlcNAc}_2$ ($\text{Man}_6(\text{Man}_3)\text{Man}\beta_4\text{GlcNA}\beta_4\text{GlcNAc}$, denominado Man3).

En algunas formas de realización, el glicano Man3 o su derivado $\text{Man}_6(\text{GlcNAc}\beta_2\text{Man}_3)\text{Man}\beta_4\text{GlcNA}\beta_4\text{GlcNAc}$ es la glucoforma principal. Cuando una fucosa está unida a la estructura central, del núcleo, preferiblemente unida en α_6 a GlcNAc de extremo reductor, el N-glicano o el núcleo de N-glicano, puede representarse como $\text{Man}_3\text{GlcNAc}_2(\text{Fuc})$. En una forma de realización, el N-glicano principal es $\text{Man}_3[\text{Man}_6(\text{Man}_3)\text{Man}_6]\text{Man}\beta_4\text{GlcNA}\beta_4\text{GlcNAc}$ (Man5).

Los N-glicanos de tipo híbrido preferidos comprenden $\text{GlcNAc}\beta_2\text{Man}_3[\text{Man}_6(\text{Man}_3)\text{Man}_6]\text{Man}\beta_4\text{GlcNA}\beta_4\text{GlcNAc}$ ("GlcNAcMan5"), o derivados b4-galactosilados de los mismos, glucoforma $\text{Gal}\beta_4\text{GlcNAcMan}_3$, G1, G2, o GalGlcNAcMan_5 .

Un "N-glicano complejo" se refiere a un N-glicano que tiene al menos un residuo de GlcNAc, opcionalmente por el residuo de $\text{GlcNAc}\beta_2$, en el brazo terminal de 1,3 manosa de la estructura central y al menos un residuo de GlcNAc, opcionalmente por el residuo de $\text{GlcNAc}\beta_2$, en el brazo de 1,6 manosa terminal de la estructura central.

Dichos N-glicanos complejos incluyen, sin limitación, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (también denominado como glucoforma GO), $\text{Gal}_1\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (también denominado como glucoforma G1), y $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (también denominado como glucoforma G2), y sus glucoformas fucosiladas centrales FG0, FG1 y FG2, respectivamente, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2(\text{Fuc})$, $\text{Gal}_1\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2(\text{Fuc})$, y $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2(\text{Fuc})$.

Como se usa en el presente documento, la expresión "N-glicano neutro" tiene su significado general en la técnica. Se refiere a N-glicanos no sialilados. En contraste, los N-glicanos sialilados son ácidos.

"Aumento" o "Actividad reducida de una enzima endógena": La célula fúngica filamentosa puede tener niveles de actividad aumentados o reducidos de diversas enzimas endógenas. Se puede proporcionar un nivel reducido de actividad mediante la inhibición de la actividad de la enzima endógena con un inhibidor, un anticuerpo, o similares. En ciertas formas de realización, la célula fúngica filamentosa está modificada genéticamente de manera que aumenta o reduce la actividad de diversas enzimas endógenas. "Modificado genéticamente" se refiere a cualquier método de ADN o ARN recombinante utilizado para crear una célula huésped procarionta o eucariota que expresa un polipéptido a niveles elevados, a niveles reducidos, o de forma mutada. En otras palabras, la célula huésped se ha transfectado, transformado o transducido con una molécula de polinucleótido recombinante y, por lo tanto, se ha alterado para que la célula altere la expresión de una proteína deseada.

Las "modificaciones genéticas" que dan como resultado una disminución o deficiencia en la expresión génica, en la función del gen, o en la función del producto génico (es decir, la proteína codificada por el gen) pueden denominarse inactivación (completa o parcial), desactivación, eliminación, alteración, interrupción, bloqueo, silenciamiento o regulación descendente, o atenuación de la expresión de un gen. Por ejemplo, una modificación genética en un gen que da como resultado una disminución en la función de la proteína codificada por dicho gen, puede ser el resultado de una eliminación completa del gen (es decir, el gen no existe, y por lo tanto, la proteína no existe), una mutación en el gen que da como resultado una traducción incompleta (interrupción) o sin traducción de la proteína (por ejemplo, la proteína no se expresa), o una mutación en el gen que disminuye o anula la función natural de la proteína (por ejemplo, se expresa una proteína que ha disminuido o no tiene actividad o acción enzimática). Más específicamente, la referencia a la disminución de la acción de las proteínas analizadas en el presente documento generalmente se refiere a cualquier modificación genética en la célula huésped en cuestión, lo que da como resultado una disminución de la expresión y/o la funcionalidad (actividad biológica) de las proteínas e incluye una actividad disminuida de las proteínas (por ejemplo, disminución de la catálisis), aumento de la inhibición o degradación de las proteínas, así como una reducción o eliminación de la expresión de las proteínas. Por ejemplo, la acción o actividad de una proteína puede disminuir bloqueando o reduciendo la producción de la proteína, reduciendo la acción de la proteína, o inhibiendo la acción de la proteína. También son posibles combinaciones de algunas de estas modificaciones. El bloqueo o la reducción de la producción de una proteína puede incluir colocar el

- gen que codifica la proteína bajo el control de un promotor que requiere la presencia de un compuesto inductor en el medio de crecimiento. Al establecer condiciones de tal forma que el inductor se agote del medio, la expresión del gen que codifica la proteína (y, por lo tanto, de la síntesis de proteínas) podría desactivarse. El bloqueo o la reducción de la acción de una proteína también podría incluir el uso de un enfoque de tecnología de escisión similar al descrito en la patente de Estados Unidos número 4.743.546. Para utilizar este enfoque, el gen que codifica la proteína de interés se clona entre secuencias genéticas específicas que permiten la escisión controlada y específica del gen del genoma. La escisión podría ser provocada, por ejemplo, por un cambio en la temperatura de cultivo del cultivo, como en la patente de Estados Unidos número 4.743.546, o por alguna otra señal física o nutricional.
- 10 En general, de acuerdo con la presente divulgación, se hace un aumento o una disminución en una característica dada de una proteína mutante o modificada (por ejemplo, actividad enzimática) con referencia a la misma característica de una proteína parental (es decir, normal, no modificada) que se deriva del mismo organismo (de la misma fuente o secuencia precursora), que se mide o se establece en las mismas condiciones o equivalentes. De manera similar, un aumento o disminución en una característica de una célula huésped modificada genéticamente (por ejemplo, la expresión y/o la actividad biológica de una proteína, o la producción de un producto) se realiza con referencia a la misma característica de una célula huésped de tipo silvestre de la misma especie, y preferiblemente la misma cepa, en las mismas condiciones o equivalentes. Dichas condiciones incluyen el ensayo o las condiciones de cultivo (por ejemplo, componentes del medio, temperatura, pH, etc.) en la que se mide la actividad de la proteína (por ejemplo, expresión o actividad biológica) u otra característica de la célula huésped, así como el tipo de ensayo utilizado, la célula huésped que se evalúa, etc. Como se ha analizado anteriormente, las condiciones equivalentes son condiciones (por ejemplo, condiciones de cultivo) que son similares, pero no necesariamente idénticas (por ejemplo, se pueden tolerar algunos cambios conservadores en las condiciones), y que no cambian sustancialmente el efecto sobre el crecimiento celular o la expresión enzimática o la actividad biológica en comparación con una comparación realizada en las mismas condiciones.
- 15 20 25 Preferiblemente, una célula huésped modificada genéticamente que tiene una modificación genética que aumenta o disminuye (reduce) la actividad de una proteína dada (por ejemplo, una proteasa) tiene un aumento o disminución, respectivamente, en la actividad o la acción (por ejemplo, expresión, producción y/o actividad biológica) de la proteína, en comparación con la actividad de la proteína en una célula huésped parental (que no tiene dicha modificación genética), de al menos aproximadamente el 5%, y más preferiblemente al menos aproximadamente el 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55 60%, 65%, 70%, 75 80%, 85 90%, 95%, o cualquier porcentaje, en todos los números enteros entre el 5% y el 100% (por ejemplo, 6%, 7%, 8%, etc.).
- 30 35 En otro aspecto de la divulgación, una célula huésped modificada genéticamente que tiene una modificación genética que aumenta o disminuye (reduce) la actividad de una proteína dada (por ejemplo, una proteasa) tiene un aumento o disminución, respectivamente, en la actividad o acción (por ejemplo, expresión, producción y/o actividad biológica) de la proteína, en comparación con la actividad de la proteína de tipo silvestre en una célula huésped parental, de al menos aproximadamente 2 veces, y más preferiblemente al menos aproximadamente 5 veces, 10 veces, 20 veces, 30 veces, 40 veces, 50 veces, 75 veces, 100 veces, 125 veces, 150 veces, o cualquier aumento en todos los números enteros comenzando desde al menos aproximadamente 2 veces (por ejemplo, 3 veces, 4 veces, 5 veces, 6 veces, etc.).
- 40 45 Como se usa en el presente documento, los términos "idéntico" o "identidad porcentual", en el contexto de dos o más secuencias de ácido nucleico o aminoácido, se refieren a dos o más secuencias o subsecuencias que son las mismas. Dos secuencias son "sustancialmente idénticas" si dos secuencias tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales (es decir, 29% de identidad, opcionalmente 30%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% o 100% de identidad en una región específica, o, cuando no se especifica, en toda la secuencia), cuando se comparan y se alinean para la máxima correspondencia sobre una ventana de comparación, o región designada de acuerdo con lo medido utilizando uno de los siguientes algoritmos de comparación de secuencias o por alineación manual e inspección visual. Opcionalmente, la identidad existe sobre una región que tiene al menos aproximadamente 50 nucleótidos (o 10 aminoácidos) de longitud, o más preferiblemente sobre una región que tiene de 100 a 500 o 1000 o más nucleótidos (o 20, 50, 200 o más aminoácidos) de longitud.
- 50 55 Para la comparación de secuencias, típicamente una secuencia actúa como una secuencia de referencia, con la que se comparan las secuencias de ensayo. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de ensayo y de referencia se introducen en un ordenador, las coordenadas de la subsecuencia se designan, si es necesario, y se designan los parámetros del programa de algoritmos de secuencia. Se pueden usar los parámetros predeterminados del programa, o se pueden designar parámetros alternativos. El algoritmo de comparación de

secuencia calcula entonces el porcentaje de identidades de secuencia para las secuencias de ensayo en relación con la secuencia de referencia, basándose en los parámetros del programa. Al comparar dos secuencias para la identidad, no es necesario que las secuencias sean contiguas, pero cualquier brecha conlleva una penalización que reducirá el porcentaje de identidad global. Para blastn, los parámetros predeterminados son penalización de apertura por Gap = 5, y penalización de extensión por Gap = 2. Para blastp, los parámetros predeterminados son penalización de apertura por Gap = 11, y penalización de extensión por Gap = 1.

Una "ventana de comparación", como se usa en el presente documento, incluye una referencia a un segmento de una cualquiera de las varias posiciones contiguas que incluyen, pero sin limitación, de 20 a 600, normalmente de aproximadamente 50 a aproximadamente 200, más normalmente de aproximadamente 100 a aproximadamente 150 en las que una secuencia puede compararse con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias estén alineadas de manera óptima. Los métodos de alineamiento de secuencias para comparación se conocen bien en la técnica. El alineamiento óptico de secuencias para comparación se puede realizar, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman (1981), mediante el algoritmo de alineamiento de homología de Needleman y Wunsch (1970) *J Mol Biol* 48(3):443-453, por el método de búsqueda de similitud de Pearson y Lipman (1988) *Proc Natl Acad Sci USA* 85(8):2444-2448, mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el paquete de software de Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI), o por alineamiento manual e inspección visual [véanse, por ejemplo, Brent et al., (2003) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (Rintou Ed)].

Dos ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul et al. (1997) *Nucleic Acids Res* 25(17):3389-3402 y Altschul et al. (1990) *J. Mol Biol* 215(3):403-410, respectivamente. El software para realizar los análisis BLAST está disponible públicamente a través del National Center for Biotechnology Information. El programa BLASTN (para secuencias de nucleótidos) utiliza como valores predeterminados una longitud de palabra (W) de 11, una expectativa (E) de 10, M = 5, N = -4 y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP utiliza como valores predeterminados una longitud de palabra de 3, y una expectativa (E) de 10, y alineamientos (B) de matriz de puntuación BLOSUM62 [véase Henikoff y Henikoff, (1992) *Proc Natl Acad Sci USA* 89(22):10915-10919] de 50, expectativa (E) de 10, M = 5, N = -4, y una comparación de ambas cadenas.

El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, (1993) *Proc Natl Acad Sci USA* 90(12):5873-5877). Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de la suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la cual una coincidencia entre dos secuencias de nucleótidos o aminoácidos se produciría por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de la suma más pequeña en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es menor de aproximadamente 0,2, más preferiblemente menor de aproximadamente 0,01, y más preferiblemente menor de aproximadamente 0,001.

"Variante funcional" o "gen homólogo funcional" como se usa en el presente documento se refiere a una secuencia codificante o una proteína que tiene una similitud de secuencia con una secuencia de referencia, típicamente, de al menos el 30%, 40%, 50%, 60%, 70%, 80%, 90% o el 95% de identidad con la secuencia codificante de referencia o proteína, y que conserva sustancialmente la misma función que dicha secuencia codificante de referencia o proteína. Una variante funcional puede conservar la misma función, pero con una actividad reducida o aumentada. Las variantes funcionales incluyen variantes naturales, por ejemplo, homólogos de diferentes especies o variantes artificiales, resultantes de la introducción de una mutación en la secuencia codificante. La variante funcional puede ser una variante con solo mutaciones modificadas de manera conservadora.

Las "mutaciones modificadas de manera conservadora", como se usan en el presente documento, incluyen sustituciones individuales, deleciones o adiciones a una secuencia de aminoácidos codificada que dan como resultado la sustitución de un aminoácido con un aminoácido químicamente similar. Las tablas de sustitución conservadora que proporcionan aminoácidos funcionalmente similares se conocen bien en la técnica. Dichas variantes modificadas de manera conservadora son, además, y no excluyen, variantes polimórficas, homólogos interespecies y alelos de la divulgación. Los siguientes ocho grupos contienen aminoácidos que son sustituciones conservadoras entre sí: 1) alanina (A), glicina (G); 2) ácido aspártico (D), ácido glutámico (E); 3) asparagina (N), glutamina (Q); 4) arginina (R), lisina (K); 5) isoleucina (I), leucina (L), metionina (M), valina (V); 6) fenilalanina (F), tirosina (Y), triptófano (W); 7) serina (S), treonina (T); y 8) cisteína (C), metionina (M) (véase, por ejemplo, Creighton,

Proteins (1984)).

Células fúngicas filamentosas

5 Como se usa en el presente documento, "células fúngicas filamentosas" incluyen células de todas las formas filamentosas de la subdivisión Eumycota y Oomycota (según se define por Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8ª edición, 1995, CAB International, University Press, Cambridge, Reino Unido). Las células fúngicas filamentosas se caracterizan en general por una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano, y otros polisacáridos complejos. El crecimiento vegetativo es por elongación de hifas y
10 el catabolismo carbónico es necesariamente aeróbico. Por el contrario, el crecimiento vegetativo por levaduras tales como *Saccharomyces cerevisiae* se produce por el brote de un talo unicelular y el catabolismo carbónico puede ser fermentativo.

Preferiblemente, la célula fúngica filamentosas no se ve afectada adversamente por la transducción de las secuencias
15 de ácido nucleico necesarias, la expresión posterior de las proteínas (por ejemplo, proteínas de mamíferos), o los compuestos intermedios resultantes. Los métodos generales para alterar genes y cultivar células fúngicas filamentosas se divulgan, por ejemplo, para *Penicillium*, en Kopke et al. (2010) Appl Environ Microbiol. 76(14):4664-74. doi: 10.1128/AEM.00670-10, para *Aspergillus*, en Maruyama y Kitamoto (2011), Methods in Molecular Biology, vol. 765, DOI10.1007/978-1-61779-197-0_27; para *Neurospora*, en Collopy et al. (2010) Methods Mol Biol.
20 2010;638:33-40. doi: 10.1007/978-1-60761-611-5_3; y para *Myceliophthora* o *Chrysosporium* PCT/NL2010/000045 y PCT/EP98/06496.

Los ejemplos de células fúngicas filamentosas adecuadas incluyen, sin limitación, células de una cepa de *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Scytalidium*,
25 *Thielavia*, *Tolyocladium*, o *Trichoderma/Hypocrea*.

En ciertas formas de realización, la célula fúngica filamentosas es de una cepa de *Trichoderma* sp., *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*,
30 *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, o *Tolyocladium*.

En algunas formas de realización, la célula fúngica filamentosas es una cepa de *Myceliophthora* o *Chrysosporium*, *Neurospora*, *Aspergillus*, *Fusarium* o *Trichoderma*.

35 Las células fúngicas de *Aspergillus* de la presente divulgación pueden incluir, sin limitación, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus clavatus*, *Aspergillus flavus*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, o *Aspergillus terreus*.

Las células fúngicas de *Neurospora* de la presente divulgación pueden incluir, sin limitación, *Neurospora crassa*.

40 Las células fúngicas de *Myceliophthora* de la presente divulgación pueden incluir, sin limitación, *Myceliophthora thermophila*.

En una forma de realización preferida, la célula fúngica filamentosas es una célula fúngica de *Trichoderma*. Las
45 células fúngicas de *Trichoderma* de la presente divulgación pueden derivarse de una cepa de *Trichoderma* de tipo silvestre o un mutante de la misma. Los ejemplos de células fúngicas de *Trichoderma* adecuadas incluyen, sin limitación, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, *Trichoderma atroviride*, *Trichoderma virens*, *Trichoderma viride*; y una forma sexual alternativa de las mismas (es decir, *Hypocrea*).

50 En una forma de realización más preferida, la célula fúngica filamentosas es *Trichoderma reesei*, y, por ejemplo, cepas derivadas de ATCC 13631 (QM 6a), ATCC 24449 (mutante de radiación 207 de QM 6a), ATCC 26921 (QM 9414; mutante de ATCC 24449), VTT-D-00775 (Selinheimo et al., FEBS J., 2006, 273: 4322-4335), Rut-C30 (ATCC 56765), RL-P37 (NRRL 15709) o aislado de *T. harzianum* T3 (Wolffhechel, H., 1989).

55 La divulgación descrita en el presente documento se refiere a una célula fúngica filamentosas, por ejemplo, seleccionada de células de *Trichoderma*, *Neurospora*, *Myceliophthora* o *Chrysosporium*, tal como una célula fúngica de *Trichoderma reesei*, que comprende:

- i. una o más mutaciones que reducen o eliminan una o más actividades de proteasa endógena en comparación con una célula fúngica filamentosa parental que no tiene dicha mutación o mutaciones,
- ii. un polinucleótido que codifica una subunidad catalítica heteróloga de la oligosacaril transferasa, y
- iii. un polinucleótido que codifica una glucoproteína heteróloga,

5

en la que dicha subunidad catalítica de oligosacaril transferasa se selecciona de subunidades catalíticas de oligosacaril transferasa de *Leishmania*.

Proteasas con actividad reducida

10

Se ha encontrado que la reducción de la actividad de la proteasa permite aumentar sustancialmente la producción de una proteína de mamífero heteróloga. De hecho, dichas proteasas encontradas en células fúngicas filamentosas que expresan una proteína heteróloga normalmente catalizan una degradación significativa de la proteína recombinante expresada. Por lo tanto, al reducir la actividad de las proteasas en las células fúngicas filamentosas que expresan una proteína heteróloga, la estabilidad de la proteína expresada aumenta, dando como resultado un mayor nivel de producción de la proteína y, en algunas circunstancias, una mejora de la calidad de la proteína producida (por ejemplo, de longitud completa en lugar de degradada).

15

Las proteasas incluyen, sin limitación, proteasas aspárticas, serina proteasas de tipo tripsina, proteasas de subtilisina, proteasas glutámicas, y proteasas de sedolisina. Dichas proteasas pueden identificarse y aislarse a partir de células fúngicas filamentosas y ensayarse para determinar si la reducción de su actividad afecta a la producción de un polipéptido recombinante a partir de la célula fúngica filamentosa. Los métodos para identificar y aislar proteasas son bien conocidos en la técnica e incluyen, sin limitación, cromatografía de afinidad, ensayos de zimogramas, y electroforesis en gel. Después, se puede ensayar una proteasa identificada eliminando el gen que codifica la proteasa identificada de una célula fúngica filamentosa que expresa un polipéptido recombinante, tal como un polipéptido heterólogo o de mamífero, y determinando si la delección produce una disminución en la actividad de la proteasa total de la célula, y un aumento en el nivel de producción del polipéptido recombinante expresado. Los métodos para eliminar genes, medir la actividad de proteasa total, y medir los niveles de proteína producida se conocen bien en la técnica e incluyen los métodos descritos en el presente documento.

20

25

Proteasas aspárticas

Las proteasas aspárticas son enzimas que usan un residuo de aspartato para la hidrólisis de los enlaces peptídicos en polipéptidos y proteínas. Típicamente, las proteasas aspárticas contienen dos residuos de aspartato altamente conservados en su sitio activo que son óptimamente activos a pH ácido. Las proteasas aspárticas de organismos eucariotas tal como los hongos *Trichoderma* incluyen pepsinas, catepsinas y reninas. Dichas proteasas aspárticas tienen una estructura de dos dominios, que se cree que surge de la duplicación génica ancestral. Consistente con tal evento de duplicación, el pliegue total de cada dominio es similar, aunque las secuencias de los dos dominios han comenzado a divergir. Cada dominio contribuye con uno de los residuos de aspartato catalíticos. El sitio activo está en una hendidura formada por los dos dominios de las proteasas aspárticas. Las proteasas aspárticas eucariotas incluyen además puentes disulfuro conservados, que pueden ayudar en la identificación de los polipéptidos como proteasas de ácido aspártico.

30

35

40

45

Se han identificado diez proteasas aspárticas en células fúngicas de *Trichoderma*: *pep1* (tre74156); *pep2* (tre53961); *pep3* (tre121133); *pep4* (tre77579), *pep5* (tre81004), y *pep7* (tre58669), *pep8* (tre122076), *pep9* (tre79807), *pep11* (121306), y *pep12* (tre119876).

50

Los ejemplos de proteasas aspárticas adecuadas incluyen, sin limitación, *pep1* de *Trichoderma reesei* (SEQ ID NO: 22), *pep2* de *Trichoderma reesei* (SEQ ID NO: 18), *pep3* de *Trichoderma reesei* (SEQ ID NO: 19); *pep4* de *Trichoderma reesei* (SEQ ID NO: 20), *pep5* de *Trichoderma reesei* (SEQ ID NO: 21) y *pep7* de *Trichoderma reesei* (SEQ ID NO:23), EGR48424 *pep8* de *Trichoderma reesei* (SEQ ID NO:85), *pep9* de *Trichoderma reesei* (SEQ ID NO:87), EGR49498 *pep11* de *Trichoderma reesei* (SEQ ID NO:86), EGR52517 *pep12* de *Trichoderma reesei* (de SEQ ID NO:35), y homólogos de las mismas. Los ejemplos de homólogos de proteasas *pep1*, *pep2*, *pep3*, *pep4*, *pep5*, *pep7*, *pep8*, *pep11* y *pep12* identificadas en otros organismos también se describen en el documento PCT/EP/2013/050186.

55 Serina proteasas de tipo tripsina

Las serina proteasas de tipo tripsina son enzimas con una especificidad de sustrato similar a la de la tripsina. Las serina proteasas de tipo tripsina usan un residuo de serina para la hidrólisis de los enlaces peptídicos en polipéptidos y proteínas. Típicamente, las serina proteasas de tipo tripsina escinden enlaces peptídicos después de

un residuo de aminoácido cargado positivamente. Las serina proteasas de tipo tripsina de organismos eucariotas tales como los hongos *Trichoderma*, incluyen la tripsina 1, la tripsina 2 y la mesotripsina. Dichas serina proteasas de tipo tripsina generalmente contienen una tríada catalítica de tres residuos de aminoácidos (tal como histidina, aspartato y serina) que forman un relé de carga que sirve para hacer que el sitio activo sea nucleófilo. Las serina proteasas de tripsina eucariotas incluyen además un "orificio de oxianión" formado por los átomos de hidrógeno de amida del esqueleto de glicina y serina, que pueden ayudar en la identificación de los polipéptidos como serina proteasas de tipo tripsina.

Se ha identificado una serina proteasa de tipo tripsina en células fúngicas de *Trichoderma*: tsp1 (tre73897). Como se analiza en el documento PCT/EP/2013/050186, se ha demostrado que tsp1 tiene un impacto significativo en la expresión de glucoproteínas recombinantes, tales como las inmunoglobulinas.

Los ejemplos de proteasas tsp1 adecuadas incluyen, sin limitación, tsp1 de *Trichoderma reesei* (SEQ ID NO: 24) y homólogos de las mismas. Los ejemplos de homólogos de proteasas tsp1 identificadas en otros organismos se describen en el documento PCT/EP/2013/050186.

Subtilisina proteasas

Las subtilisina proteasas son enzimas con especificidad de sustrato similar a la de la subtilisina. Las subtilisina proteasas usan un residuo de serina para la hidrólisis de los enlaces peptídicos en polipéptidos y proteínas. Generalmente, las subtilisina proteasas son serina proteasas que contienen una tríada catalítica de los tres aminoácidos aspartato, histidina y serina. La disposición de estos residuos catalíticos se comparte con la subtilisina prototípica de *Bacillus licheniformis*. Las subtilisina proteasas de organismos eucariotas tales como los hongos *Trichoderma*, incluyen furina, MBTPS1 y TPP2. Las serina proteasas de tipo tripsina eucariotas incluyen además un residuo de ácido aspártico en el orificio de oxianión.

Se han identificado siete subtilisina proteasas en células fúngicas de *Trichoderma*: slp1 (tre51365); slp2 (tre123244); slp3 (tre123234); slp5 (tre64719), slp6 (tre121495), slp7 (tre123865), y slp8 (tre58698). La subtilisina proteasa slp7 se parece también a la sedolisina proteasa tpp1.

Los ejemplos de proteasas slp adecuadas incluyen, sin limitación, slp1 de *Trichoderma reesei* (SEQ ID NO: 25), slp2 (SEQ ID NO: 26); slp3 (SEQ ID NO: 27); slp5 (SEQ ID NO: 28), slp6 (SEQ ID NO: 29), slp7 (SEQ ID NO: 30), y slp8 (SEQ ID NO: 31), y homólogos de las mismas. Los ejemplos de homólogos de proteasas slp identificadas en otros organismos se describen en el documento PCT/EP/2013/050186.

Proteasas glutámicas

Las proteasas glutámicas son enzimas que hidrolizan los enlaces peptídicos en polipéptidos y proteínas. Las proteasas glutámicas son insensibles a la pepstatina A, por lo que a veces se las denomina proteasas ácidas insensibles a la pepstatina. Si bien las proteasas glutámicas se agruparon previamente con las proteasas aspárticas y a menudo se denominaron proteasas ácidas, recientemente se ha encontrado que las proteasas glutámicas tienen residuos de sitios activos muy diferentes a las proteasas aspárticas.

Se han identificado dos proteasas glutámicas en células fúngicas de *Trichoderma*: gap1 (tre69555) y gap2 (tre106661).

Los ejemplos de proteasas gap adecuadas incluyen, sin limitación, gap de *Trichoderma reesei* (SEQ ID NO: 32), gap2 de *Trichoderma reesei* (SEQ ID NO: 33), y homólogos de las mismas. Los ejemplos de homólogos de proteasas gap identificadas en otros organismos se describen en el documento PCT/EP/2013/050186.

Sedolisina proteasas y homólogos de proteasas

Las sedolisina proteasas son enzimas que usan un residuo de serina para la hidrólisis de los enlaces peptídicos en polipéptidos y proteínas. Las sedolisina proteasas generalmente contienen una tríada catalítica única de serina, glutamato y aspartato. Las sedolisina proteasas también contienen un residuo de aspartato en el orificio de oxianión. Las sedolisina proteasas de organismos eucariotas tales como los hongos de *Trichoderma* incluyen la tripeptidil peptidasa.

Los ejemplos de proteasas tpp1 adecuadas incluyen, sin limitación, tpp1 de *Trichoderma reesei* tre82623 (SEQ ID

NO: 34) y homólogos de la misma. Los ejemplos de homólogos de proteasas tpp1 identificadas en otros organismos se describen en el documento PCT/EP/2013/050186.

5 Como se usa en referencia a la proteasa, el término "homólogo" se refiere a una proteína que tiene actividad de proteasa y muestra una similitud de secuencia con una secuencia de proteasa conocida (de referencia). Los homólogos pueden identificarse mediante cualquier método conocido en la técnica, preferiblemente, usando la herramienta BLAST para comparar una secuencia de referencia con una segunda secuencia o fragmento de una secuencia o con una base de datos de secuencias. Como se describe en la sección "Definiciones", BLAST comparará secuencias basadas en el porcentaje de identidad y similitud.

10 Preferiblemente, una proteasa homóloga tiene al menos un 30% de identidad con (opcionalmente un 30%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% o un 100% de identidad sobre una región especificada, o, cuando no se especifica, sobre toda la secuencia), cuando se compara con una de las secuencias de proteasa enumeradas anteriormente, incluyendo *pep1*, *pep2*, *pep3*, *pep4*, *pep5*, *pep7*, *pep8*, *pep9*, *pep11*, *pep12*,
15 *tsp1*, *slp1*, *slp2*, *slp3*, *slp5*, *slp6*, *slp7*, *slp8*, *tpp1*, *gap1* y *gap2* de *T. reesei*. Las proteasas homólogas correspondientes de *N. crassa* y *M. thermophila* se muestran en la SEQ ID NO: 136-169.

Reducción de la actividad de proteasas en la célula fúngica filamentosa de la divulgación

20 Las células fúngicas filamentosas de acuerdo con la divulgación tienen actividad reducida de al menos una proteasa endógena, típicamente 2, 3, 4, 5 o más, con el fin de mejorar la estabilidad y la producción de la proteína con una mayor ocupación del sitio de N-glucosilación en dicha célula fúngica filamentosa, preferiblemente en una célula de *Trichoderma*.

25 La actividad de proteasa total se puede medir de acuerdo con los métodos estándar en la técnica y, por ejemplo, como se describe en el presente documento utilizando el kit de ensayo de proteasa (kit de ensayo de proteasa QuantiCleave, Pierce N.º 23263) con caseína succinilada como sustrato.

La actividad de las proteasas encontradas en células fúngicas filamentosas puede reducirse por cualquier método
30 conocido por los expertos en la técnica. En algunas formas de realización, la actividad reducida de las proteasas se consigue reduciendo la expresión de la proteasa, por ejemplo, mediante la modificación del promotor o ARNi.

En formas de realización adicionales, la expresión reducida o eliminada de las proteasas es el resultado de polinucleótidos antisentido o construcciones de ARNi que son específicas para cada uno de los genes que codifican
35 cada una de las proteasas. En una forma de realización, una construcción de ARNi es específica para un gen que codifica una proteasa aspártica tal como una proteasa de tipo pep, una serina proteasas de tipo tripsina, tal como una *tsp1*, una proteasa glutámica, tal como una proteasa de tipo gap, una subtilisina proteasa, tal como una proteasa de tipo slp, o una sedolisina proteasa, tal como una proteasa tpp1 o slp7. En una forma de realización, una construcción de ARNi es específica para el gen que codifica una proteasa de tipo slp. En una forma de realización,
40 una construcción de ARNi es específica para dos o más proteasas. En una forma de realización, dos o más proteasas son una cualquiera de las proteasas de tipo pep, una cualquiera de las serina proteasas de tipo tripsina, una cualquiera de las proteasas de tipo slp, una cualquiera de las proteasas de tipo gap y/o una cualquiera de las sedolisina proteasas. En una forma de realización, dos o más proteasas son slp2, slp3, slp5 y/o slp6. En una forma
45 de realización, la construcción de ARNi comprende una cualquiera de las siguientes secuencias de ácido nucleico (véase también el documento PCT/EP/2013/050186).

Secuencia diana de ARNi
GCACACTTTCAAGATTGGC (SEQ ID NO: 15)
GTACGGTGTGCCAAGAAG (SEQ ID NO: 16)
GTTGAGTACATCGAGCGGACAGCATTGTGCACACCATGCTTCCCCTCGAGTC CAAGGACAGCATCATCGTTGAGGACTCGTGCAACGGCGAGACGGAGAAGCAG GCTCCCTGGGGTCTTGCCCGTATCTCTCACCGAGAGACGCTCAACTTTGGCTC CTTCAACAAGTACCTCTACACCGCTGATGGTGGTGAGGGTGTGATGCCTATGT CATTGACACCGGCACCAACATCGAGCACGTGCGACTTTGAGGGTCTGCGCAAGT

```
GGGGCAAGACCATCCCTGCCGGCGATGAGGACGAGGACGGCAACGGCCACG
GCACTCACTGCTCTGGTACCGTTGCTGGTAAGAAGTACGGTGTGGCCAAGAAG
GCCACGTCTACGCCGTCAAGGTGCTCCGATCCAACGGATCCGGCACCATGTC
TGACGTCGTCAAGGGCGTCGAGTACG (SEQ ID NO:17)
```

En otras formas de realización, la actividad reducida de las proteasas se consigue modificando el gen que codifica la proteasa. Los ejemplos de dichas modificaciones incluyen, sin limitación, una mutación, tal como una eliminación o alteración del gen que codifica dicha actividad de proteasa endógena.

5

Por consiguiente, la divulgación se refiere a una célula fúngica filamentosa, tal como una célula de *Trichoderma*, que tiene una mutación que reduce o elimina al menos una actividad de proteasa endógena en comparación con una célula fúngica filamentosa parental que no tiene tal mutación deficiente en proteasa, comprendiendo dicha célula fúngica filamentosa además un polinucleótido que codifica una subunidad catalítica heteróloga de la oligosacaril

10 transferasa de *Leishmania*.

La mutación por delección o interrupción incluye, sin limitación, mutación knock-out, una mutación de truncamiento, una mutación puntual, una mutación sin sentido, una mutación de sustitución, una mutación de desplazamiento de marco, una mutación de inserción, una mutación de duplicación, una mutación de amplificación, una mutación de translocación, o una mutación de inversión, y que da como resultado una reducción de la actividad de proteasa correspondiente. Los métodos para generar al menos una mutación en un gen codificante de proteasas de interés se conocen bien en la técnica e incluyen, sin limitación, mutagénesis aleatoria y cribado, mutagénesis de sitio dirigido, mutagénesis por PCR, mutagénesis por inserción, mutagénesis química e irradiación.

15

20 En ciertas formas de realización, se modifica una porción del gen que codifica la proteasa, tal como la región que codifica el dominio catalítico, la región codificante, o una secuencia de control requerida para la expresión de la región codificante. Tal secuencia de control del gen puede ser una secuencia promotora o una parte funcional de la misma, es decir, una parte que sea suficiente para afectar a la expresión del gen. Por ejemplo, una secuencia promotora puede inactivarse dando como resultado que no haya expresión o un promotor más débil puede sustituirse por la secuencia promotora nativa para reducir la expresión de la secuencia codificante. Otras secuencias de control para una posible modificación incluyen, sin limitación, una secuencia líder, una secuencia propeptídica, una secuencia señal, un terminador de la transcripción, y un activador de la transcripción.

25

Los genes que codifican proteasa que están presentes en células fúngicas filamentosas también pueden modificarse utilizando técnicas de eliminación de genes para eliminar o reducir la expresión del gen. Las técnicas de delección génica permiten la eliminación parcial o completa del gen, eliminando así su expresión. En tales métodos, la eliminación del gen se puede lograr mediante recombinación homóloga utilizando un plásmido que se ha construido para contener de forma contigua las regiones 5' y 3' que flanquean el gen.

30

35 Los genes que codifican proteasas que están presentes en células fúngicas filamentosas también pueden modificarse introduciendo, sustituyendo y/o eliminando uno o más nucleótidos en el gen, o una secuencia de control de los mismos requerida para la transcripción o traducción del gen. Por ejemplo, los nucleótidos pueden insertarse o eliminarse para la introducción de un codón de terminación, la eliminación del codón de inicio, o un desplazamiento de marco del marco de lectura abierto. Tal modificación se puede lograr mediante métodos conocidos en la técnica, que incluyen, sin limitación, mutagénesis de sitio dirigido y mutagénesis generada por peR (véanse, por ejemplo, 40 Botstein y Shortie, 1985, *Science* 229: 4719; Lo et al., 1985, *Proceedings of the National Academy of Sciences USA* 81: 2285; Higuchi et al., 1988, *Nucleic Acids Research* 16: 7351; Shimada, 1996, *Meth. Mol. Bioi.* 57: 157; Ho et al., 1989, *Gene* 77: 61; Horton et al., 1989, *Gene* 77: 61; y Sarkar y Sommer, 1990, *BioTechniques* 8: 404).

40

45 Además, los genes que codifican la proteasa que están presentes en las células fúngicas filamentosas pueden modificarse mediante técnicas de alteración génica insertando en el gen una construcción disruptiva de ácido nucleico que contiene un fragmento de ácido nucleico homólogo al gen que creará una duplicación de la región de homología e incorporará una construcción de ácido nucleico entre las regiones duplicadas. Tal alteración génica puede eliminar la expresión del gen si la construcción insertada separa el promotor del gen de la región codificante o 50 interrumpe la secuencia codificante de tal forma que da como resultado un producto génico no funcional. Una construcción disruptiva puede ser simplemente un gen marcador seleccionable acompañado de las regiones 5' y 3' homólogas al gen. El marcador seleccionable permite la identificación de transformantes que contienen el gen alterado.

50

Los genes que codifican proteasa que están presentes en células fúngicas filamentosas también pueden modificarse por el proceso de conversión génica (véase, por ejemplo, Iglesias y Trautner, 1983, Molecular General Genetics 189:5 73-76). Por ejemplo, en la conversión del gen, una secuencia de nucleótidos correspondiente al gen se mutageniza *in vitro* para producir una secuencia de nucleótidos defectuosa, que luego se transforma en una cepa de *Trichoderma* para producir un gen defectuoso. Por recombinación homóloga, la secuencia de nucleótidos defectuosa reemplaza al gen endógeno. Puede ser deseable que la secuencia de nucleótidos defectuosa también contenga un marcador para la selección de transformantes que contienen el gen defectuoso.

Los genes codificantes de proteasa de la presente divulgación que están presentes en células fúngicas filamentosas que expresan un polipéptido recombinante también pueden modificarse mediante técnicas antisentido establecidas que utilizan una secuencia de nucleótidos complementaria a la secuencia de nucleótidos del gen (véase, por ejemplo, Parish y Stoker, 1997, FEMS Microbiology Letters 154: 151-157). En particular, la expresión del gen por células fúngicas filamentosas puede reducirse o inactivarse introduciendo una secuencia de nucleótidos complementaria a la secuencia de nucleótidos del gen, que puede transcribirse en la cepa y es capaz de hibridar con el ARNm producido en las células. En condiciones que permiten que la secuencia de nucleótidos antisentido complementaria se hibride con el ARNm, por lo tanto, la cantidad de proteína traducida se reduce o elimina.

Los genes que codifican proteasas que están presentes en células fúngicas filamentosas también pueden modificarse por mutagénesis aleatoria o específica usando métodos bien conocidos en la técnica, incluyendo, sin limitación, mutagénesis química (véase, por ejemplo, Hopwood, The Isolation of Mutants in Methods in Microbiology (J.R. Norris y D.W. Ribbons, eds.) págs. 363-433, Academic Press, Nueva York, 25 1970). La modificación del gen se puede realizar sometiendo las células fúngicas filamentosas a mutagénesis y cribando las células mutantes en las que la expresión del gen se ha reducido o inactivado. La mutagénesis, que puede ser específica o aleatoria, puede realizarse, por ejemplo, mediante el uso de un agente mutagenizante físico o químico adecuado, el uso de un oligonucleótido adecuado, sometiendo la secuencia de ADN a mutagénesis generada por PCR, o cualquier combinación de las mismas. Los ejemplos de agentes mutagenizantes físicos y químicos incluyen, sin limitación, irradiación ultravioleta (UV), hidroxilamina, N-metil-N'-nitro-N-nitrosoguanidina (MNNG), N-metil-N'-nitrosoguanidina (NTG) O-metil hidroxilamina, ácido nitroso, metanosulfonato de etil (EMS), bisulfito sódico, ácido fórmico y análogos de nucleótidos. Cuando se usan dichos agentes, la mutagénesis se realiza típicamente incubando las células fúngicas filamentosas, tales como células de *Trichoderma*, a mutagenizar en presencia del agente mutagenizante de elección en condiciones adecuadas, y luego seleccionando mutantes que muestran una expresión reducida o nula del gen.

En ciertas formas de realización, la al menos una mutación o modificación en un gen codificante de proteasa de la presente divulgación da como resultado una proteasa modificada que no tiene actividad de proteasa detectable. En otras formas de realización, la al menos una modificación en un gen que codifica proteasa de la presente divulgación da como resultado una proteasa modificada que tiene al menos un 25% menos, al menos un 50% menos, al menos un 75% menos, al menos un 90%, al menos 95%, o un mayor porcentaje menos de actividad de proteasa en comparación con una proteasa no modificada correspondiente.

Las células fúngicas filamentosas o células fúngicas de *Trichoderma* de la presente divulgación pueden tener actividad proteasa reducida o no detectable de al menos tres, o al menos cuatro proteasas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, pep8, pep9, pep11, pep12, tsp1, slp1, slp2, slp3, slp5, slp6, slp7, gap1 y gap2. En una forma de realización preferida, una célula fúngica filamentosa de acuerdo con la divulgación es una célula fúngica filamentosa que tiene una delección o alteración en al menos 3 o 4 proteasas endógenas, dando como resultado una actividad no detectable para dichas proteasas endógenas eliminadas o alteradas y que además comprende un polinucleótido que codifica una subunidad catalítica heteróloga de oligosacril transferasa de *Leishmania*.

