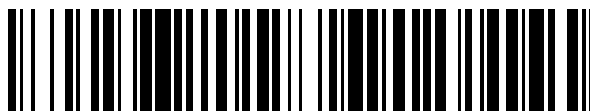


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 527 681**

51 Int. Cl.:

C12N 9/42 (2006.01)

C12N 15/56 (2006.01)

C12P 19/02 (2006.01)

C12P 7/06 (2006.01)

C12N 1/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.12.2006 E 10191152 (7)**

97 Fecha y número de publicación de la concesión europea: **03.12.2014 EP 2319920**

54 Título: **Tratamiento de material celulósico y enzimas útiles en el mismo**

30 Prioridad:

22.12.2005 FI 20051318

22.12.2005 US 753258 P

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

28.01.2015

73 Titular/es:

ROAL OY (100.0%)

Tykkimäentie 15

05200 Rajamäki, FI

72 Inventor/es:

VEHMAANPERÄ, JARI;

ALAPURANEN, MARIKA;

PURANEN, TERHI;

SIIKA-AHO, MATTI;

KALLIO, JARNO;

JÄMSÄ, SATU;

VOUTILAINEN, SANNI;

HALONEN, TEEMU y

VIIKARI, LIISA

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 527 681 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

Tratamiento de material celulósico y enzimas útiles en el mismo

Campo de la invención

La presente invención se refiere a la producción de productos hidrolizados de azúcares a partir de material celulósico. Más precisamente, la invención se refiere a la producción de azúcares fermentables a partir de material lignocelulósico mediante conversión enzimática. Los azúcares fermentables son, p. ej., útiles en la producción de bioetanol, o para otros fines. En particular, la invención se refiere a un método para el tratamiento de material celulósico con celobiohidrolasa, endoglucanasa, beta-glucosidasa, y opcionalmente xilanasa, y a preparaciones enzimáticas y a los usos de las mismas. La invención se refiere adicionalmente a nuevos polipéptidos celolíticos, los polinucleótidos que los codifican, y a vectores y células anfitrionas que contienen los polinucleótidos. También adicionalmente, la invención se refiere a los usos de los polipéptidos y a un método para su preparación.

Antecedentes de la invención

Los productos hidrolizados de azúcares se pueden utilizar para la producción microbiana de una variedad de sustancias químicas puras o biopolímeros, tales como ácidos orgánicos, por ejemplo ácido láctico, o etanol u otros alcoholes, por ejemplo n-butanol, 1,3-propanodiol, o polihidroxialcanoatos (PHA). Los productos hidrolizados de azúcares también pueden servir como materia prima para otros procesos no microbianos, por ejemplo, para el enriquecimiento, el aislamiento y la purificación de azúcares de alto valor o para diversos procedimientos de polimerización. Uno de los principales usos de los productos hidrolizados de azúcares es en la producción de biocombustibles. La producción de bioetanol y/u otros productos químicos puede tener lugar en un procedimiento integrado en una biorrefinería (Wyman 2001).

Los recursos limitados de combustibles fósiles, y las cantidades crecientes de CO₂ liberado de los mismos y que causan el fenómeno de efecto invernadero han planteado la necesidad de utilización de la biomasa como fuente de energía renovable y limpia. Una tecnología prometedora alternativa es la producción de biocombustibles, es decir, etanol a partir de materiales celulósicos. En el sector del transporte los biocombustibles son, por el momento la única opción, lo que podría reducir las emisiones de CO₂ en un orden de magnitud. El etanol puede utilizarse en vehículos y sistemas de distribución existentes y por lo tanto no requiere de inversiones en infraestructuras costosas. Los azúcares derivados de materias primas lignocelulósicas renovables se pueden utilizar también como materias primas para una variedad de productos químicos que pueden sustituir los productos químicos con una base oleosa.

La mayor parte de los carbohidratos en las plantas están en forma de lignocelulosa, que consiste esencialmente en celulosa, hemicelulosa, pectina y lignina. En un procedimiento de lignocelulosa a etanol el material lignocelulósico se pretrata en primer lugar químicamente o físicamente para hacer la fracción de celulosa más accesible a la hidrólisis. La fracción de celulosa es hidrolizada a continuación para obtener azúcares que pueden ser fermentados por levaduras a etanol. La lignina se obtiene en forma de un co-producto principal que se puede utilizar como combustible sólido.

Los costes de producción de bioetanol son altos y la producción de energía es baja, y hay una continua búsqueda de hacer el proceso más económico. La hidrólisis enzimática se considera la tecnología más prometedora para la conversión de biomasa celulósica en azúcares fermentables. Sin embargo, la hidrólisis enzimática se utiliza solo hasta una cantidad limitada a escala industrial, y especialmente cuando se utiliza material fuertemente lignificado como la madera o residuos agrícolas, la tecnología no es satisfactoria. El coste de la etapa enzimática es uno de los principales factores económicos del procedimiento. Se han realizado esfuerzos para mejorar la eficiencia de la hidrólisis enzimática del material celulósico (Badger 2002).

El documento US 2002/019 2774 A1 describe un procedimiento continuo para la conversión de biomasa lignocelulósica sólida en productos carburantes combustibles. Después del pretratamiento por oxidación húmeda o explosión de vapor de agua se separa la biomasa parcialmente en celulosa, hemicelulosa y lignina, y a continuación se somete a hidrólisis parcial utilizando una o más enzimas carbohidrasas (EC 3.2). Se proporciona como ejemplo Celluclast™, un producto comercial de Novo Nordisk A/S que contiene actividades celulasa y xilanasa.

El documento US 2004/000 5674 A1 describe nuevas mezclas de enzimas que se pueden utilizar directamente en el sustrato lignocelulósico, por medio de lo cual se pueden evitar los productos de desecho tóxicos formados durante los procedimientos de pretratamiento, y se puede ahorrar energía. La mezcla de enzimas sinérgica contiene una celulasa y una enzima auxiliar tal como celulasa, xilanasa, ligninasa, amilasa, proteasa, lipidasa o glucuronidasa, o cualquier combinación de las mismas. Se considera que la celulasa incluye endoglucanasa (EG), beta-glucosidasa (BG) y celobiohidrolasa (CBH). Los ejemplos ilustran el uso de una mezcla de preparaciones de xilanasa y celulasa de *Trichoderma*.

Kurabi *et al.* (2005) han investigado la hidrólisis enzimática de abeto de Douglas pretratado mediante explosión de vapor y organosolv con etanol mediante celulasas fúngicas novedosas y comerciales. Sometieron a ensayo dos

preparaciones de celulasa de *Trichoderma reesei* comerciales, y dos nuevas preparaciones producidas por cepas mutantes de *Trichoderma* sp. y *Penicillium* sp. La preparación de *Trichoderma* sp. mostró un funcionamiento significativamente mejor que las otras preparaciones. Se creía que el mejor funcionamiento era debido al menos en parte a una actividad beta-glucosidasa significativamente mayor, lo que alivia la inhibición por producto de celobiohidrolasa y endoglucanasa.

El documento US 2004/005 3373 A1 se refiere un método para convertir celulosa en glucosa mediante el tratamiento de un sustrato lignocelulósico pretratado con una mezcla de enzimas que comprende celulasa y una celobiohidrolasa I modificada (CBHI). La CBHI se ha modificado mediante la inactivación de su dominio de unión a celulosa (CBD). Las ventajas de la modificación de CBHI son, por ejemplo una mejor recuperación y una mayor tasa de hidrólisis con alta concentración de sustrato. La celulasa se selecciona del grupo que consiste en EG, CBH y BG. La CBHI se obtiene preferiblemente a partir de *Trichoderma*.

El documento US 2005/016 4355 A1 describe un método para degradar material lignocelulósico con una o más enzimas celulolíticas en presencia de al menos un tensioactivo. También se pueden utilizar enzimas adicionales, tales como hemicelulasas, esterasa, peroxidasa, proteasa, lacasa o mezclas de las mismas. La presencia del agente tensioactivo aumenta la degradación de material lignocelulósico en comparación con la ausencia de tensioactivo. Las enzimas celulolíticas pueden ser cualquier enzima implicada en la degradación de la lignocelulosa incluyendo CBH, EG y BG.

Existe un gran número de publicaciones que describen diversas celulastas y hemicelulasas.

Las celobiohidrolasas (CBHS) se describen por ejemplo en el documento WO 03/000 941, que se refiere a las enzimas CBHI obtenidas a partir de diversos hongos. No se proporcionan propiedades fisiológicas de las enzimas, ni ningún ejemplo de sus usos. Hong *et al.* (2003b) caracterizan CBHI de *Thermoascus aurantiacus* producida en levadura. No se describen aplicaciones de la enzima. Tuohy *et al.* (2002) describen tres formas de celobiohidrolasas de *Talaromyces emersonii*. El documento WO 2005/001065 describe celobiohidrolasas con mayor termoestabilidad, de *Humicola grisea* y *Scytalidium thermophilum* CBH1, que se utilizan junto con endoglucanasa y beta-glucosidasa en la hidrólisis de celulosa para obtener glucosa para la producción de etanol.

Las endoglucanasas de la familia cel5 (fam 5 de EG) se describen p. ej., en el documento WO 03/062 409, que se refiere a composiciones que comprenden al menos dos enzimas termoestables para su uso en aplicaciones de alimentación. Hong *et al.* (2003a) describen la producción de endo- β -1,4-glucanasa termoestable de *T. aurantiacus* en levaduras. No se explican las aplicaciones. El documento WO 01/70998 se refiere a β -glucanasa de *Talaromyces*. También describen β -glucanasas de *Talaromyces emersonii*. Se comentan las aplicaciones para alimentación, pienso, bebida, cerveza, y detergente. No se menciona la hidrólisis de lignocelulosa. El documento WO 98/06 858 describe beta-1,4-endoglucanasa de *Aspergillus niger* y comenta las aplicaciones en piensos y alimentos de la enzima. El documento WO 97/13853 describe métodos para escrutar fragmentos de ADN que codifican enzimas en bibliotecas de ADNc. La biblioteca de ADNc es de origen fúngico o de levadura, preferiblemente de *Aspergillus*. La enzima es preferiblemente una celulasa. Van Petegem *et al.* (2002) describen la estructura 3D de una endoglucanasa de la familia cel5 de *Thermoascus aurantiacus*. Parry *et al.* (2002) describen el modo de acción de una endoglucanasa de la familia cel5 de *Thermoascus aurantiacus*.

Las endoglucanasas de la familia cel7 (fam 7 de EG) se describen p. ej. en el documento US 5.912.157, que atañe a endoglucanasa de *Myceliphthora* y sus homólogos y aplicaciones de los mismos en detergentes, tejidos, y pasta de papel. El documento US 6.071.735 describe celulastas que exhiben alta actividad endoglucanasa en condiciones alcalinas. Se comentan los usos como detergente, en aplicaciones de pasta de papel y papel, y tejidos. No se menciona el bioetanol. El documento US 5.763.254 describe enzimas que degradan celulosa/hemicelulosa y que tienen residuos de aminoácidos conservados en el CBD.

Las endoglucanasas de la familia cel45 (fam 45 de EG) se describen por ejemplo, en el documento US 6.001.639, que se refiere a enzimas que tienen actividad endoglucanasa y que tienen dos secuencias de aminoácidos conservadas. Los usos en aplicaciones textiles, detergentes, y de pasta de papel y papel son comentados generalmente y se menciona el tratamiento de material lignocelulósico pero no se proporcionan ejemplos. El documento WO 2004/053039 se refiere a aplicaciones detergentes de endoglucanasas. El documento US 5.958.082 describe el uso de endoglucanasa, especialmente de *Thielavia terrestris* en aplicación textil. El documento EP 0495258 se refiere a composiciones detergentes que contienen celulasa de *Humicola*. El documento US 5.948.672 describe una preparación de celulasa que contiene endoglucanasa, especialmente de *Humicola* y su uso en aplicaciones textiles y de pasta de papel. No se menciona la hidrólisis de lignocelulosa.

Una pequeña cantidad de beta-glucosidasa (BG) mejora la hidrólisis de la biomasa a glucosa mediante hidrólisis de la celobiosa producida por celobiohidrolasas. La conversión de celobiosa a glucosa es por lo general la principal etapa limitante de la velocidad. Las beta-glucosidasas se describen, por ejemplo en el documento US 2005/021 4920, que se refiere a BG de *Aspergillus fumigatus*. La enzima ha sido producida en *Aspergillus oryzae* y *Trichoderma reesei*. Se comenta generalmente pero no se ilustra el uso de la enzima en aplicaciones de degradación de biomasa o detergentes. El documento WO 02/095 014 describe una enzima de *Aspergillus oryzae* que tiene actividad celobiasa. Se comenta generalmente pero no se ilustra el uso en la producción de etanol a partir

de biomasa. El documento WO 2005/074656 describe polipéptidos que tienen actividad potenciadora celulolítica derivados por ejemplo de *T. aurantiacus*; *A. fumigatus*; *T. terrestris* y *T. aurantiacus*. El documento WO 02/26979 describe el procesamiento enzimático de material vegetal. El documento US 6.022.725 describe la clonación y la amplificación del gen de la beta-glucosidasa de *Trichoderma reesei*, y el documento US 6.103.464 describe un método para la detección de ADN que codifica una beta-glucosidasa de un hongo filamentoso. No se proporcionan ejemplos de aplicación.

Las xilanasas se describen por ejemplo, en el documento FR2786784, que se refiere a una xilanasas estable al calor, útil, p. ej., en el tratamiento de piensos para animales y en la fabricación de pan. La enzima deriva de un hongo termófilo, particularmente del género *Thermoascus*.

El documento US 6.197.564 describe enzimas que tienen actividad xilanasas, y obtenidas a partir de *Aspergillus aculeatus*. Se ilustra su aplicación en panadería. El documento WO 02/24926 se refiere a xilanasas de *Talaromyces*. Se proporcionan ejemplos en alimentación y panadería. El documento WO 01/42433 describe xilanasas termoestables de *Talaromyces emersonii* para su uso en aplicaciones para alimentos y piensos.

Las enzimas celulolíticas mejor investigadas y más ampliamente aplicadas de origen fúngico se han obtenido de *Trichoderma reesei* (el anamorfo de *Hypocrea jecorina*). En consecuencia también la mayoría de las celulasas fúngicas disponibles comercialmente derivan de *Trichoderma reesei*. Sin embargo, la mayoría de las celulasas fúngicas menos conocidas no se han aplicado en procedimientos de importancia práctica tales como en la degradación de material celulósico, incluyendo lignocelulosa.

Existe una continua necesidad de nuevos métodos para degradar sustratos celulósicos, en particular sustratos lignocelulósicos, y de nuevas enzimas y mezclas de enzimas, que mejoren la eficacia de degradación. También existe una necesidad de procedimientos y enzimas, que trabajen a altas temperaturas, permitiendo así el uso de alta consistencia de biomasa y conduciendo a altas concentraciones de azúcar y etanol. Este enfoque puede conducir a un importante ahorro en costes de energía e inversiones. La alta temperatura también disminuye el riesgo de contaminación durante la hidrólisis. La presente invención tiene como objetivo satisfacer al menos parte de estas necesidades.

Breve descripción de la invención

Sorprendentemente, se ha encontrado ahora que las enzimas celulolíticas, y especialmente celobiohidrolasas obtenibles a partir de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum* son particularmente útiles en la hidrólisis de material celulósico. Además de celobiohidrolasas estos hongos también tienen endoglucanasas, beta-glucosidasas y xilanasas que son muy adecuadas para degradar material celulósico. Las enzimas son cinéticamente muy eficaces en un amplio intervalo de temperaturas, y aunque tienen una alta actividad a altas temperaturas, también son muy eficientes a temperaturas de hidrólisis convencionales. Esto las hace muy adecuadas para procedimientos de hidrólisis de diversos sustratos celulósicos llevados a cabo tanto a temperaturas convencionales como a temperaturas elevadas.

La presente invención proporciona un método para el tratamiento de material celulósico con celobiohidrolasa, endoglucanasa y beta-glucosidasa, con lo que dicha celobiohidrolasa comprende una secuencia de aminoácidos que tiene al menos 90% de identidad con SEQ ID NO: 6, o con un fragmento del mismo que tiene actividad celulolítica.

La invención proporciona adicionalmente una preparación enzimática que comprende celobiohidrolasa, endoglucanasa y beta-glucosidasa, en donde dicha celobiohidrolasa comprende una secuencia de aminoácidos que tiene al menos 90% de identidad con SEQ ID NO: 6, o con un fragmento del mismo que tiene actividad celulolítica.

También se proporciona el uso de dicha preparación enzimática para degradar material celulósico, así como el uso de dicho método en un procedimiento para la preparación de etanol a partir de material celulósico.

La invención también se refiere a un polipéptido que comprende un fragmento que tiene actividad celulolítica y que se selecciona del grupo que consiste en:

a) un polipéptido que comprende una secuencia de aminoácidos que tiene identidad de al menos 90% con SEQ ID NO: 6; y

b) un fragmento de a) que tiene actividad celulolítica.

Un objeto adicional de la invención es un polinucleótido aislado seleccionado del grupo que consiste en:

a) una secuencia de nucleótidos de SEQ ID NO: 5, o una secuencia que codifica un polipéptido de la invención;

b) una cadena complementaria de a); y

c) una secuencia que es degenerada como resultado del código genético con una cualquiera de las secuencias como se define en a) o b).

La invención también proporciona adicionalmente un vector, que comprende dicho polinucleótido en forma de una secuencia heteróloga, y una célula anfitriona que comprende dicho vector. La cepa de *Escherichia coli* que tiene el número de acceso DSM 16729 también se incluyen en la invención.

Otros objetos de la invención son las preparaciones enzimáticas que comprenden al menos el nuevo polipéptido, y el uso de dicha preparación de polipéptido o enzimática en la industria de combustibles, tejidos, detergentes, pasta de papel y papel, alimentos, piensos o bebidas.

Adicionalmente se proporciona un método para preparar un polipéptido que comprende un fragmento que tiene actividad celulolítica y que se selecciona del grupo que consiste en:

- a) un polipéptido que comprende una secuencia de aminoácidos que tiene identidad de al menos 90% con SEQ ID NO: 6; y
- b) un fragmento de a) que tiene actividad celulolítica,

comprendiendo dicho método transformar una célula anfitriona con un vector que codifica dicho polipéptido, y cultivar dicha célula anfitriona en condiciones que permitan la expresión de dicho polipéptido, y opcionalmente recuperar y purificar el polipéptido producido.

También se proporciona adicionalmente un método de tratamiento de material celulósico con un medio de cultivo gastado de al menos un microorganismo capaz de producir un polipéptido como se ha definido anteriormente, en donde el método comprende hacer reaccionar el material celulósico con el medio de cultivo gastado para obtener material celulósico hidrolizado.

Las realizaciones específicas de la invención se exponen en las reivindicaciones dependientes.

Breve descripción de los dibujos

Figura 1. Dependencias de la temperatura de las actividades celulasa y beta-glucosidasa en los sobrenadantes de las seis cepas fúngicas sometidas a ensayo. El tiempo de incubación en el análisis fue de 60 min a la temperatura dada, el análisis de pH fue de 5,0 (actividad MUL) o 4,8 (CMC_{asa} o BGU). La actividad obtenida a 60°C se establece como la actividad relativa de 100%. **A)** *Thermoascus aurantiacus* ALKO4239, **B)** *Thermoascus aurantiacus* ALKO4242, **C)** *Acremonium thermophilum* ALKO4245, **D)** *Talaromyces thermophilus* ALKO4246, **E)** *Chaetomium thermophilum* ALKO4261, **F)** *Chaetomium thermophilum* ALKO4265.

Figura 2. Imagen esquemática de los casetes de expresión utilizados en la transformación de protoplastos de *Trichoderma reesei* para producir las proteínas recombinantes fúngicas. Los genes recombinantes estaban bajo el control del promotor de *T. reesei cbh1* (*cel7A*) (*prom cbh1*) y la terminación de la transcripción se garantizó mediante el uso de una secuencia de terminación de *T. reesei cbh1* (*term cbh1*). El gen *amdS* se incluyó como marcador de transformación.

Figura 3 . A) Óptimos de pH de las preparaciones de proteínas CBH/Cel7 recombinantes de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 determinados sobre 4-metilumbeliferil-β-D-lactósido (MUL) a 50°C, 10 min. Los resultados se proporcionan como media (± DT) de tres mediciones separadas. **B)** Estabilidad térmica de las preparaciones de proteínas CBH/Cel7 recombinantes de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 determinada sobre 4-metilumbeliferil-β-D-lactósido (MUL) al pH óptimo durante 60 min. Los resultados se expresan como media (± DT) de tres mediciones separadas. Ambas reacciones contenían BSA (100 µg/ml) como estabilizador.

Figura 4. Hidrólisis de celulosa cristalina (Avicel) por las celobiohidrolasas recombinantes purificadas a 45°C. Concentración de sustrato 1% (p/v), pH 5,0, concentración de enzima 1,4 µM. **A)** Celobiohidrolasas que albergan un CBD, **B)** celobiohidrolasas (núcleo) sin un CBD.

Figura 5. Hidrólisis de celulosa cristalina (Avicel) por las celobiohidrolasas recombinantes purificadas a 70°C. Concentración de sustrato 1% (p/v), pH 5,0, concentración de enzima 1,4 µM. **A)** Celobiohidrolasas que albergan un CBD, **B)** celobiohidrolasas (núcleo) sin un CBD.

Figura 6. A) La dependencia del pH de la actividad de EG_40/Cel45A, de tipo EG_40/Cel45B de *Acremonium* y EG_28/Cel5A de *Thermoascus* producidas de forma heteróloga se determinó con sustrato CMC en una reacción de 10 min a 50°C. **B)** El óptimo de temperatura de EG_40/Cel45A, de tipo EG_40/Cel45B de *Acremonium* y EG_28/Cel5A de *Thermoascus* se determinó a pH 5,5, 4,8, y 6,0, respectivamente. Se realizó la reacción que contenía CMC como sustrato durante 60 minutos, a excepción de EG_28/Cel5A durante 10 min. Se añadió BSA (100 µg/ml) como estabilizador.

Figura 7. A) La dependencia del pH de la actividad de BG_101/Cel3A de *Acremonium*, BG_76/Cel3A de *Chaetomium*, y BG_81/Cel3A de *Thermoascus* producidas de forma heteróloga se determinó con sustrato 4-nitrofenil-β-D-glucopiranosido en una reacción de 10 min a 50°C. **B)** El óptimo de temperatura de BG_101/Cel3A de

Acremonium, β G₇₆/Cel3A de *Chaetomium* y β G₈₁/Cel3A de *Thermoascus* se determinó a pH 4,5, 5,5, y 4,5, respectivamente. Se realizó la reacción que contenía 4-nitrofenil- β -D-glucopiranosido como sustrato durante 60 min, se añadió BSA (100 μ g/ml) como estabilizador.

Figura 8. A) La dependencia del pH de la actividad xilanasa de XYN₃₀/Xyn10A de *Thermoascus* producida de forma heteróloga se determinó con sustrato de xilano de abedul en una reacción de 10 min a 50°C. **B)** El óptimo de temperatura de XYN₃₀/Xyn10A se determinó a pH 5,3 en una reacción de 60 min, se añadió BSA (100 μ g/ml) como estabilizador.

Figura 9. Hidrólisis de fibra de abeto sometida a explosión de vapor de agua lavada (10 mg/ml) con una mezcla de enzimas termófilas (MEZCLA 1) y enzimas de *T. reesei* a 55 y 60°C. La dosis de enzima se proporciona en en FPU/g de materia seca de sustrato, FPU analizada a 50°C, pH 5. La hidrólisis se llevó a cabo durante 72 horas a pH 5, mezclando. Los resultados se expresan como media ($\pm$ DT) de tres mediciones separadas.

Figura 10. Hidrólisis de rastrojo de maíz sometido a explosión de vapor de agua (10 mg/ml) con una mezcla de enzimas termófilas (MEZCLA 2) y enzimas de *T. reesei* a 45, 55 y 57,5°C. La dosis de enzima fue para la "MEZCLA 2" de 5 FPU/g de materia seca de sustrato y para las enzimas de *T. reesei* 5 FPU/g de materia seca Celluclast con un suplemento de 100 nkat/g de materia seca Novozym 188 (la actividad de papel de filtro se analizó a 50°C, pH 5). La hidrólisis se llevó a cabo durante 72 horas a pH 5, mezclando. Los resultados se expresan como media ($\pm$ DT) de tres mediciones separadas. El sustrato contenía azúcares reductores solubles (aprox. 0,7 mg/ml). Este contenido de azúcar de fondo se restó de los azúcares reductores formados durante la hidrólisis.

Figura 11. Hidrólisis de rastrojo de maíz sometido a explosión de vapor de agua (10 mg/ml) con una mezcla de enzimas termófilas que contenían una nueva xilanasa termófila de *Thermoascus aurantiacus* (MEZCLA 3) y enzimas de *T. reesei* a 45, 55 y 60°C. La dosis de enzima fue para la "MEZCLA 3" 5 FPU/g de materia seca de sustrato y para las enzimas de *T. reesei* 5 FPU/g de materia seca Celluclast un suplemento de 100 nkat/g de materia seca Novozym 188 (la actividad de papel de filtro se analizó a 50°C, pH 5). La hidrólisis se llevó a cabo durante 72 horas a pH 5, mezclando. Los resultados se expresan como media ($\pm$ DT) de tres mediciones separadas. El sustrato contenía azúcares reductores solubles (aprox. 0,7 mg/ml). Este contenido de azúcar de fondo se restó de los azúcares reductores formados durante la hidrólisis.

Figura 12. Hidrólisis de fibra de abeto sometida a explosión de vapor de agua (10 mg/ml) con una mezcla de enzimas termófilas que contenían una nueva xilanasa termófila XYN₃₀/Xyn10A de *Thermoascus aurantiacus* (MEZCLA 3) y enzimas de *T. reesei* a 45, 55 y 60°C. La dosis de enzima para la "MEZCLA 3" fue 5 FPU/g de materia seca de sustrato y para las enzimas de *T. reesei* 5 FPU/g de materia seca Celluclast con un suplemento de 100 nkat/g de materia seca Novozym 188 (la actividad de papel de filtro se analizó a 50°C, pH 5). La hidrólisis se llevó a cabo durante 72 horas a pH 5, mezclando. Los resultados se expresan como media ($\pm$ DT) de tres mediciones separadas.

Figura 13. Efecto de la glucosa sobre la actividad de diferentes preparaciones de β -glucosidasa. El análisis normalizado utilizando *p*-nitrofenil- β -D-glucopiranosido como sustrato se llevó a cabo en presencia de glucosa en la mezcla de ensayo. La actividad se presenta como porcentaje de la actividad obtenida sin glucosa.

Figura 14. Actividades FPU de las mezclas de enzimas a temperaturas de 50°C a 70°C, presentadas como un porcentaje de la actividad bajo condiciones convencionales (50°C, 1 h).

Figura 15. Actividad celulasa relativa de dos cepas de *T. reesei* diferentes cultivadas en medios que contienen Nutriose sin tratar (N0) o Nutriose pretratada con BG₈₁/Cel3A (NBG81) como fuente de carbono.

Descripción detallada de la invención

La celulosa es el componente estructural principal de las plantas superiores. Proporciona a las células vegetales alta resistencia a la tracción ayudándolas a resistir la tensión mecánica y la presión osmótica. La celulosa es un β -1,4-glucano compuesto por cadenas lineales de residuos de glucosa unidos por enlaces β -1,4-glucosídicos. La celobiosa es la unidad repetitiva más pequeña de la celulosa. En las paredes celulares la celulosa se empaqueta en láminas orientadas diversamente, que están embebidas en una matriz de hemicelulosa y lignina. La hemicelulosa es un grupo heterogéneo de polímeros de carbohidratos que contienen principalmente diferentes glucanos, xilanos y mananos. La hemicelulosa se compone de una cadena principal lineal con residuos unidos mediante enlaces β -1,4 sustituidos con cadenas laterales cortas que contienen usualmente acetilo, glucuronilo, arabinosilo y galactosilo. La hemicelulosa puede ser entrecruzada químicamente con la lignina. La lignina es un polímero entrecruzado complejo de unidades de *p*-hidroxifenilpropano diversamente sustituido que proporciona fuerza a la pared celular para resistir la tensión mecánica, y también protege la celulosa de la hidrólisis enzimática.

La lignocelulosa es una combinación de celulosa y hemicelulosa y polímeros de unidades de fenol propanol y lignina. Es físicamente dura, densa, e inaccesible y la sustancia bioquímica más abundante en la biosfera. Los materiales que contiene lignocelulosa son, por ejemplo: virutas de madera dura y blanda, pasta de madera, serrín y residuos forestales y de la industria de la madera; biomasa agrícola como paja de cereales, pasta de remolacha azucarera, rastrojo y mazorcas de maíz, bagazo de caña de azúcar, tallos, hojas, cáscaras, cascarillas, y similares; productos

de desecho como residuos sólidos urbanos, periódicos y papel de desecho de oficina, residuos de la molienda de p. ej. cereales; cultivos energéticos (por ejemplo, sauce, álamo, pasto varilla (*Panicum virgatum*) o alpiste arundináceo, y similares). Los ejemplos preferidos son el rastrojo de maíz, el pasto varilla, la paja de cereales, el bagazo de caña de azúcar y los materiales derivados de la madera.

"Material celulósico" según se utiliza en la presente memoria, se refiere a cualquier material que comprende celulosa, hemicelulosa y/o lignocelulosa como componente significativo. "Material lignocelulósico" significa cualquier material que comprende lignocelulosa. Tales materiales son, por ejemplo materiales vegetales tales como madera incluyendo madera blanda y dura, cultivos herbáceos, residuos agrícolas, residuos de pasta de papel y papel, papel de desecho, desechos de la industria alimenticia y de piensos, etc. Son ejemplos específicos de materiales celulósicos fibras textiles tales como el algodón, fibras derivadas de algodón, lino, cáñamo, yute y fibras celulósicas elaboradas por el hombre como modal, viscosa, liocel.

El material celulósico es degradado en la naturaleza por algunos organismos diversos, incluyendo bacterias y hongos. La celulosa es degradada típicamente por diferentes celulasas que actúan sucesivamente o simultáneamente. La conversión biológica de celulosa en glucosa requiere por lo general tres tipos de enzimas hidrolíticas: (1) Endoglucanasas que cortan enlaces beta-1,4-glucosídicos internos; (2) Exocelobiohidrolasas que cortan el disacárido celobiosa del extremo de la cadena polimérica de celulosa; (3) Beta-1,4-glucosidasas que hidrolizan la celobiosa y otros celo-oligosacáridos cortos a glucosa. En otras palabras, los tres grupos principales de celulasas son celobiohidrolasas (CBH), endoglucanasas (EG) y beta-glucosidasas (BG).

La degradación de los sustratos que contienen celulosa más complejos requiere una amplia gama de diversas enzimas. Por ejemplo la lignocelulosa es degradada por hemicelulasas, como xilanasas y mananasas. La hemicelulasa es un enzima que hidroliza hemicelulosa.

Las "enzimas celulolíticas" son enzimas que tienen "actividad celulolítica", lo que significa que son capaces de hidrolizar sustratos celulósicos o derivados de los mismos a sacáridos más pequeños. Por lo tanto las enzimas celulolíticas incluyen tanto celulasas como hemicelulasas. Las celulasas utilizadas en la presente memoria incluyen celobiohidrolasa, endoglucanasa y beta-glucosidasa.

T. reesei tiene un sistema de celulasa bien conocido y eficaz, con dos CBH, dos principales y varias secundarias EG y BG. CBHI (Cel7A) de *T. reesei* corta el azúcar desde el extremo reductor de la cadena de celulosa, tiene un dominio de unión a celulosa C-terminal (CBD) y puede constituir hasta 60% de la proteína secretada total. CBHII (Cel6A) de *T. reesei* corta el azúcar desde el extremo no reductor de la cadena de celulosa, tiene un dominio de unión a celulosa N-terminal y puede constituir hasta 20% de la proteína total secretada. Las endoglucanasas EGI (Cel7B), y EGV (Cel45A) tienen un CBD en su extremo C, EGII (Cel5A) tiene un CBD N-terminal y EGIII (Cel12A) no tiene ningún dominio de unión a celulosa. CBHI, CBHII, EGI y EGII son las llamadas "celulasas principales" de *Trichoderma* que comprenden en conjunto 80-90% del total de proteínas secretadas. Es conocido por un experto en la técnica que una enzima puede ser activa sobre varios sustratos y las actividades enzimáticas se pueden medir utilizando diferentes sustratos, métodos y condiciones. La identificación de las diferentes actividades celulolíticas es comentada, por ejemplo, por van Tilbeurgh *et al.* 1988.

Además de un dominio catalítico/núcleo que expresa la actividad celulolítica, las enzimas celulolíticas pueden comprender uno o más dominios de unión a celulosa (CBD), también denominados dominios/módulos de unión a carbohidratos (CBD/CBM), que pueden estar situados en el extremo N o C del dominio catalítico. Los CBD tienen actividad de unión a carbohidratos y median la unión de la celulasa a la celulosa cristalina, pero tienen poco o ningún efecto sobre la actividad hidrolítica celulasa de la enzima sobre sustratos solubles. Estos dos dominios están conectados típicamente a través de una región conectora flexible y altamente glicosilada.

"Celobiohidrolasa" o "CBH" según se utiliza en la presente memoria, se refiere a las enzimas que escinden la celulosa desde el extremo de la cadena de glucosa y producen principalmente celobiosa. También se denominan 1,4-beta-D-glucano celobiohidrolasas o celulosa 1,4-beta-celobiosidasas. Hidrolizan los enlaces 1,4-beta-D-glucosídicos de los extremos reductores o no reductores de un polímero que contiene dichos enlaces, tal como celulosa, con lo que se libera celobiosa. Se han aislado dos CBH diferentes de *Trichoderma reesei*, CBHI y CBHII. Tienen una estructura modular que consiste en un dominio catalítico conectado a un dominio de unión a celulosa (CBD). También hay celobiohidrolasas en la naturaleza que carecen de CBD.

"Endoglucanasa" o "EG" se refiere a enzimas que cortan enlaces glucosídicos internos de la cadena de celulosa. Se clasifican como EC 3.2.1.4. Son 1,4-beta-D-glucano 4-glucanohidrolasas y catalizan endohidrólisis de enlaces 1,4-beta-D-glucosídicos en los polímeros de glucosa tales como la celulosa y sus derivados. Algunas endoglucanasas de origen natural tienen un dominio de unión a celulosa, mientras que otras no lo tienen. Algunas endoglucanasas también tienen actividad xilanasas (Bailey *et al.*, 1993).

"Beta-glucosidasa" o "BG" o "βG" se refiere a enzimas que degradan pequeños oligosacáridos solubles incluyendo celobiosa a glucosa. Se clasifican como EC 3.2.1.21. Son beta-D-glucósido glucohidrolasas, que típicamente catalizan la hidrólisis de residuos beta-D-glucosa no reductores terminales. Estas enzimas reconocen oligosacáridos de glucosa. Los sustratos típicos son celobiosa y celotriosa. La celobiosa es un inhibidor de celobiohidrolasas, por lo

cual la degradación de celobiosa es importante para superar la inhibición por producto final de las celobiohidrolasas.

Las xilanasas son enzimas que son capaces de reconocer e hidrolizar hemicelulosa. Incluyen enzimas tanto exohidrolíticas como endohidrolíticas. Típicamente tienen actividad endo-1,4-beta-xilanasas (EC 3.2.1.8) o beta-D-xilosidasa (EC 3.2.1.37) que rompe la hemicelulosa a xilosa. "Xilanasas" o "Xyn" con respecto a la presente invención se refiere especialmente a una enzima clasificada como EC 3.2.1.8 que hidroliza polímeros de xilosa de sustrato lignocelulósico o xilano purificado.

Además de esto las celulasas se pueden clasificar en varias familias de glicosil hidrolasas de acuerdo con su secuencia primaria, apoyada por el análisis de la estructura tridimensional de algunos miembros de la familia (Henrissat 1991, Henrissat y Bairoch 1993, 1996). Algunas glicosil hidrolasas son enzimas multifuncionales que contienen dominios catalíticos que pertenecen a diferentes familias de glicosilhidrolasas. La familia 3 consta de beta-glucosidasas (CE 3.2.1.21) tales como Ta BG_81, At BG_101 y Ct BG_76 descritas en la presente memoria. La familia 5 (anteriormente conocida como celA) consiste principalmente de endoglucanasas (EC 3.2.1.4) tales como Ta EG_28 descrita en la presente memoria. La familia 7 (anteriormente familia celC de celulasas) contiene endoglucanasas (EC 3.2.1.4) y celobiohidrolasas (EC 3.2.1.91) tales como Ct EG_54, Ta CBH, At CBH_A, At CBH_C y Ct CBH descritas en la presente memoria. La familia 10 (anteriormente celF) consiste principalmente en xilanasas (EC 3.2.1.8) tales como Ta XYN_30 y At XYN_60 descritas en la presente memoria. La familia 45 (anteriormente celK) contiene endoglucanasas (EC 3.2.1.4) tales como At EG_40 y de tipo_At EG_40 descritas en la presente memoria.

Las enzimas celulolíticas útiles para hidrolizar material celulósico se pueden obtener de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*. "Obtenibles a partir de" significa que pueden ser obtenidas de dicha especie, pero no excluye la posibilidad de obtenerlas de otras fuentes. En otras palabras pueden proceder de cualquier organismo, incluidas las plantas. Preferiblemente se originan a partir de microorganismos, por ejemplo bacterias u hongos. Las bacterias pueden ser, por ejemplo, de un género seleccionado entre *Bacillus*, *Azospirillum* y *Streptomyces*. Más preferiblemente, la enzima se origina a partir de hongos (incluyendo hongos filamentosos y levaduras), por ejemplo de un género seleccionado del grupo que consiste en *Thermoascus*, *Acremonium*, *Chaetomium*, *Achaetomium*, *Thielavia*, *Aspergillus*, *Botrytis*, *Chrysosporium*, *Collybia*, *Fomes*, *Fusarium*, *Hemicola*, *Hypocrea*, *Lentinus*, *Melanocarpus*, *Myceliophthora*, *Myriococcum*, *Neurospora*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Pleurotus*, *Podospora*, *Polyporus*, *Rhizoctonia*, *Scytalidium*, *Pycnoporus*, *Trametes* y *Trichoderma*.

Las enzimas descritas se pueden obtener de la cepa de *Thermoascus aurantiacus* ALKO4242 depositada como CBS 116239, la cepa ALKO4245 depositada como CBS 116240 actualmente clasificada como *Acremonium thermophilum*, o la cepa de *Chaetomium thermophilum* ALK04265 depositada como CBS 730.95.

Otras celobiohidrolasas descritas en la presente memoria comprenden preferiblemente una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 2, 4, u 8, o un fragmento enzimáticamente activo de los mismos.

Celobiohidrolasa	Gen	Obtenible de	CBD	ácido nucleico SEQ ID NO:	aminoácido SEQ ID NO:
Ta CBH	Ta cel7A	<i>T. aurantiacus</i>	-	1	2
At CBH_A	At cel7B	<i>A. thermophilum</i>	-	3	4
At CBH_C	At cel7A	<i>A. thermophilum</i>	+	5	6
Ct CBH	Ct cel7A	<i>C. thermophilum</i>	+	7	8

Estas CBH tienen una constante de inhibición de celobiosa ventajosa en comparación con la de las CBH de *Trichoderma reesei*, y muestran mejores resultados de hidrólisis cuando se someten a ensayo diversos sustratos celulósicos. Las SEQ ID NO: 2 y 4 no comprenden un CBD. Se pueden obtener resultados de hidrólisis particularmente mejorados cuando un dominio de unión a celulosa (CBD) está unido a una CBH que no tiene CBD propio. El CBD se puede obtener por ejemplo a partir de una especie de *Trichoderma* o *Chaetomium*, y se ancla preferiblemente a la CBH través de un conector. La proteína de fusión resultante que contiene una región núcleo de CBH unida a un CBD a través de un conector puede comprender una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 28 o 30. Los polinucleótidos que comprenden una secuencia de SEQ ID NO: 27 o 29 codifican tales proteínas de fusión.

La endoglucanasa puede comprender una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 10, 12, 14 o 16, o un fragmento enzimáticamente activo de los mismos. Estas endoglucanasas tienen buena termoestabilidad.

Endoglucanasa	Gen	Obtenible de	CBD	ácido nucl. SEQ ID NO:	aminoácido SEQ ID NO:
Ta EG_28	Ta <i>cel5A</i>	<i>T. aurantiacus</i>	-	9	10
At EG_40	At <i>cel45A</i>	<i>A. thermophilum</i>	+	11	12
de tipo_At EG40	At <i>cel45B</i>	<i>A. thermophilum</i>	-	13	14
Ct EG_54	Ct <i>cel7B</i>	<i>C. thermophilum</i>	+	15	16

- 5 La beta-glucosidasa puede comprender una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 22, 24 o 26, o un fragmento enzimáticamente activo de los mismos. Estas beta-glucosidasas tienen buena resistencia a la inhibición por glucosa, lo que es ventajoso para evitar la inhibición por producto final durante la hidrólisis enzimática de material celulósico. Las beta-glucosidasas también se pueden utilizar en la preparación de soforosa, un inductor de celulasa utilizado en el cultivo de *T. reesei*.

Beta-glucosidasa	Gen	Obtenible de	ácido nucleico SEQ ID NO:	aminoácido SEQ ID NO:
Ta BG_81	Ta <i>cel3A</i>	<i>T. aurantiacus</i>	21	22
At BG_101	At <i>cel3A</i>	<i>A. thermophilum</i>	23	24
Ct BG_76	Ct <i>cel3A</i>	<i>C. thermophilum</i>	25	26

- 10 La xilanasa puede comprender una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 18 o 20, o un fragmento enzimáticamente activo de los mismos.

Xilanasa	Gen	Obtenible de	CBD	ácido nucleico SEQ ID NO:	aminoácido SEQ ID NO:
Xyn_30	Ta <i>xyn10A</i>	<i>T. aurantiacus</i>	+	17	18
Xyn_60	At <i>xyn10A</i>	<i>A. thermophilum</i>	-	19	20

- 15 Mediante el término "identidad" se quiere significar en la presente memoria la identidad global entre dos secuencias de aminoácidos en comparación con la otra desde el primer aminoácido codificado por el gen correspondiente hasta el último aminoácido. La identidad de las secuencias completas se mide mediante el uso del programa de alineamiento global de Needleman-Wunsch en el paquete de programa EMBOSS (European Molecular Biology Open Software Suite; Rice *et al.*, 2000), versión 3.0.0, con los siguientes parámetros: EMBL62, penalización por hueco 10,0, Penalización por extensión 0,5. El algoritmo es descrito por Needleman y Wunsch (1970). El experto en la técnica es consciente del hecho de que los resultados que utilizan el algoritmo de Needleman-Wunsch solo son comparables cuando se alinean los dominios correspondientes de la secuencia. Por consiguiente no se puede
- 20 realizar la comparación p. ej. de secuencias de celulasas que incluyen CBD o secuencias señal con secuencias que carecen de esos elementos.

- 25 De acuerdo con una realización de la invención, se utiliza un polipéptido celulolítico que tiene al menos 95 o 99% de identidad con SEQ ID NO: 6, o un fragmento del mismo que tiene actividad celulolítica. Otros polipéptidos celulolíticos descritos tienen 80, 85, 90, 95 o 99% de identidad con SEQ ID NO: 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 24 o 26 o al menos con su fragmento enzimáticamente activo.

Mediante el término "fragmento enzimáticamente activo" se entiende cualquier fragmento de una secuencia definida que tiene actividad celulolítica. En otras palabras un fragmento enzimáticamente activo puede ser la porción de proteína madura de la secuencia definida, o puede ser solamente un fragmento de la porción de proteína madura, siempre que todavía tenga actividad celobiohidrolasa, endoglucanasa, beta-glucosidasa o xilanasa.

Las enzimas celulolíticas son preferiblemente enzimas recombinantes, que pueden ser producidas de una manera conocida generalmente. Se aísla un fragmento de polinucleótido que comprende el gen de la enzima, el gen se inserta bajo un promotor fuerte en un vector de expresión, el vector se transfiere a células anfitrionas adecuadas y las células anfitrionas se cultivan en condiciones que provocan la producción de la enzima. Los métodos para la producción de proteínas por medio de tecnología recombinante en diferentes sistemas anfitriones son bien conocidos en la técnica (Sambrook *et al.*, 1989; Co-en, 2001; Gellissen, 2005). Preferiblemente, las enzimas son producidas como enzimas extracelulares que son secretadas al medio de cultivo, del que se pueden recuperar y aislar fácilmente. El medio de cultivo gastado del anfitrión de producción se puede utilizar tal cual, o las células anfitrionas se pueden retirar del mismo, y/o se puede concentrar, filtrar o fraccionar. También se puede secar.

El polipéptido aislado en el presente contexto puede significar simplemente que las células y los restos celulares se han eliminado del medio de cultivo que contiene el polipéptido. Convenientemente, los polipéptidos se aíslan, por ejemplo, mediante la adición de polímeros aniónicos y/o catiónicos al medio de cultivo gastado para mejorar la precipitación de las células, los restos celulares y algunas enzimas que tienen actividades secundarias no deseadas. El medio se filtra a continuación, utilizando un agente de filtrado inorgánico y un filtro para eliminar los precipitantes formados. Después de esto el producto filtrado se procesa adicionalmente utilizando una membrana semi-permeable para eliminar el exceso de sales, azúcares y productos metabólicos.

El polinucleótido heterólogo descrito en la presente memoria comprende un gen similar al incluido en un microorganismo que tiene el número de acceso DSM 16723, DSM 16728, DSM 16729, DSM 16727, DSM 17326, DSM 17324, DSM 17323, DSM 17729, DSM 16724, DSM 16726, DSM 16725, DSM 17325 o DSM 17667.