En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en pep1, tsp1, y slp1. En otras formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en gap1, slp1, y pep1. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1 y gap1. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1 y pep4. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4 y slp1. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, y slp3. En ciertas formas de realización, la célula fúngica filamentosa o la célula de

Trichoderma, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, slp3, y pep3. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, slp3, pep3 y pep2. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, slp3, pep3, pep2 y pep5. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, slp3, pep3, pep2, pep5 y tsp1. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, slp3, pep3, pep2, pep5, tsp1 y slp7. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, slp3, pep3, pep2, pep5, tsp1, slp7 y slp8. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en slp2, pep1, gap1, pep4, slp1, slp3, pep3, pep2, pep5, tsp1, slp7, slp8 y gap2. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en al menos tres proteasas endógenas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, pep8, pep9, pep11, pep12, tsp1, slp2, slp3, slp7, gap1 y gap2. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en al menos tres a seis proteasas endógenas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, tsp1, slp1, slp2, slp3, gap1 y gap2. En ciertas formas de realización, la célula fúngica filamentosa o la célula de *Trichoderma*, tiene una actividad de proteasa reducida o no detectable en al menos siete a diez proteasas endógenas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, pep7, pep8, tsp1, slp1, slp2, slp3, slp5, slp6, slp7, slp8, tpp1, gap1 y gap2.

Expresión de subunidades catalíticas heterólogas de oligosacaril transferasa en células fúngicas filamentosas

Como se usa en el presente documento, la expresión "oligosacaril transferasa" u OST se refiere al complejo enzimático que transfiere un oligosacárido de 14 azúcares de dolicol a la proteína naciente. Es un tipo de glucosiltransferasa. El azúcar Glc3Man9GlcNAc2 está unido a un residuo de asparagina (Asn) en la secuencia Asn-X-Ser o Asn-X-Thr, donde X es cualquier aminoácido excepto prolina. Esta secuencia se denomina un secuón de glucosilación. La reacción catalizada por OST es la etapa central en la ruta de glucosilación ligada a N.

En la mayoría de los eucariotas, OST es un complejo hetero-oligomérico compuesto por ocho proteínas diferentes, en el que se cree que el componente STT3 es la subunidad catalítica.

De acuerdo con la presente descripción, la subunidad catalítica heteróloga de oligosacaril transferasa se selecciona de subunidades catalíticas de oligosacaril transferasa de *Leishmania*. Hay cuatro parálogos de STT3 en el protozoo parásito *Leishmania*, denominados STT3A, STT3B, STT3C y STT3D.

En una forma de realización, la subunidad catalítica heteróloga de oligosacaril transferasa es STT3D de *Leishmania major* (que tiene la secuencia de aminoácidos que se expone en la SEQ ID No:1).

En otra forma de realización, la subunidad catalítica heteróloga de oligosacaril transferasa es STT3D de *Leishmania infantum* (que tiene la secuencia de aminoácidos que se expone en la SEQ ID No:8).

En otra forma de realización, la subunidad catalítica heteróloga de oligosacaril transferasa es STT3D de *Leishmania braziliensis* (que tiene la secuencia de aminoácidos que se expone en la SEQ ID No:89).

En otra forma de realización, la subunidad catalítica heteróloga de oligosacaril transferasa es STT3D de *Leishmania mexicana* (que tiene la secuencia de aminoácidos que se expone en la SEQ ID No:91).

En otra forma de realización más, la subunidad catalítica heteróloga de la oligosacaril transferasa es un polipéptido variante funcional que tiene al menos un 50%, preferiblemente al menos un 60%, incluso más preferiblemente al menos un 70%, 80%, 90%, 95% de identidad con SEQ ID NO: 1, SEQ ID NO: 8, SEQ ID NO: 89 o SEQ ID NO: 91.

En otra forma de realización más, la subunidad catalítica heteróloga de la oligosacaril transferasa es un polipéptido variante funcional que tiene al menos un 50%, preferiblemente al menos un 60%, incluso más preferiblemente al menos un 70%, 80%, 90%, 95% de identidad con SEQ ID NO: 1 o SEQ ID NO: 8.

En una forma de realización de la presente divulgación, el polinucleótido que codifica la subunidad catalítica

heteróloga de oligosacaril transferasa comprende la SEQ ID NO:2.

SEQ ID NO:2 es una versión optimizada por codón del gen STT3D de *L major* (gi389594572|XM_003722461.1).

- 5 En una forma de realización de la presente divulgación, el polinucleótido que codifica la subunidad catalítica heteróloga de oligosacaril transferasa comprende la SEQ ID NO:9.

SEQ ID NO:9 es una versión optimizada por codón del gen STT3D de *L major* (gi339899220|XM_003392747.1).

- 10 En una forma de realización de la divulgación, el polinucleótido que codifica la subunidad catalítica heteróloga de oligosacaril transferasa comprende la SEQ ID NO:88 o una variante o la SEQ ID NO: 88 que se ha optimizado por codón para la expresión en células fúngicas filamentosas tal como *Trichoderma reesei*.

- En una forma de realización de la divulgación, el polinucleótido que codifica la subunidad catalítica heteróloga de oligosacaril transferasa comprende la SEQ ID NO:90 o una variante o la SEQ ID NO: 90 que se ha optimizado por codón para la expresión en células fúngicas filamentosas tal como *Trichoderma reesei*.

- En una forma de realización de la divulgación, el polinucleótido que codifica una subunidad catalítica heteróloga de oligosacaril transferasa comprende un polinucleótido que codifica un polipéptido variante funcional de STT3D de *Leishmania major*, *Leishmania infantum*, *Leishmania braziliensis* o *Leishmania mexicana* que tiene al menos un 50%, preferiblemente al menos un 60%, incluso más preferiblemente al menos un 70%, 80%, 90%, 95% de identidad con la SEQ ID NO:1, SEQ ID NO:8, SEQ ID NO: 89 o SEQ ID NO: 91.

- En una forma de realización de la divulgación, el polinucleótido que codifica una subunidad catalítica heteróloga de oligosacaril transferasa comprende un polinucleótido que codifica un polipéptido variante funcional de STT3D de *Leishmania major* o *Leishmania infantum* que tiene al menos un 50%, preferiblemente al menos un 60%, incluso más preferiblemente al menos un 70%, 80%, 90%, 95% de identidad con la SEQ ID NO:1 o SEQ ID NO:8.

- En una forma de realización de la divulgación, el polinucleótido que codifica una subunidad catalítica heteróloga de oligosacaril transferasa está bajo el control de un promotor para la expresión constitutiva de dicha oligosacaril transferasa es dicha célula fúngica filamentosas.

- Los promotores que se pueden usar para la expresión de la oligosacaril transferasa incluyen promotores constitutivos tales como *gpd* o *cDNA1*, promotores de enzimas de glucosilación endógena y glucosiltransferasas tales como manosiltransferasas que sintetizan N-glicanos en el Golgi o el RE, y promotores inducibles de proteínas endógenas de alto rendimiento, tal como el promotor *cbh1*.

En una forma de realización de la divulgación, dicho promotor es el promotor *cDNA1* de *Trichoderma reesei*.

40 Aumento de la ocupación del sitio de N-glucosilación en células fúngicas filamentosas de la divulgación

Las células fúngicas filamentosas de acuerdo con la divulgación tienen una actividad aumentada de oligosacárido transferasa, con el fin de aumentar la ocupación del sitio de N-glucosilación.

- 45 La ocupación del sitio de N-glucosilación se puede medir mediante métodos estándar en la técnica (por ejemplo, Schulz y Aebi (2009) Analysis of Glycosylation Site Occupancy Reveals a Role for Ost3p and Ost6p in Site-specific N-Glycosylation Efficiency, Molecular & Cellular Proteomics, 8:357-364, o Millward et al. (2008), Effect of constant and variable domain glycosylation on pharmacokinetics of therapeutic antibodies in mice, Biologicals, 36:41-47, Forno et al. (2004) N- and O-linked carbohydrates and glycosylation site occupancy in recombinant human granulocyte-macrophage colony-stimulating factor secreted by a Chinese hamster ovary cell line, Eur. J. Biochem. 271: 907-919) o métodos como se describen en el presente documento en los Ejemplos.

- La ocupación del sitio de N-glucosilación se refiere al porcentaje molar (o % en moles) de las glucoproteínas heterólogas que están N-glucosiladas con respecto al número total de glucoproteínas heterólogas producidas por la célula fúngica filamentosas (como se describe en el Ejemplo 1 a continuación).

En una forma de realización de la divulgación, la ocupación del sitio de N-glucosilación es de al menos el 95%, y las glucoformas Man3, Man5, G0, G1 y/o G2 representan al menos el 50% de los N-glicanos neutros totales de la glucoproteína heteróloga.

El porcentaje de diversas glucoformas con respecto a los N-glicanos neutros totales de la glucoproteína heteróloga se puede medir, por ejemplo, como se describe en el documento WO2012069593.

- 5 En una forma de realización, la proteína heteróloga con ocupación del sitio de N-glicosilación aumentada se selecciona del grupo que consiste en:
- a) una inmunoglobulina, tal como IgG,
 - b) una cadena ligera o pesada de una inmunoglobulina,
 - 10 c) una cadena pesada o una cadena ligera de un anticuerpo,
 - d) un anticuerpo monocatenario,
 - e) un anticuerpo camélido,
 - f) un anticuerpo monomérico o multimérico de un solo dominio,
 - 15 g) un fragmento FAb, un fragmento FAb2, y
 - h) sus fragmentos de unión a antígeno.

Métodos para producir glucoproteínas con un aumento de la ocupación del sitio de N-glicosilación y N-glicanos de tipo mamífero

- 20 Las células fúngicas filamentosas de acuerdo con la presente divulgación pueden ser útiles, en particular, para producir una composición de glucoproteínas heteróloga, tal como una composición de anticuerpo, con una mayor ocupación del sitio de N-glicosilación y N-glicanos de tipo mamífero, tales como complejos de N-glicanos. Por consiguiente, en un aspecto, la célula fúngica filamentosa se modifica genéticamente adicionalmente para producir un N-glicano de tipo mamífero, permitiendo de este modo la producción *in vivo* de glucoproteína o composición de anticuerpos con una mayor ocupación del sitio de N-glicosilación y con N-glicano de tipo mamífero como glucoformas principales de dicha glucoproteína o anticuerpo.

En ciertas formas de realización, este aspecto incluye métodos para producir glucoproteínas o anticuerpos con N-glicanos de tipo mamífero en una célula de *Trichoderma*.

- 30 En cierta forma de realización, la glucoproteína o anticuerpo comprende, como glucoforma principal, el N-glicano de tipo mamífero que tiene la fórmula $[[\text{Gal}\beta 4]_x \text{GlcNAc}\beta 2]_z \text{Man}\alpha 3([\{\text{Gal}\beta 4\}_y \text{GlcNAc}\beta 2]_w \text{Man}\alpha 6) \text{Man}\beta 4 \text{GlcNAc}\beta [\text{Fuc}\alpha 6]_a \text{GlcNAc}$, donde $()$ define una ramificación en la estructura, donde $[\]$ o $\{ \}$ definen una parte de la estructura del glicano presente o ausente en una secuencia lineal, y donde a, x, y, z y w son 0 o 1, independientemente. En una forma de realización, w y z son 1, y x e y son 0 para una estructura G0 no galactosilada; tanto x como y son 1 para una estructura G2; y solo uno de x o y es 1 para una estructura G1. Cuando a es 1, la estructura está fucosilada en el núcleo, tal como un glicano FG0, FG1 o FG2.

- 40 En ciertas formas de realización, la glucoproteína o el anticuerpo comprende, como una glucoforma principal, un N-glicano de tipo mamífero seleccionado del grupo que consiste en:

- i. $\text{Man}\alpha 3[\text{Man}\alpha 6(\text{Man}\alpha 3)\text{Man}\alpha 6]\text{Man}\beta 4 \text{GlcNA}\beta 4 \text{GlcNAc}$ (glucoforma Man5);
- 45 ii. $\text{GlcNAc}\beta 2 \text{Man}\alpha 3[\text{Man}\alpha 6(\text{Man}\alpha 3)\text{Man}\alpha 6]\text{Man}\beta 4 \text{GlcNA}\beta 4 \text{GlcNAc}$ (glucoforma GlcNAcMan5);
- iii. $\text{Man}\alpha 6(\text{Man}\alpha 3)\text{Man}\beta 4 \text{GlcNA}\beta 4 \text{GlcNAc}$ (glucoforma Man3);
- iv. $\text{Man}\alpha 6(\text{GlcNAc}\beta 2 \text{Man}\alpha 3)\text{Man}\beta 4 \text{GlcNA}\beta 4 \text{GlcNAc}$ (GlcNAcMan3) o,
- v. N-glicanos de tipo complejo seleccionados de la glucoforma G0, G1, o G2.

- 50 En una forma de realización, la glucoproteína o la composición del anticuerpo con N-glicanos de tipo mamífero, preferiblemente producidos por una cepa knock-out de *alg3*, incluyen glucoformas que esencialmente carecen o están desprovistas de glicanos $\text{Man}\alpha 3[\text{Man}\alpha 6(\text{Man}\alpha 3)\text{Man}\alpha 6]\text{Man}\beta 4 \text{GlcNA}\beta 4 \text{GlcNAc}$ (Man5). En formas de realización específica, la célula fúngica filamentosa produce glucoproteínas heterólogas o anticuerpos con, como glucoforma principal, la estructura de N-glicano trimanosilo $\text{Man}\alpha 3[\text{Man}\alpha 6]\text{Man}\beta 4 \text{GlcNAc}\beta 4 \text{GlcNAc}$. En otras formas de realización, la célula fúngica filamentosa produce glucoproteínas o anticuerpos con, como glucoforma principal, la estructura de N-glicano GO $\text{GlcNAc}\beta 2 \text{Man}\alpha 3[\text{GlcNAc}\beta 2 \text{Man}\alpha 6]\text{Man}\beta 4 \text{GlcNAc}\beta 4 \text{GlcNAc}$.

En ciertas formas de realización, la célula fúngica filamentosa de la divulgación produce una glucoproteína o composición de anticuerpo con una mezcla de diferentes N-glicanos.

En algunas formas de realización, el N-glicano $\text{Man}_3\text{GlcNAc}_2$ (es decir, $\text{Man}\alpha_3[\text{Man}\alpha_6]\text{Man}\beta_4\text{GlcNAc}\beta_4\text{GlcNAc}$) representa al menos el 10%, al menos el 20%, al menos el al menos el 30%, al menos el 40%, al menos el 50%, al menos el 60%, al menos el 70%, al menos el 80%, al menos el 90% o más de los N-glicanos neutros totales (% en moles) de la glucoproteína heteróloga o anticuerpo, como se expresa en las células fúngicas filamentosas de la
5 divulgación.

En otras formas de realización, el N-glicano $\text{GlcNAc}_2\text{Man}_3$ (por ejemplo, $\text{G}_0\text{GlcNAc}\beta_2\text{Man}\alpha_3[\text{GlcNAc}\beta_2\text{Man}\alpha_6]\text{Man}\beta_4\text{GlcNAc}\beta_4\text{GlcNAc}$) representa al menos el 10%, al menos el 20%, al menos el al menos el 30%, al menos el 40%, al menos el 50%, al menos el 60%, al menos el 70%, al menos el 80%,
10 al menos el 90% o más de los N-glicanos neutros totales (% en moles) de la glucoproteína heteróloga o anticuerpo, como se expresa en las células fúngicas filamentosas de la divulgación.

En otras formas de realización, el N-glicano $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (por ejemplo, N-glicano G_1) representa al menos el 10%, al menos el 20%, al menos el al menos el 30%, al menos el 40%, al menos el 50%, al menos el 60%,
15 al menos el 70%, al menos el 80%, al menos el 90% o más de los N-glicanos neutros totales (% en moles) de la glucoproteína heteróloga o anticuerpo, como se expresa en las células fúngicas filamentosas de la divulgación.

En otras formas de realización, el N-glicano $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (por ejemplo, N-glicano G_2) representa al menos el 10%, al menos el 20%, al menos el al menos el 30%, al menos el 40%, al menos el 50%, al menos el 60%,
20 al menos el 70%, al menos el 80%, al menos el 90% o más de los N-glicanos neutros totales (% en moles) de la glucoproteína heteróloga o anticuerpo, como se expresa en las células fúngicas filamentosas de la divulgación.

En otras formas de realización, el N-glicano de tipo complejo representa al menos el 10%, al menos el 20%, al menos el al menos el 30%, al menos el 40%, al menos el 50%, al menos el 60%, al menos el 70%, al menos el 80%,
25 al menos el 90% o más de los N-glicanos neutros totales (% en moles) de una glucoproteína heteróloga o anticuerpo, como se expresa en las células fúngicas filamentosas de la divulgación.

En otras formas de realización, el N-glicano de tipo híbrido representa al menos el 10%, al menos el 20%, al menos el al menos el 30%, al menos el 40%, al menos el 50%, al menos el 60%, al menos el 70%, al menos el 80%,
30 al menos el 90% o más de los N-glicanos neutros totales (% en moles) de una glucoproteína heteróloga o anticuerpo, como se expresa en las células fúngicas filamentosas de la divulgación.

En otras formas de realización, menos del 0,5%, 0,1%, 0,05%, o menos del 0,01% del N-glicano de la composición de glucoproteína heteróloga o la composición de anticuerpo producida por la célula huésped de la divulgación,
35 comprende galactosa. En ciertas formas de realización, ninguno de los N-glicanos comprende galactosa.

Las estructuras Neu5Gc y Gala- ($\text{Gal}\alpha_3\text{Gal}\beta_4\text{GlcNAc}$ terminal de extremo no reductor) son modificaciones xenoantigénicas (derivadas de animales) conocidas de anticuerpos que se producen en células animales, tales como células CHO. Las estructuras pueden ser antigénicas y, por lo tanto, dañinas incluso a bajas concentraciones.
40 Los hongos filamentosos de la presente divulgación carecen de rutas biosintéticas para producir las estructuras terminales Neu5Gc y Gala. En una forma de realización que puede combinarse con las formas de realización anteriores, menos del 0,1%, 0,01%, 0,001% o el 0% de los N-glicanos y/o O-glicanos de la composición de glucoproteína o anticuerpo comprende la estructura Neu5Gc y/o Gala. En una forma de realización que puede combinarse con las formas de realización anteriores, menos del 0,1%, 0,01%, 0,001% o el 0% de los N-glicanos y/o
45 O-glicanos de la composición de glucoproteína heteróloga o anticuerpo comprende la estructura Neu5Gc y/o Gala.

Las células fúngicas filamentosas de la presente divulgación carecen de genes para producir proteínas heterólogas fucosiladas. En una forma de realización que puede combinarse con las formas de realización anteriores, menos del 0,1%, 0,01%, 0,001% o el 0% del N-glicano de la composición de glucoproteína o anticuerpo comprende estructuras
50 de fucosa central.

La estructura $\text{Gal}\beta_4\text{GlcNAc}$ terminal de N-glicano de glicanos producidos por las células de mamífero afecta a la bioactividad de los anticuerpos y la $\text{Gal}\beta_3\text{GlcNAc}$ puede ser una estructura xenoantigénica de las proteínas producidas por células vegetales. En una forma de realización que puede combinarse con una o más de las formas
55 de realización anteriores, menos del 0,1%, 0,01%, 0,001% o el 0% de N-glicano de la composición de glucoproteína heteróloga o de anticuerpo comprende epítomos de galactosa terminales $\text{Gal}\beta_3/4\text{GlcNAc}$.

La glicación es una modificación postraduccional común de proteínas, que es resultado de la reacción química entre los azúcares reductores, tal como glucosa, y los grupos amino primarios en la proteína. La glicación se produce

normalmente en pH neutro o ligeramente alcalino en condiciones de cultivo celular, por ejemplo, cuando se producen anticuerpos en células CHO y se analizan Zhang et al. (2008) Unveiling a glycation hot spot in a recombinant humanized monoclonal antibody. Anal Chem. 80(7):2379-2390). Dado que los hongos filamentosos de la presente divulgación se cultivan típicamente en pH ácido, se reduce la aparición de glicación. En una forma de realización que puede combinarse con las formas de realización anteriores, menos del 1,0%, 0,5%, 0,1%, 0,01%, 0,001%, o el 0% de la composición de glucoproteína heteróloga o de anticuerpo comprende estructuras de glicación.

En una forma de realización, la composición de glucoproteína, tal como un anticuerpo carece de una, dos, tres, cuatro, cinco o seis de las estructuras seleccionadas del grupo de Neu5Gc, Gal α 3Gal β 4GlcNAc terminal, Gal β 4GlcNAc terminal, Gal β 3GlcNAc terminal, fucosa ligada al núcleo y estructuras de glicación.

En ciertas formas de realización, dicha proteína de glucoproteína con N-glicano de tipo mamífero, como se produce en la célula fúngica filamentosa de la divulgación, es una proteína terapéutica. Las proteínas terapéuticas pueden incluir inmunoglobulina, o una fusión de proteínas que comprende un fragmento Fc u otras glucoproteínas terapéuticas, tales como anticuerpos, eritropoyetinas, interferones, hormonas del crecimiento, albúminas o albúmina sérica, enzimas, o factores de coagulación de la sangre y pueden ser útiles en el tratamiento de seres humanos o animales. Por ejemplo, las glucoproteínas con N-glicano de tipo mamífero producidas por la célula fúngica filamentosa de acuerdo con la divulgación pueden ser una glucoproteína terapéutica tal como rituximab.

Los procedimientos para producir glucoproteínas con N-glicanos de tipo mamífero en células fúngicas filamentosas también se describen, por ejemplo, en el documento WO2012/069593.

En un aspecto, la célula fúngica filamentosa de acuerdo con la divulgación que se ha descrito anteriormente, se modifica genéticamente además para imitar la ruta tradicional de las células de mamíferos, a partir de N-glicanos Man5 como sustrato aceptor para GnTI, y va seguido secuencialmente de etapas de reacción de GnT1, manosidasa II y GnTII (en lo sucesivo aquí denominada la "ruta tradicional" para producir glucoformas GO). En una variante, se usa una sola enzima recombinante que comprende los dominios catalíticos de GnTI y GnTII.

Como alternativa, en un segundo aspecto, la célula fúngica filamentosa de acuerdo con la divulgación como se ha descrito anteriormente se modifica genéticamente adicionalmente para tener una expresión reducida en alg3, permitiendo la producción de N-glicanos Man₃GlcNAc₂ centrales, como sustrato aceptor para reacciones posteriores de GnTI y GnTII y evitando la necesidad de las enzimas manosidasa α 1,2 o manosidasa II (la ruta "alg3" reducida). En una variante, se usa una sola enzima recombinante que comprende los dominios catalíticos de GnTI y GnTII.

En dichas formas de realización para imitar la ruta tradicional para producir glucoproteínas con N-glicanos de tipo mamífero, una célula fúngica filamentosa que expresa Man₅, tal como la cepa de *T. reesei*, puede transformarse con una enzima de fusión GnTI o GnTII/GnTI usando integración aleatoria o por integración dirigida a un sitio conocido que no afecte a la glucosilación de Man5. Se seleccionan las cepas que sintetizan N-glicano GlcNAcMan5 para la producción de proteínas que tienen uno o más glicanos de tipo híbrido. Las cepas seleccionadas se transforman adicionalmente con un dominio catalítico de una manosidasa de tipo manosidasa II capaz de escindir las estructuras Man5 para generar GlcNAcMan3 para la producción de proteínas que tienen la correspondiente glucoforma GlcNAcMan3 o su derivado o derivados. En ciertas formas de realización, las enzimas de tipo manosidasa II pertenecen a la familia de glucósido hidrolasa 38 (cazy.org/GH38_all.html). Las enzimas caracterizadas incluyen las enzimas enumeradas en cazy.org/GH38_characterized.html. Las enzimas especialmente útiles son las enzimas de tipo Golgi que escinden las glucoproteínas, tales como las de la subfamilia α -manosidasa II (Man2A1; ManA2). Los ejemplos de dichas enzimas incluyen la enzima humana AAC50302, una enzima de *D. melanogaster* (Van den Elsen J.M. et al (2001) EMBO J. 20: 3008-3017), aquellas con la estructura 3D de acuerdo con la referencia 1BTY de PDB, y otras en referencia al dominio catalítico en PDB. Para la expresión citoplásmica, el dominio catalítico de la manosidasa se fusiona típicamente con un péptido de direccionamiento N-terminal (por ejemplo, como se divulga en la Sección anterior) o se expresa con estructuras de direccionamiento de Golgi endógenas animales o vegetales de enzimas manosidasa II de animales o plantas. Después de la transformación con el dominio catalítico de una manosidasa de tipo manosidasa II, se seleccionan cepas que producen GlcNAcMan3 (si se expresa GnTI), o se seleccionan cepas que producen eficazmente GlcNAc2Man3 (si se expresa una fusión de GnTI y GnTII). Para las cepas que producen GlcNAcMan3, dichas cepas se transforman además con un polinucleótido que codifica un dominio catalítico de GnTII y se seleccionan cepas transformantes que son capaces de producir GlcNAc2Man3GlcNAc2.

En dicha forma de realización para imitar la ruta tradicional, la célula fúngica filamentosa es una célula fúngica filamentosa como se define en secciones anteriores, y comprende además uno o más polinucleótidos que codifican

un polipéptido seleccionado del grupo que consiste en:

- i) α 1,2 manosidasa,
 - ii) dominio catalítico de N-acetilglucosaminiltransferasa I,
 - 5 iii) α manosidasa II,
 - iv) dominio catalítico de N-acetilglucosaminiltransferasa II,
 - v) β 1,4 galactosiltransferasa, y,
 - vi) fucosiltransferasa.
- 10 En formas de realización que utilizan la ruta *alg3* reducida, la célula fúngica filamentosa, tal como una célula de *Trichoderma*, tiene un nivel reducido de actividad de una dolcil-P-Man:Man(5)GlcNAc(2)-PP-dolcil manosiltransferasa en comparación con el nivel de actividad en una célula huésped parental. Dolcil-P-Man:Man(5)GlcNAc(2)-PP-dolcil manosiltransferasa (EC 2.4.1.130) transfiere un residuo de alfa-D-manosil de dolcil-fosfato D-manose en un oligosacárido unido a lípidos de membrana. Típicamente, la enzima dolcil-P-
- 15 Man:Man(5)GlcNAc(2)-PP-dolcil manosiltransferasa se codifica por un gen *alg3*. En ciertas formas de realización, la célula fúngica filamentosa para producir glucoproteínas con N-glicanos de tipo mamífero tiene un nivel reducido de expresión de un gen *alg3* en comparación con el nivel de expresión en una cepa parental.

Más preferiblemente, la célula fúngica filamentosa comprende una mutación de *alg3*. El gen *ALG3* puede mutarse por cualquier medio conocido en la técnica, tal como mutaciones puntuales o la delección de todo el gen *alg3*. Por ejemplo, la función de la proteína *alg3* se reduce o se elimina mediante la mutación de *alg3*. En ciertas formas de realización, el gen *alg3* se altera o se elimina de la célula fúngica filamentosa, tal como la célula de *Trichoderma*. En ciertas formas de realización, la célula fúngica filamentosa es una célula de *T. reesei*. SEQ ID NOs: 36 y 37 proporcionan las secuencias de ácidos nucleicos y aminoácidos del gen *alg3* en *T. reesei*, respectivamente. En una

20 forma de realización, la célula fúngica filamentosa se usa para la producción de una glucoproteína, en la que el uno o más glicanos comprenden o consisten en Man α 3[Man α 6]Man β 4GlcNAc β 4GlcNAc, y/o una variante alargada del extremo no reductor del mismo.

En ciertas formas de realización, la célula fúngica filamentosa tiene un nivel reducido de actividad de una alfa-1,6-manosiltransferasa en comparación con el nivel de actividad en una cepa parental. La alfa-1,6-manosiltransferasa (EC 2.4.1.232) transfiere un residuo de alfa-D-manosilo de GDP-manosa a un oligosacárido unido a proteínas, formando una unión alfa-(1- \rightarrow 6)-D-manosil-D-manosa de inicio del alargamiento en el aparato de Golgi. Típicamente, la enzima alfa-1,6-manosiltransferasa está codificada por un gen *och1*. En ciertas formas de realización, la célula fúngica filamentosa tiene un nivel reducido de expresión de un gen *och1* en comparación con el nivel de expresión en una célula fúngica filamentosa parental. En ciertas formas de realización, el gen *och1* se elimina de la célula fúngica filamentosa.

30

35

Las células fúngicas filamentosas utilizadas en los métodos para producir glucoproteína con N-glicanos de tipo mamífero pueden contener además un polinucleótido que codifica un dominio catalítico de N-acetilglucosaminiltransferasa I (GnTI) que cataliza la transferencia de N-acetilglucosamina a un Man α 3 terminal y un polinucleótido que codifica un dominio catalítico de N-acetilglucosaminiltransferasa II (GnTII), que cataliza la N-acetilglucosamina a un residuo de Man α 6 terminal de un glicano aceptor para producir un N-glicano complejo. En una forma de realización, dichos polinucleótidos que codifican GnTI y GnTII están unidos para producir una sola fusión de proteína que comprende ambos dominios catalíticos de GnTI y GnTII.

40

Como se divulga en el presente documento, N-acetilglucosaminiltransferasa I (GlcNAc-TI; GnTI; EC 2.4.1.101) cataliza la reacción UDP-N-acetil-D-glucosamina + 3-(alfa-D-manosil)-beta-D-manosil-R \rightleftharpoons UDP + 3-(2-(N-acetil-beta-D-glucosaminil)-alfa-D-manosil)-beta-D-manosil-R, donde R representa el resto del oligosacárido unido a N en el receptor de glicano. Un dominio catalítico de N-acetilglucosaminiltransferasa I es cualquier porción de una enzima N-acetilglucosaminiltransferasa I que es capaz de catalizar esta reacción. Las enzimas GnTI se enumeran en la base de datos CAZy en la familia de glucosiltransferasa 13 (cazy.org/GT13_all). La especie caracterizada enzimáticamente incluye *A. thaliana* AAR78757.1 (documento US6 653 459), *C. elegans* AAD03023.1 (Chen S. et al J. Biol.Chem 1999;274(1):288-97), *D. melanogaster* AAF57454.1 (Sarkar & Schachter Biol Chem. 2001 Feb;382(2):209-17); *C. griseus* AAC52872.1 (Puthalakath H. et al J. Biol.Chem 1996 271(44):27818-22); *H. sapiens* AAA52563.1 (Kumar R. et al Proc Natl Acad Sci U S A. 1990 Dec;87(24):9948-52); *M. auratus* AAD04130.1 (Opat As et al Biochem J. 1998 Dec 15;336 (Pt 3):593-8), (incluyendo un ejemplo de mutante de desactivación), Conejo, *O. cuniculus* AAA31493.1 (Sarkar M et al. Proc Natl Acad Sci U S A. 1991 Jan 1;88(1):234-8). Las secuencias de aminoácidos para las enzimas N-acetilglucosaminiltransferasa I de diversos organismos se describen, por ejemplo, en el documento PCT/EP2011/070956. Se pueden encontrar ejemplos adicionales de enzimas activas

45

50

55

caracterizadas en cazy.org/GT13_characterized. La estructura 3D del dominio catalítico de GnTI de conejo se definió mediante cristalografía de rayos X en Unligil UM et al. EMBO J. 2000 Oct 16;19(20):5269-80. Las estructuras del Protein Data Bank (PDB) para GnTI son 1FO8, 1FO9, 1FOA, 2AM3, 2AM4, 2AM5, y 2APC. En ciertas formas de realización, el dominio catalítico de N-acetilglucosaminiltransferasa I es de la enzima N-acetilglucosaminiltransferasa I humana (SEQ ID NO: 38) o variantes de la misma. En ciertas formas de realización, el dominio catalítico de N-acetilglucosaminiltransferasa I contiene una secuencia al menos un 70%, al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99%, o un 100% idéntica a los residuos de aminoácidos 84-445 de SEQ ID NO: 38. En algunas formas de realización, puede usarse una secuencia más corta como dominio catalítico (por ejemplo, residuos de aminoácidos 105-445 de la enzima humana o residuos de aminoácidos 107-447 de la enzima de conejo; Sarkar et al. (1998) Glycoconjugate J 15:193-197). Las secuencias adicionales que se pueden usar como dominio catalítico de GnTI incluyen residuos de aminoácidos de aproximadamente el aminoácido 30 a 445 de la enzima humana o cualquier dominio del tallo C-terminal que comienza entre el residuo de aminoácido 30 a 105 y continúa hasta aproximadamente el aminoácido 445 de la enzima humana, o la secuencia homóloga correspondiente de otra GnTI o una variante catalíticamente activa, o mutante de la misma. El dominio catalítico puede incluir partes N-terminales de la enzima, tal como la totalidad o parte del dominio del tallo, el dominio transmembrana o el dominio citoplásmico.

Como se divulga en el presente documento, N-acetilglucosaminiltransferasa II (GlcNAc-TII; GnTII; EC 2.4.1.143) cataliza la reacción $\text{UDP-N-acetil-D-glucosamina} + 6\text{-(alfa-D-manosil)-beta-D-manosil-R} \rightleftharpoons \text{UDP} + 6\text{-(2-(N-acetil-beta-D-glucosaminil)-alfa-D-manosil)-beta-D-manosil-R}$, donde R representa el resto del oligosacárido unido a N en el receptor de glicano. Un dominio catalítico de N-acetilglucosaminiltransferasa II es cualquier porción de una enzima N-acetilglucosaminiltransferasa II que es capaz de catalizar esta reacción. Las secuencias de aminoácidos para enzimas N-acetilglucosaminiltransferasa II de diversos organismos se enumeran en el documento WO2012069593. En ciertas formas de realización, el dominio catalítico de N-acetilglucosaminiltransferasa II es de la enzima N-acetilglucosaminiltransferasa II humana (SEQ ID NO: 39) o variantes de la misma. Se enumeran especies GnTII adicionales en la base de datos CAZY en la familia de glucosiltransferasa 16 (cazy.org/GT16_all). Las especies caracterizadas enzimáticamente incluyen GnTII de *C. elegans*, *D. melanogaster*, *Homo sapiens* (NP_002399.1), *Rattus norvegicus*, *Sus scrofa* (cazy.org/GT16_characterized). En ciertas formas de realización, el dominio catalítico de N-acetilglucosaminiltransferasa II contiene una secuencia que es al menos un 70%, al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99%, o un 100% idéntica a los residuos de aminoácidos de aproximadamente 30 a aproximadamente 447 de SEQ ID NO: 39. El dominio catalítico puede incluir partes N-terminales de la enzima, tal como la totalidad o parte del dominio del tallo, el dominio transmembrana o el dominio citoplásmico.

En las formas de realización en las que la célula fúngica filamentosa contiene una proteína de fusión de la divulgación, la proteína de fusión puede contener además un espaciador entre el dominio catalítico de N-acetilglucosaminiltransferasa I y el dominio catalítico de N-acetilglucosaminiltransferasa II. En ciertas formas de realización, el espaciador es un espaciador EGIV, un espaciador 2xG4S, un espaciador 3xG4S, o un espaciador CBHI. En otras formas de realización, el espaciador contiene una secuencia de un dominio del tallo.

Para la expresión en el RE/Golgi, el dominio catalítico de N-acetilglucosaminiltransferasa I y/o de N-acetilglucosaminiltransferasa II se fusiona típicamente con un péptido de direccionamiento o una parte de una proteína del RE o de Golgi temprana, o se expresa con estructuras de direccionamiento de RE endógenas de una enzima N-acetilglucosaminiltransferasa animal o vegetal. En ciertas formas de realización preferidas, el dominio catalítico de N-acetilglucosaminiltransferasa I y/o N-acetilglucosaminiltransferasa II contiene cualquiera de los péptidos de direccionamiento de la divulgación como se describe en la sección titulada "Secuencias de direccionamiento". Preferiblemente, el péptido de direccionamiento está unido al extremo N-terminal del dominio catalítico. En algunas formas de realización, el péptido de direccionamiento contiene cualquiera de los dominios del tallo de la divulgación como se describe en la sección titulada "Secuencias de direccionamiento". En ciertas formas de realización preferidas, el péptido de direccionamiento es un péptido de direccionamiento Kre2/Mnt1. En otras formas de realización, el péptido de direccionamiento contiene además un dominio transmembrana unido al extremo N-terminal del dominio del tallo o un dominio citoplásmico unido al extremo N-terminal del dominio del tallo. En formas de realización en las que el péptido de direccionamiento contiene además un dominio transmembrana, el péptido de direccionamiento puede contener además un dominio citoplásmico unido al extremo N-terminal del dominio transmembrana.

Las células fúngicas filamentosas también pueden contener un polinucleótido que codifica un transportador UDP-GlcNAc. El polinucleótido que codifica el transportador UDP-GlcNAc puede ser endógeno (es decir, naturalmente presente) en la célula huésped, o puede ser heterólogo a la célula fúngica filamentosa.

En ciertas formas de realización, la célula fúngica filamentosa puede contener además un polinucleótido que codifica una α -1,2-manosidasa. El polinucleótido que codifica la α -1,2-manosidasa puede ser endógeno en la célula huésped, o puede ser heterólogo a la célula huésped. Los polinucleótidos heterólogos son especialmente útiles para una célula huésped que expresa glicanos ricos en manosa transferidos desde el Golgi al RE sin una escisión eficaz de la exo- α -2-manosidasa. La α -1,2-manosidasa puede ser una enzima del tipo manosidasa I que pertenece a la familia de glucósido hidrolasa 47 (cazy.org/GH47_all.html). En ciertas formas de realización, la α -1,2-manosidasa es una enzima enumerada en cazy.org/GH47_characterized.html. En particular, la α -1,2-manosidasa puede ser una enzima de tipo RE que escinde las glucoproteínas, tales como las enzimas en la subfamilia de las enzimas α -manosidasa I de RE EC 3.2.1.113. Los ejemplos de tales enzimas incluyen α -2-manosidasa 1B humana (AAC26169), una combinación de manosidasas de RE de mamífero, o una enzima fúngica filamentosa tal como α -1,2-manosidasa (MDS1) (*T. reesei* AAF34579; Maras M et al J Biotech. 77, 2000, 255, o Trire 45717). Para la expresión en el RE, el dominio catalítico de la manosidasa se fusiona típicamente con un péptido de direccionamiento, tal como HDEL, KDEL, o parte de una proteína de RE o de Golgi temprana, o se expresa con estructuras de direccionamiento al RE endógenas de una enzima manosidasa I animal o vegetal.

En ciertas formas de realización, la célula fúngica filamentosa también puede contener además un polinucleótido que codifica una galactosiltransferasa. Las galactosiltransferasas transfieren residuos galactosilo unidos a β a un residuo N-acetilglucosaminilo terminal. En ciertas formas de realización, la galactosiltransferasa es una β -1,4-galactosiltransferasa. Generalmente, las β -1,4-galactosiltransferasas pertenecen a la familia de glucosiltransferasa 7 de CAZy (cazy.org/GT7_all.html) e incluyen β -N-acetilglucosaminil-gluco péptido β -1,4-galactosiltransferasa (EC 2.4.1.38), que también se conoce como N-acetilactosamina sintasa (EC 2.4.1.90). Las subfamilias útiles incluyen β 4-GalT1, β 4-GalT-II, -III, -IV, -V, y -VI, tal como β 4-GalT1 o β 4GalT-II, -III, -IV, -V, y -VI de mamífero o ser humano, o cualquier combinación de las mismas. β 4-GalT1, β 4-GalTII, o β 4-GalTIII son especialmente útiles para la galactosilación de las estructuras de GlcNAc β 2 terminal en N-glicanos, tales como GlcNAcMan3, GlcNAc2Man3, o GlcNAcMan5 (Guo S. et al. Glycobiology 2001, 11:813-20). La estructura tridimensional de la región catalítica es conocida (por ejemplo, (2006) J.Mol.Biol. 357: 1619-1633), y la estructura se ha representado en la base de datos del PDB con el código 2FYD. La base de datos CAZy incluye ejemplos de ciertas enzimas. Las enzimas caracterizadas también se enumeran en la base de datos CAZy en cazy.org/GT7_characterized.html. Los ejemplos de enzimas β 4GalT útiles incluyen β 4GalT1, por ejemplo, la enzima bovina de *Bos taurus* AAA30534.1 (Shaper N.L. et al Proc. Natl. Acad. Sci. U.S.A. 83 (6), 1573-1577 (1986)), una enzima humana (Guo S. et al. Glycobiology 2001, 11:813-20), y una enzima de *Mus musculus* AAA37297 (Shaper, N.L. et al. 1998 J. Biol. Chem. 263 (21), 10420-10428); enzimas β 4GalTII tales como β 4GalTII humana BAA75819.1, *Cricetulus griseus* de hámster chino AAM77195, enzima de *Mus musculus* BAA34385, y pez Medaka japonés *Oryzias latipes* BAH36754; y enzimas β 4GalTIII tales como β 4GalTIII humana BAA75820.1, *Cricetulus griseus* de hámster chino AAM77196 y enzimas de *Mus musculus* AAF22221.

La galactosiltransferasa puede expresarse en la membrana plasmática de la célula huésped. Puede usarse un péptido de orientación heterólogo, tal como un péptido Kre2 descrito en Schwientek J.Biol. Chem 1996 3398. Los promotores que se pueden usar para la expresión de la galactosiltransferasa incluyen promotores constitutivos tales como *gpd*, promotores de enzimas de glucosilación endógena y glucosiltransferasas tales como manosiltransferasas que sintetizan N-glicanos en el Golgi o el RE, y promotores inducibles de proteínas endógenas de alto rendimiento, tal como el promotor *cbh1*.