El anfitrión de producción puede ser cualquier organismo capaz de expresar la enzima celulolítica. Preferiblemente, el anfitrión es una célula microbiana, más preferiblemente un hongo. Lo más preferiblemente, el anfitrión es un hongo filamentos. Preferiblemente, el anfitrión recombinante se modifica para expresar y secretar enzimas celulolíticas como actividad principal o una de sus actividades principales. Esto se puede realizar suprimiendo los principales genes homólogos secretados, p. ej., las cuatro principales celulasas de *Trichoderma* y dirigiendo genes heterólogos a un locus que se ha modificado para asegurar altos niveles de expresión y producción. Los anfitriones preferidos para la producción de las enzimas celulolíticas son en particular cepas del género *Trichoderma* o *Aspergillus*.

Las enzimas necesarias para la hidrólisis del material celulósico se pueden añadir en una cantidad enzimáticamente eficaz ya sea simultáneamente, p. ej., en forma de una mezcla de enzimas, o sucesivamente, o como una parte de la sacarificación y fermentación simultáneas (SSF). Se puede utilizar cualquier combinación de celobiohidrolasas que comprenden una secuencia de aminoácidos que tiene identidad de al menos 90% con SEQ ID NO: 6, o su fragmento activo, y celobiohidrolasas que comprenden una secuencia de aminoácidos que tiene identidad de al menos 80% con SEQ ID NO: 2, 4, u 8 o un fragmento enzimáticamente activo de los mismos junto con cualquier combinación de endoglucanasas y beta-glucosidasas. Si el material celulósico comprende hemicelulosa, se utilizan adicionalmente hemicelulasas, preferiblemente xilanasas para la degradación. Las endoglucanasas, beta-glucosidasas y xilanasas se pueden seleccionar entre las descritas en la presente memoria, pero no se limitan a las mismas. También pueden ser por ejemplo preparaciones enzimáticas asequibles comercialmente. Además de las celulasas y las hemicelulasas opcionales se pueden utilizar una o más enzimas, por ejemplo proteasas, amilasas, lacasas, lipasas, pectinasas, esterases y/o peroxidases. Se puede llevar a cabo otro tratamiento enzimático antes, durante o después del tratamiento con celulasa.

El término "preparación enzimática" denota una composición que comprende al menos una de las enzimas deseadas. La preparación puede contener las enzimas en forma al menos parcialmente purificada y aislada. Puede consistir incluso esencialmente en la enzima o enzimas deseadas. Alternativamente, la preparación puede ser un medio de cultivo gastado o producto filtrado que contiene una o más enzimas celulolíticas. Además de la actividad celulolítica, la preparación puede contener aditivos, tales como mediadores, estabilizantes, tampones, conservantes, tensioactivos y/o componentes del medio de cultivo. Los aditivos preferidos son tales que se utilizan comúnmente en preparaciones de enzimas destinadas a una aplicación concreta. La preparación enzimática puede estar en forma de líquido, polvo o granulado. Preferiblemente, la preparación enzimática es medio de cultivo gastado. "Medio de cultivo gastado" se refiere al medio de cultivo del anfitrión que comprende las enzimas producidas. Preferiblemente, las células anfitrionas se separan de dicho medio después de la producción.

De acuerdo con una realización de la invención, la preparación enzimática comprende una mezcla de CBH, EG y BG, opcionalmente junto con xilanasas y/u otras enzimas. La CBH comprende una secuencia de aminoácidos que tiene una identidad de al menos 90% con SEQ ID NO: 6 o un fragmento enzimáticamente activo del mismo. Otras CBH descritas tienen una identidad de al menos 80% con SEQ ID NO: 2, 4, u 8 o un fragmento enzimáticamente activo de los mismos, y puede obtenerse a partir de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*, mientras que EG, BG y la xilanasas pueden tener cualquier origen, incluyendo a partir de dichos organismos. Otras enzimas que pueden estar presentes en la preparación son, por ejemplo proteasas, amilasas, lacasas, lipasas, pectinasas, esterases y/o peroxidases.

Se pueden utilizar diferentes mezclas de enzimas y combinaciones para adaptarse a diferentes condiciones del proceso. Por ejemplo, si el proceso de degradación se va a llevar a cabo a una temperatura alta, se eligen enzimas

termoestables. Una combinación de una CBH de la familia 7 con una endoglucanasa de la familia 45, opcionalmente combinadas con una BG de la familia 3 y/o una xilanasa de la familia 10 tuvo un excelente funcionamiento de hidrólisis tanto a 45°C, como a temperaturas elevadas.

Las enzimas celulolíticas de *Trichoderma reesei* se utilizan convencionalmente a temperaturas en el intervalo de aproximadamente 40-50°C en la hidrólisis, y de 30-40°C en SSF. Las CBH, EG, BG y Xyn obtenibles a partir de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum* son eficientes a estas temperaturas también, pero además la mayoría de las mismas también funcionan extremadamente bien a temperaturas entre 50°C y 75°C, o incluso hasta 80°C y 85°C, tal como entre 55°C y 70°C, por ejemplo, entre 60°C y 65°C. Para tiempos de incubación cortos las mezclas de enzimas son funcionales hasta incluso 85°C, para la hidrólisis completa se utilizan normalmente temperaturas más bajas.

El método para el tratamiento de material celulósico con CBH, EG, BG y Xyn es especialmente adecuado para la producción de azúcares fermentables a partir de material lignocelulósico. Los azúcares fermentables pueden ser fermentados a continuación por la levadura a etanol, y utilizados como combustible. También se pueden utilizar como productos intermedios o materias primas para la producción de diversos productos químicos o elementos esenciales para los procedimientos de la industria química, por ejemplo en la denominada biorrefinería. El material lignocelulósico puede ser pretratado antes de la hidrólisis enzimática para interrumpir la estructura de fibra de los sustratos celulósicos y hacer la fracción de celulosa más accesible a las enzimas celulolíticas. Los pretratamientos actuales incluyen procedimientos mecánicos, químicos o térmicos y combinaciones de los mismos. El material puede ser pretratado por ejemplo mediante explosión de vapor de agua o hidrólisis ácida.

Se encontraron varios nuevos polipéptidos celulolíticos en *Thermoascus aurantiacus*, *Acremonium thermophilum*, y *Chaetomium thermophilum*. Los nuevos polipéptidos pueden comprender un fragmento que tiene actividad celulolítica y ser seleccionados del grupo que consiste en un polipéptido que comprende una secuencia de aminoácidos que tiene al menos 66%, preferiblemente 70% o 75%, de identidad con SEQ ID NO: 4, 90% de identidad con SEQ ID NO: 6, 78% de identidad con SEQ ID NO: 12, 68%, preferiblemente 70% o 75%, de identidad con SEQ ID NO: 14, 72%, preferiblemente 75%, de identidad con SEQ ID NO: 16, 68%, preferiblemente 70% o 75%, de identidad con SEQ ID NO: 20, 74% de identidad con SEQ ID NO: 22 o 24, o 78% de identidad con SEQ ID NO: 26.

También se describen las variantes de los nuevos polipéptidos. Una "variante" puede ser un polipéptido que se produce naturalmente, p. ej., en forma de una variante alélica dentro de la misma cepa, especie o género, o que puede haber sido generado por mutagénesis. Puede comprender sustituciones, delecciones o inserciones de aminoácidos, pero todavía funcionar de una manera sustancialmente similar a las enzimas anteriormente definidas es decir, comprende un fragmento que tiene actividad celulolítica.

Los polipéptidos celulolíticos se producen normalmente en la célula en forma de polipéptidos inmaduros que comprenden una secuencia señal que se escinde durante la secreción de la proteína. También se pueden procesar adicionalmente durante la secreción en el extremo N-terminal y/o C-terminal para proporcionar una proteína madura, enzimáticamente activa. Un polipéptido "que comprende un fragmento que tiene actividad celulolítica" significa por tanto que el polipéptido puede estar en forma inmadura o madura, preferiblemente está en forma madura, es decir, el procesamiento ha tenido lugar.

Los nuevos polipéptidos descritos pueden ser adicionalmente un "fragmento de los polipéptidos o variantes" mencionados anteriormente. El fragmento puede ser la forma madura de las proteínas mencionadas anteriormente, o puede ser solo una parte activa enzimáticamente de la proteína madura. De acuerdo con una realización de la invención, el polipéptido tiene una identidad de al menos 90% con SEQ ID NO: 6. Los otros polipéptidos pueden tener una secuencia de aminoácidos que tiene una identidad de al menos 80, 85, 90, 95, o 99% con SEQ ID NO: 4, 12, 14, 16, 20, 22, 24 o 26, o un fragmento celulolíticamente activo de los mismos. También puede ser una variante del mismo, o un fragmento del mismo que tiene actividad celobiohidrolasa, endoglucanasa, xilanasa, o beta-glucosidasa. El polipéptido puede consistir esencialmente en un fragmento celulolíticamente activo de una secuencia de SEQ ID NO: 4, 6, 12, 14, 16, 20, 22, 24 o 26.

Los nuevos polinucleótidos pueden comprender una secuencia de nucleótidos de SEQ ID NO: 3, 5, 11, 13, 15, 19, 21, 23 o 25, o una secuencia que codifica un nuevo polipéptido como se ha definido anteriormente, incluyendo sus cadenas complementarias. El polinucleótido según se utiliza en la presente memoria se refiere tanto a ARN como a ADN, y puede ser monocatenario o bicatenario. El polinucleótido también puede ser un fragmento de dichos polinucleótidos que comprende al menos 20 nucleótidos, p. ej. al menos 25, 30 o 40 nucleótidos. Puede tener al menos 100, 200 o 300 nucleótidos de longitud. Adicionalmente, el polinucleótido puede ser degenerado como resultado del código genético para una cualquiera de las secuencias definidas anteriormente. Esto significa que diferentes codones pueden codificar el mismo aminoácido.

El polinucleótido puede estar "comprendido en" SEQ ID NO: 3, 5, 11, 13, 15, 19, 21, 23 o 25, lo que significa que la secuencia tiene al menos parte de la secuencia mencionada. El polinucleótido puede comprender un gen similar al incluido en un microorganismo que tiene el número de acceso DSM 16728, DSM 16729, DSM 17324, DSM 17323, DSM 17729, DSM 16726, DSM 16725, DSM 17325 o DSM 17667.

Las nuevas proteínas/polipéptidos se pueden preparar como se ha descrito anteriormente. Los nuevos polinucleótidos pueden insertarse en un vector, que es capaz de expresar el polipéptido codificado por la secuencia heteróloga, y el vector se pueden insertar en una célula anfitriona capaz de expresar dicho polipéptido. La célula anfitriona es preferiblemente del género *Trichoderma* o *Aspergillus*.

- 5 Un gen heterólogo que codifica los nuevos polipéptidos se ha introducido en un plásmido en una cepa de *Escherichia coli* que tiene el número acceso DSM 16728, DSM 16729, DSM 17324, DSM 17323, DSM 17729, DSM 16726, DSM 16725, DSM 17325 o DSM 17667.

- 10 Las nuevas enzimas pueden ser componentes de una preparación enzimática. La preparación enzimática puede comprender uno o más de los nuevos polipéptidos, y puede estar, p. ej., en forma de medio de cultivo gastado, polvo, gránulos o líquido. De acuerdo con una realización de la invención ésta comprende actividad celobiohidrolasa, endoglucanasa, beta-glucosidasa, y opcionalmente xilanasas y/u otras actividades enzimáticas. Puede comprender adicionalmente cualquier aditivo convencional.

- 15 Las nuevas enzimas se pueden aplicar en cualquier procedimiento que implique enzimas celulolíticas, tal como en la industria de los combustibles, tejidos, detergentes, pasta de papel y papel, alimentos, piensos o bebidas, y especialmente en la hidrólisis de material celulósico para la producción de biocombustible que comprende etanol. En la industria de la pasta de papel y el papel se pueden utilizar para modificar fibra celulósica, por ejemplo, en el tratamiento de pasta kraft, pasta mecánica, o papel reciclado.

La invención se ilustra mediante los siguientes ejemplos no limitantes.

Ejemplos

- 20 Ejemplo 1. Escrutinio de cepas que expresan actividad celulolítica y su cultivo para la purificación

Se sometieron a ensayo aproximadamente 25 cepas de hongos de la colección de cultivos Roal Oy para determinar la actividad celulolítica incluyendo beta-glucosidasas. Después del escrutinio preliminar se seleccionaron seis cepas para estudios adicionales. Estas fueron *Thermoascus aurantiacus* ALKO4239 y ALKO4242, *Acremonium thermophilum* ALKO4245, *Talaromyces thermophilus* ALKO4246 y *Chaetomium thermophilum* ALKO4261 y ALKO4265.

- 25 Las cepas ALKO4239, ALKO4242 y ALKO4246 se cultivaron en matraces oscilantes a 42°C durante 7 días en el medio 3 x B, que contiene en g/litro: celulosa Solka Floc 18, bagazo de destilería 18, xilano de granza de avena 9, CaCO₃ 2, harina de soja 4,5, (NH₄)HPO₄ 4,5, salvado de trigo 3,0, KH₂PO₄ 1,5, MgSO₄ · H₂O 1,5, NaCl 0,5, KNO₃ 0,9, goma garrofin 9,0, disolución de elementos traza Núm. 1 0,5, disolución de oligoelementos Núm. 2 0,5 y Struktol (Stow, OH, USA) antiespumante 0,5 ml; el pH se ajustó a 6,5. La disolución de elementos traza Núm. 1 tiene en g/litro: MnSO₄ 1,6, ZnSO₄ · 7 H₂O 3,45 y CoCl₂ · 6H₂O 2,0; La disolución de oligoelementos Núm. 2 tiene en g/litro: FeSO₄ · 7 H₂O 5,0 con dos gotas de H₂SO₄ concentrado.

- 30 La cepa ALKO4261 se cultivó en matraces oscilantes en el medio 1XB, que tiene un tercio de cada uno de los constituyentes del medio 3 x B (anterior), excepto que tiene mismas concentraciones de CaCO₃, NaCl y los elementos traza. La cepa se cultivó a 45°C durante 7 días.

La cepa ALKO4265 se cultivó en matraces oscilantes en el siguiente medio, g/l: celulosa Solka Floc 40, Pharmamedia™ (Traders Protein, Memphis, TN, USA) 10, macerado de maíz en polvo 5, (NH₄)₂SO₄ 5 y KH₂PO₄ 15; el pH se ajustó a 6,5. La cepa se cultivó a 45°C durante 7 días.

- 40 Después del cultivo, las células y otros sólidos se recogieron por centrifugación y el sobrenadante se recuperó. Para los cultivos en matraz oscilante, se añadieron inhibidores de proteasa PMSF (fluoruro de fenilmetilsulfonilo) y pepstatina A a 1 mM y 10 g/ml, respectivamente. Si no se utilizaron inmediatamente, las preparaciones se almacenaron en alícuotas a -20°C.

- 45 Para la estimación de la termoactividad de las enzimas, se realizaron análisis de las preparaciones de cultivo de los matraces oscilantes a 50°C, 60°C, 65°C, 70°C y 75°C durante 1 h, en presencia de 100 µg de albúmina de suero bovino (BSA)/ml como estabilizador. Se realizaron ensayos preliminares a 50°C y 65°C a dos valores de pH diferentes (4,8/5,0 o 6,0) con el fin de aclarar que pH era más apropiado para el análisis de termoactividad.

Todos los sobrenadantes del matraz oscilante se analizaron para determinar las siguientes actividades:

Actividad de tipo celobiohidrolasa I ('CBHI') y actividad de tipo endoglucanasa I ('EGI'):

- 50 Estas se midieron en tampón de acetato de Na 50 mM con MUL (4-metilumbeliferil-beta-D-lactósido) 0,5 mM como sustrato. Se añadió glucosa (100 mM) para inhibir cualquier actividad beta-glucosidasa de interferencia. El 4-metilumbeliferilo liberado se midió a 370 nm. Las actividades 'CBHI' y 'EGI' se distinguieron midiendo la actividad en presencia y ausencia de celobiosa (5 mM). La actividad que no es inhibida por la celobiosa representa la actividad 'EGI' y la actividad MUL restante representa la actividad 'CBHI' (van Tilbeurgh *et al*, 1988). El análisis se realizó a pH 5,0 o 6,0 (véase más adelante).

Actividad endoglucanasa (CMCasa):

Ésta se analizó con carboximetilcelulosa (CMC) al 2% (p/v) como sustrato en tampón de citrato 50 mM esencialmente como describen Bailey y Nevalainen 1981; Haakana *et al.* 2004. Los azúcares reductores se midieron con el reactivo DNS. El análisis se realizó a pH 4,8 o 6,0 (véase más adelante).

5 Actividad beta-glucosidasa (BGU):

Ésta se analizó con 4-nitrofenil-β-D-glucopiranosido (1 mM) en tampón de citrato 50 mM como describen Bailey y Nevalainen 1981. El 4-nitrofenol liberado se midió a 400 nm. El análisis se realizó a pH 4,8 o 6,0 (véase más adelante).

10 Las actividades relativas de las enzimas se presentan en Figura 1. Las actividades relativas se presentaron estableciendo la actividad a 60°C como 100% (Figura 1). Todas las cepas produjeron enzimas, que tenían alta actividad a altas temperaturas (65°C-75°C).

15 Para las purificaciones de proteínas. ALKO4242 también se cultivó en un biorreactor de 2 litros (Braun Biostat® B, Braun, Melsungen, Alemania) en el siguiente medio, g/litro: celulosa Solka Floc 40, harina de soja 10, NH₄NO₃ 5, KH₂PO₄ 5, MgSO₄ · 7 H₂O 0,5, CaCl₂ · 2 H₂O 0,05, disolución de elementos traza Núm. 1 0,5, disolución de elementos traza Núm. 2 0,5. La aireación fue de 1 vvm, control antiespumante con Struktol, la agitación de 200-800 rpm y la temperatura de 47°C. Se ejecutaron dos lotes, uno a pH 4,7 ± 0,2 (NH₃/H₂SO₄) y el otro con pH inicial de pH 4,5. El tiempo de cultivo fue de 7 d. Después del cultivo las células y otros sólidos se eliminaron mediante centrifugación.

20 La cepa ALKO4245 se cultivó en un biorreactor de 2 litros (Braun Biostat® B, Braun, Melsungen, Alemania) en el siguiente medio, g/litro: celulosa Solka Floc 40, macerado de maíz en polvo 15, bagazo de destilería 5, xilano de granza de avena 3, goma garrofín 3, (NH₄)₂SO₄ 5 y KH₂PO₄ 5. El intervalo de pH fue de 5,2 ± 0,2 (NH₃/H₂SO₄), la aireación de 1 vvm, la agitación de 300-600 rpm, el control antiespumante con Struktol y la temperatura de 42°C. El tiempo de cultivo fue 4 d. Después del cultivo las células y otros sólidos se eliminaron mediante centrifugación.

25 Para la purificación de la enzima, ALKO4261 se cultivó en un biorreactor de 10 litros (Braun Biostat® ED, Braun, Melsungen, Alemania) en el siguiente medio, g/litro: celulosa Solka Floc 30, bagazo de destilería 10, xilano de granza de avena 5, CaCO₃ 2, harina de soja 10, salvado de trigo 3,0, (NH₄)₂SO₄ 5, KH₂PO₄ 5, MgSO₄ · 7 H₂O 0,5, NaCl 0,5, KNO₃ 0,3, disolución de elementos traza Núm. 1 0,5 y disolución de elementos traza Núm. 2 0,5. El intervalo de pH fue de 5,2 ± 0,2 (NH₃/H₂SO₄), la aireación de 1 vvm, la agitación de 200-600 rpm, el control antiespumante con Struktol y la temperatura de 42°C. El tiempo de cultivo fue de 5 d. Un segundo lote se cultivó en condiciones similares, excepto que se añadió Solka Floc a 40 g/l y bagazo gastado a 15 g/l. Los sobrenadantes se recuperaron por centrifugación y se filtraron a través de filtros Seitz K-150 y EK (Pall SeitzSchenk Filtersystems GmbH, Bad Kreuznach, Alemania). El último sobrenadante se concentró aproximadamente diez veces utilizando el sistema de ultrafiltración Pellicon Mini (NMWL del filtro de 10 kDa; Millipore, Billerica, MA, USA).

35 Para la purificación de la enzima, ALKO4265 también se cultivó en un biorreactor de 10 litros (Braun Biostat® ED, Braun, Melsungen, Alemania) en el mismo medio que antes, excepto que se añadió KH₂PO₄ a 2,5 g/l. El intervalo de pH fue de 5,3 ± 0,3 (NH₃/H₃PO₄), la aireación de 0,6 vvm, la agitación de 500 rpm, el control antiespumante con Struktol y la temperatura de 43°C. El tiempo de cultivo fue de 7 d. Los sobrenadantes se recuperaron mediante centrifugación y se filtraron a través de filtros Seitz K-150 y EK (Pall SeitzSchenk Filtersystems GmbH, Bad Kreuznach, Alemania). El último sobrenadante se concentró aproximadamente 20 veces utilizando el sistema de ultrafiltración Pellicon Mini (NMWL del filtro de 10 kDa; Millipore, Billerica, MA, USA).

Ejemplo 2. Purificación y caracterización de celobiohidrolasas de *Acremonium thermophilum* ALKO4245 y *Chaetomium thermophilum* ALKO4265

45 Se cultivaron *Acremonium thermophilum* ALKO4245 y *Chaetomium thermophilum* ALKO4265 como se describe en el Ejemplo 1. Las principales celobiohidrolasas se purificaron utilizando una columna de afinidad basada en p-aminobencil-1-tio-β-celobiosido, preparada como describen Tomme *et al.*, 1988.

50 Los sobrenadantes del cultivo se tamponaron primero en tampón de acetato de sodio 50 mM de pH 5,0, que contenía δ-gluconolactona 1 mM y glucosa 0,1 M con el fin de retardar la hidrólisis del ligando en presencia de β-glucosidasas. Las celobiohidrolasas se eluyeron con lactosa 0,1 M y finalmente se purificaron mediante cromatografía de filtración en gel utilizando columnas Superdex 200 HR 10/30 en el sistema ÄKTA (Amersham Pharmacia Biotech). El tampón utilizado en la filtración en gel fue fosfato de sodio 50 mM de pH 7,0, que contenía cloruro de sodio 0,15 M.

55 Las celobiohidrolasas purificadas se analizaron mediante electroforesis en gel de SDS-poliacrilamida y se determinó que la masa molecular de ambas proteínas era aproximadamente 70 kDa evaluada sobre la base de los patrones de masa molecular (Kit de calibración de bajo peso molecular, Amersham Biosciences). Las celobiohidrolasas de *Acremonium* y *Chaetomium* purificadas fueron denominadas At Cel7A y Ct Cel7A, respectivamente, siguiendo el esquema en Henrissat *et al.* (1998) (Henrissat, 1991; Henrissat y Bairoch, 1993).

La actividad específica de las preparaciones se determinó utilizando 4-metilumbeliferil- β -D-lactósido (MUL), 4-metilumbeliferil- β -D-celobiosido (Mug2) o 4-metilumbeliferil- β -D-celotriósido (mug3) como sustrato (van Tilbeurghet *et al.*, a 50°C durante 10 min 1988) en tampón de citrato de sodio 0,05 M de pH 5. Las actividades endoglucanasa y xilanasa se determinaron mediante procedimientos convencionales (de acuerdo con la IUPAC, 1987) utilizando carboximetil celulosa (CMC) y glucuronoxilano de abedul (Bailey *et al.*, 1992) como sustratos. La actividad específica frente a Avicel se calculó sobre la base de los azúcares reductores formados en una reacción de 24 h a 50°C, pH 5,0, con sustrato al 1% y dosificación de enzima 0,25 μ M. El contenido de proteína de las preparaciones enzimáticas purificadas se midió de acuerdo con Lowry *et al.*, 1951. Para caracterizar los productos finales de la hidrólisis, los azúcares solubles liberados en un experimento de hidrólisis de 24 h, como se ha descrito anteriormente, se analizaron mediante HPLC (Dionex). La celobiohidrolasa I (CBHI/Cel7A) purificada de *Trichoderma reesei* se utilizó como referencia.

Las actividades específicas de las enzimas purificadas y la de CBHI/Cel7A de *T. reesei* se presentan en la Tabla 1. Las celobiohidrolasas At Cel7A y Ct Cel7A purificadas poseen actividades específicas mayores contra sustratos sintéticos pequeños en comparación con CBHI/Cel7A de *T. reesei*. La actividad específica contra Avicel fue claramente superior con las enzimas descritas en la presente memoria. Las bajas actividades de las preparaciones enzimáticas purificadas frente a xilano y CMC pueden ser debidas o bien a las propias propiedades de las proteínas, o al menos parcialmente a las pequeñas cantidades restantes de enzimas contaminantes. El principal producto final de la hidrólisis de la celulosa de todas las enzimas purificadas fue la celobiosa que es típica de las celobiohidrolasas.

Tabla 1. Actividades específicas (nkat/mg) de las celobiohidrolasas purificadas y la enzima de referencia de *T. reesei* (50°C, pH 5,0, 24 h).

Sustrato	Cel7A de <i>A. thermophilum</i> ALK04245	Cel7A de <i>C. thermophilum</i> ALK04265	Cel7A de <i>T. reesei</i>
Xilano	11,3	6,7	1,3
CMC	26,2	5,5	1,0
MUG2	9,2	18,9	4,3
MUG3	1,3	1,5	0,9
MUL	21,5	54,0	21,9
Avicel	1,8	1,4	0,6

La estabilidad térmica de las celobiohidrolasas purificadas se determinó a diferentes temperaturas. La reacción se realizó en presencia de BSA al 0,1% a pH 5,0 durante 60 minutos utilizando 4-metilumbeliferil- β -D-lactósido como sustrato. CBHI/Cel7A de *C. thermophilum* ALK04265 y CBHI/Cel7A de *A. thermophilum* ALK04245 fueron estables hasta 65°C y 60°C, respectivamente. La enzima de referencia de *T. reesei* (CBHI/Cel7A) conservó 100% de la actividad hasta 55°C.

Ejemplo 3. Purificación y caracterización de una endoglucanasa de *Acremonium thermophilum* ALK04245

Se cultivó *Acremonium thermophilum* ALK04245 como se describe en el Ejemplo 1. El sobrenadante del cultivo se incubó a 70°C durante 24 horas después de lo cual se concentró mediante ultrafiltración. La endoglucanasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba y de intercambio catiónico seguida de filtración en gel. La actividad endoglucanasa de las fracciones recogidas durante la purificación se determinó utilizando carboximetilcelulosa (CMC) como sustrato (Procedimiento de la IUPAC 1987). El contenido de proteína se midió mediante BioRad Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

El sobrenadante de cultivo concentrado se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Butyl FF equilibrada con tampón de fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 1 M. Las proteínas unidas se eluyeron con el gradiente lineal del tampón anterior a fosfato de potasio 5 mM, pH 6,0. Las fracciones se recogieron y la actividad endoglucanasa se determinó como se ha descrito anteriormente. La actividad endoglucanasa se eluyó en una amplia área de conductividad de 120 a 15 mS/cm.

Las fracciones combinadas se aplicaron a una columna de intercambio catiónico HiTrap SP XL equilibrada con acetato de sodio 8 mM, pH 4,5. Las proteínas unidas se eluyeron con un gradiente lineal de NaCl de 0 a 0,25 M en el tampón de equilibrado. La proteína que contenía actividad endoglucanasa se eluyó en la zona de conductividad de 3-7 mS/cm. Se repitió la cromatografía de intercambio catiónico y el eluato de proteína se concentró mediante

lío-filización.

La muestra disuelta se cargó en una columna de filtración en gel Superdex 75 HR10/30 equilibrada con tampón de fosfato de sodio 20 mM de pH 7,0, que contenía NaCl 0,15 M. La fracción principal de la proteína eluyó de la columna con el volumen de retención de 13,3 ml. Se estimó que el eluato de proteína era puro mediante electroforesis en gel de poliacrilamida-SDS y se evaluó que el peso molecular era de 40 kDa. Se determinó que la actividad específica de la proteína purificada, denominada At EG_40, a 50°C era de 450 nkat/mg (Procedimiento de la IUPAC 1987, utilizando CMC como sustrato).

La estabilidad térmica de la endoglucanasa purificada se determinó a diferentes temperaturas. La reacción se realizó en presencia de BSA de 0,1 mg/ml a pH 5,0 durante 60 minutos utilizando carboximetilcelulosa como sustrato. EG_40/Cel45A de *A. thermophilum* fue estable hasta 80°C. Las enzimas de referencia EGI (Cel7B) y EGII (Cel5A) de *T. reesei* conservaron 100% de la actividad hasta 60°C y 65°C, respectivamente.

Ejemplo 4. Purificación de una endoglucanasa de *Chaetomium thermophilum* ALKO4261

Se cultivó *Chaetomium thermophilum* ALKO4261 como se describe en el Ejemplo 1. La endoglucanasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba y de intercambio catiónico seguida por filtración en gel. La actividad endoglucanasa de las fracciones recogidas durante la purificación se determinó utilizando carboximetil celulosa (CMC) como sustrato (Procedimiento de IUPAC 1987).

Se añadió sulfato amónico al sobrenadante de cultivo para alcanzar a la misma conductividad que fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 1 M. La muestra se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Phenyl FF equilibrada con fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 1 M. La elución se llevó a cabo con un gradiente lineal de fosfato de potasio de 20 a 0 mM, pH 6,0, seguido de fosfato de potasio 5 mM, pH 6,0 y agua. Las proteínas unidas se eluyeron con un gradiente lineal de Urea de 0 a 6 M. Las fracciones se recogieron y la actividad endoglucanasa se analizó como se ha descrito anteriormente. La proteína que contenía actividad endoglucanasa se eluyó al principio del gradiente de urea.

Las fracciones se combinaron, se equilibraron con Tris-HCl 16 mM, pH 7,5 ($I = 1,4$ mS/cm) mediante una columna 10DG (Bio-Rad) y se aplicaron a una columna de intercambio aniónico HiTrap DEAE FF equilibrada con Tris-HCl 20 mM, pH 7,5. Las proteínas unidas se eluyeron con un gradiente lineal de NaCl de 0 a 1 M en el tampón de equilibrado. Las fracciones se recogieron y analizaron para determinar la actividad endoglucanasa como se ha descrito anteriormente. La proteína se eluyó en el intervalo de 10-20 mS/cm.

La muestra se equilibró con acetato de sodio 15 mM, pH 4,5 mediante una columna 10DG (Bio-Rad) y se aplicó a una columna de intercambio catiónico HiTrap SP XL equilibrada con acetato de sodio 20 mM de pH 4,5. Las proteínas se eluyeron con un gradiente lineal de acetato de sodio de 0 a 0,4 M, pH 4,5. La actividad endoglucanasa se eluyó en el intervalo de 1-10 mS/cm. La muestra recogida se liofilizó.

La muestra se disolvió en agua y se aplicó a una columna de filtración en gel Superdex 75 HR 10/30 equilibrada con fosfato de sodio 20 mM de pH 6,0, que contenía NaCl 0,15 M. Las fracciones se recogieron y las que contenían actividad endoglucanasa se combinaron. Se estimó que el eluato de proteína era puro mediante electroforesis en gel de poliacrilamida-SDS y se evaluó que la masa molecular sobre la base de patrones de masa molecular (patrones de SDS-PAGE preteñidos, Intervalo Amplio, Bio-Rad) era de 54 kDa. Se determinó que el pI de la proteína purificada, denominada Ct EG_54 con PhastSystem (Pharmacia) era de aprox. 5,5.

Ejemplo 5. Purificación de una endoglucanasa de *Thermoascus aurantiacus* ALKO4242

Se cultivó *Thermoascus aurantiacus* ALKO4242 como se describe en el Ejemplo 1. La endoglucanasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba y de intercambio aniónico seguida de filtración en gel. La actividad endoglucanasa de las fracciones recogidas durante la purificación se determinó utilizando carboximetil celulosa (CMC) como sustrato (Procedimiento de la IUPAC 1987). El contenido de proteína se midió mediante BioRad Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

El sobrenadante de cultivo se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Butyl equilibrada con tampón fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 0,7 M. Las proteínas unidas se eluyeron con $(\text{NH}_4)_2\text{SO}_4$ 0,2 M ($I = 39$ mS/cm). Las fracciones que contenían actividad endoglucanasa se combinaron y se concentraron mediante ultrafiltración.

La muestra se desaló en las columnas 10DG (Bio-Rad) y se aplicó a una columna de intercambio aniónico HiTrap DEAE FF equilibrada con Tris-HCl 15 mM, pH 7,0. Las proteínas unidas se eluyeron con un gradiente lineal de NaCl de 0 a 0,4 M en el tampón de equilibrado. La proteína que contenía actividad endoglucanasa se eluyó en la zona de conductividad de 15-21 mS/cm. Las fracciones recogidas se combinaron y se concentraron como antes.

La muestra se aplicó a una columna de filtración en gel Sephacryl S-100 HR 26/60 equilibrada con tampón acetato de sodio 50 mM de pH 5,0, que contenía NaCl 0,05 M. La fracción de proteína que contenía actividad endoglucanasa se eluyó de la columna con un volumen de retención correspondiente a un peso molecular de 16

kDa. Las fracciones recogidas se combinaron, se concentraron y se repitió la filtración en gel. Se estimó que el eluato de proteína era puro mediante electroforesis en gel de poliacrilamida-SDS y se evaluó que el peso molecular era de 28 kDa. Se determinó que el pI de la proteína purificada, denominada Ta EG_28, en un gel IEF (PhastSystem, Pharmacia) era de aproximadamente 3,5. Se determinó que la actividad específica de Ta EG_28 a 50°C era 4290 nkat/mg (Procedimiento de la IUPAC 1987, utilizando CMC como sustrato).

Ejemplo 6. Purificación y caracterización de una β -glucosidasa de *Acremonium thermophilum* ALKO4245

Se cultivó *Acremonium thermophilum* ALKO4245 como se describe en el Ejemplo 1. La β -glucosidasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba e intercambio aniónico seguida de filtración en gel. La actividad β -glucosidasa de las fracciones recogidas durante la purificación se determinó utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato (Bailey y Linko, 1990). El contenido de proteína se midió mediante BioRad Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

El sobrenadante de cultivo se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Phenyl Sepharose FF equilibrada con fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 1 M. Las proteínas unidas se eluyeron con un gradiente lineal del tampón de equilibrado a fosfato potásico 5 mM en la zona de conductividad de 137-16 mS/cm. Las fracciones recogidas se combinaron y se concentraron mediante ultrafiltración.

La muestra se desaló en columnas 10DG (Bio-Rad) y se aplicó a una columna de intercambio aniónico HiTrap DEAE FF equilibrada con fosfato de potasio 10 mM de pH 7,0. Las proteínas unidas se eluyeron con un gradiente lineal del tampón de equilibrado al mismo tampón que contenía NaCl 0,25 M en la zona de conductividad de 1,5-12 mS/cm. La cromatografía de intercambio aniónico se repitió como antes, excepto que se utilizó el tampón fosfato de potasio 4 mM de pH 7,2. Las proteínas se eluyeron en la zona de conductividad de 6-9 mS/cm. Las fracciones que contenían la actividad β -glucosidasa se recogieron, se combinaron y se concentraron.

El material activo de la cromatografía de intercambio aniónico se aplicó a una columna Sephacryl S-300 HR 26/60 equilibrada con fosfato de sodio 20 mM de pH 6,5, que contenía NaCl 0,15 M. La proteína con actividad β -glucosidasa se eluyó con un volumen de retención correspondiente a un peso molecular de 243 kDa. Se estimó que la proteína era pura mediante electroforesis en gel de poliacrilamida-SDS y se evaluó que el peso molecular era de 101 kDa. Se determinó que el pI de la proteína purificada, denominada At β G_101, en un gel de IEF (PhastSystem, Pharmacia) estaba en la zona de 5,6 a 4,9. Se determinó que la actividad específica de At β G_101 a 50°C era de 1100 nkat/mg (utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato, Bailey y Linko, 1990).

La estabilidad térmica de la β -glucosidasa purificada se determinó a diferentes temperaturas. La reacción se realizó en presencia de BSA de 0,1 mg/ml a pH 5,0 durante 60 minutos utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato. A. *thermophilum* β G_101 era estable hasta 70°C. La enzima de referencia de *Aspergillus* (Novozym 188) conserva 100% de la actividad hasta 60°.

Ejemplo 7. Purificación de una β -glucosidasa de *Chaetomium thermophilum* ALKO4261

Se cultivó *Chaetomium thermophilum* ALKO4261 como se describe en el Ejemplo 1. La β -glucosidasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba, de intercambio aniónico y catiónico seguido de filtración en gel. La actividad β -glucosidasa de las fracciones recogidas durante la purificación se determinó utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato (Bailey y Linko, 1990).

El sobrenadante de cultivo se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Phenyl Sepharose FF equilibrada con fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 0,8 M. La elución se llevó a cabo con un gradiente lineal del tampón de equilibrado de fosfato potásico 3 mM, pH 6,0, seguido de elución con agua y urea 6 M. Las primeras fracciones con actividad β -glucosidasa se eluyeron en el área de conductividad de 80-30 mS/cm. La segunda actividad β -glucosidasa se eluyó con urea 6 M. Las fracciones activas eluidas mediante urea se reunieron y se desalaron en columnas 10DG (BioRad) equilibradas con Tris-HCl 10 mM de pH 7,0.

Después de la desalación, la muestra se aplicó a una columna de intercambio aniónico HiTrap DEAE FF equilibrada con Tris-HCl 15 mM de pH 7,0. La proteína no se unió a la columna, pero se eluyó durante la alimentación de la muestra. Esta fracción de flujo continuo se desaló en las columnas 10DG (Bio-Rad) equilibradas con acetato de Na 7 mM, pH 4,5.

La muestra de la cromatografía de intercambio aniónico se aplicó a una columna de intercambio catiónico HiTrap SP FF equilibrada con acetato de sodio 10 mM de pH 4,5. Las proteínas unidas se eluyeron con un gradiente lineal de acetato de sodio de 10 mM a 400 mM, pH 4,5. Se recogieron las fracciones con actividad β -glucosidasa eluyendo en la zona de conductividad de 6,5 a 12 mS/cm, se desalaron en columnas 10DG (Bio-Rad) equilibradas con acetato de sodio 7 mM, pH 4,5 y se liofilizaron.

La muestra liofilizada se diluyó hasta 100 μ l de agua y se aplicó a una columna de filtración en gel Superdex 75 HF10/30 equilibrada con fosfato de sodio 20 mM de pH 4,5, que contenía NaCl 0,15 M. La actividad β -glucosidasa se eluyó a un volumen de retención de 13,64 ml. Las fracciones recogidas se combinaron, se liofilizaron y se disolvieron en agua. Se estimó que la proteína era pura mediante electroforesis en gel de poliacrilamida-SDS y se

evaluó que el peso molecular era de 76 kDa. La proteína se denominó Ct β G₇₆.

Ejemplo 8. Purificación y caracterización de una β -glucosidasa de *Thermoascus aurantiacus* ALKO4242

Se cultivó *Thermoascus aurantiacus* ALKO4242 como se describe en el Ejemplo 1. La β -glucosidasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba, e intercambio aniónico y catiónico seguido de filtración en gel. La actividad β -glucosidasa de las fracciones recogidas durante la purificación se determinó utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato (Bailey y Linko, 1990). El contenido de proteína se midió mediante BioRad Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

El sobrenadante de cultivo se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Phenyl Sepharose FF equilibrada con fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 0,7 M. Las proteínas unidas se eluyeron con un gradiente lineal de $(\text{NH}_4)_2\text{SO}_4$ 0,2 M a fosfato de potasio 5 mM, pH 6,0. La actividad β -glucosidasa se eluyó durante el gradiente en la zona de conductividad de 28,0 a 1,1 mS/cm. Las fracciones se combinaron y se concentraron mediante ultrafiltración.

La muestra se desaló en columnas 10DG (Bio-Rad) y se aplicó a una columna de intercambio aniónico HiTrap DEAE FF equilibrada con Tris-HCl 20 mM de pH 7,0. La enzima se eluyó con un gradiente lineal de NaCl de 0 a 0,2 M en el tampón de equilibrado y con elución retardada mediante Tris-HCl 20 mM, que contenía NaCl 0,4 M. La muestra que eluyó en la zona de conductividad de aprox. 10-30 mS/cm se concentró mediante ultrafiltración y se desaló mediante la columna 10DG (Bio-Rad).

La muestra se aplicó a una columna de intercambio catiónico HiTrap SP XL equilibrada con acetato de sodio 9 mM de pH 4,5. La enzima se eluyó con un gradiente lineal de NaAc de 10 mM a 400 mM y mediante elución retardada utilizando NaAc 400 mM de pH 4,5. Las proteínas con actividad β -glucosidasa se eluyeron ampliamente durante el gradiente lineal en la zona de conductividad de 5,0 a 11,3 mS/cm.

El material activo de la cromatografía de intercambio catiónico se aplicó a una columna Sephacryl S-300 HR 26/60 equilibrada con fosfato de sodio 20 mM de pH 7,0, que contenía NaCl 0,15 M. La proteína con actividad β -glucosidasa se eluyó con un volumen de retención correspondiente a un peso molecular de 294 kDa. Las fracciones recogidas se combinaron, se liofilizaron y se disolvieron en agua. Se estimó que la proteína era pura mediante electroforesis en gel de poli(acrilamida)-SDS y se evaluó que el peso molecular era de 81 kDa, lo que representa más probablemente la forma monomérica de la proteína. El isoelectroenfoque (IEF) se llevó a cabo utilizando un gel de pl 3-9. Después de la tinción con plata, se tiñó una amplia zona por encima de pl 5,85 además de una banda estrecha que correspondía a pl 4,55. Se determinó que la actividad específica de la proteína purificada, denominada Ta β G₈₁, a 50°C era de 600 nkat/mg utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato (Bailey y Linko, 1990).

La estabilidad térmica de la β -glucosidasa purificada se determinó a diferentes temperaturas. La reacción se realizó en presencia de BSA de 0,1 mg/ml a pH 5,0 durante 60 minutos utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato. β G₈₁ de *T. aurantiacus* era estable hasta 75°C. La enzima de referencia de *Aspergillus* (Novozym 188) conserva 100% de la actividad hasta 60°C.

Ejemplo 9. Purificación de una xilanasa a partir de *Acremonium thermophilum* ALKO4245

Se cultivó *Acremonium thermophilum* ALKO4245 como se describe en el Ejemplo 1. El sobrenadante del cultivo se incubó a 70°C durante 24 horas después de lo cual, se concentró mediante ultrafiltración. La xilanasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba y de intercambio catiónico seguida de filtración en gel. La actividad xilanasa se determinó utilizando xilano de abedul como sustrato (Procedimiento de la IUPAC 1987). La proteína se analizó mediante BioRad Protein Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

El sobrenadante de cultivo concentrado se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Butyl FF equilibrada con tampón fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 1 M. Las proteínas unidas se eluyeron con el gradiente lineal del tampón anterior a fosfato de potasio 5 mM, pH 6,0. La fracción de proteína se eluyó en una amplia zona de conductividad de 120 a 15 mS/cm.

La muestra de la columna de interacción hidrófoba se aplicó a una columna de intercambio catiónico HiTrap SP XL equilibrada con acetato de sodio 8 mM, pH 4,5. La proteína no se unió a esta columna pero se eluyó en el flujo continuo durante la alimentación de la muestra. Este eluto se concentró mediante ultrafiltración. La cromatografía hidrófoba se repitió como se ha descrito anteriormente. Las proteínas no unidas se recogieron y se liofilizaron.

La muestra disuelta se cargó en la columna de filtración en gel Superdex 75 HR10/30 equilibrada con tampón de fosfato de sodio 20 mM de pH 7,0, que contenía NaCl 0,15 M. Se estimó que la proteína eluida de la columna con el volumen de retención de 11,2 ml era pura mediante electroforesis en gel de poli(acrilamida)-SDS. La masa molecular de la proteína purificada evaluada sobre la base de patrones de masa molecular (patrones de SDS-PAGE preteñidos, Amplio Intervalo, Bio-Rad) era de 60 kDa. Se determinó que la actividad específica de la proteína, denominada At XYN₆₀, a 50°C era de 1800 nkat/mg (Procedimiento de la IUPAC 1987, utilizando como sustrato xilano de abedul). La actividad relativa aumentó aproximadamente 1,2 veces a 60°C y 1,65 veces a 70°C (10 min,

pH 5,0) en comparación con 50°C. Las actividades específicas contra MUG2 (4-metilumbeliferil-β-D-celobiósido), MUL (4-metilumbeliferil-beta-D-lactósido) y MUG3 (4-metilumbeliferil-β-D-celotriósido) fueron 54, 33 y 78 nkat/mg (50°C pH 5,0 10 min), respectivamente. Esto está de acuerdo con el hecho de que las xilanasas de la familia 10 también muestran actividad contra los arilglucopiranosidos (Biely *et al.* 1997).

5 Ejemplo 10. Purificación de una xilanasa de *Thermoascus aurantiacus* ALKO4242

Se cultivó *Thermoascus aurantiacus* ALKO4242 como se describe en el Ejemplo 1. La xilanasa pura se obtuvo mediante purificación sucesiva con cromatografía de interacción hidrófoba, e intercambio aniónico y catiónico seguido de filtración en gel. La actividad xilanasa se determinó utilizando xilano de abedul como sustrato (Procedimiento de IUPAC 1987). La proteína se analizó mediante BioRad Protein Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

El sobrenadante de cultivo se aplicó a una columna de interacción hidrófoba HiPrep 16/10 Phenyl Sepharose FF equilibrada con tampón fosfato de potasio 20 mM de pH 6,0, que contenía (NH₄)₂SO₄ 0,7 M. Las proteínas unidas se eluyeron con un protocolo de elución de dos etapas. La elución se llevó a cabo dejando caer la concentración de sal primero a (NH₄)₂SO₄ 0,2 M y después de eso se aplicó un gradiente lineal de fosfato de potasio 20 mM de pH 6,0, que contenía (NH₄)₂SO₄ 0,2 M a fosfato de potasio 5 mM de pH 6,0. La proteína se eluyó con (NH₄)₂SO₄ 0,2 M (I = 39 mS/cm).

La muestra se desaló en columnas 10DG (Bio-Rad) y se aplicó a una columna de intercambio aniónico HiTrap DEAE FF equilibrada con Tris-HCl 15 mM, pH 7,0. La proteína no se unió a la columna de intercambio aniónico pero se eluyó en el flujo continuo. La conductividad de la muestra se ajustó para que correspondiera a la de acetato de sodio 20 mM, pH 4,5 mediante la adición de agua y el pH se ajustó a 4,5 durante la concentración mediante ultrafiltración.

La muestra se aplicó a una columna de intercambio catiónico HiTrap SP XL equilibrada con acetato de sodio 20 mM, pH 4,5. Las proteínas unidas se eluyeron con un gradiente lineal del tampón de equilibrado al mismo tampón que contenía NaCl 1 M. La enzima se eluyó en la zona de conductividad de 1-7 mS/cm. La muestra se liofilizó y después de eso se disolvió en agua.

La muestra liofilizada se disolvió en agua y se aplicó a una columna de filtración en gel Superdex 75 HR 10/30 equilibrada con fosfato de sodio 20 mM de pH 7,0, que contenía NaCl 0,15 M. La proteína se eluyó de la columna con un volumen de retención correspondiente a un peso molecular de 26 kDa. Se estimó que la proteína era pura mediante electroforesis en gel de poliacrilamida-SDS. La masa molecular de la proteína pura fue de 30 kDa según se evaluó sobre la base de patrones de masa molecular (patrones de SDS-PAGE preteñidos, Intervalo Amplio, Bio-Rad). Se determinó que el pI de la proteína purificada, denominada Ta XYN_30 con PhastSystem (Pharmacia) era de aprox. 6.8. Se determinó que la actividad específica de Ta XYN_30 a 50°C era de 4800 nkat/mg (Procedimiento de la IUPAC 1987, utilizando como sustrato de xilano de abedul).

Ejemplo 11. Secuenciación de aminoácidos internos

Los péptidos internos se secuenciaron mediante ionización de electropulverización combinada con espectrometría de masas en tándem (ESI-MS/MS) utilizando el aparato Q-TOF1 (Micromass). La proteína se alquiló primero y se digirió a péptidos tripticos. Los péptidos generados se desalaron y se separaron parcialmente mediante nano-cromatografía líquida (de fase inversa) antes de aplicar el aparato Q-TOF1. Las secuencias de péptidos internos para las celobiohidrolasas de *Chaetomium thermophilum* y *Acremonium thermophilum* se muestran en la Tabla 2. Los péptidos de CBH de *Chaetomium* se obtuvieron después de haber clonado el correspondiente gen *cbh*. Los péptidos determinados a partir de CBH de *Acremonium* no se utilizaron en la clonación del gen correspondiente.

Tabla 2. Secuencias de péptidos internos determinadas a partir de CBH (1_C - 4_C) de *Chaetomium thermophilum* ALKO4265 y CBH (1_A - 4_A) de *Acremonium thermophilum* ALKO4245.

Péptido	Secuencia
Péptido 1_C	T P S T N D A N A G F G R
Péptido 2_C	V A F S N T D D F N R
Péptido 3_C	F S N T D D F N R K
Péptido 4_C	P G N S L / I T Q E Y C D A Q / K K
Péptido 1_A	V T Q F I / L T G

Péptido	Secuencia
Péptido 2_A	M G D T S F Y G P G
Péptido 3_A	C D P D G C D F N
Péptido 4_A	S G N S L I T T D F
I/L = la leucina y la isoleucina tienen la misma masa molecular y no se pueden distinguir en el análisis ESI-MS/MS Q/K = la masa molecular de la glutamina y lisina difiere solo en 0,036 Da y no se pueden distinguir en el análisis ESI-MS/MS	

Las secuencias de péptidos internos de las endoglucanasas, β -glucosidasas y xilanasas purificadas de *Acremonium thermophilum* ALKO4245, *Chaetomium thermophilum* ALKO4261 y *Thermoascus aurantiacus* ALKO4242 se enumeran en la Tabla 3, la Tabla 4 y la Tabla 5.

- 5 Tabla 3. Secuencias de péptidos internos determinadas a partir de endoglucanasas EG_40 de *Acremonium thermophilum* ALKO4245, EG_54 de *Chaetomium thermophilum* ALKO4261 y EG_28 de *Thermoascus aurantiacus* ALKO4242.

Proteína	Péptido	Secuencia ^a
At EG_40	Péptido 1	Q S C S S F P A P L K P G C Q W R
	Péptido 2	Y A L T F N S G P V A G K
	Péptido 3	V Q C P S E L T S R
	Péptido 4	N Q P V F S C S A D W Q R
	Péptido 5	Y W D C C K P S C G W P G K
	Péptido 6	P T F T
Ct EG_54	Péptido 1	E P E P E V T Y Y V
	Péptido 2	Y Y L L D Q T E Q Y
	Péptido 3	R Y C A C M D L W E A N S R
	Péptido 4	P G N T P E V H P Q / K
	Péptido 5	S I / L A P H P C N Q / K
	Péptido 6	Q Q Y E M F R
Ta EG_28	Péptido 7	A L N D D F C R
	Péptido 8	W G N P P P R
	Péptido 1	I / L T S A T Q W L R
	Péptido 2	G C A I / L S A T C V S S T I / L G Q E R

Proteína	Péptido	Secuencia ^(a)
	Péptido 3	P F M M E R
	Péptido 4	Q Y A V V D P H N Y G R
(a I/L = la leucina y la isoleucina tienen la misma masa molecular y no se pueden distinguir en el análisis ESI-MS/MS, Q/K = la masa molecular de la glutamina y lisina difiere solo en 0,036 Da y no se puede distinguir en el análisis ESI-MS/MS.		

Tabla 4. Secuencias de péptidos internos determinadas a partir de beta-glucosidasas β G_101 de *Acremonium thermophilum* ALKO4245, β G_76 de *Chaetomium thermophilum* ALKO4261 y β G_81 de *Thermoascus aurantiacus* ALKO4242.

Proteína	Péptido	Secuencia ^(a)
At β G_101	Péptido 1	S P F T W G P T R
	Péptido 2	V V V G D D A G N P C
	Péptido 3	A F V S Q L T L L E K
	Péptido 4	G T D V L I Y T P N N K
	Péptido 5	Q P N P A G P N A C V L I R
Ct β G_76	Péptido 1	E G L F I D Y R
	Péptido 2	P G Q S G T A T F R
	Péptido 3	E T M S S N V D D R
	Péptido 4	I A L V G S A A V V
	Péptido 5	M W L C E N D R
	Péptido 6	Y P Q L C L Q D G P L G I R
	Péptido 7	E L N G Q N S G Y P S I
Ta β G_81	Péptido 1	T P F T W G K
	Péptido 2	L C L Q D S L P G V R
	Péptido 3	G V D V Q L G P V A G V A P R
	Péptido 4	V N L T L E
	Péptido 5	F T G V F G E D V V G
	Péptido 6	N D L P L T G Y E K
(a I/L = la leucina y la isoleucina tienen la misma masa molecular y no se pueden distinguir en el análisis ESI-MS/MS)		

Tabla 5. Secuencias de péptidos internos determinadas a partir de xilanasas XYN_60 de *Acremonium thermophilum* ALKO4245 y XYN_30 de *Thermoascus aurantiacus* ALKO4242.

Proteína	Péptido	Secuencia
At XYN_60	Péptido 1	YNDYNLEYNQK
	Péptido 2	FGQVTPEN
	Péptido 3	VDGDATYMSYVNNK
	Péptido 4	KPAWTSVSSVLA AK
	Péptido 5	SQGDIVPRAK
Ta XYN_30	Péptido 1	VYFGVATDQNR
	Péptido 2	NAAIIQADFGQVTPENSMK
	Péptido 3	GHTLVWHSQLPSWVSSITDK
	Péptido 4	NHITTLMT R
	Péptido 5	AWDVVNEAFNEDGSLR
	Péptido 6	LYINDYNLDSASYPK
	Péptido 7	ASTTPLLLFDGNFNPKPAYNAIVQDLQQ
	Péptido 8	QTVFLNVIGEDIPIAFQTAR

5 Ejemplo 12. Construcción de bibliotecas genómicas para *Thermoascus aurantiacus*, *Chaetomium thermophilum* y *Acremonium thermophilum*

La biblioteca genómica de *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 se elaboraron para el vector Lambda DASH[®]II (Stratagene, USA) de acuerdo con las instrucciones del proveedor. Los ADN cromosómicos, aislados por el método de Raeder y Broda (1985), se digirieron parcialmente con Sau3A. Los ADN digeridos se fraccionaron por tamaños y los fragmentos del tamaño elegido (≈ 5 -23 kb) se desfosforilaron y se ligaron a los brazos del vector lambda digerido con BamHI. Las mezclas de ligación se empaquetaron utilizando extractos de empaquetamiento Gigapack III Gold de acuerdo con las instrucciones del fabricante (Stratagene, USA). Los títulos de las bibliotecas genómicas de *Chaetomium thermophilum* y *Acremonium thermophilum* fueron $3,6 \times 10^6$ pfu/ml y $3,7 \times 10^5$ pfu/ml y los de las bibliotecas amplificadas fueron $6,5 \times 10^{10}$ pfu/ml y $4,2 \times 10^8$ pfu/ml, respectivamente.

Se utilizó Lambda FIX[®] II/XhoI Partial Fill-In Vector Kit (Stratagene, USA) en la construcción de las bibliotecas genómicas para *Thermoascus aurantiacus* ALKO4242 y *Chaetomium thermophilum* ALKO4261 de acuerdo con las instrucciones del proveedor. Los ADN cromosómicos, aislados por el método de Raeder y Broda (1985), se digirieron parcialmente con Sau3A. Los ADN digeridos se fraccionan por tamaños y los fragmentos del tamaño seleccionado (≈ 6 -23 kb) se rellenaron y se ligaron a los brazos del vector Lambda Fix II digeridos con Xho I. Las mezclas de ligación se empaquetaron utilizando extractos de empaquetamiento Gigapack III Gold de acuerdo con las instrucciones del fabricante (Stratagene, USA). Los títulos de las bibliotecas genómicas de *Thermoascus aurantiacus* ALKO4242 y *Chaetomium thermophilum* ALKO4261 fueron $0,2 \times 10^6$ y $0,3 \times 10^6$ pfu/ml y los de las bibliotecas amplificadas fueron $1,8 \times 10^9$ y $3,8 \times 10^9$ pfu/ml, respectivamente.

25 Ejemplo 13. Clonación de los genes de celobiohidrolasa (*cbh/cel7*) de *Thermoascus aurantiacus*, *Chaetomium thermophilum* y *Acremonium thermophilum*

Se utilizaron métodos convencionales de biología molecular en el aislamiento y los tratamientos enzimáticos del ADN (plásmidos, fragmentos de ADN), en transformaciones en *E. coli*, etc. Los métodos básicos utilizados se describen en los manuales de biología molecular convencionales, por ejemplo, Sambrook *et al.* (1989) y Sambrook y Russell (2001).

- 5 Las sondas para el escrutinio de las bibliotecas genómicas construidas como se describe en el Ejemplo 12 se amplificaron mediante PCR utilizando los ADN genómicos de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 como moldes en las reacciones. Se diseñaron varios cebadores sometidos a ensayo en reacciones de PCR de acuerdo con la secuencia de nucleótidos publicada (documento WO 03/000941, Hong *et al.*, 2003b). Las mezclas de reacción de PCR contenían Tris-HCl 50 mM, pH 9,0, (NH₄)₂SO₄ 15 mM, Triton X-100 al 0,1%, MgCl₂ 1,5 mM, dNTP 0,2 mM, 5 µM de cada cebador y 1 unidad de ADN polimerasa Dynazyme EXT (Finnzymes, Finlandia) y ≈ 0,5-1 µg de ADN genómico. Las condiciones para las reacciones de PCR fueron las siguientes: 5 min de desnaturalización inicial a 95°C, seguido de 30 ciclos de 1 min a 95°C, recocido de 1 min a 62°C (± 8°C gradiente) para moldes de *Thermoascus* ALKO4242 y *Chaetomium* ALKO4265 o recocido de 1 min a 58°C (± 6°C gradiente) para el molde de *Acremonium* ALKO4245, 2 min de extensión a 72°C y una extensión final a 72°C durante 10 minutos.

Los productos de ADN de los tamaños esperados (calculados a partir de secuencias de *cbh* publicadas) se obtuvieron de todos los moldes genómicos utilizados. Los fragmentos de ADN de los tamaños esperados se aislaron de las reacciones de PCR más específicas y se clonaron al vector pCR® Blunt-TOPO® (Invitrogen, USA). Los insertos se caracterizaron por secuenciación y realizando hibridaciones mediante transferencia Southern de los ADN genómicos digeridos con varias enzimas de restricción. Los fragmentos de PCR, que fueron seleccionados para ser utilizados como sondas para la detección de las bibliotecas genómicas de *Thermoascus aurantiacus*, *Chaetomium thermophilum* y *Acremonium thermophilum* se presentan en la Tabla 6.

Tabla 6. Cebadores utilizados en las reacciones de PCR y las sondas seleccionadas para el escrutinio de los genes *cbh/cel7* de las bibliotecas genómicas de *Thermoascus aurantiacus*, *Chaetomium thermophilum* y *Acremonium thermophilum*. Se muestran el molde de ADN genómico y el nombre del plásmido que contiene el fragmento de la sonda.

Gen	Cebador directo	Cebador inverso	Molde de ADN	Fragmento (kb)	Plásmido
Ta <i>cbh</i>	TCEL11 atgcgaactggcgttggtcc	TCEL12 gaatttgagctagtgctgacg	<i>Thermoascus</i> ALKO4242	0,8 kb	pALK1633
Ct <i>cbh</i>	TCEL7 cgatgccaaactggcgctggac	TCEL8 ttcttggtggtgctgacggtc	<i>Chaetomium</i> ALKO4265	0,8 kb	pALK1632
At <i>cbh</i>	TCEL13 aggtcgaccaactgctacacg	TCEL4 accgtgaacttctgctggtg	<i>Acremonium</i> ALKO4245	0,7 kb	pALK1634

Las secuencias de aminoácidos deducidas de todas estas sondas tenían homología con varias secuencias de CBH publicadas (programa BLAST, versión 2.2.9 en el NCBI, National Center for Biotechnology Information; Altschul *et al.*, 1990), de las glucósido hidrolasas de la familia 7 (Henrissat, 1991; Henrissat y Bairoch, 1993).

Los insertos de los plásmidos enumerados en la Tabla 6 fueron marcados con digoxigenina de acuerdo con las instrucciones del proveedor (Roche, Alemania), y las bibliotecas genómicas amplificadas (2×10^5 - 3×10^5 placas) se escrutaron con los fragmentos de la sonda marcada. La temperatura de hibridación para los filtros fue de 68°C y los filtros se lavaron 2 x 5 min a RT utilizando 2 x SSC - SDS al 0,1% seguido de 2 x 15 min a 68°C utilizando 0,1 x SSC - SDS al 0,1% con las sondas homólogas utilizadas. Se obtuvieron varias placas positivas de cada una de las hibridaciones. En el escrutinio de las bibliotecas genómicas de *Acremonium* ALKO4245, algunas de las placas positivas hibridaron fuertemente con la sonda en cuestión, pero, además, hubo una cantidad de placas que hibridaron más débilmente con las sondas. Esto sugirió que podrían estar presentes otro u otros genes de celobiohidrolasa en el genoma, causando reacción cruzada. Se purificaron de cuatro a cinco placas con fuerte hibridación de los escrutinios de las bibliotecas genómicas de *Thermoascus* ALKO4242 y *Chaetomium* ALKO4265. En el caso de *Acremonium thermophilum* ALKO4245, cuatro de las seis placas purificadas hibridaron débilmente con la sonda utilizada. Los ADN de los fagos se aislaron y se caracterizaron mediante hibridaciones de transferencia Southern. Los fragmentos de restricción seleccionados que hibridaban con la sonda se subclonaron en el vector pBluescript II KS+ y se secuenciaron las regiones relevantes de los clones.

En total se clonaron cuatro genes *cbh/cel7*; uno de *Thermoascus aurantiacus* ALKO4242, uno de *Chaetomium thermophilum* ALKO4265 y dos de *Acremonium thermophilum* ALKO4245 (en la fase temprana del trabajo, éstos

tenían los códigos At_cbh_C y At_cbh_A, y a continuación se denominaron At cel7A y At cel7B, respectivamente). La Tabla 7 resume la información sobre las sondas utilizadas para la detección de los genes, los clones de fagos de los que se aislaron los genes, los fragmentos de restricción seleccionados que contenían los genes completos con sus regiones promotora y terminadora, los nombres de plásmido, y los números de depósito DSM para las cepas de *E. coli* que portan estos plásmidos.

Tabla 7. Sondas utilizadas para la clonación de los genes *cbh/cel7*, clon de fago y subclones seleccionados, número de plásmido y número del depósito de la cepa de *E. coli* correspondiente.

Gen	Sonda utilizada en el escrutinio	Clon de fago	Fragmento subclonado en pBluescript II	Núm. de Plásmido	Núm. de depósito de <i>E. coli</i>
Ta cel7A	pALK1633	F12	XbaI de 3,2 kb	pALK1635	DSM 16723
Ct cel7A	pALK1632	F36	PvuI - HindIII de 2,3 kb	pALK1642	DSM 16727
At cel7B	pALK1634	F6	EcoRI de 3,1 kb	pALK1646	DSM 16728
At cel7A	pALK1634	F2	XhoI de 3,4 kb	pALK1861	DSM 16729

La información relevante sobre los genes y las secuencias de proteínas deducidas (SEQ ID NO: 1-8) se resumen en la Tabla 8 y la Tabla 9, respectivamente.

Las secuencias peptídicas de las proteínas CBH purificadas de *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 (Tabla 2) se encontraron partir de las secuencias de aminoácidos deducidas de los clones que contenían los genes Ct cel7A y At cel7A. Por lo tanto, se puede concluir que se clonaron los genes que codifican las proteínas CBH/Cel7 purificadas de *Chaetomium thermophilum* y *Acremonium thermophilum*.

Tabla 8. Compendio de los genes *cbh/cel7* aislados de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245.

Gen Cbh	Longitud con intrones (pb) ^(a)	Región codificante (pb) ^(b)	Núm. de intrones	Longitud de los intrones (pb)	SEQ ID NO:
Ta cel7A	1439	1371	1	65	1
Ct cel7A	1663	1596	1	64	7
At cel7B	1722	1377	3	134, 122, 87	3
At cel7A	1853	1569	4	88, 53, 54, 86	5

(a) Está incluido el codón de PARADA.

(b) No está incluido el codón de PARADA.

Tabla 9. Resumen de las secuencias de aminoácidos deducidas de las secuencias de los genes *cbh/cel7* de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245. ss, secuencia señal.

Proteína CBH	Núm. de aa	Longitud de NN/HMM ^(a)	ss	CBD C-terminal ^(b)	PM pronosticado (Da, ss no incl.) ^(c)	pl pronosticado (ss no incl.)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
Ta Cel7A	457	17/17		NO	46873	4,44	2	2

Proteína CBH	Núm. de aa	Longitud de ss NN/HMM ^(a)	CBD C-terminal ^(b)	PM pronosticado (Da, ss no incl.) ^(c)	pl pronosticado (ss no incl.)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
Ct Cel7A	532	18/18	SI, T497 a L532	54564	5,05	3	8
At Cel7B	459	21/21	NO	47073	4,83	2	4
At Cel7A	523	17/17	SI, Q488 a L523	53696	4,67	4	6

(a) La predicción de la secuencia señal se realizó utilizando el programa SignalP V3.0 (Nielsen *et al.*, 1997; Bendtsen *et al.*, 2004); el valor NN se obtuvo utilizando redes neuronales y valor HMM utilizando modelos ocultos de Markov.

(b) Dominio de unión a celulosa (CBD), se indican los aminoácidos de la región CBD C-terminal (M1 (Met Núm. 1) incluida en la numeración)

(c) No se incluyó la secuencia señal pronosticada. La predicción se realizó utilizando la herramienta Compute pl/MW del servidor ExPASy (Gasteiger *et al.*, 2003).

(d) Número de secuencias NXS/T.

5 Las secuencias de aminoácidos deducidas de Cel7A de *Thermoascus aurantiacus* y de Cel7A de *Acremonium thermophilum* (núcleo, sin el CBD) fueron más homólogas entre sí (analizadas mediante alineamiento global de Needleman-Wunsch, EMBOSS 3.0.0 Needle, con Matriz EBLOSUM62, Penalización por Hueco 10,0 y Penalización por extensión 0,5; Needleman y Wunsch, 1970). Además, la Cel7A de *Acremonium thermophilum* deducida tenía una identidad inferior a Cel7A de *Chaetomium thermophilum* deducida. Cel7B de *Acremonium thermophilum* era muy distinta de las secuencias de CBH/Cel7 descritas aquí.

10 La secuencia de Cel7A de *Chaetomium* deducida poseía las identidades más altas (analizada mediante alineamiento global de Needleman-Wunsch, EMBOSS Needle, véase más arriba) con polipéptidos de CBHI de *Chaetomium thermophilum*, *Scytalidium thermophilum* y *Thielavia australiensis* descritos en el documento WO 03/000941. Del mismo modo, la secuencia de Cel7A de *Thermoascus aurantiacus* deducida era altamente idéntica a la de CBHI publicada de *Thermoascus aurantiacus* (documento WO 03/000941, Hong *et al.*, 2003b). Cel7B de *Acremonium thermophilum* tenía identidades significativamente inferiores a las secuencias previamente publicadas, estando más estrechamente relacionado con el polipéptido CBHI de *Oryza sativa*. Los mayores homólogos de la secuencia de Cel7A de *Acremonium thermophilum* deducida fueron los polinucleótidos de CBHI de *Exidia gladiosa* y *Acremonium thermophilum* (documento WO 03/000941). El alineamiento indica que las secuencias clonadas de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALK04245 codifican las proteínas CBH que tenían alta homología con los polipéptidos de la familia 7 de glucósido hidrolasa, por lo tanto éstos fueron denominados Cel7A o Cel7B (Henrissat *et al.* 1998).

20 La comparación de las secuencias de aminoácidos deducidas de los genes *cbh/cel7* de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 *Thielavia* entre sí, y adicionalmente con las secuencias encontradas a partir de las bases de datos, se muestran en la Tabla 10.

25 Tabla 10. Secuencias de homología más alta con las secuencias de aminoácidos deducidas de los genes *cbh/cel7* de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALK04265 y *Acremonium thermophilum* ALK04245. El alineamiento se realizó utilizando alineamiento global de Needleman-Wunsch (EMBOSS Needle, Penalización por hueco 10,0, Penalización por extensión 0,5). *Indica una secuencia de aminoácidos derivada de uno de los genes de celobiohidrolasa clonados en este trabajo. 'Núcleo' indica el alineamiento sin el CBD.

Organismo, enzima y número de acceso	Identidad, (%)
* <i>Thermoascus aurantiacus</i> Cel7A	100,0

Organismo, enzima y número de acceso	Identidad, (%)
<i>Thermoascus aurantiacus</i> , AY840982	99,6
<i>Thermoascus aurantiacus</i> , AX657575	99,1
<i>Thermoascus aurantiacus</i> , AF421954	97,8
<i>Talaromyces emersonii</i> , AY081766	79,5
<i>Chaetomidium pingtungium</i> , AX657623	76,4
<i>Trichophaea saccata</i> , AX657607	73,4
* <i>Acremonium thermophilum</i> Cel7A (núcleo)	70,6
<i>Emericella nidulans</i> , AF420020 (núcleo)	70,4
* <i>Chaetomium thermophilum</i> Cel7A (núcleo)	66,4
* <i>Chaetomium thermophilum</i> Cel7A	100,0
<i>Chaetomium thermophilum</i> , AY861347	91,9
<i>Chaetomium thermophilum</i> , AX657571	91,7
<i>Scytalidium thermophilum</i> , AX657627	74,7
<i>Thielavia australiensis</i> , AX657577	74,6
<i>Acremonium thermophilum</i> , AX657569	72,3
<i>Exidia glandulosa</i> , AX657613	68,0
* <i>Acremonium thermophilum</i> Cel7A	66,9
* <i>Thermoascus aurantiacus</i> Cel7A (núcleo)	66,4
<i>Exidia glandulosa</i> , AX657615	60,8
<i>Chaetomium pingtungium</i> , AX657623	60,7
* <i>Acremonium thermophilum</i> Cel7B (núcleo)	60,2
* <i>Acremonium thermophilum</i> Cel7B	100,0
<i>Oryza sativa</i> , AK108948	66,1
<i>Exidia glandulosa</i> , AX657615	65,0
<i>Acremonium thermophilum</i> , AX657569 (núcleo)	64,8
<i>Thermoascus aurantiacus</i> , AX657575	64,8

Organismo, enzima y número de acceso	Identidad, (%)
* <i>Acremonium thermophilum</i> Cel7A	64,6
* <i>Thermoascus aurantiacus</i> Cel7A	64,4
<i>Trichophaea saccata</i> , AX657607	63,6
* <i>Chaetomium thermophilum</i> Cel7A (núcleo)	60,2
* <i>Acremonium thermophilum</i> Cel7A	100,0
<i>Exidia glandulosa</i> , AX657613	77,9
<i>Exidia glandulosa</i> , AX657615	77,9
<i>Acremonium thermophilum</i> , AX657569	77,5
<i>Thielavia australiensis</i> , AX657577	71,0
* <i>Thermoascus aurantiacus</i> Cel7A (núcleo)	70,6
<i>Scytalidium Thermophilum</i> , AX657627	67,5
<i>Chaetomium thermophilum</i> , AX657571	67,5
<i>Chaetomium pingtungium</i> , AX657623	67,3
* <i>Chaetomium thermophilum</i> Cel7A	66,9
* <i>Acremonium thermophilum</i> Cel7B (núcleo)	64,6

Ejemplo 14. Producción de proteínas CBH/Cel7 recombinantes en *Trichoderma reesei*

Se construyeron plásmidos de expresión para la producción de las proteínas CBH/Cel7 recombinantes de *Thermoascus aurantiacus* (Ta Cel7A), *Chaetomium thermophilum* (Ct Cel7A) y *Acremonium thermophilum* (At Cel7A, At Cel7B; en la fase inicial del trabajo estas proteínas tenían los códigos temporales At CBH_C y CBH_A, respectivamente). Los plásmidos de expresión contruidos se enumeran en la Tabla 11. Los genes *cbh/cel7* recombinantes, incluyendo sus propias secuencias señal, se fusionaron exactamente al promotor *cbh1* (*cel7A*) de *T. reesei* mediante PCR. La terminación de la transcripción fue asegurada por el terminador *cel7A* de *T. reesei* y se utilizó el gen marcador *amdS* de *A. nidulans* para la selección de los transformantes como describen Paloheimo *et al.* (2003). Los casetes de expresión lineal (Fig. 2), se aislaron de las cadenas principales del vector después de la digestión con EcoRI y se transformaron en protoplastos de *T. reesei* A96 y A98 (ambas cepas tienen los genes que codifican las cuatro principales celulasas CBHI/Cel7A, CBHII/Cel6A, EGI/Cel7B y EGII/Cel5A suprimidos). Las transformaciones se realizaron como en Penttilä *et al.* (1987) con las modificaciones descritas por Karhunen *et al.* (1993), seleccionando con acetamida como única fuente de nitrógeno. Los transformantes se purificaron en placas de selección a través de conidios individuales antes de esporularlos en PD.

Tabla 11. Casetes de expresión contruidos para producir las proteínas CBH/Cel7 de *Thermoascus aurantiacus* ALKO4242 (Ta Cel7A), *Chaetomium thermophilum* ALK04265 (Ct Cel7A), y *Acremonium thermophilum* ALKO4245 (At Cel7A, At Cel7B) en *Trichoderma reesei*. La estructura general de las casetes de expresión fue la descrita en la Fig. 2. Los genes *cbh/cel7* clonados se fusionaron exactamente al promotor *cbh1/cel7A* de *T. reesei*.

CBH/Cel7	Plásmido de expresión	Tamaño de casete de expresión ^(a)	Terminador cel7A ^(b)
Ta Cel7A	pALK1851	9,0 kb	245 bp (<i>Xba</i> I)
Ct Cel7A	pALK1857	9,2 kb	240 bp (<i>Hind</i> III)
At Cel7B	pALK1860	9,4 kb	361 bp (<i>Eco</i> RI)
At Cel7A	pALK1865	9,5 kb	427 bp (<i>Eco</i> RV)

(a) El casete de expresión para la transformación en *T. reesei* se aisló de la cadena principal del vector utilizando la digestión con *Eco*RI.

(b) Número de los nucleótidos de la región del terminador *cbh1/cel7A* genómico después del codón de PARADA. El sitio de restricción en el extremo 3', utilizado en la escisión del fragmento de gen genómico, se incluye en el paréntesis.

5

La producción de CBH/Cel7 de los transformantes se analizó a partir de los sobrenadantes de cultivo de los cultivos en matraz oscilante (50 ml). Los transformantes se cultivaron durante 7 días a 28°C en un medio inductor de celulasas basado en complejo de lactosa (Joutsjoki *et al.* 1993) tamponado con KH_2PO_4 al 5%. La actividad celobiohidrolasa se analizó utilizando sustrato de 4-metilumbeliferil- β -D-lactósido (MUL) de acuerdo con van Tilbeurgh *et al.*, 1988. Los genotipos de los transformantes seleccionados se confirmaron mediante el uso de transferencias Southern en las que se incluyeron varios productos digeridos genómicos y el casete de expresión respectivo se utilizó como sonda. La expresión heteróloga de las proteínas Ta Cel7A, Ct Cel7A, At Cel7A y At Cel7B se analizó mediante SDS-PAGE con tinción de Coomassie subsiguiente. Los hallazgos de que no se pudieron detectar actividad celobiohidrolasa o producción de proteínas heterólogas en la SDS-PAGE para los transformantes At Cel7B que contenían el casete de expresión integrado, sugieren que At Cel7B se produce por debajo de los niveles de detección en *Trichoderma* utilizando el diseño experimental descrito.

10

15

Las preparaciones de enzimas CBH/Cel7 recombinantes se caracterizaron en términos de óptimo de pH y estabilidad térmica. El óptimo de pH de las proteínas CBH/Cel7 recombinantes de *Thermoascus aurantiacus*, *Chaetomium thermophilum*, y *Acremonium thermophilum* se determinó en el tampón de Mcllvaine universal dentro de un intervalo de pH de 3,0-8,0 utilizando 4-metilumbeliferil- β -D-lactósido (MUL) como sustrato (Fig 3 A). El óptimo de pH para las enzimas Ct Cel7A y At Cel7A está en 5,5, por encima del cual la actividad empieza a disminuir gradualmente. El pH óptimo de la Ta Cel7A bruta recombinante está en 5,0 (Fig 3 A). La estabilidad térmica de las enzimas Cel7 recombinantes se determinó midiendo la actividad MUL en tampón universal de Mcllvaine al pH óptimo con el tiempo de reacción de 1 h. Como se muestra a partir de los resultados Ta Cel7A y Ct Cel7A conservaron más de 60% de sus actividades a 70°C, mientras que Cel7A demostró ser claramente menos estable a las temperaturas más altas ($\geq 65^\circ\text{C}$) (Figura 3 B).

20

25

Los transformantes CBH/Cel7 seleccionados se cultivaron en biorreactores de laboratorio a 28°C en el medio indicado anteriormente durante 3-4 días con el control de pH $4,4 \pm 0,2$ ($\text{NH}_3/\text{H}_3\text{PO}_4$) para obtener material para las pruebas de aplicación. Los sobrenadantes se recuperaron mediante centrifugación y se filtraron a través de filtros Seitz K-150 y EK (Pall SeitzSchenk Filtersystems GmbH, Bad Kreuznach, Alemania).

30

Ejemplo 15. Producción de las proteínas de fusión Cel7A+CBD de *Thermoascus aurantiacus* recombinantes en *T. reesei*

Cel7A de *Thermoascus aurantiacus* (AF478686, Hong *et al.*, 2003b; SEQ ID. NO: 1) se fusionó al conector y al CBD de CBH/Cel7A de *Trichoderma reesei* (AR088330, Srisodsuk *et al.* 1993) (= Tr CBD), seguido de la producción de la proteína de fusión (SEQ ID NO: 28 que correspondía a SEQ ID NO: 27 de ácido nucleico) en *T. reesei* como se describe en FI20055205/US 11/119.526; presentada el 29 de Abril de 2005. Además, Cel7A de *Thermoascus aurantiacus* se fusionó al conector y al CBD de Cel7A de *Chaetomium thermophilum* (SEQ ID NO.: 7) (Ct CBD). Para ello, la secuencia codificante del conector y el CBD de Cel7A de *Chaetomium thermophilum* se sintetizaron mediante PCR utilizando los siguientes cebadores:

35

5'-TTAAACATATGTTATCTACTCCAACATCAAGGTCGGACCCATCGGCTCGACCGTCCCTGGCCTTGAC-3' (secuencia directa)

40

y

5'-TATATGCGGCCGCAAGCTTTACCATCAAGTTACTCCAGCAAATCAGGGAAGT-3' (secuencia inversa)

La mezcla de reacción de PCR contenía tampón de reacción EXT 1x DyNAzyme™ (Finnzymes, Finlandia), Mg²⁺ 15 mM, dNTP 0,2 mM, 2 mM de cada cebador, 0,6 unidades de DyNAzyme™ ADN polimerasa EXT (Finnzymes, Finlandia), y aproximadamente el 75 ng/30 µl de ADN molde, que contenía el gen *cel7A* completo de *Chaetomium thermophilum*. Las condiciones para la reacción de PCR fueron las siguientes: 2 min de desnaturalización a 98°C, seguido de 30 ciclos de 30 segundos a 98°C, 30 s de recocido a 68°C (gradiente de ±4°C), extensión de 30 segundos a 72°C y una extensión final a 72°C durante 10 minutos. El fragmento de ADN específico en la reacción de PCR se obtuvo en el intervalo de temperatura de hibridación de 64°C a 68,5°C. El fragmento CBD sintetizado *Chaetomium thermophilum* se ligó después del gen *cel7A* de *Thermoascus aurantiacus* dando como resultado un punto de unión de GPIGST entre los dominios. El fragmento amplificado por PCR en el plásmido se confirmó mediante secuenciación (SEQ ID NO: 29). La gen *cel7A* de fusión construido se fusionó exactamente al promotor *cbh1* (*cel7A*) de *T. reesei*. La terminación de la transcripción fue asegurada por el terminador *cel7A* de *T. reesei* y el gen marcador *amdS* de *A. nidulans* se utilizó para la selección de los transformantes como describen Paloheimo *et al.* (2003).

El casete de expresión lineal se aisló de la cadena principal del vector después de la digestión con *NotI* y se transformó en protoplastos de *T. reesei* A96. Las transformaciones se realizaron como en Penttilä *et al.* (1987) con las modificaciones descritas por Karhunen *et al.* (1993), seleccionando con acetamida como única fuente de nitrógeno. Los transformantes se purificaron en placas de selección a través de conidios individuales antes de esporularlos en PD.

La producción de los transformantes de Cel7A + CBD de *Thermoascus aurantiacus* (SEQ ID NO: 28 y 30) se analizó a partir de los sobrenadantes de cultivo de los cultivos en matraz oscilante (50 ml). Los transformantes se cultivaron durante 7 días en un medio inductor de celulasa complejo (Joutsjoki *et al.* 1993) tamponado con 5% de KH₂PO₄ a pH 5,5. La actividad celobiohidrolasa se analizó utilizando sustrato de 4-metilumbeliferil-β-D-lactósido (MUL) de acuerdo con van Tilbeurgh *et al.*, 1988. Los genotipos de los transformantes seleccionados se confirmaron mediante el uso de transferencias de Southern en las que se incluyeron diversos productos digeridos genómicos y el casete de expresión se utilizó como sonda. Los análisis de SDS-PAGE mostraron que las enzimas Cel7A+CBD de *Thermoascus aurantiacus* recombinantes se produjeron en forma de proteínas de fusión estables en *T. reesei*.

El transformante seleccionado que producía la proteína de fusión Ta Cel7A+Tr CBD (SEQ ID NO: 28) también se cultivó en un biorreactor de 2 litros a 28°C en el medio indicado anteriormente durante 3-4 días con control de pH de 4,4 ± 0,2 (NH₃/H₃PO₄) para obtener material para los ensayos de aplicación. Los sobrenadantes se recuperaron mediante centrifugación y se filtraron a través de filtros Seitz K-150 y EK (Pall SeitzSchenk Filtersystems GmbH, Bad Kreuznach, Alemania).

Ejemplo 16. Comparación de las constantes de Michaelis-Menten y de inhibición de celobiosa de celobiohidrolasas recombinantes purificadas

Las constantes de Michaelis-Menten y de inhibición de celobiosa se determinaron a partir de las celobiohidrolasas producidas de forma heteróloga en *T. reesei* (Ejemplos 14 y 15). Las enzimas se purificaron como se describe en el Ejemplo 2. Se midieron las concentraciones de proteína de las enzimas purificadas por su absorción a 280 nm utilizando un coeficiente de extinción molar teórico, que se calcularon a partir de las secuencias de aminoácidos (Gill y von Hippel, 1989).

Las constantes cinéticas (valores Km y Kcat) y la constante de inhibición de celobiosa (Ki) para Tr CBH/Cel7A, Ta CBH/Cel7A, At CBH/Cel7A y Ct CBH/Cel7A, se midieron utilizando CNPLac (2-Cloro-4-nitrofenil-β-D-lactósido) como sustrato a temperatura ambiente (22°C) en tampón de fosfato de sodio 50 mM, pH 5,7. Para la determinación de la constante de inhibición (Ki), se utilizaron ocho concentraciones de sustrato diferentes (31 a 4.000 µM) en presencia de un intervalo de cinco concentraciones de inhibidor (0-100 µM o 0-400 µM), que enmarcan el valor de Ki. Todos los experimentos se realizaron en placas de microtitulación y el volumen de reacción total fue de 200 µl. Las tasas iniciales se midieron en cada caso controlando continuamente la liberación del anión cloro-nitrofenolato (CNP, 2-cloro-4-nitrofenolato) a través de mediciones a 405 nm utilizando un lector de placas de microtitulación Varioscan (ThermoLabsystems). Los resultados se calcularon a partir curva patrón de CNP (de 0 a 100 µM). Las concentraciones de enzimas utilizadas fueron: Tr CBH/Cel7A 2,46 µM, Ta CBH/Cel7A 1,58 µM, Ct CBH/Cel7A 0,79 µM y At CBH/Cel7A 3 µM. Las constantes Km y kcat se calcularon a partir del ajuste de la ecuación de Michaelis-Menten utilizando el programa de Origin. Se utilizaron diagramas de Lineweaver-Burk, diagramas secundarios (pendiente LWB frente a [Glc2; celobiosa]) y diagramas de Hanes para distinguir entre la inhibición de tipo competitivo y de tipo mixto y para determinar las constantes de inhibición (Ki).

Los resultados de las mediciones cinéticas se muestran en la Tabla 12 y la Tabla 13. Como se puede observar, Ct CBH/Cel7A tiene claramente el número de recambio más alto (kcat) en CNPLac y también la constante de especificidad (kcat/Km) es mayor en comparación con CBH/Cel7A de *T. reesei*. La celobiosa (Glc2) es un inhibidor competitivo para todas las celulasas medidas, y Tr CBH/Cel7A (utilizada como control) tiene la inhibición más fuerte

(es decir, el valor más bajo de K_i) por celobiosa. At CBH/Cel7A tenía una constante de inhibición más de 7 veces mayor en comparación con la de Tr CBH/Cel7A. Estos resultados indican que las tres nuevas celobiohidrolasas podría funcionar mejor en la hidrólisis de celulosa debido a la disminución de la inhibición por celobiosa en comparación con la celobiohidrolasa I Cel7A de *Trichoderma reesei*.

- 5 Tabla 12. Comparación de las constantes de inhibición por celobiosa de cuatro celobiohidrolasas GH de la familia 7, medidas en CNPLac en tampón de fosfato de sodio 50 mM de pH 5,7, a 22°C.

Ct Cel7A	39	competitiva
Ta Cel7A	107	competitiva
At Cel7A	141	competitiva
Tr Cel7A	19	competitiva

- 10 Tabla 13. Comparación de las constantes cinéticas de Michaelis-Menten de celobiohidrolasa Cel7A de *Chaetomium thermophilum* con CBH/Cel7A de *T. reesei*, medidas en CNPLac en tampón de fosfato de sodio 50 mM de pH 5,7, a 22°C.

Enzima	kcat (min^{-1})	Km (μM)	kcat /Km ($\text{min}^{-1} \text{M}^{-1}$)
Ct Cel7A	18,8	1960	9,5 10 ³
Tr Cel7A	2,6	520	5,0 10 ³

Ejemplo 17. Hidrólisis de celulosa cristalina (Avicel) por las celobiohidrolasas recombinantes

- 15 Los celobiohidrolasas recombinantes purificadas Ct Cel7A, Ta Cel7A, Ta Cel7A + Tr CBD, Ta Cel7A + Ct CBD, At Cel7A así como la versión núcleo de Ct Cel7A (véase más adelante) se sometieron a ensayo en cantidades equimolares en la hidrólisis de celulosa cristalina a dos temperaturas, 45°C y 70°C; Tr Cel7A de *T. reesei* purificada y su versión núcleo (véase más adelante) se utilizaron como comparación. Los análisis de hidrólisis de celulosa cristalina (Ph 101, Avicel; Fluka, Bucsh, Suiza) se realizaron en acetato de sodio 50 mM con escala de tubo de 1,5 ml, pH 5,0. Se sacudió Avicel a 45°C o a 70°C, con la disolución de enzima (1,4 μM), y el volumen final de la mezcla de reacción fue de 325 μl . La hidrólisis prosiguió hasta 24 horas tomando muestras en seis momentos diferentes y deteniendo la reacción mediante la adición de 163 μl de reactivo de parada que contenía 9 vol de etanol del 94% y 1 vol de glicina 1 M (pH 11). La disolución se filtró a través de una unidad de filtración de 0,22 μM Millex GV13 (Millipore, Billerica, MA, USA). La formación de azúcares reductores solubles en el sobrenadante se determinó por el método de la hidrazida de ácido para-hidroxibenzoico (PAHBAH) (Lever, 1972) utilizando una curva patrón de celobiosa (celobiosa 50 a 1600 μM). Se añadió una disolución de PAHBAH 0,1 M (Sigma-Aldrich, St. Louis, MO, USA) en NaOH 0,5 M (100 μl) recién elaborada a 150 μl de la muestra filtrada y se hirvió durante 10 minutos después de lo cual la disolución se enfrió en hielo. Se midió la absorbancia de las muestras a 405 nm.

- 20 Las versiones núcleo de las celobiohidrolasas que albergan un CBD en su forma nativa se obtuvieron como sigue: Ct Cel7A y Tr Cel7A se expusieron a digestión proteolítica para eliminar el dominio de unión a celulosa. La digestión con papaína (Papaya Latex, 14 U/mg, Sigma) de las celobiohidrolasas nativas se realizó a 37°C durante 24 h en una mezcla de reacción compuesta de L-cisteína 10 mM y EDTA 2 mM en tampón de acetato de sodio 50 mM (pH 5,0) con la adición de papaína (se sometieron a ensayo dos concentraciones de papaína: una quinta o una décima parte la cantidad de papaína de la cantidad total de la Cel7A en la mezcla de reacción). La proteína núcleo resultante se purificó con columna de intercambio aniónico DEAE Sepharose FF (Pharmacia, Uppsala, Suecia) como se describió anteriormente. El producto se analizó en SDS-PAGE.

- 35 Los resultados de la hidrólisis a 45°C y 70°C se muestran en la Figura 4 y la Figura 5, respectivamente. Los resultados muestran claramente que todas las celobiohidrolasas muestran una hidrólisis más rápida y más completa a ambas temperaturas en comparación con la celobiohidrolasa Cel7A de *T. reesei* del estado de la técnica. A 70°C, las celobiohidrolasas termoestables de *Thermoascus aurantiacus* ALKO4242 y *Chaetomium thermophilum* ALKO4265 son superiores en comparación con Cel7A de *T. reesei*, también en el caso en el que el núcleo Cel7A de *Thermoascus* está conectado al CBD de Cel7A de *T. reesei* (Ta Cel7A + Tr CBD). Fue sorprendente que las

celobiohidrolasas aisladas y clonadas en este trabajo sean superiores, cuando albergan un CBD, en velocidad y formación de producto en la hidrólisis de celulosa cristalina también a la temperatura de hidrólisis convencional de 45°C en comparación con la celobiohidrolasa Cel7A (CBHI) de *T. reesei* del estado de la técnica a la misma concentración de enzima. Los resultados también están de acuerdo con las preparaciones enzimáticas (At Cel7A y Ct Cel7A), que fueron purificadas a partir de los anfitriones originales y sometidas a ensayo en la hidrólisis de Avicel (50°C, 24 h) (Ejemplo 2, Tabla 1).

Ejemplo 18. Clonación de genes de endoglucanasa de *Acremonium thermophilum* ALKO4245, *Chaetomium thermophilum* ALKO4261, y *Thermoascus aurantiacus* ALKO4242

Se utilizaron métodos de biología molecular convencionales como se describe en el Ejemplo 13. La construcción de las bibliotecas genómicas de *Acremonium*, *Chaetomium*, y *Thermoascus* se ha descrito en el Ejemplo 12.