En ciertas formas de realización de la divulgación, donde la célula fúngica filamentosa contiene un polinucleótido que codifica una galactosiltransferasa, la célula fúngica filamentosa también contiene un polinucleótido que codifica una epimerasa UDP-Gal 4 y/o un transportador UDP-Gal. En ciertas formas de realización de la divulgación en las que la célula fúngica filamentosa contiene un polinucleótido que codifica una galactosiltransferasa, puede usarse lactosa como fuente de carbono en lugar de glucosa cuando se cultiva la célula huésped. El medio de cultivo puede estar entre pH 4,5 y 7,0 o entre 5,0 y 6,5. En ciertas formas de realización de la divulgación donde la célula fúngica filamentosa contiene un polinucleótido que codifica una galactosiltransferasa y un polinucleótido que codifica una epimerasa UDP-Gal 4 y/o un transportador UDP-Gal, se puede añadir un catión divalente tal como Mn²⁺, Ca²⁺ o Mg²⁺ al medio de cultivo celular.

Por consiguiente, en ciertas formas de realización, la célula fúngica filamentosa de la divulgación, por ejemplo, seleccionada de células de *Neurospora*, *Trichoderma*, *Myceliophthora*, *Aspergillus*, *Fusarium* o *Chrysosporium*, y más preferiblemente una célula de *Trichoderma reesei*, puede comprender las siguientes características:

- a) una mutación en al menos una proteasa endógena que reduce o elimina la actividad de dicha proteasa

endógena, preferiblemente la actividad de proteasa de dos o tres o más proteasas endógenas se reduce, por ejemplo, proteasas pep1, tsp1, gap1 y/o slp1, con el fin de mejorar la producción o estabilidad de una glucoproteína heteróloga a producir,

b) un polinucleótido que codifica una subunidad catalítica heteróloga de oligosacáril transferasa, preferiblemente de SEQ ID NO:2 o NO:9,

c) un polinucleótido que codifica una glucoproteína que tiene al menos una asparagina, preferiblemente una glucoproteína heteróloga, tal como una inmunoglobulina, un anticuerpo o una fusión de proteínas que comprende el fragmento Fc de una inmunoglobulina.

d) opcionalmente, una delección o alteración del gen alg3,

e) opcionalmente, un polinucleótido que codifica el dominio catalítico de la N-acetilglucosaminiltransferasa I y un polinucleótido que codifica el dominio catalítico de la N-acetilglucosaminiltransferasa II,

f) opcionalmente, un polinucleótido que codifica β 1,4 galactosiltransferasa,

g) opcionalmente, un polinucleótido o polinucleótidos que codifican la epimerasa y/o el transportador UDP-Gal 4.

Secuencias de direccionamiento

En ciertas formas de realización, las enzimas recombinantes, tales como α 1,2 manosidasas, GnTI, u otras glucosiltransferasas introducidas en las células fúngicas filamentosas, incluyen un péptido de direccionamiento unido a los dominios catalíticos. El término "unido" como se usa en el presente documento significa que dos polímeros de residuos de aminoácidos en el caso de un polipéptido, o dos polímeros de nucleótidos en el caso de un polinucleótido están acoplados directamente adyacentes entre sí o están dentro del mismo polipéptido o polinucleótido, pero están separados por residuos de aminoácidos o nucleótidos intermedios. Un "péptido de direccionamiento", como se usa en el presente documento, se refiere a cualquier número de residuos de aminoácidos consecutivos de la proteína recombinante que son capaces de localizar la proteína recombinante en el retículo endoplásmico (RE) o el aparato de Golgi (Golgi) dentro de la célula huésped. El péptido de direccionamiento puede ser N-terminal o C-terminal a los dominios catalíticos. En ciertas formas de realización, el péptido de direccionamiento es N-terminal a los dominios catalíticos. En ciertas formas de realización, el péptido de direccionamiento proporciona unión a un componente del RE o de Golgi, tal como a una enzima manosidasa II. En otras formas de realización, el péptido de direccionamiento proporciona una unión directa a la membrana de RE o Golgi.

Los componentes del péptido de direccionamiento pueden provenir de cualquier enzima que normalmente reside en el RE o el aparato de Golgi. Dichas enzimas incluyen manosidasas, manosiltransferasas, glucosiltransferasas, proteínas de Golgi tipo 2, y enzimas MNN2, MNN4, MNN6, MNN9, MNN10, MNS1, KRE2, VAN1, y OCH1. Dichas enzimas pueden provenir de una levadura o de una especie de hongo tales como las de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filobasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, y *Trichoderma*. Las secuencias para dichas enzimas se pueden encontrar en la base de datos de secuencias del GenBank.

En ciertas formas de realización, el péptido de direccionamiento proviene de la misma enzima y organismo que uno de los dominios catalíticos de la proteína recombinante. Por ejemplo, si la proteína recombinante incluye un dominio catalítico de GnTII humano, el péptido de direccionamiento de la proteína recombinante es de la enzima GnTII humana. En otras formas de realización, el péptido de direccionamiento puede provenir de una enzima y/u organismo diferente como los dominios catalíticos de la proteína recombinante.

Los ejemplos de diversos péptidos de direccionamiento para su uso en el direccionamiento de proteínas al RE o Golgi que pueden usarse para dirigir las enzimas recombinantes incluyen: Péptido N-terminal Kre2/Mnt1 fusionado a galactosiltransferasa (Schwientek, JBC 1996, 3398), HDEL para la localización de manosidasa en el RE de células de levadura para producir Man5 (Chiba, JBC 1998, 26298-304; Callewaert, FEBS Lett 2001, 173-178), péptido dirigido a OCH1 fusionado al dominio catalítico de GnTI (Yoshida et al, Glycobiology 1999, 53-8), péptido de levadura N-terminal de Mns1 fusionado a α 2-manosidasa (Martinet et al, Biotech Lett 1998, 1171), porción N-terminal de Kre2 unido a dominio catalítico de GnTI o β 4GalT (Vervecken, Appl. Environ Microb 2004, 2639-46), diversos enfoques revisados en Wildt and Gerngross (Nature Rev Biotech 2005, 119), GnTI de longitud completa en *Aspergillus nidulans* (Kalsner et al, Glycocon. J 1995, 360-370), GnTI de longitud completa en *Aspergillus oryzae* (Kasajima et al, Biosci Biotech Biochem 2006, 2662-8), porción de estructura de localización Sec12 de levadura fusionada a GnTI de *C. elegans* en *Aspergillus* (Kainz et al 2008), porción N-terminal de Mnn9 de levadura fusionada a GnTI humano en *Aspergillus* (Kainz et al 2008), porción N-terminal de Mnn10 de *Aspergillus* fusionada a GnTI

humano (Kainz et al, Appl. Environ Microb 2008, 1076-86), y GnTI de longitud completa humano en *T. reesei* (Maras et al, FEBS Lett 1999, 365-70).

En ciertas formas de realización, el péptido de direccionamiento es una porción N-terminal del péptido de direccionamiento Mnt1/Kre2 que tiene la secuencia de aminoácidos de SEQ ID NO: 40 (por ejemplo, codificada por el nucleótido de la SEQ ID NO:41). En ciertas formas de realización, el péptido de direccionamiento se selecciona de GNT2, KRE2, de tipo KRE2, Och1, Anp1, Van1 humanos como se muestra en la Tabla 1 a continuación:

Tabla 1: Secuencia de aminoácidos de péptidos de direccionamiento

Proteína	TreID	Secuencia de aminoácidos
GNT2 humano	-	MRFRIYKRKVLILTLVVAACGFVLWSSNGRQR
		KNEALAPPLLD AEPARGAGGRGGDHP (SEQ ID NO:42)
KRE2	21576	MASTNARYVRYLLIAFFTILVFYFVSNKYEGV
		DLNKGFTTAPDSTKTPK (SEQ ID NO:43)
de tipo KRE2	69211	MAIARPVRALGGLAAILWCFFLYQLLRPSSSY
		NSPGDRYINFERDPNLDPTG (SEQ ID NO:44)
Och1	65646	MLNPRRALIAAAFILTVFFLISRSHNSEASTS (SEQ ID NO:45)
Anp1	82551	MMPRHSSGFSNGYPRADTFEISPHRFQPR
		TLPPHRKRKRTAIRVGI VVVILVVLWFGQPR
		SVASLISLGILSGYDDLKLE (SEQ ID NO:46)
Van1	81211	MLLPKGGLDWR SARAQIPTRALWNAVTRTR
		FILLVGITGLILLW RGVSTSASE (SEQ ID NO:47)

Otros ejemplos de secuencias que pueden usarse para los péptidos de direccionamiento incluyen las secuencias de direccionamiento como se describe en el documento WO2012/069593.

5

Las secuencias no caracterizadas pueden ensayarse para su uso como péptidos de direccionamiento expresando enzimas de la ruta de glucosilación en una célula huésped, donde una de las enzimas contiene la secuencia no caracterizada como único péptido de direccionamiento, y midiendo los glicanos producidos en vista de la localización del citoplasma de la biosíntesis de glucanos (por ejemplo, como en Schwientek JBC 1996 3398), o expresando una proteína indicadora fluorescente fusionada con el péptido de direccionamiento, y analizando la localización de la proteína en el Golgi mediante inmunofluorescencia o fraccionando las membranas citoplásmicas del Golgi y midiendo la localización de la proteína.

15 **Métodos para producir una glucoproteína que tiene un aumento de la ocupación del sitio de N-glucosilación**

Las células fúngicas filamentosas como se han descrito anteriormente son útiles en métodos para producir una composición de glucoproteína con un aumento de la ocupación del sitio de N-glucosilación.

Por consiguiente, en otro aspecto, la divulgación se refiere a un método para producir una composición de glucoproteína con un aumento de la ocupación del sitio de N-glucosilación, que comprende

- 25 a) proporcionar una célula fúngica filamentosas, por ejemplo, una célula de *Trichoderma*, que tiene un gen STT3D de *Leishmania* que codifica una subunidad catalítica de oligosacaril transferasa, o una variante funcional de la misma, y un polinucleótido que codifica una glucoproteína heteróloga,
- b) cultivar la célula en condiciones apropiadas para la expresión del gen STT3D o su variante funcional, y la producción de la glucoproteína heteróloga; y,
- c) recuperar dicha composición de glucoproteína y, opcionalmente, purificar la composición de glucoproteína heteróloga.

En formas de realización específicas del método, la célula fúngica filamentosa comprende una o más mutaciones que reducen o eliminan una o más actividades de proteasa endógena en comparación con una célula fúngica filamentosa parental que no tiene dicha mutación o mutaciones, como se ha descrito anteriormente.

- 5 En los métodos de la divulgación, ciertos medios de crecimiento incluyen, por ejemplo, medios comunes preparados en el mercado tales como caldo Luria-Bertani (LB), caldo Sabouraud Dextrose (SD) o caldo de medio de levadura (YM). También se pueden usar otros medios de crecimiento definidos o sintéticos, y un experto en la técnica de microbiología o ciencia de la fermentación conocerá el medio apropiado para el crecimiento de la célula huésped particular. El medio de cultivo típicamente tiene el medio mínimo de *Trichoderma reesei* (Penttilä et al., 1987, Gene
10 61, 155-164) como base, complementado con sustancias que inducen el promotor de producción, tal como lactosa, celulosa, grano usado o soforosa. Los intervalos de temperatura y otras condiciones adecuadas para el crecimiento se conocen en la técnica (véase, por ejemplo, Bailey y Ollis 1986). En ciertas formas de realización, el pH del cultivo celular está entre 3,5 y 7,5, entre 4,0 y 7,0, entre 4,5 y 6,5, entre 5 y 5,5, o en 5.5. En ciertas formas de realización, para producir un anticuerpo, la célula fúngica filamentosa o la célula fúngica de *Trichoderma* se cultiva en un
15 intervalo de pH seleccionado de 4,7 a 6,5; pH 4,8 a 6,0; pH 4,9 a 5,9; y pH 5,0 a 5,8.

En algunas formas de realización de la divulgación, el método comprende cultivar en un medio que comprende uno o dos inhibidores de proteasa.

- 20 En una forma de realización específica de la divulgación, el método comprende cultivar en un medio que comprende uno o dos inhibidores de la proteasa seleccionados de SBTI y quimiostatina.

En algunas formas de realización, la glucoproteína es una glucoproteína heteróloga, preferiblemente una glucoproteína de mamífero. En otras formas de realización, la glucoproteína heteróloga es una glucoproteína de
25 origen no mamífero.

En ciertas formas de realización, una glucoproteína de mamífero se selecciona de una inmunoglobulina, una cadena pesada o ligera de inmunoglobulina o anticuerpo, un anticuerpo monoclonal, un fragmento Fab, un fragmento de anticuerpo F(ab')₂, un anticuerpo monocatenario, un anticuerpo de dominio único monomérico o multimérico, un
30 anticuerpo camélido, o sus fragmentos de unión a antígeno.

Un fragmento de una proteína, como se usa en el presente documento, consiste en al menos 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 aminoácidos consecutivos de una proteína de referencia.

- 35 Como se usa en el presente documento, una "inmunoglobulina" se refiere a una proteína multimérica que contiene una cadena pesada y una cadena ligera unidas de forma covalente y capaces de combinarse específicamente con el antígeno. Las moléculas de inmunoglobulina son una gran familia de moléculas que incluyen varios tipos de moléculas tales como IgM, IgD, IgG, IgA e IgE.

40 Como se usa en el presente documento, un "anticuerpo" se refiere a moléculas de inmunoglobulina intactas, así como a fragmentos de las mismas que son capaces de unirse a un antígeno. Estas incluyen moléculas de anticuerpos híbridos (quiméricos) (véanse, por ejemplo, Winter et al. Nature 349:293-99225, 1991; y la Pat. de Estados Unidos N.º 4.816.567 226); moléculas F(ab')₂; heterodímeros no covalentes; construcciones de fragmentos de anticuerpos diméricos y triméricos; moléculas de anticuerpos humanizados (véanse, por ejemplo, Riechmann et
45 al. Nature 332, 323-27, 1988; Verhoevan et al. Science 239, 1534-36, 1988; y el documento GB 2.276.169); y cualquier fragmento funcional obtenido a partir de tales moléculas, así como anticuerpos obtenidos a través de procesos no convencionales, tales como presentación de fagos o ratones transgénicos. Preferiblemente, los anticuerpos son anticuerpos clásicos con región Fc. Los métodos de fabricación de anticuerpos se conocen bien en la técnica.

50

En formas de realización adicionales, el rendimiento de la glucoproteína de mamífero, por ejemplo, el anticuerpo, es de al menos 0,5, al menos 1, al menos 2, al menos 3, al menos 4 o al menos 5 gramos por litro.

En ciertas formas de realización, la glucoproteína de mamífero es un anticuerpo, opcionalmente, IgG1, IgG2, IgG3 o
55 IgG4. En formas de realización adicionales, el rendimiento del anticuerpo es de al menos 0,5, al menos 1, al menos 2, al menos 3, al menos 4 o al menos 5 gramos por litro. En formas de realización adicionales, la glucoproteína de mamífero es un anticuerpo, y el anticuerpo contiene al menos el 70%, al menos el 80%, al menos el 90%, al menos el 95%, o al menos el 98% de un extremo C y extremo N de anticuerpo natural sin residuos de aminoácidos adicionales. En otras formas de realización, la glucoproteína de mamífero es un anticuerpo, y el anticuerpo contiene

al menos el 70%, al menos el 80%, al menos el 90%, al menos el 95%, o al menos el 98% de un extremo C y un extremo N de anticuerpo natural que no carece de ningún residuo de aminoácido C-terminal o N-terminal.

En ciertas formas de realización donde la glucoproteína de mamífero (por ejemplo, el anticuerpo) se purifica del cultivo celular, el cultivo que contiene la glucoproteína de mamífero contiene fragmentos de polipéptidos que constituyen un porcentaje en masa que es menor del 50%, menor del 40%, menor del 30%, menor del 20% o menor del 10% de la masa de los polipéptidos producidos. En ciertas formas de realización preferidas, la glucoproteína de mamífero es un anticuerpo, y los fragmentos de polipéptidos son fragmentos de cadena pesada y/o fragmentos de cadena ligera. En otras formas de realización, cuando la glucoproteína de mamífero es un anticuerpo y el anticuerpo se purifica a partir del cultivo celular, el cultivo que contiene el anticuerpo contiene cadenas pesadas libres y/o cadenas ligeras libres que constituyen un porcentaje en masa que es inferior al 50%, inferior al 40%, inferior al 30%, inferior al 20% o inferior al 10% de la masa del anticuerpo producido. Los métodos para determinar el porcentaje en masa de fragmentos de polipéptidos se conocen bien en la técnica e incluyen, medir la intensidad de la señal de un gel SDS.

En otras formas de realización, la glucoproteína heteróloga (por ejemplo, el anticuerpo) con un aumento de la ocupación del sitio de N-glicosilación, por ejemplo, el anticuerpo, comprende la estructura de N-glicano trimanosilo $\text{Man}\alpha 3[\text{Man}\alpha 6]\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$. En algunas formas de realización, la estructura $\text{Man}\alpha 3[\text{Man}\alpha 6]\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$ representa al menos el 20%, 30%; 40%, 50%; 60%, 70%, 80% (% en moles) o más, de los N-glicanos totales de la composición de glucoproteína heteróloga (por ejemplo, el anticuerpo) obtenida por los métodos de la divulgación. En otras formas de realización, la glucoproteína heteróloga (por ejemplo, el anticuerpo) comprende la estructura de N-glicano $\text{G0 GlcNAc}\beta 2\text{Man}\alpha 3[\text{GlcNAc}\beta 2\text{Man}\alpha 6]\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$. En otras formas de realización, la estructura de la glucoforma G0 no fucosilada representa al menos el 20%, 30%; 40%, 50%; 60%, 70%, 80% (% en moles) o más, de los N-glicanos totales de la composición de glucoproteína heteróloga (por ejemplo, el anticuerpo) obtenida por los métodos de la divulgación. En otras formas de realización, los N-glicanos galactosilados representan menos (% en moles) del 0,5%, 0,1%, 0,05%, 0,01% de los N-glicanos totales del cultivo, y/o de la glucoproteína heteróloga con un aumento de la ocupación del sitio de N-glicosilación. En ciertas formas de realización, el cultivo o la glucoproteína heteróloga, por ejemplo, un anticuerpo, no comprende N-glicanos galactosilados.

En ciertas formas de realización de cualquiera de los métodos descritos, el método incluye la etapa adicional de proporcionar uno o más, dos o más, tres o más, cuatro o más, o cinco o más inhibidores de la proteasa. En ciertas formas de realización, los inhibidores de la proteasa son péptidos que se expresan conjuntamente con la glucoproteína de mamífero. En otras formas de realización, los inhibidores inhiben al menos dos, al menos tres o al menos cuatro proteasas de una familia de proteasas seleccionada de proteasas aspárticas, serina proteasas de tipo tripsina, subtilisina proteasas y proteasas glutámicas.

En ciertas formas de realización de cualquiera de los métodos divulgados, la célula fúngica filamentosa o la célula fúngica de *Trichoderma* también contiene una proteína portadora. Como se usa en el presente documento, una "proteína portadora" es porción de una proteína que es endógena y altamente secretada por una célula fúngica filamentosa o una célula fúngica de *Trichoderma*. Las proteínas portadoras adecuadas incluyen, sin limitación, las de mananasa I de *T. reesei* (Man5A o MANI), celobiohidrolasa II de *T. reesei* (Cel6A o CBHII) (véase, por ejemplo, Paloheimo et al Appl. Environ. Microbiol. 2003 December; 69(12): 7073-7082) o celobiohidrolasa de *T. reesei* (CBHI). En algunas formas de realización, la proteína portadora es CBH1. En otras formas de realización, la proteína portadora es una proteína CBH1 de *T. reesei* truncada que incluye la región central de CBH1 y parte de la región enlazadora de CBH1. En algunas formas de realización, un vehículo tal como una celobiohidrolasa o su fragmento, se fusiona con una cadena ligera de anticuerpo y/o una cadena pesada de anticuerpo. En algunas formas de realización, un polipéptido de fusión vehículo-anticuerpo comprende un sitio de escisión Kex2. En ciertas formas de realización, Kex2, u otra enzima de escisión de vehículo, es endógeno a una célula fúngica filamentosa. En ciertas formas de realización, la proteasa de escisión de vehículo es heteróloga a la célula fúngica filamentosa, por ejemplo, otra proteína Kex2 derivada de levadura o una proteasa TEV. En ciertas formas de realización, la enzima de escisión de vehículo se sobreexpresa. En ciertas formas de realización, el vehículo consiste en aproximadamente 469 a 478 aminoácidos de la parte N-terminal de la proteína CBH1 de *T. reesei* GenBank N.º de acceso EGR44817.1.

En una forma de realización, el polinucleótido que codifica la glucoproteína heteróloga (por ejemplo, el anticuerpo) comprende además un polinucleótido que codifica el dominio catalítico de CBH1 y el enlazador como una proteína portadora y/o el promotor *cbh1*.

En ciertas formas de realización, la célula fúngica filamentosa de la divulgación sobreexpresa la proteasa KEX2. En

una forma de realización, la glucoproteína heteróloga (por ejemplo, el anticuerpo) se expresa como una construcción de fusión que comprende un polipéptido fúngico endógeno, un sitio de proteasa tal como un sitio de escisión Kex2, y la proteína heteróloga tal como una cadena pesada y/o ligera del anticuerpo. Se han descrito combinaciones útiles de 2-7 aminoácidos que preceden al sitio de escisión de Kex2, por ejemplo, en Mikosch et al. (1996) *J. Biotechnol.* 52:97-106; Goller et al. (1998) *Appl Environ Microbiol.* 64:3202-3208; Spencer et al. (1998) *Eur. J. Biochem.* 258:107-112; Jalving et al. (2000) *Appl. Environ. Microbiol.* 66:363-368; Ward et al. (2004) *Appl. Environ. Microbiol.* 70:2567-2576; Ahn et al. (2004) *Appl. Microbiol. Biotechnol.* 64:833-839; Paloheimo et al. (2007) *Appl Environ Microbiol.* 73:3215-3224; Paloheimo et al. (2003) *Appl Environ Microbiol.* 69:7073-7082; y Margolles-Clark et al. (1996) *Eur J Biochem.* 237:553-560.

10

La descripción se refiere además a la composición de glucoproteínas, por ejemplo, la composición de anticuerpo, obtenible u obtenida por el método como se ha divulgado anteriormente.

En otras formas de realización específicas, dicha composición de glucoproteína o de anticuerpo comprende además como el 50%, 60%, 70% o el 80% (% en moles de N-glicano neutro), de la siguiente glucoforma:

15

- (i) Man α 3[Man α 6(Man α 3)Man α 6]Man β 4GlcNA β 4GlcNAc (glucoforma Man5);
- (ii) GlcNAc β 2Man α 3[Man α 6(Man α 3)Man α 6]Man β 4GlcNA β 4GlcNAc, o variante β 4-galactosilada de la misma;
- (iii) Man α 6(Man α 3)Man β 4GlcNA β 4GlcNAc;
- (iv) Man α 6(GlcNAc β 2Man α 3)Man β 4GlcNA β 4GlcNAc, o variante β 4-galactosilada de la misma; o,
- (v) N-glicanos de tipo complejo seleccionados de la glucoforma G0, G1, o G2.

20

En algunas formas de realización, la glucoforma de N-glicano de acuerdo con iii-v comprende menos del 15%, 10%, 7%, 5%, 3%, 1% o 0,5%, o está desprovista de glicano Man5 como se ha definido en i) anteriormente.

25

EJEMPLOS

ENSAYOS FUNCIONALES

30

Ensayo para medir la actividad proteasa total de las células de la divulgación

Las concentraciones de proteína se determinaron a partir de muestras de sobrenadante del día 2-7 de 1x-7x cepas deficientes en proteasa (descritas en el documento PCT/EP2013/050126) de acuerdo con el kit de ensayo de proteasa EnzChek (Molecular probes N.º E6638, sustrato de caseína verde fluorescente). En resumen, los sobrenadantes se diluyeron en tampón de citrato de sodio para igualar la concentración de proteína total y se añadieron cantidades iguales de los sobrenadantes diluidos en una placa negra de 96 pocillos, utilizando 3 pocillos duplicados por muestra. La solución de caseína FL diluida hecha en tampón de citrato de sodio se añadió a cada sobrenadante que contenía un pocillo y las placas se incubaron cubiertas en una bolsa de plástico a 37°C. La fluorescencia de los pocillos se midió después de 2, 3 y 4 horas. Las lecturas se hicieron en el lector de placas fluorescente de Varioskan utilizando una excitación de 485 nm y una emisión de 530 nm. Algunas mediciones de la actividad de la proteasa se realizaron utilizando caseína succinilada (kit de ensayo de proteasa QuantiCleave, Pierce N.º 23263) de acuerdo con el protocolo del fabricante.

35

40

La delección individual de pep1 redujo la actividad de la proteasa 1,7 veces, la delección doble de pep1/tsp1 redujo la actividad de la proteasa 2 veces, la triple delección de pep1/tsp1/slp1 redujo la actividad de la proteasa 3,2 veces, la delección cuádruple de pep1/tsp1/slp1/gap1 redujo la actividad de la proteasa 7,8 veces en comparación con la cepa M124 de tipo silvestre, la delección quintuple de pep1/tsp1/slp1/gap1/gap2 redujo la actividad de la proteasa 10 veces, la delección séxtuple de pep1/tsp1/slp1/gap1/gap2/pep4 redujo la actividad de la proteasa 15,9 veces, y la delección séptuple de pep1/tsp1/slp1/gap1/gap2/pep4/pep3 redujo la actividad de la proteasa 18,2 veces.

50

La Figura 5 representa gráficamente datos de actividad de proteasa normalizados de sobrenadantes de cultivo de cada uno de los sobrenadantes de delección de proteasas (mutante de delección de 1 a 7 veces) y la cepa parental sin delecciones de proteasas. La actividad de la proteasa se midió a pH 5,5 en las primeras 5 cepas y a pH 4,5 en las últimas tres cepas de delección. La actividad de proteasa es contra la caseína verde fluorescente. La cepa de delección de proteasas séxtuple tiene solo el 6% de la cepa parental de tipo silvestre y la actividad de la proteasa de la cepa de delección de proteasas séptuple fue aproximadamente un 40% menor que la actividad de la cepa de delección de proteasas séxtuple.

55

60

Ensayo para medir la ocupación del sitio de N-glucosilación en una composición de glucoproteína

Se digieren 10 - 30 µg de anticuerpo con 13,4 - 30 U de FabRICATOR (Genovis), +37°C, 60 min durante una noche, produciendo un fragmento F(ab')₂ y un fragmento Fc por una molécula de anticuerpo. Las muestras digeridas se purifican utilizando la placa de filtro Poros R1 (Glyken corp.) y los fragmentos Fc se analizan para determinar la ocupación del sitio de N-glicano utilizando MALDI-TOF MS. El porcentaje de ocupación de un Fc en el sitio es el promedio de dos valores: el que se obtiene a partir de los valores de intensidad de los picos (carga simple y doble) y el otro del área de los picos (carga simple y doble); ambos valores se calculan como señal glucosilada dividida por la suma de señales no glucosiladas y glucosiladas.

10

EJEMPLO 1 - Generación de STT3 de *L. major* que expresa *T. reesei*

La secuencia codificante de oligosacaril transferasa 3D de *Leishmania major* (anterior GenBank N.º XP_843223.1, nuevo XP_003722509.1; SEQ ID NO: 1) se optimizó por codón para determinar la expresión de *Trichoderma reesei* (secuencia de ácidos nucleicos optimizada por codón SEQ ID NO: 2). La secuencia codificante optimizada se sintetizó junto con el promotor cDNA1 (SEQ ID NO: 3) y la secuencia flanqueante del terminador TrpC (SEQ ID NO: 4). El gen STT3 de *Leishmania major* se eliminó del vector de clonación optimizado utilizando la digestión con enzimas de restricción Pacl. El vector de entrada de expresión también se digirió con Pacl y se desfosforiló con fosfatasa alcalina de ternera. El gen STT3 y el vector digerido se separaron con electroforesis en gel de agarosa y los fragmentos correctos se aislaron del gel con un kit de extracción de gel (Qiagen) de acuerdo con el protocolo del fabricante. El gen STT3 de *Leishmania major* purificado se ligó en el vector de expresión con ADN ligasa T4. La reacción de ligación se transformó en *E. coli* DH5α químicamente competente y se hizo crecer en placas de selección de ampicilina (100 µg/ml). Las preparaciones de plásmidos Miniprep se hicieron a partir de varias colonias. La presencia del inserto del gen STT3 de *Leishmania major* se comprobó mediante la digestión de los plásmidos preparados con digestión con Pacl y se secuenciaron varios clones positivos para verificar la orientación del gen. Se eligió un clon correctamente orientado como el vector final pTTv201.

El casete de expresión contenía el promotor constitutivo cDNA1 de *Trichoderma reesei* para dirigir la expresión de STT3 de *Leishmania major*. La secuencia del terminador incluida en el casete era el terminador TrpC de *Aspergillus niger*. El casete de expresión se dirigió al locus xilanasa 1 (*xyn1*, tre74223) utilizando la secuencia de la xilanasa 1 de los flancos 5' y 3' del gen (SEQ ID NO: 5 y SEQ ID NO: 6). Estas secuencias se incluyeron en el casete para permitir que el casete se integrara en el locus *xyn1* a través de recombinación homóloga. El casete contenía un marcador de bucle pequeño pyr4 para la selección. El gen *pyr4* codifica la orotidina-5'-monofosfato (OMP) descarboxilasa de *T. reesei* (Smith, J.L., et al., 1991, Current Genetics 19:27-33) y es necesario para la síntesis de la uridina. Las cepas deficientes para la actividad de la OMP descarboxilasa no pueden crecer en un medio mínimo sin la complementación con uridina (es decir, son auxotrofas de uridina).

Para preparar el vector para la transformación, el vector se cortó con PmeI para liberar el casete de expresión (Figura 1). La digestión se separó con electroforesis en gel de agarosa y el fragmento correcto se aisló del gel con un kit de extracción de gel (Qiagen) de acuerdo con el protocolo del fabricante. El ADN del casete de expresión purificado (5 µg) se transformó luego en protoplastos de la cepa M317 de *Trichoderma reesei* (M317 se ha descrito en la Solicitud de Patente Internacional N.º PCT/EP2013/050126; M317 es pyr4⁻ de M304 y comprende una cadena ligera de MAB01 fusionada al vehículo CBH1 truncado de *T. reesei* con la secuencia de escisión de Kex2 NVISKR, cadena pesada de MAB01 fusionada a vehículo CBH1 truncado de *T. reesei* con la secuencia de escisión de Kex2 AXE1 [DGETVVKR], Δ*pep1*Δ*tsp1*Δ*slp1*, y sobreexpresión de KEX2 de *T. reesei*). La preparación de protoplastos y la transformación se realizaron de acuerdo con los métodos en Penttilä et al. (1987, Gene 61:155-164) y Gruber et al (1990, Curr. Genet. 18:71-76) para la selección de pyr4. Los protoplastos transformados se colocaron en placas de medio mínimo de *Trichoderma* (TrMM).

Los transformantes se sembraron luego en placas TrMM con TritonX-100 al 0,1%. Los transformantes que crecieron rápidamente a medida que las siembras selectivas se cribaron por PCR utilizando los cebadores enumerados en la Tabla 1. El ADN de los micelios se purificó y se analizó mediante PCR para observar la integración de los flancos 5' y 3' del casete y la existencia del ORF de xilanasa 1. El casete se dirigió al locus de xilanasa 1; por lo tanto, el marco de lectura abierto no estaba presente en los transformantes integrados de manera positiva. Para cribar la integración 5', se usó la secuencia fuera del flanco de integración 5' para crear un cebador directo que amplificará el ADN genómico que flanquea *xyn1* y el cebador inverso se realizó a partir de la secuencia en el promotor de ADNc del casete. Para verificar la correcta integración del casete en el flanco 3', se realizó un cebador directo desde la secuencia fuera del flanco de integración 3' que amplificaría el ADN genómico que flanquea *xyn1* y el cebador inverso se realizó a partir de la secuencia en el marcador *pyr4*. Por lo tanto, un cebador amplificará la secuencia del

ADN genómico fuera del casete y el otro amplificará la secuencia del ADN en el casete. Las secuencias de cebador se enumeran en la Tabla 1. Las cuatro cepas finales que muestran una integración adecuada y una delección de orf xyn1 se denominaron M420-M423.

- 5 Se realizaron cultivos en matraz de agitación para cuatro de las cepas productoras de STT3 (M420-M423) para evaluar las características de crecimiento y para proporcionar muestras para el análisis de ocupación del sitio de glucosilación. Los cultivos en matraz de agitación se realizaron en TrMM, 40 g/l de lactosa, 20 g/l de extracto de grano usado, 9 g/l de casaminoácidos, PIPPS 100 mM, pH 5,5. La expresión de STT3 de *L. major* no afectó al crecimiento negativamente cuando se comparó con la cepa parental M304 (Tablas 2 y 3). El peso seco de las células para los transformantes de expresión de STT3 parece ser ligeramente superior en comparación con la cepa parental M304.

Tabla 1: Lista de cebadores usados para el cribado por PCR de los transformantes STT3.

Cebadores de cribado del flanco 5':	Producto de 1205 pb
T403_Xyn1_5'flank_fwd	CCGCGTTGAACGGCTTCCCA ID NO:48)
T140_cDNA1_promoter_rev	TAACCTGTACGCTCTCAGTTTCGAG (SEQ ID NO:49)
Cebadores de cribado del flanco 3':	Producto de 1697 pb
T404_Xyn1_3'flank_fwd	GCGACGGCGACCCATTAGCA ID NO:50)
T028_Pyr4_flank_rev	CATCCTCAAGGCCTCAGAC (SEQ ID NO:51)
Cebadores de orf de xilanasa 1:	Producto de 589 pb
T405_Xyn1_orf_screen_fwd	TGCGCTCTCACCAGCATCGC ID NO:52)
T406_Xyn1_orf_screen_rev	GTCTGGGCGAGTTCGCAC (SEQ ID NO:53)

15

Tabla 2: Peso seco celular de cultivos en matraz de agitación grandes.

Peso seco celular (g/l)	Peso seco celular (g/l)		
	día 3	día 5	día 7
M304	2,3	3,3	4,3
M420	3,7	4,3	5,4
M421	3,7	4,6	6,3
M422	3,8	4,5	5,4
M423	3,7	4,6	5,7

Tabla 3: Valores de pH de cultivos en matraz de agitación grandes.

Valores de pH	Valores de pH		
	día 3	día 5	día 7
M304	5,6	6,1	6,2
M420	6,1	6,1	6,1
M421	6,0	5,9	6,0
M422	6,1	6,1	6,2
M423	6,1	6,1	6,1

Análisis de ocupación del sitio

20

Se cultivaron cuatro transformantes [pTTv201; 17A-a (M420), 26B-a (M421), 65B-a (M422) y 97A-a (M423)] y su cepa parental (M317) en matraces de agitación y las muestras se recogieron en los puntos de tiempo del día 5 y 7. El anticuerpo MAB01 se purificó a partir de sobrenadantes de cultivo utilizando una placa de 96 pocillos Protein G HP MultiTrap (GE Healthcare) de acuerdo con las instrucciones del fabricante. El anticuerpo se eluyó con tampón de citrato 0,1 M, pH 2,6 y se neutralizó con Tris 2 M, pH 9. La concentración se determinó a través de la absorbancia UV en un espectrofotómetro frente a la curva estándar MAB01. Se digirieron 10 µg de anticuerpo con 13,4 U de FabRICATOR (Genovis), + 37°C, 60 min, produciendo un fragmento F(ab')₂ y un fragmento Fc. Las muestras digeridas se purificaron utilizando una placa de filtro Poros R1 (Glyken corp.) y los fragmentos Fc se analizaron para determinar la ocupación del sitio de N-glicano utilizando MALDI-TOF MS (Figura 2).

30

La sobreexpresión de STT3 de *Leishmania major* mejoró la cobertura del sitio en comparación con la cepa parental. El mejor clon se volvió a cultivar en tres matraces de agitación paralelos cada uno y los resultados del análisis fueron comparables con el primer análisis. En comparación con la cepa parental, las señales Fc y Fc + K están prácticamente ausentes en los clones de STT3.

35

La diferencia en la ocupación del sitio entre la cepa parental y todos los clones de STT3 de *L. major* fue significativa (Figura 2). Debido a que las señales provenientes de Fc o Fc + K estaban prácticamente ausentes, la ocupación del sitio de N-glicano de MAB01 en estos cultivos en matraz de agitación fue del 100% (Tabla 4).

5 **Tabla 4: Análisis de la ocupación del sitio de la cepa parental M317 y cuatro transformantes de STT3 de *L. major*. Los promedios se han calculado a partir del área y la intensidad a partir de señales de carga simple y doble de tres muestras paralelas.**

	M317	17A-a	26B-a	65B-a	97A-a
Estado de glucosilación	% promedio	% promedio	% promedio	% promedio	% promedio
No glucosilado	13,0	0,0	0,0	0,0	0,0
Glucosilado	87,0	100,0	100,0	100,0	100,0

Cultivos de fermentador

10

Se cultivaron tres clones de STT3 (*L. major*) (M420, M421 y M422), así como la cepa parental M304 en un fermentador. Las muestras en los puntos de tiempo de los días 3, 4, 5, 6 y 7 se recogieron y el análisis de ocupación del sitio se realizó para el anticuerpo purificado. Las cepas de sobreexpresión de STT3 y la respectiva cepa de control (M304) se cultivaron en fermentaciones discontinuas durante 7 días, en medio que contenía extracto de levadura al 2%, celulosa al 4%, celobiosa al 4%, sorbosa al 2%, 5 g/l de KH₂PO₄ y 5 g/l de (NH₄)₂SO₄. El pH del cultivo se controló a pH 5,5 (ajustado con NH₃OH). La temperatura se cambió de 28°C a 22°C a las 48 horas transcurridas del tiempo de proceso. Las fermentaciones se realizaron en 4 reactores de vasos de vidrio paralelos de 2 l con un volumen de cultivo de 1 l. Se tomaron muestras de sobrenadante de cultivo durante el transcurso de las ejecuciones y se almacenaron a -20°C. El anticuerpo MAB01 se purificó y se digirió con FABRICATOR como se ha descrito anteriormente. Los títulos de anticuerpo se muestran en la Tabla 5.

15
20

Resultados

La ocupación del sitio en la cepa parental M304 fue inferior al 60%, pero en todos los clones de STT3 analizados, la ocupación del sitio había aumentado hasta el 98% (Tabla 6).

25

Tabla 5: Títulos de anticuerpo MAB01 de las cepas M420, M421 y M422 de LmSTT3 y su cepa parental M304.

Cepa	Título g/l				
	d3	d4	d5	d6	d7
M304	0,225	0,507	0,981	1,52	1,7
M420	0,758	1,21	1,55	1,71	1,69
M421	0,76	1,24	1,54	1,67	1,6
M422	0,65	1,07	1,43	1,56	1,54

Tabla 6: Las ocupaciones del sitio de N-glucosilación del anticuerpo MAB01 de las cepas de LmSTT3 M420, M421 y M422 y su cepa parental M304.

30

Cepa	% de ocupación del sitio				
	d3	d4	d5	d6	d7
M304	48,0	47,7	47,7	46,3	55,4
M420	97,8	97,5	96,9	94,3	94,6
M421	96,1	90,8	91,5	89,7	95,6
M422	94,4	88,5	80,9	83,6	75,2

En conclusión, la sobreexpresión del gen STT3D de *L. major* aumentó la ocupación del sitio de N-glucosilación del 46%-87% en la cepa parental al 98%-100% en transformantes que tenían STT3 de *Leishmania* en condiciones de matraz de agitación o de cultivo de fermentación.

35

La sobreexpresión del gen STT3D de *L. major* aumentó significativamente la ocupación del sitio de N-glucosilación en cepas productoras de un anticuerpo como una proteína heteróloga. Los títulos de anticuerpo no variaron significativamente entre los transformantes que tenían STT3 y la cepa parental.

40 **EJEMPLO 2 - Generación de cepas de *T. reesei* que expresan STT3 a partir de *T. vaginalis*, *L. infantums* o *E. histolytica***

Las secuencias codificantes de la oligosacaril transferasa de *Trichomonas vaginalis*, *Leishmania infantum* y *Entamoeba histolytica* (STT3; secuencias de aminoácidos *T. vaginalis* SEQ ID NO: 7, *L. infantum* SEQ ID NO: 8, y *E. histolytica* SEQ ID NO: 10) se optimizaron por codón para determinar la expresión de *Trichoderma reesei* (ácido nucleico de *L. infantum* optimizado por codón SEQ ID NO: 9). Las secuencias codificantes optimizadas se sintetizaron junto con la secuencia flanqueante del terminador *cbh1* de *T. reesei* (SEQ ID NO: 11). Los plásmidos que contenían los genes STT3 bajo el promotor constitutivo *cDNA1*, con terminador *cbh1*, el marcador de bucle pequeño *pyr4*, y las regiones flanqueantes *alg3* (SEQ ID NO: 12 y SEQ ID NO: 13) se clonaron mediante recombinación homóloga de levadura como se describe en el documento WO2012/069593. El fragmento *NotI* del plásmido pTTv38 se usó como esqueleto del vector. Este vector contiene los flancos *alg3* (tre104121) 5' y 3' del gen para permitir que el casete de expresión se integre en el locus *alg3* mediante recombinación homóloga en *T. reesei* y el plásmido se ha descrito en el documento WO2012/069593. Los genes *STT3* se eliminaron de los vectores de clonación utilizando la digestión con enzimas de restricción *SfiI*. El promotor *cdna1* y los fragmentos del terminador *cbh1* se crearon mediante PCR, utilizando plásmidos pTTv163 y pTTv166 como plantillas, respectivamente. El marcador de bucle pequeño *pyr4* se extrajo del plásmido pTTv142 por digestión con *NotI* (el plásmido pTTv142 que tiene un dominio catalítico GNT2 humano fusionado con el péptido de direccionamiento MNT1/KRE2 de *T. reesei* se ha descrito en el documento WO2012/069593). El gen *pyr4* codifica la orotidina-5'-monofosfato (OMP) descarboxilasa de *T. reesei* (Smith, J.L., et al., 1991, Current Genetics 19:27-33) y es necesario para la síntesis de la uridina. Las cepas deficientes para la actividad de la OMP descarboxilasa no pueden crecer en un medio mínimo sin la complementación con uridina (es decir, son auxotrofas de uridina). Los cebadores usados para la clonación se enumeran en la Tabla 7. Los fragmentos digeridos y los productos de PCR se separaron con electroforesis en gel de agarosa y los fragmentos correctos se aislaron del gel con un kit de extracción de gel (Qiagen) de acuerdo con el protocolo del fabricante. Los plásmidos se construyeron utilizando el método de recombinación homóloga de levadura, utilizando oligonucleótidos superpuestos para la recombinación del gap entre el marcador *pyr4* y el flanco 3' de *alg3* como se describe en el documento WO2012/069593. El ADN plasmídico se rescató de la levadura y se transformó en *E. coli* TOP10 electrocompetente que se cultivó en placas de selección de ampicilina (100 µg/ml). Las preparaciones de plásmidos Miniprep se hicieron a partir de varias colonias. La presencia de los genes STT3 de *Trichomonas vaginalis* y *Leishmania infantum* se confirmó digiriendo los plásmidos preparados con *BglII-KpnI*, mientras que el plásmido *Entamoeba histolytica* se digirió con *HindIII-KpnI*. Los clones positivos se secuenciaron para verificar las secuencias plasmídicas. Se eligió un clon correcto de *Trichomonas vaginalis* para ser el vector final pTTv321, y los clones correctos de *Leishmania infantum* y *Entamoeba histolytica* se eligieron para ser los vectores pTTv322 y pTTv323, respectivamente. Los cebadores utilizados para la secuenciación de los vectores se enumeran en la Tabla 8.