Los péptidos derivados de endoglucanasas de *Acremonium* y *Chaetomium* purificadas comparten homología con varias endoglucanasas de la familia 45 de glicosil hidrolasa tales como la endoglucanasa de *Melanocarpus albomyces* Cel45A (AJ515703) y la endoglucanasa de *Humicola insolens* (A35275), respectivamente. Los péptidos derivados de la endoglucanasa de *Thermoascus* comparten casi 100% de identidad con la secuencia de endoglucanasa de *Thermoascus aurantiacus* EG1 (AF487830) publicada. Para amplificar una sonda para el escrutinio de las bibliotecas genómicas de *Acremonium* y *Chaetomium*, se diseñaron cebadores degenerados basándose en las secuencias de péptidos. El orden de los péptidos en la secuencia de la proteína y la naturaleza efectora o anti-sentido correspondiente de los cebadores se dedujo de la comparación con las endoglucanasas homólogas publicadas. Las secuencias de los cebadores y los péptidos correspondientes se enumeran en la Tabla 14. Debido a casi 100% de identidad de los péptidos de *Thermoascus* con la secuencia publicada, el gen de endoglucanasa se amplificó mediante PCR directamente a partir del ADN genómico.

Tabla 14. Oligonucleótidos sintetizados y utilizados como cebadores de PCR para amplificar una sonda para el escrutinio de los genes *cel45A* (EG_40) de *Acremonium thermophilum* y *cel7B* (EG_54) de *Chaetomium thermophilum* de las bibliotecas genómicas correspondientes.

Proteína	Péptido	Localización del cebador ^(a)	Secuencia del cebador ^(b)
At EG_40	Péptido 5	1-6	TAYTGGGAYTGYTGYAARCC
	WFQNADN ^(c)		RTTRTCNGCRTTYTGRAACCA
Ct EG_54	Péptido 7	3-7	GCAAGCTTCGRCARAARTCRTRTT ^(d)
	Péptido 2	5-9	GGAATTCGAYCARACNGARCARTA ^(e)

(a) Aminoácidos del péptido utilizado para el diseño de la secuencia del cebador

(b) N = A, C, G o T; R = A o G; Y = C o T

(c) El péptido no deriva de la proteína de *Acremonium* EG_40 purificada, pero se origina a partir de la secuencia de Cel45A de *M. albomyces* (AJ515703) homóloga a EG_40.

(d) Se añadió un sitio de restricción *HindIII* al extremo 5' del oligonucleótido

(e) Se añadió un sitio de restricción *EcoRI* al extremo 5' del oligonucleótido

Se amplificó la sonda específica del gen de *Acremonium thermophilum* *cel45A* para explorar la biblioteca genómica con los cebadores directo (TAYTGGGAYTGYTGYAARCC) e inverso (RTTRTCNGCRTTYTGRAACCA) utilizando ADN genómico como molde. Las mezclas de reacción de PCR contenían Tris-HCl 50 mM, pH 9,0, (NH₄)₂SO₄ 15 mM, Triton X-100 al 0,1%, MgCl₂ 1,5 mM, dNTP 0,1 mM, 0,5 µg de cada cebador, 1 unidad de ADN polimerasa Dynazyme EXT (Finnzymes, Finlandia) y aproximadamente 0,5 µg de ADN genómico de *Acremonium*. Las condiciones para las reacciones de PCR fueron las siguientes: 5 min de desnaturalización inicial a 95°C, seguido de 30 ciclos de 1 min a 95°C, recocido de 1 min a 50-60°C, extensión de 2 min a 72°C y una extensión final a 72°C durante 10 min. Para la amplificación de la sonda específica del gen *cel7B* de *Chaetomium thermophilum* (que codificaba Ct EG_54), se utilizaron un cebador directo (GGAATTCGAYCARACNGARCARTA) y un cebador inverso (GCAAGCTTCGRCARAARTCRTRTT). Las mezclas de reacción de PCR contenían Tris-HCl 10 mM, pH 8,8, KCl 50 mM, Triton X-100 al 0,1%, MgCl₂ 1,5 mM, dNTP 0,2 mM, 250 pmol de cada cebador, 2 unidad de ADN polimerasa Dynazyme II (Finnzymes, Finlandia) y aproximadamente 2 µg de ADN genómico de *Chaetomium*. Las

condiciones para la reacción de PCR fueron las descritas anteriormente, excepto que el recocido se realizó a 45-50°C.

Se obtuvieron dos productos de PCR a partir de la reacción de PCR de *Acremonium*. Se aislaron fragmentos de ADN de aproximadamente 0,6 kb y 0,8 kb a partir de gel de agarosa y se clonaron en el vector pCR4-TOPO® TA (Invitrogen, USA), dando como resultado los plásmidos pALK1710 y pALK1711, respectivamente. Los productos de ADN se caracterizaron mediante secuenciación y realizando hibridaciones de transferencia Southern con el ADN genómico de *Acremonium* digerido con varias enzimas de restricción. Los patrones de hibridación obtenidos con los dos fragmentos en condiciones de lavado rigurosas sugieren que se podrían escrutar dos supuestos genes de endoglucanasa a partir de la biblioteca genómica de *Acremonium*. Las secuencias de aminoácidos deducidas de los dos productos de PCR tienen homología con varias secuencias de endoglucanasa de la familia 45 de glicosil hidrolasa publicadas (programa BLAST, National Center for Biotechnology Information; Altschul *et al.*, 1990).

Se obtuvo un producto de PCR del tamaño esperado (estimado a partir de la secuencia de endoglucanasa de *Humicola insolens*, A35275 homóloga) a partir de la reacción de PCR de *Chaetomium*. Este fragmento de ADN de aproximadamente 0,7 kb se clonó en el vector pCR4-TOPO® TA (Invitrogen, USA), dando como resultado el plásmido pALK2005 y se analizó como se ha descrito anteriormente. La secuencia de aminoácidos deducida del producto de la PCR tiene homología con varias secuencias de celulasa publicadas de la familia de glicosil hidrolasa 7 (programa BLAST, versión 2.2.9 en NCBI, National Center for Biotechnology Information; Altschul *et al.*, 1990).

El inserto de los plásmidos pALK1710, pALK1711 y pALK2005 se aisló por digestión con enzimas de restricción y se marcó con digoxigenina de acuerdo con las instrucciones del proveedor (Roche, Alemania). Se escrutaron aproximadamente $1-2 \times 10^5$ placas de la biblioteca genómica de *Acremonium* o *Chaetomium* amplificada. La temperatura de hibridación fue 68°C y los filtros se lavaron 2 x 5 min a RT utilizando 2 x SSC - SDS al 0,1% seguido de 2 x 15 min a 68°C utilizando 0,1 x SSC - SDS al 0,1%. Se obtuvieron varias placas positivas, de las cuales de cinco a seis placas de fuerte hibridación se purificaron a partir de cada escrutinio. Los ADN de los fagos se aislaron y se analizaron mediante hibridación de transferencia Southern. Los fragmentos de restricción que hibridaron con la sonda se subclonaron en el vector pBluescript II KS+ (Stratagene, USA) y las porciones relevantes se secuenciaron. En todos los casos, el fragmento de fago subclonado contenía el gen completo de interés. La Tabla 15 resume la información de las sondas utilizadas para el escrutinio de los genes de endoglucanasa, los clones de fagos de los que se aislaron los genes, los fragmentos de restricción seleccionados que contenían los genes completos con sus regiones promotora y terminadora, los nombres de los plásmidos que contenían el fragmento de fago subclonado, y los números de depósito en la colección de cultivos de Deutsche Sammlung von und Mikroorganismen Zellkulturen GmbH (DSMZ) para las cepas de *E. coli* portadoras de estos plásmidos.

Tabla 15. Sondas utilizadas para la clonación del gen de endoglucanasa, clon de fago y subclón seleccionado, nombre del plásmido y el número de depósito correspondiente de la cepa de *E. coli*.

Gen	Biblioteca genómica	Sonda utilizada en el escrutinio	Clon de fago	Fragmento subclonado	Plásmido	Núm. de depósito de <i>E. coli</i>
At <i>cel45A</i>	<i>A. thermophilum</i> ALK04245	pALK1710	P24	SmaI de 5,5 kb	pALK1908	DSM 17324
At <i>cel45B</i>	<i>A. thermophilum</i> ALK04245	pALK1711	P41	XhoI de 6,0 kb	pALK1904	DSM 17323
Ct <i>cel7B</i>	<i>C. thermophilum</i> ALK04261	pALK2005	P55	BamHI de 5,1 kb	pALK2010	DSM 17729

El gen *cel5A* de *Thermoascus aurantiacus* (que codificaba EG_28) (SEQ ID NO: 9) se amplificó directamente a partir del ADN genómico aislado por medio de reacción PCR. Los cebadores directo (ATTAACCGCGGACTGCGCATCATGAAGCTCGGCTCTCTCGTGCTC) e inverso (AACTGAGGCATAGAACTGACGTCATATT) que se utilizaron para la amplificación fueron diseñados basándose en el gen eg.1 de *T. aurantiacus* (AF487830) publicado. Las mezclas de reacción de PCR contenían 1 x tampón Phusion HF, dNTP 0,3 mM, 0,5 µM de cada cebador, 2 unidades de ADN polimerasa Phusion TM (Finnzymes, Finlandia) y aproximadamente 0,25 µg de ADN genómico de *Thermoascus*. Las condiciones para las reacciones de PCR fueron las siguientes: 5 min de desnaturalización inicial a 95°C, seguido de 25 ciclos de 30 s a 95°C, recocido de 30 s a 57-67°C, extensión de 2,5 min a 72°C y una extensión final a 72°C durante 5 min. El producto de 1,3 kb amplificado que contenía el gen exacto (de codón de INICIO a PARADA) se clonó como un fragmento SacII-PstI en el vector pBluescript II KS+. Se secuenciaron dos clones independientes y un clon se seleccionó y se denominó pALK1926. El número de depósito de la cepa de *E. coli* que contenía pALK1926 en la colección de cultivos Deutsche Sammlung von und Mikroorganismen Zellkulturen GmbH es DSM 17326.

La información relevante de los genes y las secuencias de proteínas deducidas (SEQ ID NO: 9-16) se resumen en la

Tabla 16 y la Tabla 17, respectivamente. Las secuencias de péptido de las endoglucanasas EG_40 (gen *AceI* 45A) de *Acremonium*, EG_54 (gen *Ct cel7B*) de *Chaetomium*, y EG_28 (gen *Ta Cel5A*) de *Thermoascus* purificadas se encontraron en las correspondientes secuencias de aminoácidos deducidas de los genes clonados confirmando que se clonaron los genes apropiados.

- 5 Tabla 16. Resumen de los genes de endoglucanasa aislados a partir de *Acremonium thermophilum*, *Chaetomium thermophilum*, y *Thermoascus aurantiacus*.

Gen de endoglucanasa	Longitud con intrones (pb) ^(a)	Región codificante (pb) ^(b)	Núm. de intrones	Longitudes de los intrones (pb)	SEQ ID NO:
At <i>cel45A</i>	1076	891	2	59, 123	11
At <i>cel45B</i>	1013	753	2	155, 102	13
Ct <i>cel7B</i>	1278	1275	-	-	15
Ta <i>cel5A</i>	1317	1005	5	55, 60, 59, 74, 61	9

(a) Está incluido el codón de PARADA.

(b) No está incluido el codón de PARADA.

Tabla 17. Resumen de las secuencias de endoglucanasa deducidas de *Acremonium thermophilum*, *Chaetomium thermophilum*, y *Thermoascus aurantiacus*. ss, secuencia señal.

Proteína endoglucanasa	Núm. de aa	Longitud de ss NN/HMM ^(a)	CBD ^(b)	PM pronosticado (Da, sin incluir la ss) ^(c)	pl pronosticado (sin incluir la ss)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
At EG_40	297	21/21	Si, K265 a L297	28625	4.79	2	12
De tipo_EG_40	251	20/20	No	23972	6.11	2	14
Ct EG_54	425	17/17	No	45358	5.44	1	16
Ta EG_28	335	30(e)	No	33712	4.30	1	10

(a) La predicción de la secuencia señal se realizó utilizando el programa SignalP V3.0 (Nielsen *et al.*, 1997; Bendtsen *et al.*, 2004); el valor NN se obtuvo utilizando redes neuronales y el valor HMM utilizando modelos ocultos de Markov.

(b) Presencia de un dominio de unión a celulosa en la proteína, se indican los aminoácidos del CBD C-terminal (numeración de acuerdo con el polipéptido completo)

(c) La secuencia señal pronosticada no está incluida. La predicción se realizó utilizando la herramienta Compute pI/MW del servidor ExPASy (Gasteiger *et al.*, 2003).

(d) Los supuestos sitios de N-glicosilación NXS/T se pronosticaron utilizando el programa NetNGlyc 1.0 (Gupta *et al.*, 2004).

(e) De acuerdo con Hong *et al.* 2003a

Las secuencias de proteínas deducidas de las endoglucanasas EG_40 (At Cel45A) y de tipo_EG_40 (At Cel45B) de *Acremonium*, *Chaetomium* EG_54 (Ct Cel7B), y EG_28 de *Thermoascus* (Ta Cel5A) comparten homología con celulasas de la familia 45 de glicosil hidrolasas (*Acremonium*), familia 7 (*Chaetomium*), y familia 5 (*Thermoascus*), identificando así los genes aislados como miembros de estas familias de genes. Las homologías más cercanas de

las endoglucanasas EG_40/Cel45A y de tipo_EG_40/Cel45B de *Acremonium* son las endoglucanasas de *Thielavia terrestris* (CQ827970, 77,3% de identidad) y *Myceliophthora thermophila* (AR094305, 66,9% de identidad), respectivamente (Tabla 18). Las dos endoglucanasas aisladas de la familia 45 de *Acremonium* comparten solo una identidad de 53,7% entre sí. De estas enzimas solo EG_40/Cel45A contiene un dominio de unión a celulosa (CBD).

- 5 La homología más cercana para la secuencia de proteína pronosticada de la endoglucanasa EG_54/Cel7B de *Chaetomium* se encuentra en la secuencia de celulasa Cel7A (AJ515704) de *Melanocarpus albomyces*. La identidad entre estas dos secuencias de proteínas es de 70,6%.

La secuencia de proteína de la endoglucanasa de *Thermoascus aurantiacus* aislada es completamente idéntica a la de EGI de *T. aurantiacus* (AF487830, Tabla 18) publicada. La homología más cercana se encuentra en una secuencia de β -glucanasa de *Talaromyces emersonii* (AX254752, identidad de 71,1%).

10

Tabla 18. Comparación de las endoglucanasas EG_40, de tipo_EG_40/Cel45B de *Acremonium thermophilum*, EG_54/Cel7B de *Chaetomium thermophilum* y EG_28/Cel5A de *Thermoascus aurantiacus* deducidas con sus contrapartes homólogas. El alineamiento se realizó utilizando el programa Needle del paquete de programas EMBOSS. * Indica una endoglucanasa codificada por un gen clonado en este trabajo.

Organismo, enzima, y número de acceso	Identidad (%)
<i>Acremonium thermophilum</i> EG_40	100,0
<i>Thielavia terrestris</i> EG45, CQ827970	77,3
<i>Melanocarpus albomyces</i> Cel45A, AJ515703	75,3
<i>Neurospora crassa</i> , hipotética XM_324477	68,9
<i>Humicola grisea</i> var <i>thermoidea</i> , EGL3, AB003107	67,5
<i>Humicola insolens</i> EG5, A23635	67,3
<i>Myceliophthora thermophila</i> fam 45, AR094305	57,9
* <i>Acremonium thermophilum</i> de tipo_EG_40	53,7
<i>Acremonium thermophilum</i> de tipo_EG_40	100,0
<i>Myceliophthora thermophila</i> fam 45, AR094305	66,9
<i>Magnaporthe grisea</i> 70-15 hipotética, XM_363402	61,9
<i>Thielavia terrestris</i> EG45, CQ827970	
* <i>Acremonium thermophilum</i> EG_40	56,8
<i>Melanocarpus albomyces</i> Cel45A, AJ515703	53,7
	52,8
<i>Chaetomium thermophilum</i> EG_54	100,0
<i>Melanocarpus albomyces</i> Cel7A, AJ515704	70,6
<i>Humicola grisea</i> var <i>thermoidea</i> EGI, D63516	68,8
<i>Humicola insolens</i> EGI, AR012244	67,7

Organismo, enzima, y número de acceso	Identidad (%)
<i>Myceliophthora thermophila</i> EGI, AR071934	61,7
<i>Fusarium oxysporum</i> var <i>lycopercisi</i> EGI, AF29210	53,5
<i>Fusarium oxysporum</i> EGI, AR012243	52,6
<i>Thermoascus aurantiacus</i> EG_28	100,0
<i>Thermoascus aurantiacus</i> EG, AX812161	100,0
<i>Thermoascus aurantiacus</i> EGI, AY055121	99,4
<i>Talaromyces emersonii</i> β -glucanasa, AX254752	71,1
<i>Talaromyces emersonii</i> EG, AF440003	70,4
<i>Aspergillus niger</i> EG, A69663	70,1
<i>Aspergillus niger</i> EG, A62441	69,9
<i>Aspergillus niger</i> EG, AF331518	69,6
<i>Aspergillus aculeatus</i> EGV, AF054512	68,5

Ejemplo 19. Producción de endoglucanasas recombinantes en *Trichoderma reesei*

5 Los plásmidos de expresión se construyeron para la producción de las proteínas EG_40/Cel45A, de tipo EG_40/Cel45B de *Acremonium*, y EG_28/Cel5A de *Thermoascus* recombinantes como se describe en el Ejemplo 14. Los casetes de expresión lineales (Tabla 19) se aislaron a partir de la cadena principal del vector mediante digestión con enzimas de restricción, se transformaron en *T. reesei* A96 y los transformantes se purificaron como se describe en el Ejemplo 14.

10 Tabla 19. Casetes de expresión contruidos para la producción de las endoglucanasas EG_40/Cel45A, de tipo EG_40/Cel45B de *Acremonium thermophilum*, y EG_28/Cel5A de *Thermoascus aurantiacus* en *Trichoderma reesei*. La estructura esquemática de los casetes de expresión se describe en la Figura 2.

Endoglucanasa	Plásmido de expresión	Tamaño de casete de expresión ^(a)	Terminador heterólogo ^(b)
At EG_40	pALK1920	<i>NotI</i> de 10,9 kb	156 pb (<i>HindIII</i>)
de tipo_At EG_40	pALK1921	<i>EcoRI</i> de 8,6 kb	282 pb (<i>SspI</i>)
Ta EG_28	pALK1930	<i>NotI</i> de 8,6 kb	ninguno

(a) El casete de expresión para la transformación en *T. reesei* se aisló a partir de la cadena principal del vector mediante digestión con *EcoRI* o *NotI*.

(b) Se indica el número de nucleótidos después del codón de PARADA del gen clonado que se incluyen en el casete de expresión. El sitio de restricción en la región 3' del gen que se utilizó en la construcción del casete de expresión se indica entre paréntesis.

La producción de endoglucanasa de los transformantes se analizó a partir de los sobrenadantes de cultivo de cultivos en matraz oscilante (50 ml). Los transformantes se cultivaron como en el Ejemplo 14 y la actividad enzimática de la proteína recombinante se midió a partir del sobrenadante de cultivo como la liberación de azúcares

reductores de la carboximetilcelulosa (2% (p/v) CMC) a 50°C en tampón de citrato 50 mM pH 4,8 esencialmente como describen Bailey y Nevalainen 1981; Haakana *et al.* 2004. La producción de las proteínas recombinantes también se detectó a partir de sobrenadantes de cultivo mediante electroforesis en gel de poliacrilamida-SDS. Se produjeron anticuerpos policlonales específicos de EG_40 de *Acremonium* en conejos (Universidad de Helsinki, Finlandia). La expresión de EG_40 se verificó mediante análisis de transferencia Western con anticuerpos anti_EG_40 utilizando el sistema ProtoBlot Western blot AP (Promega). Los genotipos de los transformantes seleccionados se analizaron mediante transferencia Southern utilizando el casete de expresión como una sonda.

El pH óptimo de las endoglucanasas producidas de forma heteróloga se determinó en el tampón de Mcllvaine universal dentro de un intervalo de pH de 4,0-8,0 utilizando carboximetilcelulosa como sustrato. Como se muestra en la Figura 6A el intervalo de pH más amplio (4,5-6,0) es el de la proteína EG_40/Cel45A de *Acremonium*, siendo el óptimo a pH 5,5. Los óptimos de pH para las otras endoglucanasas producidas de forma heteróloga son pH 5,0-5,5 y 6,0 para la de tipo_EG_40/Cel45B de *Acremonium* y EG_28/Cel5A de *Thermoascus*, respectivamente. La temperatura óptima para la actividad enzimática de estas endoglucanasas se determinó en el intervalo de temperatura de 50-85°C como se describió anteriormente. Se determinó que la mayor actividad de las enzimas era a 75°C, 60°C y 75°C para EG_40/Cel45A, de tipo_EG_40/Cel45B de *Acremonium*, y EG_28/Cel5A de *Thermoascus*, respectivamente (Figura 6B).

Los transformantes seleccionados se cultivaron, como se describe en el Ejemplo 14, en un biorreactor de 2 litros durante cuatro días (28°C, pH 4,2) para obtener material para los ensayos de aplicación.

Ejemplo 20. Clonación de genes de beta-glucosidasas de *Acremonium thermophilum* ALKO4245, *Chaetomium thermophilum* ALKO4261 y *Thermoascus aurantiacus* ALKO4242

Se utilizaron métodos convencionales de biología molecular como se describe en el Ejemplo 13. La construcción de las bibliotecas genómicas de *Acremonium*, *Chaetomium*, y *Thermoascus* se ha descrito en el Ejemplo 12.

Los péptidos derivados de las β -glucosidasas de *Acremonium*, *Chaetomium*, y *Thermoascus* purificadas compartían homología con varias β -glucosidasas de la familia 3 de glicosil hidrolasas tales como las β -glucosidasas de *Acremonium cellulolyticus* (BD168028), *Trichoderma viride* (AY368687), y *Talaromyces emersonii* (AY072918), respectivamente. Para amplificar una sonda para el escrutinio de las bibliotecas genómicas de *Acremonium*, *Chaetomium*, o *Thermoascus*, se diseñaron cebadores degenerados basándose en las secuencias de péptidos. El orden de los péptidos en la secuencia de la proteína y la naturaleza efectora o anti-sentido correspondiente de los cebadores se dedujo a partir de la comparación con los β -glucosidasas homólogas publicadas. Las secuencias de los cebadores y los péptidos correspondientes se enumeran en la Tabla 20.

Tabla 20. Oligonucleótidos sintetizados y utilizados como cebadores de PCR para amplificar una sonda para el escrutinio de los genes *ce3A* (β G_101) de *Acremonium thermophilum*, *ce3A* (β G_76) de *Chaetomium thermophilum*, y *ce3A* (β G_81) de *Thermoascus aurantiacus* a partir de las bibliotecas genómicas correspondientes.

Proteína	Péptido	Localización del cebador ^(a)	Secuencia del cebador ^(b)
At β G_101	EKVNL ^(c)		GARAARGTNAAYCTNAC
	Péptido 4	6-11	YTTRCCRTRRTTSGGRGTRTA
Ct β G_76	Péptido 6	4-9	TNTGYCTNCARGAYGG
	Péptido 1	3-8	TCRAARTGSCGRTARTCRATRAASAG
Ta β G_81	Péptido 3	1-5	AARGGYGTSGAYGTSCAR
	Péptido 1	2-7	YTTRCCCCASGTRAASGG

(a) Aminoácidos del péptido utilizado para el diseño de la secuencia del cebador

(b) Para reducir la degeneración, algunos codones se seleccionaron de acuerdo con la preferencia fúngica N = A, C, G, o T; R = A o G; S = C o G; Y = C o T

(c) El péptido no deriva de la proteína β G_101 de *Acremonium* purificada, pero se origina a partir de la secuencia de β -glucosidasa de *A. cellulolyticus* (BD168028) homóloga a β G_101.

Las sondas para el escrutinio de las bibliotecas genómicas construidas fueron amplificadas con las combinaciones de cebadores enumeradas (Tabla 20) utilizando ADN genómico de *Acremonium*, *Chaetomium*, o *Thermoascus* como molde. Las mezclas de reacción de PCR contenían Tris-HCl 50 mM, pH 9,0, (NH₄)₂SO₄ 15 mM, Triton X-100 al 0,1%, MgCl₂ 1,5 mM, dNTP 0,1-0,2 mM, 0,25 µg de cada cebador, 1 unidad de ADN polimerasa Dynazyme EXT (Finnzymes, Finlandia) y aproximadamente 0,5 µg de ADN genómico. Las condiciones para las reacciones de PCR fueron las siguientes: 5 min de desnaturalización inicial a 95°C, seguido de 30 ciclos de 1 min a 95°C, recocido de 1 min a 40°C (ADN de *Acremonium* como molde), a 50°C (ADN de *Chaetomium* como molde), o a 63°C (ADN de *Thermoascus* como molde), extensión de 2-3 min a 72°C y una extensión final a 72°C durante 5-10 min.

Los productos de PCR específicos de tamaño esperado (estimado a partir de las secuencias de β-glucosidasa homólogas BD168028, AY072918, y AY368687) se aislaron del gel de agarosa. Los fragmentos de ADN de aproximadamente 1,8 kb (*Acremonium*), 1,5 kb (*Chaetomium*), y 1,52 kb (*Thermoascus*) se clonaron en el vector pCR4-TOPO® TA (Invitrogen, USA), dando como resultado los plásmidos pALK1924, pALK1935 y pALK1713, respectivamente. Los productos de ADN se caracterizaron por secuenciación y realizando hibridaciones de transferencia Southern con el ADN genómico digerido con varias enzimas de restricción. Los patrones de hibridación en condiciones de lavado rigurosas sugieren que se podría aislar un supuesto gen de β-glucosidasa de la biblioteca genómica de *Acremonium*, *Chaetomium*, y *Thermoascus*. Las secuencias de aminoácidos deducidas de los tres productos de PCR tienen homología con varias secuencias de β-glucosidasa publicadas de la familia 3 de glicosil hidrolasa (programa BLAST, National Center for Biotechnology Information; Altschul *et al.*, 1990).

El inserto de los plásmidos pALK1713, pALK1924 y pALK1935 se aisló mediante digestión con enzimas de restricción y se marcó con digoxigenina de acuerdo con las instrucciones del proveedor (Roche, Alemania). Se escrutaron aproximadamente 1-2 x 10⁵ placas de la biblioteca genómica de *Acremonium*, *Chaetomium*, o *Thermoascus* amplificada como se describe en el Ejemplo 18. Se obtuvieron varias placas positivas, de las cuales se purificaron de cinco a seis placas de fuerte hibridación a partir de cada escrutinio. Los ADN de los fagos se aislaron y se analizaron mediante hibridación de transferencia Southern. Los fragmentos de restricción que hibridaban con la sonda se subclonaron en el vector pBluescript II KS+ (Stratagene, USA) y las porciones relevantes se secuenciaron. En todos los casos, el fragmento de fago subclonado contiene el gen completo de interés. La Tabla 21 resume la información de las sondas utilizadas para el escrutinio de los genes de β-glucosidasa, los clones de fagos a partir de los que se aislaron los genes, los fragmentos de restricción seleccionados que contienen los genes completos con sus regiones promotora y terminadora, los nombres de los plásmidos que contienen el fragmento de fago subclonado, y los números de depósito en la Deutsche Sammlung von und Mikroorganismen Zellkulturen GmbH colección de cultivos (DSM) para las cepas de *E. coli* que portan estos plásmidos.

Tabla 21. Sondas utilizadas para la clonación del gen de la β-glucosidasa, clon de fago y subclón elegido, nombre del plásmido y número de depósito correspondiente de la cepa de *E. coli*.

Gen	Biblioteca genómica	Sonda utilizada en el escrutinio	Clon de fago	Fragmento subclonado	Plásmido	Núm. de depósito de <i>E. coli</i>
At <i>cel3A</i>	<i>A. thermophilum</i> ALKO4245	pALK1924	P44	<i>Hind</i> III de 6,0 kb	pALK1925	DSM 17325
Ct <i>cel3A</i>	<i>C. thermophilum</i> ALKO4261	pALK1935	P51	<i>Xba</i> I de 7,0 kb	pALK2001	DSM 17667
Ta <i>cel3A</i>	<i>T. aurantiacus</i> ALKO4242	pALK1713	P21	<i>Bam</i> HI de 5,3 kb	pALK1723	DSM 16725

La información relevante de los genes y las secuencias de proteína deducidas (SEQ ID NO: 21-26) se resumen en la Tabla 22 y la Tabla 23, respectivamente. Las secuencias peptídicas de las proteínas βG₁₀₁ (At *Cel3A*) de *Acremonium*, βG₇₆ (Ct *Cel3A*) de *Chaetomium*, y βG₈₁ (Ta *Cel3A*) de *Thermoascus* purificadas se encontraron en las secuencias de aminoácidos deducidas correspondientes de los genes clonados confirmando que se clonaron los genes correspondientes.

Tabla 22. Resumen de los genes β-glucosidasa aislados a partir de *Acremonium thermophilum*, *Chaetomium thermophilum*, y *Thermoascus aurantiacus*.

Gen de β-glucosidasa	Longitud con intrones (pb) ^(a)	Región codificante (pb) ^(b)	Núm. de intrones	Longitud de los intrones (pb)	SEQ ID NO:
At <i>cel3A</i>	2821	2583	3	92, 74, 69	23

Gen de β -glucosidasa	Longitud con intrones (pb) ^(a)	Región codificante pb) ^(b)	Núm. de intrones	Longitud de los intrones (pb)	SEQ ID NO:
Ct <i>ce3A</i>	2257	2202	1	52	25
Ta <i>ce3A</i>	3084	2529	7	134, 67, 56, 64, 59, 110, 62	21

(a) Está incluido el codón de PARADA.

(b) No está incluido el codón de PARADA.

Tabla 23. Resumen de las secuencias β -glucosidasa deducidas de *Acremonium thermophilum*, *Chaetomium thermophilum*, y *Thermoascus aurantiacus*. ss, secuencia señal.

Proteína β -glucosidasa	Núm. de aa	Longitud de ss NN/HMM ^(a)	CBD ^(b)	PM pronosticado (Da, sin incluir ss) ^(c)	pl pronosticado sin incluir ss)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
At β G_101	861	19/18	No	91434	5,46	8	24
Ct β G_76	734	20/20	No	76457	6,3	2	26
Ta β G_81	843	19/19	No	89924	4,95	8	22

(a) La predicción de la secuencia señal se realizó utilizando el programa SignalP V3.0 (Nielsen *et al.*, 1997; Bendtsen *et al.*, 2004); el valor NN se obtuvo utilizando redes neuronales y el valor HMM utilizando modelos ocultos de Markov.

(b) Presencia de un dominio de unión a celulosa en la proteína

(c) La secuencia señal pronosticada no está incluida. La predicción se realizó utilizando la herramienta Compute pI/MW del servidor ExPASy (Gasteiger *et al.*, 2003).

(d) Los supuestos sitios de N-glicosilación NXS/T se pronosticaron utilizando el programa NetNGlyc 1.0 (Gupta *et al.*, 2004).

- 5 Las secuencias de proteínas deducidas de las β -glucosidasas β G_101/Cel3A de *Acremonium*, β G_76/Cel3A de *Chaetomium*, y β G_B1/Cel3A de *Thermoascus* comparten homología con las enzimas de la familia 3 de glicosil hidrolasas, identificando así que los genes aislados pertenecen a esta familia de genes. Las contrapartes más cercanas de las β -glucosidasas de *Acremonium*, *Chaetomium*, y *Thermoascus* son las de *Magnaporthe grisea* (β -glucosidasa, AY849670), *Neurospora crassa* (hipotética, XM_324308), y *Talaromyces emersonii* (β -glucosidasa, AY072918), respectivamente (Tabla 24). Se encontró que la identidad de secuencia más alta (73,2%) fue la de β G_76/Cel3A de *C. thermophilum* con la proteína hipotética de *N. crassa* lo que indica que se clonaron nuevos genes de enzimas.
- 10

Tabla 24. Comparación de las β -glucosidasas β G_101/Cel3A de *Acremonium thermophilum*, β G_76/Cel3A de *Chaetomium thermophilum* y β G_81/Cel3A de *Thermoascus aurantiacus* deducidas con sus contrapartes homólogas. El alineamiento se realizó utilizando el programa Needle del paquete de programas EMBOSS. * Indica una β -glucosidasa codificada por un gen clonado en este trabajo.

Organismo, enzima, y número de acceso	Identidad (%)
* <i>Acremonium thermophilum</i> βG_101	100,0
<i>Magnaporthe grisea</i> β -glucosidasa, AY849670	73,1
<i>Neurospora crassa</i> hipotética, XM_330871	71,1
<i>Trichoderma reesei</i> Cel3B, AY281374	65,2
* <i>Thermoascus aurantiacus</i> β G_81	62,2
<i>Aspergillus aculeatus</i> β -glucosidasa, D64088	59,5
<i>Talaromyces emersonii</i> β -glucosidasa, AY072918	58,9
<i>Aspergillus oryzae</i> , AX616738	58,2
<i>Acremonium cellulolyticus</i> β -glucosidasa, BD168028	57,2
* <i>Chaetomium thermophilum</i> β G_76	40,9
<i>Chaetomium thermophilum</i> βG_76	100,0
<i>Neurospora crassa</i> , hipotética XM_324308	76,9
<i>Magnaporthe grisea</i> , hipotética XM_364573	70,2
<i>Trichoderma viridae</i> BGI, AY368687	65,8
<i>Acremonium cellulolyticus</i> β -glucosidasa, BD168028	41,2
* <i>Acremonium thermophilum</i> β G_101	40,9
<i>Trichoderma reesei</i> Cel3B, AY281374	40,0
* <i>Thermoascus aurantiacus</i> β G_81	39,9
* <i>Thermoascus aurantiacus</i> βG_81	100,0
<i>Talaromyces emersonii</i> β -glucosidasa, AY072918	73,2
<i>Aspergillus oryzae</i> , AX616738	69,5
<i>Aspergillus aculeatus</i> β -glucosidasa, D64088	68,0
<i>Acremonium cellulolyticus</i> β -glucosidasa, BD168028	65,7
* <i>Acremonium thermophilum</i> β G_101	62,2

Organismo, enzima, y número de acceso	Identidad (%)
<i>Trichoderma reesei</i> Cel3B, AY281374	57,9
* <i>Chaetomium thermophilum</i> β G_76	39,9

Ejemplo 21. Producción de beta-glucosidasas recombinantes en *Trichoderma reesei*

Se construyeron plásmidos de expresión para la producción recombinante de las proteínas β G_101/Cel3A de *Acremonium*, β G_76/Cel3A de *Chaetomium*, y β G_81/Cel3A de *Thermoascus* como se describe en el Ejemplo 14. Los casetes de expresión lineales (Tabla 25) se aislaron a partir de la cadena principal de vector mediante digestión con enzimas de restricción, se transformaron en A96 o A33 de *T. reesei* (ambas cepas tienen suprimidos los genes que codifican las cuatro principales celulasas CBHI/Cel7A, CBHI/Cel6A, EGI/Cel7B y EGI/Cel5A) y se purificaron los transformantes purificados como se describe en el Ejemplo 14.

Tabla 25. Casetes de expresión contruidos para la producción de β -glucosidasas β G_101/Cel3A de *Acremonium thermophilum*, β G_76/Cel3A de *Chaetomium thermophilum* y β G_81/Cel3A de *Thermoascus aurantiacus* en *Trichoderma reesei*. La estructura esquemática de los casetes de expresión se describe en la Figura 2.

β -glucosidasa	Plásmido de expresión	Tamaño del casete de expresión ^(a)	Terminador heterólogo ^(b)
At β G_101	pALK1933	<i>NotI</i> de 10,5 kb	300 pb (<i>HindIII</i>)
Ct β G_76	pALK2004	<i>EcoRI</i> de 10,1 kb	528 pb (<i>XbaI</i>)
Ta β G_81	pALK1914	<i>EcoRI</i> de 10,9 kb	452 pb (<i>ApoI</i>)

(a) El casete de expresión para la transformación en *T. reesei* se aisló a partir de la cadena principal del vector mediante digestión con *EcoRI* o *NotI*.

(b) Se indica el número de nucleótidos después del codón de PARADA del gen clonado que se incluye en el casete de expresión. El sitio de restricción en la región 3' del gen que se utilizó en la construcción del casete de expresión se indica entre paréntesis.

La producción de beta-glucosidasa de los transformantes se analizó a partir de los sobrenadantes de cultivo de los cultivos en matraz oscilante (50 ml). Los transformantes se cultivaron como en el Ejemplo 14 y se midió la actividad enzimática de la proteína recombinante a partir del sobrenadante de cultivo utilizando sustrato de 4-nitrofenil- β -D-glucopiranosido como describen Bailey y Nevalainen 1981. La producción de las proteínas recombinantes también se detectó a partir de los sobrenadantes de cultivo mediante electroforesis en gel de poliacrilamida-SDS. Además, la expresión de β G_81 de *Thermoascus* se verificó mediante análisis de transferencia Western con anticuerpos anti- β G_81 como se describe en el Ejemplo 19. Los genotipos de los transformantes seleccionados se analizaron mediante transferencia Southern utilizando el casete de expresión como sonda.

El óptimo de pH de las β -glucosidasas producidas de manera heteróloga se determinó en el tampón de Mcllvaine universal dentro de un intervalo de pH de 3,0-8,0 utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato. Los óptimos de pH para β G_101 de *Acremonium*, β G_76 de *Chaetomium* y β G_81 de *Thermoascus* son un pH de 4,5, 5,5 y 4,5, respectivamente (Figura 7A). La temperatura óptima para la actividad enzimática de estas β -glucosidasas se determinó en el intervalo de temperatura de 50-85°C como se describió anteriormente. Se determinó que la mayor actividad de las enzimas era a 70°C, 65°C y 75°C para β G_101/Cel3A de *Acremonium*, β G_76/Cel3A de *Chaetomium* y β G_81/Cel3A de *Thermoascus*, respectivamente (Figura 7B).

Los transformantes seleccionados se cultivaron, tal como se describe en el Ejemplo 14, en un biorreactor de 2 litros durante cuatro días (28°C, pH 4,2) para obtener material para los ensayos de aplicación.

Ejemplo 22. Clonación de los genes de xilanasa de *Acremonium thermophilum* ALK04245 y *Thermoascus aurantiacus* ALK04242

Se utilizaron métodos de biología molecular convencionales como se describe en el Ejemplo 13. La construcción de la biblioteca genómica de *Acremonium* se ha descrito en el Ejemplo 12.

Los péptidos derivados de la xilanasas de *Acremonium* purificada comparten homología con las xilanasas de la familia 10 de glicosil hidrolasas tales como XYNI (AB001030) de *Humicola grisea*. Todos los péptidos derivados de la xilanasas de *Thermoascus* fueron completamente idénticos a la secuencia de XynA (AJ132635) de *Thermoascus aurantiacus* publicada identificando así la proteína purificada como la misma enzima. Debido a esto el gen de xilanasas de *Thermoascus* se amplificó mediante PCR a partir del ADN genómico.

Para amplificar una sonda para el escrutinio del gen de xilanasas de *Acremonium* a partir de la biblioteca genómica, se diseñaron cebadores degenerados sobre la base de las secuencias de péptidos (Ejemplo 11, Tabla 5). El orden de los péptidos en la secuencia de la proteína y la naturaleza efectora o antisentido correspondiente de los cebadores se dedujo de la comparación con la secuencia XYNI (AB001030) de *Humicola insolens* homóloga. La secuencia del cebador efector (GAYGGYGAYGCSACYTAYATG) se basa en el péptido 3 (aminoácidos 2-8) y el cebador anti-sentido (YTTYTGRTCTAYTCSAGRTTSTA) en el péptido 1 (aminoácidos 4-11).

Se obtuvo de la reacción un producto de PCR del tamaño esperado (estimado a partir de la secuencia XYNI AB001030 de *Humicola insolens* homóloga). Este fragmento de ADN de aproximadamente 0,7 kb se clonó en el vector pCR4-TOPO® TA (Invitrogen, USA), dando como resultado el plásmido pALK1714, y se caracterizó mediante secuenciación. La secuencia de aminoácidos deducida del producto de la PCR tiene homología con varias secuencias de xilanasas publicadas de la familia 10 de glicosil hidrolasas (programa BLAST, National Center for Biotechnology Information; Altschul *et al.*, 1990).

El inserto del plásmido pALK1714 se aisló mediante digestión con enzimas de restricción y se marcó con digoxigenina de acuerdo con las instrucciones del proveedor (Roche, Alemania). Se seleccionaron aproximadamente $1-2 \times 10^5$ placas de la biblioteca genómica de *Acremonium* amplificada como se describe en el Ejemplo 18. Se obtuvieron varias placas positivas, de las cuales se purificaron cinco placas de fuerte hibridación. Los ADN de los fagos se aislaron y se analizaron mediante hibridación de transferencia Southern. Un fragmento de restricción de *XbaI* de 3,0 kb que hibridaba con la sonda se subclonó en el vector pBluescript II KS+ (Stratagene, USA), dando como resultado el plásmido pALK1725. Las porciones relevantes de pALK1725 se secuenciaron y se encontró que contenían el gen *xyn* 10A completo de *Acremonium thermophilum* (SEQ ID NO: 19). El número de depósito de la cepa de *E. coli* que contiene pALK1725 en la colección de cultivos Deutsche Sammlung von und Mikroorganismen Zellkulturen GmbH es DSM 16726.

El gen *xyn* 10A de *Thermoascus aurantiacus* (SEQ ID NO: 17) se amplificó directamente a partir del ADN genómico aislado por reacción mediante PCR. Los cebadores directo (TTATACCGCGGGAAGCCATGGTTCGACCAACGATCCTAC) e inverso (TTATAGGATCCACCGGTCTATACTCACTGCTGCAGGTCCTG) que se utilizaron en la amplificación del gen se diseñaron sobre la base del gen *XynA* (AJ132635) de *T. aurantiacus* publicado. Las mezclas de reacción de PCR contenían Tris-HCl 50 mM, pH 9,0, $(\text{NH}_4)_2\text{SO}_4$ 15 mM, Triton X-100 al 0,1%, MgCl_2 1,5 mM, dNTP 0,3 mM, 1 μM de cada cebador, 1 unidad de ADN polimerasa Dynazyme EXT (Finnzymes, Finlandia) y aproximadamente 0,5 μg de ADN genómico de *Thermoascus*. Las condiciones para las reacciones de PCR fueron las siguientes: 5 min de desnaturalización inicial a 95°C, seguido de 30 ciclos de 1 min a 95°C, recocido de 1 min a 60-66°C, extensión de 3 min a 72°C y una extensión final a 72°C durante 10 min. El producto de 1,9 kb amplificado que contenía el gen exacto (del codón de INICIO al de PARADA) se clonó como un fragmento *SacII-BamHI* en el vector pBluescript II KS+. Se secuenciaron tres clones independientes y se seleccionó un clon y se denominó pALK1715. El número de depósito de la cepa de *E. coli* que contiene pALK1715 en la colección de cultivos Deutsche Sammlung von und Mikroorganismen Zellkulturen GmbH es DSM 16724.

La información relevante de los genes y las secuencias de proteína deducidas (SEQ ID NO: 17-20) se resumen en la Tabla 26 y la Tabla 27, respectivamente. Las secuencias peptídicas de las proteínas XYN_60 de *Acremonium* y XYN_30 de *Thermoascus* purificadas se encontraron en las secuencias de aminoácidos deducidas correspondientes de los genes clonados (At *xyn*10A y Ta *xyn*10A, respectivamente) confirmando que se clonaron los genes correspondientes.

Tabla 26. Resumen de los genes de xilanasas aislados de *Acremonium thermophilum* y *Thermoascus aurantiacus*.

Gen de xilanasas	Longitud con intrones (pb) ^(a)	Región codificante (pb) ^(b)	Núm. de intrones	Longitud de los intrones (pb)	SEQ ID NO:
At <i>xyn</i> 10A	1471	1248	2	135, 85	19
Ta <i>xyn</i> 10A	1913	987	10	73, 74, 68, 103, 69, 65, 93, 66, 100, 212	17

(a) Está incluido el codón de PARADA.

(b) No está incluido el codón de PARADA.

Tabla 27. Resumen de las secuencias deducidas de xilanasas de *Acremonium thermophilum* y *Thermoascus aurantiacus*. ss, secuencia señal.

Proteína xilanasas	Núm. de aa	Longitud de ss NN/HMM ^(a)	CBD ^(b)	PM pronosticado (Da, sin incluir la) ^(c)	pl pronosticado (sin incluir la ss)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
At XYN_60	416	19/19	Si, W385 a L416	42533	6,32	1-2	20
Ta XYN_30	329	26(e)	No	32901	5,81	0	18

(a) La predicción de la secuencia señal se realizó utilizando el programa SignalP V3.0 (Nielsen *et al.*, 1997; Bendtsen *et al.*, 2004); el valor NN se obtuvo utilizando redes neuronales y valor HMM utilizando modelos ocultos de Markov.

(b) Presencia de un dominio de unión a carbohidrato CBD, se indican los aminoácidos del CBD C-terminal (numeración de acuerdo con el polipéptido completo)

(c) No está incluida la secuencia señal pronosticada. La predicción se realizó utilizando la herramienta Compute pI/MW del servidor ExPASy (Gasteiger *et al.*, 2003).

(d) Los supuestos sitios de N-glicosilación NXS/T se pronosticaron utilizando el programa NetNGlyc 1,0 (Gupta *et al.*, 2004).

(e) Según Lo Leggio *et al.*, 1999.

- 5 Las secuencias de proteínas deducidas de las xilanasas de *Acremonium* y *Thermoascus* comparten homología con varias enzimas de la familia 10 de glicosil hidrolasas, identificando los genes correspondientes como miembros de la familia 10 de xilanasas. La contraparte más cercana para XYN_60/Xyn10A de *Acremonium* encontrado es XYLI (AB001030) de *Humicola grisea* que muestra 67,1% de identidad con XYN_60 (Tabla 28). La secuencia de proteína pronosticada xilanasas XYN_30/Xyn10A de *Thermoascus aurantiacus* aislada es completamente idéntica a la de XynA de *T. aurantiacus* (P23360, Tabla 28) publicada. La homología más cercana se encontró en una secuencia de xilanasas de *Aspergillus niger* (A62445, identidad de 69,7%).
- 10

Tabla 28. Comparación de las xilanasas XYN_60/Xyn10A de *Acremonium thermophilum* y XYN_30/Xyn10A de *Thermoascus aurantiacus* deducidas con sus contrapartes homólogas. El alineamiento se realizó utilizando el programa Needle del paquete de programas EMBOSS. * Indica una xilanasas codificada por un gen clonado en este trabajo.

Organismo, enzima, y número de acceso	identidad (%)
* <i>Thermoascus aurantiacus</i> XYN_30	100,0
<i>Thermoascus aurantiacus</i> XynA, P23360	100,0
<i>Thermoascus aurantiacus</i> XynA, AF127529	99,4
<i>Aspergillus niger</i> xilanasas, A62445	69,7
<i>Aspergillus aculeatus</i> xilanasas, AR137844	69,9
<i>Aspergillus terreus</i> fam 10 xyn, DQ087436	65,0
<i>Aspergillus sojae</i> , XynXI AB040414	63,8
<i>Penicillium chrysogenum</i> xilanasas, AY583585	62,5
* <i>Acremonium thermophilum</i> XYN_60	100,0

Organismo, enzima, y número de acceso	identidad (%)
<i>Humicola grisea</i> XYL I, AB001030	67,1
<i>Magnaporthe grisea</i> 70-15, hipotética XM_364947	63,8
<i>Aspergillus aculeatus</i> xilanasa, AR149839	53,7
<i>Talaromyces emersonii</i> xilanasa, AX403831	51,8
<i>Gibberella zeae</i> xilanasa, AY575962	51,4
<i>Magnaporthe grisea</i> XYL5, AY144348	48,5
<i>Talaromyces emersonii</i> , AX172287	46,9

Ejemplo 23. Producción de xilanasas recombinantes en *Trichoderma reesei*

Se construyeron plásmidos de expresión para la producción de las proteínas XYN_60/Xyn10A de *Acremonium* y XYN_30/Xyn10A de *Thermoascus* recombinantes como se describe en el Ejemplo 14. Los casetes de expresión lineales (Tabla 29) se aislaron a partir de la cadena principal del vector mediante digestión con enzimas de restricción, se transformaron en *T. reesei* A96, y los transformantes se purificaron como se describe en el Ejemplo 14.

Tabla 29. Casetes de expresión contruidos para la producción de las xilanasas XYN_60/Xyn10A de *Acremonium thermophilum* y XYN_30/Xyn10A de *Thermoascus aurantiacus* en *Trichoderma reesei*. La estructura esquemática de los casetes de expresión se describe en la Figura 2.

Xilanasa	Plásmido de expresión	Tamaño del casete de expresión ^(a)	Terminador heterólogo ^(b)
At XYN_60	pALK1912	9,0 kb	150 pb (<i>Bam</i> HI)
Ta XYN_30	pALK1913	9,3 kb	ninguno

(a) El casete de expresión para la transformación en *T. reesei* se aisló a partir de la cadena principal del vector mediante digestión con *Eco*RI.

(b) Se indica el número de nucleótidos después del codón de parada del gen clonado que se incluyen en el casete de expresión. El sitio de restricción en la región 3' del gen que se utilizó en la construcción del casete de expresión se indica entre paréntesis.

La producción de xilanasa de los transformantes se analizó a partir de los sobrenadantes de cultivo de cultivos en matraz oscilante (50 ml). Los transformantes se cultivaron como en el Ejemplo 14 y la actividad enzimática de la proteína recombinante se midió a partir del sobrenadante de cultivo como la liberación de azúcares reductores a partir de xilano de abedul (1% p/v) a 50°C en tampón de citrato 50 mM de pH 5,3 tal como describen Bailey y Poutanen 1989. También se analizó la producción de la proteína recombinante del sobrenadante de cultivo mediante electroforesis en gel de poliacrilamida-SDS. Además, la expresión de ambas xilanasas se determinó mediante análisis de transferencia Western con anticuerpos anti-XYN_30 o anti-XYN_60 como se describe en el Ejemplo 19. Los genotipos de los transformantes seleccionados se analizaron mediante transferencia Southern utilizando el casete de expresión como sonda.

Se produjo XYN_30/Xyn10A de *Thermoascus* en *T. reesei* y el pH óptimo de la proteína heteróloga producida se determinó en tampón de McIlvaine universal dentro de un intervalo de pH de 3,0-8,0 utilizando xilano de abedul como sustrato (Figura 8A). Se determinó que el pH óptimo era 4,5. Se determinó que la temperatura óptima para la actividad enzimática de XYN_30 era 75°C (Figura 8B).

Los transformantes seleccionados se cultivaron, tal como se describe en el Ejemplo 14, en un biorreactor de 2 litros durante cuatro días (28°C, pH 4,2) para obtener material para los ensayos de aplicación.

Ejemplo 24. Funcionamiento de las celobiohidrolasas recombinantes en la hidrólisis

El funcionamiento de las celobiohidrolasas recombinantes purificadas se evaluó en los estudios de hidrólisis con enzimas de *T. reesei* purificadas. La hidrólisis se llevó a cabo con mezclas controladas de enzimas purificadas sobre varios sustratos pretratados. Los productos filtrados de los cultivos de *T. reesei*, que contenían diferentes enzimas CBH/Cel7 clonadas se obtuvieron como se describe en los Ejemplos 14 y 15, y las enzimas CBH se purificaron mediante cromatografía de afinidad como se describe en el Ejemplo 2. Además, se utilizaron celulasas de *T. reesei* puras (purificadas como describen Suurnäkki *et al.*, 2000) en las mezclas de enzimas. Las celobiohidrolasas utilizadas en el experimento fueron:

Thermoascus aurantiacus ALK04242 CBH (Ta Cel7A)

- 10 *Thermoascus aurantiacus* ALK04242 CBH (Ta Cel7A) con CBD anclado genéticamente de *Trichoderma reesei* (Ta Cel7A + Tr CBD)

Thermoascus aurantiacus ALK04242 CBH (Ta Cel7A) con CBD anclado genéticamente de *Chaetomium thermophilum* (Ta Cel7A + Ct CBD)

Acremonium thermophilum ALK04245 CBH (At Cel7A)

- 15 *Chaetomium thermophilum* ALK04265 CBH (Ct Cel7A).

Cada CBH/Cel7 que se iba a someter a ensayo (dosis 14,5 mg/g de materia seca de sustrato) se utilizó junto con EGII/Cel5A de *T. reesei* (3,6 mg/g) o con una mezcla que contenía EGI/Cel7B de *T. reesei* (1,8 mg/g), EGII/Cel5A (1,8 mg/g), xilanasa pl 9 (Tenkanen *et al.* 1992) (5000 nkat/g) y acetil xilano esterasa (AXE) (Sundberg y Poutanen, 1991) (250 nkat/g). A todas las mezclas se les añadió un suplemento de β -glucosidasa adicional de una preparación enzimática comercial Novozym 188 (176 nkat/g p.s.). Los tubos por triplicado que contenían la mezcla de enzima y 10 mg (materia seca)/ml del sustrato suspendido en acetato de sodio 0,05 M, se incubaron en la mezcla por agitación magnética a 45°C durante 48 h. También se prepararon muestras de referencia con enzimas inactivadas y sustratos correspondientes. La liberación de productos de la hidrólisis se midió en forma de azúcares reductores con el método DNS utilizando glucosa como patrón (Tabla 30).

- 25 Los siguientes sustratos se utilizaron en el experimento:

Celulosa cristalina (Avicel)

Fibra de abeto pretratada con vapor de agua lavada (impregnación con SO₂ al 3% p/p durante 20 minutos, seguido de pretratamiento con vapor de agua a 215°C durante 5 min), 25,9% de materia seca (ABETO).

Fibra rastrojo de maíz oxidada mojada lavada ("WOCS en sus siglas inglesas").

- 30 Fibra de sauce pretratada con vapor de agua lavada (pre-tratamiento durante 14 min a 210°C), 23,0% de materia seca (SAUCE).

Tabla 30. Productos de la hidrólisis con enzimas CBH (45°C, pH 5,0). Los productos de reacción después de 48 h de hidrólisis como azúcares reductores (mg/ml), glucosa medida como patrón. Abreviaturas: CBH = celobiohidrolasa; EGI = endoglucanasa I (Cel7B) de *T. reesei*, EGII = endoglucanasa II (Cel5A) de *T. reesei*; bG = β -glucosidasa (de Novozym 188); XYL = xilanasa pl 9 (Xyn II) de *T. reesei*, AXE = acetil xilano esterasa de *T. reesei*; nd = no realizado.

Enzimas		Sustratos			
CBH	Enzimas adicionales	Avicel	ABETO	WOCS	SAUCE
Ta Cel7A	EGII, bG	2,0	2,0	2,8	2,0
Ta Cel7A +Tr CBD	EGII, bG	5,8	4,0	4,4	4,0
Ta Cel7A +Ct CBD	EGII, bG	4,9	3,7	4,6	3,7
At Cel7A	EGII, bG	5,3	3,3	4,5	3,3
Ct Cel7A	EGII, bG	6,0	2,6	3,4	2,6
Cel7A de <i>T. reesei</i>	EGII, bG	4,7	2,9	2,9	2,9

Enzimas		Sustratos			
CBH	Enzimas adicionales	Avicel	ABETO	WOCS	SAUCE
Ta Cel7A	EGII, EGI, XYL, AXE, bG	nd	nd	4,3	2,8
Ta Cel7A +Tr CBD	EGII, EGI, XYL, AXE, bG	nd	nd	7,2	5,9
Ta Cel7A +Ct CBD	EGII, EGI, XYL, AXE, bG	nd	nd	7,2	5,6
At Cel7A	EGII, EGI, XYL, AXE, bG	nd	nd	6,4	5,4
Ct Cel7A	EGII, EGI, XYL, AXE, bG	nd	nd	5,6	4,0
Cel7A de <i>T. reesei</i>	EGII, EGI, XYL, AXE, bG	nd	nd	6,0	4,1

En la Tabla 30 se han comparado las diferentes celobiohidrolasas basándose en la misma dosis de proteína en la hidrólisis. Los resultados muestran que en sustratos celulósicos (Avicel y fibra de abeto) Cel7A de *Thermoascus aurantiacus* con el CBD anclado genéticamente mostró una hidrólisis claramente superior que CBHI/Cel7A de *T. reesei*. Sin CBD, Cel7A de *T. aurantiacus* fue menos eficiente en estos sustratos. El rendimiento de celobiohidrolasas de *Acremonium thermophilum* y *Chaetomium thermophilum* también fue mejor que el de CBHI/Cel7A de *T. reesei* sobre diversos sustratos; en particular, Cel7A de *C. thermophilum* mostró una alta eficacia sobre celulosa pura (Avicel).

En el caso de los sustratos que contienen cantidades notables de hemicelulosa (sauce y rastrojo de maíz) las enzimas CBH/Cel7 claramente necesitaron adicionalmente hemicelulasas y endoglucanasas para funcionar eficazmente. Si no estaban presentes hemicelulasas adicionales, Cel7A de *T. aurantiacus* con CBD anclado genéticamente volvió a mostrar una hidrólisis claramente más alta. Con las enzimas que degradan la hemicelulosa más importantes (xilanasas, acetil xilano esterasa y EGI) Cel7A de *T. aurantiacus* con CBD anclado genéticamente funcionó de nuevo con la más alta eficacia. Cel7A de *A. thermophilum* fue más eficaz que la enzima de *T. reesei* y Cel7A de *C. thermophilum* produjo productos de hidrólisis al mismo nivel que CBHI/Cel7A de *T. reesei*. El dominio de unión a celulosa de *T. reesei* parecía proporcionar una eficacia ligeramente mejor que el CBD de *C. thermophilum* en el funcionamiento hidrolítico de Cel7A de *T. aurantiacus*, aunque la diferencia fue bastante pequeña.

Se puede concluir que cuando se reemplazó CBHI/Cel7A en la mezcla de enzimas de *Trichoderma* por las celobiohidrolasas producidas en la presente memoria, la eficacia de la hidrólisis a juzgar por estas disposiciones experimentales mejoró claramente en el caso de Cel7A de *T. aurantiacus* con el CBD anclado genéticamente, y también mejoró en el caso de Cel7A de *A. thermophilum* y Cel7A de *C. thermophilum*. Considerando también la mejor estabilidad frente a la temperatura de las celobiohidrolasas producidas de la presente memoria, los resultados indican que el funcionamiento de las mezclas de enzimas celulasas a temperaturas más altas que 45°C puede ser claramente mejorado mediante el uso de las celobiohidrolasas producidas en la presente memoria.

Ejemplo 25. Rendimiento de las endoglucanasas recombinantes en la hidrólisis

Las preparaciones que contenían las endoglucanasas se compararon en estudios de hidrólisis mezcladas con las enzimas CBH/Cel7 y CBH/Cel6 purificadas sobre varios sustratos pretratados. Los productos filtrados de los cultivos de *T. reesei*, que contenían diferentes enzimas endoglucanasas clonadas se obtuvieron como se describe en el Ejemplo 19. Las enzimas fueron enriquecidas mediante la eliminación de las proteínas termolábiles a partir de las mezclas mediante un tratamiento térmico (60°C, 2 h, pH 5) y el sobrenadante se utilizó para los estudios de hidrólisis. Además, las celulasas de *T. reesei* puras (purificadas como describen Suurnäkki *et al.*, 2000) se utilizaron en las mezclas de enzimas. Las endoglucanasas utilizadas en el experimento fueron:

Endoglucanasa A EG_40/Cel45A (ALKO4245 EG_40) de *Acremonium thermophilum* ALKO4245

Endoglucanasa A de tipo_EG_40/Cel45B (ALKO4245 de tipo_EG_40) de *Acremonium thermophilum* ALKO4245

Endoglucanasa Ta EG_28/Cel5A (ALKO4242 EG_28) de *Thermoascus aurantiacus* ALKO4242.

Los siguientes sustratos se utilizaron en el experimento:

Fibra de abeto pretratada con vapor de agua lavada (impregnación con SO₂ al 3% durante 20 min, seguido de pretratamiento con vapor a 215°C durante 5 min), 25,9% de materia seca (ABETO).

Rastrojo de maíz sometido a explosión de vapor de agua (pretratamiento con vapor a agua a 210°C durante 5 min), materia seca 31,0% (SECS).

Las endoglucanasas a estudiar (dosis 840 nkat/g de materia seca, basada de la actividad endoglucanasa contra HEC de acuerdo con IUPAC, 1987) se utilizaron, con celobiohidrolasas de *T. reesei* (CBHI/Cel7A, 8,1 mg/g m.s. y CBHII/Cel6A, 2,0 mg/g m.s.) o con Cel7A de *Thermoascus aurantiacus* con CBD anclado genéticamente de *T. reesei* (10,1 mg/g m.s.). También se incluyeron EGI (Cel7B) y EGII (Cel5A) de *T. reesei* purificadas (Suurnäkki *et al.*, 2000) en los experimentos para la comparación. A todas las mezclas se les añadió un suplemento de β -glucosidasa adicional de Novozym 188 (para hacer la dosificación total de β -glucosidasa de 560 nkat/g p.s., se utilizó la dosificación relativamente alta para compensar las diferencias en las actividades de fondo de las diferentes preparaciones de EG). Se incubaron tubos por triplicado en la mezcla a 45°C durante 48 h y se prepararon muestras de referencia con enzimas inactivadas y sustratos correspondientes. La liberación de productos de la hidrólisis se midió en forma de azúcares reductores con el método DNS utilizando glucosa como patrón (Tabla 31).

Tabla 31. Productos de hidrólisis con diferentes preparaciones de endoglucanasa cuando se utilizan junto con celobiohidrolasas de *T. reesei* o con Cel7A de *T. aurantiacus* que alberga CBD de *T. reesei*. Productos de reacción después de 48 h de hidrólisis (45°C, pH 5,0) en forma de azúcares reductores (mg/ml), glucosa medida como patrón. Abreviaturas: CBHI = celobiohidrolasa I (Cel7A) de *T. reesei*; CBHII = celobiohidrolasa II (Cel6A) de *T. reesei*; EGI = endoglucanasa I (Cel7B) de *T. reesei*, EGII = endoglucanasa II (Cel5A) de *T. reesei*; BG = β -glucosidasa (de Novozym 188); nd. = no realizado.

Enzimas		Sustrato	
Endoglucanasa	CBH/Cel7	ABETO	SECS
EG no añadida	CBHI and CBHII de <i>T. reesei</i>	2,4	3,2
EGI	CBHI and CBHII de <i>T. reesei</i>	3,5	4,6
EGII	CBHI and CBHII de <i>T. reesei</i>	3,8	3,5
At EG_40	CBHI and CBHII de <i>T. reesei</i>	4,9	4,3
de tipo _At EG_40	CBHI and CBHII de <i>T. reesei</i>	4,5	4,8
Ta EG_28	CBHI and CBHII de <i>T. reesei</i>	3,0	3,9
Sin añadir EG	Cel7A + Tr CBD de <i>T. aurantiacus</i>	1,8	2,1
EG I	Cel7A + Tr CBD de <i>T. aurantiacus</i>	nd,	4,2
EG II	Cel7A + Tr CBD de <i>T. aurantiacus</i>	3,2	nd,
At EG_40	Cel7A + Tr CBD de <i>T. aurantiacus</i>	4,8	4,0
Ta EG_28	Cel7A + Tr CBD de <i>T. aurantiacus</i>	1,5	nd,

En la Tabla 31 se han comparado las diferentes endoglucanasas basándose en la misma dosificación de actividad en la hidrólisis. Esto puede favorecer las enzimas con baja actividad específica contra el sustrato (hidroxietilcelulosa) utilizado en el análisis y subestimar la eficacia de las enzimas con alta actividad específica contra hidroxietilcelulosa. En cualquier caso, los resultados muestran que las endoglucanasas de *Acremonium thermophilum* funcionan muy bien en la hidrólisis cuando ejercen su efecto junto con ambas celobiohidrolasas utilizadas en la mezcla. Las endoglucanasas de *A. thermophilum* tienen un rendimiento similar a EGI/Cel7B de *T. reesei* que es una enzima muy eficaz en el sustrato de rastrojo de maíz que contiene hemicelulosa debido a su fuerte actividad secundaria xilanas. La endoglucanasa Cel5A (ALK04242 EG_28) de *T. aurantiacus* mostró menor hidrólisis que las enzimas de *T. reesei*.

Se puede concluir que las endoglucanasas de *A. thermophilum* funcionan con una eficacia comparable o mejorada en comparación con las enzimas de *Trichoderma* correspondientes en la hidrólisis como se determina por esta

disposición experimental. Considerando también la estabilidad frente a la temperatura de las endoglucanasas descritas en la presente memoria, los resultados indican que el rendimiento de mezclas de enzimas celulasas a temperaturas más altas que 45°C se puede mejorar mediante el uso de las endoglucanasas descritas en la presente memoria.

5 Ejemplo 26. Hidrólisis de abeto pretratado con vapor de agua a altas temperaturas

Se suspendió fibra de abeto sometida a explosión de vapor de agua lavada (impregnación con SO₂ al 3% p/p durante 20 min, seguido de pretratamiento con vapor de agua a 215°C durante 5 min), con materia seca de 25,9% en 5 ml de tampón de acetato de sodio 0,05 M a la consistencia de 10 mg/ml. Este sustrato se hidrolizó utilizando diferentes mezclas de enzimas en tubos de ensayo con agitación magnética en el baño de agua ajustado a diferentes temperaturas durante 72 h. Para cada punto de muestra, se retiraron tubos de ensayo por triplicado de la hidrólisis, se hirvieron durante 10 min con el fin de terminar la hidrólisis enzimática, se centrifugaron, y el sobrenadante se analizó para determinar los productos de reacción de la hidrólisis. Los blancos que contenían el sustrato solo (solo se añadió tampón en lugar de las enzimas) también se incubaron en las condiciones correspondientes.

15 Se preparó una mezcla de celulasas termófilas utilizando los siguientes componentes:

Preparación de CBH/Cel7 termófila que contenía Cel7A de *Thermoascus aurantiacus* ALKO4242 con CBD anclado genéticamente de CBHI/Cel7A de *T. reesei*. La preparación de proteína se produjo como se describe en el Ejemplo 15 y se purificó de acuerdo con el Ejemplo 2 dando como resultado la preparación de Ta Cel7A + Tr CBD purificada con un contenido de proteína de 5,6 mg/ml.

20 Preparación de endoglucanasa termófila que contenía endoglucanasa A EG_40/Cel45A de *Acremonium thermophilum* ALKO4245. La proteína se produjo en *T. reesei* como se describe en el Ejemplo 19. Con el fin de enriquecer los componentes termófilos, el medio de cultivo gastado se trató con calor (60°C durante 2 horas). La preparación obtenida contenía 4,9 mg/ml de proteína y una actividad endoglucanasa (de acuerdo con la IUPAC, 1987) de 422 nkat/ml.

25 Preparación de β-glucosidasa termófila preparada como se ha descrito en el Ejemplo 21 que contenía β-glucosidasa Ta βG_81/Cel3A de *Thermoascus aurantiacus* ALKO4242. Con el fin de enriquecer los componentes termófilos, el caldo del fermentador se trató con calor (65°C durante 2 horas). La preparación obtenida contenía 4,3 mg/ml de proteína y una actividad β-glucosidasa de 6270 nkat/ml (de acuerdo con Bailey y Linko, 1990).

30 Estas preparaciones de enzimas se combinaron como sigue (por 10 ml de la mezcla): 4,51 ml de preparación de CBH/Cel7, 5,19 ml de preparación de endoglucanasa y 0,29 ml de preparación de β-glucosidasa. Esta mezcla se usó como "MEZCLA 1" de enzimas termófilas.

A modo de comparación y de referencia, se construyó una mezcla del estado de la técnica de enzimas de *Trichoderma reesei* comerciales combinando (por 10 ml): 8,05 ml de Celluclast 1,5 L FG (de Novozymes A/S) y 1,95 ml de Novozym 188 (de Novozymes A/S). Esto se denominó "ENZIMAS DE T. REESEI".

35 Las enzimas se dosificaron sobre la base de la actividad FPU de las mezclas: "MEZCLA 1", utilizando la dosis de 5,5 FPU por 1 gramo de materia seca en el sustrato de abeto, y "ENZIMAS DE T. REESEI" utilizando 5,8 FPU por 1 gramo de materia seca en el sustrato de abeto.

40 Se tomaron muestras de la hidrólisis después de 24, 48 y 72 h y se trataron como se ha descrito anteriormente. Los productos de hidrólisis se cuantificaron utilizando el análisis de azúcares reductores (Bernfeld, 1955), utilizando glucosa como patrón. La cantidad de productos de hidrólisis en forma de azúcares reductores se presenta en la Figura 9.

45 Los resultados muestran claramente un mejor funcionamiento de las enzimas descritas en la presente memoria en comparación con las enzimas de *Trichoderma* del estado de la técnica a 55°C y 60°C en el sustrato de abeto. Basándose en el análisis de HPLC el máximo rendimiento de los azúcares del sustrato sería 5,67 mg por 10 mg de sustrato de abeto seco. Debido a la relativamente baja dosis de enzima los rendimientos finales de azúcar fueron claramente inferiores. Para enzimas termoestables el rendimiento de azúcar basado en el análisis de azúcares reductores fue de 66% y 57% del teórico a 55°C y 60°C, respectivamente. Para las enzimas de *Trichoderma* del estado de la técnica fue solo de 31% y de 11% a 55°C y 60°C, respectivamente.

Ejemplo 27. Hidrólisis de rastrojo de maíz pretratado con vapor de agua a altas temperaturas

50 Se suspendió fibra de rastrojo de maíz sometida a explosión de vapor de agua (tratamiento a 195°C durante 5 min), con materia seca de 45,3% en 5 ml de tampón de acetato de sodio 0,05 M a la consistencia de 10 mg/ml. Los tratamientos y las mediciones se realizaron como se describe en el Ejemplo 26.

Se construyó una mezcla de las celulasas termófilas descritas en la presente memoria utilizando los siguientes componentes:

Preparación de CBH termófila que contenía Cel7A de *Thermoascus aurantiacus* ALKO4242 con CBD anclado genéticamente de CBH1/Cel7A de *T. reesei* (Ta Cel7A + Tr CBD, Ejemplo 15). El contenido de proteína de la preparación fue de 31 mg/ml.

- 5 La preparación de endoglucanasa termófila que contenía la endoglucanasa At EG_40/Cel45A de *Acremonium thermophilum* ALKO4245 se obtuvo como se describe en el Ejemplo 19. La preparación enzimática concentrada contenía una actividad endoglucanasa (de acuerdo con la IUPAC, 1987) de 2057 nkat/ml.

La preparación de β -glucosidasa termófila que contenía la β -glucosidasa Ta β G_81/Cel3A de *Thermoascus aurantiacus* ALKO 4242 se obtuvo como se describe en el Ejemplo 21 conteniendo una actividad β -glucosidasa (de acuerdo con Bailey y Linko, 1990) de 11500 nkat/ml.

- 10 Producto de xilanasa termófila que contenía una xilanasa AM24 procedente de *Nonomuraea flexuosa* DSM43186. El producto se preparó mediante el uso de una cepa de *Trichoderma reesei* recombinante que había sido transformada con el casete de expresión pALK1502, como se describe en el documento WO2005/100557. El producto sólido se disolvió en agua para elaborar una disolución al 10% y se obtuvo una preparación enzimática con actividad xilanasa (analizada de acuerdo con Bailey *et al.*, 1992) de 208000 nkat/ml.
- 15 Estas preparaciones enzimáticas se combinaron como sigue (por 10 ml de mezcla): 7,79 ml de preparación de CBH/Cel7, 0,96 ml de preparación de endoglucanasa, 1,14 ml de preparación de β -glucosidasa y 0,31 ml de preparación de xilanasa. Esta mezcla se usó como "MEZCLA 2" de enzimas termófilas.

- 20 A modo de comparación y de referencia, se construyó una mezcla del estado de la técnica de enzimas de *Trichoderma reesei* comerciales combinando (por 10 ml) 8,05 ml de Celluclast 1,5 L FG (de Novozymes A/S) y 1,95 ml de Novozym 188 (de Novozymes A/S). Esto se denominó "ENZIMAS DE *T. REESEI*".

Se tomaron muestras de la hidrólisis después de 24, 48 y 72 h y se trataron como se ha descrito anteriormente. Los productos de hidrólisis se cuantificaron utilizando el análisis de azúcares reductores (Bernfeld, 1955), utilizando glucosa como patrón. Los resultados de los blancos de sustrato se restaron de las muestras con enzimas, y la concentración de productos de hidrólisis en forma de azúcares reductores se presenta en la Figura 10.

- 25 Los resultados muestran claramente un mejor funcionamiento de las enzimas descritas en la presente memoria en comparación con las enzimas de *Trichoderma* del estado de la técnica. A 45°C la mezcla de enzimas termófilas mostró una hidrólisis más eficaz en comparación con las enzimas de *T. reesei*. La hidrólisis fue más rápida y también se obtuvieron rendimientos más altos de azúcar. Sobre la base del análisis HPLC el máximo rendimiento de azúcares (incluyendo azúcares solubles libres en el sustrato sin lavar que se utilizó) del sustrato sería de 5,73 mg
- 30 por 10 mg de sustrato seco. Por lo tanto, la hidrólisis por la MEZCLA 2 de enzimas fue casi completa en el plazo de 48 horas. A 55°C y 57,5°C las enzimas termófilas descritas en la presente memoria mostraron también claramente un mejor funcionamiento en la hidrólisis en comparación con las enzimas de *Trichoderma* del estado de la técnica.

Ejemplo 28. Hidrólisis del rastrojo de maíz pretratado a altas temperaturas utilizando una mezcla con una xilanasa termoestable

- 35 Se repitió el procedimiento explicado en el Ejemplo 27 excepto que el producto de xilanasa XT 02026A3 fue reemplazado por una preparación de xilanasa termófila que contenía la xilanasa Ta XYN_30/Xyn10A de *Thermoascus aurantiacus* ALKO4242 producida en *T. reesei*. El caldo del fermentador, producido como se describe en el Ejemplo 23 contenía una actividad xilanasa de 132000 nkat/ml (analizada de acuerdo con Bailey *et al.*, 1992).

- 40 Estas preparaciones enzimáticas se combinaron como sigue (por 10 ml de la mezcla): 7,64 ml de preparación de CBH/Cel7, 0,96 ml de preparación de endoglucanasa, 1,15 ml de preparación de β -glucosidasa y 0,25 ml de preparación de xilanasa. Esta mezcla se utilizó como "MEZCLA 3" de enzimas termófilas.

A modo de comparación y de referencia, se construyó una mezcla del estado de la técnica de enzimas de *Trichoderma reesei* comerciales combinando (por 10 ml) 8,05 ml de Celluclast 1,5 L FG (de Novozymes A/S) y 1,95 ml de Novozym 188 (de Novozymes A/S). Esto se denominó "ENZIMAS DE *T. REESEI*".

- 45 Se tomaron muestras de la hidrólisis después de 24, 48 y 72 h y se trataron como se ha descrito anteriormente. Los productos de hidrólisis se cuantificaron utilizando el análisis de azúcares reductores (Bernfeld, 1955), utilizando glucosa como patrón. Los resultados de los blancos de sustrato se restaron de las muestras con enzimas, y la concentración de productos de hidrólisis en forma de azúcares reductores se presenta en la Figura 11.

- 50 Los resultados muestran claramente un mejor funcionamiento de la mezcla de las enzimas descritas en la presente memoria en comparación con las enzimas de *Trichoderma* del estado de la técnica. A 45°C la mezcla de enzimas termófilas mostró una hidrólisis más eficaz en comparación con las enzimas *T. reesei*. A 55°C y 60°C, las enzimas termófilas descritas en la presente memoria muestran claramente un mejor funcionamiento en la hidrólisis en comparación con las enzimas de *Trichoderma* del estado de la técnica. El funcionamiento de la nueva mezcla enzimática a 60°C estaba al mismo nivel que el funcionamiento de las enzimas del estado de la técnica a 45°C.

Ejemplo 29. Hidrólisis de abeto pretratado a altas temperaturas utilizando una mezcla con una xilanasa termoestable

Se repitió el procedimiento descrito en el Ejemplo 28 con fibra de abeto sometida a explosión de vapor de agua lavada (impregnación con SO₂ al 3% p/p durante 20 min, seguido de pretratamiento con vapor a 215°C durante 5 min, con materia seca de 25,9%) como sustrato utilizando temperaturas de hidrólisis 45°C, 55°C y 60°C. Se tomaron muestras de la hidrólisis después de 24, 48 y 72 h y se trataron como se ha descrito anteriormente. Los productos de hidrólisis se cuantificaron utilizando el análisis de azúcares reductores (Bernfeld, 1955), utilizando glucosa como patrón. Los resultados de los blancos de sustrato se restaron de las muestras con enzimas, y la concentración de productos de hidrólisis en forma de azúcares reductores se presenta en Figura 12.

Los resultados muestran claramente un mejor funcionamiento de la mezcla de las enzimas descritas en la presente memoria, en comparación con las enzimas de *Trichoderma* del estado de la técnica a todas las temperaturas estudiadas. A 45°C la mezcla de enzimas termófilas mostró una hidrólisis más eficiente en comparación con las enzimas de *T. reesei*, evidentemente debido a la mejor estabilidad en la hidrólisis a largo plazo. A 55°C la eficacia de la mezcla de las enzimas descritas en la presente memoria estaba todavía al mismo nivel que a 45°C, mientras que la mezcla del estado de la técnica era ineficaz con el sustrato utilizado a esta temperatura. A 60°C las enzimas termófilas descritas en la presente memoria mostraron una disminución de la hidrólisis aunque la hidrólisis estaba casi al mismo nivel que el rendimiento de las enzimas del estado de la técnica a 45°C.

Ejemplo 30. Evaluación de la inhibición por glucosa de las β-glucosidasas de *Acremonium thermophilum* ALKO4245, *Chaetomium thermophilum* ALKO4261 y *Thermoascus aurantiacus* ALKO4242

Los productos filtrados de cultivo producidos por las cepas *Acremonium thermophilum* ALKO4245, *Chaetomium thermophilum* ALKO4261 y *Thermoascus aurantiacus* ALKO4242 se describen en el Ejemplo 1. Las actividades β-glucosidasa (medidas de acuerdo con Bailey y Linko, 1990) de estas preparaciones fueron 21,4 nkat/ml, 5,6 nkat/ml y 18,6 nkat/ml, respectivamente. Para la comparación, también se incluyeron en el experimento las enzimas comerciales β-glucosidasa Celluclast 1,5L de 534 nkat/ml) y (β-glucosidasa Novozym 188 de 5840 nkat/ml).

Con el fin de evaluar la sensibilidad de las diferentes β-glucosidasas con miras a la inhibición por glucosa, el procedimiento de análisis de actividad convencional se realizó en presencia de diferentes concentraciones de glucosa. A las disoluciones de sustrato (p-nitrofenil-β-D-glucopiranosido) para el análisis de la actividad β-glucosidasa se les añadió un suplemento de glucosa de manera que la concentración de glucosa en la mezcla de análisis se ajustara a los valores de 0 a 0,5 M. Excepto esta adición de glucosa el análisis se realizó utilizando el procedimiento convencional (Bailey y Linko, 1990). Las actividades en la presencia de diferentes concentraciones de glucosa como un porcentaje de la actividad sin glucosa se presentan en la Figura 13.

Los resultados muestran que las β-glucosidasas de *C. thermophilum* y *T. aurantiacus* resultan menos afectadas por la inhibición por glucosa que las β-glucosidasas presentes en las enzimas comerciales: β-glucosidasa derivada de *Aspergillus* en Novozym 188 o β-glucosidasa derivada de *Trichoderma* en Celluclast 1,5L. La enzima de *A. thermophilum* mostró un comportamiento comparable a la enzima de *T. reesei* de Celluclast. Especialmente la enzima de *C. thermophilum* resultó claramente menos afectada por la alta concentración de glucosa. Por lo tanto, estos resultados indican que la inhibición por glucosa teniendo en cuenta el uso de la nueva β-glucosidasa, especialmente a partir de cepas *Acremonium thermophilum* ALKO4242 y *Chaetomium thermophilum* ALKO4261, proporcionaría claras ventajas en la hidrólisis en condiciones industriales con alta concentración de glucosa.

Ejemplo 31. Actividad sobre papel de filtro de mezclas de enzimas a altas temperaturas

La actividad sobre papel de filtro de las preparaciones enzimáticas se midió de acuerdo con el método de la IUPAC (1987) como se describe en el procedimiento, excepto que la reacción enzimática se realizó a temperaturas de 50°C a 70°C. La actividad FPU calculada se basa en la cantidad de enzima requerida para hidrolizar 4% del sustrato de papel de filtro en 1 h en las condiciones experimentales. Se considera que la actividad FPU representa la actividad de la celulasa global total de una preparación enzimática.

Las mezclas de enzimas fueron la MEZCLA 2 preparada como se ha descrito en el Ejemplo 27, la MEZCLA 3 preparada como se ha descrito en el Ejemplo 28, y la MEZCLA 4. La MEZCLA 4 se preparó combinando las preparaciones enzimáticas descritas en el Ejemplo 27 como sigue (por 10 ml de mezcla): 7,84 ml de preparación de CBH/Cel7, 0,99 ml de preparación de endoglucanasa y 1,17 ml de preparación de β-glucosidasa.

Las mezclas de enzimas utilizadas como referencia, que representan las mezclas del estado de la técnica, fueron:

"ENZIMAS A de T. REESEI" preparadas como preparación de "ENZIMAS de T. REESEI" descrita en el Ejemplo 26.

Las "ENZIMAS B de T. REESEI" se construyeron combinando (por 10 ml) 8,05 ml de Econase CE (una preparación de celulasa de *T. reesei* comercial de AB Enzymes Oy, Rajamäki, Finlandia) y 1,95 ml de Novozym 188 (de Novozymes A/S).

Las actividades FPU medidas para las preparaciones enzimáticas a diferentes temperaturas se presentan en la Figura 14 en forma de porcentajes de la actividad en condiciones convencionales (IUPAC, 1987) (a 50°C).

Los resultados muestran claramente que las mezclas aquí descritas muestran una mayor actividad celulasa general a elevadas (60-70°) temperaturas en comparación con las mezclas del estado de la técnica basadas en enzimas de *Trichoderma* y *Aspergillus*.

Ejemplo 32. Uso de las nuevas beta-glucosidasas en la preparación de soforosa

- 5 Una alta concentración de la mezcla de producto hidrolizado de almidón (Nutriose 74/968, Roquette) se trató con la preparación enzimática enriquecida en β G_81/Cel3A de *Thermoascus aurantiacus* producida como se ha descrito en el Ejemplo 21 para producir una mezcla de azúcar que contenía cantidades apreciables de inductor de celulasa (soforosa) para superar la represión por glucosa .

- 10 La preparación enzimática enriquecida en Ta β G_81/Cel3A se añadió a una disolución al 70% (p/p) de Nutriose a una concentración final de 1 g de proteína total/litro. El recipiente de la mezcla se incubó en un baño de agua a 65°C durante 3 días con agitación constante y se utilizó como fuente de carbono en un medio de matraz oscilante para dos diferentes cepas de *Trichoderma* (A47 y Rut-C30). El efecto del tratamiento enzimático se midió como la actividad endoglucanasa formada durante un cultivo en matraz oscilante de 7 días. Como referencia los cultivos se realizaron en las mismas condiciones con Nutriose sin tratar como fuente de carbono. Se obtuvo un aumento de más del doble en las actividades en los cultivos en matraces oscilantes realizados en medios Nutriose pretratados con Ta β G_81/Cel3A con las cepas sometidas a ensayo. Los resultados se muestran en la Figura 15.

15

Lista de los organismos depositados

Cepa	Plásmido contenido	Autoridad de Depósito	Fecha de Depósito	Número de depósito
<i>Acremonium thermophilum</i> ALKO4245	-	CBS ⁽¹⁾	20 Sep 2004	CBS 116240
<i>Thermoascus aurantiacus</i> ALKO4242	-	CBS ⁽¹⁾	20 Sep 2004	CBS 116239
<i>Chaetomium thermophilum</i> ALKO4265	-	CBS ⁽²⁾	8 Nov 1995	CBS 730.95 ⁽⁴⁾
<i>Escherichia coli</i>	pALK1635	DSMZ ⁽³⁾	16 Sep 2004	DSM 16723
<i>Escherichia coli</i>	pALK1642	DSMZ	16 Sep 2004	DSM 16727
<i>Escherichia coli</i>	pALK1646	DSMZ	16 Sep 2004	DSM 16728
<i>Escherichia coli</i>	pALK1861	DSMZ	16 Sep 2004	DSM 16729
<i>Escherichia coli</i>	pALK1715	DSMZ	16 Sep 2004	DSM 16724
<i>Escherichia coli</i>	pALK1723	DSMZ	16 Sep 2004	DSM 16725
<i>Escherichia coli</i>	pALK1725	DSMZ	16 Sep 2004	DSM 16726
<i>Escherichia coli</i>	pALK1904	DSMZ	13 Mayo 2005	DSM 17323
<i>Escherichia coli</i>	pALK1908	DSMZ	13 Mayo 2005	DSM 17324
<i>Escherichia coli</i>	pALK1925	DSMZ	13 Mayo 2005	DSM 17325
<i>Escherichia coli</i>	pALK1926	DSMZ	13 Mayo 2005	DSM 17326
<i>Escherichia coli</i>	pALK2001	DSMZ	18 Oct 2005	DSM 17667
<i>Escherichia coli</i>	pALK2010	DSMZ	18 Nov 2005	DSM 17729

⁽¹⁾ Centraalbureau Voor Schimmelcultures at Uppsalalaan 8, 3584 CT, Utrecht, Holanda

⁽²⁾ Centraalbureau Voor Schimmelcultures at Oosterstraat 1, 3742 SK BAARN, Holanda

⁽³⁾ Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1 b, D-38124 Braunschweig, Alemania

⁽⁴⁾ [Después de la finalización del período de depósito actual, las muestras se almacenarán bajo los acuerdos para hacer que la cepa esté disponible más allá del tiempo ejecutorio de la patente.]

Referencias

- Altschul S., Gish W., Miller W., Myers E. W. y Lipman D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- 5 Badger, P.C. (2002) Ethanol from cellulose: a general review. En *Trends in new crops and new uses*. J. Janick y A. Whipkey (eds.). ASHS Press, Alexandria, VA, USA, pp. 17-21.
- Bailey M. J. and K. M. H. Nevalainen (1981) Induction, isolation and testing of stable *Trichoderma reesei* mutants with improved production of solubilizing cellulase. *Enz Microbiol Technol.* 3: 153-157.
- Bailey, M.J., Biely, P. y Poutanen, K. (1992) Interlaboratory testing for assay of xylanase activity. *J. Biotechnol.* 23:257-270.
- 10 Bailey, M. J. y Linko, M. (1990) Production of β -galactosidase by *Aspergillus oryzae* in submerged bioreactor cultivation. *J. Biotechnol.* 16:57-66.
- Bailey M. J. y Poutanen K. (1989) Production of xylanases by strains of *Aspergillus*. *Appl. Microbiol. Biotechnol.* 30:5-10.
- 15 Bailey M., Siika-aho M., Valkeajärvi A. y Penttilä M. (1993) Hydrolytic properties of two cellulases of *Trichoderma reesei* expressed in yeast. *Biotechnol. Appl. Biochem* 17: 65-76.
- Bendtsen J. D., Nielsen H., von Heijne G. y Brunak S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340: 783-795.
- Bernfeld, B. (1955) Amylases, α and β . En: *Methods in Enzymology*, vol. 1, Eds. Colowick, S.P. y Kaplan, N.O. Academic Press, New York, pp. 149-158.
- 20 Biely P., Vrsanska M., Tenkanen M., Kluepfel D. (1997) Endo-beta-1,4-xylanase families: differences in catalytic properties. *Journal of Biotechnology* 57: 151-166.
- Coen, D.M. (2001) The polymerase chain reaction. En: Ausubel, F.M., Brent, R., Kingston, R.E., More, D.D., Seidman, J.G., Smith, K. y Struhl, K. (eds.) *Current protocols in molecular biology*. John Wiley & Sons. Inc., Hoboken, USA.
- 25 Gasteiger, E., Gattiker A., Hoogland C., Ivanyi I., Appel R. D. y Bairoch A. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31: 3784-3788.
- Gellissen, G. (ed.) (2005) *Production of recombinant proteins. Novel microbial and eukaryotic expression systems*. Wiley-VCH Verlag GmbH&Co. Weinheim, Germany.
- 30 Gill, S.C. y von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182: 319-326.
- Gupta, R., E. Jung y S. Brunak. (2004) Prediction of N-glycosylation sites in human proteins, In preparation. www.cbs.dtu.dk/services/NetNGlyc/
- 35 Haakana H., Miettinen-Oinonen A., Joutsjoki V., Mäntylä A., Suominen P. y Vehmaanperä J. (2004) Cloning of cellulase genes from *Melanocarpus albomyces* and their efficient expression in *Trichoderma reesei*. *Enz. Microbiol. Technol.* 34: 159-167.
- Henrissat B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280: 309-316.
- Henrissat B. y Bairoch A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 293: 781-788.
- 40 Henrissat B. y Bairoch A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316: 695-696
- Henrissat B., Teeri T. T. y Warren R. A. J. (1998) A scheme for designating enzymes that hydrolyse the polysaccharides in the cell wall of plants. *FEBS Letters* 425: 352-354.
- 45 Hong J., H. Tamaki, K. Yamamoto, y Kumagai H. (2003a) Cloning of a gene encoding a thermo-stabile endo- β -1,4-glucanase from *Thermoascus aurantiacus* and its expression in yeast. *Biotech. Letters* 25: 657-661.
- Hong J., Tamaki H., Yamamoto K. y Kumagai H. (2003b) Cloning of a gene encoding thermostable cellobiohydrolase from *Thermoascus aurantiacus* and its expression in yeast. *Appl. Microbiol. Biotechnol.* 63: 42-50.