Tabla 7: Lista de cebadores usados para los vectores de clonación pTTv321, pTTv322 y pTTv323.

Fragmento	Cebador	Secuencia de cebador
Promotor <i>cDNA1</i> , pTTv321	T1177_pTTv321_1	AGATTTTCAGTCTCTCACCCTCACCTGAGTTGCCT CTCTCGGTCTGAAGGACGTGGAATGATG (SEQ ID NO:54)
	T1178_pTTv321_2	GCAGGGTGATGAGCTGGATCACCTTGACGGTGTT
		GCCCATGTTGAGAGAAGTTGTTGGATTGATCA (SEQ ID NO:55)
Promotor <i>cDNA1</i> , pTTv322	T1177_pTTv321_1	AGATTTTCAGTCTCTCACCCTCACCTGAGTTGCCT CTCTCGGTCTGAAGGACGTGGAATGATG (SEQ ID NO:56)
	T1183_pTTv322_1	

		CAGAGCCGCTATCGCCGAGGAGGTTGCCCTTCTT GCCCATGTTGAGAGAAGTTGTTGGATTGATCA (SEQ ID NO:57)
Promoter <i>cDNA1</i> , pTTv323	T1177_pTTv321_1	AGATTTCAGTCTCTCACCCTCACCTGAGTTGCCT CTCTCGGTCTGAAGGACGTGGAATGATG (SEQ ID NO:58)
	T1184_pTTv323_1	TCTTGAGGATGAGCTGGACGAGGGTCTTGAAAA GCCCATGTTGAGAGAAGTTGTTGGATTGATCA (SEQ ID NO:59)
Terminador <i>cbh1</i>	T1179_pTTv321_3	AGCTCCGTGGCGAAAGCCTGA (SEQ ID NO:60)
	T1180_pTTv321_4	CAGCCGCAGCCTCAGCCTCTCTCAGCCTCATCAG CCGCGGCCGCCAACTTTGCGTCCCTTGTGACG (SEQ ID NO:61)
oligos de solapamiento en flanco 3' <i>pyr4-alg3</i>	T1181_pTTv321_5	GCAACGAGAGCAGAGCAGCAGTAGTCGATGCTA GGCGGCCGCGGGCAGTATGCCGGATGGCTGGCT TATACAGGCA (SEQ ID NO:62)
	T1182_pTTv321_6	TGCCTGTATAAGCCAGCCATCCGGCATACTGCC GCGGCCGCCTAGCATCGACTACTGCTGCTCTGCT CTCGTTGC (SEQ ID NO:63)

Tabla 8: Lista de cebadores usados para los vectores de secuenciación pTTv321, pTTv322 y pTTv323.

Cebador	Secuencia
T027_Pyr4_orf_start_rev	TGCGTCGCCGTCTCGCTCCT (SEQ ID NO:64)
T061_pyr4_orf_screen_2F	TTAGGCGACCTCTTTTTCCA (SEQ ID NO:65)
T143_cDNA1_promoter_seqF3	CGAGGAAGTCTCGTGAGGAT (SEQ ID NO:66)
T410_alg3_5-flank_F	CAGCTAAACCGACGGGCCA (SEQ ID NO:67)
T1153_cbh1_term_start_rev	GACCGTATATTTGAAAAGGG (SEQ ID NO:68)

Para preparar los vectores para la transformación, los vectores se cortaron con *PmeI* para liberar los casetes de expresión (Figura 3). Los fragmentos se separaron con electroforesis en gel de agarosa y el fragmento correcto se aisló del gel con un kit de extracción de gel (Qiagen) de acuerdo con el protocolo del fabricante. El ADN del casete de expresión purificado se transformó entonces en protoplastos de M317 de *Trichoderma reesei*. La preparación de protoplastos y la transformación se realizaron esencialmente de acuerdo con los métodos en Penttilä et al. (1987,

Gene 61:155-164) y Gruber et al (1990, Curr. Genet. 18:71-76) para la selección de *pyr4*. Los protoplastos transformados se colocaron en placas de medio mínimo de *Trichoderma* (TrMM) que contenían sorbitol.

Los transformantes se sembraron luego en placas TrMM con TritonX-100 al 0,1%. Los transformantes que crecieron rápidamente a medida que las siembras selectivas se cribaron por PCR utilizando los cebadores enumerados en la Tabla 9. El ADN de los micelios se purificó y se analizó mediante PCR para observar la integración de los flancos 5' y 3' del casete y la existencia del ORF de *alg3*. El casete se dirigió al locus *alg3*; por lo tanto, el marco de lectura abierto no estaba presente en los transformantes integrados de manera positiva, purificados en clones de células individuales. Para cribar la integración 5', se usó la secuencia fuera del flanco de integración 5' para crear un cebador directo que amplificará el ADN genómico que flanquea *alg3* y el cebador inverso se realizó a partir de la secuencia en el promotor de *cDNA1* del casete. Para verificar la correcta integración del casete en el flanco 3', se realizó un cebador inverso a partir de una secuencia fuera del flanco de integración 3' que amplificaría el ADN genómico que flanquea *alg3* y el cebador directo se realizó a partir de la secuencia en el marcador *pyr4*. Por lo tanto, un cebador amplificará la secuencia del ADN genómico fuera del casete y el otro amplificará la secuencia del ADN en el casete.

Tabla 9: Lista de cebadores utilizados para el cribado por PCR de transformantes de *T. reesei*.

Cebadores de cribado del flanco 5':	Producto de 1165 pb
T066_104121_5int	GATGTTGCGCCTGGGTTGAC (SEQ ID NO:69)
T140_cDNA1_promoter_seqR1	TAAC TTGTACGCTCTCAGTT CGA (SEQ ID NO:70)
Cebadores de cribado del flanco 3':	Producto de 1469 pb
T026_Pyr4_orf_5rev2	CCATGAGCTTGAACAGGTAA (SEQ ID NO:71)
T068_104121_3int	GATTGTCATGGTGTACGTGA (SEQ ID NO:72)
Cebadores ORF <i>alg3</i> :	Producto de 689 pb
T767_ <i>alg3</i> _del_F	CAAGATGGAGGGCGGCACAG (SEQ ID NO:73)
T768_ <i>alg3</i> _del_R	GCCAGTAGCGTGATAGAGAAGC (SEQ ID NO:74)
Cebadores ORF <i>alg3</i> :	Producto de 1491 pb
T069_104121_5orf_pcr	GCGTCACTCATCAAAACTGC (SEQ ID NO:75)
T070_104121_3orf_pcr	CTTCGGCTTCGATGTTTCA (SEQ ID NO:76)

Se cultivaron cuatro cepas finales, cada una de las cuales mostró una integración adecuada y una eliminación de ORF *alg3* en matraces de agitación grandes en medio TrMM complementado con 40 g/l de lactosa, 20 g/l de extracto de grano usado, 9 g/l de casaminoácidos y PIPPS 100 mM, pH 5,5. El crecimiento de las cepas pTTv321 y pTTv323 fue algo más lento que en la cepa parental M304 (Tabla 10). Tres de cada cuatro clones de *Leishmania infantum* pTTv322 crecieron algo mejor que la cepa parental.

Tabla 10. Mediciones del peso seco celular (en g/l) de las cepas parentales M304 y cepas que expresan STT3.

Cepa	3 días	5 días	7 días
M304	3,06	3,34	4,08
pTTv321 #18-9-2	2,54	2,89	2,52
pTTv321 #18-9-10	2,44	3,03	2,65
pTTv321#18-12-1	2,43	3,12	2,86
pTTv321 #18-12-2	2,84	3,49	3,39
pTTv322#60-2	3,02	3,42	3,63
pTTv322#60-6	3,37	4,45	4,68
pTTv322#60-12	3,30	4,15	4,29
pTTv322#60-14	2,92	3,90	4,39
pTTv323#37-4-1	2,29	2,27	2,59
pTTv323#37-4-14	1,88	2,08	2,69
pTTv323#37-11-3	2,15	2,27	2,62
pTTv323#37-11-8	1,92	2,25	2,62

Ocupación del sitio y análisis de glicanos

A partir de muestras de sobrenadante del día 5, se purificó MAB01 usando una placa de filtro de 96 pocillos Protein G HP MultiTrap (GE Healthcare) de acuerdo con las instrucciones del fabricante. Se cargaron aprox. 1,4 ml de sobrenadante de cultivo y el volumen de elución fue de 230 µl. Las concentraciones de anticuerpos se determinaron a través de la absorbancia UV contra la curva estándar MAB01.

Para el análisis de ocupación del sitio, se tomaron 16-20 µg de anticuerpo MAB01 purificado y los anticuerpos se digirieron, se purificaron y se analizaron como se describe en el ejemplo 1. La ocupación del sitio al 100% se logró con los clones STT3 de *Leishmania infantum* 60-6, 60-12 y 60-14 (Tabla 11). En los transformantes STT3 de *T. vaginalis* y *E. histolytica*, la ocupación del sitio fue baja y en estos últimos, los anticuerpos parecieron degradarse, lo que resultó en que no se pudo realizar un análisis de la ocupación del sitio para una cepa.

Tabla 11. Ocupación del sitio de N-glucosilación de anticuerpos de variantes de STT3 y M304 parental el día 5.

M304				
Estado de glucosilación	%			
No glucosilado	8			
Glucosilado	92			
STT3 de <i>Trichomonas vaginalis</i>, Δalg3				
	18-9-2	18-9-10	18-12-1	18-12-2
Estado de glucosilación	%	%	%	%
No glucosilado	75	71	69	64
Glucosilado	25	29	31	36
Glucosilado	18	n.d.	27	14
STT3 de <i>Leishmania infantum</i>, Δalg3				
	60-2	60-6	60-12	60-14
Estado de glucosilación	%	%	%	%
No glucosilado	38	0	0	0
Glucosilado	62	100	100	100
STT3 de <i>Entamoeba histolytica</i>, Δalg3				
	37-4-1	37-4-14	37-11-3	37-11-8
Estado de glucosilación	%	%	%	%
No glucosilado	82	n.d.	73	86

10

Estos resultados muestran que la sobreexpresión de la subunidad catalítica de *Leishmania infantum* es capaz de aumentar la ocupación del sitio de N-glucosilación en células de fúngicas filamentosas, hasta el 100%.

En contraste, los genes STT3 de *Trichomonas vaginalis* o *Entamoeba histolytica* no producen una alta ocupación del sitio de N-glucosilación.

Los N-glicanos se analizaron a partir de tres de los clones de STT3 de *Leishmania infantum*. Las reacciones de PNGasa F se realizaron con 20 µg de anticuerpo MAB01 como se describe en los ejemplos y los N-glicanos liberados se analizaron con MALDI-TOF MS. Las tres cepas produjeron aproximadamente el 25% de N-glicano Man3 unido a MAB01, mientras que la glucoforma Hex6 representa aproximadamente el 60% de los N-glicanos unidos a MAB01 (Tabla 12).

Tabla 12: N-glicanos neutros y análisis de ocupación del sitio de MAB01 de clones de STT3 de *L. infantum* STT3 el día 5.

STT3 de <i>Leishmania infantum</i>, Δalg3				
Clones		60-6	60-12	60-14
Corto	m/z	%	%	%
Man3	933,3	25,9	26,4	25,9
Man4	1095,4	9,4	9,3	9,0
Man5	1257,4	6,5	6,1	7,6
Hex6	1419,5	58,3	58,2	57,5
Fc		0	0	0
Fc+Gn		0	0	0
Glucosilado		100	100	100

25

Esto muestra que las glucoformas Man3, G0, G1 y/o G2 representan al menos el 25% del total de N-glicanos neutros de MAB01 en 3 clones diferentes que sobreexpresan STT3 de *L. infantum*. La Figura 4 muestra las estructuras de glicano de Man3, Man4, Man5 y Hex6 producidas en las cepas Δalg3. "Fc" significa un fragmento Fc (sin ningún N-

glicano) y "Fc+Gn" significa un fragmento Fc con una N-acetilglucosamina unida (la posible actividad de la enzima Endo T podría escindir los N-glicanos de un Fc+Gn resultante de Fc).

EJEMPLO 3 - Generación de cepas $\Delta alg3$ de cepas que expresan MAB01

5 El marcador de acetamida del plásmido de delección *alg3* de pTTv38 se cambió al marcador *pyr4*. Los vectores pTTv38 y pTTv142 se digirieron con *NotI* y los fragmentos se separaron con electroforesis en gel de agarosa. Los fragmentos correctos se aislaron del gel con un kit de extracción de gel (Qiagen) de acuerdo con el protocolo del fabricante. El marcador de bucle pequeño *pyr4* purificado de pTTv142 se ligó al plásmido pTTv38 con ADN ligasa
10 T4. La reacción de ligación se transformó en *E. coli* TOP10 electrocompetente y se hizo crecer en placas de selección de ampicilina (100 μ g/ml). Las preparaciones de plásmidos Miniprep se hicieron a partir de cuatro colonias. La orientación del marcador se confirmó secuenciando los clones con los cebadores enumerados en la Tabla 13. Se eligió un clon con el marcador en dirección invertida para ser el vector final pTTv324.

15 **Tabla 13. Lista de cebadores usados para los vectores de secuenciación pTTv324.**

Cebador	Secuencia
T027_Pyr4_orf_start_rev	TGCGTCGCCGTCTCGCTCCT (SEQ ID NO:77)
T060_pyr4_orf_screen_1F	TGACGTACCAGTTGGGATGA (SEQ ID NO:78)

Se generó una cepa *pyr4* de la cepa de expresión de STT3 de *Leishmania major* M420 en bucle corto con el marcador *pyr4* mediante la selección de 5-FOA como se describe en la Solicitud de Patente Internacional N.º PCT/EP2013/050126. Una cepa *pyr4* fue designada con el número M602.

20 Para preparar los vectores para la transformación, el vector pTTv324 se cortó con *PmeI* para liberar el casete de delección. Los fragmentos se separaron con electroforesis en gel de agarosa y el fragmento correcto se aisló del gel con un kit de extracción de gel (Qiagen) de acuerdo con el protocolo del fabricante. El ADN del casete de delección purificado se transformó entonces en protoplastos de M317 y M602 de *Trichoderma reesei*. La preparación de
25 protoplastos, la transformación y la colocación en placas de los protoplastos se realizaron como se ha descrito anteriormente.

Los transformantes se sembraron luego en placas TrMM con TritonX-100 al 0,1%. Los transformantes que crecieron rápidamente a medida que las siembras selectivas se cribaron por PCR utilizando los cebadores enumerados en la
30 Tabla 14. El ADN de los micelios se purificó y se analizó mediante PCR para observar la integración de los flancos 5' y 3' del casete y la existencia del ORF de *alg3*. El casete se dirigió al locus *alg3*; por lo tanto, el marco de lectura abierto no estaba presente en los transformantes integrados de manera positiva, purificados en clones de células individuales. Para cribar la integración 5', se usó la secuencia fuera del flanco de integración 5' para crear un
35 cebador directo que amplificará el ADN genómico que flanquea *alg3* y el cebador inverso se realizó a partir de la secuencia en el marcador *pyr4* del casete. Para verificar la correcta integración del casete en el flanco 3', se realizó un cebador inverso a partir de una secuencia fuera del flanco de integración 3' que amplificaría el ADN genómico que flanquea *alg3* y el cebador directo se realizó a partir de la secuencia en el marcador *pyr4*. Por lo tanto, un cebador amplificará la secuencia del ADN genómico fuera del casete y el otro amplificará la secuencia del ADN en el
40 casete.

Tabla 14: Lista de cebadores utilizados para el cribado por PCR de transformantes de *T. reesei*.

Cebadores de cribado del flanco 5':	Producto de 1455 pb
T066_104121_5int	GATGTTGCGCCTGGGTTGAC (SEQ ID NO:79)
T060_pyr4_orf_screen_1F	TGACGTACCAGTTGGGATGA (SEQ ID NO:80)
Cebadores de cribado del flanco 3':	Producto de 1433 pb
T027_Pyr4_orf_start_rev	TGCGTCGCCGTCTCGCTCCT (SEQ ID NO:81)
T068_104121_3int	GATTGTCATGGTGTACGTGA (SEQ ID NO:82)
Cebadores ORF <i>alg3</i> :	Producto de 689 pb
T767_ <i>alg3</i> _del_F	CAAGATGGAGGGCGGCACAG (SEQ ID NO:83)
T768_ <i>alg3</i> _del_R	GCCAGTAGCGTGATAGAGAAGC (SEQ ID NO:84)

Se cultivaron dos cepas M602 y siete cepas M317 que mostraban una integración adecuada y una eliminación de ORF *alg3* en matraces de agitación grandes en medio TrMM complementado con 40 g/l de lactosa, 20 g/l de extracto
45 de grano usado, 9 g/l de casaminoácidos y PIPPS 100 mM, pH 5,5 (Tabla 15). La cepa M317 19.13 y 19.20 se designaron con los números M697 y M698, respectivamente, y las cepas M602 1.22 y 11.18 se designaron con los

números M699 y M700, respectivamente.

Tabla 15 Mediciones de peso seco celular (en g/l) de las cepas parentales M304 y la cepa de expresión de STT3 M420 y los transformantes de delección *alg3*.

Cepa	3 días	5 días	7 días
M602 1.22	3,63	3,23	3,79
M602 11.18	3,52	3,74	4,12
M317 19.1	3,64	3,84	4,22
M317 19.5	3,54	3,87	4,31
M317 19.6	3,72	3,66	4,78
M317 19.13	3,63	3,21	4,06
M317 19.20	3,97	4,28	5,09
M317 19.43	3,77	4,02	4,18
M317 19.44	3,58	3,78	4,17
M420	3,31	3,69	5,57
M304	2,55	2,99	4,09

5

Ocupación del sitio y análisis de glicanos

Dos transformantes de la sobreexpresión de STT3 de *Leishmania major* en la cepa de delección de *alg3* [pTTv324; 1.22 (M699) y 11.18 (M700)] y siete transformantes con delección de *alg3* [M317, pyr4- de M304; clones 19.1, 19.5, 10 19.6, 19.13 (M697), 19.20 (M698), 19.43 y 19.44], y sus cepas parentales M420 y M304 se cultivaron en matraces de agitación en TrMM, lactosa al 4%, extracto de grano usado al 2%, casaminoácidos al 0,9%, PIPPS 100 mM, pH 5,5. El anticuerpo MAB01 se purificó y se analizó a partir de los sobrenadantes del cultivo del día 5 como se describe en el Ejemplo 1, excepto que se digirieron 30 µg de anticuerpo con 80,4 U de FabRICATOR (Genovis), +37°C, durante una noche, para producir los fragmentos F(ab')₂ y Fc.

15

En ambos clones con delección de *alg3* y sobreexpresión de LmSTT3, la ocupación del sitio fue del 100% (Tabla 16). Sin LmSTT3, la cobertura del sitio varió entre el 56 y el 71% en los clones de delección de *alg3*. La ocupación del sitio mejorada se mostró también en la cepa parental M420 en comparación con M304, ambas con glucosilación de tipo silvestre.

20

Tabla 16: La ocupación del sitio de las muestras de matraz de agitación. El análisis falló en los clones M317 19.5 y 19.6.

Cepa	Clon	Explicación	% de ocupación del sitio
M602	1,22	M304 LmSTT3 Δ <i>alg3</i>	100
M602	11,18	M304 LmSTT3 Δ <i>alg3</i>	100
M317	19,1	M304 Δ <i>alg3</i>	71
M317	19,13	M304 Δ <i>alg3</i>	62
M317	19,2	M304 Δ <i>alg3</i>	56
M317	19,43	M304 Δ <i>alg3</i>	63
M317	19,44	M304 Δ <i>alg3</i>	60
M420		Cepa parental M304 LmSTT3	100
M304		Cepa parental	89

Para el análisis de N-glucano, se purificó MAB01 a partir de los sobrenadantes del cultivo del día 7 como se ha descrito anteriormente y se liberaron N-glicanos de EtOH precipitado y anticuerpo desnaturalizado con SDS usando PNGase F (Prozyme) en tampón de fosfato de sodio 20 mM, pH 7,3, en una reacción de una noche a +37°C. Los N-glicanos liberados se purificaron con Hypersep C18 e Hypersep Hypercarb (Thermo Scientific) y se analizaron con MALDI-TOF MS.

30 Los niveles de Man3 estaban en el intervalo del 21 al 49%, mientras que la principal glucoforma en los clones de M602 y M317 era Hex6 (Tabla 17). Los niveles de Man5 fueron aproximadamente del 73% en las cepas que expresaban glucosilación de tipo silvestre (M304) y LmSTT3 (M420).

Tabla 17. Proporciones relativas de N-glucanos neutros del anticuerpo purificado de los clones M602 y M317 y cepas parentales M420 y M304.

Composición	Corto	m/z	M602			M317			Cepas parentales		
			1,22 %	11,18 %	19,1 %	19,13 %	19,2 %	19,43 %	19,44 %	M420 %	M304 %
Hex3HexNAc2	Man3	933,3	21,1	27,3	45,4	37,5	34,9	24,6	48,6	0,0	0,0
Hex4HexNAc2	Man4	1095,4	9,5	8,7	6,2	7,6	7,1	7,5	9,4	0,8	0,0
Hex5HexNAc2	Man5	1257,4	5,8	7,0	8,1	7,6	6,7	5,6	6,6	72,5	72,8
Hex6HexNAc2	Man6/Hex6	1419,5	63,1	56,6	39,7	45,8	51,4	61,8	34,6	15,6	16,4
Hex7HexNAc2	Man7/Hex7	1581,5	0,5	0,5	0,6	0,8	0,0	0,5	0,7	7,2	7,9
Hex8HexNAc2	Man8/Hex8	1743,6	0,0	0,0	0,0	0,6	0,0	0,0	0,0	3,2	2,4
Hex9HexNAc2	Man9/Hex9	1905,6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	0,5

Fermentación y ocupación del sitio

La cepa de delección *alg3* de STT3 de *L. major* M699 (pTTV324; clon 1.22) y la cepa M698 con delección de *alg3* [M317, *pyr4-* de M304; clon 19.20], y la cepa parental M304 se fermentaron en YE al 2%, celulosa al 4%, celobiosa al 8%, sorbosa al 4%. Las muestras se recogieron el día 3, 4, 5 y 6. El anticuerpo MAB01 se purificó y se analizó a partir de los sobrenadantes del cultivo del día 5 como se describe en el Ejemplo 1, excepto que se digirieron 30 µg de anticuerpo con 80,4 U de FabRICATOR (Genovis), +37°C, durante una noche, para producir los fragmentos F(ab')₂ y Fc.

10 Resultados

En la cepa M699, la ocupación del sitio fue más del 90% en todos los puntos de tiempo (Tabla 18). Sin LmSTT3, la cobertura del sitio varió entre el 29-37% en la cepa M698. En la cepa parental M304, la cobertura del sitio varió entre el 45-57%. El día 6, los títulos de MAB01 fueron 1,2 y 1,3 g/l para las cepas M699 y M698, respectivamente, y 1,8 g/l en la cepa parental M304.

Tabla 18. Títulos de anticuerpo MAB01 y resultados del análisis de ocupación del sitio de las cepas fermentadas M699 y M698 y la cepa parental M304.

M699	d3	d4	d5	d6
Título g/l	0,206	0,361	0,685	1,22
Estado de glucosilación	%	%	%	%
No glucosilado	2,4	6,8	8,0	8,5
Glucosilado	97,6	93,2	92,0	91,5
Fc + Gn	0,0	0,0	0,0	0,0
M698	d3	d4	d5	d6
Título g/l	0,252	0,423	0,8	1,317
Estado de glucosilación	%	%	%	%
No glucosilado	63,0	70,8	64,3	65,8
Glucosilado	37,0	29,2	35,7	34,2
Fc + Gn	0,0	0,0	0,0	0,0
M304	d3	d4	d5	d6
Título g/l	0,589	0,964	1,41	1,79
Estado de glucosilación	%	%	%	%
No glucosilado	45,9	43,3	n.d.	54,9
Glucosilado	54,1	56,7	n.d.	45,1
Fc + Gn	0,0	0,0	n.d.	0,0

20 En conclusión, la sobreexpresión de la subunidad catalítica de STT3 de *Leishmania* es capaz de aumentar la ocupación del sitio de N-glucosilación en células fúngicas filamentosas $\Delta alg3$ hasta el 91,5-100%.

La Tabla 19 a continuación recapitula las diferentes cepas utilizadas en los Ejemplos:

Base de datos	Vector	Clon	Cepa transformada	Locus, aleatorio o K/o	Proteasas k/o	Descripción	Selección de tr.	Marcadores en la cepa
M44					Ninguno	Cepa base		Ninguno
M124				K/o mus53	Ninguno	delección mus53 de M44		pyr4
M127				pyr4- de M124	Ninguno	cepa negativa a pyr4 de M124		pyr4-
M181	pTTv71	9-20A-1	M127	K/o pep1	pep1	delección pep1	pyr4	pyr4
M194	pTTv42	13-172D	M181	K/o tsp1	pep1 tsp1	delección pep1 tsp1	bar	bar/pyr4
M252	pTTv99/67	6.14A	M194	loci cbh1 eg11	pep1 tsp1	MAB01 LC NVISKR/HC AXE1	AmdS/HvqR	AmdS/HygR/bar/pyr4
M284	5-FOA de M252	3A	pyr4- de M252	Mutación espontánea	pep1 tsp1	Cepa negativa a pyr4 de M252	ninguno	AmdS/Hyg R/bar/pyr4-
M304	pTTv128	12A	M284	K/o slp1, Kex2 o/e	pep1 tsp1 slp1	Sobreexpresión de Kex2 nativo, del slp1	pyr4	AmdS/HyaR/bar/pyr4
M317	5-FOA de M304	1A	pyr4- de M304	bucle pequeño de pyr4	pep1 tsp1 slp1	cepa negativa a pyr4 de M304	Ninguno	AmdS/Hyg R/bar/pyr4-
M420	pTTv201	17A-a	M317	xilanasas 1	pep1 tsp1 slp1	Oligosacaril transferasa de stt3 de <i>Leishmania major</i>	pyr4	AmdS/HygR/bar/pyr4
M421	pTTv201	26B-a	M317	xilanasas 1	pep1 tsp1 slp1	Oligosacaril transferasa de stt3 de <i>Leishmania major</i>	pyr4	AmdS/Hyg R/bar/pyr4
M422	pTTv201	65B-a	M317	xilanasas 1	pep1 tsp1 slp1	Oligosacaril transferasa de stt3 de <i>Leishmania major</i>	pyr4	AmdS/Hyg R/bar/pyr4
M423	pTTv201	97A-a	M317	xilanasas 1	pep1 tsp1 slp1	Oligosacaril transferasa de stt3 de <i>Leishmania major</i>	pyr4	AmdS/HygR/bar/pyr4
M602	5-FOA de M420	2A	pyr4- de M420	bucle pequeño de pyr4	pep1 tsp1 slp1	cepa negativa a pyr4 de M420	ninguno	AmdS/HygR/bar/pyr4-
M698	pTTv324	19,20	M317	alg3	pep1 tsp1 slp1	Delección de alg3	pyr4	AmdS/Hyg R/bar/pyr4
M699	pTTv324	1,22	M602	alg3	pep1 tsp1 slp1	Delección de alg3	pyr4	AmdS/HygR/bar/pyr4
M800	pTTv322	60-6	M317	alg3	pep1 tsp1 slp1	Leishmania infantum STT3, cDNA1p_cbh1t	pyr4	AmdS/Hyg R/bar/pyr4

M801	pTTv322	60-12	M317	alg3	pep1 tsp1 slp1	Leishmania infantum STT3, cDNA1p cbh1t	pyr4	AmdS/HygR/bar/pyr4
M802	pTTv322	60-14	M317	alg3	pep1 tsp1 slp1	Leishmania infantum STT3, cDNA1p cbh1t	pyr4	AmdS/Hyg R/bar/pyr4

Las cepas de *Trichoderma* que tienen STT3 (M420-M423) son deficientes en proteasa triple (pep1, tsp1, slp1), así como deficientes en xilanasasa 1, cbh1, y egl1.
Las formas de realización incluyen también cepas deficientes en proteasas de orden superior.

LISTA DE SECUENCIAS

<110> Novartis AG

5 <120> Producción de glucoproteínas que tienen una ocupación del sitio de N-glicosilación aumentada

<130> NOVA 0005 WO

<160> 91

10 <170> PatentIn versión 3.5

<210> 1

<211> 857

15 <212> PRT

<213> Leishmania major

<400> 1

Met Gly Lys Arg Lys Gly Asn Ser Leu Gly Asp Ser Gly Ser Ala Ala
 1 5 10 15

Thr Ala Ser Arg Glu Ala Ser Ala Gln Ala Glu Asp Ala Ala Ser Gln
 20 25 30

Thr Lys Thr Ala Ser Pro Pro Ala Lys Val Ile Leu Leu Pro Lys Thr
 35 40 45

Leu Thr Asp Glu Lys Asp Phe Ile Gly Ile Phe Pro Phe Pro Phe Trp
 50 55 60

Pro Val His Phe Val Leu Thr Val Val Ala Leu Phe Val Leu Ala Ala
 65 70 75 80

Ser Cys Phe Gln Ala Phe Thr Val Arg Met Ile Ser Val Gln Ile Tyr
 85 90 95

Gly Tyr Leu Ile His Glu Phe Asp Pro Trp Phe Asn Tyr Arg Ala Ala
 100 105 110

Glu Tyr Met Ser Thr His Gly Trp Ser Ala Phe Phe Ser Trp Phe Asp
 115 120 125

Tyr Met Ser Trp Tyr Pro Leu Gly Arg Pro Val Gly Ser Thr Thr Tyr
 130 135 140

Pro Gly Leu Gln Leu Thr Ala Val Ala Ile His Arg Ala Leu Ala Ala
 145 150 155 160

Ala Gly Met Pro Met Ser Leu Asn Asn Val Cys Val Leu Met Pro Ala
 165 170 175

20

ES 2 711 165 T3

Trp Phe Gly Ala Ile Ala Thr Ala Thr Leu Ala Phe Cys Thr Tyr Glu
180 185 190

Ala Ser Gly Ser Thr Val Ala Ala Ala Ala Ala Ala Leu Ser Phe Ser
195 200 205

Ile Ile Pro Ala His Leu Met Arg Ser Met Ala Gly Glu Phe Asp Asn
210 215 220

Glu Cys Ile Ala Val Ala Ala Met Leu Leu Thr Phe Tyr Cys Trp Val
225 230 235 240

Arg Ser Leu Arg Thr Arg Ser Ser Trp Pro Ile Gly Val Leu Thr Gly
245 250 255

Val Ala Tyr Gly Tyr Met Ala Ala Ala Trp Gly Gly Tyr Ile Phe Val
260 265 270

Leu Asn Met Val Ala Met His Ala Gly Ile Ser Ser Met Val Asp Trp
275 280 285

Ala Arg Asn Thr Tyr Asn Pro Ser Leu Leu Arg Ala Tyr Thr Leu Phe
290 295 300

Tyr Val Val Gly Thr Ala Ile Ala Val Cys Val Pro Pro Val Gly Met
305 310 315 320

Ser Pro Phe Lys Ser Leu Glu Gln Leu Gly Ala Leu Leu Val Leu Val
325 330 335

Phe Leu Cys Gly Leu Gln Val Cys Glu Val Leu Arg Ala Arg Ala Gly
340 345 350

Val Glu Val Arg Ser Arg Ala Asn Phe Lys Ile Arg Val Arg Val Phe
355 360 365

Ser Val Met Ala Gly Val Ala Ala Leu Ala Ile Ser Val Leu Ala Pro
370 375 380

Thr Gly Tyr Phe Gly Pro Leu Ser Val Arg Val Arg Ala Leu Phe Val
385 390 395 400

Glu His Thr Arg Thr Gly Asn Pro Leu Val Asp Ser Val Ala Glu His
405 410 415

Gln Pro Ala Ser Pro Glu Ala Met Trp Ala Phe Leu His Val Cys Gly
420 425 430

ES 2 711 165 T3

Val Thr Trp Gly Leu Gly Ser Ile Val Leu Ala Val Ser Thr Phe Val
 435 440 445

His Tyr Ser Pro Ser Lys Val Phe Trp Leu Leu Asn Ser Gly Ala Val
 450 455 460

Tyr Tyr Phe Ser Thr Arg Met Ala Arg Leu Leu Leu Leu Ser Gly Pro
 465 470 475 480

Ala Ala Cys Leu Ser Thr Gly Ile Phe Val Gly Thr Ile Leu Glu Ala
 485 490 495

Ala Val Gln Leu Ser Phe Trp Asp Ser Asp Ala Thr Lys Ala Lys Lys
 500 505 510

Gln Gln Lys Gln Ala Gln Arg His Gln Arg Gly Ala Gly Lys Gly Ser
 515 520 525

Gly Arg Asp Asp Ala Lys Asn Ala Thr Thr Ala Arg Ala Phe Cys Asp
 530 535 540

Val Phe Ala Gly Ser Ser Leu Ala Trp Gly His Arg Met Val Leu Ser
 545 550 555 560

Ile Ala Met Trp Ala Leu Val Thr Thr Thr Ala Val Ser Phe Phe Ser
 565 570 575

Ser Glu Phe Ala Ser His Ser Thr Lys Phe Ala Glu Gln Ser Ser Asn
 580 585 590

Pro Met Ile Val Phe Ala Ala Val Val Gln Asn Arg Ala Thr Gly Lys
 595 600 605

Pro Met Asn Leu Leu Val Asp Asp Tyr Leu Lys Ala Tyr Glu Trp Leu
 610 615 620

Arg Asp Ser Thr Pro Glu Asp Ala Arg Val Leu Ala Trp Trp Asp Tyr
 625 630 635 640

Gly Tyr Gln Ile Thr Gly Ile Gly Asn Arg Thr Ser Leu Ala Asp Gly
 645 650 655

Asn Thr Trp Asn His Glu His Ile Ala Thr Ile Gly Lys Met Leu Thr
 660 665 670

Ser Pro Val Val Glu Ala His Ser Leu Val Arg His Met Ala Asp Tyr

ES 2 711 165 T3

675	680	685
Val Leu Ile Trp Ala Gly Gln Ser Gly Asp Leu Met Lys Ser Pro His 690	695	700
Met Ala Arg Ile Gly Asn Ser Val Tyr His Asp Ile Cys Pro Asp Asp 705	710	715
Pro Leu Cys Gln Gln Phe Gly Phe His Arg Asn Asp Tyr Ser Arg Pro 725	730	735
Thr Pro Met Met Arg Ala Ser Leu Leu Tyr Asn Leu His Glu Ala Gly 740	745	750
Lys Arg Lys Gly Val Lys Val Asn Pro Ser Leu Phe Gln Glu Val Tyr 755	760	765
Ser Ser Lys Tyr Gly Leu Val Arg Ile Phe Lys Val Met Asn Val Ser 770	775	780
Ala Glu Ser Lys Lys Trp Val Ala Asp Pro Ala Asn Arg Val Cys His 785	790	795
Pro Pro Gly Ser Trp Ile Cys Pro Gly Gln Tyr Pro Pro Ala Lys Glu 805	810	815
Ile Gln Glu Met Leu Ala His Arg Val Pro Phe Asp Gln Val Thr Asn 820	825	830
Ala Asp Arg Lys Asn Asn Val Gly Ser Tyr Gln Glu Glu Tyr Met Arg 835	840	845
Arg Met Arg Glu Ser Glu Asn Arg Arg 850	855	

<210> 2

<211> 2575

<212> ADN

<213> Leishmania major

<400> 2

aatgggcaag cgcaagggca acagcctcgg cgacagcggc agcgcgcgcca cgcctcacg	60
agaggcctct gcccaggccg aggacgccgc cagccagacc aagaccgcca gccccctgc	120
caaggtcatc ctctgcccc agaccctcac cgacgagaag gacttcacgc gcatcttccc	180
gttcccgttc tggcccgtcc acttcgtcct caccgtcgtc gccctcttcg tctcgcgcgc	240
cagctgcttc caggccttca ccgtccgcat gatcagcgtc cagatctacg gctacctcat	300

5

10

ES 2 711 165 T3

ccacgagttc gaccctggt tcaactaccg agccgccgag tacatgagca cccacggctg 360
 gtccgccttt ttcagctggt tcgactacat gagctggtat ccgctcggcc gaccctcggg 420
 cagcaccacc taccctggcc tccagctcac cgccgtggcc atccatcgag cctcgcgcgc 480
 tgccggcatg cctatgagcc tcaacaacgt ctgctcctc atgcccgcct ggttcggcgc 540
 cattgcgacc gccaccctcg cgttctgcac ctacgaggcc agcggctcta cagtggccgc 600
 tgccgcggct gccctcagct tcagcatcat ccccgcccac ctcatgcgct ccatggccgg 660
 cgagttcgac aacgagtgca ttgcccgcgc cgccatgctc ctcaccttct actgctgggt 720
 ccgagcctc cgcacgcgca gcagctggcc catcgcgctc ctgaccggcg tcgcctacgg 780
 ctacatggct gccgcctggg gcggtacat cttcgtcctc aacatggtgg ccatgcacgc 840
 cgcatcagc agcatggtcg actgggccc caacacctac aaccctcagc tgctccgcgc 900
 ctacaccctc ttctacgtcg tcggcaccgc cattgcccgc tgcgtcccc ccgtcggcat 960
 gagccccttc aagagcctcg agcagctcgg cgccctcctc gtctgtgtct ttctgtcggg 1020
 cctccaggtc tgcgaggtcc tccgagccc agccggcgtc gaggtccgct ctgcgcctaa 1080
 cttcaagatc cgcgtccgcg tctttagcgt catggccggc gtggccgccc tcgccatctc 1140
 tgtcctcggc cccaccggct acttcggccc cctcagcgtc cgagtgcgcg cctgttctgt 1200
 cgagcacacc cgcaccggca acccctcgt cgacagcgtc gccgagcacc agcccgcag 1260
 ccccgaggcc atgtgggcct ttctccacgt ctgcccgtc acctggggcc tcggcagcat 1320
 cgtcctggcc gtcagcacct tcgtccacta cagcccagc aaggtctttt ggctcctcaa 1380
 ctctggcgcc gtctactact tctcgaccgc aatggcccgc ctctcctcc tgtccggccc 1440
 tgccgcctgc ctgagcaccg gcatcttctg cggcacgac ctcgaggccg ccgtccagct 1500
 cagcttctg gacagcgacg ccaccaaggc caagaagcag cagaagcagg cccagcgcga 1560
 ccagcgaggc gctggcaagg gctctggccc cgacgacgcc aagaacgcga cgaccgcccg 1620
 agccttctgc gacgtctttg ccggcagcag cctcgcctgg ggccaccgca tggctcctc 1680
 gatcgccatg tggcgctcgc tcacgacaac ggccgtcagc ttcttcagca gcgagttcgc 1740
 cagccacagc accaagttcg ccgagcagag cagcaacccc atgatcgtct ttgccgccgt 1800
 cgtccagaac cgcgccaccg gcaagccgat gaacctcctc gtcgacgact acctcaaggc 1860
 ctacgagtgg ctcccgaca gcaccctga ggacgcccgc gtctggcct ggtgggacta 1920
 cggctaccag atcacggca tcggcaaccg caccagcctc gccgacggca acacctggaa 1980
 ccacgagcac attgccacca tcggcaagat gtcaccagc ccggtcgtcg aggccacag 2040
 cctcgtccgc cacatggccg actacgtcct catctgggct ggccagagcg gcgacctcat 2100
 gaagtcccc cacatggccc gcatcggcaa cagcgtctac cagcactct gccccgacga 2160
 cccctctgc cagcagttcg gcttccaccg caacgactac agccgcccc ccccgatgat 2220

ES 2 711 165 T3

gcgcgccagc ctctctaca acctccacga ggccggcaag cgaaagggcg tcaaggtcaa 2280
cccctcgctg ttccaggagg tctacagcag caagtacggc ctggtccgca tcttcaaggt 2340
catgaacgtc agcgccgaga gcaagaagtg ggtcgccgat cccgcccaacc gagtctgcc 2400
ccccctggc agctggatct gccctggcca gtaccctccc gcccaaggaaa tccaggagat 2460
gctcgccac cgcgctcccg tgcaccaggt caccaacgcc gaccgcaaga acaacgtcgg 2520
cagctaccaa gaggagtaca tgcgccgat gcgcgagagc gagaaccgcc gctag 2575

5 <210> 3
<211> 50
<212> ADN
<213> Trichoderma reesei

10 <400> 3
accaagact tttgatcaa tccaacaact tcttcaact taattaaatc 50

15 <210> 4
<211> 48
<212> ADN
<213> Aspergillus niger

20 <400> 4
ttaattaaga tccactaac gttactgaaa tcatcaaca gcttgacg 48

<210> 5
<211> 1000
<212> ADN
<213> Trichoderma reesei

<400> 5
caagtcttcg tactctatcg aagtctcgcc ttacgtactt gatctgctgt ctttcgtgtc 60
cggccaacat atactcgcac acattagccc cagcagaaca tgtcgtcggc ataaaaggcc 120
aattcagatc gcagataaca aaatgctacc agcatctgtc tagttgtgga gatatgaagg 180
ggatattcag gctttccttg tgggaataaa gagagaaaga gagacttaca ggagctctag 240
gcttcgtagc ccctgcgttc ttagttcgca atgccgtgaa agcagctaca tctaccaaga 300
cactcgtgca tcgtctatct tatttggtac atgctgggaa tttccgggac attgtttaag 360
gatgactagg ttcagccggt aaagaatgga aggccatggc ttgtccctct gtggcaagtc 420
attgcactcc aaggccttct cctgtactag toctacaatt ctgcagcaaa tggcctcaag 480
caactacgta aaactccatg agattgcaga tgcggcccaac tggaaataca catcctccgc 540
aagtcggaca tgaagcccct tgacttgatt ggcaggctaa atgcgacatc ttagccggat 600
gcaccccaga tctggggaac gcgccgcttg aggcccgaaag cgccgggttc gatgcattac 660
tgccatattt cagcagttaa ctaggaccgg cttgtgtcga tattgcgggt ggcgttcaat 720
25 ctattccggc actcctatgc cgtttgatcc gatacctgga gggcgtgctt taggcaaaat 780

ES 2 711 165 T3

gccaaagcttc gaggatactg tacgagccgc tttcaacctc acttgatgat gtctgagttt 840
 catcaagaga attgaagtca aagctcaaat catgatgtga agaggttttg aatgtggaag 900
 aattctgcat atataaagcc atggaagaag acgtaaaact gagacagcaa gctcaactgc 960
 atagtatcga cttcaaggaa aacacgcaca aataatcatc 1000

<210> 6
 <211> 1000
 <212> ADN
 <213> Trichoderma reesei

5

<400> 6
 aggggtttga gctggtatgt agtattgggg tggttagtga gttaacttga cagactgcac 60
 tttggcaaca gagccgacga ttaagagatt gctgtcatgt aactaaagta gcctgccttt 120
 gacgctgtat gctcatgata catgctgac atcgaaatat atcagccaaa gtatccgtcc 180
 ggcgacatgc ccatcaacta tattgaagtc agaaacacac tgtccctctt ccctcctatg 240
 cttttacaag ctgctcctct atccgcccc acagtccctt gttcatatac cccgaaagcc 300
 aaaagtttcc atccttgtcc ttgcccata tggggaagcc gtttggtagc acgatacccc 360
 actgattatt ctgtatatag atcggatgaac ccgatttccc accctcccta ctgggctgaa 420
 gcacagctgc agaaaagtcc aagtccaaca gctttgcctt gcccgaattt gacaacgtaa 480
 tcatgtgcat gttgccgttg ccgaagaaag gcggaatcct cccgctagat cctcgccaca 540
 tagcgaaaaa ggcttctacc tgagaccgag ttcccagttc ttgaatcgcg gttcgagtag 600
 cagcagcaat ataactcagc ggcttctcaa atatgtggtg caccggcagt agcagcttga 660
 tgaagccggt accggttgag acatatggca cccctttcgg cagcagatcc gtctctagac 720
 actttcgtag agagtatgcg ttgttgatga caaccgtcct ctggctattc gctggcagat 780
 gtgaagtggc aactttgatc caccagggcg agagaacatc gccttcagtc aagaaagtgt 840
 tttctgogcc ctoggactca agctcactga ttgcctcttt gcgaaggttc tcaatgaaag 900
 atccaggaac acaaagcatg cgattctctt gcgctcggaa gagatcgagg acattgttga 960
 tcccatactg ggccagccca aacattgaca agcgcggaga 1000

10

<210> 7
 <211> 688
 <212> PRT
 <213> Trichomonas vaginalis

15

<400> 7
 Met Gly Asn Thr Val Lys Val Ile Gln Leu Ile Thr Leu Leu Leu Ser
 1 5 10 15
 Cys Leu Leu Ala Phe Leu Ile Arg Gln Phe Ala Asn Val Val Asn Glu

ES 2 711 165 T3

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

ES 2 711 165 T3

Phe Thr Phe Ile Lys Ser Arg Phe Ser Pro Thr Thr Tyr Asn Ser Val
 275 280 285
 Ala Ile Thr Ser Ile Leu Ile Leu Pro Ser Phe Leu Leu Leu Met Ile
 290 295 300
 Thr Val Gly Met Ser Thr Gly Leu Leu Gly Gly Phe Ser Gly Arg Leu
 305 310 315 320
 Leu Gln Met Phe Asp Pro Thr Tyr Ala Ala Lys Asn Val Pro Ile Ile
 325 330 335
 Asn Ser Val Ala Glu His Gln Pro Thr Ala Trp Val Lys Tyr Tyr Ser
 340 345 350
 Asp Cys Glu Leu Phe Ile Phe Phe Phe Pro Leu Gly Ala Tyr Ile Val
 355 360 365
 Ile Ser Ser Leu Ile Arg Thr Gln Lys Thr Lys Asp Gln Thr Glu Leu
 370 375 380
 Lys Arg Ala Glu Thr Leu Leu Leu Leu Phe Ile Tyr Gly Phe Ser Thr
 385 390 395 400
 Leu Tyr Phe Ala Ser Ile Met Val Arg Leu Val Leu Val Phe Thr Pro
 405 410 415
 Ala Leu Val Phe Val Ala Gly Ile Ala Ile His Gln Leu Leu Arg Glu
 420 425 430
 Ser Phe Lys Gln Lys Ser Phe Leu His Pro Val Ser Leu Thr Met Ile
 435 440 445
 Ile Leu Thr Phe Ile Ile Cys Leu His Gly Val Leu His Ala Thr His
 450 455 460
 Phe Ala Cys Tyr Ser Tyr Ser Gly Asp His Leu His Phe Asn Ile Met
 465 470 475 480
 Thr Pro Arg Gly Val Glu Thr Ser Asp Asp Tyr Arg Glu Gly Tyr Arg
 485 490 495
 Trp Leu Thr Glu Asn Thr Tyr Arg Asp Asp Ile Val Met Ser Trp Trp
 500 505 510
 Asp Tyr Gly Tyr Gln Ile Thr Ser Met Gly Asn Arg Gly Cys Ile Ala
 515 520 525

ES 2 711 165 T3

Asp Gly Asn Thr Asn Asn Phe Thr His Ile Gly Ile Ile Gly Met Ala
530 535 540

Met Ser Ser Pro Glu Pro Ile Ser Trp Arg Ile Ala Arg Leu Met Asn
545 550 555 560

Val Lys Tyr Met Leu Val Ile Phe Gly Gly Ala Ala Gln Tyr Ser Gly
565 570 575

Asp Asp Ile Asn Lys Phe Leu Trp Met Pro Arg Ile Ala His Gln Thr
580 585 590

Phe Asp Asn Ile Thr Gly Glu Met Tyr Gln Ile Pro Tyr Arg His Ile
595 600 605

Val Gly Glu Ser Met Thr Lys Asn Met Thr Leu Ser Met Met Phe Lys
610 615 620

Phe Cys Tyr Asn Asn Tyr Lys Tyr Tyr Gln Pro His Pro Gln Phe Pro
625 630 635 640

Thr Gly Tyr Asp Leu Thr Arg Arg Thr Ser Ile Pro Asn Ile Lys Asp
645 650 655

Ile Ser Met Ser Gln Phe Thr Glu Ala Phe Thr Thr Lys Asn Trp Ile
660 665 670

Val Arg Ile Tyr Lys Val Gly Asp Asp Pro Gln Trp Asn Arg Val Tyr
675 680 685

<210> 8
<211> 836
<212> PRT
<213> Leishmania infantum

<400> 8
Met Gly Lys Lys Gly Asn Leu Leu Gly Asp Ser Gly Ser Ala Ala Thr
1 5 10 15

Ala Ser Pro Pro Ala Asn Met Ile Leu Leu Pro Lys Thr Pro Ile Asp
20 25 30

Thr Lys Asp Phe Ile Gly Ile Phe Ser Phe Pro Phe Trp Pro Val Arg
35 40 45

Phe Val Val Thr Val Val Ala Leu Phe Val Val Gly Ala Ser Cys Phe
50 55 60

5

10

ES 2 711 165 T3

Gln Ala Phe Thr Val Arg Met Thr Ser Val Gln Ile Tyr Gly Tyr Leu
65 70 75 80

Ile His Glu Phe Asp Pro Trp Phe Asn Tyr Arg Ala Ala Glu Tyr Met
85 90 95

Ser Thr His Gly Trp Ser Ala Phe Phe Ser Trp Phe Asp Tyr Met Ser
100 105 110

Trp Tyr Pro Leu Gly Arg Pro Val Gly Ser Thr Thr Tyr Pro Gly Leu
115 120 125

Gln Leu Thr Ala Val Ala Ile His Arg Ala Leu Ala Ala Ala Gly Met
130 135 140

Pro Met Ser Leu Asn Asn Val Cys Val Leu Met Pro Ala Trp Phe Gly
145 150 155 160

Ala Ile Ala Thr Ala Thr Leu Ala Phe Cys Thr Tyr Glu Ala Ser Gly
165 170 175

Ser Thr Val Ala Ala Ala Ala Ala Ala Leu Ser Phe Ser Ile Ile Pro
180 185 190

Ala His Leu Met Arg Ser Met Ala Gly Glu Phe Asp Asn Glu Cys Ile
195 200 205

Ala Val Ala Ala Met Leu Leu Thr Phe Tyr Cys Trp Val Arg Ser Leu
210 215 220

Arg Thr Arg Ser Ser Trp Pro Ile Gly Val Leu Thr Gly Val Ala Tyr
225 230 235 240

Gly Tyr Met Val Ala Ala Trp Gly Gly Tyr Ile Phe Val Leu Asn Met
245 250 255

Val Ala Met His Ala Gly Ile Ser Ser Met Val Asp Trp Ala Arg Asn
260 265 270

Thr Tyr Asn Pro Ser Leu Leu Arg Ala Tyr Thr Leu Phe Tyr Val Val
275 280 285

Gly Thr Ala Ile Ala Val Cys Val Pro Pro Val Gly Met Ser Pro Phe
290 295 300

Lys Ser Leu Glu Gln Leu Gly Ala Leu Leu Val Leu Val Phe Leu Cys
305 310 315 320

ES 2 711 165 T3

Gly Leu Gln Ala Cys Glu Val Phe Arg Ala Arg Ala Gly Val Glu Val
 325 330 335
 Arg Ser Arg Ala Asn Phe Lys Ile Arg Val Arg Val Phe Ser Val Met
 340 345 350
 Ala Gly Val Ala Ala Leu Ala Ile Ala Val Leu Ala Pro Thr Gly Tyr
 355 360 365
 Phe Gly Pro Leu Ser Val Arg Val Arg Ala Leu Phe Val Glu His Thr
 370 375 380
 Arg Thr Gly Asn Pro Leu Val Asp Ser Val Ala Glu His Gln Pro Ala
 385 390 395 400
 Gly Pro Glu Ala Met Trp Ser Phe Leu His Val Cys Gly Val Thr Trp
 405 410 415
 Gly Leu Gly Ser Ile Val Leu Ala Leu Ser Thr Phe Val His Tyr Ala
 420 425 430
 Pro Ser Lys Leu Phe Trp Leu Leu Asn Ser Gly Ala Val Tyr Tyr Phe
 435 440 445
 Ser Thr Arg Met Ala Arg Leu Leu Leu Ser Gly Pro Ala Ala Cys
 450 455 460
 Leu Ser Thr Gly Ile Phe Val Gly Thr Ile Leu Glu Ala Ala Val Gln
 465 470 475 480
 Leu Ser Phe Trp Asp Ser Asp Ala Thr Lys Ala Arg Lys Gln Gln Lys
 485 490 495
 Pro Ala Gln Arg His Arg Arg Gly Ala Gly Lys Asp Ser Asp Arg Asp
 500 505 510
 Asp Ala Glu Ser Ala Thr Thr Ala Arg Thr Leu Cys Asp Val Phe Ala
 515 520 525
 Gly Ser Pro Leu Ala Trp Gly His Arg Met Val Leu Phe Ile Ala Val
 530 535 540
 Trp Ala Leu Val Thr Thr Thr Ala Val Ser Phe Phe Ser Ser Asp Phe
 545 550 555 560
 Ala Ser His Ser Thr Thr Phe Ala Glu Gln Ser Ser Asn Pro Met Ile

ES 2 711 165 T3

				565					570					575	
Val	Phe	Ala	Ala	Val	Val	Gln	Asn	Arg	Ala	Thr	Gly	Lys	Pro	Met	Asn
			580					585					590		
Ile	Leu	Val	Asp	Asp	Tyr	Leu	Arg	Ser	Tyr	Ile	Trp	Leu	Arg	Asp	Asn
		595					600					605			
Thr	Pro	Glu	Asp	Ala	Arg	Ile	Leu	Ala	Trp	Trp	Asp	Tyr	Gly	Tyr	Gln
	610					615					620				
Ile	Thr	Gly	Ile	Gly	Asn	Arg	Thr	Ser	Leu	Ala	Asp	Gly	Asn	Thr	Trp
625					630					635					640
Asn	His	Glu	His	Ile	Ala	Thr	Ile	Gly	Lys	Met	Leu	Thr	Ser	Pro	Val
				645					650					655	
Ala	Glu	Ala	His	Ser	Leu	Val	Arg	His	Met	Ala	Asp	Tyr	Val	Leu	Ile
			660					665					670		
Trp	Ala	Gly	Gln	Ser	Gly	Asp	Leu	Met	Lys	Ser	Pro	His	Met	Ala	Arg
		675					680						685		
Ile	Gly	Asn	Ser	Val	Tyr	His	Asp	Ile	Cys	Pro	His	Asp	Pro	Leu	Cys
	690					695					700				
Gln	Gln	Phe	Gly	Phe	Tyr	Arg	Asn	Asp	Tyr	Ser	Arg	Pro	Thr	Pro	Met
705					710					715					720
Met	Arg	Ala	Ser	Leu	Leu	Tyr	Asn	Leu	His	Glu	Val	Gly	Lys	Thr	Lys
				725					730					735	
Gly	Val	Lys	Val	Asp	Pro	Ser	Leu	Phe	Gln	Glu	Val	Tyr	Ser	Ser	Lys
			740					745					750		
Tyr	Gly	Leu	Val	Arg	Val	Phe	Lys	Val	Met	Asn	Val	Ser	Glu	Glu	Ser
		755					760					765			
Lys	Lys	Trp	Val	Ala	Asp	Pro	Ala	Asn	Arg	Val	Cys	His	Pro	Pro	Gly
	770					775					780				
Ser	Trp	Ile	Cys	Pro	Gly	Gln	Tyr	Pro	Pro	Ala	Lys	Glu	Ile	Gln	Glu
785					790					795					800
Met	Leu	Ala	His	Arg	Val	Pro	Phe	Asp	Gln	Val	Glu	Lys	Val	Asp	Arg
			805						810					815	

ES 2 711 165 T3

Lys Asn His Val Gly Ser Tyr His Glu Glu Tyr Met Arg Arg Met Arg
 820 825 830

Glu Ser Glu Ser
 835

<210> 9
 <211> 2511
 <212> ADN
 <213> Leishmania infantum

5

<400> 9
 atgggcaaga agggcaacct cctcggcgat agcggctctg ctgccaccgc cagccccct 60
 gccaacatga tcctgctccc caagaccccc atcgacacca aggacttcat cggcatcttc 120
 agcttccogt tctggcccgt ccgcttcgtc gtcaccgtcg tcgccctctt cgtcgtcggc 180
 gccagctgct tccaggcctt caccgtccgc atgaccagcg tccagatcta oggctacctc 240
 atccacgagt tcgaccctg gttcaactac cgagccgccc agtacctgag caccacggc 300
 tggctccgct ttttcagctg gttcgactat atgagctggg atcccctcgg ccgaccctgc 360
 ggcagcacca cctaccccgg cctccagctc accgctgtcg ccatccaccg agccctcgtc 420
 gcggctggca tgcccattgag cctcaacaac gtctgcgtcc tcatgcccgc ctggttcggc 480
 gccattgcga ccgccaccct cgcgttctgc acctacgagg ccagcggcag cacagtggct 540
 gctgccgctg cggccctcag cttcagcatc atccccgccc acctcatgag cagcatggcc 600
 ggcgagttcg acaacgagtg cattgccgtc gccgccatgc tcctcacctt ctactgctgg 660
 gtccgctccc tccgcaccgg cagcagctgg cccatcggcg tcctcaccgg ggtcgcctac 720
 ggctacatgg tggccgcctg gggcggctac atcttcgtcc tcaacatggt cgccatgcac 780
 gccggcatca gcagcatggt cgactgggcc cgcaacacct acaaccccag cctgctccgc 840
 gcctacaccc tcttctacgt cgtcggcacc gccattgccg tctgcgtccc ccccgtcggc 900
 atgagcccct tcaagagcct cgagcagctc ggagcgtgctc tcgtcctggt ctttctgtgc 960
 ggcctccagg cctgcgaggt ctttcgcgcc cgagccggcg tcgaggtccg cagccgcgcc 1020
 aactcaaga tccgcgtccg cgtgttcagc gtcattggcg gcgtcggcgc cttggtatc 1080
 gccgtcctcg cccccaccgg ctacttcggc ccctcagcg tccgcgtgcg cgccctgttc 1140
 gtcgagcaca cccgcaccgg caatcccctg gtcgacagcg tcgccgagca ccagcctgcc 1200
 ggcctgagg ccatgtggtc gttcctccac gtctcggggc tcacctgggg cctcggatcc 1260
 atcgtcctgg ccctcagcac cttcgtccac tacgccccca gcaagctgtt ctggctcctc 1320
 aactctggcg ccgtctacta cttctcgacc cgaatggccc gcctcctgct cctcagcggc 1380
 cctgccgctt gcctcagcac cggcatcttc gtgggcacca tcctcgaggc cgccgtccag 1440
 ctcagcttct gggacagcga cgccaccaag gcccgcaagc agcagaagcc tgcccagcgc 1500

10

ES 2 711 165 T3

caccgacggg gagccggcaa ggatagcgac cgcgacgacg ccgagtctgc caccaccgcc 1560
 cgcaccctct gcgacgtctt tgccggcagc cccctcgcct ggggccaccg catggtcctc 1620
 ttcattgccg tgtgggccct cgtaacgacg accgccgtca gcttcttcag cagcgacttc 1680
 gccagccaca gcaccacctt cggcgagcag agcagcaacc ccatgatcgt ctttgccgcc 1740
 gtcgtccaga accgcgccac cggcaagccg atgaacatcc tcgtcgacga ctacctcgcg 1800
 agctacatct ggctccgcga caacaccccc gaggacgccc gcatcctcgc ctggtgggac 1860
 tacggctacc agatcaccgg catcggcaac cgcaccagcc tcgccgacgg caacacctgg 1920
 aaccacgagc acattgccac catcggcaag atgctcacca gccccgtcgc cgaggccccac 1980
 agcctcgtcc gccacatggc cgactacgtc ctcatctggg ctggccagag cggcgacctc 2040
 atgaagtccc cccacatggc ccgcatcggc aacagcgtct accacgacat ctgccccac 2100
 gacccccctc gccagcagtt cggcttctac cgcaacgact acagccgccc cccccgatg 2160
 atgcgcgcca gcctcctcta caacctccac gaggtcggca agaccaaggg cgtcaaggtc 2220
 gaccccagcc tcttccaaga ggtctacagc agcaagtaag gcctcgtgcg cgtgttcaag 2280
 gtcatgaacg tcagcgaaga gtccaagaag tgggtcgcgg accccgcaa cagggtctgc 2340
 cccccctg gcagctggat ctgccctggc cagtaccctc ccgccaaga gatccaagag 2400
 atgctcgcgc accgcgtccc gttcgaccag gtcgagaagg tcgaccgcaa gaaccacgtc 2460
 ggctcctacc acgaagagta catgcgccgc atgcgcgaga gcgagagctg a 2511

<210> 10
 <211> 721
 <212> PRT
 <213> Entamoeba histolytica

<400> 10
 Met Gly Phe Phe Lys Thr Leu Val Gln Leu Ile Leu Lys Asn Ile Gly
 1 5 10 15
 Ile Thr Leu Ile Cys Ile Ile Ala Phe Ser Ser Arg Leu Tyr Ser Ile
 20 25 30
 Ile Met Tyr Glu Ala Ile Ile His Glu Phe Asp Pro Tyr Phe Asn Phe
 35 40 45
 Arg Ala Thr Lys Tyr Leu Val Glu His Gly Pro Thr Ala Phe Met Asn
 50 55 60
 Trp Phe Asp Pro Asp Ser Trp Tyr Pro Leu Gly Arg Asn Ile Gly Thr
 65 70 75 80

5

10

ES 2 711 165 T3

Thr Val Phe Pro Gly Leu Met Phe Thr Ser Ala Phe Ile Phe Lys Phe
 85 90 95
 Leu Ala Tyr Phe Asn Leu Ile Ile Asp Val Arg Leu Ile Cys Val Cys
 100 105 110
 Met Gly Pro Ile Tyr Ser Val Ile Thr Cys Ile Val Ala Tyr Leu Phe
 115 120 125
 Gly Ser Arg Val His Ser Asp Arg Ala Gly Leu Phe Ala Ala Ala Leu
 130 135 140
 Ile Ser Val Val Pro Gly Tyr Met Ser Arg Ser Val Ala Gly Ser Tyr
 145 150 155 160
 Asp Tyr Glu Cys Ile Ser Ile Thr Ile Leu Ile Leu Thr Phe Tyr Leu
 165 170 175
 Trp Ile Glu Ala Val His Asn Asn Ser Pro Ile Leu Ser Ala Val Thr
 180 185 190
 Ala Leu Ser Tyr Phe Tyr Met Ala Ser Thr Trp Gly Ala Tyr Val Phe
 195 200 205
 Ile Asn Asn Ile Ile Pro Leu His Val Leu Ile Ser Ile Phe Cys Gly
 210 215 220
 Phe Tyr Asn Lys Lys Leu Tyr Ser Cys Tyr Ser Ile Tyr Tyr Ile Phe
 225 230 235 240
 Ala Thr Ile Leu Ser Met Gln Val Pro Phe Ile Asn Tyr Val Pro Ile
 245 250 255
 Arg Ser Ser Glu His Ile Gly Ala Met Gly Val Phe Gly Ile Cys Gln
 260 265 270
 Leu Ile Glu Leu Tyr Ser Leu Ile His Lys Leu Leu Gly Gln Lys Lys
 275 280 285
 Thr Val Glu Leu Ile Lys Lys Val Leu Met Gly Ser Val Ile Ile Gly
 290 295 300
 Ile Ile Met Val Leu Ile Leu Ile Lys Lys Gly Tyr Ile Ser Ala Trp
 305 310 315 320
 Ser Gly Arg Phe Tyr Ala Leu Phe Asp Pro Thr Phe Ala Lys Lys Asn
 325 330 335

ES 2 711 165 T3

Ile Pro Leu Ile Val Ser Val Ser Glu His Gln Pro Ala Asn Trp Ala
 340 345 350

Ser Tyr Phe Phe Asp Leu His Cys Leu Ile Val Ile Ala Pro Ala Gly
 355 360 365

Leu Tyr Tyr Cys Phe Lys Lys Phe Asp Phe Asn Met Leu Phe Leu Ile
 370 375 380

Ile Tyr Ser Val Ser Val Phe Tyr Phe Ser Cys Val Met Ser Arg Leu
 385 390 395 400

Val Leu Ile Leu Ala Pro Ala Ile Cys Leu Leu Ser Gly Ile Ala Leu
 405 410 415

Ala Glu Phe Phe Thr Gln Ile Gln Lys Gln Leu Glu Ser Thr Leu Lys
 420 425 430

Met Val Phe Lys Ser Asn Lys Lys Gln Gln Gln Gln Gln Ser Asn Glu
 435 440 445

Pro Thr Thr Lys Ile Glu Lys Glu Lys Arg Lys Ile His Pro Pro Lys
 450 455 460

Lys Glu Gln Asn Asn Glu Lys Ser Phe Ile Ser Glu Phe Ile Ile Phe
 465 470 475 480

Ile Ile Met Thr Ile Val Gly Ile Leu Leu Ile Ile Phe Leu Phe Lys
 485 490 495

Phe Phe Glu Tyr Ser Ile Gln Met Ser Lys Asn Tyr Ser Ser Pro Ser
 500 505 510

Val Val Leu Tyr Gly Asn His Gly Gly Lys Gln Ile Ala Phe Asp Asp
 515 520 525

Tyr Arg Glu Ala Tyr Arg Trp Leu Ala His Asn Thr Pro Glu Gly Ser
 530 535 540

Arg Val Met Ser Trp Trp Asp Tyr Gly Tyr Gln Ile Ser His Leu Ala
 545 550 555 560

Asn Arg Thr Val Ile Val Asp Asn Asn Thr Trp Asn Asn Ser His Ile
 565 570 575

Ala Leu Thr Gly Asn Val Met Ala Ser Arg Glu Glu Asp Ala Met Lys
 580 585 590

ES 2 711 165 T3

Thr Ile Arg Asp Leu Asp Val Asp Tyr Leu Leu Val Val Phe Gly Gly
595 600 605

Tyr Leu Gly Tyr Ser Ser Asp Asp Ile Asn Lys Phe Leu Trp Met Ile
610 615 620

Arg Ile Gly Ala Gly Val Asn Pro Ser Leu Asn Glu Asn Asn Tyr Tyr
625 630 635 640

Asn His Asn Ala Tyr Thr Val Ala Asp Pro Ser Asp Thr Phe Lys Tyr
645 650 655

Ser Met Met Tyr Lys Met Cys Tyr His Asn Phe Tyr Lys Ala Ser Asn
660 665 670

Gly Tyr Arg Ala Gly Met Asp Ala Val Arg Arg Glu Val Ile Glu Glu
675 680 685

Gln Thr Tyr Phe Lys Asn Ile Gln Glu Ala Phe Thr Ser Gln His Trp
690 695 700

Val Val Arg Ile Tyr Lys Val Asn Lys Pro Asn Pro Ile Asp Ser Leu
705 710 715 720

Leu

<210> 11

<211> 40

5 <212> ADN

<213> Trichoderma reesei

<400> 11

10 agctccgtgg cgaaagcctg acgcaccggt agattcttgg 40

<210> 12

<211> 1000

<212> ADN

15 <213> Artificial

<220>

<223> Cebador

<400> 12

gttgggctga ggccgtatcg gagggacggg gtgaggattg aggcggagga gatgaagggg 60

gatgatgggg agacggtggt tgttgtgcat aattatgggc atgcgggatg ggggtatcag 120

gggtcgtatg ggtgtgcgga gagggttgtc gagttgtgg aggggattgt gaggggatga 180

20 gcggatgttt ttgatgtttt gactgctcgc ctttgactcg attctgatac ggacactttt 240

ES 2 711 165 T3

cgacctttgt ttctccaaga tggccctgta cagtcagatt gatagaggag catgtataat 300
 tcattgccgg ttgccgtccc gttccaagc agaaagccac tgttgagaag caacgtgctt 360
 tgacgaaagt cgtggctcac tactcaaate tctccacact catacattgt gtttcagtca 420
 aaacactttg gcaaccaaga cgtgggaggg agtatctgca tcttttctca tcggcaagct 480
 atctgactcg attgagaaga tgcgtggttc atatcacctg gccgttgag gtttcttct 540
 aggcagtcgc tctgttctcc ttctataaag aactccatcg ttcttgaata cctctttggc 600
 cttcaagctc gatagtattg aaccattct tcaactatgc tgctcatcat tccacctccc 660
 tcaagttggg tgctggtgag tacctagtgt acataagcgg gtctatgcat ttaaaggggt 720
 atcttcacca ccagcaatat ccacacttct aggcctcacg ttgcacataa cgaaaccaa 780
 acagctaaac cgacgggcca atttcacgag catcttcacg gacgaagcga gcgacagcga 840
 agccgatacg caaatcctct tcagacaagc tcaactcggc caagcctcat gttttgcaa 900
 cggaaacctg cacaagtcgg ctggcattaa agaggaaagg agaacagaaa gagagtgagc 960
 agatttcagt ctctcaccac tcacctgagt tgcctctctc 1000

<210> 13
 <211> 1001
 <212> ADN
 <213> Artificial

5

<220>
 <223> Cebador

10

<400> 13
 gggcagtatg ccggatggct ggcttataca ggcaaaaacc accttcttca ttcttcattc 60
 ttctgtcttct tcttcttctt cctcctcatc gtcggtaggc ggcagctttc ccacattgga 120
 gtcgctctcc tcgctcgtga gttcctcgac cgtcttttctg aattcctttg gcctggagtc 180
 atcataatag ttaatacac gtttagagta tagagagaaa aaataagggg gaaaaagacg 240
 caaatcatac cagtacggct gcttccgcca gagcttctcg tcgcgcacga ccttgataat 300
 ctgcgcaaaag gccctgctgt cctgctgccc gacgggatat ccctcggcca ggacgtagcc 360
 gccgcgcttg atccagacgg tgttgccgag ggcgtgggag agctcagga cgacgtgctt 420
 cttgtggtggc atctcgcagg ttagaggtt gttgccctcg ggcttcagga cgcgcacaaat 480
 ggcctgcgtc ggctcagagg catcgggagg cgtgagggcc tcttgtgtgg cggcaaggac 540
 atttcgcttc ggcttaccba tggctgcgag tctttggggt cgattcgggt atactatctg 600
 atcccaagaa aaaagagaca aaatttcatt gttgttgatt ggaaaataaa ctggggccgt 660
 gatggagggg cagctttatc gataggacgg ggatttctcg aataggaaaa taaaaccctt 720
 ccgccgtcc cgctctccgg cacggtggtt cccattcgg cgaaaccgct tcagggacca 780

ES 2 711 165 T3

	aactagaagt aaggtaccta tccataagct atcacgatga tatagaaggc atggatgtat	840
	tgcaaaagcg aattgtaga cgcccaatg ggaggcttg tggggttatc ggtttacgaa	900
	atacttgaat caatgcatta ttaatctatc cattaggcat tttggcgttc accagaccgt	960
	ttgactcacc gatatcgttc gtggtggtac tgggccagat g	1001
5	<210> 14 <211> 19 <212> ADN <213> Artificial	
10	<220> <223> Cebador	
	<400> 14 gcacacttc aagattggc	19
15	<210> 15 <211> 19 <212> ADN <213> Artificial	
20	<220> <223> Cebador	
	<400> 15 gcacacttc aagattggc	19
25	<210> 16 <211> 19 <212> ADN <213> Artificial	
30	<220> <223> Cebador	
35	<400> 16 gtacggtgt gccagaag	19
40	<210> 17 <211> 448 <212> ADN <213> Artificial	
	<220> <223> Cebador	
45	<400> 17 gttgagtaca tcgagcgcga cagcattgtg cacacatgc ttcccctcga gtccaaggac	60
	agcatcatcg ttgaggactc gtgcaacggc gagacggaga agcaggctcc ctggggtctt	120
	gcccgtatct ctcaccgaga gacgctcaac tttggctcct tcaacaagta cctctacacc	180
	gctgatggtg gtgagggtgt tgatgcctat gtcattgaca ccggcaccaa catcgagcac	240
	gtcgactttg agggctcgtc caagtggggc aagaccatcc ctgccggcga tgaggacgag	300

ES 2 711 165 T3

gacggcaacg gccacggcac tactgtctct ggtaccgttg ctggaagaa gtacggtgtt 360
 gccaaagaagg cccacgtcta cgccgtcaag gtgctccgat ccaacggatc cggcaccatg 420
 tctgacgtcg tcaagggcgt cgagtacg 448

<210> 18
 <211> 399
 <212> PRT
 <213> Trichoderma reesei

5

<400> 18
 Met Gln Pro Ser Phe Gly Ser Phe Leu Val Thr Val Leu Ser Ala Ser
 1 5 10 15
 Met Ala Ala Gly Ser Val Ile Pro Ser Thr Asn Ala Asn Pro Gly Ser
 20 25 30
 Phe Glu Ile Lys Arg Ser Ala Asn Lys Ala Phe Thr Gly Arg Asn Gly
 35 40 45
 Pro Leu Ala Leu Ala Arg Thr Tyr Ala Lys Tyr Gly Val Glu Val Pro
 50 55 60
 Lys Thr Leu Val Asp Ala Ile Gln Leu Val Lys Ser Ile Gln Leu Ala
 65 70 75 80
 Lys Arg Asp Ser Ala Thr Val Thr Ala Thr Pro Asp His Asp Asp Ile
 85 90 95
 Glu Tyr Leu Val Pro Val Lys Ile Gly Thr Pro Pro Gln Thr Leu Asn
 100 105 110
 Leu Asp Phe Asp Thr Gly Ser Ser Asp Leu Trp Val Phe Ser Ser Asp
 115 120 125
 Val Asp Pro Thr Ser Ser Gln Gly His Asp Ile Tyr Thr Pro Ser Lys
 130 135 140
 Ser Thr Ser Ser Lys Lys Leu Glu Gly Ala Ser Trp Asn Ile Thr Tyr
 145 150 155 160
 Gly Asp Arg Ser Ser Ser Ser Gly Asp Val Tyr His Asp Ile Val Ser
 165 170 175
 Val Gly Asn Leu Thr Val Lys Ser Gln Ala Val Glu Ser Ala Arg Asn
 180 185 190

10

ES 2 711 165 T3

Val Ser Ala Gln Phe Thr Gln Gly Asn Asn Asp Gly Leu Val Gly Leu
 195 200 205

Ala Phe Ser Ser Ile Asn Thr Val Lys Pro Thr Pro Gln Lys Thr Trp
 210 215 220

Tyr Asp Asn Ile Val Gly Ser Leu Asp Ser Pro Val Phe Val Ala Asp
 225 230 235 240

Leu Arg His Asp Thr Pro Gly Ser Tyr His Phe Gly Ser Ile Pro Ser
 245 250 255

Glu Ala Ser Lys Ala Phe Tyr Ala Pro Ile Asp Asn Ser Lys Gly Phe
 260 265 270

Trp Gln Phe Ser Thr Ser Ser Asn Ile Ser Gly Gln Phe Asn Ala Val
 275 280 285

Ala Asp Thr Gly Thr Thr Leu Leu Leu Ala Ser Asp Asp Leu Val Lys
 290 295 300

Ala Tyr Tyr Ala Lys Val Gln Gly Ala Arg Val Asn Val Phe Leu Gly
 305 310 315 320

Gly Tyr Val Phe Asn Cys Thr Thr Gln Leu Pro Asp Phe Thr Phe Thr
 325 330 335

Val Gly Glu Gly Asn Ile Thr Val Pro Gly Thr Leu Ile Asn Tyr Ser
 340 345 350

Glu Ala Gly Asn Gly Gln Cys Phe Gly Gly Ile Gln Pro Ser Gly Gly
 355 360 365

Leu Pro Phe Ala Ile Phe Gly Asp Ile Ala Leu Lys Ala Ala Tyr Val
 370 375 380

Ile Phe Asp Ser Gly Asn Lys Gln Val Gly Trp Ala Gln Lys Lys
 385 390 395

<210> 19
 <211> 452
 <212> PRT
 <213> Trichoderma reesei

<400> 19
 Met Glu Ala Ile Leu Gln Ala Gln Ala Lys Phe Arg Leu Asp Arg Gly
 1 5 10 15

5

10

ES 2 711 165 T3

Leu Gln Lys Ile Thr Ala Val Arg Asn Lys Asn Tyr Lys Arg His Gly
 20 25 30

Pro Lys Ser Tyr Val Tyr Leu Leu Asn Arg Phe Gly Phe Glu Pro Thr
 35 40 45

Lys Pro Gly Pro Tyr Phe Gln Gln His Arg Ile His Gln Arg Gly Leu
 50 55 60

Ala His Pro Asp Phe Lys Ala Ala Val Gly Gly Arg Val Thr Arg Gln
 65 70 75 80

Lys Val Leu Ala Lys Lys Val Lys Glu Asp Gly Thr Val Asp Ala Gly
 85 90 95

Gly Ser Lys Thr Gly Glu Val Asp Ala Glu Asp Gln Gln Asn Asp Ser
 100 105 110

Glu Tyr Leu Cys Glu Val Thr Ile Gly Thr Pro Gly Gln Lys Leu Met
 115 120 125

Leu Asp Phe Asp Thr Gly Ser Ser Asp Leu Trp Val Phe Ser Thr Glu
 130 135 140

Leu Ser Lys His Leu Gln Glu Asn His Ala Ile Phe Asp Pro Lys Lys
 145 150 155 160

Ser Ser Thr Phe Lys Pro Leu Lys Asp Gln Thr Trp Gln Ile Ser Tyr
 165 170 175

Gly Asp Gly Ser Ser Ala Ser Gly Thr Cys Gly Ser Asp Thr Val Thr
 180 185 190

Leu Gly Gly Leu Ser Ile Lys Asn Gln Thr Ile Glu Leu Ala Ser Lys
 195 200 205

Leu Ala Pro Gln Phe Ala Gln Gly Thr Gly Asp Gly Leu Leu Gly Leu
 210 215 220

Ala Trp Pro Gln Ile Asn Thr Val Gln Thr Asp Gly Arg Pro Thr Pro
 225 230 235 240

Ala Asn Thr Pro Val Ala Asn Met Ile Gln Gln Asp Asp Ile Pro Ser
 245 250 255

Asp Ala Gln Leu Phe Thr Ala Ala Phe Tyr Ser Glu Arg Asp Glu Asn
 260 265 270

ES 2 711 165 T3

Ala Glu Ser Phe Tyr Thr Phe Gly Tyr Ile Asp Gln Asp Leu Val Ser
275 280 285

Ala Ser Gly Gln Glu Ile Ala Trp Thr Asp Val Asp Asn Ser Gln Gly
290 295 300

Phe Trp Met Phe Pro Ser Thr Lys Thr Thr Ile Asn Gly Lys Asp Ile
305 310 315 320

Ser Gln Glu Gly Asn Thr Ala Ile Ala Asp Thr Gly Thr Thr Leu Ala
325 330 335

Leu Val Ser Asp Glu Val Cys Glu Ala Leu Tyr Lys Ala Ile Pro Gly
340 345 350

Ala Lys Tyr Asp Asp Asn Gln Gln Gly Tyr Val Phe Pro Ile Asn Thr
355 360 365

Asp Ala Ser Ser Leu Pro Glu Leu Lys Val Ser Val Gly Asn Thr Gln
370 375 380

Phe Val Ile Gln Pro Glu Asp Leu Ala Phe Ala Pro Ala Asp Asp Ser
385 390 395 400

Asn Trp Tyr Gly Gly Val Gln Ser Arg Gly Ser Asn Pro Phe Asp Ile
405 410 415

Leu Gly Asp Val Phe Leu Lys Ser Val Tyr Ala Ile Phe Asp Gln Gly
420 425 430

Asn Gln Arg Phe Gly Ala Val Pro Lys Ile Gln Ala Lys Gln Asn Leu
435 440 445

Gln Pro Pro Gln
450

<210> 20

<211> 395

5

<212> PRT

<213> Trichoderma reesei

<400> 20

Met Lys Ser Ala Leu Leu Ala Ala Ala Ala Leu Val Gly Ser Ala Gln
1 5 10 15

Ala Gly Ile His Lys Met Lys Leu Gln Lys Val Ser Leu Glu Gln Gln
20 25 30

10

ES 2 711 165 T3

Leu Glu Gly Ser Ser Ile Glu Ala His Val Gln Gln Leu Gly Gln Lys
 35 40 45
 Tyr Met Gly Val Arg Pro Thr Ser Arg Ala Glu Val Met Phe Asn Asp
 50 55 60
 Lys Pro Pro Lys Val Gln Gly Gly His Pro Val Pro Val Thr Asn Phe
 65 70 75 80
 Met Asn Ala Gln Tyr Phe Ser Glu Ile Thr Ile Gly Thr Pro Pro Gln
 85 90 95
 Ser Phe Lys Val Val Leu Asp Thr Gly Ser Ser Asn Leu Trp Val Pro
 100 105 110
 Ser Gln Ser Cys Asn Ser Ile Ala Cys Phe Leu His Ser Thr Tyr Asp
 115 120 125
 Ser Ser Ser Ser Ser Thr Tyr Lys Pro Asn Gly Ser Asp Phe Glu Ile
 130 135 140
 His Tyr Gly Ser Gly Ser Leu Thr Gly Phe Ile Ser Asn Asp Val Val
 145 150 155 160
 Thr Ile Gly Asp Leu Lys Ile Lys Gly Gln Asp Phe Ala Glu Ala Thr
 165 170 175
 Ser Glu Pro Gly Leu Ala Phe Ala Phe Gly Arg Phe Asp Gly Ile Leu
 180 185 190
 Gly Leu Gly Tyr Asp Thr Ile Ser Val Asn Gly Ile Val Pro Pro Phe
 195 200 205
 Tyr Gln Met Val Asn Gln Lys Leu Ile Asp Glu Pro Val Phe Ala Phe
 210 215 220
 Tyr Leu Gly Ser Ser Asp Glu Gly Ser Glu Ala Val Phe Gly Gly Val
 225 230 235 240
 Asp Asp Ala His Tyr Glu Gly Lys Ile Glu Tyr Ile Pro Leu Arg Arg
 245 250 255
 Lys Ala Tyr Trp Glu Val Asp Leu Asp Ser Ile Ala Phe Gly Asp Glu
 260 265 270
 Val Ala Glu Leu Glu Asn Thr Gly Ala Ile Leu Asp Thr Gly Thr Ser
 275 280 285