- IUPAC (International Union of Pure and Applied Chemistry) (1987). Measurement of cellulase activities. *Pure and Appl. Chem.* 59: 257-268.
- 5 Joutsjoki, V. V., Torkkeli T.K. y Nevalainen K. M. H. (1993) Transformation of *Trichoderma reesei* with the *Hormoconis resinae* glucoamylase P (gamP) gene: production of a heterologous glucoamylase by *Trichoderma reesei*. *Curr. Genet.* 24: 223-228.
- Karhunen T., Mäntylä A., Nevalainen K. M. H. y Suominen P. L. (1993) High frequency one-step gene replacement in *Trichoderma reesei*. I. Endoglucanase I overproduction. *Mol. Gen. Genet.* 241: 515-522.
- 10 Kurabi A., Berlin A., Gilkes N., Kilburn D., Markov A., Skomarovsky A., Gusakov A., Okunev O., Sinitsyn A., Gregg D. Xie D. y Saddler J. (2005) Enzymatic hydrolysis of steam-exploded and ethanol organosolv-pretreated Douglas-Fir by novel and commercial fungal cellulases. *Appl. Biochem and Biotechnol.* Vol 121-124: 219-229.
- Lever, M. (1972) A new reaction for colorimetric determination of carbohydrates. *Anal. Biochem.*, 47: 276-279.
- Lo Leggio, L., Kalogiannis S., Bhat M.K., y Pickersgill R.W. (1999) High resolution structure and sequence of the *T. aurantiacus* xylanase I: implications for evolution of thermostability in family 10 xylanases and enzymes with (beta) alpha-barrel architecture. *Proteins* 36(3): 295-306.
- 15 Lowry, O., Rosenbrough, N., Farr, A. y Randall, R. (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193: 265-275.
- Needleman S. y Wunsch C. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* 48, 443-453.
- 20 Nielsen H., Engelbrecht J., Brunak S. y von Heijne G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* 10: 1-6.
- Paloheimo M., Mäntylä A., Kallio J., y Suominen P. (2003) High-yield production of a bacterial xylanase in the filamentous fungus *Trichoderma reesei* requires a carrier polypeptide with an intact domain structure. *Appl. Env. Microbiol.* 69: 7073-7082.
- 25 Parry N., Beever D., Owen E., Nerinckx W. Claeyssens M, Van Beeumen J. y Bhat M. (2002) Biochemical characterization and mode of action of a thermostable endoglucanase purified from *Thermoascus aurantiacus*, *Arch. of Biochem. and Biophys.* 404: 243-253.
- Penttilä M., Nevalainen H., Rättö M., Salminen E. y Knowles J. (1987) A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. *Gene* 61:155-164.
- Raeder U. y Broda P. (1985) Rapid preparation of DNA from filamentous fungi. *Lett. Appl. Microbiol.* 1: 17-20.
- 30 Rice P, Longden I y Bleasby A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics* 16:276-277.
- Sambrook J., Fritsch E. F. y Maniatis T. (1989) *Molecular cloning, a laboratory manual*. Cold Spring Harbor Laboratory, New York, US.
- 35 Sambrook J. y Russell D. W. (2001) *Molecular cloning, a laboratory manual*. Cold Spring Harbor Laboratory, New York, US.
- Srisodsuk M, Reinikainen T, Penttilä M y Teeri T. (1993) Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. *J. Biol. Chem.* Oct 5;268(28): 20756-61.
- Sundberg, M., y Poutanen, K. (1991) Purification and properties of two acetylxyylan esterases of *Trichoderma reesei*. *Biotechnol. Appl. Biochem.* 13: 1-11.
- 40 Suurnäkki, A., Tenkanen M., Siika-aho, M., Niku-Paavola, M.-L., Viikari, L. y Buchert, J. (2000) *Trichoderma reesei* cellulases and their core domains in the hydrolysis and modification of chemical pulp. *Cellulose* 7: 189-209.
- Tenkanen, M., Puls, J. y Poutanen, K (1992) Two major xylanases of *Trichoderma reesei*. *Enzyme Microbiol. Technol.* 14: 566-574.
- 45 Tomme, P. McRae, S., Wood, T. y Claeyssens, M. (1988) Chromatographic separation of cellulolytic enzymes. *Methods in Enzymol.* 160: 187-192.
- Tuohy M., Walsh J., Murray P., Claeyssens M., Cuffe M., Savage A. y Coughan M. (2002) Kinetic parameters and mode of action of cellobiohydrolases produced by *Talaromyces emersonii*. *Biochem. Biophys. Acta* 1596: 366-380 (abstract).

Van Petegem et al (2002) Atomic resolution structure of major endoglucanase from *Thermoascus aurantiacus*, *Biochem. and Biophys. Res. Comm.* 296, pp. 161-166.

Van Tilbeurgh, H., Loontjens, F., de Bruyne, C. y Claeysens, M. (1988) Fluorogenic and chromogenic glycosides as substrates and ligands of carbohydrases. *Methods Enzymol.* 160:45-59.

- 5 Wyman, C.E. (2001) Twenty years of trials, tribulations, and research progress in bioethanol technology. *Applied Biochemistry and Biotechnology* 91-93: 5-21.

LISTA DE SECUENCIAS

<110> Roal

<120> Tratamiento de material celulósico y enzimas útiles en el mismo

5 <130> 2051999

<160> 30

10 <170> PatentIn versión 3.1

<210> 1

<211> 3192

<212> ADN

15 <213> Thermoascus aurantiacus

<220>

<221> CDS

<222> (1514)..(2122)

20 <223>

<220>

<221> Intrón

<222> (2123)..(2187)

25 <223>

<220>

<221> CDS

<222> (2188)..(2949)

30 <223>

<400> 1

ctagaccttt atcctttcat ccgaccagac ttcccctttg accttggcgc cctgttgact 60

acctacctac ctaggtaggt aacgtcgtcg accctcttga atgacccctcg tcacactgca 120

aacatccgaa acatacggca aaagatgatt gggcatggat gcaggagaca tcgaatgagg	180
gcttagaagg aaatgaaaac ctgggaccag gacgctaggt acgatgaaat ccgccaatgg	240
tgaaacttta agtcgtgcct acagcacagg ctctgtgaag attgcgctgt tcagacttaa	300
tcttctcatc acagtccaag tctttatgaa aaggaaaaga gagagaagag cgctatttcg	360
agctgtcggc ctcatagggg gacagtcgag cataccagcg gtatcgacgt tagactcaac	420
caagaataat gacgagaata aacacagaag tcaaccttga actgtatatc agggttccag	480
cagcagatag ttacttgcatt aaagacaact ccccgagggc tctctgcata caccaggatg	540
ttccggaatt attcactgct cgtttccgac gtggcgtcag tgatccgtct ccacagaacc	600
tctacctggg gaataaccca ggggaggaat ctgcaagtaa gaacttaata ccaatccccg	660
gggctgccgg ggtgaatcaa atctcccgcg ggaaattaaa cccatacgat gtttttgcac	720
cacatgcatg cttggcacga tttctccgca agggagtcac agagaaagac atatttcgca	780
tactactgtg actctgcaga gttacatatt actcaggata cattgcagat cattgtccga	840
gcatacaaca tggacctgca ggatcaacgg cccgacaaaa cacaagtggc taaagctggg	900
ggatgcccgga acccgctgcg caatatcatt gatggatgtt ccccccatt tttaaaacat	960
cgacggatcg gcccgcatatc taatcctttt atcaacaaaa agttccactc gactagagaa	1020
aaaaaggcca agggcactaa ttgcagtcgg atactggtct tttcgccgtc caacaccttc	1080
atccatgatc cccttagcca ccaatgcccc acataatata tgttgacata ggtacgtagc	1140
tctgttatcc aatcgcatcc gaacctcttt aacggacccc tcctacacac cttatcctaa	1200
cttcaggaga ctgttgcccc ttggggattg aggaggtccg ggttgcagga tgcgttctag	1260
gctaaattct cggccggtag ccatctcgaa tctctcgtga agccttcatt tgaacggtg	1320
gcggcccgtc aagccgatga ccatgggttc ctgatagagc ttgtgcctga ccggccttgg	1380
cggcatagac gagctgaaca catcagggtat gaacagatca gatataaagt cggattgagt	1440
cctagtacga agcaatccgc caccacaaaa tcaagcaacg agcgacagca ataacaatat	1500
caatcgaatc gca atg tat cag cgc gct ctt ctc ttc ttc ttc ctc	1549
Met Tyr Gln Arg Ala Leu Leu Phe Ser Phe Phe Leu	
1 5 10	
gcc gcc gcc cgc gcg cag cag gcc ggt acc gta acc gca gag aat cac	1597
Ala Ala Ala Arg Ala Gln Gln Ala Gly Thr Val Thr Ala Glu Asn His	
15 20 25	
cct tcc ctg acc tgg cag caa tgc tcc agc ggc ggt agt tgt acc acg	1645
Pro Ser Leu Thr Trp Gln Gln Cys Ser Ser Gly Gly Ser Cys Thr Thr	
30 35 40	
cag aat gga aaa gtc gtt atc gat gcg aac tgg cgt tgg gtc cat acc	1693
Gln Asn Gly Lys Val Val Ile Asp Ala Asn Trp Arg Trp Val His Thr	
45 50 55 60	
acc tct gga tac acc aac tgc tac acg ggc aat acg tgg gac acc agt	1741
Thr Ser Gly Tyr Thr Asn Cys Tyr Thr Gly Asn Thr Trp Asp Thr Ser	
65 70 75	
atc tgt ccc gac gac gtg acc tgc gct cag aat tgt gcc ttg gat gga	1789

ES 2 527 681 T3

Ile	Cys	Pro	Asp	Asp	Val	Thr	Cys	Ala	Gln	Asn	Cys	Ala	Leu	Asp	Gly	
			80					85					90			
gcg	gat	tac	agt	ggc	acc	tat	ggt	gtt	acg	acc	agt	ggc	aac	gcc	ctg	1837
Ala	Asp	Tyr	Ser	Gly	Thr	Tyr	Gly	Val	Thr	Thr	Ser	Gly	Asn	Ala	Leu	
		95					100					105				
aga	ctg	aac	ttt	gtc	acc	caa	agc	tca	ggg	aag	aac	att	ggc	tcg	cgc	1885
Arg	Leu	Asn	Phe	Val	Thr	Gln	Ser	Ser	Gly	Lys	Asn	Ile	Gly	Ser	Arg	
	110					115					120					
ctg	tac	ctg	ctg	cag	gac	gac	acc	act	tat	cag	atc	ttc	aag	ctg	ctg	1933
Leu	Tyr	Leu	Leu	Gln	Asp	Asp	Thr	Thr	Tyr	Gln	Ile	Phe	Lys	Leu	Leu	
	125				130					135					140	
ggt	cag	gag	ttt	acc	ttc	gat	gtc	gac	gtc	tcc	aat	ctc	cct	tgc	ggg	1981
Gly	Gln	Glu	Phe		Phe	Asp	Val	Asp	Val	Ser	Asn	Leu	Pro	Cys	Gly	
				145					150					155		
ctg	aac	ggc	gcc	ctc	tac	ttt	gtg	gcc	atg	gac	gcc	gac	ggc	gga	ttg	2029
Leu	Asn	Gly	Ala	Leu	Tyr	Phe	Val	Ala	Met	Asp	Ala	Asp	Gly	Gly	Leu	
			160					165					170			
tcc	aaa	tac	cct	ggc	aac	aag	gca	ggc	gct	aag	tat	ggc	act	ggt	tac	2077
Ser	Lys	Tyr	Pro	Gly	Asn	Lys	Ala	Gly	Ala	Lys	Tyr	Gly	Thr	Gly	Tyr	
		175					180					185				
tgc	gac	tct	cag	tgc	cct	cgg	gat	ctc	aag	ttc	atc	aac	ggt	cag		2122
Cys	Asp	Ser	Gln	Cys	Pro	Arg	Asp	Leu	Lys	Phe	Ile	Asn	Gly	Gln		
	190					195					200					
gtacgtcaga agtgataact agccagcaga gcccatgaat cattaactaa cgctgtcaaa																2182
tacag	gcc	aat	gtt	gaa	ggc	tgg	cag	ccg	tct	gcc	aac	gac	cca	aat	gcc	2232
	Ala	Asn	Val	Glu	Gly	Trp	Gln	Pro	Ser	Ala	Asn	Asp	Pro	Asn	Ala	
		205					210					215				
ggc	gtt	ggt	aac	cac	ggt	tcc	tgc	tgc	gct	gag	atg	gat	gtc	tgg	gaa	2280
Gly	Val	Gly	Asn	His	Gly	Ser	Cys	Cys	Ala	Glu	Met	Asp	Val	Trp	Glu	
	220					225					230					
gcc	aac	agc	atc	tct	act	gcg	gtg	acg	cct	cac	cca	tgc	gac	acc	ccc	2328
Ala	Asn	Ser	Ile	Ser	Thr	Ala	Val	Thr	Pro	His	Pro	Cys	Asp	Thr	Pro	
	235				240					245					250	
ggc	cag	acc	atg	tgc	cag	gga	gac	gac	tgt	ggt	gga	acc	tac	tcc	tcc	2376
Gly	Gln	Thr	Met	Cys	Gln	Gly	Asp	Asp	Cys	Gly	Gly	Thr	Tyr	Ser	Ser	
				255					260					265		
act	cga	tat	gct	ggt	acc	tgc	gac	cct	gat	ggc	tgc	gac	ttc	aat	cct	2424
Thr	Arg	Tyr	Ala	Gly	Thr	Cys	Asp	Pro	Asp	Gly	Cys	Asp	Phe	Asn	Pro	
		270						275					280			
tac	cgc	cag	ggc	aac	cac	tcg	ttc	tac	ggc	ccc	ggg	cag	atc	gtc	gac	2472
Tyr	Arg	Gln	Gly	Asn	His	Ser	Phe	Tyr	Gly	Pro	Gly	Gln	Ile	Val	Asp	
		285					290					295				
acc	agc	tcc	aaa	ttc	acc	gtc	gtc	acc	cag	ttc	atc	acc	gac	gac	ggg	2520
Thr	Ser	Ser	Lys	Phe	Thr	Val	Val	Thr	Gln	Phe	Ile	Thr	Asp	Asp	Gly	
	300					305					310					
acc	ccc	tcc	ggc	acc	ctg	acg	gag	atc	aaa	cgc	ttc	tac	gtc	cag	aac	2568
Thr	Pro	Ser	Gly	Thr	Leu	Thr	Glu	Ile	Lys	Arg	Phe	Tyr	Val	Gln	Asn	
	315				320					325					330	
ggc	aag	gta	atc	ccc	cag	tcg	gag	tcg	acg	atc	agc	ggc	gtc	acc	ggc	2616
Gly	Lys	Val	Ile	Pro	Gln	Ser	Glu	Ser	Thr	Ile	Ser	Gly	Val	Thr	Gly	
				335					340					345		

aac tca atc acc acc gag tat tgc acg gcc cag aag gcc gcc ttc ggc 2664
 Asn Ser Ile Thr Thr Glu Tyr Cys Thr Ala Gln Lys Ala Ala Phe Gly
 350 355 360

gac aac acc ggc ttc ttc acg cac ggc ggg ctt cag aag atc agt cag 2712
 Asp Asn Thr Gly Phe Phe Thr His Gly Gly Leu Gln Lys Ile Ser Gln
 365 370 375

gct ctg gct cag ggc atg gtc ctc gtc atg agc ctg tgg gac gat cac 2760
 Ala Leu Ala Gln Gly Met Val Leu Val Met Ser Leu Trp Asp Asp His
 380 385 390

gcc gcc aac atg ctc tgg ctg gac agc acc tac ccg act gat gcg gac 2808
 Ala Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asp Ala Asp
 395 400 405 410

ccg gac acc cct ggc gtc gcg cgc ggt acc tgc ccc acg acc tcc ggc 2856
 Pro Asp Thr Pro Gly Val Ala Arg Gly Thr Cys Pro Thr Thr Ser Gly
 415 420 425

gtc ccg gcc gac gtt gag tcg cag tac ccc aat tca tat gtt atc tac 2904
 Val Pro Ala Asp Val Glu Ser Gln Tyr Pro Asn Ser Tyr Val Ile Tyr
 430 435 440

tcc aac atc aag gtc gga ccc atc aac tcg acc ttc acc gcc aac 2949
 Ser Asn Ile Lys Val Gly Pro Ile Asn Ser Thr Phe Thr Ala Asn
 445 450 455

taagtaagta actggcactc taccaccgag agcttcgtga agatacaggg gtggttggga 3009

gattgtcgtg tacaggggac atgcgatgct caaaaatcta catcagtttg ccaattgaac 3069

catgaaaaaa agggggagat caaagaagtc tgtcaaaaga ggggggctgt ggcagcttaa 3129

gccttggtgt agatcgagtc gacgccctat agtgagtcgt attagagctc gcggccgcga 3189

gct 3192

<210> 2
 <211> 457
 <212> PRT
 <213> Thermoascus aurantiacus

<400> 2
 Met Tyr Gln Arg Ala Leu Leu Phe Ser Phe Phe Leu Ala Ala Ala Arg
 1 5 10 15

Ala Gln Gln Ala Gly Thr Val Thr Ala Glu Asn His Pro Ser Leu Thr
 20 25 30

Trp Gln Gln Cys Ser Ser Gly Gly Ser Cys Thr Thr Gln Asn Gly Lys
 35 40 45

Val Val Ile Asp Ala Asn Trp Arg Trp Val His Thr Thr Ser Gly Tyr
 50 55 60

5

ES 2 527 681 T3

Thr Asn Cys Tyr Thr Gly Asn Thr Trp Asp Thr Ser Ile Cys Pro Asp
 65 70 75 80
 Asp Val Thr Cys Ala Gln Asn Cys Ala Leu Asp Gly Ala Asp Tyr Ser
 85 90 95
 Gly Thr Tyr Gly Val Thr Thr Ser Gly Asn Ala Leu Arg Leu Asn Phe
 100 105 110
 Val Thr Gln Ser Ser Gly Lys Asn Ile Gly Ser Arg Leu Tyr Leu Leu
 115 120 125
 Gln Asp Asp Thr Thr Tyr Gln Ile Phe Lys Leu Leu Gly Gln Glu Phe
 130 135 140
 Thr Phe Asp Val Asp Val Ser Asn Leu Pro Cys Gly Leu Asn Gly Ala
 145 150 155 160
 Leu Tyr Phe Val Ala Met Asp Ala Asp Gly Gly Leu Ser Lys Tyr Pro
 165 170 175
 Gly Asn Lys Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln
 180 185 190
 Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly
 195 200 205
 Trp Gln Pro Ser Ala Asn Asp Pro Asn Ala Gly Val Gly Asn His Gly
 210 215 220
 Ser Cys Cys Ala Glu Met Asp Val Trp Glu Ala Asn Ser Ile Ser Thr
 225 230 235 240
 Ala Val Thr Pro His Pro Cys Asp Thr Pro Gly Gln Thr Met Cys Gln
 245 250 255
 Gly Asp Asp Cys Gly Gly Thr Tyr Ser Ser Thr Arg Tyr Ala Gly Thr
 260 265 270
 Cys Asp Pro Asp Gly Cys Asp Phe Asn Pro Tyr Arg Gln Gly Asn His
 275 280 285
 Ser Phe Tyr Gly Pro Gly Gln Ile Val Asp Thr Ser Ser Lys Phe Thr
 290 295 300
 Val Val Thr Gln Phe Ile Thr Asp Asp Gly Thr Pro Ser Gly Thr Leu
 305 310 315 320
 Thr Glu Ile Lys Arg Phe Tyr Val Gln Asn Gly Lys Val Ile Pro Gln
 325 330 335

ES 2 527 681 T3

Ser Glu Ser Thr Ile Ser Gly Val Thr Gly Asn Ser Ile Thr Thr Glu
340 345 350

Tyr Cys Thr Ala Gln Lys Ala Ala Phe Gly Asp Asn Thr Gly Phe Phe
355 360 365

Thr His Gly Gly Leu Gln Lys Ile Ser Gln Ala Leu Ala Gln Gly Met
370 375 380

Val Leu Val Met Ser Leu Trp Asp Asp His Ala Ala Asn Met Leu Trp
385 390 400

Leu Asp Ser Thr Tyr Pro Thr Asp Ala Asp Pro Asp Thr Pro Gly Val
405 410 415

Ala Arg Gly Thr Cys Pro Thr Thr Ser Gly Val Pro Ala Asp Val Glu
420 425 430

Ser Gln Tyr Pro Asn Ser Tyr Val Ile Tyr Ser Asn Ile Lys Val Gly
435 440 445

Pro Ile Asn Ser Thr Phe Thr Ala Asn
450 455

<210> 3
<211> 3055
5 <212> ADN
<213> Acremonium thermophilum

<220>
<221> CDS
10 <222> (972)..(1595)
<223>

<220>
<221> Intrón
15 <222> (1596)..(1729)
<223>

<220>
<221> CDS
20 <222> (1730)..(2290)
<223>

<220>
<221> Intrón
25 <222> (2291)..(2412)
<223>

<220>
<221> CDS
30 <222> (2413)..(2540)
<223>

<220>
<221> Intrón
35 <222> (2541)..(2627)
<223>

<220>
<221> CDS
40 <222> (2628)..(2691)
<223>

<400> 3

gaattcggat cacaccgaga gcttcgcat ggccagctgt ctcagcttgt acccgtctac	60
caacgttccg catcttcgtt accttgatag ctgcggtttg ctggactgct ttgtgagggg	120
actgtgccac gcctgggaga cgggtgccgt accatcggtt actgcgcaga ctgagaaccg	180
tcgttgccga aacagccagg caggaagcct gtccaccttc atgtatcttc atatggaccc	240
cagcgcgcc ctctctttct cctcatttct tgcccaccac gatggacacc atgccaatct	300
atttcttgat cccttgactc ctcagcccc cagcagtccg acaatgtaca gtgatgggca	360
tctctttctg tacatacgtc ccctctcgcg gtgtccacgc gcggccgggg atgcctggga	420
cggagtgcc cccgcagggg acgagacttg gctgatgggg tgcggtgcat ggtggcacia	480
gagatccagg ccccccgatc tcgttctcgc acgtatcctt ccccgccgg cgatgccaa	540
gtgggaagtc ttcggagcgg caccagggc catcttgccg atgccggca cggctctggc	600
ggttgccctc atctatcgtg gctgcacatc cgccgtgccc ccattgggaa agcaggcttt	660
gttcttcccg tctgtcgatc gtctcccacc taccctccct cctcgcaagg gcttaccctg	720

ES 2 527 681 T3

gccccctcact	gctgtcttcac	ctcactgctg	cttcccgcga	atgccccctc	gccccccccc		780
ccccccctctc	ctttgcagta	cagatctaca	taatatcgag	acgcccccca	agctgtttct		840
ctggcacagc	cctctcgcgc	gtggtgcaag	agcaagtcat	agtatcaatt	cccccatctc		900
tcattctcagc	ccttctgccc	tgggtccacc	gacattcttg	gcccgtagcc	aagaccgatc		960
cgctcttcac	c atg cac aag cgg gcg gcc acc ctc tcc gcc ctc gtc gtc Met His Lys Arg Ala Ala Thr Leu Ser Ala Leu Val Val 1 5 10						1010
gcc ala gcc ala gcc ala ggc gly ttc phe gcc ala cgc arg ggc gly cag gln ggc gly gtg val ggc gly acg thr cag gln acg thr	15	20	25				1058
gag glu acg thr cac his ccc pro aag lys ctc leu acc thr ttc phe cag gln aag lys tgc cys tcc ser gcc ala gcc ala ggc gly agc ser	30	35	40				1106
tgc cys acg thr acc thr cag gln aac asn ggc gly gag glu gtg val gtc val atc ile gac asp gcc ala aac asn tgg trp cgc arg tgg trp	50	55	60				1154
gtg val cac his gac asp aag lys aac asn ggc gly tac tyr acc thr aac asn tgc cys tac tyr acg thr ggc gly aac asn gag glu tgg trp	65	70	75				1202
aac asn acc thr acc thr atc ile tgc cys gcc ala gac asp gcc ala gcc ala tgc cys gcc ala agc ser aac asn tgc cys gtc val	80	85	90				1250
gtc val gac asp ggc gly gcc ala gac asp tac tyr cag gln ggc gly acc thr tac tyr ggc gly gcc ala tcc ser acc thr tcc ser ggc gly	95	100	105				1298
aac asn gcc ala ctg leu acc thr ctc leu aag lys ttc phe gtc val acc thr aag lys ggc gly agc ser tac thr gcc ala acc thr aac asn	110	115	120				1346
atc ile ggc gly tgc ser cgc arg atg met tac tyr ctg leu atg met gcc ala agc ser ccc pro acc thr aag lys tac thr gcc ala atg met	130	135	140				1394
ttc phe acc thr ctg leu ctg leu ggc gly cac his gag glu ttc phe gcc ala ttc phe gac asp gtc val gac asp ctg leu agc ser aag lys	145	150	155				1442
ctg leu ccc pro tgc cys ggc gly ctc leu aac asn ggc gly gcc ala gtc val tac tyr ttc phe gtc val agc ser atg met gac asp gag glu	160	165	170				1490
gac asp ggc gly ggc gly acc thr agc ser aag lys tac tyr ccc pro tcc ser aac asn aag lys gcc ala ggc gly gcc ala aag lys tac tyr	175	180	185				1538
ggc gly acg thr ggc gly tac tyr tgc cys gac asp tcg ser cag gln tgt cys ccg pro cgc arg gac asp ctc leu aag lys ttt phe atc ile	190	195	200				1586
gac asp ggc gly aag lys gtgagaacct gcactagcgt cccgccttcc gtgtccctcc							1635
ttttgccttc	ttcgaccgcc	ctcttccctg	cgaggccagg	tcgctggggg	gctgtccctc		1695
ttttctggtgg	gcagcggtgc	tgatccgcgc	ccaq gcc	aac tcg gcc	agc tgg cag		1750

ES 2 527 681 T3

																Ala	Asn	Ser	Ala	Ser	Trp	Gln	
																210							215
ccc	tgc	tcc	aac	gac	cag	aac	gcc	ggc	gtg	ggc	ggc	atg	ggc	tgc	tgc	1798							
Pro	Ser	Ser	Asn	Asp	Gln	Asn	Ala	Gly	Val	Gly	Gly	Met	Gly	Ser	Cys								
				220												230							
tgc	gcc	gag	atg	gac	atc	tgg	gag	gcc	aac	tcc	gtc	tcc	gcc	gcc	tac	1846							
Cys	Ala	Glu	Met	Asp	Ile	Trp	Glu	Ala	Asn	Ser	Val	Ser	Ala	Ala	Tyr								
				235												245							
acg	ccg	cac	ccg	tgc	cag	aac	tac	cag	cag	cac	agc	tgc	agc	ggc	gac	1894							
Thr	Pro	His	Pro	Cys	Gln	Asn	Tyr	Gln	Gln	His	Ser	Cys	Ser	Gly	Asp								
				250												260							
gac	tgc	ggc	ggc	acc	tac	tgc	gcc	acc	cgc	ttc	gcc	ggc	gac	tgc	gac	1942							
Asp	Cys	Gly	Gly	Thr	Tyr	Ser	Ala	Thr	Arg	Phe	Ala	Gly	Asp	Cys	Asp								
				265												275							
ccg	gac	ggc	tgc	gac	tgg	aac	gcc	tac	cgc	atg	ggc	gtg	cac	gac	ttc	1990							
Pro	Asp	Gly	Cys	Asp	Trp	Asn	Ala	Tyr	Arg	Met	Gly	Val	His	Asp	Phe								
				280												285	290	295					
tac	ggc	aac	ggc	aag	acc	gtc	gac	acc	ggc	aag	aag	ttc	tcc	atc	gtc	2038							
Tyr	Gly	Asn	Gly	Lys	Thr	Val	Asp	Thr	Gly	Lys	Lys	Phe	Ser	Ile	Val								
				300												305	310						
acc	cag	ttc	aag	ggc	tcc	ggc	tcc	acc	ctg	acc	gag	atc	aag	cag	ttc	2086							
Thr	Gln	Phe	Lys	Gly	Ser	Gly	Ser	Thr	Leu	Thr	Glu	Ile	Lys	Gln	Phe								
				315												320	325						
tac	gtc	cag	gac	ggc	agg	aag	atc	gag	aac	ccc	aac	gcc	acc	tgg	ccc	2134							
Tyr	Val	Gln	Asp	Gly	Arg	Lys	Ile	Glu	Asn	Pro	Asn	Ala	Thr	Trp	Pro								
				330												340							
ggc	ctc	gag	ccc	ttc	aac	tcc	atc	acc	ccg	gac	ttc	tgc	aag	gcc	cag	2182							
Gly	Leu	Glu	Pro	Phe	Asn	Ser	Ile	Thr	Pro	Asp	Phe	Cys	Lys	Ala	Gln								
				345												355							
aag	cag	gtc	ttc	ggc	gac	ccc	gac	cgc	ttc	aac	gac	atg	ggc	ggc	ttc	2230							
Lys	Gln	Val	Phe	Gly	Asp	Pro	Asp	Arg	Phe	Asn	Asp	Met	Gly	Gly	Phe								
				360												365	370	375					
acc	aac	atg	gcc	aag	gcc	ctg	gcc	aac	ccc	atg	gtc	ctg	gtg	ctg	tgc	2278							
Thr	Asn	Met	Ala	Lys	Ala	Leu	Ala	Asn	Pro	Met	Val	Leu	Val	Leu	Ser								
				380												385	390						
ctg	tgg	gac	gac	gtgagccatt	ttcgccattct	ctcctgactc	tcctccgctg									2330							
Leu	Trp	Asp	Asp																				
				395																			
ccatcaccac ctcttccacc accgccacga ggggtgtagct tgatctccgc tgactgacgt																2390							
gtgcccacac ccccgtttct ag cac tac tcc aac atg ctg tgg ctc gac tct																2442							
				His Tyr Ser Asn Met Leu Trp Leu Asp Ser																			
				400 405																			
acc	tac	ccg	acc	gac	gcc	gat	ccc	agc	gcg	ccc	ggc	aag	gga	cgt	ggc	2490							
Thr	Tyr	Pro	Thr	Asp	Ala	Asp	Pro	Ser	Ala	Pro	Gly	Lys	Gly	Arg	Gly								
				410												415	420						
acc	tgc	gac	acc	agc	agc	ggc	gtg	cca	agc	gac	gtg	gag	tgc	aag	aat	2538							
Thr	Cys	Asp	Thr	Ser	Ser	Gly	Val	Pro	Ser	Asp	Val	Glu	Ser	Lys	Asn								
				425												435							
gg	gtgagtcgga	tcttctgcat	gcggcccggt	ttccgagcat	tgcttgggggt											2590							
Gly																							

cctccctcag gctgacacac gcgcgccttc gatacag c gat gcg acc gtc atc 2643
 Asp Ala Thr Val Ile
 440

tac tcc aac atc aag ttt ggg ccg ctg gac tcc acc tac acg gct tcc 2691
 Tyr Ser Asn Ile Lys Phe Gly Pro Leu Asp Ser Thr Tyr Thr Ala Ser
 445 450 455

tgagcagccg ctttgggttc ggtggggccg aagcacaaca agtgtgtgcg tagctgagat 2751

gatggccgat ctctgtcctt tgtctcctag tgtctctctt atcgaacaac cccccgacct 2811

gcagcgtcgg cgggcatcgt atagtctggt gtaactgtat atagctctgt gcgtgtgaat 2871

cgaacgagca ccgacgaaat gtggtgtttc atgctatcgt acatgctctt gcgagatctg 2931

aagtcgtcaa ttagacattg ccaccatcca acttggcgac tgtccacccg gtccatttgt 2991

atcactggct cttccgagac ccggtctctc tcacaccgta atcactgcaa gcagagtga 3051

attc 3055

<210> 4

<211> 459

<212> PRT

<213> Acremonium thermophilum

<400> 4

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

Gly Phe Ala Arg Gly Gln Gly Val Gly Thr Gln Gln Thr Glu Thr His
 20 25 30

Pro Lys Leu Thr Phe Gln Lys Cys Ser Ala Ala Gly Ser Cys Thr Thr
 35 40 45

Gln Asn Gly Glu Val Val Ile Asp Ala Asn Trp Arg Trp Val His Asp
 50 55 60

Lys Asn Gly Tyr Thr Asn Cys Tyr Thr Gly Asn Glu Trp Asn Thr Thr
 65 70 75 80

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

Ala Asp Tyr Gln Gly Thr Tyr Gly Ala Ser Thr Ser Gly Asn Ala Leu
 100 105 110

Thr Leu Lys Phe Val Thr Lys Gly Ser Tyr Ala Thr Asn Ile Gly Ser
 115 120 125

ES 2 527 681 T3

Arg Met Tyr Leu Met Ala Ser Pro Thr Lys Tyr Ala Met Phe Thr Leu
 130 135 140
 Leu Gly His Glu Phe Ala Phe Asp Val Asp Leu Ser Lys Leu Pro Cys
 145 150 155 160
 Gly Leu Asn Gly Ala Val Tyr Phe Val Ser Met Asp Glu Asp Gly Gly
 165 170 175
 Thr Ser Lys Tyr Pro Ser Asn Lys Ala Gly Ala Lys Tyr Gly Thr Gly
 180 185 190
 Tyr Cys Asp Ser Gln Cys Pro Arg Asp Leu Lys Phe Ile Asp Gly Lys
 195 200 205
 Ala Asn Ser Ala Ser Trp Gln Pro Ser Ser Asn Asp Gln Asn Ala Gly
 210 215 220
 Val Gly Gly Met Gly Ser Cys Cys Ala Glu Met Asp Ile Trp Glu Ala
 225 230 235 240
 Asn Ser Val Ser Ala Ala Tyr Thr Pro His Pro Cys Gln Asn Tyr Gln
 245 250 255
 Gln His Ser Cys Ser Gly Asp Asp Cys Gly Gly Thr Tyr Ser Ala Thr
 260 265 270
 Arg Phe Ala Gly Asp Cys Asp Pro Asp Gly Cys Asp Trp Asn Ala Tyr
 275 280 285
 Arg Met Gly Val His Asp Phe Tyr Gly Asn Gly Lys Thr Val Asp Thr
 290 295 300
 Gly Lys Lys Phe Ser Ile Val Thr Gln Phe Lys Gly Ser Gly Ser Thr
 305 310 315 320
 Leu Thr Glu Ile Lys Gln Phe Tyr Val Gln Asp Gly Arg Lys Ile Glu
 325 330 335
 Asn Pro Asn Ala Thr Trp Pro Gly Leu Glu Pro Phe Asn Ser Ile Thr
 340 345 350
 Pro Asp Phe Cys Lys Ala Gln Lys Gln Val Phe Gly Asp Pro Asp Arg
 355 360 365
 Phe Asn Asp Met Gly Gly Phe Thr Asn Met Ala Lys Ala Leu Ala Asn
 370 375 380
 Pro Met Val Leu Val Leu Ser Leu Trp Asp Asp His Tyr Ser Asn Met
 385 390 395 400
 Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asp Ala Asp Pro Ser Ala Pro
 405 410 415
 Gly Lys Gly Arg Gly Thr Cys Asp Thr Ser Ser Gly Val Pro Ser Asp
 420 425 430
 Val Glu Ser Lys Asn Gly Asp Ala Thr Val Ile Tyr Ser Asn Ile Lys
 435 440 445
 Phe Gly Pro Leu Asp Ser Thr Tyr Thr Ala Ser
 450 455

<210> 5
 <211> 3401
 <212> ADN
 <213> Acremonium thermophilum
 5
 <220>
 <221> CDS
 <222> (891)..(1299)
 <223>
 10
 <220>
 <221> Intrón
 <222> (1300)..(1387)
 <223>
 15
 <220>
 <221> CDS
 <222> (1388)..(1442)
 <223>
 20
 <220>
 <221> Intrón
 <222> (1443)..(1495)
 <223>
 25
 <220>
 <221> CDS
 <222> (1496)..(1643)
 <223>
 30
 <220>
 <221> Intrón
 <222> (1644)..(1697)
 <223>
 35
 <220>
 <221> CDS
 <222> (1698)..(1928)
 <223>
 40
 <220>
 <221> Intrón
 <222> (1929)..(2014)
 <223>
 45
 <220>
 <221> CDS
 <222> (2015)..(2740)
 <223>
 50
 <400> 5
 ctcgagtttc cctgggtcggc cactctctgc tcattctcgt ctgcgcccct ggatgtgccg 60
 tgtgtccagt cgtgtatctc ttgactgcac gacgtgttcc tcgcgactcg tctcgcgccg 120
 gtggatgccc gtccactcat ttgtccgtct actgggtcag cctctcgtct cgaacgagct 180
 tccacggccc actccccgga caacctcggc tctggatggc cctcctcccc ctccgtgtct 240
 cccctctcgc ggggtccgct tgccctggc tgcattgctc acatcgcttg atcacgctgc 300
 gagccaccgc agagccccat ctccaaagcg accgtggcag cactacctct gtttctggga 360
 tggggccccc gtcgatggcc tggcatccct tgccaccctc ctccatcccc ctgacctcac 420
 tcccaaccga taggagaagt ggtcatgggc acgaccccg gcacgtcttg gactcgacga 480
 gcttgatcgg gccggaagcc gtcaacgacg ggggagccgt gtcttgccac gcgtggccgt 540
 ccttcgacag tggacagcga gaaaactggt ggggaagagg gctgctacag tcttgtcttg 600

cgaggccccga cgctcctagt ccgagaacca cctacgtgtt tctcgcgaag acggggccag	660
cttagcggcc aaatttgccc cccgggccta ggggtctagcg atggggatga tgaactgggtg	720
tcgacgatgt ctatataacg acggcgatct cctgtctctg agatcccatc ctttcatctc	780
caaccactt catcccttcc tctctctctc cccctccctt ctctgacata ccgagtcctc	840
agaagcctcg tccgtcgtca cctattctca cttccccgcg aactccggcc atg tat Met Tyr 1	896
acc aag ttc gcc gcc ctc gcc gcc ctc gtg gcc acc gtc cgc ggc cag Thr Lys Phe Ala Ala Leu Ala Ala Leu Val Ala Thr Val Arg Gly Gln 5 10 15	944
gcc gcc tgc tcg ctc acc gcc gag acc cac ccg tcg ctg cag tgg cag Ala Ala Cys Ser Leu Thr Ala Glu Thr His Pro Ser Leu Gln Trp Gln 20 25 30	992
aag tgc acc gcg ccc ggc agc tgc acc acc gtc agc ggc cag gtc acc Lys Cys Thr Ala Pro Gly Ser Cys Thr Thr Val Ser Gly Gln Val Thr 35 40 45 50	1040
atc gac gcc aac tgg cgc tgg ctg cac cag acc aac agc agc acc aac Ile Asp Ala Asn Trp Arg Trp Leu His Gln Thr Asn Ser Ser Thr Asn 55 60 65	1088
tgc tac acc ggc aac gag tgg gac acc agc atc tgc agc tcc gac acc Cys Tyr Thr Gly Asn Glu Trp Asp Thr Ser Ile Cys Ser Ser Asp Thr 70 75 80	1136
gac tgc gcc acc aag tgc tgc ctc gac ggc gcc gac tac acc ggc acc Asp Cys Ala Thr Lys Cys Cys Leu Asp Gly Ala Asp Tyr Thr Gly Thr 85 90 95	1184
tac ggc gtc acc gcc agc ggc aac tcg ctc aac ctc aag ttc gtc acc Tyr Gly Val Thr Ala Ser Gly Asn Ser Leu Asn Leu Lys Phe Val Thr 100 105 110	1232
cag ggg ccc tac tcc aag aac atc ggc tcg cgc atg tac ctc atg gag Gln Gly Pro Tyr Ser Lys Asn Ile Gly Ser Arg Met Tyr Leu Met Glu 115 120 125 130	1280
tcg gag tcc aag tac cag g gtagcatat agatcacatc tttcgtcact Ser Glu Ser Lys Tyr Gln 135	1329
tgcgtccggt tcgcacggca agcgggtccag acgctaacgg gacgggtctc ttctctag	1387
gc ttc act ctc ctc ggt cag gag ttt acc ttt gac gtg gac gtc tcc Gly Phe Thr Leu Leu Gly Gln Glu Phe Thr Phe Asp Val Asp Val Ser 140 145 150	1434
aac ctc gg gtaggtgatg acttctcccg catgagaaga gctctgctaa Asn Leu Gly	1482
ccgtgttggtc cag c tgc ggt ctg aac gga gcg ctc tac ttc gtg tcc atg Cys Gly Leu Asn Gly Ala Leu Tyr Phe Val Ser Met 160 165	1532
gac ctc gac ggc ggc gtg tcc aag tac acc acc aac aag gcc ggc gcc Asp Leu Asp Gly Gly Val Ser Lys Tyr Thr Thr Asn Lys Ala Gly Ala 170 175 180	1580
aag tac ggc acc ggc tac tgc gac tcc cag tgc ccg cgg gat ctc aag	1628

ES 2 527 681 T3

Lys	Tyr	Gly	Thr	Gly	Tyr	Cys	Asp	Ser	Gln	Cys	Pro	Arg	Asp	Leu	Lys	
185						190					195					
ttc	atc	aac	ggc	cag	gtgggctcgag	agaccctctt	cccctctcag	tgaacgatgt								1683
Phe	Ile	Asn	Gly	Gln												
200																
ctg	acc	ctt	ctag	gac	aac	atc	gac	ggc	tgg	caa	ccg	tcg	tcc	aac	gac	1733
				Ala	Asn	Ile	Asp	Gly	Trp	Gln	Pro	Ser	Ser	Asn	Asp	
				205					210					215		
gcc	aac	ggc	ggc	ctc	ggg	aac	cac	ggc	agc	tgc	tgc	tcc	gag	atg	gac	1781
Ala	Asn	Ala	Gly	Leu	Gly	Asn	His	Gly	Ser	Cys	Cys	Ser	Glu	Met	Asp	
			220					225					230			
atc	tgg	gag	ggc	aac	aag	gtc	tcc	ggc	ggc	tac	acg	ccg	cac	ccc	tgc	1829
Ile	Trp	Glu	Ala	Asn	Lys	Val		Ala	Ala	Tyr	Thr	Pro	His	Pro	Cys	
		235					240					245				
acc	acc	atc	ggc	cag	acc	atg	tgc	acc	ggc	gac	gac	tgc	ggc	ggc	acc	1877
Thr	Thr	Ile	Gly	Gln	Thr	Met	Cys	Thr	Gly	Asp	Asp	Cys	Gly	Gly	Thr	
		250				255					260					
tat	tcg	tcg	gac	cgc	tat	gcc	ggc	atc	tgc	gac	ccc	gac	ggg	tgc	gat	1925
Tyr	Ser	Ser	Asp	Arg	Tyr	Ala	Gly	Ile	Cys	Asp	Pro	Asp	Gly	Cys	Asp	
265					270				275					280		
ttt	gtaggttctt	tctctcgccg	ctccctgacg	acctatatgt	gtgaaggac											1978
Phe																
gcacagaaaa	gacaagggtca	aagctgacca	gagcag	aac	tcg	tac	cgc	atg	ggc							2032
				Asn	Ser	Tyr	Arg	Met	Gly							
							285									
gac	acc	agc	ttc	tac	ggc	ccc	ggc	aag	acg	gtc	gac	acc	ggc	tcc	aag	2080
Asp	Thr	Ser	Phe	Tyr	Gly	Pro	Gly	Lys	Thr	Val	Asp	Thr	Gly	Ser	Lys	
		290					295					300				
ttc	acc	gtc	gtg	acc	cag	ttc	ctc	acg	ggc	tcc	gac	ggc	aac	ctc	agc	2128
Phe	Thr	Val	Val	Thr	Gln	Phe	Leu	Thr	Gly	Ser	Asp	Gly	Asn	Leu	Ser	
		305				310					315					
gag	atc	aag	cgc	ttc	tac	gtg	cag	aac	ggc	aag	gtc	atc	ccc	aac	tcc	2176
Glu	Ile	Lys	Arg	Phe	Tyr	Val	Gln	Asn	Gly	Lys	Val	Ile	Pro	Asn	Ser	
320					325					330				335		
gag	tcc	aag	atc	ggc	ggc	gtc	tcc	ggc	aac	tcc	atc	acc	acc	gac	ttc	2224
Glu	Ser	Lys	Ile	Ala	Gly	Val	Ser	Gly	Asn	Ser	Ile	Thr	Thr	Asp	Phe	
				340				345						350		
tgc	acc	ggc	cag	aag	acc	ggc	ttc	ggc	gac	acc	aac	gtc	ttc	gag	gag	2272
Cys	Thr	Ala	Gln	Lys	Thr	Ala	Phe	Gly	Asp	Thr	Asn	Val	Phe	Glu	Glu	
			355					360					365			
cgc	ggc	ggc	ctc	ggc	cag	atg	ggc	aag	ggc	ctg	ggc	gag	ccc	atg	gtc	2320
Arg	Gly	Gly	Leu	Ala	Gln	Met	Gly	Lys	Ala	Leu	Ala	Glu	Pro	Met	Val	
		370					375					380				
ctg	gtc	ctg	tcc	gtc	tgg	gac	gac	cac	ggc	gtc	aac	atg	ctc	tgg	ctc	2368
Leu	Val	Leu	Ser	Val	Trp	Asp	Asp	His	Ala	Val	Asn	Met	Leu	Trp	Leu	
		385				390					395					
gac	tcc	acc	tac	ccc	acc	gac	agc	acc	aag	ccc	ggc	ggc	ggc	cgc	ggc	2416
Asp	Ser	Thr	Tyr	Pro	Thr	Asp	Ser	Thr	Lys	Pro	Gly	Ala	Ala	Arg	Gly	
400					405					410				415		
gac	tgc	ccc	atc	acc	tcc	ggc	gtg	ccc	ggc	gac	gtc	gag	tcc	cag	gcg	2464

Asp Cys Pro Ile Thr Ser Gly Val Pro Ala Asp Val Glu Ser Gln Ala	
420 425 430	
ccc aac tcc aac gtc atc tac tcc aac atc cgc ttc ggc ccc atc aac	2512
Pro Asn Ser Asn Val Ile Tyr Ser Asn Ile Arg Phe Gly Pro Ile Asn	
435 440 445	
tcc acc tac acc ggc acc ccc agc ggc ggc aac ccc ccc ggc ggc ggc	2560
Ser Thr Tyr Thr Gly Thr Pro Ser Gly Gly Asn Pro Pro Gly Gly Gly	
450 455 460	
acc acc acc acc acc acc acc acc acc tcc aag ccc tcc ggc ccc acc	2608
Thr Thr Thr Thr Thr Thr Thr Thr Thr Ser Lys Pro Ser Gly Pro Thr	
465 470 475	
acc acc acc aac ccc tcc ggt ccg cag cag acg cac tgg ggt cag tgc	2656
Thr Thr Thr Asn Pro Ser Gly Pro Gln Gln Thr His Trp Gly Gln Cys	
480 485 490 495	
ggc ggc cag gga tgg acc ggc ccc acg gtc tgc cag agc ccc tac acc	2704
Gly Gly Gln Gly Trp Thr Gly Pro Thr Val Cys Gln Ser Pro Tyr Thr	
500 505 510	
tgc aag tac tcc aac gac tgg tac tcc cag tgc ctg taagccataa	2750
Cys Lys Tyr Ser Asn Asp Trp Tyr Ser Gln Cys Leu	
515 520	
gccccctgta cgttcgggaag acggtggcaa cagacaaacc cctccccga gcacaccccc	2810
cagggatcta aggggggttgt ggtaagaca taagaatgcg ccgtggcttg gcctacgcca	2870
cggtcatgaa agtgcagtga aaatgggggc aagagtcgga aaaagtgagt ttgcttgcaa	2930
gggagagagg atgtcagagag gtgatgactt cgtttgtaca tagttggctc ttcgtgattg	2990
ggaacgggag gagtgtcggg gggagccctc cagactcctt ggccctctccg ctcgttccat	3050
ctttctcagt acatatacat ctgcattttc atccacgtct ctggcgtctc tggatgtgaa	3110
cgaatccgac aactgggtggg ctgagatgaa tcgcaaggag agtatcttgc gaggatatca	3170
cagtcagaaa gtagcatttg agccactact aaaagggtcaa ccagtatgcg aagcttagca	3230
attatataca gcagctcaac ttcagaacga agtattgcat gtggcagaga atcttgggaa	3290
atgagccatg aagacctcgt cgagagagta cctctcaccg ccaaataacc agctagcggg	3350
ttgggagagg agcaatagga cgagcgcgat ggacagatat acgaactcga g	3401