ES 2 711 165 T3

Leu Asn Val Leu Pro Ser Gly Leu Ala Glu Leu Leu Asn Ala Glu Ile
 290 295 300

Gly Ala Lys Lys Gly Phe Gly Gly Gln Tyr Thr Val Asp Cys Ser Lys
 305 310 315 320

Arg Asp Ser Leu Pro Asp Ile Thr Phe Ser Leu Ala Gly Ser Lys Tyr
 325 330 335

Ser Leu Pro Ala Ser Asp Tyr Ile Ile Glu Met Ser Gly Asn Cys Ile
 340 345 350

Ser Ser Phe Gln Gly Met Asp Phe Pro Glu Pro Val Gly Pro Leu Val
 355 360 365

Ile Leu Gly Asp Ala Phe Leu Arg Arg Tyr Tyr Ser Val Tyr Asp Leu
 370 375 380

Gly Arg Asp Ala Val Gly Leu Ala Lys Ala Lys
 385 390 395

<210> 21

<211> 426

<212> PRT

<213> Trichoderma reesei

<400> 21

Met Lys Phe His Ala Ala Ala Leu Thr Leu Ala Cys Leu Ala Ser Ser
 1 5 10 15

Ala Ser Ala Gly Val Ala Gln Pro Arg Ala Asp Glu Val Glu Ser Ala
 20 25 30

Glu Gln Gly Lys Thr Phe Ser Leu Glu Gln Ile Pro Asn Glu Arg Tyr
 35 40 45

Lys Gly Asn Ile Pro Ala Ala Tyr Ile Ser Ala Leu Ala Lys Tyr Ser
 50 55 60

Pro Thr Ile Pro Asp Lys Ile Lys His Ala Ile Glu Ile Asn Pro Asp
 65 70 75 80

Leu His Arg Lys Phe Ser Lys Leu Ile Asn Ala Gly Asn Met Thr Gly
 85 90 95

Thr Ala Val Ala Ser Pro Pro Pro Gly Ala Asp Ala Glu Tyr Val Leu
 100 105 110

5

10

ES 2 711 165 T3

Pro Val Lys Ile Gly Thr Pro Pro Gln Thr Leu Pro Leu Asn Leu Asp
 115 120 125
 Thr Gly Ser Ser Asp Leu Trp Val Ile Ser Thr Asp Thr Tyr Pro Pro
 130 135 140
 Gln Val Gln Gly Gln Thr Arg Tyr Asn Val Ser Ala Ser Thr Thr Ala
 145 150 155 160
 Gln Arg Leu Ile Gly Glu Ser Trp Val Ile Arg Tyr Gly Asp Gly Ser
 165 170 175
 Ser Ala Asn Gly Ile Val Tyr Lys Asp Arg Val Gln Ile Gly Asn Thr
 180 185 190
 Phe Phe Asn Gln Gln Ala Val Glu Ser Ala Val Asn Ile Ser Asn Glu
 195 200 205
 Ile Ser Asp Asp Ser Phe Ser Ser Gly Leu Leu Gly Ala Ala Ser Ser
 210 215 220
 Ala Ala Asn Thr Val Arg Pro Asp Arg Gln Thr Thr Tyr Leu Glu Asn
 225 230 235 240
 Ile Lys Ser Gln Leu Ala Arg Pro Val Phe Thr Ala Asn Leu Lys Lys
 245 250 255
 Gly Lys Pro Gly Asn Tyr Asn Phe Gly Tyr Ile Asn Gly Ser Glu Tyr
 260 265 270
 Ile Gly Pro Ile Gln Tyr Ala Ala Ile Asn Pro Ser Ser Pro Leu Trp
 275 280 285
 Glu Val Ser Val Ser Gly Tyr Arg Val Gly Ser Asn Asp Thr Lys Tyr
 290 295 300
 Val Pro Arg Val Trp Asn Ala Ile Ala Asp Thr Gly Thr Thr Leu Leu
 305 310 315 320
 Leu Val Pro Asn Asp Ile Val Ser Ala Tyr Tyr Ala Gln Val Lys Gly
 325 330 335
 Ser Thr Phe Ser Asn Asp Val Gly Met Met Leu Val Pro Cys Ala Ala
 340 345 350
 Thr Leu Pro Asp Phe Ala Phe Gly Leu Gly Asn Tyr Arg Gly Val Ile

ES 2 711 165 T3

355 360 365
 Pro Gly Ser Tyr Ile Asn Tyr Gly Arg Met Asn Lys Thr Tyr Cys Tyr
 370 375 380
 Gly Gly Ile Gln Ser Ser Glu Asp Ala Pro Phe Ala Val Leu Gly Asp
 385 390 395
 Ile Ala Leu Lys Ala Gln Phe Val Val Phe Asp Met Gly Asn Lys Val
 405 410 415
 Val Gly Phe Ala Asn Lys Asn Thr Asn Val
 420 425
 <210> 22
 <211> 407
 <212> PRT
 <213> Trichoderma reesei
 <400> 22
 Met Gln Thr Phe Gly Ala Phe Leu Val Ser Phe Leu Ala Ala Ser Gly
 1 5 10 15
 Leu Ala Ala Ala Leu Pro Thr Glu Gly Gln Lys Thr Ala Ser Val Glu
 20 25 30
 Val Gln Tyr Asn Lys Asn Tyr Val Pro His Gly Pro Thr Ala Leu Phe
 35 40 45
 Lys Ala Lys Arg Lys Tyr Gly Ala Pro Ile Ser Asp Asn Leu Lys Ser
 50 55 60
 Leu Val Ala Ala Arg Gln Ala Lys Gln Ala Leu Ala Lys Arg Gln Thr
 65 70 75 80
 Gly Ser Ala Pro Asn His Pro Ser Asp Ser Ala Asp Ser Glu Tyr Ile
 85 90 95
 Thr Ser Val Ser Ile Gly Thr Pro Ala Gln Val Leu Pro Leu Asp Phe
 100 105 110
 Asp Thr Gly Ser Ser Asp Leu Trp Val Phe Ser Ser Glu Thr Pro Lys
 115 120 125
 Ser Ser Ala Thr Gly His Ala Ile Tyr Thr Pro Ser Lys Ser Ser Thr
 130 135 140
 Ser Lys Lys Val Ser Gly Ala Ser Trp Ser Ile Ser Tyr Gly Asp Gly

5

10

ES 2 711 165 T3

145	150	155	160
Ser Ser Ser Ser	Gly Asp Val Tyr Thr Asp	Lys Val Thr Ile	Gly Gly
	165	170	175
Phe Ser Val	Asn Thr Gln Gly Val	Glu Ser Ala Thr Arg	Val Ser Thr
	180	185	190
Glu Phe Val	Gln Asp Thr Val	Ile Ser Gly Leu Val	Gly Leu Ala Phe
	195	200	205
Asp Ser Gly	Asn Gln Val Arg Pro His Pro	Gln Lys Thr Trp Phe Ser	
	210	215	220
Asn Ala Ala	Ser Ser Leu Ala Glu Pro Leu	Phe Thr Ala Asp Leu Arg	
	225	230	235
His Gly Gln	Asn Gly Ser Tyr Asn Phe	Gly Tyr Ile Asp Thr Ser Val	
	245	250	255
Ala Lys Gly	Pro Val Ala Tyr Thr Pro	Val Asp Asn Ser Gln Gly Phe	
	260	265	270
Trp Glu Phe	Thr Ala Ser Gly Tyr Ser Val	Gly Gly Gly Lys Leu Asn	
	275	280	285
Arg Asn Ser	Ile Asp Gly Ile Ala Asp Thr	Gly Thr Thr Leu Leu Leu	
	290	295	300
Leu Asp Asp	Asn Val Val Asp Ala Tyr Tyr	Ala Asn Val Gln Ser Ala	
	305	310	315
Gln Tyr Asp	Asn Gln Gln Glu Gly Val	Val Phe Asp Cys Asp Glu Asp	
	325	330	335
Leu Pro Ser	Phe Ser Phe Gly Val Gly Ser Ser	Thr Ile Thr Ile Pro	
	340	345	350
Gly Asp Leu	Leu Asn Leu Thr Pro Leu Glu Glu	Gly Ser Ser Thr Cys	
	355	360	365
Phe Gly Gly	Leu Gln Ser Ser Ser Gly Ile Gly	Ile Asn Ile Phe Gly	
	370	375	380
Asp Val Ala	Leu Lys Ala Ala Leu Val Val	Phe Asp Leu Gly Asn Glu	
	385	390	395
Arg Leu Gly	Trp Ala Gln Lys		
	405		

ES 2 711 165 T3

<210> 23
 <211> 446
 <212> PRT
 <213> Trichoderma reesei

5

<400> 23
 Met Thr Leu Pro Val Pro Leu Arg Glu His Asp Leu Pro Phe Leu Lys
 1 5 10 15

 Glu Lys Arg Lys Leu Pro Ala Asp Asp Ile Pro Ser Gly Thr Tyr Thr
 20 25 30

 Leu Pro Ile Ile His Ala Arg Arg Pro Lys Leu Ala Ser Arg Ala Ile
 35 40 45

 Glu Val Gln Val Glu Asn Arg Ser Asp Val Ser Tyr Tyr Ala Gln Leu
 50 55 60

 Asn Ile Gly Thr Pro Pro Gln Thr Val Tyr Ala Gln Ile Asp Thr Gly
 65 70 75 80

 Ser Phe Glu Leu Trp Val Asn Pro Asn Cys Ser Asn Val Gln Ser Ala
 85 90 95

 Asp Gln Arg Phe Cys Arg Ala Ile Gly Phe Tyr Asp Pro Ser Ser Ser
 100 105 110

 Ser Thr Ala Asp Val Thr Ser Gln Ser Ala Arg Leu Arg Tyr Gly Ile
 115 120 125

 Gly Ser Ala Asp Val Thr Tyr Val His Asp Thr Ile Ser Leu Pro Gly
 130 135 140

 Ser Gly Ser Gly Ser Lys Ala Met Lys Ala Val Gln Phe Gly Val Ala
 145 150 155 160

 Asp Thr Ser Val Asp Glu Phe Ser Gly Ile Leu Gly Leu Gly Ala Gly
 165 170 175

 Asn Gly Ile Asn Thr Glu Tyr Pro Asn Phe Val Asp Glu Leu Ala Ala
 180 185 190

 Gln Gly Val Thr Ala Thr Lys Ala Phe Ser Leu Ala Leu Gly Ser Lys
 195 200 205

ES 2 711 165 T3

Ala Glu Glu Glu Gly Val Ile Ile Phe Gly Gly Val Asp Thr Ala Lys
 210 215 220

Phe His Gly Glu Leu Ala His Leu Pro Ile Val Pro Ala Asp Asp Ser
 225 230 235 240

Pro Asp Gly Val Ala Arg Tyr Trp Val Lys Met Lys Ser Ile Ser Leu
 245 250 255

Thr Pro Pro Pro Pro Ser Ser Ser Gly Ser Thr Asp Asp Asn Asn Asn
 260 265 270

Lys Pro Val Ala Phe Pro Gln Thr Ser Met Thr Val Phe Leu Asp Ser
 275 280 285

Gly Ser Thr Leu Thr Leu Leu Pro Pro Ala Leu Val Arg Gln Ile Ala
 290 295 300

Ser Ala Leu Gly Ser Thr Gln Thr Asp Glu Ser Gly Phe Phe Val Val
 305 310 315 320

Asp Cys Ala Leu Ala Ser Gln Asp Gly Thr Ile Asp Phe Glu Phe Asp
 325 330 335

Gly Val Thr Ile Arg Val Pro Tyr Ala Glu Met Ile Arg Gln Val Ser
 340 345 350

Thr Leu Pro Pro His Cys Tyr Leu Gly Met Met Gly Ser Thr Gln Phe
 355 360 365

Ala Leu Leu Gly Asp Thr Phe Leu Arg Ser Ala Tyr Ala Val Phe Asp
 370 375 380

Leu Thr Ser Asn Val Val His Leu Ala Pro Tyr Ala Asn Cys Gly Thr
 385 390 395 400

Asn Val Lys Ser Ile Thr Ser Thr Ser Ser Leu Ser Asn Leu Val Gly
 405 410 415

Thr Cys Asn Asp Pro Ser Lys Pro Ser Ser Ser Pro Ser Pro Ser Gln
 420 425 430

Thr Pro Ser Ala Ser Pro Ser Ser Thr Ala Thr Gln Lys Ala
 435 440 445

<210> 24
 <211> 259
 <212> PRT
 <213> Trichoderma reesei

ES 2 711 165 T3

<400> 24

Met Ala Pro Ala Ser Gln Val Val Ser Ala Leu Met Leu Pro Ala Leu
 1 5 10 15

Ala Leu Gly Ala Ala Ile Gln Pro Arg Gly Ala Asp Ile Val Gly Gly
 20 25 30

Thr Ala Ala Ser Leu Gly Glu Phe Pro Tyr Ile Val Ser Leu Gln Asn
 35 40 45

Pro Asn Gln Gly Gly His Phe Cys Gly Gly Val Leu Val Asn Ala Asn
 50 55 60

Thr Val Val Thr Ala Ala His Cys Ser Val Val Tyr Pro Ala Ser Gln
 65 70 75 80

Ile Arg Val Arg Ala Gly Thr Leu Thr Trp Asn Ser Gly Gly Thr Leu
 85 90 95

Val Gly Val Ser Gln Ile Ile Val Asn Pro Ser Tyr Asn Asp Arg Thr
 100 105 110

Thr Asp Phe Asp Val Ala Val Trp His Leu Ser Ser Pro Ile Arg Glu
 115 120 125

Ser Ser Thr Ile Gly Tyr Ala Thr Leu Pro Ala Gln Gly Ser Asp Pro
 130 135 140

Val Ala Gly Ser Thr Val Thr Thr Ala Gly Trp Gly Thr Thr Ser Glu
 145 150 155 160

Asn Ser Asn Ser Ile Pro Ser Arg Leu Asn Lys Val Ser Val Pro Val
 165 170 175

Val Ala Arg Ser Thr Cys Gln Ala Asp Tyr Arg Ser Gln Gly Leu Ser
 180 185 190

Val Thr Asn Asn Met Phe Cys Ala Gly Leu Thr Gln Gly Gly Lys Asp
 195 200 205

Ser Cys Ser Gly Asp Ser Gly Gly Pro Ile Val Asp Ala Asn Gly Val
 210 215 220

Leu Gln Gly Val Val Ser Trp Gly Ile Gly Cys Ala Glu Ala Gly Phe
 225 230 235 240

Pro Gly Val Tyr Thr Arg Ile Gly Asn Phe Val Asn Tyr Ile Asn Gln
 245 250 255

Asn Leu Ala

ES 2 711 165 T3

5

<210> 25
 <211> 882
 <212> PRT
 <213> Trichoderma reesei

<400> 25
 Met Val Arg Ser Ala Leu Phe Val Ser Leu Leu Ala Thr Phe Ser Gly
 1 5 10 15
 Val Ile Ala Arg Val Ser Gly His Gly Ser Lys Ile Val Pro Gly Ala
 20 25 30
 Tyr Ile Phe Glu Phe Glu Asp Ser Gln Asp Thr Ala Asp Phe Tyr Lys
 35 40 45
 Lys Leu Asn Gly Glu Gly Ser Thr Arg Leu Lys Phe Asp Tyr Lys Leu
 50 55 60
 Phe Lys Gly Val Ser Val Gln Leu Lys Asp Leu Asp Asn His Glu Ala
 65 70 75 80
 Lys Ala Gln Gln Met Ala Gln Leu Pro Ala Val Lys Asn Val Trp Pro
 85 90 95
 Val Thr Leu Ile Asp Ala Pro Asn Pro Lys Val Glu Trp Val Ala Gly
 100 105 110
 Ser Thr Ala Pro Thr Leu Glu Ser Arg Ala Ile Lys Lys Pro Pro Ile
 115 120 125
 Pro Asn Asp Ser Ser Asp Phe Pro Thr His Gln Met Thr Gln Ile Asp
 130 135 140
 Lys Leu Arg Ala Lys Gly Tyr Thr Gly Lys Gly Val Arg Val Ala Val
 145 150 155 160
 Ile Asp Thr Gly Ile Asp Tyr Thr His Pro Ala Leu Gly Gly Cys Phe
 165 170 175
 Gly Arg Gly Cys Leu Val Ser Phe Gly Thr Asp Leu Val Gly Asp Asp
 180 185 190

ES 2 711 165 T3

Tyr Thr Gly Phe Asn Thr Pro Val Pro Asp Asp Asp Pro Val Asp Cys
 195 200 205

Ala Gly His Gly Ser His Val Ala Gly Ile Ile Ala Ala Gln Glu Asn
 210 215 220

Pro Tyr Gly Phe Thr Gly Gly Ala Pro Asp Val Thr Leu Gly Ala Tyr
 225 230 235 240

Arg Val Phe Gly Cys Asp Gly Gln Ala Gly Asn Asp Val Leu Ile Ser
 245 250 255

Ala Tyr Asn Gln Ala Phe Glu Asp Gly Ala Gln Ile Ile Thr Ala Ser
 260 265 270

Ile Gly Gly Pro Ser Gly Trp Ala Glu Glu Pro Trp Ala Val Ala Val
 275 280 285

Thr Arg Ile Val Glu Ala Gly Val Pro Cys Thr Val Ser Ala Gly Asn
 290 295 300

Glu Gly Asp Ser Gly Leu Phe Phe Ala Ser Thr Ala Ala Asn Gly Lys
 305 310 315 320

Lys Val Ile Ala Val Ala Ser Val Asp Asn Glu Asn Ile Pro Ser Val
 325 330 335

Leu Ser Val Ala Ser Tyr Lys Ile Asp Ser Gly Ala Ala Gln Asp Phe
 340 345 350

Gly Tyr Val Ser Ser Ser Lys Ala Trp Asp Gly Val Ser Lys Pro Leu
 355 360 365

Tyr Ala Val Ser Phe Asp Thr Thr Ile Pro Asp Asp Gly Cys Ser Pro
 370 375 380

Leu Pro Asp Ser Thr Pro Asp Leu Ser Asp Tyr Ile Val Leu Val Arg
 385 390 395 400

Arg Gly Thr Cys Thr Phe Val Gln Lys Ala Gln Asn Val Ala Ala Lys
 405 410 415

Gly Ala Lys Tyr Leu Leu Tyr Tyr Asn Asn Ile Pro Gly Ala Leu Ala
 420 425 430

Val Asp Val Ser Ala Val Pro Glu Ile Glu Ala Val Gly Met Val Asp
 435 440 445

ES 2 711 165 T3

Asp Lys Thr Gly Ala Thr Trp Ile Ala Ala Leu Lys Asp Gly Lys Thr
 450 455 460

Val Thr Leu Thr Leu Thr Asp Pro Ile Glu Ser Glu Lys Gln Ile Gln
 465 470 475 480

Phe Ser Asp Asn Pro Thr Thr Gly Gly Ala Leu Ser Gly Tyr Thr Thr
 485 490 495

Trp Gly Pro Thr Trp Glu Leu Asp Val Lys Pro Gln Ile Ser Ser Pro
 500 505 510

Gly Gly Asn Ile Leu Ser Thr Tyr Pro Val Ala Leu Gly Gly Tyr Ala
 515 520 525

Thr Leu Ser Gly Thr Ser Met Ala Cys Pro Leu Thr Ala Ala Ala Val
 530 535 540

Ala Leu Ile Gly Gln Ala Arg Gly Thr Phe Asp Pro Ala Leu Ile Asp
 545 550 555 560

Asn Leu Leu Ala Thr Thr Ala Asn Pro Gln Leu Phe Asn Asp Gly Glu
 565 570 575

Lys Phe Tyr Asp Phe Leu Ala Pro Val Pro Gln Gln Gly Gly Gly Leu
 580 585 590

Ile Gln Ala Tyr Asp Ala Ala Phe Ala Thr Thr Leu Leu Ser Pro Ser
 595 600 605

Ser Leu Ser Phe Asn Asp Thr Asp His Phe Ile Lys Lys Lys Gln Ile
 610 615 620

Thr Leu Lys Asn Thr Ser Lys Gln Arg Val Thr Tyr Lys Leu Asn His
 625 630 635 640

Val Pro Thr Asn Thr Phe Tyr Thr Leu Ala Pro Gly Asn Gly Tyr Pro
 645 650 655

Ala Pro Phe Pro Asn Asp Ala Val Ala Ala His Ala Asn Leu Lys Phe
 660 665 670

Asn Leu Gln Gln Val Thr Leu Pro Ala Gly Arg Ser Ile Thr Val Asp
 675 680 685

Val Phe Pro Thr Pro Pro Arg Asp Val Asp Ala Lys Arg Leu Ala Leu

ES 2 711 165 T3

690

695

700

Trp Ser Gly Tyr Ile Thr Val Asn Gly Thr Asp Gly Thr Ser Leu Ser
705 710 715 720

Val Pro Tyr Gln Gly Leu Thr Gly Ser Leu His Lys Gln Lys Val Leu
725 730 735

Tyr Pro Glu Asp Ser Trp Ile Ala Asp Ser Thr Asp Glu Ser Leu Ala
740 745 750

Pro Val Glu Asn Gly Thr Val Phe Thr Ile Pro Ala Pro Gly Asn Ala
755 760 765

Gly Pro Asp Asp Lys Leu Pro Ser Leu Val Val Ser Pro Ala Leu Gly
770 775 780

Ser Arg Tyr Val Arg Val Asp Leu Val Leu Leu Ser Ala Pro Pro His
785 790 795 800

Gly Thr Lys Leu Lys Thr Val Lys Phe Leu Asp Thr Thr Ser Ile Gly
805 810 815

Gln Pro Ala Gly Ser Pro Leu Leu Trp Ile Ser Arg Gly Ala Asn Pro
820 825 830

Ile Ala Trp Thr Gly Glu Leu Ser Asp Asn Lys Phe Ala Pro Pro Gly
835 840 845

Thr Tyr Lys Ala Val Phe His Ala Leu Arg Ile Phe Gly Asn Glu Lys
850 855 860

Lys Lys Glu Asp Trp Asp Val Ser Glu Ser Pro Ala Phe Thr Ile Lys
865 870 875 880

Tyr Ala

<210> 26

<211> 541

5

<212> PRT

<213> Trichoderma reesei

<400> 26

Met Arg Ser Val Val Ala Leu Ser Met Ala Ala Val Ala Gln Ala Ser
1 5 10 15

10

Thr Phe Gln Ile Gly Thr Ile His Glu Lys Ser Ala Pro Val Leu Ser

ES 2 711 165 T3

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

ES 2 711 165 T3

Ala Ala Leu Ser His Ile Glu Gln Val Lys Lys Ala Lys Lys Gly Lys
 275 280 285

Arg Lys Gly Phe Lys Gly Ser Val Ala Asn Met Ser Leu Gly Gly Gly
 290 295 300

Lys Thr Gln Ala Leu Asp Ala Ala Val Asn Ala Ala Val Arg Ala Gly
 305 310 315 320

Val His Phe Ala Val Ala Ala Gly Asn Asp Asn Ala Asp Ala Cys Asn
 325 330 335

Tyr Ser Pro Ala Ala Ala Thr Glu Pro Leu Thr Val Gly Ala Ser Ala
 340 345 350

Leu Asp Asp Ser Arg Ala Tyr Phe Ser Asn Tyr Gly Lys Cys Thr Asp
 355 360 365

Ile Phe Ala Pro Gly Leu Ser Ile Gln Ser Thr Trp Ile Gly Ser Lys
 370 375 380

Tyr Ala Val Asn Thr Ile Ser Gly Thr Ser Met Ala Ser Pro His Ile
 385 390 395 400

Cys Gly Leu Leu Ala Tyr Tyr Leu Ser Leu Gln Pro Ala Gly Asp Ser
 405 410 415

Glu Phe Ala Val Ala Pro Ile Thr Pro Lys Lys Leu Lys Glu Ser Val
 420 425 430

Ile Ser Val Ala Thr Lys Asn Ala Leu Ser Asp Leu Pro Asp Ser Asp
 435 440 445

Thr Pro Asn Leu Leu Ala Trp Asn Gly Gly Gly Cys Ser Asn Phe Ser
 450 455 460

Gln Ile Val Glu Ala Gly Ser Tyr Thr Val Lys Pro Lys Gln Asn Lys
 465 470 475 480

Gln Ala Lys Leu Pro Ser Thr Ile Glu Glu Leu Glu Glu Ala Ile Glu
 485 490 495

Gly Asp Phe Glu Val Val Ser Gly Glu Ile Val Lys Gly Ala Lys Ser
 500 505 510

Phe Gly Ser Lys Ala Glu Lys Phe Ala Lys Lys Ile His Asp Leu Val
 515 520 525

Glu Glu Glu Ile Glu Glu Phe Ile Ser Glu Leu Ser Glu
 530 535 540

ES 2 711 165 T3

5

<210> 27
 <211> 391
 <212> PRT
 <213> Trichoderma reesei

<400> 27
 Met Arg Leu Ser Val Leu Leu Ser Val Leu Pro Leu Val Leu Ala Ala
 1 5 10 15

 Pro Ala Ile Glu Lys Arg Ala Glu Pro Ala Pro Leu Leu Val Pro Thr
 20 25 30

 Thr Lys His Gly Leu Val Ala Asp Lys Tyr Ile Val Lys Phe Lys Asp
 35 40 45

 Gly Ser Ser Leu Gln Ala Val Asp Glu Ala Ile Ser Gly Leu Val Ser
 50 55 60

 Asn Ala Asp His Val Tyr Gln His Val Phe Arg Gly Phe Ala Ala Thr
 65 70 75 80

 Leu Asp Lys Glu Thr Leu Glu Ala Leu Arg Asn His Pro Glu Val Asp
 85 90 95

 Tyr Ile Glu Gln Asp Ala Val Val Lys Ile Asn Ala Tyr Val Ser Gln
 100 105 110

 Thr Gly Ala Pro Trp Gly Leu Gly Arg Ile Ser His Lys Ala Arg Gly
 115 120 125

 Ser Thr Thr Tyr Val Tyr Asp Asp Ser Ala Gly Ala Gly Thr Cys Ser
 130 135 140

 Tyr Val Ile Asp Thr Gly Val Asp Ala Thr His Pro Asp Phe Glu Gly
 145 150 155 160

 Arg Ala Thr Leu Leu Arg Ser Phe Val Ser Gly Gln Asn Thr Asp Gly
 165 170 175

 Asn Gly His Gly Thr His Val Ser Gly Thr Ile Gly Ser Arg Thr Tyr
 180 185 190

 Gly Val Ala Lys Lys Thr Gln Ile Tyr Gly Val Lys Val Leu Asp Asn
 195 200 205

ES 2 711 165 T3

Ser Gly Ser Gly Ser Phe Ser Thr Val Ile Ala Gly Met Asp Tyr Val
 210 215 220

Ala Ser Asp Ser Gln Thr Arg Asn Cys Pro Asn Gly Ser Val Ala Asn
 225 230 235 240

Met Ser Leu Gly Gly Gly Tyr Thr Ala Ser Val Asn Gln Ala Ala Ala
 245 250 255

Arg Leu Ile Gln Ala Gly Val Phe Leu Ala Val Ala Ala Gly Asn Asp
 260 265 270

Gly Val Asp Ala Arg Asn Thr Ser Pro Ala Ser Glu Pro Thr Val Cys
 275 280 285

Thr Val Gly Ala Ser Thr Ser Ser Asp Ala Arg Ala Ser Phe Ser Asn
 290 295 300

Tyr Gly Ser Val Val Asp Ile Phe Ala Pro Gly Gln Asp Ile Leu Ser
 305 310 315 320

Thr Trp Pro Asn Arg Gln Thr Asn Thr Ile Ser Gly Thr Ser Met Ala
 325 330 335

Thr Pro His Ile Val Gly Leu Gly Ala Tyr Leu Ala Gly Leu Glu Gly
 340 345 350

Phe Ser Asp Pro Gln Ala Leu Cys Ala Arg Ile Gln Ser Leu Ala Asn
 355 360 365

Arg Asn Leu Leu Ser Gly Ile Pro Ser Gly Thr Ile Asn Ala Ile Ala
 370 375 380

Phe Asn Gly Asn Pro Ser Gly
 385 390

<210> 28

<211> 387

<212> PRT

<213> Trichoderma reesei

<400> 28

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

Tyr Ile Val Val Tyr Asn Asn Ser Phe Gly Glu Glu Ala Ile Ser Ala
 20 25 30

5

10

ES 2 711 165 T3

Lys Gln Ala Gln Phe Ala Ala Lys Ile Ala Lys Arg Asn Leu Gly Lys
 35 40 45
 Arg Gly Leu Phe Gly Asn Glu Leu Ser Thr Ala Ile His Ser Phe Ser
 50 55 60
 Met His Thr Trp Arg Ala Met Ala Leu Asp Ala Asp Asp Ile Met Ile
 65 70 75 80
 Lys Asp Ile Phe Asp Ala Glu Glu Val Ala Tyr Ile Glu Ala Asp Thr
 85 90 95
 Lys Val Gln His Ala Ala Leu Val Ala Gln Thr Asn Ala Ala Pro Gly
 100 105 110
 Leu Ile Arg Leu Ser Asn Lys Ala Val Gly Gly Gln Asn Tyr Ile Phe
 115 120 125
 Asp Asn Ser Ala Gly Ser Asn Ile Thr Ala Tyr Val Val Asp Thr Gly
 130 135 140
 Ile Arg Ile Thr His Ser Glu Phe Glu Gly Arg Ala Thr Phe Gly Ala
 145 150 155 160
 Asn Phe Val Asn Asp Asp Thr Asp Glu Asn Gly His Gly Ser His Val
 165 170 175
 Ala Gly Thr Ile Gly Gly Ala Thr Phe Gly Val Ala Lys Asn Val Glu
 180 185 190
 Leu Val Ala Val Lys Val Leu Asp Ala Asp Gly Ser Gly Ser Asn Ser
 195 200 205
 Gly Val Leu Asn Gly Met Gln Phe Val Val Asn Asp Val Gln Ala Lys
 210 215 220
 Lys Arg Ser Gly Lys Ala Val Met Asn Met Ser Leu Gly Gly Ser Phe
 225 230 235 240
 Ser Thr Ala Val Asn Asn Ala Ile Thr Ala Leu Thr Asn Ala Gly Ile
 245 250 255
 Val Pro Val Val Ala Ala Gly Asn Glu Asn Gln Asp Thr Ala Asn Thr
 260 265 270
 Ser Pro Gly Ser Ala Pro Gln Ala Ile Thr Val Gly Ala Ile Asp Ala
 275 280 285

ES 2 711 165 T3

Thr Thr Asp Ile Arg Ala Gly Phe Ser Asn Phe Gly Thr Gly Val Asp
290 295 300

Ile Tyr Ala Pro Gly Val Asp Val Leu Ser Val Gly Ile Lys Ser Asp
305 310 315 320

Ile Asp Thr Ala Val Leu Ser Gly Thr Ser Met Ala Ser Pro His Val
325 330 335

Ala Gly Leu Ala Ala Tyr Leu Met Ala Leu Glu Gly Val Ser Asn Val
340 345 350

Asp Asp Val Ser Asn Leu Ile Lys Asn Leu Ala Ala Lys Thr Gly Ala
355 360 365

Ala Val Lys Gln Asn Ile Ala Gly Thr Thr Ser Leu Ile Ala Asn Asn
370 375 380

Gly Asn Phe
385

<210> 29

<211> 409

<212> PRT

<213> Trichoderma reesei

<400> 29

Met Ala Ser Leu Arg Arg Leu Ala Leu Tyr Leu Gly Ala Leu Leu Pro
1 5 10 15

Ala Val Leu Ala Ala Pro Ala Val Asn Tyr Lys Leu Pro Glu Ala Val
20 25 30

Pro Asn Lys Phe Ile Val Thr Leu Lys Asp Gly Ala Ser Val Asp Thr
35 40 45

Asp Ser His Leu Thr Trp Val Lys Asp Leu His Arg Arg Ser Leu Gly
50 55 60

Lys Arg Ser Thr Ala Gly Val Glu Lys Thr Tyr Asn Ile Asp Ser Trp
65 70 75 80

Asn Ala Tyr Ala Gly Glu Phe Asp Glu Glu Thr Val Lys Gln Ile Lys
85 90 95

Ala Asn Pro Asp Val Ala Ser Val Glu Pro Asp Tyr Ile Met Trp Leu
100 105 110

5

10

ES 2 711 165 T3

Ser Asp Ile Val Glu Asp Lys Arg Ala Leu Thr Thr Gln Thr Gly Ala
 115 120 125

Pro Trp Gly Leu Gly Thr Val Ser His Arg Thr Pro Gly Ser Thr Ser
 130 135 140

Tyr Ile Tyr Asp Thr Ser Ala Gly Ser Gly Thr Phe Ala Tyr Val Val
 145 150 155 160

Asp Ser Gly Ile Asn Ile Ala His Gln Gln Phe Gly Gly Arg Ala Ser
 165 170 175

Leu Gly Tyr Asn Ala Ala Gly Gly Asp His Val Asp Thr Leu Gly His
 180 185 190

Gly Thr His Val Ser Gly Thr Ile Gly Gly Ser Thr Tyr Gly Val Ala
 195 200 205

Lys Gln Ala Ser Leu Ile Ser Val Lys Val Phe Gln Gly Asn Ser Ala
 210 215 220

Ser Thr Ser Val Ile Leu Asp Gly Tyr Asn Trp Ala Val Asn Asp Ile
 225 230 235 240

Val Ser Arg Asn Arg Ala Ser Lys Ser Ala Ile Asn Met Ser Leu Gly
 245 250 255

Gly Pro Ala Ser Ser Thr Trp Ala Thr Ala Ile Asn Ala Ala Phe Asn
 260 265 270

Lys Gly Val Leu Thr Ile Val Ala Ala Gly Asn Gly Asp Ala Leu Gly
 275 280 285

Asn Pro Gln Pro Val Ser Ser Thr Ser Pro Ala Asn Val Pro Asn Ala
 290 295 300

Ile Thr Val Ala Ala Leu Asp Ile Asn Trp Arg Thr Ala Ser Phe Thr
 305 310 315 320

Asn Tyr Gly Ala Gly Val Asp Val Phe Ala Pro Gly Val Asn Ile Leu
 325 330 335

Ser Ser Trp Ile Gly Ser Asn Thr Ala Thr Asn Thr Ile Ser Gly Thr
 340 345 350

Ser Met Ala Thr Pro His Val Val Gly Leu Ala Leu Tyr Leu Gln Ala

ES 2 711 165 T3

355 360 365
 Leu Glu Gly Leu Ser Thr Pro Thr Ala Val Thr Asn Arg Ile Lys Ala
 370 375 380

Leu Ala Thr Thr Gly Arg Val Thr Gly Ser Leu Asn Gly Ser Pro Asn
 385 390 395 400

Thr Leu Ile Phe Asn Gly Asn Ser Ala
 405

<210> 30
 <211> 555
 <212> PRT
 <213> Trichoderma reesei

5

<400> 30
 Met Arg Ala Cys Leu Phe Leu Gly Ile Thr Ala Leu Ala Thr Ala
 1 5 10 15

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

Thr Gln Leu Ala Lys Val Ser Ile Ala Leu Gln Pro Glu Cys Arg Glu
 35 40 45

Leu Leu Glu Gln Ala Leu His His Leu Ser Asp Pro Ser Ser Pro Arg
 50 55 60

Tyr Gly Arg Tyr Leu Gly Arg Glu Glu Ala Lys Ala Leu Leu Arg Pro
 65 70 75 80

Arg Arg Glu Ala Thr Ala Ala Val Lys Arg Trp Leu Ala Arg Ala Gly
 85 90 95

Val Pro Ala His Asp Val Leu Thr Asp Gly Gln Phe Ile His Val Arg
 100 105 110

Thr Leu Ala Glu Lys Ala Gln Ala Leu Leu Gly Phe Glu Tyr Asn Ser
 115 120 125

Thr Leu Gly Ser Gln Thr Ile Ala Ile Ser Thr Leu Pro Gly Lys Ile
 130 135 140

Arg Lys His Val Met Thr Val Gln Tyr Val Pro Leu Trp Thr Glu Ala
 145 150 155 160

Asp Trp Glu Glu Cys Lys Thr Ile Ile Thr Pro Ser Cys Leu Lys Arg

10

ES 2 711 165 T3

				165						170						175
Leu	Tyr	His	Val	Asp	Ser	Tyr	Arg	Ala	Lys	Tyr	Glu	Ser	Ser	Ser	Leu	
			180					185					190			
Phe	Gly	Ile	Val	Gly	Phe	Ser	Gly	Gln	Ala	Ala	Gln	His	Asp	Glu	Leu	
		195					200					205				
Asp	Lys	Phe	Leu	His	Asp	Phe	Ala	Pro	Tyr	Ser	Thr	Asn	Ala	Asn	Phe	
	210					215					220					
Ser	Ile	Glu	Ser	Val	Asn	Gly	Gly	Gln	Ser	Pro	Gln	Gly	Met	Asn	Glu	
225					230					235					240	
Pro	Ala	Ser	Glu	Ala	Asn	Gly	Asp	Val	Gln	Tyr	Ala	Val	Ala	Met	Gly	
				245					250					255		
Tyr	His	Val	Pro	Val	Arg	Tyr	Tyr	Ala	Val	Gly	Gly	Glu	Asn	His	Asp	
			260					265					270			
Ile	Ile	Pro	Asp	Leu	Asp	Leu	Val	Asp	Thr	Thr	Glu	Glu	Tyr	Leu	Glu	
		275					280					285				
Pro	Phe	Leu	Glu	Phe	Ala	Ser	His	Leu	Leu	Asp	Leu	Asp	Asp	Asp	Glu	
	290					295					300					
Leu	Pro	Arg	Val	Val	Ser	Ile	Ser	Tyr	Gly	Ala	Asn	Glu	Gln	Leu	Phe	
305					310					315					320	
Pro	Arg	Ser	Tyr	Ala	His	Gln	Val	Cys	Asp	Met	Phe	Gly	Gln	Leu	Gly	
				325					330					335		
Ala	Arg	Gly	Val	Ser	Ile	Val	Val	Ala	Ala	Gly	Asp	Leu	Gly	Pro	Gly	
			340					345					350			
Val	Ser	Cys	Gln	Ser	Asn	Asp	Gly	Ser	Ala	Arg	Pro	Lys	Phe	Ile	Pro	
		355					360					365				
Ser	Phe	Pro	Ala	Thr	Cys	Pro	Tyr	Val	Thr	Ser	Val	Gly	Ser	Thr	Arg	
	370					375					380					
Gly	Ile	Met	Pro	Glu	Val	Ala	Ala	Ser	Phe	Ser	Ser	Gly	Gly	Phe	Ser	
385					390					395					400	
Asp	Tyr	Phe	Ala	Arg	Pro	Ala	Trp	Gln	Asp	Arg	Ala	Val	Gly	Ala	Tyr	
				405					410					415		