<210> 6

<211> 523

<212> PRT

<213> Acremonium thermophilum

<400> 6

Met Tyr Thr Lys Phe Ala Ala Leu Ala Ala Leu Val Ala Thr Val Arg
1 5 10 15

Gly Gln Ala Ala Cys Ser Leu Thr Ala Glu Thr His Pro Ser Leu Gln
20 25 30

ES 2 527 681 T3

Trp Gln Lys Cys Thr Ala Pro Gly Ser Cys Thr Thr Val Ser Gly Gln
 35 40 45
 Val Thr Ile Asp Ala Asn Trp Arg Trp Leu His Gln Thr Asn Ser Ser
 50 55 60
 Thr Asn Cys Tyr Thr Gly Asn Glu Trp Asp Thr Ser Ile Cys Ser Ser
 65 70 75 80
 Asp Thr Asp Cys Ala Thr Lys Cys Cys Leu Asp Gly Ala Asp Tyr Thr
 85 90 95
 Gly Thr Tyr Gly Val Thr Ala Ser Gly Asn Ser Leu Asn Leu Lys Phe
 100 105 110
 Val Thr Gln Gly Pro Tyr Ser Lys Asn Ile Gly Ser Arg Met Tyr Leu
 115 120 125
 Met Glu Ser Glu Ser Lys Tyr Gln Gly Phe Thr Leu Leu Gly Gln Glu
 130 135 140
 Phe Thr Phe Asp Val Asp Val Ser Asn Leu Gly Cys Gly Leu Asn Gly
 145 150 155 160
 Ala Leu Tyr Phe Val Ser Met Asp Leu Asp Gly Gly Val Ser Lys Tyr
 165 170 175
 Thr Thr Asn Lys Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser
 180 185 190
 Gln Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Ile Asp
 195 200 205
 Gly Trp Gln Pro Ser Ser Asn Asp Ala Asn Ala Gly Leu Gly Asn His
 210 215 220
 Gly Ser Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Lys Val Ser
 225 230 235 240
 Ala Ala Tyr Thr Pro His Pro Cys Thr Thr Ile Gly Gln Thr Met Cys
 245 250 255
 Thr Gly Asp Asp Cys Gly Gly Thr Tyr Ser Ser Asp Arg Tyr Ala Gly
 260 265 270
 Ile Cys Asp Pro Asp Gly Cys Asp Phe Asn Ser Tyr Arg Met Gly Asp
 275 280 285
 Thr Ser Phe Tyr Gly Pro Gly Lys Thr Val Asp Thr Gly Ser Lys Phe
 290 295 300

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

Ile Lys Arg Phe Tyr Val Gln Asn Gly Lys Val Ile Pro Asn Ser Glu
325 330 335

Ser Lys Ile Ala Gly Val Ser Gly Asn Ser Ile Thr Thr Asp Phe Cys
340 345 350

Thr Ala Gln Lys Thr Ala Phe Gly Asp Thr Asn Val Phe Glu Glu Arg
355 360 365

Gly Gly Leu Ala Gln Met Gly Lys Ala Leu Ala Glu Pro Met Val Leu
370 375 380

Val Leu Ser Val Trp Asp Asp His Ala Val Asn Met Leu Trp Leu Asp
385 390 395 400

Ser Thr Tyr Pro Thr Asp Ser Thr Lys Pro Gly Ala Ala Arg Gly Asp
405 410 415

Cys Pro Ile Thr Ser Gly Val Pro Ala Asp Val Glu Ser Gln Ala Pro
420 425 430

Asn Ser Asn Val Ile Tyr Ser Asn Ile Arg Phe Gly Pro Ile Asn Ser
435 440 445

Thr Tyr Thr Gly Thr Pro Ser Gly Gly Asn Pro Pro Gly Gly Gly Thr
450 455 460

Thr Thr Thr Thr Thr Thr Thr Thr Ser Lys Pro Ser Gly Pro Thr Thr
465 470 475 480

Thr Thr Asn Pro Ser Gly Pro Gln Gln Thr His Trp Gly Gln Cys Gly
485 490 495

Gly Gln Gly Trp Thr Gly Pro Thr Val Cys Gln Ser Pro Tyr Thr Cys
500 505 510

Lys Tyr Ser Asn Asp Trp Tyr Ser Gln Cys Leu
515 520

<210> 7

<211> 3649

5 <212> ADN

<213> Chaetomium thermophilum

<220>

<221> CDS

10 <222> (1290)..(2879)

<223>

<220>

<221> Intrón

15 <222> (2880)..(2943)

<223>

<220>

<221> CDS

20 <222> (2944)..(2949)

<223>

<400> 7

tctagagctg tgcacgcggc cgcgtaatac gactcactat agggcgaaga attcggatcg	60
gactagagct cgtcacgggc tcgcgccgac gaggcgatga ggacgaaggg ccgacataat	120
ccgtacttta cgctacatga cgactctcga aaattgtaaa gggccggcat ttcggagcga	180
gtgctgcgag ggcgcattcg cggcgtacct ggaattcctg gaatggtaag caatggccag	240
caatgggcca ggtatggacc agcttgaatc ctggttgctg cgtcaccagg cccagcatgg	300
tgcccagaat ggcccaccgt ggcccacgt cctaagaaac aagctgcgtc ccgcgatcca	360
aaaacgtcgt cttcggcgca cgtcctccgt ggtcccccg gctggacacc ctggctggcc	420
ctccaatgag cggcatttgc ccctgtcgag cgtgtcggca accttaatcg actccatctc	480
tcggctccac gccgtccatc ctgtcctcga cctcgtcatc tgtgtctccc ttgccctccc	540
ttgcccttcc ttgcctccgc cagcagctgc cacaatgtga ccctgctgcc cggagcgccc	600
agcgccatgc accgtttggg cttgtcgccg tgtcgccagt ctccatcgag cgattcgacc	660
gtgtgcctct ctccaccagc gttccccgcg ctctccatag tccatgctac ttcgagccgt	720
tgcctcaca gctgccagcg gcatggctct gtcggtctcg cctctccttt tccggtgaag	780
cgtgccata caattctccg tctgccccag tccttgaggc gccgctattc ccaatcggcc	840
atggcactgg ccagcccgat ccatgttcga tcgagcttcg acgggcccgtg agccgtctgc	900
acggaggagc ttgcgagcct gcgaacctgg cggacctgga gaagcctggc ccatctccct	960
ggatggagat actgggtgcg ctagcaccac ggcgtgccac ggccaagctc cggccgaccc	1020
ggaggcggga agaggggttc gttgctgtct tcggcggctg tcagggcaaa gggtaatcgt	1080
caatgtggga aaaggggctc atctccatga gattcatgac tcggacatcg tctatataag	1140

tcgagtcgcc catcctccaa cagccgattc tgctcctcat cccatcacca ccctcgtcca	1200
caaccacgca gttgtgtaca tcaaaacaag ttcgctcctt ttacatcttc accacaacaa	1260
cagcacatcc tctcctttcg gctttcaag atg atg tat aag aag ttc gcc gct Met Met Tyr Lys Lys Phe Ala Ala	1313
ctc gcc gcc ctc gtg gct ggc gcc tcc gcc cag cag gct tgc tcc ctc Leu Ala Ala Leu Val Ala Gly Ala Ser Ala Gln Gln Ala Cys Ser Leu	1361
acc gct gag aac cac cct agc ctc acc tgg aag cgc tgc acc tct ggc Thr Ala Glu Asn His Pro Ser Leu Thr Trp Lys Arg Cys Thr Ser Gly	1409
ggc agc tgc tcg acc gtg aac ggc gcc gtc acc atc gat gcc aac tgg Gly Ser Cys Ser Thr Val Asn Gly Ala Val Thr Ile Asp Ala Asn Trp	1457
cgc tgg act cac acc gtc tcc ggc tcg acc aac tgc tac acc ggc aac Arg Trp Thr His Thr Val Ser Gly Ser Thr Asn Cys Tyr Thr Gly Asn	1505
cag tgg gat acc tcc ctc tgc act gat ggc aag agc tgc gcc cag acc Gln Trp Asp Thr Ser Leu Cys Thr Asp Gly Lys Ser Cys Ala Gln Thr	1553
tgc tgc gtc gat ggc gct gac tac tct tcg acc tat ggt atc acc acc Cys Cys Val Asp Gly Ala Asp Tyr Ser Ser Thr Tyr Gly Ile Thr Thr	1601
agc ggt gac tcc ctg aac ctc aag ttc gtc acc aag cac cag tac ggc Ser Gly Asp Ser Leu Asn Leu Lys Phe Val Thr Lys His Gln Tyr Gly	1649
acc aac gtc ggc tcc cgt gtc tat ctg atg gag aac gac acc aag tac Thr Asn Val Gly Ser Arg Val Tyr Leu Met Glu Asn Asp Thr Lys Tyr	1697
cag atg ttc gag ctc ctc ggc aac gag ttc acc ttc gat gtc gat gtc Gln Met Phe Glu Leu Leu Gly Asn Glu Phe Thr Phe Asp Val Asp Val	1745
tcc aac ctg ggc tgc ggt ctc aac ggc gcc ctc tac ttc gtt tcc atg Ser Asn Leu Gly Cys Gly Leu Asn Gly Ala Leu Tyr Phe Val Ser Met	1793
gat gct gat ggt ggc atg agc aaa tac tct ggc aac aag gct ggc gcc Asp Ala Asp Gly Gly Met Ser Lys Tyr Ser Gly Asn Lys Ala Gly Ala	1841
aag tac ggt acc ggc tac tgc gat gct cag tgc ccg cgc gac ctc aag Lys Tyr Gly Thr Gly Tyr Cys Asp Ala Gln Cys Pro Arg Asp Leu Lys	1889
ttc atc aac ggc gag gcc aac gtt ggg aac tgg acc ccc tcg acc aac Phe Ile Asn Gly Glu Ala Asn Val Gly Asn Trp Thr Pro Ser Thr Asn	1937
gat gcc aac gcc ggc ttc ggc cgc tat ggc agc tgc tgc tct gag atg Asp Ala Asn Ala Gly Phe Gly Arg Tyr Gly Ser Cys Cys Ser Glu Met	1985
gat gtc tgg gag gcc aac aac atg gct act gcc ttc act cct cac cct Asp Val Trp Glu Ala Asn Asn Met Ala Thr Ala Phe Thr Pro His Pro	2033

ES 2 527 681 T3

tgc acc acc gtt ggc cag agc cgc tgc gag gcc gac acc tgc ggt ggc Cys Thr Thr Val Gly Gln Ser Arg Cys Glu Ala Asp Thr Cys Gly Gly	2081
acc tac agc tct gac cgc tat gct ggt gtt tgc gac cct gat ggc tgc Thr Tyr Ser Ser Asp Arg Tyr Ala Gly Val Cys Asp Pro Asp Gly Cys	2129
gac ttc aac gcc tac cgc caa ggc gac aag acc ttc tac ggc aag ggc Asp Phe Asn Ala Tyr Arg Gln Gly Asp Lys Thr Phe Tyr Gly Lys Gly	2177
atg act gtc gac acc aac aag aag atg acc gtc gtc acc cag ttc cac Met Thr Val Asp Thr Asn Lys Lys Met Thr Val Val Thr Gln Phe His	2225
aag aac tcg gct ggc gtc ctc agc gag atc aag cgc ttc tac gtc cag Lys Asn Ser Ala Gly Val Leu Ser Glu Ile Lys Arg Phe Tyr Val Gln	2273
gac ggc aag atc att gcc aac gct gag tcc aag atc ccc ggc aac ccc Asp Gly Lys Ile Ile Ala Asn Ala Glu Ser Lys Ile Pro Gly Asn Pro	2321
gga aac tcc att acc cag gag tat tgc gat gcc cag aag gtc gcc ttc Gly Asn Ser Ile Thr Gln Glu Tyr Cys Asp Ala Gln Lys Val Ala Phe	2369
agt aac acc gat gac ttc aac cgc aag ggc ggt atg gct cag atg agc Ser Asn Thr Asp Asp Phe Asn Arg Lys Gly Gly Met Ala Gln Met Ser	2417
aag gcc ctc gca ggc ccc atg gtc ctg gtc atg tcc gtc tgg gat gac Lys Ala Leu Gly Pro Met Val Leu Val Met Ser Val Trp Asp Asp	2465
cac tac gcc aac atg ctc tgg ctc gac tcg acc tac ccc atc gac cag His Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Ile Asp Gln	2513
gcc ggc gcc ccc ggc gcc gag cgc ggt gct tgc ccg acc acc tcc ggt Ala Gly Ala Pro Gly Ala Glu Arg Gly Ala Cys Pro Thr Thr Ser Gly	2561
gtc cct gcc gag atc gag gcc cag gtc ccc aac agc aac gtc atc ttc Val Pro Ala Glu Ile Glu Ala Gln Val Pro Asn Ser Asn Val Ile Phe	2609
tcc aac atc cgt ttc ggc ccc atc ggc tcg acc gtc cct ggc ctt gac Ser Asn Ile Arg Phe Gly Pro Ile Gly Ser Thr Val Pro Gly Leu Asp	2657
ggc agc aac ccc ggc aac ccg acc acc acc gtc gtt cct ccc gct tct Gly Ser Asn Pro Gly Asn Pro Thr Thr Val Val Pro Pro Ala Ser	2705
acc tcc acc tcc cgt ccg acc agc agc act agc tct ccc gtt tcg acc Thr Ser Thr Ser Arg Pro Thr Ser Thr Ser Ser Pro Val Ser Thr	2753
ccg act ggc cag ccc ggc ggc tgc acc acc cag aag tgg ggc cag tgc Pro Thr Gly Gln Pro Gly Gly Cys Thr Thr Gln Lys Trp Gly Gln Cys	2801
ggc ggt atc ggc tac acc ggc tgc act aac tgc gtt gct ggc acc acc Gly Gly Ile Gly Tyr Thr Gly Cys Thr Asn Cys Val Ala Gly Thr Thr	2849

tgc act cag ctc aac ccc tgg tac agc cag gtatgtttct cttccccctt 2899
Cys Thr Gln Leu Asn Pro Trp Tyr Ser Gln
525 530

ctagactcgc ttggatttga cagttgctaa catctgctca acag tgc ctg 2949
Cys Leu

taaacaactc gcttcgtccg cacgacggag gagggccatg agaaagaatg ggcaacatag 3009
attcttttgcg cggttgtgga ctacttgggt attttctgga tgtacatagt tttatcacgt 3069
catgaggctg tcatgtgggg atgtgtatct ttttcgcttc ttcgtacata aatttacgca 3129
ttgagctttt cccccccaa aaacagttcc ctgatttgct ggagtaactt gatggtaaag 3189
cttggtcata agctcttcaa tggaaaaaac gatacagtca tgccttgaca catcctccca 3249
aagtcttcgt ccatgacatc acggtcgatc cttaagcaca agttcaataa ccccatgtgg 3309
cgttgccttg tcctgaaaca cagatgagat cttcagccca gccgcatcgg ccacttcctt 3369
gaactgagcc aacgagcggt ctttcccgcc gattgagagc atcgcatagt ccttgaaggc 3429
tgcatagaga ggaatagggg gcttgtttcc ggtagttggg ctgccggaac tcggatctgt 3489
tggcgcaagg gggtcagggg tgatctgctc ggcgatgagg acgcgtccat cggggtttgt 3549
tagtgcacga gcgacattgc gcaggatggt gactgccaca gggtcggagt aatcgcgagg 3609
gatgtggcgg aggtagtaga ccagtgcacc tggaatcgat 3649

<210> 8

<211> 532

<212> PRT

<213> Chaetomium thermophilum

<400> 8

Met Met Tyr Lys Lys Phe Ala Ala Leu Ala Ala Leu Val Ala Gly Ala
1 5 10 15

Ser Ala Gln Gln Ala Cys Ser Leu Thr Ala Glu Asn His Pro Ser Leu
20 25 30

Thr Trp Lys Arg Cys Thr Ser Gly Gly Ser Cys Ser Thr Val Asn Gly
35 40 45

Ala Val Thr Ile Asp Ala Asn Trp Arg Trp Thr His Thr Val Ser Gly
50 55 60

Ser Thr Asn Cys Tyr Thr Gly Asn Gln Trp Asp Thr Ser Leu Cys Thr
65 70 75 80

Asp Gly Lys Ser Cys Ala Gln Thr Cys Cys Val Asp Gly Ala Asp Tyr
85 90 95

ES 2 527 681 T3

Ser Ser Thr Tyr Gly Ile Thr Thr Ser Gly Asp Ser Leu Asn Leu Lys
 100 105 110
 Phe Val Thr Lys His Gln Tyr Gly Thr Asn Val Gly Ser Arg Val Tyr
 115 120 125
 Leu Met Glu Asn Asp Thr Lys Tyr Gln Met Phe Glu Leu Leu Gly Asn
 130 135 140
 Glu Phe Thr Phe Asp Val Asp Val Ser Asn Leu Gly Cys Gly Leu Asn
 145 150 155 160
 Gly Ala Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Met Ser Lys
 165 170 175
 Tyr Ser Gly Asn Lys Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp
 180 185 190
 Ala Gln Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Glu Ala Asn Val
 195 200 205
 Gly Asn Trp Thr Pro Ser Thr Asn Asp Ala Asn Ala Gly Phe Gly Arg
 210 215 220
 Tyr Gly Ser Cys Cys Ser Glu Met Asp Val Trp Glu Ala Asn Asn Met
 225 230 235 240
 Ala Thr Ala Phe Thr Pro His Pro Cys Thr Thr Val Gly Gln Ser Arg
 245 250 255
 Cys Glu Ala Asp Thr Cys Gly Gly Thr Tyr Ser Ser Asp Arg Tyr Ala
 260 265 270
 Gly Val Cys Asp Pro Asp Gly Cys Asp Phe Asn Ala Tyr Arg Gln Gly
 275 280 285
 Asp Lys Thr Phe Tyr Gly Lys Gly Met Thr Val Asp Thr Asn Lys Lys
 290 295 300
 Met Thr Val Val Thr Gln Phe His Lys Asn Ser Ala Gly Val Leu Ser
 305 310 315 320
 Glu Ile Lys Arg Phe Tyr Val Gln Asp Gly Lys Ile Ile Ala Asn Ala
 325 330 335
 Glu Ser Lys Ile Pro Gly Asn Pro Gly Asn Ser Ile Thr Gln Glu Tyr
 340 345 350
 Cys Asp Ala Gln Lys Val Ala Phe Ser Asn Thr Asp Asp Phe Asn Arg
 355 360 365

ES 2 527 681 T3

Lys Gly Gly Met Ala Gln Met Ser Lys Ala Leu Ala Gly Pro Met Val
 370 375 380
 Leu Val Met Ser Val Trp Asp Asp His Tyr Ala Asn Met Leu Trp Leu
 385 390 395 400
 Asp Ser Thr Tyr Pro Ile Asp Gln Ala Gly Ala Pro Gly Ala Glu Arg
 405 410 415
 Gly Ala Cys Pro Thr Thr Ser Gly Val Pro Ala Glu Ile Glu Ala Gln
 420 425 430
 Val Pro Asn Ser Asn Val Ile Phe Ser Asn Ile Arg Phe Gly Pro Ile
 435 440 445
 Gly Ser Thr Val Pro Gly Leu Asp Gly Ser Asn Pro Gly Asn Pro Thr
 450 455 460
 Thr Thr Val Val Pro Pro Ala Ser Thr Ser Thr Ser Arg Pro Thr Ser
 465 470 475 480
 Ser Thr Ser Ser Pro Val Ser Thr Pro Thr Gly Gln Pro Gly Gly Cys
 485 490 495
 Thr Thr Gln Lys Trp Gly Gln Cys Gly Gly Ile Gly Tyr Thr Gly Cys
 500 505 510
 Thr Asn Cys Val Ala Gly Thr Thr Cys Thr Gln Leu Asn Pro Trp Tyr
 515 520 525

Ser Gln Cys Leu
530

<210> 9

<211> 1339

<212> ADN

5 <213> Thermoascus aurantiacus

<220>

<221> CDS

<222> (17)..(122)

<223>

10 <220>

<221> Intrón

<222> (123)..(177)

<223>

<220>

15 <221> CDS

<222> (178)..(236)

<223>

<220>

<221> Intrón

20 <222> (237)..(296)

<223>

<220>

<221> CDS

<222> (297)..(449)

25 <223>

<220>

<221> Intrón

<222> (450)..(508)

<223>

30 <220>

<221> CDS

<222> (509)..(573)

	<223>		
	<220>		
	<221> Intrón		
	<222> (574)..(647)		
5	<223>		
	<220>		
	<221> CDS		
	<222> (648)..(745)		
	<223>		
10	<220>		
	<221> Intrón		
	<222> (746)..(806)		
	<223>		
	<220>		
15	<221> CDS		
	<222> (807)..(1330)		
	<223>		
	<400> 9		
	ccgcggactg cgcac	atg aag ctc ggc tct ctc gtg ctc gct ctc agc gca	52
		Met Lys Leu Gly Ser Leu Val Leu Ala Leu Ser Ala	
		1 5 10	
	gct agg ctt aca ctg tcg gcc cct ctc gca gac agg aag cag gag acc		100
	Ala Arg Leu Thr Leu Ser Ala Pro Leu Ala Asp Arg Lys Gln Glu Thr		
		15 20 25	
	aag cgt gcg aaa gta ttc caa t gttcgtaca tccacgtctg gcttgctggc		152
	Lys Arg Ala Lys Val Phe Gln		
		30 35	
	ttactggcaa ctgacaatgg cgaag gg ttc ggt tca aac gag tcc ggt gct		203
		Trp Phe Gly Ser Asn Glu Ser Gly Ala	
		40	
	gaa ttc gga agc cag aac ctt cca gga gtc gag gtcagcatgc ctgtactctc		256
	Glu Phe Gly Ser Gln Asn Leu Pro Gly Val Glu		
		45 50 55	
	tgcattatat taatatctca agaggcttac tctttcgcag gga aag gat tat ata		311
		Gly Lys Asp Tyr Ile	
		60	
	tgg cct gat ccc aac acc att gac aca ttg atc agc aag ggg atg aac		359
	Trp Pro Asp Pro Asn Thr Ile Asp Thr Leu Ile Ser Lys Gly Met Asn		
		65 70 75	
	atc ttt cgt gtc ccc ttt atg atg gag aga ttg gtt ccc aac tca atg		407
	Ile Phe Arg Val Pro Phe Met Met Glu Arg Leu Val Pro Asn Ser Met		
		80 85 90	
	acc ggc tct ccg gat ccg aac tac ctg gca gat ctc ata gcg		449
	Thr Gly Ser Pro Asp Pro Asn Tyr Leu Ala Asp Leu Ile Ala		
		95 100 105	
	gtacatttca attccacat gtttggagct gtcttcgttg tgctgacatt taatggtag		508

ES 2 527 681 T3

act gta aat gca atc acc cag aaa ggt gcc tac gcc gtc gtc gat cct 556
Thr Val Asn Ala Ile Thr Gln Lys Gly Ala Tyr Ala Val Val Asp Pro
110 115 120

cat aac tac ggc aga ta gtgaggtccc cggttctggt attgctgctg 603
His Asn Tyr Gly Arg Tyr
125

tatatctaag tagatatgtg tttctaacat ttccacgatt tcag c tac aat tct 657
Tyr Asn Ser
130

ata atc tcg agc cct tcc gat ttc cag acc ttc tgg aaa acg gtc gcc 705
Ile Ile Ser Ser Pro Ser Asp Phe Gln Thr Phe Trp Lys Thr Val Ala
135 140 145

tca cag ttt gct tcg aat cca ctg gtc atc ttc gac act a gtaagctgaa 755
Ser Gln Phe Ala Ser Asn Pro Leu Val Ile Phe Asp Thr
150 155 160

cacccgaaat taactgagtc tgagcatgtc tgacaagacg atccatgaaa g at aac 811
Asn Asn

gaa tac cac gat atg gac cag acc tta gtc ctc aat ctc aac cag gcc 859
Glu Tyr His Asp Met Asp Gln Thr Leu Val Leu Asn Leu Asn Gln Ala
165 170 175

gct atc gac ggc atc cgt tcc gcc gga gcc act tcc cag tac atc ttt 907
Ala Ile Asp Gly Ile Arg Ser Ala Gly Ala Thr Ser Gln Tyr Ile Phe
180 185 190

gtc gag ggc aat tcg tgg acc ggg gca tgg acc tgg acg aac gtg aac 955
Val Glu Gly Asn Ser Trp Thr Gly Ala Trp Thr Trp Thr Asn Val Asn
195 200 205 210

gat aac atg aaa agc ctg acc gac cca tct gac aag atc ata tac gag 1003
Asp Asn Met Lys Ser Leu Thr Asp Pro Ser Asp Lys Ile Ile Tyr Glu
215 220 225

atg cac cag tac ctg gac tct gac gga tcc ggg aca tca gcg acc tgc 1051
Met His Gln Tyr Leu Asp Ser Asp Gly Ser Gly Thr Ser Ala Thr Cys
230 235 240

gta tct tcg acc atc ggt caa gag cga atc acc agc gca acg caa tgg 1099
Val Ser Ser Thr Ile Gly Gln Glu Arg Ile Thr Ser Ala Thr Gln Trp
245 250 255

ctc agg gcc aac ggg aag aag ggc atc atc ggc gag ttt gcg ggc gga 1147
Leu Arg Ala Asn Gly Lys Lys Gly Ile Ile Gly Glu Phe Ala Gly Gly
260 265 270

gcc aac gac gtc tgc gag acg gcc atc acg ggc atg ctg gac tac atg 1195
Ala Asn Asp Val Cys Glu Thr Ala Ile Thr Gly Met Leu Asp Tyr Met
275 280 285 290

gcc cag aac acg gac gtc tgg act ggc gcc atc tgg tgg gcg gcc ggg 1243
Ala Gln Asn Thr Asp Val Trp Thr Gly Ala Ile Trp Trp Ala Ala Gly
295 300 305

ccg tgg tgg gga gac tac ata ttc tcc atg gag ccg gac aat ggc atc 1291
Pro Trp Trp Gly Asp Tyr Ile Phe Ser Met Glu Pro Asp Asn Gly Ile
310 315 320

gcg tat cag cag ata ctt cct att ttg act ccg tat ctt tgactgcag 1339
Ala Tyr Gln Gln Ile Leu Pro Ile Leu Thr Pro Tyr Leu
325 330 335

<210> 10

<211> 335

<212> PRT

<213> Thermoascus aurantiacus

<400> 10

ES 2 527 681 T3

Met Lys Leu Gly Ser Leu Val Leu Ala Leu Ser Ala Ala Arg Leu Thr
 1 5 10 15
 Leu Ser Ala Pro Leu Ala Asp Arg Lys Gln Glu Thr Lys Arg Ala Lys
 20 25 30
 Val Phe Gln Trp Phe Gly Ser Asn Glu Ser Gly Ala Glu Phe Gly Ser
 35 40 45
 Gln Asn Leu Pro Gly Val Glu Gly Lys Asp Tyr Ile Trp Pro Asp Pro
 50 55 60
 Asn Thr Ile Asp Thr Leu Ile Ser Lys Gly Met Asn Ile Phe Arg Val
 65 70 75 80
 Pro Phe Met Met Glu Arg Leu Val Pro Asn Ser Met Thr Gly Ser Pro
 85 90 95
 Asp Pro Asn Tyr Leu Ala Asp Leu Ile Ala Thr Val Asn Ala Ile Thr
 100 105 110
 Gln Lys Gly Ala Tyr Ala Val Val Asp Pro His Asn Tyr Gly Arg Tyr
 115 120 125
 Tyr Asn Ser Ile Ile Ser Ser Pro Ser Asp Phe Gln Thr Phe Trp Lys
 130 135 140
 Thr Val Ala Ser Gln Phe Ala Ser Asn Pro Leu Val Ile Phe Asp Thr
 145 150 155 160
 Asn Asn Glu Tyr His Asp Met Asp Gln Thr Leu Val Leu Asn Leu Asn
 165 170 175
 Gln Ala Ala Ile Asp Gly Ile Arg Ser Ala Gly Ala Thr Ser Gln Tyr
 180 185 190
 Ile Phe Val Glu Gly Asn Ser Trp Thr Gly Ala Trp Thr Trp Thr Asn
 195 200 205
 Val Asn Asp Asn Met Lys Ser Leu Thr Asp Pro Ser Asp Lys Ile Ile
 210 215 220
 Tyr Glu Met His Gln Tyr Leu Asp Ser Asp Gly Ser Gly Thr Ser Ala
 225 230 235 240
 Thr Cys Val Ser Ser Thr Ile Gly Gln Glu Arg Ile Thr Ser Ala Thr
 245 250 255
 Gln Trp Leu Arg Ala Asn Gly Lys Lys Gly Ile Ile Gly Glu Phe Ala
 260 265 270
 Gly Gly Ala Asn Asp Val Cys Glu Thr Ala Ile Thr Gly Met Leu Asp
 275 280 285
 Tyr Met Ala Gln Asn Thr Asp Val Trp Thr Gly Ala Ile Trp Trp Ala
 290 295 300
 Ala Gly Pro Trp Trp Gly Asp Tyr Ile Phe Ser Met Glu Pro Asp Asn
 305 310 315 320
 Gly Ile Ala Tyr Gln Gln Ile Leu Pro Ile Leu Thr Pro Tyr Leu
 325 330 335

<210> 11
 <211> 2334
 <212> ADN
 5 <213> Acremonium thermophilum

 <220>
 <221> rasgo_misc
 <222> (13)..(13)
 10 <223> N = no identificado

 <220>
 <221> CDS
 <222> (715)..(797)
 15 <223>

 <220>
 <221> Intrón
 <222> (798)..(856)
 20 <223>

 <220>
 <221> CDS
 <222> (857)..(1105)
 25 <223>

 <220>
 <221> Intrón
 <222> (1106)..(1228)
 30 <223>

 <220>
 <221> CDS
 <222> (1229)..(1787)
 35 <223>

 <400> 11
 tctgtctctt gtntcagaac agatctcctg gcggcctgct ttgccggtcc gaattgcat 60
 cgatgcaacg tcgattgcat acgagctaag cccgtctcgt gataaccgca aggggtcttc 120
 cgagtttctg tctgcgaccc aggcattttc cgatttgtgt gcggggaccc aactgtcttc 180
 tggggagtac ctggtgacaa aagcacagat aaacagatgg atgacggtat tgctgtgata 240
 tcgccgtggc gctgaatcct ttctcttcgc taccaagata ttatttcccc gttgtgaaat 300
 cttctattca gcccattcca tccggcaaca cgcattctgct tttcgttcg gcattccgat 360
 acctggttcc tggagtgcct accgagcctc gcttcctggg atcgggctgt gcaccccgcc 420
 aaaccctatg ccccaaacg tacggacaag gatgccggac cccggttttg tccagaaagg 480
 ttgcattcct acccacctcg ctggagccac aacatgcaga tcaccgcccg agggaggaca 540
 tgtgtggtgc agggacgttg gcaactctgc tgtgtctgaa gtatatgagg ccgatggttc 600
 tccttgacaa aagcagagaa tggagtagcc agctcctcct caccagagtc gcctttgcag 660
 cgtctcggca ttgcaggctc cccatcgtca gcatttcaact tctcagcaac gaac atg 717
 Met
 1

 cgc tcc tca ccc ttt ctc cgc gca gct ctg gct gcc gct ctg cct ctg 765
 Arg Ser Ser Pro Phe Leu Arg Ala Ala Leu Ala Ala Ala Leu Pro Leu
 5 10 15

 agc gcc cat gcc ctc gac gga aag tcg acg ag gtatgccaat cctcgtacct 817
 Ser Ala His Ala Leu Asp Gly Lys Ser Thr Arg
 20 25

 ctgccctctg tagaaacaag tgaccgactg caaagacag a tac tgg gac tgc tgc 872

ES 2 527 681 T3

Tyr Trp Asp Cys Cys																
30																
aag	ccg	tcc	tgc	ggc	tgg	ccg	gga	aag	gcc	tgc	gtg	aac	cag	ccc	gtc	920
Lys	Pro	Ser	Cys	Gly	Trp	Pro	Gly	Lys	Ala	Ser	Val	Asn	Gln	Pro	Val	
35						40					45					
ttc	tgc	tgc	tgc	gcc	gac	tgg	cag	cgc	atc	agc	gac	ttc	aac	gcg	aag	968
Phe	Ser	Cys	Ser	Ala	Asp	Trp	Gln	Arg	Ile	Ser	Asp	Phe	Asn	Ala	Lys	
50					55					60					65	
tgc	ggc	tgc	gac	gga	ggc	tcc	gcc	tac	tgc	tgc	gcc	gac	cag	acg	ccc	1016
Ser	Gly	Cys	Asp	Gly	Gly	Ser	Ala	Tyr	Ser	Cys	Ala	Asp	Gln	Thr	Pro	
				70					75					80		
tgg	gcg	gtc	aac	gac	aac	ttc	tgc	tac	ggc	ttc	gca	gcc	acg	gcc	atc	1064
Trp	Ala	Val	Asn	Asp	Asn	Phe	Ser	Tyr	Gly	Phe	Ala	Ala	Thr	Ala	Ile	
			85					90					95			
gcc	ggc	ggc	tcc	gag	tcc	agc	tgg	tgc	tgc	gcc	tgc	tat	gc			1105
Ala	Gly	Gly	Ser	Glu	Ser	Ser	Trp	Cys	Cys	Ala	Cys	Tyr	Ala			
			100				105					110				
gtgagttctc	tgcaagccgc	ttcccccccc	cgctttctgt	gcaggccgct	tccccctac											1165
ccaccacatt	cccccccccc	gcctctgtga	tggggcatcc	gagctaagtt	gcgtgtcgtc											1225
cag	a	ctc	acc	ttc	aac	tgc	ggc	ccc	gtc	gcg	ggc	aag	acc	atg	gtg	1274
	Leu	Thr	Phe	Asn	Ser	Gly	Pro	Val	Ala	Gly	Lys	Thr	Met	Val	Val	
				115					120					125		
cag	tgc	acc	agc	acc	ggc	ggc	gac	ctg	ggc	agc	aac	cag	ttc	gac	ctc	1322
Gln	Ser	Thr	Ser	Thr	Gly	Gly	Asp	Leu	Gly	Ser	Asn	Gln	Phe	Asp	Leu	
			130					135					140			
gcc	atc	ccc	ggc	ggc	ggc	gtg	ggc	atc	ttc	aac	ggc	tgc	gcc	tcc	cag	1370
Ala	Ile	Pro	Gly	Gly	Gly	Val	Gly	Ile	Phe	Asn	Gly	Cys	Ala	Ser	Gln	
		145					150					155				
ttc	ggc	ggc	ctc	ccc	ggc	ggc	cag	tac	ggc	ggc	atc	agc	gac	cgc	agc	1418
Phe	Gly	Gly	Leu	Pro	Gly	Ala	Gln	Tyr	Gly	Gly	Ile	Ser	Asp	Arg	Ser	
	160					165					170					
cag	tgc	tgc	tcc	ttc	ccc	gcg	ccg	ctc	cag	ccg	ggc	tgc	cag	tgg	cgc	1466
Gln	Cys	Ser	Ser	Phe	Pro	Ala	Pro	Leu	Gln	Pro	Gly	Cys	Gln	Trp	Arg	
					180					185					190	
ttc	gac	tgg	ttc	cag	aac	gcc	gac	aac	ccc	acc	ttc	acc	ttc	cag	cgc	1514
Phe	Asp	Trp	Phe	Gln	Asn	Ala	Asp	Asn	Pro	Thr	Phe	Thr	Phe	Gln	Arg	
				195					200					205		
gtg	cag	tgc	ccg	tcc	gag	ctc	acg	tcc	cgc	acg	ggc	tgt	aag	cgc	gac	1562
Val	Gln	Cys	Pro	Ser	Glu	Leu	Thr	Ser	Arg	Thr	Gly	Cys	Lys	Arg	Asp	
			210					215					220			
gac	gac	gcc	agc	tat	ccc	gtc	ttc	aac	ccg	cct	agc	ggt	ggc	tcc	ccc	1610
Asp	Asp	Ala	Ser	Tyr	Pro	Val	Phe	Asn	Pro	Pro	Ser	Gly	Gly	Ser	Pro	
		225				230						235				
agc	acc	acc	agc	acc	acc	acc	agc	tcc	ccg	tcc	ggt	ccc	acg	ggc	aac	1658
Ser	Thr	Thr	Ser	Thr	Thr	Thr	Ser	Ser	Pro	Ser	Gly	Pro	Thr	Gly	Asn	
			240			245					250					
cct	cct	gga	ggc	ggt	ggc	tgc	act	gcc	cag	aag	tgg	gcc	cag	tgc	ggc	1706
Pro	Pro	Gly	Gly	Gly	Gly	Cys	Thr	Ala	Gln	Lys	Trp	Ala	Gln	Cys	Gly	
					260					265					270	
ggc	act	ggc	ttc	acg	ggc	tgc	acc	acc	tgc	gtc	tgc	ggc	acc	acc	tgc	1754

gly Thr Gly Phe Thr Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys
275 280 285
cag gtg cag aac cag tgg tat tcc cag tgt ctg tgagcgggag ggttggtggg 1807
Gln Val Gln Asn Gln Trp Tyr Ser Gln Cys Leu
290 295
gtccgtttcc ctagggtga ggctgacgtg aactgggtcc tcttggtccgc cccatcacgg 1867
gttcgtattc gcgcgcttag ggagaggagg atgcagtttg agggggccac attttgaggg 1927
ggacgcagtc tggggtcgaa gcttgctcgt tagggctgcc gtgacgtggt agagcagatg 1987
ggaccaagtg cggagctagg cagggtgggtg gttgtggtgg tggcttacct tctgtaacgc 2047
aatggcatct catctcactc gcctgctccc tgattggtgg ctctgttcgg cctggcgctt 2107
tttgggaccg ctggctggaa tggattgctc cggaacgcca ggttgagctg ggctggcgcg 2167
agtagattgg ccgctccgag ctgcaacat aataaaattt tcggaccctg taagccgcac 2227
ccgaccaggt ctccattggc ggacatgcac gacgtccttc gcaggcacgg cctgcccgcc 2287
tctgatcacc cgagttttc gtaccgtcag accagataca agccccg 2334
<210> 12
<211> 297
<212> PRT
5 <213> Acremonium thermophilum
<220>
<221> rasgo_misc
<222> (13)..(13)
10 <223> N = no identificado
<400> 12
Met Arg Ser Ser Pro Phe Leu Arg Ala Ala Leu Ala Ala Ala Leu Pro
1 5 10 15
Leu Ser Ala His Ala Leu Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
20 25 30
Cys Lys Pro Ser Cys Gly Trp Pro Gly Lys Ala Ser Val Asn Gln Pro
35 40 45
Val Phe Ser Cys Ser Ala Asp Trp Gln Arg Ile Ser Asp Phe Asn Ala
50 55 60
Lys Ser Gly Cys Asp Gly Gly Ser Ala Tyr Ser Cys Ala Asp Gln Thr
65 70 75 80
Pro Trp Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ala
85 90 95

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

<210> 13
<211> 2033
5 <212> ADN
<213> Acremonium thermophilum

<220>
<221> CDS
10 <222> (259)..(702)
<223>

<220>
<221> Intrón
15 <222> (703)..(857)
<223>

<220>
<221> CDS
20 <222> (858)..(888)
<223>

<220>
<221> Intrón
25 <222> (889)..(990)
<223>

<220>
 <221> CDS
 <222> (991)..(1268)
 <223>

5

```

<400> 13
ctcgaggaga ggaaccgagt ttgaaagatg ctatatatcg atagactacc ggcgtcgcct      60
cgccctgtcc gctctcttgc attccccctg ttgatgagac gagacaaaat tcctgggttag      120
aaaagatccg tcgccgagat ttcaccagtg gtaagtcccg agaattggtc attcgacggt      180
caatatgagt gtcaaagcta tgggtcctaa caaagaagga agcaagagct ttaaagagac      240
agaataacag cagcaaag atg  cgt  ctc  cca  cta  ccg  act  ctg  ctc  gcc  ctc      291
                    Met Arg Leu Pro Leu Pro Thr Leu Leu Ala Leu
                    1      5      10

ttg ccc tac tac ctc gaa gtg tcc gct cag ggg gca tcc gga acc ggc      339
Leu Pro Tyr Tyr Leu Glu Val Ser Ala Gln Gly Ala Ser Gly Thr Gly
                    15      20      25

acg aca aca cgt tac tgg gat tgc tgc aag ccg agc tgc gcg tgg cct      387
Thr Thr Thr Arg Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro
                    30      35      40

ctg aag ggc aat tcg ccc agc ccg gtg cag act tgc gac aag aat gac      435
    
```

Leu	Lys	Gly	Asn	Ser	Pro	Ser	Pro	Val	Gln	Thr	Cys	Asp	Lys	Asn	Asp		
45						50					55						
agg	ccg	ctg	aac	gat	ggg	gga	aac	acc	aag	tcc	ggc	tgc	gac	aac	ggt	483	
Arg	Pro	Leu	Asn	Asp	Gly	Gly	Asn	Thr	Lys	Ser	Gly	Cys	Asp	Asn	Gly		
60					65					70					75		
ggc	ggg	gcc	ttc	atg	tgc	tca	tcc	cag	agt	ccc	tgg	gcc	gtc	aat	gag	531	
Gly	Gly	Ala	Phe	Met	Cys	Ser	Ser	Gln	Ser	Pro	Trp	Ala	Val	Asn	Glu		
				80					85					90			
acc	acc	agc	tac	ggc	tgg	gca	gcc	gtt	cgt	atc	gcc	ggc	agt	acc	gag	579	
Thr	Thr	Ser	Tyr	Gly	Trp	Ala	Ala	Val	Arg	Ile	Ala	Gly	Ser	Thr	Glu		
				95				100					105				
tcg	gcc	tgg	tgc	tgt	gcc	tgc	tac	gag	ctc	acc	ttc	acc	agt	ggg	ccc	627	
Ser	Ala	Trp	Cys	Cys	Ala	Cys	Tyr	Glu	Leu	Thr	Phe	Thr	Ser	Gly	Pro		
		110					115					120					
gtc	agt	gga	aag	aag	ctc	ata	gtc	cag	gcc	acg	aac	act	ggt	gga	gac	675	
Val	Ser	Gly	Lys	Lys	Leu	Ile	Val	Gln	Ala	Thr	Asn	Thr	Gly	Gly	Asp		
		125				130					135						
ctt	ggg	agc	aac	cac	ttt	gac	ctt	gcg	gtatgtgggg	tttttctttc						722	
Leu	Gly	Ser	Asn	His	Phe	Asp	Leu	Ala									
140					145												
ttcatcatcg	ctctcaccat	ggattcctcg	gcgcaaggac	caagattgag	aagcgtcaat											782	
gccgggttgg	acacgggagc	cgggatagga	acacagaggc	cgtttaagac	cgtcagctga											842	
cagcagagca	attag	att	ccc	gga	ggt	ggt	ggt	ggt	cag	tcc	aat	g				888	
			Ile	Pro	Gly	Gly	Gly	Val	Gly	Gln	Ser	Asn					
				150				155									
gtaggttcct	tcctgaagt	accggcaaca	gcctgtgcgt	tgctgtatac	cccttttaat											948	
catagcatct	tcctgctgga	tacaagccaa	cccattttct	ag	ct	tgc	acg	aac								1001	
						Ala	Cys	Thr	Asn								
							160										
cag	tat	ggt	gcg	ccc	ccg	aac	ggc	tgg	ggc	gac	agg	tat	ggt	ggc	gtg	1049	
Gln	Tyr	Gly	Ala	Pro	Pro	Asn	Gly	Trp	Gly	Asp	Arg	Tyr	Gly	Gly	Val		
		165					170					175					
cac	tcg	cgg	agc	gac	tgc	gac	agc	ttc	ccc	gcg	gcg	ctc	aag	gcc	ggc	1097	
His	Ser	Arg	Ser	Asp	Cys	Asp	Ser	Phe	Pro	Ala	Ala	Leu	Lys	Ala	Gly		
		180				185					190						
tgc	tac	tgg	cga	ttc	gac	tgg	ttc	cag	ggc	gcc	gac	aac	ccg	tcc	gtg	1145	
Cys	Tyr	Trp	Arg	Phe	Asp	Trp	Phe	Gln	Gly	Ala	Asp	Asn	Pro	Ser	Val		
					200				205						210		
agc	ttc	aaa	cag	gta	gcc	tgc	ccg	gca	gcc	atc	aca	gct	aag	agc	ggc	1193	
Ser	Phe	Lys	Gln	Val	Ala	Cys	Pro	Ala	Ala	Ile	Thr	Ala	Lys	Ser	Gly		
				215				220						225			
tgt	act	cgc	cag	aac	gat	gcc	atc	aac	gag	act	ccg	act	ggg	ccc	agc	1241	
Cys	Thr	Arg	Gln	Asn	Asp	Ala	Ile	Asn	Glu	Thr	Pro	Thr	Gly	Pro	Ser		
			230					235					240				
act	gtg	cct	acc	tac	acc	gcg	tca	ggc	tgaaagtcgg	ctggggcacc						1288	
Thr	Val	Pro	Thr	Tyr	Thr	Ala	Ser	Gly									
		245				250											
attgcccagg	tgatggttgg	gcatgtgtta	gtctcactca	ccagggacat	ttgtcgcgac											1348	
ctgatcatag	gcgccagggg	agttgaaagg	ggttgccgta	cgagaagaca	ttttgtcgcc											1408	

gtcttactcc cagccacttc tgtacatatt caatgacatt acatagcccg caaatatggt 1468
catatatcgt ggccgccc aa accgccccgg tttgcttagg ctggagctga agtggctcgc 1528
cgatggctgt caaaggcagt cggaatattc ctcgttgctt cggcaacacg gtagctgctt 1588
gaaccgtacc cagcattaga acaccccccg ccgaggggctt gctacgtcaa tggcggggtc 1648
tccaacccct gcgcggcaca aaaccaacca cgccctcgtc ttttatgatg tcctcgtcca 1708
aacgtcccgt gacgacactc cgctcatggg ctggctctct gatgtagaag gggtagggtca 1768
gccgatgggtc gtcaccgtcg tcaatgcttc cctcaagctt cttgcggcct ttatcctcca 1828
actcttccca catgagaact ccatctttcc gccttttcac aaagccactg ccctccttgt 1888
caagggccaa aaaccaacgc cgctgatgaa tgcttccgat cgtgtttgac gcgccccggg 1948
tatgcatttg gttcggcgca cttttttcgt cctccagctc ccttaactcc cgttccatct 2008
gagaggggtga ctcgtctact cgact 2033