ES 2 711 165 T3

Leu Gly Ala His Gly Glu Glu Trp Glu Gly Phe Tyr Asn Pro Ala Gly
 420 425 430

Arg Gly Phe Pro Asp Val Ala Ala Gln Gly Val Asn Phe Arg Phe Arg
 435 440 445

Ala His Gly Asn Glu Ser Leu Ser Ser Gly Thr Ser Leu Ser Ser Pro
 450 455 460

Val Phe Ala Ala Leu Ile Ala Leu Leu Asn Asp His Arg Ser Lys Ser
 465 470 475 480

Gly Met Pro Pro Met Gly Phe Leu Asn Pro Trp Ile Tyr Thr Val Gly
 485 490 495

Ser His Ala Phe Thr Asp Ile Ile Glu Ala Arg Ser Glu Gly Cys Pro
 500 505 510

Gly Gln Ser Val Glu Tyr Leu Ala Ser Pro Tyr Ile Pro Asn Ala Gly
 515 520 525

Trp Ser Ala Val Pro Gly Trp Asp Pro Val Thr Gly Trp Gly Thr Pro
 530 535 540

Leu Phe Asp Arg Met Leu Asn Leu Ser Leu Val
 545 550 555

<210> 31

<211> 388

<212> PRT

<213> Trichoderma reesei

<400> 31

Met Ala Trp Leu Lys Lys Leu Ala Leu Val Leu Leu Ala Ile Val Pro
 1 5 10 15

Tyr Ala Thr Ala Ser Pro Ala Leu Ser Pro Arg Ser Arg Glu Ile Leu
 20 25 30

Ser Leu Glu Asp Leu Glu Ser Glu Asp Lys Tyr Val Ile Gly Leu Lys
 35 40 45

Gln Gly Leu Ser Pro Thr Asp Leu Lys Lys His Leu Leu Arg Val Ser
 50 55 60

Ala Val Gln Tyr Arg Asn Lys Asn Ser Thr Phe Glu Gly Gly Thr Gly
 65 70 75 80

5

10

ES 2 711 165 T3

Val Lys Arg Thr Tyr Ala Ile Gly Asp Tyr Arg Ala Tyr Thr Ala Val
85 90 95

Leu Asp Arg Asp Thr Val Arg Glu Ile Trp Asn Asp Thr Leu Glu Lys
100 105 110

Pro Pro Trp Gly Leu Ala Thr Leu Ser Asn Lys Lys Pro His Gly Phe
115 120 125

Leu Tyr Arg Tyr Asp Lys Ser Ala Gly Glu Gly Thr Phe Ala Tyr Val
130 135 140

Leu Asp Thr Gly Ile Asn Ser Lys His Val Asp Phe Glu Gly Arg Ala
145 150 155 160

Tyr Met Gly Phe Ser Pro Pro Lys Thr Glu Pro Thr Asp Ile Asn Gly
165 170 175

His Gly Thr His Val Ala Gly Ile Ile Gly Gly Lys Thr Phe Gly Val
180 185 190

Ala Lys Lys Thr Gln Leu Ile Gly Val Lys Val Phe Leu Asp Asp Glu
195 200 205

Ala Thr Thr Ser Thr Leu Met Glu Gly Leu Glu Trp Ala Val Asn Asp
210 215 220

Ile Thr Thr Lys Gly Arg Gln Gly Arg Ser Val Ile Asn Met Ser Leu
225 230 235 240

Gly Gly Pro Tyr Ser Gln Ala Leu Asn Asp Ala Ile Asp His Ile Ala
245 250 255

Asp Met Gly Ile Leu Pro Val Ala Ala Ala Gly Asn Lys Gly Ile Pro
260 265 270

Ala Thr Phe Ile Ser Pro Ala Ser Ala Asp Lys Ala Met Thr Val Gly
275 280 285

Ala Ile Asn Ser Asp Trp Gln Glu Thr Asn Phe Ser Asn Phe Gly Pro
290 295 300

Gln Val Asn Ile Leu Ala Pro Gly Glu Asp Val Leu Ser Ala Tyr Val
305 310 315 320

Ser Thr Asn Thr Ala Thr Arg Val Leu Ser Gly Thr Ser Met Ala Ala
325 330 335

ES 2 711 165 T3

Pro His Val Ala Gly Leu Ala Leu Tyr Leu Met Ala Leu Glu Glu Phe
 340 345 350

Asp Ser Thr Gln Lys Leu Thr Asp Arg Ile Leu Gln Leu Gly Met Lys
 355 360 365

Asn Lys Val Val Asn Leu Met Thr Asp Ser Pro Asn Leu Ile Ile His
 370 375 380

Asn Asn Val Lys
 385

<210> 32

<211> 256

<212> PRT

<213> Trichoderma reesei

<400> 32

Met Phe Ile Ala Gly Val Ala Leu Ser Ala Leu Leu Cys Ala Asp Thr
 1 5 10 15

Val Leu Ala Gly Val Ala Gln Asp Arg Gly Leu Ala Ala Arg Leu Ala
 20 25 30

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

Ala Thr Val Gln Ser Asn Trp Gly Gly Ala Ile Leu Glu Gly Ser Gly
 50 55 60

Phe Thr Ala Ala Ser Ala Thr Val Asn Val Pro Arg Gly Gly Gly Gly
 65 70 75 80

Ser Asn Ala Ala Gly Ser Ala Trp Val Gly Ile Asp Gly Ala Ser Cys
 85 90 95

Gln Thr Ala Ile Leu Gln Thr Gly Phe Asp Trp Tyr Gly Asp Gly Thr
 100 105 110

Tyr Asp Ala Trp Tyr Glu Trp Tyr Pro Glu Phe Ala Ala Asp Phe Ser
 115 120 125

Gly Ile Asp Ile Arg Gln Gly Asp Gln Ile Ala Met Ser Val Val Ala
 130 135 140

Thr Ser Leu Thr Gly Gly Ser Ala Thr Leu Glu Asn Leu Ser Thr Gly
 145 150 155 160

5

10

ES 2 711 165 T3

Gln Lys Val Thr Gln Asn Phe Asn Arg Val Thr Ala Gly Ser Leu Cys
 165 170 175

Glu Thr Ser Ala Glu Phe Ile Ile Glu Asp Phe Glu Glu Cys Asn Ser
 180 185 190

Asn Gly Ser Asn Cys Gln Pro Val Pro Phe Ala Ser Phe Ser Pro Ala
 195 200 205

Ile Thr Phe Ser Ser Ala Thr Ala Thr Arg Ser Gly Arg Ser Val Ser
 210 215 220

Leu Ser Gly Ala Glu Ile Thr Glu Val Ile Val Asn Asn Gln Asp Leu
 225 230 235 240

Thr Arg Cys Ser Val Ser Gly Ser Ser Thr Leu Thr Cys Ser Tyr Val
 245 250 255

<210> 33

<211> 236

<212> PRT

<213> Trichoderma reesei

<400> 33

Met Asp Ala Ile Arg Ala Arg Ser Ala Ala Arg Arg Ser Asn Arg Phe
 1 5 10 15

Gln Ala Gly Ser Ser Lys Asn Val Asn Gly Thr Ala Asp Val Glu Ser
 20 25 30

Thr Asn Trp Ala Gly Ala Ala Ile Thr Thr Ser Gly Val Thr Glu Val
 35 40 45

Ser Gly Thr Phe Thr Val Pro Arg Pro Ser Val Pro Ala Gly Gly Ser
 50 55 60

Ser Arg Glu Glu Tyr Cys Gly Ala Ala Trp Val Gly Ile Asp Gly Tyr
 65 70 75 80

Ser Asp Ala Asp Leu Ile Gln Thr Gly Val Leu Trp Cys Val Glu Asp
 85 90 95

Gly Glu Tyr Leu Tyr Glu Ala Trp Tyr Glu Tyr Leu Pro Ala Ala Leu
 100 105 110

Val Glu Tyr Ser Gly Ile Ser Val Thr Ala Gly Ser Val Val Thr Val
 115 120 125

5

10

ES 2 711 165 T3

Thr Ala Thr Lys Thr Gly Thr Asn Ser Gly Val Thr Thr Leu Thr Ser
130 135 140

Gly Gly Lys Thr Val Ser His Thr Phe Ser Arg Gln Asn Ser Pro Leu
145 150 155 160

Pro Gly Thr Ser Ala Glu Trp Ile Val Glu Asp Phe Thr Ser Gly Ser
165 170 175

Ser Leu Val Pro Phe Ala Asp Phe Gly Ser Val Thr Phe Thr Gly Ala
180 185 190

Thr Ala Val Val Asn Gly Ala Thr Val Thr Ala Gly Gly Asp Ser Pro
195 200 205

Val Ile Ile Asp Leu Glu Asp Ser Arg Gly Asp Ile Leu Thr Ser Thr
210 215 220

Thr Val Ser Gly Ser Thr Val Thr Val Glu Tyr Glu
225 230 235

<210> 34

<211> 612

<212> PRT

<213> Trichoderma reesei

<400> 34

Met Ala Lys Leu Ser Thr Leu Arg Leu Ala Ser Leu Leu Ser Leu Val
1 5 10 15

Ser Val Gln Val Ser Ala Ser Val His Leu Leu Glu Ser Leu Glu Lys
20 25 30

Leu Pro His Gly Trp Lys Ala Ala Glu Thr Pro Ser Pro Ser Ser Gln
35 40 45

Ile Val Leu Gln Val Ala Leu Thr Gln Gln Asn Ile Asp Gln Leu Glu
50 55 60

Ser Arg Leu Ala Ala Val Ser Thr Pro Thr Ser Ser Thr Tyr Gly Lys
65 70 75 80

Tyr Leu Asp Val Asp Glu Ile Asn Ser Ile Phe Ala Pro Ser Asp Ala
85 90 95

Ser Ser Ser Ala Val Glu Ser Trp Leu Gln Ser His Gly Val Thr Ser
100 105 110

5

10

ES 2 711 165 T3

Tyr Thr Lys Gln Gly Ser Ser Ile Trp Phe Gln Thr Asn Ile Ser Thr
 115 120 125
 Ala Asn Ala Met Leu Ser Thr Asn Phe His Thr Tyr Ser Asp Leu Thr
 130 135 140
 Gly Ala Lys Lys Val Arg Thr Leu Lys Tyr Ser Ile Pro Glu Ser Leu
 145 150 155 160
 Ile Gly His Val Asp Leu Ile Ser Pro Thr Thr Tyr Phe Gly Thr Thr
 165 170 175
 Lys Ala Met Arg Lys Leu Lys Ser Ser Gly Val Ser Pro Ala Ala Asp
 180 185 190
 Ala Leu Ala Ala Arg Gln Glu Pro Ser Ser Cys Lys Gly Thr Leu Val
 195 200 205
 Phe Glu Gly Glu Thr Phe Asn Val Phe Gln Pro Asp Cys Leu Arg Thr
 210 215 220
 Glu Tyr Ser Val Asp Gly Tyr Thr Pro Ser Val Lys Ser Gly Ser Arg
 225 230 235 240
 Ile Gly Phe Gly Ser Phe Leu Asn Glu Ser Ala Ser Phe Ala Asp Gln
 245 250 255
 Ala Leu Phe Glu Lys His Phe Asn Ile Pro Ser Gln Asn Phe Ser Val
 260 265 270
 Val Leu Ile Asn Gly Gly Thr Asp Leu Pro Gln Pro Pro Ser Asp Ala
 275 280 285
 Asn Asp Gly Glu Ala Asn Leu Asp Ala Gln Thr Ile Leu Thr Ile Ala
 290 295 300
 His Pro Leu Pro Ile Thr Glu Phe Ile Thr Ala Gly Ser Pro Pro Tyr
 305 310 315 320
 Phe Pro Asp Pro Val Glu Pro Ala Gly Thr Pro Asn Glu Asn Glu Pro
 325 330 335
 Tyr Leu Gln Tyr Tyr Glu Phe Leu Leu Ser Lys Ser Asn Ala Glu Ile
 340 345 350
 Pro Gln Val Ile Thr Asn Ser Tyr Gly Asp Glu Glu Gln Thr Val Pro
 355 360 365

ES 2 711 165 T3

Arg Ser Tyr Ala Val Arg Val Cys Asn Leu Ile Gly Leu Leu Gly Leu
 370 375 380

Arg Gly Ile Ser Val Leu His Ser Ser Gly Asp Glu Gly Val Gly Ala
 385 390 395 400

Ser Cys Val Ala Thr Asn Ser Thr Thr Pro Gln Phe Asn Pro Ile Phe
 405 410 415

Pro Ala Thr Cys Pro Tyr Val Thr Ser Val Gly Gly Thr Val Ser Phe
 420 425 430

Asn Pro Glu Val Ala Trp Ala Gly Ser Ser Gly Gly Phe Ser Tyr Tyr
 435 440 445

Phe Ser Arg Pro Trp Tyr Gln Gln Glu Ala Val Gly Thr Tyr Leu Glu
 450 455 460

Lys Tyr Val Ser Ala Glu Thr Lys Lys Tyr Tyr Gly Pro Tyr Val Asp
 465 470 475 480

Phe Ser Gly Arg Gly Phe Pro Asp Val Ala Ala His Ser Val Ser Pro
 485 490 495

Asp Tyr Pro Val Phe Gln Gly Gly Glu Leu Thr Pro Ser Gly Gly Thr
 500 505 510

Ser Ala Ala Ser Pro Val Val Ala Ala Ile Val Ala Leu Leu Asn Asp
 515 520 525

Ala Arg Leu Arg Glu Gly Lys Pro Thr Leu Gly Phe Leu Asn Pro Leu
 530 535 540

Ile Tyr Leu His Ala Ser Lys Gly Phe Thr Asp Ile Thr Ser Gly Gln
 545 550 555 560

Ser Glu Gly Cys Asn Gly Asn Asn Thr Gln Thr Gly Ser Pro Leu Pro
 565 570 575

Gly Ala Gly Phe Ile Ala Gly Ala His Trp Asn Ala Thr Lys Gly Trp
 580 585 590

Asp Pro Thr Thr Gly Phe Gly Val Pro Asn Leu Lys Lys Leu Leu Ala
 595 600 605

Leu Val Arg Phe
 610

ES 2 711 165 T3

<211> 477
 <212> PRT
 <213> Trichoderma reesei

5

<400> 35
 Met Arg Phe Val Gln Tyr Val Ser Leu Ala Gly Leu Phe Ala Ala Ala
 1 5 10 15
 Thr Val Ser Ala Gly Val Val Thr Val Pro Phe Glu Lys Arg Asn Leu
 20 25 30
 Asn Pro Asp Phe Ala Pro Ser Leu Leu Arg Arg Asp Gly Ser Val Ser
 35 40 45
 Leu Asp Ala Ile Asn Asn Leu Thr Gly Gly Gly Tyr Tyr Ala Gln Phe
 50 55 60
 Ser Val Gly Thr Pro Pro Gln Lys Leu Ser Phe Leu Leu Asp Thr Gly
 65 70 75 80
 Ser Ser Asp Thr Trp Val Asn Ser Val Thr Ala Asp Leu Cys Thr Asp
 85 90 95
 Glu Phe Thr Gln Gln Thr Val Gly Glu Tyr Cys Phe Arg Gln Phe Asn
 100 105 110
 Pro Arg Arg Ser Ser Ser Tyr Lys Ala Ser Thr Glu Val Phe Asp Ile
 115 120 125
 Thr Tyr Leu Asp Gly Arg Arg Ile Arg Gly Asn Tyr Phe Thr Asp Thr
 130 135 140
 Val Thr Ile Asn Gln Ala Asn Ile Thr Gly Gln Lys Ile Gly Leu Ala
 145 150 155 160
 Leu Gln Ser Val Arg Gly Thr Gly Ile Leu Gly Leu Gly Phe Arg Glu
 165 170 175
 Asn Glu Ala Ala Asp Thr Lys Tyr Pro Thr Val Ile Asp Asn Leu Val
 180 185 190
 Ser Gln Lys Val Ile Pro Val Pro Ala Phe Ser Leu Tyr Leu Asn Asp
 195 200 205
 Leu Gln Thr Ser Gln Gly Ile Leu Leu Phe Gly Gly Val Asp Thr Asp

ES 2 711 165 T3

210 Lys Phe His Gly Gly Leu Ala Thr Leu Pro Leu Gln Ser Leu Pro Pro 225 Ser Ile Ala Glu Thr Gln Asp Ile Val Met Tyr Ser Val Asn Leu Asp 245 Gly Phe Ser Ala Ser Asp Val Asp Thr Pro Asp Val Ser Ala Lys Ala 260 Val Leu Asp Ser Gly Ser Thr Ile Thr Leu Leu Pro Asp Ala Val Val 275 Gln Glu Leu Phe Asp Glu Tyr Asp Val Leu Asn Ile Gln Gly Leu Pro 290 Val Pro Phe Ile Asp Cys Ala Lys Ala Asn Ile Lys Asp Ala Thr Phe 305 Asn Phe Lys Phe Asp Gly Lys Thr Ile Lys Val Pro Ile Asp Glu Met 325 Val Leu Asn Asn Leu Ala Ala Ala Ser Asp Glu Ile Met Ser Asp Pro 340 Ser Leu Ser Lys Phe Phe Lys Gly Trp Ser Gly Val Cys Thr Phe Gly 355 Met Gly Ser Thr Lys Thr Phe Gly Ile Gln Ser Asp Glu Phe Val Leu 370 Leu Gly Asp Thr Phe Leu Arg Ser Ala Tyr Val Val Tyr Asp Leu Gln 385 Asn Lys Gln Ile Gly Ile Ala Gln Ala Thr Leu Asn Ser Thr Ser Ser 405 Thr Ile Val Glu Phe Lys Ala Gly Ser Lys Thr Ile Pro Gly Pro Ala 420 Ser Thr Gly Asp Asp Ser Asp Asp Ser Ser Asp Asp Ser Asp Glu Asp 435 Ser Ala Gly Ala Ala Leu His Pro Thr Phe Ser Ile Ala Leu Ala Gly 450 Thr Leu Phe Thr Ala Val Ser Met Met Met Ser Val Leu 465 470 475	215 230 265 280 295 310 330 345 360 375 390 410 425 440 455 470	220 235 255 270 285 300 315 335 350 365 380 395 415 430 445 460 475
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

ES 2 711 165 T3

<210> 36
 <211> 1263
 <212> ADN
 <213> Trichoderma reesei

5

<400> 36
 atggcgtcac tcatcaaaac tgccgtggac attgccaaacg gccgccatgc gctgtccaga 60
 tatgtcatct ttgggctctg gcttgccgat gcgggtgctgt gcgggctgat tatctggaaa 120
 gtgccttata cggaaatcga ctgggtcgcc tacatggagc aagtcaccca gttcgtccac 180
 ggagagcgag actaccccaa gatggagggc ggcacagggc ccctggtgta tcccgcggcc 240
 catgtgtaca tctacacagg gctctactac ctgacgaaca agggcaccga catcctgctg 300
 gcgcagcagc tctttgccgt gctctacatg gctactctgg cggtcgtcat gacatgctac 360
 tccaaggcca aggtcccgcc gtacatcttc ccgcttctca tcctctcaa aagacttcac 420
 agcgtcttcg tcttgagatg cttcaacgac tgcttcgccg ccttcttcct ctggctctgc 480
 atcttcttct tccagaggcg agagtggacc atcggagctc tcgcatacag catcggcctg 540
 ggcgtcaaaa tgtcgtctct actggttctc ccgcctggg tcatcgtcct ctacctcggc 600
 cgcggcttca agggcgccct gcggctgctc tggctcatgg tgcaggtcca gtcctcctc 660
 gccataccct tcatcacgac aaattggcgc ggctacctcg gccgtgcatt cgagctctcg 720
 aggcagttca agtttgaatg gacagtcaat tggcgcagtc tgggcgagga tctgttcctc 780
 agccggggct tctctatcac gctactggca tttcacgcca tcttctcctc cgcctttatc 840
 ctcggccggg ggctgaagat tagggaacgg accgtactcg ggatgatccc ctatgtcatc 900
 cgattcagat cgccctttac cgagcaggaa gagcgcgcca tctccaaccg cgtcgtcacg 960
 cccggctatg tcatgtccac catcttgctg gccaacgtgg tgggactgct gtttgcccg 1020
 tctctgcact accagttcta tgcatatctg gcgtgggcga cccctatct cctgtggacg 1080
 gcctgcccga atcttttggg ggtggcccc ctctgggcgg cgcaagaatg ggcctggaac 1140
 gtcttcccga gcacgcctct tagctcgagc gtcgtggtga gcgtgctggc cgtgacggtg 1200
 gccatggcgt ttgcaggttc aaatccgcag ccacgtgaaa catcgaagcc gaagcagcac 1260
 taa 1263

10

<210> 37
 <211> 420
 <212> PRT
 <213> Trichoderma reesei

15

<400> 37
 Met Ala Ser Leu Ile Lys Thr Ala Val Asp Ile Ala Asn Gly Arg His

ES 2 711 165 T3

Asp Leu Phe Leu Ser Arg Gly Phe Ser Ile Thr Leu Leu Ala Phe His
 260 265 270

Ala Ile Phe Leu Leu Ala Phe Ile Leu Gly Arg Trp Leu Lys Ile Arg
 275 280 285

Glu Arg Thr Val Leu Gly Met Ile Pro Tyr Val Ile Arg Phe Arg Ser
 290 295 300

Pro Phe Thr Glu Gln Glu Glu Arg Ala Ile Ser Asn Arg Val Val Thr
 305 310 315 320

Pro Gly Tyr Val Met Ser Thr Ile Leu Ser Ala Asn Val Val Gly Leu
 325 330 335

Leu Phe Ala Arg Ser Leu His Tyr Gln Phe Tyr Ala Tyr Leu Ala Trp
 340 345 350

Ala Thr Pro Tyr Leu Leu Trp Thr Ala Cys Pro Asn Leu Leu Val Val
 355 360 365

Ala Pro Leu Trp Ala Ala Gln Glu Trp Ala Trp Asn Val Phe Pro Ser
 370 375 380

Thr Pro Leu Ser Ser Ser Val Val Val Ser Val Leu Ala Val Thr Val
 385 390 395 400

Ala Met Ala Phe Ala Gly Ser Asn Pro Gln Pro Arg Glu Thr Ser Lys
 405 410 415

Pro Lys Gln His
 420

<210> 38

<211> 445

<212> PRT

<213> Homo sapiens

<400> 38

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

Phe Val Ala Trp Asn Ala Leu Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30

Ala Pro Gly Arg Pro Pro Ser Val Ser Ala Leu Asp Gly Asp Pro Ala
 35 40 45

5

10

ES 2 711 165 T3

Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala Glu Val Glu
50 55 60

Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Gly Asp Ala Leu Ser
65 70 75 80

Ser Gln Arg Gly Arg Val Pro Thr Ala Ala Pro Pro Ala Gln Pro Arg
85 90 95

Val Pro Val Thr Pro Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala
100 105 110

Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu His Tyr
115 120 125

Arg Pro Ser Ala Glu Leu Phe Pro Ile Ile Val Ser Gln Asp Cys Gly
130 135 140

His Glu Glu Thr Ala Gln Ala Ile Ala Ser Tyr Gly Ser Ala Val Thr
145 150 155 160

His Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His
165 170 175

Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala
180 185 190

Leu Gly Gln Val Phe Arg Gln Phe Arg Phe Pro Ala Ala Val Val Val
195 200 205

Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe Arg Ala
210 215 220

Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val Ser Ala
225 230 235 240

Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ala Ser Arg Pro Glu
245 250 255

Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu
260 265 270

Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala Phe Trp
275 280 285

Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala Cys Ile
290 295 300

ES 2 711 165 T3

Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser
305 310 315 320

His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln
325 330 335

Gln Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg Glu
340 345 350

Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro Gln Leu
355 360 365

Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly Glu Val
370 375 380

Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys Ala
385 390 395 400

Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala Gly Tyr
405 410 415

Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His Leu Ala
420 425 430

Pro Pro Leu Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
435 440 445

<210> 39

<211> 447

<212> PRT

<213> Homo sapiens

<400> 39

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

Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
20 25 30

Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
35 40 45

Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
50 55 60

Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro
65 70 75 80

5

10

ES 2 711 165 T3

Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
85 90 95

Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
100 105 110

Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
115 120 125

Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
130 135 140

Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
145 150 155 160

Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
165 170 175

Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
180 185 190

Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
195 200 205

Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
210 215 220

Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
225 230 235 240

Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
245 250 255

Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
260 265 270

Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
275 280 285

Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
290 295 300

Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
305 310 315 320

Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
325 330 335

ES 2 711 165 T3

Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350

Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365

Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380

Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400

Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
 405 410 415

Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430

Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440 445

<210> 40

<211> 85

<212> PRT

<213> Trichoderma reesei

5

<400> 40

Met Ala Ser Thr Asn Ala Arg Tyr Val Arg Tyr Leu Leu Ile Ala Phe
 1 5 10 15

Phe Thr Ile Leu Val Phe Tyr Phe Val Ser Asn Ser Lys Tyr Glu Gly
 20 25 30

Val Asp Leu Asn Lys Gly Thr Phe Thr Ala Pro Asp Ser Thr Lys Thr
 35 40 45

Thr Pro Lys Pro Pro Ala Thr Gly Asp Ala Lys Asp Phe Pro Leu Ala
 50 55 60

Leu Thr Pro Asn Asp Pro Gly Phe Asn Asp Leu Val Gly Ile Ala Pro
 65 70 75 80

Gly Pro Arg Met Asn
 85

10

<210> 41

<211> 255

<212> ADN

<213> Trichoderma reesei

15

<400> 41

ES 2 711 165 T3

atggcgtcaa caaatgcgcg ctatgtgctg tatctactaa tcgccttctt cacaatcctc 60
 gtcttctact ttgtctccaa ttcaaagtat gagggcgtcg atctcaacaa gggcaccttc 120
 acagctccgg attcgaccaa gacgacacca aagccgccag ccactggcga tgccaaagac 180
 tttcctctgg ccctgacgcc gaacgatcca ggcttcaacg acctcgtcgg catcgtctcc 240
 ggccctcgaa tgaac 255

<210> 42
 <211> 58
 <212> PRT
 <213> Homo sapiens

5

<400> 42
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45

Gly Ala Gly Gly Arg Gly Gly Asp His Pro
 50 55

10

<210> 43
 <211> 51
 <212> PRT
 <213> Trichoderma reesei

15

<400> 43
 Met Ala Ser Thr Asn Ala Arg Tyr Val Arg Tyr Leu Leu Ile Ala Phe
 1 5 10 15
 Phe Thr Ile Leu Val Phe Tyr Phe Val Ser Asn Ser Lys Tyr Glu Gly
 20 25 30
 Val Asp Leu Asn Lys Gly Thr Phe Thr Ala Pro Asp Ser Thr Lys Thr
 35 40 45

Thr Pro Lys
 50

20

<210> 44
 <211> 52
 <212> PRT
 <213> Trichoderma reesei

<400> 44

ES 2 711 165 T3

Met Ala Ile Ala Arg Pro Val Arg Ala Leu Gly Gly Leu Ala Ala Ile
 1 5 10 15

Leu Trp Cys Phe Phe Leu Tyr Gln Leu Leu Arg Pro Ser Ser Ser Tyr
 20 25 30

Asn Ser Pro Gly Asp Arg Tyr Ile Asn Phe Glu Arg Asp Pro Asn Leu
 35 40 45

Asp Pro Thr Gly
 50

<210> 45

<211> 33

<212> PRT

<213> Trichoderma reesei

5

<400> 45

Met Leu Asn Pro Arg Arg Ala Leu Ile Ala Ala Ala Phe Ile Leu Thr
 1 5 10 15

Val Phe Phe Leu Ile Ser Arg Ser His Asn Ser Glu Ser Ala Ser Thr
 20 25 30

Ser

<210> 46

<211> 84

<212> PRT

<213> Trichoderma reesei

10

<400> 46

Met Met Pro Arg His His Ser Ser Gly Phe Ser Asn Gly Tyr Pro Arg
 1 5 10 15

Ala Asp Thr Phe Glu Ile Ser Pro His Arg Phe Gln Pro Arg Ala Thr
 20 25 30

Leu Pro Pro His Arg Lys Arg Lys Arg Thr Ala Ile Arg Val Gly Ile
 35 40 45

Ala Val Val Val Ile Leu Val Leu Val Leu Trp Phe Gly Gln Pro Arg
 50 55 60

Ser Val Ala Ser Leu Ile Ser Leu Gly Ile Leu Ser Gly Tyr Asp Asp
 65 70 75 80

Leu Lys Leu Glu

<210> 47

<211> 55

<212> PRT

<213> Trichoderma reesei

20

25

ES 2 711 165 T3

<400> 47

Met Leu Leu Pro Lys Gly Gly Leu Asp Trp Arg Ser Ala Arg Ala Gln
 1 5 10 15

Ile Pro Pro Thr Arg Ala Leu Trp Asn Ala Val Thr Arg Thr Arg Phe
 20 25 30

Ile Leu Leu Val Gly Ile Thr Gly Leu Ile Leu Leu Leu Trp Arg Gly
 35 40 45

Val Ser Thr Ser Ala Ser Glu
 50 55

5 <210> 48
 <211> 20
 <212> ADN
 <213> Artificial

10 <220>
 <223> Cebador

<400> 48
 ccgcggtgaa cggcttccca 20

15 <210> 49
 <211> 24
 <212> ADN
 <213> Artificial

20 <220>
 <223> Cebador

<400> 49
 taactgtac gctctcagtt cgag 24

25 <210> 50
 <211> 20
 <212> ADN
 <213> Artificial

30 <220>
 <223> Cebador

<400> 50
 gcgacggcga cccattagca 20

35 <210> 51
 <211> 19
 <212> ADN
 <213> Artificial

40 <220>
 <223> Cebador

45 <400> 51
 catcctcaag gcctcagac 19

ES 2 711 165 T3

<210> 52
 <211> 20
 <212> ADN
 <213> Artificial
 5
 <220>
 <223> Cebador
 <400> 52
 10 tgcgctctca ccagcatcgc 20
 <210> 53
 <211> 20
 <212> ADN
 15 <213> Artificial
 <220>
 <223> Cebador
 20 <400> 53
 gtcctgggcg agttccgcac 20
 <210> 54
 <211> 63
 25 <212> ADN
 <213> Artificial
 <220>
 <223> Cebador
 30 <400> 54
 agatttcagt ctctcaccac tcacctgagt tgcctctctc ggtctgaagg acgtggaatg 60
 atg 63
 <210> 55
 35 <211> 66
 <212> ADN
 <213> Artificial
 <220>
 40 <223> Cebador
 <400> 55
 gcagggtgat gagctggatc accttgacgg tgttgcccat gttgagagaa gttggttgat 60
 tgatca 66
 <210> 56
 <211> 63
 <212> ADN
 <213> Artificial
 50 <220>
 <223> Cebador
 <400> 56
 agatttcagt ctctcaccac tcacctgagt tgcctctctc ggtctgaagg acgtggaatg 60
 atg 63

ES 2 711 165 T3

5 <210> 57
 <211> 66
 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador

 10 <400> 57
 cagagccgct atcgccgagg aggttgcct tcttgcccat gttgagagaa gttggtgat 60
 tgatca 66

 15 <210> 58
 <211> 63
 <212> ADN
 <213> Artificial

 20 <220>
 <223> Cebador

 <400> 58
 agatttcagt ctctcaccac tcacctgagt tgcctctctc ggtctgaagg acgtggaatg 60
 atg 63

 25 <210> 59
 <211> 66
 <212> ADN
 <213> Artificial

 30 <220>
 <223> Cebador

 <400> 59
 tcttgaggat gagctggacg aggttcttga aaaagcccat gttgagagaa gttggtgat 60
 tgatca 66

 35 <210> 60
 <211> 21
 <212> ADN
 <213> Artificial

 40 <220>
 <223> Cebador

 45 <400> 60
 agctccgtgg cgaaagcctg a 21

 50 <210> 61
 <211> 66
 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador

 <400> 61

ES 2 711 165 T3

	cagccgcagc ctcagcctct ctcagcctca tcagccggcg ccgccaactt tgcgtccctt	60
	gtgacg	66
5	<210> 62 <211> 76 <212> ADN <213> Artificial	
10	<220> <223> Cebador	
	<400> 62 gcaacgagag cagagcagca gtagtcgatg ctaggcggcc gcgggcagta tgccggatgg	60
	ctggcttata caggca	76
15	<210> 63 <211> 76 <212> ADN <213> Artificial	
20	<220> <223> Cebador	
	<400> 63 tgcctgtata agccagccat ccggcatact gcccgcgccc gcctagcatc gactactgct	60
	gctctgctct cgttgc	76
25	<210> 64 <211> 20 <212> ADN <213> Artificial	
30	<220> <223> Cebador	
35	<400> 64 tgcgtcgccg tctcgctct	20
40	<210> 65 <211> 20 <212> ADN <213> Artificial	
45	<220> <223> Cebador	
	<400> 65 ttagcgacc tcttttcca	20
50	<210> 66 <211> 20 <212> ADN <213> Artificial	
	<220> <223> Cebador	

ES 2 711 165 T3

	<400> 66 cgaggaagtc tcgtgaggat	20
5	<210> 67 <211> 19 <212> ADN <213> Artificial	
10	<220> <223> Cebador	
	<400> 67 cagctaaacc gacgggcca	19
15	<210> 68 <211> 20 <212> ADN <213> Artificial	
20	<220> <223> Cebador	
	<400> 68 gaccgtatat ttgaaaaggg	20
25	<210> 69 <211> 20 <212> ADN <213> Artificial	
30	<220> <223> Cebador	
	<400> 69 gatgttcgc ctgggtgac	20
35	<210> 70 <211> 23 <212> ADN <213> Artificial	
40	<220> <223> Cebador	
	<400> 70 taactgtac gctctcagtt cga	23
45	<210> 71 <211> 20 <212> ADN <213> Artificial	
50	<220> <223> Cebador	
	<400> 71 ccatgagctt gaacaggtaa	20
55	<210> 72	

<211> 20
 <212> ADN
 <213> Artificial
 5 <220>
 <223> Cebador
 <400> 72
 gattgtcatg ggtacgtga 20
 10 <210> 73
 <211> 20
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Cebador
 <400> 73
 caagatggag ggcggcacag 20
 <210> 74
 <211> 22
 <212> ADN
 <213> Artificial
 25 <220>
 <223> Cebador
 <400> 74
 gccagtagcg tgatagagaa gc 22
 <210> 75
 <211> 20
 <212> ADN
 <213> Artificial
 35 <220>
 <223> Cebador
 <400> 75
 gcgtcactca tcaaaactgc 20
 <210> 76
 <211> 19
 <212> ADN
 <213> Artificial
 40 <220>
 <223> Cebador
 <400> 76
 cttcggcttc gatgttca 19
 50 <210> 77
 <211> 20
 <212> ADN
 <213> Artificial
 55

ES 2 711 165 T3

<220>
 <223> Cebador

 5 <400> 77
 tgcgtcgccg tctcgctct 20

 10 <210> 78
 <211> 20
 <212> ADN
 <213> Artificial

 15 <220>
 <223> Cebador

 <400> 78
 tgacgtacca gttgggatga 20

 20 <210> 79
 <211> 20
 <212> ADN
 <213> Artificial

 25 <220>
 <223> Cebador

 <400> 79
 gatgttcgc ctgggtgac 20

 30 <210> 80
 <211> 20
 <212> ADN
 <213> Artificial

 35 <220>
 <223> Cebador

 <400> 80
 tgacgtacca gttgggatga 20

 40 <210> 81
 <211> 20
 <212> ADN
 <213> Artificial

 45 <220>
 <223> Cebador

 <400> 81
 tgcgtcgccg tctcgctct 20

 50 <210> 82
 <211> 20
 <212> ADN
 <213> Artificial

 55 <220>
 <223> Cebador

 <400> 82

ES 2 711 165 T3

gattgtcatg gtgtacgtga 20
 <210> 83
 <211> 20
 5 <212> ADN
 <213> Artificial
 <220>
 10 <223> Cebador
 <400> 83
 caagatggag ggcggcacag 20
 <210> 84
 15 <211> 22
 <212> ADN
 <213> Artificial
 <220>
 20 <223> Cebador
 <400> 84
 gccagtagcg tgatagagaa gc 22
 25 <210> 85
 <211> 488
 <212> PRT
 <213> Trichoderma reesei
 30 <400> 85
 Met Arg Ala Ser Pro Leu Ala Val Ala Gly Val Ala Leu Ala Ser Ala
 1 5 10 15

ES 2 711 165 T3

Ala Gln Ala Gln Val Val Gln Phe Asp Ile Glu Lys Arg His Ala Pro
 20 25 30

Arg Leu Ser Arg Arg Asp Gly Thr Ile Asp Gly Thr Leu Ser Asn Gln
 35 40 45

Arg Val Gln Gly Gly Tyr Phe Ile Asn Val Gln Val Gly Ser Pro Gly
 50 55 60

Gln Asn Ile Thr Leu Gln Leu Asp Thr Gly Ser Ser Asp Val Trp Val
 65 70 75 80

Pro Ser Ser Thr Ala Ala Ile Cys Thr Gln Val Ser Glu Arg Asn Pro
 85 90 95

Gly Cys Gln Phe Gly Ser Phe Asn Pro Asp Asp Ser Asp Thr Phe Asp
 100 105 110

Glu Val Gly Gln Gly Leu Phe Asp Ile Thr Tyr Val Asp Gly Ser Ser
 115 120 125

Ser Lys Gly Asp Tyr Phe Gln Asp Asn Phe Gln Ile Asn Gly Val Thr
 130 135 140

Val Lys Asn Leu Thr Met Gly Leu Gly Leu Ser Ser Ser Ile Pro Asn
 145 150 155 160

Gly Leu Ile Gly Val Gly Tyr Met Asn Asp Glu Ala Ser Val Ser Thr
 165 170 175

Thr Arg Ser Thr Tyr Pro Asn Leu Pro Ile Val Leu Gln Gln Gln Lys
 180 185 190

Leu Ile Asn Ser Val Ala Phe Ser Leu Trp Leu Asn Asp Leu Asp Ala
 195 200 205

Ser Thr Gly Ser Ile Leu Phe Gly Gly Ile Asp Thr Glu Lys Tyr His
 210 215 220

Gly Asp Leu Thr Ser Ile Asp Ile Ile Ser Pro Asn Gly Gly Lys Thr
 225 230 235 240

Phe Thr Glu Phe Ala Val Asn Leu Tyr Ser Val Gln Ala Thr Ser Pro
 245 250 255

Ser Gly Thr Asp Thr Leu Ser Thr Ser Glu Asp Thr Leu Ile Ala Val
 260 265 270

ES 2 711 165 T3

Leu Asp Ser Gly Thr Thr Leu Thr Tyr Leu Pro Gln Asp Met Ala Glu
275 280 285

Glu Ala Trp Asn Glu Val Gly Ala Glu Tyr Ser Asn Glu Leu Gly Leu
290 295 300

Ala Val Val Pro Cys Ser Val Gly Asn Thr Asn Gly Phe Phe Ser Phe
305 310 315 320

Thr Phe Ala Gly Thr Asp Gly Pro Thr Ile Asn Val Thr Leu Ser Glu
325 330 335

Leu Val Leu Asp Leu Phe Ser Gly Gly Pro Ala Pro Arg Phe Ser Ser
340 345 350

Gly Pro Asn Lys Gly Gln Ser Ile Cys Glu Phe Gly Ile Gln Asn Gly
355 360 365

Thr Gly Ser Pro Phe Leu Leu Gly Asp Thr Phe Leu Arg Ser Ala Phe
370 375 380

Val Val Tyr Asp Leu Val Asn Asn Gln Ile Ala Ile Ala Pro Thr Asn
385 390 395 400

Phe Asn Ser Thr Arg Thr Asn Val Val Ala Phe Ala Ser Ser Gly Ala
405 410 415

Pro Ile Pro Ser Ala Thr Ala Ala Pro Asn Gln Ser Arg Thr Gly His
420 425 430

Ser Ser Ser Thr His Ser Gly Leu Ser Ala Ala Ser Gly Phe His Asp
435 440 445

Gly Asp Asp Glu Asn Ala Gly Ser Leu Thr Ser Val Phe Ser Gly Pro
450 455 460

Gly Met Ala Val Val Gly Met Thr Ile Cys Tyr Thr Leu Leu Gly Ser
465 470 475 480

Ala Ile Phe Gly Ile Gly Trp Leu
485

<210> 86

<211> 761

<212> PRT

<213> Trichoderma reesei

<400> 86

5

ES 2 711 165 T3

Met Arg Ser Thr Leu Tyr Gly Leu Ala Ala Leu Pro Leu Ala Ala Gln
1 5 10 15

Ala Leu Glu Phe Ile Asp Asp Thr Val Ala Gln Gln Asn Gly Ile Met
20 25 30

Arg Tyr Thr Leu Thr Thr Thr Lys Gly Ala Thr Ser Lys His Leu His
35 40 45

Arg Arg Gln Asp Ser Ala Asp Leu Met Ser Gln Gln Thr Gly Tyr Phe
50 55 60

Tyr Ser Ile Gln Leu Glu Ile Gly Thr Pro Pro Gln Ala Val Ser Val
65 70 75 80

Asn Phe Asp Thr Gly Ser Ser Glu Leu Trp Val Asn Pro Val Cys Ser
85 90 95

Lys Ala Thr Asp Pro Ala Phe Cys Lys Thr Phe Gly Gln Tyr Asn His
100 105 110

Ser Thr Thr Phe Val Asp Ala Lys Ala Pro Gly Gly Ile Lys Tyr Gly
115 120 125

Thr Gly Phe Val Asp Phe Asn Tyr Gly Tyr Asp Tyr Val Gln Leu Gly
130 135 140

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

Phe Ala Ser Val Gly Ile Leu Gly Ala Gly Pro Asp Leu Ser Gly Trp
165 170 175

Thr Ser Pro Tyr Pro Phe Val Ile Asp Asn Leu Val Lys Gln Gly Phe
180 185 190

Ile Lys Ser Arg Ala Phe Ser Leu Asp Ile Arg Gly Leu Asp Ser Asp
195 200 205

Arg Gly Ser Val Thr Tyr Gly Gly Ile Asp Ile Lys Lys Phe Ser Gly
210 215 220

Pro Leu Ala Lys Lys Pro Ile Ile Pro Ala Ala Gln Ser Pro Asp Gly
225 230 235 240

Tyr Thr Arg Tyr Trp Val His Met Asp Gly Met Ser Ile Thr Lys Glu
245 250 255

ES 2 711 165 T3

Asp Gly Ser Lys Phe Glu Ile Phe Asp Lys Pro Asn Gly Gln Pro Val
 260 265 270
 Leu Leu Asp Ser Gly Tyr Thr Val Ser Thr Leu Pro Gly Pro Leu Met
 275 280 285
 Asp Lys Ile Leu Glu Ala Phe Pro Ser Ala Arg Leu Glu Ser Thr Ser
 290 295 300
 Gly Asp Tyr Ile Val Asp Cys Asp Ile Ile Asp Thr Pro Gly Arg Val
 305 310 315 320
 Asn Phe Lys Phe Gly Asn Val Val Val Asp Val Glu Tyr Lys Asp Phe
 325 330 335
 Ile Trp Gln Gln Pro Asp Leu Gly Ile Cys Lys Leu Gly Val Ser Gln
 340 345 350
 Asp Asp Asn Phe Pro Val Leu Gly Asp Thr Phe Leu Arg Ala Ala Tyr
 355 360 365
 Val Val Phe Asp Trp Asp Asn Gln Glu Val His Ile Ala Ala Asn Glu
 370 375 380
 Asp Cys Gly Asp Glu Leu Ile Pro Ile Gly Ser Gly Pro Asp Ala Ile
 385 390 395 400
 Pro Ala Ser Ala Ile Gly Lys Cys Ser Pro Ser Val Lys Thr Asp Thr
 405 410 415
 Thr Thr Ser Val Ala Glu Thr Thr Ala Thr Ser Ala Ala Ala Ser Thr
 420 425 430
 Ser Glu Leu Ala Ala Thr Thr Ser Glu Ala Ala Thr Thr Ser Ser Glu
 435 440 445
 Ala Ala Thr Thr Ser Ala Ala Ala Glu Thr Thr Ser Val Pro Leu Asn
 450 455 460
 Thr Ala Pro Ala Thr Thr Gly Leu Leu Pro Thr Thr Ser His Arg Phe
 465 470 475 480
 Ser Asn Gly Thr Ala Pro Tyr Pro Ile Pro Ser Leu Ser Ser Val Ala
 485 490 495
 Ala Ala Ala Gly Ser Ser Thr Val Pro Ser Glu Ser Ser Thr Gly Ala

ES 2 711 165 T3

5

<210> 87
 <211> 526
 <212> PRT
 <213> Trichoderma reesei

<400> 87
 Met Arg Pro Asn Ser Val Leu Leu Ala Pro Leu Ala Leu Tyr Ala Ser
 1 5 10 15

 Gly Ala Leu Ala Phe Tyr Pro Tyr Thr Pro Pro Trp Leu Lys Glu Leu
 20 25 30

 Glu Glu His Asn Ala Gly Glu Ala Lys Arg Ser Ala Asp Asn Gly Leu
 35 40 45

 Thr Phe Asp Ile Lys Arg Arg Ala Ser Arg Arg Ala Pro Ala Ser Gln
 50 55 60

 Glu Glu Lys Ala Ala Trp Gln Ala Ala Leu Leu Ser His Lys Tyr Ser
 65 70 75 80

 Glu Ser Val Thr Pro Ser Pro Ser Pro Asp Thr Thr Leu Ser Lys Arg
 85 90 95

 Asp Asn Gln Phe Ser Ile Leu Lys Ala Val Asp Pro Asp Ala Pro Asn
 100 105 110

 Thr Ala Gly Leu Ala Gln Asp Gly Thr Asp Tyr Ser Tyr Phe Val Gln
 115 120 125

 Ala Ser Leu Gly Ser Lys Lys Thr Lys Leu Tyr Met Leu Leu Asp Thr
 130 135 140

 Gly Ala Gly Ser Ser Trp Val Met Gly Thr Asp Cys Val Ser Glu Ala
 145 150 155 160

 Cys Ser Leu His Asp Ser Phe Gly Pro Glu Asp Ser Asp Thr Leu Lys
 165 170 175

 Thr Ser Thr Lys Asp Phe Ser Ile Ala Tyr Gly Ser Gly Ala Val Ser
 180 185 190

 Gly Ser Leu Val Asn Asp Thr Ile Glu Val Ala Gly Met Ser Leu Thr
 195 200 205

ES 2 711 165 T3

Tyr Gln Phe Gly Leu Ala His Asn Thr Ser Ser Asp Phe Val His Phe
 210 215 220
 Ala Phe Asp Gly Ile Leu Gly Met Ser Met Asn Ser Gly Ala Asn Glu
 225 230 235 240
 Asn Phe Leu Ser Ala Leu Glu Gly Ala Gly Leu Leu Asp Lys Ser Ile
 245 250 255
 Phe Ser Val Ala Leu Ala Arg Ala Ser Asp Gly His Asn Asp Gly Glu
 260 265 270
 Val Thr Phe Gly Ala Thr Asn Pro Ser Arg Tyr Thr Gly Asp Ile Thr
 275 280 285
 Tyr Thr Pro Ile Pro Ser Gly Thr Asp Trp Ser Ile Pro Leu Asp Asp
 290 295 300
 Met Ser Tyr Asn Gly Lys Lys Gly Asn Val Gly Gly Ile Asn Ala Tyr
 305 310 315 320
 Ile Asp Thr Gly Thr Ser Tyr Met Phe Gly Pro Ser Lys Asn Val Lys
 325 330 335
 Ala Leu His Ala Val Ile Asp Gly Ala Lys Ser Ser Asp Gly Ile Thr
 340 345 350
 Trp Thr Val Pro Cys Asp Thr Thr Thr Pro Leu Val Val Thr Phe Ser
 355 360 365
 Gly Val Asp Phe Ala Ile Ser Pro Lys Asp Trp Ile Ser Pro Lys Asp
 370 375 380
 Ser Ser Gly Lys Cys Thr Ser Asn Val Tyr Gly Tyr Glu Val Val Ser
 385 390 395 400
 Gly Ser Trp Leu Phe Gly Asp Thr Phe Leu Lys Asn Val Tyr Ala Val
 405 410 415
 Phe Asp Lys Glu Gln Met Arg Ile Gly Lys Thr Ser Pro Arg Ala Thr
 420 425 430
 Ser Pro Ser Ser Pro Ala Pro Thr Arg Thr Pro Ser Pro Ala Thr Thr
 435 440 445
 Ser Pro Ser Ser Ala Ser Thr Pro Gly Ser Thr Pro Thr Thr Ser Ser
 450 455 460

ES 2 711 165 T3

Thr Arg Thr Ala Arg Pro Ser Thr Ser Ala Pro Ser Gly Thr Ser Ser
465 470 475 480

Thr Gly Ala Pro Ser Pro Ser Ala Ser Ala Asn Arg Asp Val Leu Arg
485 490 495

Ala Lys Arg Ile Asn Met Leu Lys Ser Ile Ser Ser Phe Trp His Asp
500 505 510

Pro Cys Cys Cys Leu Phe Leu His Val Ser Ile Ser Ser Thr
515 520 525

<210> 88
<211> 2559
<212> ADN
<213> Leishmania mexicana

<400> 88
atggggaaaa ataaggcaaa ttcagtggcc gactccggct ctgctggcaac cgcacctcgt 60
gaagctcctg cccaagccaa agatgccgcc ccacaagccc agaccgcatc tccaccgcct 120
aagaagactt tggtgcccaa aacgctaaca gatgagacgg aatttgctcg catccttccg 180
ttccctttct ggccagtacg gttcgtcgtt acggtggtgg cactcttcg cttaggcgcc 240
agctgcctcc aagccttcac ggctcgcgatg acctcggtta agatttacgg atacctgatc 300
cacgagttcg acccgtggtt caactaccgc gctgccgagt acatgtccac gcacggctgg 360
tccgccttct tcagctggtt cgactacatg agctggtacc cgctgggccc ccccgctggc 420
tccaccacgt acccgggcct gcagttcact gccgtcgcca ttcaccgcgc actggcggct 480
gccggcatcc cgatgtctct caacgacgtg tgtgtgctga tcccggcgtg gtttgcgccc 540
atcgetaccg ctcttctggc tctttgcacg tacgaagcca gtgggtcgac ggtggcggcc 600
gccgctgccg ccctctcctt ctccatcatc ccagcccacc tgatgcggtc catggcgggt 660
gagttcgaca acgagtgcat cgccgtcgcc gccatgctgc tcaccttcta ctgctgggtg 720
cgctcgctgc gcacgcggtc ctctggccc atcggcgtcc tcaccggtgt cgcctacggc 780
tacatggtgg cggcgtgggg cggctacatt ttcgtgctca acatggttgc catgcatgcc 840
ggcatatcat cgatggtgga ctgggccgc aacacgtaca acccgtcgct gctgcgtgca 900
tacacgctgt tctacgttgt cggcaccgcc atcgcctgtg gcgtgccgcc agtggggatg 960
tcgcccttca agtcgctgga gcagctgggt gcgctgctgg tgcttctct cctgtgcggg 1020
ctgcaggtgt gcgaggtgct gcgggcacgc gccggtgtg agttctgctc tcgcgcgaac 1080
ttcaagatcc gcgcgcgcgt cttcagcgcg atggctggcg gggctgcgct tgcaatcgcg 1140
ctgctggcac cgagggggta cttcgggccc ctttcggctc gtgtgcgtgc gctgttcgtg 1200

5

10

ES 2 711 165 T3

gagcacacgc gcaactggcaa tccgctggtc gactcggtcg ccgaacatca acccgccagc 1260
 cctgaggcaa tgtggtcggt tcttcacgtg tgcggcgtga catggggctt gggcttcatt 1320
 gtgcttgctg tctcaacggt cgtgaactac tccccgtcga aggtcttctg ggtactgaac 1380
 tctggtgccg tgtactactt cagcaaccgc atggctcggc tgctgcttct ctccggtccc 1440
 gctgcgtgtc tgtccactgg cattttcgtg ggggcaattc tggaagcagc ggtgcagctc 1500
 agcttttggg acagtgatgc gacaaaggcc aagccccaga agcagacca acgccaccag 1560
 aggggggctc gtaaggacaa caagcgaaat gacgctgaga gcggaatgac cgcgctctca 1620
 ctttgcgaca tcgtgtccgg tagctctctg gcttggggcc atcgtatggt gctgtgcatc 1680
 gctatgtggg ctctcgtgac gacaaccgtg gtgacctca tcagttccgg tttcgcgtcc 1740
 cactactaa aatttgccga gcagtcgtca aatccgatga ttgttttcgc ggcctccgtg 1800
 ccaaaccgtg caacaggcaa gcctatgatg atattggtgg atgactacct gcacagctat 1860
 ctctggctgc gcgataaac acccaggagt gcgcgcattt tggcctggtg ggactacggc 1920
 taccagatca caggcatcgg caaccgcacc tcgctggccg atggcaacac ctggaaccac 1980
 gagcacatcg ccaccatcgg caagatggtg acgtcgcccg tggcggaggc gcactcgtg 2040
 gtgcgccaca tggccgacta cgtcctcctc tgggctgggc agagcggaga cttgatgaag 2100
 tcaccgcaca tggcgcgcat cggcaacagt gtgtaccacg acatctgccc caacgacccg 2160
 ctgtgccagc aattcggctt ttacagaaat gattaccatc gtccaacacc gatgatgctg 2220
 gcgtcgtgc tgtacaacct gcacgaggcc gggaaaacag cggcctgaa ggtggacca 2280
 tccctctttc aggaggtgta ctctccaag tacggcctgg tgcgcatctt caaggtcatg 2340
 aacgtgagcg cggagagcaa gaagtgggtt gctgaccggc caaacccgct gtgccgccc 2400
 cctgggtcgt ggatctgccc cgggcagtac ccgccggcga aggagatcca ggagatgctg 2460
 gcacaccggg tctccttcga tcaggtggac aaggacaaga agcgaaggc gacgtaccac 2520
 gaggagtaca tgcgccgat gcgtgaaaac gagatctga 2559

<210> 89
 <211> 852
 <212> PRT
 <213> Leishmania mexicana

<400> 89
 Met Gly Lys Asn Lys Ala Asn Ser Val Ala Asp Ser Gly Ser Ala Ala
 1 5 10 15
 Thr Ala Pro Arg Glu Ala Pro Ala Gln Ala Lys Asp Ala Ala Pro Gln
 20 25 30

Ala Gln Thr Ala Ser Pro Pro Pro Lys Lys Thr Leu Leu Pro Lys Thr

5

10

ES 2 711 165 T3

Ala Arg Asn Thr Tyr Asn Pro Ser Leu Leu Arg Ala Tyr Thr Leu Phe
 290 295 300

Tyr Val Val Gly Thr Ala Ile Ala Val Cys Val Pro Pro Val Gly Met
 305 310 315 320

Ser Pro Phe Lys Ser Leu Glu Gln Leu Gly Ala Leu Leu Val Leu Val
 325 330 335

Phe Leu Cys Gly Leu Gln Val Cys Glu Val Leu Arg Ala Arg Ala Gly
 340 345 350

Val Glu Val Arg Ser Arg Ala Asn Phe Lys Ile Arg Ala Arg Val Phe
 355 360 365

Ser Ala Met Ala Gly Gly Ala Ala Leu Ala Ile Ala Leu Leu Ala Pro
 370 375 380

Arg Gly Tyr Phe Gly Pro Leu Ser Ala Arg Val Arg Ala Leu Phe Val
 385 390 395 400

Glu His Thr Arg Thr Gly Asn Pro Leu Val Asp Ser Val Ala Glu His
 405 410 415

Gln Pro Ala Ser Pro Glu Ala Met Trp Ser Phe Leu His Val Cys Gly
 420 425 430

Val Thr Trp Gly Leu Gly Phe Ile Val Leu Ala Val Ser Thr Phe Val
 435 440 445

Asn Tyr Ser Pro Ser Lys Val Phe Trp Val Leu Asn Ser Gly Ala Val
 450 455 460

Tyr Tyr Phe Ser Thr Arg Met Ala Arg Leu Leu Leu Leu Ser Gly Pro
 465 470 475 480

Ala Ala Cys Leu Ser Thr Gly Ile Phe Val Gly Ala Ile Leu Glu Ala
 485 490 495

Ala Val Gln Leu Ser Phe Trp Asp Ser Asp Ala Thr Lys Ala Lys Pro
 500 505 510

Gln Lys Gln Thr Gln Arg His Gln Arg Gly Ala Arg Lys Asp Asn Lys
 515 520 525

Arg Asn Asp Ala Glu Ser Gly Met Thr Ala Leu Ser Leu Cys Asp Ile
 530 535 540

ES 2 711 165 T3

Val Ser Gly Ser Ser Leu Ala Trp Gly His Arg Met Val Leu Cys Ile
545 550 555 560

Ala Met Trp Ala Leu Val Thr Thr Thr Val Val Thr Phe Ile Ser Ser
565 570 575

Gly Phe Ala Ser His Ser Leu Lys Phe Ala Glu Gln Ser Ser Asn Pro
580 585 590

Met Ile Val Phe Ala Ala Ser Val Pro Asn Arg Ala Thr Gly Lys Pro
595 600 605

Met Met Ile Leu Val Asp Asp Tyr Leu His Ser Tyr Leu Trp Leu Arg
610 615 620

Asp Asn Thr Pro Arg Ser Ala Arg Ile Leu Ala Trp Trp Asp Tyr Gly
625 630 635 640

Tyr Gln Ile Thr Gly Ile Gly Asn Arg Thr Ser Leu Ala Asp Gly Asn
645 650 655

Thr Trp Asn His Glu His Ile Ala Thr Ile Gly Lys Met Leu Thr Ser
660 665 670

Pro Val Ala Glu Ala His Ser Leu Val Arg His Met Ala Asp Tyr Val
675 680 685

Leu Ile Trp Ala Gly Gln Ser Gly Asp Leu Met Lys Ser Pro His Met
690 695 700

Ala Arg Ile Gly Asn Ser Val Tyr His Asp Ile Cys Pro Asn Asp Pro
705 710 715 720

Leu Cys Gln Gln Phe Gly Phe Tyr Arg Asn Asp Tyr His Arg Pro Thr
725 730 735

Pro Met Met Arg Ala Ser Leu Leu Tyr Asn Leu His Glu Ala Gly Lys
740 745 750

Thr Ala Ala Val Lys Val Asp Pro Ser Leu Phe Gln Glu Val Tyr Ser
755 760 765

Ser Lys Tyr Gly Leu Val Arg Ile Phe Lys Val Met Asn Val Ser Ala
770 775 780

Glu Ser Lys Lys Trp Val Ala Asp Pro Ala Asn Arg Val Cys Arg Pro
785 790 795 800

ES 2 711 165 T3

Pro Gly Ser Trp Ile Cys Pro Gly Gln Tyr Pro Pro Ala Lys Glu Ile
 805 810 815

Gln Glu Met Leu Ala His Arg Val Ser Phe Asp Gln Val Asp Lys Asp
 820 825 830

Lys Lys Arg Lys Ala Thr Tyr His Glu Glu Tyr Met Arg Arg Met Arg
 835 840 845

Glu Asn Glu Ile
 850

<210> 90

<211> 2565

<212> ADN

<213> Leishmania braziliensis

<400> 90

atgggtaaga agaaagcaat tccgtcgggc agcgtcggcc ctgcgacaac cacctcccgt 60
 gaagctccag gcaaagacga aggtgcctcc caaccgcga agactgcagc tctgccggtg 120
 aagccctttg tgttgcccaa cacgctgaca gacgaggagg agtttggttg catccttccc 180
 tgccctttct ggccagtgcg atttgtcatc acagtgatgg cactcgtcct cttgggtgcc 240
 agctgtatcc ggcgcttcac gattcgcgatg ctatccgttc agctttatgg ctacatcatc 300
 cacgagttcg acccgtgggt caactaccgc gccgcccagt acatgtccgc gcacggctgg 360
 tccgccttct tcagctgggt cgactacatg agctggtacc cgctgggccc ccccggtggc 420
 accaccacgt acccgggcct gcagctcacc gccgttgcca tccaccgcgc attggcggct 480
 gccggggtgc cgatgtctct caacaacgtg tgcgtgctga tccccgcgtg gtatggtgcc 540
 atcgtactg ctatcctggc ctttgcgct tacgaggta gtaggtcaat ggtagcggcg 600
 gctgttgctg cactctcatt ctccatcatt ccagcacacc tgatgcggtc catggcgggc 660
 gagttcgaca acgagtgcat cgccgttgca gccatgctcc tcaccttcta cttgtgggta 720
 cgctcgtgc gcacgcggtg ctctggccc atcgcatcc tcaccggtat cgcctacggc 780
 tacatggtgg cggcgtgggg cggatacatt tttgtgctca acatggttgc catgcacgcc 840
 ggcatatcat cgatggtcga ctgggctcgc aacacgtaca acccgtcgt gctgcgcgca 900
 tacgcgctgt tctacgttgt cggcaccgcc atcgccacgc gcgtgccgcc tgtgggatg 960
 tcgcccttca ggtcgtgga gcagctgggt gcgctggcgg tgctcctct cctgtgcggg 1020
 ctgcaggcct gcgaggtgt tcgcgcacgg gccgacgtg agttcgtc ccgcgcgaac 1080
 ttcaagatcc gcatgcgtgc cttcagcgtg atggctggcg tgggtgcgt tgcaatcgcg 1140
 gtgctgtcgc cgaccgggta ctttgcccc ctcacggctc gtgtgcgtgc gctgttcatg 1200

5

10

ES 2 711 165 T3

gagcacacgc gcaactggcaa tccgctggtc gactcggtcg ctgagcacca ccccgccagt 1260
 cctgaggcga tgtggacatt tcttcacgtg tgcggcgtga cttggggttt gggctccatt 1320
 gttcttcttg tgtcgttgct ggtggactac tctcggcaa agctcttttg gctgatgaac 1380
 tctggtgccg tgtactatth cagcaccgcg atgtcacgac tgctgcttct cacgggcccc 1440
 gctgcgtgtc tgtccactgg ctgtttcgtg gggacattac tggaaagcggc gatacagttc 1500
 accttctggt ccagcgtatg acaaaggcc aaaaaacagc aagagacaca acttcaccaa 1560
 aagggcgcgc gcaagcatag cgaccggagt aactctaaga atgcaactgac tgtgcgtaaca 1620
 ttgggcgacg tcttgaggag tacctctctg gcatggggtc atcgcatggt gctctgcttc 1680
 gctatgtggg ctcttggtat tacagtcgcg gtgtgcctct tgggttccga tttcacttcc 1740
 catgcaacga tgtttgcaag gcagacgtcg aaccgcgtga ttgtctttgc aaccgtgctg 1800
 cgagaccgcg ctaccggcaa gccaacacag gtattggtgg atgactacct gcgcagctat 1860
 ctctggctgc gcgacaacac gccagaaat gcgcgcgtgc tgtcctggtg ggactacggc 1920
 taccagatca caggtatcgg caaccgcacc tcgctggccg atggcaacac ctggaaccac 1980
 gagcacatcg ccaccatcgg caagatgctg acgtcgcgagc tggcggaggc gcactcactg 2040
 gtgcgccaca tggcggacta cgtcctcctc tgggctgggc agggcggaga cttgatgaag 2100
 tcgccgcaca tggcgcgcat tggcaacagc gtgtaccacg acatctgccc caacgacccg 2160
 ctttgccagc atttcggctt ttacaagaac gatcgcaatc gcccaaaacc gatgatgctc 2220
 gcgtcgtgct tgtacaacct gcacgaggcc ggacgaagcg cgggtgtgaa ggtggacccg 2280
 tccctctttc aggaagtgta ctcatccaag tacggcctgg tgcgcatctt caaggtcatg 2340
 aacgtgagcg cggagagcaa gaagtgggtg gctgaccgag caaacgcgct gtgccacccg 2400
 cctgggtcgt ggatctgccc cggcagctac ccgccggcga aggagatcca ggagatgctg 2460
 gcgcaccgcg tcccctttga ccatgtgaac agcttcagtc ggaaaaggc cgggtcttat 2520
 catgaagaat acatgcgccc gatgcgtgaa gagcaggacc gatga 2565

<210> 91
 <211> 854
 <212> PRT
 <213> Leishmania braziliensis

<400> 91
 Met Gly Lys Lys Lys Ala Ile Pro Ser Gly Ser Val Gly Pro Ala Thr
 1 5 10 15

Thr Thr Ser Arg Glu Ala Pro Gly Lys Asp Glu Gly Ala Ser Gln Pro
 20 25 30

5

10

ES 2 711 165 T3

Ala Lys Thr Ala Ala Leu Pro Val Lys Pro Phe Val Leu Pro Asn Thr
 35 40 45

Leu Thr Asp Glu Glu Glu Phe Val Gly Ile Phe Pro Cys Pro Phe Trp
 50 55 60

Pro Val Arg Phe Val Ile Thr Val Met Ala Leu Val Leu Leu Gly Ala
 65 70 75 80

Ser Cys Ile Arg Ala Phe Thr Ile Arg Met Leu Ser Val Gln Leu Tyr
 85 90 95

Gly Tyr Ile Ile His Glu Phe Asp Pro Trp Phe Asn Tyr Arg Ala Ala
 100 105 110

Glu Tyr Met Ser Ala His Gly Trp Ser Ala Phe Phe Ser Trp Phe Asp
 115 120 125

Tyr Met Ser Trp Tyr Pro Leu Gly Arg Pro Val Gly Thr Thr Thr Tyr
 130 135 140

Pro Gly Leu Gln Leu Thr Ala Val Ala Ile His Arg Ala Leu Ala Ala
 145 150 155 160

Ala Gly Val Pro Met Ser Leu Asn Asn Val Cys Val Leu Ile Pro Ala
 165 170 175

Trp Tyr Gly Ala Ile Ala Thr Ala Ile Leu Ala Leu Cys Ala Tyr Glu
 180 185 190

Val Ser Arg Ser Met Val Ala Ala Val Ala Ala Leu Ser Phe Ser
 195 200 205

Ile Ile Pro Ala His Leu Met Arg Ser Met Ala Gly Glu Phe Asp Asn
 210 215 220

Glu Cys Ile Ala Val Ala Ala Met Leu Leu Thr Phe Tyr Leu Trp Val
 225 230 235 240

Arg Ser Leu Arg Thr Arg Cys Ser Trp Pro Ile Gly Ile Leu Thr Gly
 245 250 255

Ile Ala Tyr Gly Tyr Met Val Ala Ala Trp Gly Gly Tyr Ile Phe Val
 260 265 270

Leu Asn Met Val Ala Met His Ala Gly Ile Ser Ser Met Val Asp Trp
 275 280 285

ES 2 711 165 T3

Ala Arg Asn Thr Tyr Asn Pro Ser Leu Leu Arg Ala Tyr Ala Leu Phe
 290 295 300

Tyr Val Val Gly Thr Ala Ile Ala Thr Arg Val Pro Pro Val Gly Met
 305 310 315 320

Ser Pro Phe Arg Ser Leu Glu Gln Leu Gly Ala Leu Ala Val Leu Leu
 325 330 335

Phe Leu Cys Gly Leu Gln Ala Cys Glu Val Phe Arg Ala Arg Ala Asp
 340 345 350

Val Glu Val Arg Ser Arg Ala Asn Phe Lys Ile Arg Met Arg Ala Phe
 355 360 365

Ser Val Met Ala Gly Val Gly Ala Leu Ala Ile Ala Val Leu Ser Pro
 370 375 380

Thr Gly Tyr Phe Gly Pro Leu Thr Ala Arg Val Arg Ala Leu Phe Met
 385 390 395 400

Glu His Thr Arg Thr Gly Asn Pro Leu Val Asp Ser Val Ala Glu His
 405 410 415

His Pro Ala Ser Pro Glu Ala Met Trp Thr Phe Leu His Val Cys Gly
 420 425 430

Val Thr Trp Gly Leu Gly Ser Ile Val Leu Leu Val Ser Leu Leu Val
 435 440 445

Asp Tyr Ser Ser Ala Lys Leu Phe Trp Leu Met Asn Ser Gly Ala Val
 450 455 460

Tyr Tyr Phe Ser Thr Arg Met Ser Arg Leu Leu Leu Leu Thr Gly Pro
 465 470 475 480

Ala Ala Cys Leu Ser Thr Gly Cys Phe Val Gly Thr Leu Leu Glu Ala
 485 490 495

Ala Ile Gln Phe Thr Phe Trp Ser Ser Asp Ala Thr Lys Ala Lys Lys
 500 505 510

Gln Gln Glu Thr Gln Leu His Gln Lys Gly Ala Arg Lys His Ser Asp
 515 520 525

Arg Ser Asn Ser Lys Asn Ala Leu Thr Val Arg Thr Leu Gly Asp Val
 530 535 540

ES 2 711 165 T3

Leu Arg Ser Thr Ser Leu Ala Trp Gly His Arg Met Val Leu Cys Phe
 545 550 555 560

Ala Met Trp Ala Leu Val Ile Thr Val Ala Val Cys Leu Leu Gly Ser
 565 570 575

Asp Phe Thr Ser His Ala Thr Met Phe Ala Arg Gln Thr Ser Asn Pro
 580 585 590

Leu Ile Val Phe Ala Thr Val Leu Arg Asp Arg Ala Thr Gly Lys Pro
 595 600 605

Thr Gln Val Leu Val Asp Asp Tyr Leu Arg Ser Tyr Leu Trp Leu Arg
 610 615 620

Asp Asn Thr Pro Arg Asn Ala Arg Val Leu Ser Trp Trp Asp Tyr Gly
 625 630 635 640

Tyr Gln Ile Thr Gly Ile Gly Asn Arg Thr Ser Leu Ala Asp Gly Asn
 645 650 655

Thr Trp Asn His Glu His Ile Ala Thr Ile Gly Lys Met Leu Thr Ser
 660 665 670

Pro Val Ala Glu Ala His Ser Leu Val Arg His Met Ala Asp Tyr Val
 675 680 685

Leu Ile Trp Ala Gly Gln Gly Gly Asp Leu Met Lys Ser Pro His Met
 690 695 700

Ala Arg Ile Gly Asn Ser Val Tyr His Asp Ile Cys Pro Asn Asp Pro
 705 710 715 720

Leu Cys Gln His Phe Gly Phe Tyr Lys Asn Asp Arg Asn Arg Pro Lys
 725 730 735

Pro Met Met Arg Ala Ser Leu Leu Tyr Asn Leu His Glu Ala Gly Arg
 740 745 750

Ser Ala Gly Val Lys Val Asp Pro Ser Leu Phe Gln Glu Val Tyr Ser
 755 760 765

Ser Lys Tyr Gly Leu Val Arg Ile Phe Lys Val Met Asn Val Ser Ala
 770 775 780

Glu Ser Lys Lys Trp Val Ala Asp Pro Ala Asn Arg Val Cys His Pro

REIVINDICACIONES

1. Una célula de *Trichoderma* o *Myceliophthora* que comprende
- 5 i. una o más mutaciones que reducen o eliminan una o más actividades de proteasa endógena en comparación con una célula de *Trichoderma* o *Myceliophthora* parental que no tiene dicha mutación o mutaciones,
- 10 ii. un polinucleótido que codifica una subunidad catalítica heteróloga de la oligosacaril transferasa, y
- 15 iii. un polinucleótido que codifica una glucoproteína heteróloga,
- en la que dicha subunidad catalítica de oligosacaril transferasa se selecciona de subunidades catalíticas de oligosacaril transferasa de *Leishmania*, en la que la ocupación del sitio de N-glucosilación de la glucosilación heteróloga expresada en la célula de *Trichoderma* o *Myceliophthora* es de al menos un 90%, y en la que dicha célula de *Trichoderma* o *Myceliophthora* tiene al menos una reducción de dos veces en la actividad de la proteasa total en
- 20 comparación con una célula de *Trichoderma* o *Myceliophthora* parental que no tiene la actividad de proteasa reduciendo o eliminando la mutación o mutaciones, y en la que la glucoproteína heteróloga es una glucoproteína de mamífero y en la que el rendimiento de la glucoproteína de mamífero es de al menos 1 gramo por litro.
2. La célula de *Trichoderma* o *Myceliophthora* de la reivindicación 1, en la que dicho polinucleótido que
- 25 codifica la subunidad catalítica heteróloga de oligosacaril transferasa comprende un ácido nucleico seleccionado del grupo que consiste en la SEQ ID NO: 2, SEQ ID NO: 9, SEQ ID NO: 88 y SEQ ID NO: 90 o un polinucleótido que codifica un polipéptido variante funcional que tiene al menos un 50%, al menos un 60%, al menos un 70% de identidad, al menos un 80% de identidad, al menos un 90% de identidad, o al menos un 95% de identidad con la SEQ ID NO: 1, SEQ ID NO: 8, SEQ ID NO: 89 o SEQ ID NO: 91, teniendo dicho polipéptido variante funcional
- 30 actividad de oligosacariltransferasa.
3. La célula de *Trichoderma* o *Myceliophthora* de una cualquiera de las reivindicaciones anteriores, en la que la ocupación del sitio de N-glucosilación de la glucoproteína heteróloga es de al menos el 95% y las glucoformas Man3, Man5, G0, G1 y/o G2 representan al menos el 50% de los N-glicanos neutros totales de la glucoproteína
- 35 heteróloga.
4. La célula de *Trichoderma* o *Myceliophthora* de una cualquiera de las reivindicaciones anteriores, en la que dicha célula es, por ejemplo, *Trichoderma reesei*, y dicha célula comprende mutaciones que reducen o eliminan la actividad de
- 40
- las tres proteasas endógenas pep1, tsp1, y slp1;
 - las tres proteasas endógenas gap1, slp1, y pep1;
 - las tres proteasas endógenas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, pep8, pep9, pep11, pep12, tsp1, slp1, slp2, slp3, slp7, gap1 y gap2;
 - de tres a seis proteasas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, tsp1, slp1, slp2, slp3, gap1 y gap2; o,
 - de siete a diez proteasas seleccionadas del grupo que consiste en pep1, pep2, pep3, pep4, pep5, pep7, pep8, tsp1, slp1, slp2, slp3, slp5, slp6, slp7, slp8, tpp1, gap1 y gap2.
- 45 5. La célula de *Trichoderma* o *Myceliophthora* de una cualquiera de las reivindicaciones anteriores, en la que la célula de *Trichoderma* comprende además una mutación en el gen que codifica ALG3 que reduce o elimina la expresión de ALG3 correspondiente en comparación con el nivel de expresión del gen ALG3 en una célula parental que no tiene tal mutación.
- 50 6. La célula de *Trichoderma* o *Myceliophthora* de una cualquiera de las reivindicaciones anteriores, que comprende además un polinucleótido que codifica un dominio catalítico de N-acetilglucosaminiltransferasa I y un dominio catalítico de N-acetilglucosaminiltransferasa II.
7. La célula de *Trichoderma* o *Myceliophthora* de una cualquiera de las reivindicaciones anteriores, que
- 55 comprende además uno o más polinucleótidos que codifican un polipéptido seleccionado del grupo que consiste en:
- i. α 1, 2 manosidasa;
 - ii. dominio catalítico de N-acetilglucosaminiltransferasa I;
 - iii. α -manosidasa II;

- iv. dominio catalítico de N-acetilglucosaminiltransferasa II;
- v. β 1,4 galactosiltransferasa; y,
- vi. fucosiltransferasa.

5 8. Un método para producir una glucoproteína heteróloga con una mayor ocupación del sitio de N-glucosilación, que comprende

- 10 a) proporcionar una célula de *Trichoderma* o *Myceliophthora*, que tiene un gen STT3D de *Leishmania* que codifica una subunidad catalítica de oligosacaril transferasa, y un polinucleótido que codifica dicha glucoproteína heteróloga,
- b) cultivar la célula en condiciones apropiadas para la expresión del gen STT3D, y la producción de la glucoproteína heteróloga; y,
- c) recuperar y, opcionalmente, purificar la glucoproteína heteróloga,

15 en el que dicha célula de *Trichoderma* o *Myceliophthora* comprende una o más mutaciones que reducen o eliminan una o más actividades de proteasa endógena en comparación con una célula de *Trichoderma* o *Myceliophthora* parental que no tiene dicha mutación o mutaciones, y en el que la ocupación del sitio de N-glucosilación de la glucoproteína heteróloga expresada en la célula de *Trichoderma* o *Myceliophthora* es de al menos el 90 %, y en el que la glucoproteína heteróloga es una glucoproteína de mamífero, y en el que el rendimiento de la glucoproteína de
20 mamífero es de al menos 1 gramo por litro.

9. El método de la reivindicación 8, en el que dicha glucoproteína es un anticuerpo.

10. El método de una cualquiera de las reivindicaciones 8 a 9, en el que dicha célula huésped de
25 *Trichoderma* o *Myceliophthora* es la célula como se define en cualquiera de las reivindicaciones 1-7.

11. El método de una cualquiera de las reivindicaciones 8 a 10, en el que dicho gen STT3D de *Leishmania* que codifica una subunidad catalítica de oligosacaril transferasa comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en la SEQ ID NO:2, SEQ ID NO:9, SEQ ID NO: 88 y SEQ ID NO: 90 o un
30 polinucleótido que codifica un polipéptido variante funcional que tiene al menos un 50%, al menos un 60%, al menos un 70% de identidad, al menos un 80% de identidad, al menos un 90% de identidad, o al menos un 95% de identidad con la SEQ ID NO:1, SEQ ID NO:8, SEQ ID NO: 89 o SEQ ID NO: 91, teniendo dicho polipéptido variante funcional actividad de oligosacariltransferasa.

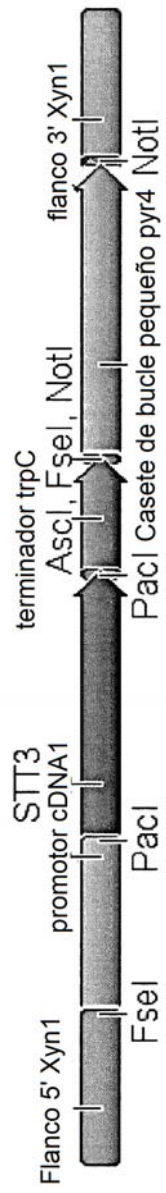


Figura 1

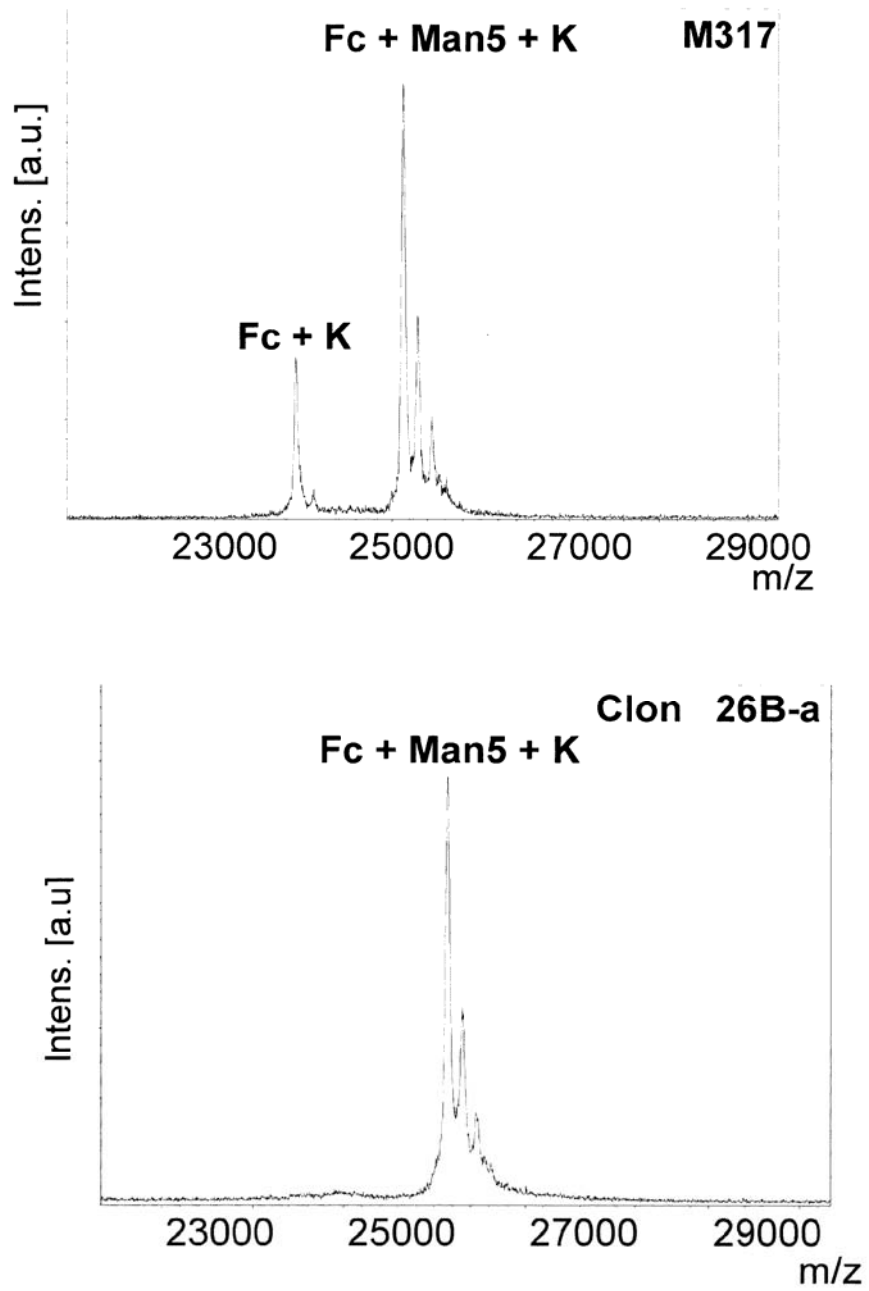


Figura 2

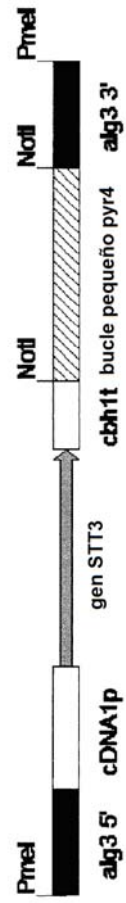


Figura 3

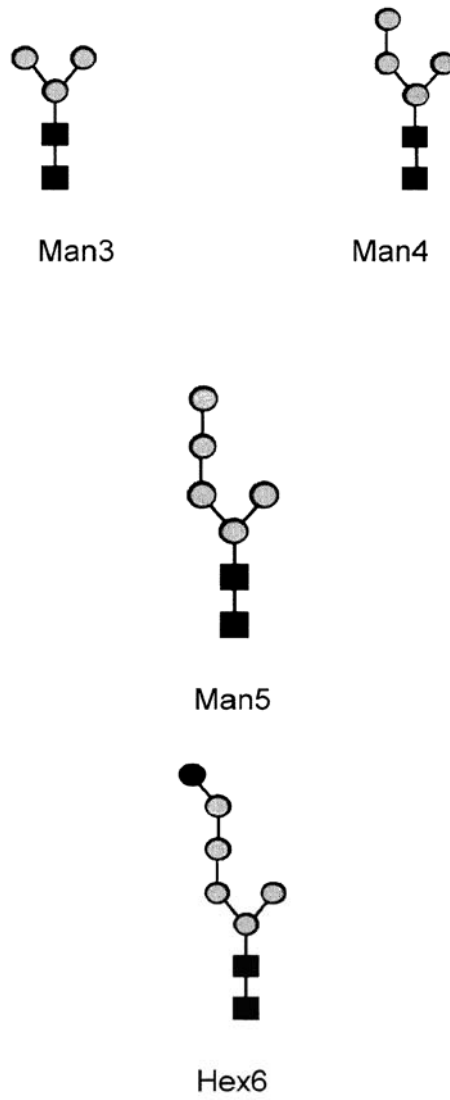


Figura 4

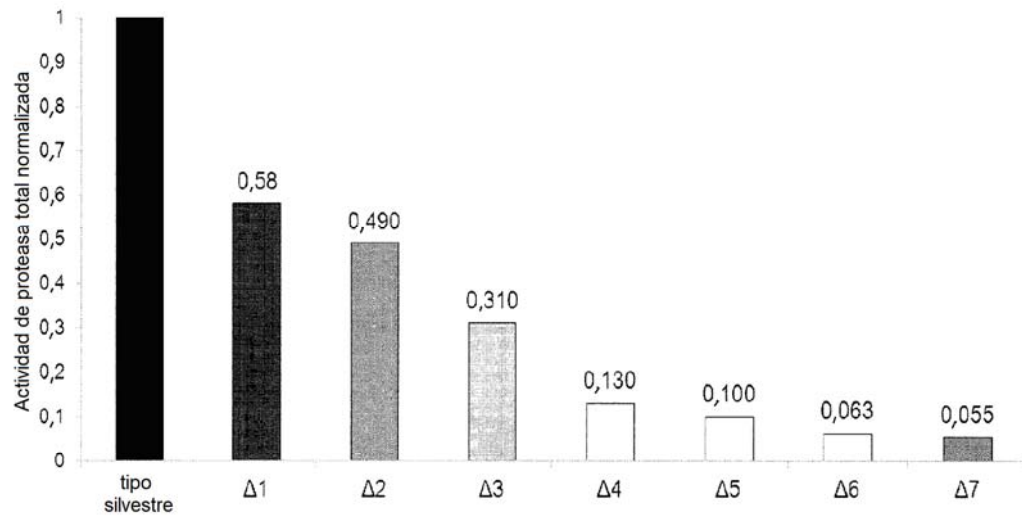


Figura 5