<210> 14

<211> 251

<212> PRT

<213> Acremonium thermophilum

<400> 14

Met Arg Leu Pro Leu Pro Thr Leu Leu Ala Leu Leu Pro Tyr Tyr Leu
1 5 10 15

Glu Val Ser Ala Gln Gly Ala Ser Gly Thr Gly Thr Thr Thr Arg Tyr
20 25 30

Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Leu Lys Gly Asn Ser
35 40 45

Pro Ser Pro Val Gln Thr Cys Asp Lys Asn Asp Arg Pro Leu Asn Asp
50 55 60

Gly Gly Asn Thr Lys Ser Gly Cys Asp Asn Gly Gly Gly Ala Phe Met
65 70 75 80

Cys Ser Ser Gln Ser Pro Trp Ala Val Asn Glu Thr Thr Ser Tyr Gly
85 90 95

Trp Ala Ala Val Arg Ile Ala Gly Ser Thr Glu Ser Ala Trp Cys Cys
100 105 110

Ala Cys Tyr Glu Leu Thr Phe Thr Ser Gly Pro Val Ser Gly Lys Lys
115 120 125

ES 2 527 681 T3

Leu Ile Val Gln Ala Thr Asn Thr Gly Gly Asp Leu Gly Ser Asn His
130 135 140

Phe Asp Leu Ala Ile Pro Gly Gly Gly Val Gly Gln Ser Asn Ala Cys
145 150 155 160

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

Gly Val His Ser Arg Ser Asp Cys Asp Ser Phe Pro Ala Ala Leu Lys
180 185 190

Ala Gly Cys Tyr Trp Arg Phe Asp Trp Phe Gln Gly Ala Asp Asn Pro
195 200 205

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

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

Pro Ser Thr Val Pro Thr Tyr Thr Ala Ser Gly
245 250

<210> 15

<211> 2800

<212> ADN

5 <213> Chaetomium thermophilum

<220>

<221> rasgo_misc

<222> (2786)..(2786)

10 <223> N = no identificado

<220>

<221> CDS

<222> (768)..(2042)

15 <223>

<400> 15

ggatccaaga cccatcccgga ggattctcgg attatgtttg catctcacc tccgaaaccg 60

catgaaaaat tgaaatgggc aactgtcgct gtgtttaatg ctttgcacat catgggatca 120

tgttcacccg ctctaattct tcattctcca gatcttatct atcctccgca tctagccggc 180

ttcttgcttg tgatccaaag ccctgatccc acgcggcttc tagacgcttt agaaattaca	240
ccgaatctcc ccatgccctt cttgcaatat cttcccgacc aggaacttcg ggtgctcaac	300
atccgcgagc ttgacgacga cccttcttgg ccggcttggc atgcgactct gttcgggact	360
caatgcaact ctgggccctt caatgccgcg catgaccgtt actgaggctt agccgcccc	420
atcgcttggc acggtacctt gcagacggaa tcccgggccc gttgtccgat ctgctttggt	480
tccggtagag aagcctcggg ggaagagaca cacggacaca acgattgcgg gccccaatgc	540
gctgctccta attgaggctc cgaggctcgtg tgccgtgtgg agaggccgcg actgggtctg	600
gggtgctggg gattgcggag atgaagataa tctgggtgca accgtggata cataaaagg	660
agtagttctc ccctctgtga aaccttcttc cccaggattc tcctcgcttc taagagtcca	720
aagtcattca agacatccta cagcgggggc agtgagattc cataatc atg act cgc Met Thr Arg 1	776
aag ttc gca ctc gtt ccc ctc ctt ctg ggt ctt gcc tcg gcc cag aaa Lys Phe Ala Leu Val Pro Leu Leu Gly Leu Ala Ser Ala Gln Lys 5 10 15	824
ccc ggc aac act cca gaa gtc cac ccc aag atc acc act tac cgc tgc Pro Gly Asn Thr Pro Glu Val His Pro Lys Ile Thr Thr Tyr Arg Cys 20 25 30 35	872
agc cac cgc cag gga tgc cgc ccg gag acg aac tac atc gtc ctc gac Ser His Arg Gln Gly Cys Arg Pro Glu Thr Asn Tyr Ile Val Leu Asp 40 45 50	920
tcc ctc acc cat ccc gtg cac cag ttg aac tcc aac gcg aac tgc ggc Ser Leu Thr His Pro Val His Gln Leu Asn Ser Asn Ala Asn Cys Gly 55 60 65	968
gac tgg ggt aac ccg ccc ccg cgc agc gtc tgc cct gat gtc gag acc Asp Trp Gly Asn Pro Pro Pro Arg Ser Val Cys Pro Asp Val Glu Thr 70 75 80	1016
tgc gcg cag aat tgc atc atg gag ggc atc caa gac tac tcc acc tac Cys Ala Gln Asn Cys Ile Met Glu Gly Ile Gln Asp Tyr Ser Thr Tyr 85 90 95	1064
ggc gtg acc acc tct ggc tct tcc ctt cgc ctg aag cag atc cac cag Gly Val Thr Thr Ser Gly Ser Ser Leu Arg Leu Lys Gln Ile His Gln 100 105 110 115	1112
ggc cgc gtc acc tct cct cgt gtc tac ctc ctc gac aag acg gag cag Gly Arg Val Thr Ser Pro Arg Val Tyr Leu Leu Asp Lys Thr Glu Gln 120 125 130	1160
cag tat gag atg atg cgt ctc acc ggc ttc gag ttc act ttc gac gtc Gln Tyr Glu Met Met Arg Leu Thr Gly Phe Glu Phe Thr Phe Asp Val 135 140 145	1208
gac acc acc aag ctc ccc tgc ggc atg aac gct gcg ctc tat ctc tcc Asp Thr Thr Lys Leu Pro Cys Gly Met Asn Ala Ala Leu Tyr Leu Ser 150 155 160	1256
gag atg gac gct acc ggc gct cgc tcc cgc ctc aac cct ggc ggt gcc Glu Met Asp Ala Thr Gly Ala Arg Ser Arg Leu Asn Pro Gly Gly Ala 165 170 175	1304
tac tac ggc acg ggt tac tgc gat gca cag tgc ttc gtc acc ccc ttc	1352

ES 2 527 681 T3

Tyr 180	Tyr	Gly	Thr	Gly	Tyr 185	Cys	Asp	Ala	Gln	Cys 190	Phe	Val	Thr	Pro	Phe 195	
atc Ile	aat Asn	ggc Gly	atc Ile	ggc Gly 200	aac Asn	atc Ile	gag Glu	ggc Gly	aag Lys 205	ggc Gly	tgc Ser	tgc Cys	tgc Cys	aac Asn 210	gag Glu	1400
atg Met	gac Asp	att Ile	tgg Trp 215	gag Glu	gcc Ala	aac Asn	tgc Ser	cgt Arg 220	agt Ser	cag Gln	tcc Ser	att Ile	gct Ala 225	ccg Pro	cac His	1448
ccc Pro	tgc Cys	aac Asn 230	aag Lys	cag Gln	ggt Gly	ctg Leu	tac Tyr 235	atg Met	tgc Cys	tcc Ser	ggc Gly	cag Gln 240	gag Glu	tgc Cys	gag Glu	1496
ttc Phe	gac Asp 245	ggc Gly	gtc Val	tgc Cys	gac Asp	gag Glu 250	tgg Trp	gga Gly	tgc Cys	aca Thr	tgg Trp 255	aac Asn	ccg Pro	tac Tyr	aag Lys	1544
gtc Val 260	aac Asn	gtt Val	acc Thr	gac Asp	tac Tyr 265	tat Tyr	ggc Gly	cgc Arg	ggt Gly	ccg Pro 270	cag Gln	ttc Phe	aag Lys	gtc Val	gac Asp 275	1592
acg Thr	acc Thr	cgt Arg	ccc Pro	ttc Phe 280	acc Thr	gtc Val	atc Ile	aca Thr	cag Gln 285	ttt Phe	cca Pro	gcc Ala	gac Asp	cag Gln 290	aac Asn	1640
ggc Gly	aag Lys	ctg Leu 295	acg Thr	tgc Ser	atc Ile	cat His	cgc Arg	atg Met 300	tat Tyr	gtg Val	caa Gln	gat Asp	ggc Gly 305	aag Lys	ttg Leu	1688
atc Ile	gag Glu 310	gcg Ala	cat His	acc Thr	gtc Val	aac Asn	ctg Leu 315	ccg Pro	ggt Gly	tat Tyr	cct Pro	caa Gln 320	gtg Val	aac Asn	gcg Ala	1736
ctg Leu 325	aac Asn	gat Asp	gac Asp	ttc Phe	tgc Cys	cgt Arg 330	gcc Ala	acg Thr	gga Gly	gcc Ala	gcg Ala 335	acg Thr	aag Lys	tat Tyr	ctt Leu	1784
gaa Glu 340	ctg Leu	ggt Gly	gcc Ala	act Thr	gcg Ala 345	ggt Gly	atg Met	ggc Gly	gag Glu	gct Ala 350	ctg Leu	agg Arg	cgt Arg	ggt Gly	atg Met 355	1832
gtg Val	ctg Leu	gct Ala	atg Met	agc Ser 360	atc Ile	tgg Trp	tgg Trp	gat Asp	gag Glu 365	agc Ser	ggc Gly	ttc Phe	atg Met	aac Asn 370	tgg Trp	1880
ctt Leu	gat Asp	agc Ser	ggc Gly 375	gag Glu	tct Ser	ggg Gly	ccg Pro	tgc Cys 380	aac Asn	ccg Pro	aac Asn	gag Glu	ggt Gly 385	aac Asn	cca Pro	1928
cag Gln	aac Asn	att Ile 390	cgc Arg	cag Gln	att Ile	gag Glu	ccc Pro 395	gag Glu	ccg Pro	gag Glu	gtt Val	acc Thr 400	tat Tyr	agc Ser	aac Asn	1976
ctg Leu 405	cgc Arg	tgg Trp	ggt Gly	gag Glu	att Ile	ggg Gly 410	tgc Ser	act Thr	tat Tyr	aag Lys	cac His 415	aac Asn	ctg Leu	aag Lys	ggc Gly	2024
ggg Gly 420	tgg Trp	act Thr	ggc Gly	agg Arg	aac Asn 425	taagtgttgg ggattagagc ctgtgattgg										2072
atacctgtgg gttaaacggg gctcggtttg agagggttgt tgaaatttat ttctcgtaca																2132
tagttggcgt cttggcgaat atatgcccc aggactttga tccagtcctc gtccatttct																2192
ctgtgactta gttggtgcaa gtatcattgt tatgtcctgg gtgagacaaa gcaatctctt																2252

ES 2 527 681 T3

cagtggcatc gggtaaataa tctacaggct gtgaatggcg ttgcgtcagc ctcatctaact 2312
 taaacgattg gactccccctt ttcctaataca tcgccgttgc cgtgtaactc tcctagatct 2372
 cttgttgat atggcttcaa ctcgaagtga agaaaaatgg atacggcgac ctctttgtgc 2432
 caattttctt gctgttcttc cggatttgac cctcggcaag acaactatgg ccaatattct 2492
 gttatagtcg gcagttagtg ttgtgtcgt caagtcgtgc gggagcaata ctcaacagcc 2552
 gcccttaata tgggtattta cgccacgacg cacttcatta cacggctttg gggggtatat 2612
 attccgttca actctatccc tcattcgggtg tgattgaacg tctccaacag tgaaagtata 2672
 agtctgacaa aaatgcccaa ccgccatgcc actgatgac ctgttgagat gctcgtggtc 2732
 tataacatcc tgtctaagtg ttacctccct aatgttagcc ccagttctgc tctncttgtc 2792
 tcgacagc 2800

<210> 16

<211> 425

5 <212> PRT

<213> Chaetomium thermophilum

<220>

<221> rasgo_misc

10 <222> (2786)..(2786)

<223> N = no identificado

<400> 16

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

Ala Gln Lys Pro Gly Asn Thr Pro Glu Val His Pro Lys Ile Thr Thr
 20 25 30

Tyr Arg Cys Ser His Arg Gln Gly Cys Arg Pro Glu Thr Asn Tyr Ile
 35 40 45

Val Leu Asp Ser Leu Thr His Pro Val His Gln Leu Asn Ser Asn Ala
 50 55 60

Asn Cys Gly Asp Trp Gly Asn Pro Pro Pro Arg Ser Val Cys Pro Asp
 65 70 75 80

Val Glu Thr Cys Ala Gln Asn Cys Ile Met Glu Gly Ile Gln Asp Tyr
 85 90 95

Ser Thr Tyr Gly Val Thr Thr Ser Gly Ser Ser Leu Arg Leu Lys Gln
 100 105 110

ES 2 527 681 T3

Ile His Gln Gly Arg Val Thr Ser Pro Arg Val Tyr Leu Leu Asp Lys
115 120 125

Thr Glu Gln Gln Tyr Glu Met Met Arg Leu Thr Gly Phe Glu Phe Thr
130 135 140

Phe Asp Val Asp Thr Thr Lys Leu Pro Cys Gly Met Asn Ala Ala Leu
145 150 155 160

Tyr Leu Ser Glu Met Asp Ala Thr Gly Ala Arg Ser Arg Leu Asn Pro
165 170 175

Gly Gly Ala Tyr Tyr Gly Thr Gly Tyr Cys Asp Ala Gln Cys Phe Val
180 185 190

Thr Pro Phe Ile Asn Gly Ile Gly Asn Ile Glu Gly Lys Gly Ser Cys
195 200 205

Cys Asn Glu Met Asp Ile Trp Glu Ala Asn Ser Arg Ser Gln Ser Ile
210 215 220

Ala Pro His Pro Cys Asn Lys Gln Gly Leu Tyr Met Cys Ser Gly Gln
225 230 235 240

Glu Cys Glu Phe Asp Gly Val Cys Asp Glu Trp Gly Cys Thr Trp Asn
245 250 255

Pro Tyr Lys Val Asn Val Thr Asp Tyr Tyr Gly Arg Gly Pro Gln Phe
260 265 270

Lys Val Asp Thr Thr Arg Pro Phe Thr Val Ile Thr Gln Phe Pro Ala
275 280 285

Asp Gln Asn Gly Lys Leu Thr Ser Ile His Arg Met Tyr Val Gln Asp
290 295 300

Gly Lys Leu Ile Glu Ala His Thr Val Asn Leu Pro Gly Tyr Pro Gln
305 310 315 320

Val Asn Ala Leu Asn Asp Asp Phe Cys Arg Ala Thr Gly Ala Ala Thr
325 330 335

Lys Tyr Leu Glu Leu Gly Ala Thr Ala Gly Met Gly Glu Ala Leu Arg
340 345 350

Arg Gly Met Val Leu Ala Met Ser Ile Trp Trp Asp Glu Ser Gly Phe
355 360 365

Met Asn Trp Leu Asp Ser Gly Glu Ser Gly Pro Cys Asn Pro Asn Glu
370 375 380

Gly Asn Pro Gln Asn Ile Arg Gln Ile Glu Pro Glu Pro Glu Val Thr
385 390 395 400

Tyr Ser Asn Leu Arg Trp Gly Glu Ile Gly Ser Thr Tyr Lys His Asn
405 410 415

Leu Lys Gly Gly Trp Thr Gly Arg Asn
420 425

<210> 17
<211> 1943
<212> ADN
<213> Thermoascus aurantiacus

5 <220>
 <221> CDS
 <222> (13)..(256)
 <223>

10 <220>
 <221> Intrón
 <222> (257)..(329)
 <223>

15 <220>
 <221> CDS
 <222> (330)..(370)
 <223>

20 <220>
 <221> Intrón
 <222> (371)..(444)
 <223>

25 <220>
 <221> CDS
 <222> (445)..(493)
 <223>

30 <220>
 <221> Intrón
 <222> (494)..(561)
 <223>

35 <220>
 <221> CDS
 <222> (562)..(683)
 <223>

40 <220>
 <221> Intrón
 <222> (684)..(786)
 <223>

45 <220>
 <221> CDS
 <222> (787)..(932)
 <223>

50 <220>
 <221> Intrón
 <222> (933)..(1001)
 <223>

55 <220>
 <221> CDS
 <222> (1002)..(1090)
 <223>

60 <220>
 <221> Intrón
 <222> (1091)..(1155)
 <223>

65 <220>
 <221> CDS
 <222> (1156)..(1174)
 <223>

<220>
 <221> Intrón
 <222> (1175)..(1267)
 <223>
 5
 <220>
 <221> CDS
 <222> (1268)..(1295)
 <223>
 10
 <220>
 <221> Intrón
 <222> (1296)..(1361)
 <223>
 15
 <220>
 <221> CDS
 <222> (1362)..(1451)
 <223>
 20
 <220>
 <221> Intrón
 <222> (1452)..(1551)
 <223>
 25
 <220>
 <221> CDS
 <222> (1552)..(1617)
 <223>
 30
 <220>
 <221> Intrón
 <222> (1618)..(1829)
 <223>
 35
 <220>
 <221> CDS
 <222> (1830)..(1922)
 <223>
 40
 <400> 17
 ccgcgggaag cc atg gtt cga cca acg atc cta ctt act tca ctc ctg cta 51
 Met Val Arg Pro Thr Ile Leu Leu Thr Ser Leu Leu Leu
 1 5 10
 gct ccc ttc gca gct gcg agc cct atc ctc gag gaa cgc caa gct gca 99
 Ala Pro Phe Ala Ala Ala Ser Pro Ile Leu Glu Glu Arg Gln Ala Ala
 15 20 25
 cag agt gtc gac caa ctg atc aag gct cgc ggc aag gtg tac ttt ggc 147
 Gln Ser Val Asp Gln Leu Ile Lys Ala Arg Gly Lys Val Tyr Phe Gly
 30 35 40 45
 gtc gcc acg gac caa aac cgg ctg acg acc ggc aag aat gcg gct atc 195
 Val Ala Thr Asp Gln Asn Arg Leu Thr Thr Gly Lys Asn Ala Ala Ile
 50 55 60
 atc cag gct gat ttc ggc cag gtc acg ccg gag aat agt atg aaa tgg 243
 Ile Gln Ala Asp Phe Gly Gln Val Thr Pro Glu Asn Ser Met Lys Trp
 65 70 75
 gac gct act gaa c gtgcgtgaga aagataattt gatttttttc ttctatgacc 296
 Asp Ala Thr Glu
 80
 gctcggaccg ttctgactag gtttataata tag ct tct caa gga aac ttc aac 349
 Pro Ser Gln Gly Asn Phe Asn
 85
 ttt gcc ggt gct gat tac ctt gtacgtacat acgaccactt gacgtttctt 400
 Phe Ala Gly Ala Asp Tyr Leu
 90 95
 gcacgcaact gcgattgagg agaagataact aatcttcttg aaag gtc aat tgg gcc 456
 Val Asn Trp Ala
 cag caa aat gga aag ctg atc cgt ggc cat act ctt g gtttagtagaa 503

Gln Gln Asn Gly Lys Leu Ile Arg Gly His Thr Leu
100 105 110

cgccaacctg cttccctaac ttactgaaga aggaaaaccg aattgaccgt cccccaag 561

ta tgg cac tcg cag ctg ccc tcg tgg gtg agc tcc atc acc gac aag 608
Val Trp His Ser Gln Leu Pro Ser Trp Val Ser Ser Ile Thr Asp Lys
115 120 125

aat acg ctg acc aac gtg atg aaa aat cac atc acc acc ttg atg acc 656
Asn Thr Leu Thr Asn Val Met Lys Asn His Ile Thr Thr Leu Met Thr
130 135 140

cgg tac aag ggc aag atc cgt gca tgg gtcagtcac ctaccctaag 703
Arg Tyr Lys Gly Lys Ile Arg Ala Trp
145 150

ctgcgtttca atgaagagac aaataagaac acacgtattt gcccgggcgt ttcagaatca 763

gaactgacag aatcactgaa tag gac gtg gtg aac gag gca ttc aac gag gat 816
Asp Val Val Asn Glu Ala Phe Asn Glu Asp
155 160

ggc tcc ctc cgc cag act gtc ttc ctc aac gtc atc ggg gag gat tac 864
Gly Ser Leu Arg Gln Thr Val Phe Leu Asn Val Ile Gly Glu Asp Tyr
165 170 175

atc ccg att gct ttc cag acc gcc cgc gcc gct gac ccg aat gcc aag 912
Ile Pro Ile Ala Phe Gln Thr Ala Arg Ala Ala Asp Pro Asn Ala Lys
180 185 190

ctg tac atc aac gat tac aa gtaagattta aggctcagtg atattccatt 962
Leu Tyr Ile Asn Asp Tyr Asn
195 200

tagtgtgaga agcattgctt atgagcatct gtattacag c ctc gac agt gcc tcg 1017
Leu Asp Ser Ala Ser
205

tac ccc aag acg cag gcc att gtc aac cgc gtc aag caa tgg cgt gca 1065
Tyr Pro Lys Thr Gln Ala Ile Val Asn Arg Val Lys Gln Trp Arg Ala
210 215 220

gct gga gtc ccg att gac ggc ata g gtatgtctct ctttctgttt 1110
Ala Gly Val Pro Ile Asp Gly Ile
225 230

gtgatgtgac cgatttgaaa ccagtctaac gttagctggg tctag ga tcg caa acg 1166
Gly Ser Gln Thr
235

cac ctc ag gtaaataatc gggaatgcct cggagaataa aagagaaaaa 1214
His Leu Ser
235

aatgattgt cttatcagat cgtatcgact gactcatggc ttgtccaaaa tag c gct 1271
Ala
240

ggc cag gga gcc ggt gtt cta caa taagtgtccc cctcccctat tttttactat 1325
Gly Gln Gly Ala Gly Val Leu Gln
240 245

tattgcgaga gcggaatagg ctgacaaccc caaacg gct ctt ccg ctc ctt gct 1379
Ala Leu Pro Leu Leu Ala
250

agt gcc gga act ccc gag gtc gct atc acg gaa ctg gac gtg gct ggt 1427

ES 2 527 681 T3

Ser Ala Gly Thr Pro Glu Val Ala Ile Thr Glu Leu Asp Val Ala Gly
255 260 265

gct agc ccg acg gat tac gtc aat gtatgtacct cgttgccct atcccccttg 1481
Ala Ser Pro Thr Asp Tyr Val Asn
270 275

gatactttgt ataattatta tcttcccgga gcctgttgat cagatctgac gatcatttct 1541

cgttttttag gtc gtg aac gct tgc ctc aac gtg cag tcc tgc gtg ggc 1590
Val Val Asn Ala Cys Leu Asn Val Gln Ser Cys Val Gly
280 285

atc acc gtc tgg ggc gtg gca gat ccg gtaagcgcgg ttcttccgta 1637
Ile Thr Val Trp Gly Val Ala Asp Pro
290 295

ctccgtaccc aactagagtt cgggctgtca cgtcatgtct tagtcgtctt cagtcaggcc 1697

aaggccaaga cacaggacct gaaacgggca ggcagcagct gtagcagcc caagaagcag 1757

ccacatgatg catgattatt attattatat ctccgagttc tgggctaacg attggtgata 1817

ataaataaat ag gac tca tgg cgt gct agc acg acg cct ctc ctc ttc gac 1868
Asp Ser Trp Arg Ala Ser Thr Thr Pro Leu Leu Phe Asp
300 305 310

ggc aac ttc aac ccg aag ccg gcg tac aac gcc att gtg cag gac ctg 1916
Gly Asn Phe Asn Pro Lys Pro Ala Tyr Asn Ala Ile Val Gln Asp Leu
315 320 325

cag cag tgagtataga ccggtggatc c 1943
Gln Gln
<210> 18
<211> 329
<212> PRT
5 <213> Thermoascus aurantiacus

<400> 18
Met Val Arg Pro Thr Ile Leu Leu Thr Ser Leu Leu Leu Ala Pro Phe
1 5 10 15

Ala Ala Ala Ser Pro Ile Leu Glu Glu Arg Gln Ala Ala Gln Ser Val
20 25 30

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

Asp Gln Asn Arg Leu Thr Thr Gly Lys Asn Ala Ala Ile Ile Gln Ala
50 55 60

Asp Phe Gly Gln Val Thr Pro Glu Asn Ser Met Lys Trp Asp Ala Thr
65 70 75 80

ES 2 527 681 T3

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

<210> 19
 <211> 2955
 5 <212> ADN
 <213> Acremonium thermophilum

<220>
 <221> CDS
 10 <222> (1335)..(1671)
 <223>

<220>
 <221> Intrón
 15 <222> (1672)..(1806)
 <223>

ES 2 527 681 T3

5 <220>
<221> CDS
<222> (1807)..(2032)
<223>

10 <220>
<221> Intrón
<222> (2033)..(2117)
<223>

15 <220>
<221> CDS
<222> (2118)..(2802)
<223>

```
<400> 19
tctagagctg tcgacgcggc cgcgtaatac gactcactat agggcgaaga attcggatca    60
cgtttgcttc agcaagtcgt tcgctacgac accacgtcca tgatggaggc cctgattcaa    120
tcataccaag gacggggcat gatggctgat ggctggactc gaagtgagtg gcccgtggct    180
gaattttcct tcccgttctc tacagtcctt ccctcagcga cacatccgca gttttgacag    240
cggaaatcgt caggatgctc cgccttctct cgcaacctga gtgccaggc gtctcggcca    300
ccgtctctta tatatggccg ctgggtccgc ctttcgatcg gttttcgatt tggctctctc    360
```

tagttccctc agctgacccg ggatatacgt tgtgggtccg aaacctcacc atcccagacg	420
agcaagttct ccgcagtcca cctcagctca tccggccctt ggtagcatcg cagcgacccc	480
agacgaaggc accaaagaag catactatat attaggctaa atcgagcccc acgtggaata	540
tttgccatcg aggaggggtg gttgggcttc ttgtcctcgc aggtgctgcg cctgtaccta	600
cctgggtgctc cagctgggtg tcccgtggt gctgttccag tcgccgtctg gcccgaatgc	660
tctgtatctc ggttcgtccc gcactccttt cgccaagcgc taccaatgct ttgacgaacc	720
cggtaaattt gcagtggacc tgcagctggg caaaccgcga gtgggaacca cagacctggt	780
tcgttcgaca cactccaatc gcaacccgc ccgcgcaaac cttgcaccac atgtcgcccc	840
tttcccagtt gggtcctga agacacggag ccacttccgt gatcgctggc tccccaagcc	900
gacagtcgga cgctgcaata ggatgccagc acccggtgat ccaagggcca gtgaccccaa	960
ctctttcgcg gtattctggc cctcccaaag gtatgccagg acttccctgt ctttgctacc	1020
accagctctc ctccacggcg gaacggatac gccgtctcgc cggctcttgc tcgacaacat	1080
gcgagggggc gcgaaggcta ggttgtagcg atgcgacggt gcgatgtcac catttggcag	1140
tgatgttttc cgttgctccc ttctccacc tgcccggtt cctcaaagac gcccacaacca	1200
taaatacgat gcgacgcaa cttcatgtg ttctgtggcat cttgcctgac cagtctcagc	1260
aagaaacctg tggcggcgcg attgtcttga ctttctgatt gaaaacggat ctgcgtcctc	1320
ctcgatagcc gacc atg cgc gcc aag caa ctc ctg gcg gcc ggc ctg ctg Met Arg Ala Lys Gln Leu Leu Ala Ala Gly Leu Leu 1 5 10	1370
gcc ccc gcg tcc gtc tcg gcc cag ctc aac agc ctc gcc gtg gcg gct Ala Pro Ala Ser Val Ser Ala Gln Leu Asn Ser Leu Ala Val Ala Ala 15 20 25	1418
ggc ctc aag tac ttc ggc acg gcc gtg cgg gag gcc aac gtc aac ggc Gly Leu Lys Tyr Phe Gly Thr Ala Val Arg Glu Ala Asn Val Asn Gly 30 35 40	1466
gac gcc acc tac atg tcg tac gtc aac aac aag tcc gag ttc ggc cag Asp Ala Thr Tyr Met Ser Tyr Val Asn Asn Lys Ser Glu Phe Gly Gln 45 50 55 60	1514
gtg acg ccc gag aac ggc cag aag tgg gat tcc acc gag ccc agc cag Val Thr Pro Glu Asn Gly Gln Lys Trp Asp Ser Thr Glu Pro Ser Gln 65 70 75	1562
ggc cag ttc agc tac agc cag gcc gac atc gtc ccc gcc gtc gcg aag Gly Gln Phe Ser Tyr Ser Gln Gly Asp Ile Val Pro Gly Val Ala Lys 80 85 90	1610
aag aac ggc cag gtg ctg cgc tgc cac acc ctg gtg tgg tac agc cag Lys Asn Gly Gln Val Leu Arg Cys His Thr Leu Val Trp Tyr Ser Gln 95 100 105	1658
ctc ccc agc tgg g gtcagtgact ctctctttct ctctgtcttt ctctttgtct Leu Pro Ser Trp 110	1711
ttctctcttt ctctctctct ctctctctct ctctctctct ctctctccca tccagcatcg	1771
actgctgac ttgctgacca gaagctcgtg tgcag tg tca tcc gga agt tgg	1823

ES 2 527 681 T3

Val Ser Ser Gly Ser Trp 115																	
acc Thr 120	cgc Arg 120	gcg Ala	acg Thr	ctt Leu	cag Gln 125	tcc Ser 125	gtc Val	atc Ile	gag Glu	acg Thr 130	cac His 130	atc Ile	tcg Ser	aac Asn	gtg Val	1871	
atg Met 135	ggc Gly	cac His	tac Tyr	aag Lys 140	ggc Gly 140	cag Gln	tgc Cys	tac Tyr	gcc Ala 145	tgg Trp 145	gac Asp	gtg Val	gtc Val	aac Asn	gag Glu 150	1919	
gcc Ala	atc Ile	aac Asn	gac Asp 155	gac Asp 155	ggc Gly	acg Thr	tgg Trp	cgg Arg	acc Thr 160	agc Ser	gtc Val	ttc Phe	tac Tyr	aac Asn 165	acc Thr	1967	
ttc Phe	aac Asn	acc Thr 170	gac Asp 170	tac Tyr	ctg Leu	gcc Ala	att Ile 175	gcc Ala 175	ttc Phe	aac Asn	gcc Ala	gcg Ala	aag Lys 180	aag Lys	gcc Ala	2015	
gat Asp	gcg Ala 185	ggc Gly 185	gcg Ala	aag Lys	ct Leu	gtaggtgtcg gcctttacgt tgccgcagcg										2062	
cacctccgcg acatgagccc cagagcgcgt ggctaatagt tcctcacgca cgacag g																	2118
tac Tyr 190	tac Tyr 190	aac Asn	gac Asp	tac Tyr	aat Asn	ctc Leu 195	gag Glu	tac Tyr	aac Asn	ggc Gly 200	gcc Ala 200	aag Lys	acc Thr	aac Asn	acg Thr	2166	
gcc Ala 205	gtg Val	cag Gln	ctg Leu	gtg Val	cag Gln 210	atc Ile	gtg Val	cag Gln	cag Gln	gcc Ala 215	ggc Gly	gcg Ala	ccc Pro	atc Ile	gac Asp 220	2214	
ggg Gly	gtg Val	ggc Gly	ttc Phe	cag Gln 225	ggc Gly	cac His	ctg Leu	atc Ile	gtg Val 230	ggg Gly	tca Ser	acg Thr	ccg Pro	tcg Ser 235	cg Arg	2262	
agc Ser	tcc Ser	ctg Leu	gcc Ala 240	acg Thr	gcg Ala	ctg Leu	aag Lys	cgc Arg 245	ttc Phe	acg Thr	gcg Ala	ctt Leu	ggc Gly 250	ctg Leu	gag Glu	2310	
gtg Val	gcg Ala	tac Tyr 255	acg Thr	gag Glu	ctg Leu	gac Asp	atc Ile 260	cgg Arg	cac His	tcg Ser	agc Ser	ctg Leu 265	ccg Pro	ccg Pro	tcg Ser	2358	
tcg Ser	gcg Ala 270	gcg Ala	ctg Leu	gcg Ala	acg Thr	cag Gln 275	ggc Gly	aac Asn	gac Asp	ttc Phe	gcc Ala 280	agc Ser	gtg Val	gtg Val	ggc Gly	2406	
tcg Ser 285	tgc Cys	ctc Leu	gac Asp	gtg Val	gcg Ala 290	ggc Gly	tgc Cys	gtg Val	ggc Gly	atc Ile 295	acc Thr	atc Ile	tgg Trp	ggg Gly	ttc Phe 300	2454	
acg Thr	gac Asp	aag Lys	tac Tyr	agc Ser 305	tgg Trp	gtg Val	ccc Pro	gac Asp	acg Thr 310	ttc Phe	ccc Pro	ggc Gly	tcg Ser	ggc Gly 315	gcg Ala	2502	
gcg Ala	ctg Leu	ctg Leu	tac Tyr 320	gac Asp	gcg Ala	aac Asn	tac Tyr	agc Ser 325	aag Lys	aag Lys	ccg Pro	gcg Ala	tgg Trp 330	acg Thr	tcg Ser	2550	
gtc Val	tcg Ser	tcg Ser 335	gtg Val	ctg Leu	gcg Ala	gcc Ala	aag Lys 340	gcg Ala	acg Thr	aac Asn	ccg Pro	ccc Pro 345	ggc Gly	ggc Gly	ggg Gly	2598	
aac Asn	cca Pro 350	ccc Pro	ccc Pro	gtc Val	acc Thr	acc Thr 355	acg Thr	acc Thr	acg Thr	acc Thr	acg Thr 360	acc Thr	acg Thr	tcg Ser	aag Lys	2646	

ES 2 527 681 T3

ccg tcg cag ccc acc acc acg acc acg acc acc agc ccg cag ggt ccg 2694
Pro Ser Gln Pro Thr Thr Thr Thr Thr Thr Thr Ser Pro Gln Gly Pro
365 370 375 380

cag cag acg cac tgg ggc cag tgc ggc ggc atc ggc tgg acg ggc ccg 2742
Gln Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly Trp Thr Gly Pro
385 390 395

cag tcg tgc cag agc ccg tgg acg tgc cag aag cag aac gac tgg tac 2790
Gln Ser Cys Gln Ser Pro Trp Thr Cys Gln Lys Gln Asn Asp Trp Tyr
400 405 410

tct cag tgc ctg tgaccaccac ggctgaccag ctgccattcc gaccacgggg 2842
Ser Gln Cys Leu
415

cccgactac aaaaagaggg gacggtgtaa ataaagagcc gaacgggtct acgtacactg 2902

ttttgacctt ttctccgcag acgtatatta tcaattatag ttggatttct aga 2955

<210> 20

<211> 416

5 <212> PRT

<213> Acremonium thermophilum

<400> 20

Met Arg Ala Lys Gln Leu Leu Ala Ala Gly Leu Leu Ala Pro Ala Ser
1 5 10 15

Val Ser Ala Gln Leu Asn Ser Leu Ala Val Ala Ala Gly Leu Lys Tyr
20 25 30

Phe Gly Thr Ala Val Arg Glu Ala Asn Val Asn Gly Asp Ala Thr Tyr
35 40 45

Met Ser Tyr Val Asn Asn Lys Ser Glu Phe Gly Gln Val Thr Pro Glu
50 55 60

Asn Gly Gln Lys Trp Asp Ser Thr Glu Pro Ser Gln Gly Gln Phe Ser
65 70 75 80

Tyr Ser Gln Gly Asp Ile Val Pro Gly Val Ala Lys Lys Asn Gly Gln
85 90 95

Val Leu Arg Cys His Thr Leu Val Trp Tyr Ser Gln Leu Pro Ser Trp
100 105 110

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

Thr His Ile Ser Asn Val Met Gly His Tyr Lys Gly Gln Cys Tyr Ala
130 135 140

Trp Asp Val Val Asn Glu Ala Ile Asn Asp Asp Gly Thr Trp Arg Thr
 145 150 155 160
 Ser Val Phe Tyr Asn Thr Phe Asn Thr Asp Tyr Leu Ala Ile Ala Phe
 165 170 175
 Asn Ala Ala Lys Lys Ala Asp Ala Gly Ala Lys Leu Tyr Tyr Asn Asp
 180 185 190
 Tyr Asn Leu Glu Tyr Asn Gly Ala Lys Thr Asn Thr Ala Val Gln Leu
 195 200 205
 Val Gln Ile Val Gln Gln Ala Gly Ala Pro Ile Asp Gly Val Gly Phe
 210 215 220
 Gln Gly His Leu Ile Val Gly Ser Thr Pro Ser Arg Ser Ser Leu Ala
 225 230 235 240
 Thr Ala Leu Lys Arg Phe Thr Ala Leu Gly Leu Glu Val Ala Tyr Thr
 245 250 255
 Glu Leu Asp Ile Arg His Ser Ser Leu Pro Pro Ser Ser Ala Ala Leu
 260 265 270
 Ala Thr Gln Gly Asn Asp Phe Ala Ser Val Val Gly Ser Cys Leu Asp
 275 280 285
 Val Ala Gly Cys Val Gly Ile Thr Ile Trp Gly Phe Thr Asp Lys Tyr
 290 295 300
 Ser Trp Val Pro Asp Thr Phe Pro Gly Ser Gly Ala Ala Leu Leu Tyr
 305 310 315 320
 Asp Ala Asn Tyr Ser Lys Lys Pro Ala Trp Thr Ser Val Ser Ser Val
 325 330 335
 Leu Ala Ala Lys Ala Thr Asn Pro Pro Gly Gly Gly Asn Pro Pro Pro
 340 345 350
 Val Thr Thr Thr Thr Thr Thr Thr Thr Thr Ser Lys Pro Ser Gln Pro
 355 360 365
 Thr Thr Thr Thr Thr Thr Thr Ser Pro Gln Gly Pro Gln Gln Thr His
 370 375 380
 Trp Gly Gln Cys Gly Gly Ile Gly Trp Thr Gly Pro Gln Ser Cys Gln
 385 390 395 400
 Ser Pro Trp Thr Cys Gln Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu
 405 410 415

<210> 21

<211> 5092

<212> ADN

5 <213> Thermoascus aurantiacus

<220>

<221> CDS

<222> (669)..(728)

10 <223>

<220>

<221> Intrón

<222> (729)..(872)
 <223>

 <220>
 5 <221> CDS
 <222> (873)..(1015)
 <223>

 <220>
 10 <221> Intrón
 <222> (1016)..(1082)
 <223>

 <220>
 15 <221> CDS
 <222> (1083)..(1127)
 <223>

 <220>
 20 <221> Intrón
 <222> (1128)..(1183)
 <223>

 <220>
 25 <221> CDS
 <222> (1184)..(1236)
 <223>

 <220>
 30 <221> Intrón
 <222> (1237)..(1300)
 <223>

 <220>
 35 <221> CDS
 <222> (1301)..(1717)
 <223>

 <220>
 40 <221> Intrón
 <222> (1718)..(1776)
 <223>

 <220>
 45 <221> CDS
 <222> (1777)..(2489)
 <223>

 <220>
 50 <221> Intrón
 <222> (2490)..(2599)
 <223>

 <220>
 55 <221> CDS
 <222> (2600)..(3469)
 <223>

 <220>
 60 <221> Intrón
 <222> (3470)..(3531)
 <223>

 <220>
 65 <221> CDS
 <222> (3532)..(3759)

ES 2 527 681 T3

<223>

<400> 21

ggatccgtcc gcggaacacag gcagagagac ggcacgggga ctcgacctga tcctcccagg	60
gcgggggtgtt gtttgtggcg agggagcgat gctgatgttc ttccagctcc gttgctacct	120
tcccacggcc atttagccgg cggacggcat gtaacatgtc aaacatgtgg gctcggcagt	180
gggggctga gacgcagcac ctgacccggc ggcgcggcgc ttgcagggtc cagggacagc	240
cggccgtggt cgtttgcggg gaaggcgaca cagacgactt ggcgcggccc gccggaaggc	300
gaggaatcat gagtgcgacg gagacatggc aagaccacgg ccttcctggc gaagaagaag	360
atgaataatc gcaggggcag tgtggcatgg accgcacggc cgccagggac ctgccccgtg	420
aggtttctcg ggtgtttcca ctggttccat cgctgggggc gatcccagac ccgtgtgccc	480
gtgtaactat tattgacgat caacatgcc a tggccagcca gcttctataa taatcatata	540
taacaccccc cgttctcccg ctgcccttgc ccgtggctct cctggctctg cttgagggtc	600
acgagtctcc ttgcatggtc aactcgtcct ctgcttcatc cgctgcttga ctccgtacct	660
cagcaacc atg agg ctt ggg tgg ctg gag ctg gcc gtc gcg gcg gcc gca	710
Met Arg Leu Gly Trp Leu Glu Leu Ala Val Ala Ala Ala Ala	
1 5 10	
acc gtc gcc agc gcc aag gtgcgtcaga ccctcccccg gatcgacctt	758
Thr Val Ala Ser Ala Lys	
15 20	
taggtgcttc ttcagcaagt gcgcgccggc cgcgacatcc gccgccgctg ccctcaccga	818
cgcagcacc atatgcagca ggagagaagg catctctgac gaaagctccc ccag gat	875
Asp	
gac ttg gcc tac tcg ccg cct ttc tac ccg tcg cca tgg atg aac gga	923
Asp Leu Ala Tyr Ser Pro Pro Phe Tyr Pro Ser Pro Trp Met Asn Gly	
25 30 35	
aac gga gag tgg gcg gag gcc tac cgc agg gct gtc gac ttc gtc tcg	971

Asn Gly Glu Trp Ala Glu Ala Tyr Arg Arg Ala Val Asp Phe Val Ser
40 45 50

cag ctg acc ctc gcg gag aag gtc aac ctg acg acc ggt gtc gg 1015
Gln Leu Thr Leu Ala Glu Lys Val Asn Leu Thr Thr Gly Val Gly
55 60 65

gtgagtcctat tgacctctac cgagcccccg ttccatgtcc attgagcaat tggctgacgt 1075

cttgaag c tgg atg cag gag aaa tgt gtc ggt gaa acg ggc agc att ccg 1125
Trp Met Gln Glu Lys Cys Val Gly Glu Thr Gly Ser Ile Pro
70 75 80

ag gtaggctcac ttcccaatgc cgctgcaaag gaggtgtcta aactggaata aatcag 1183
Arg

a ctg ggg ttc cgt gga ctg tgc ctc caa gac tcg ccc ctt ggt gtc aga 1232
Leu Gly Phe Arg Gly Leu Cys Leu Gln Asp Ser Pro Leu Gly Val Arg
85 90 95

ttt g gtaggtcttt caacagagaa caagggtcgt cgcgggagag atgctgatcg 1286
Phe
100

atacctacttt ttag ct gac tac gtt tct gcc ttc ccc gcc ggt gtc aat 1335
Ala Asp Tyr Val Ser Ala Phe Pro Ala Gly Val Asn
105 110

gtc gct gca acg tgg gat aag aac ctc gcc tac ctt cgt ggg aag gcg 1383
Val Ala Ala Thr Trp Asp Lys Asn Leu Ala Tyr Leu Arg Gly Lys Ala
115 120 125

atg ggt gag gaa cac cgt ggt aag ggc gtc gac gtc cag ctg gga cct 1431
Met Gly Glu Glu His Arg Gly Lys Gly Val Asp Val Gln Leu Gly Pro
130 135 140

gtc gcc ggc cct ctt ggc aga cac ccc gac ggt ggc aga aac tgg gag 1479
Val Ala Gly Pro Leu Gly Arg His Pro Asp Gly Gly Arg Asn Trp Glu
145 150 155 160

ggt ttc tct cct gac ccc gtc ctg acc ggt gtg ctt atg gcg gag acg 1527
Gly Phe Ser Pro Asp Pro Val Leu Thr Gly Val Leu Met Ala Glu Thr
165 170 175

atc aag ggt atc cag gat gcc ggt gtg att gct tgc gcc aag cac ttc 1575
Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Cys Ala Lys His Phe
180 185 190

att ggt aac gag atg gag cac ttc cgg caa gcc ggt gag gct gtt ggc 1623
Ile Gly Asn Glu Met Glu His Phe Arg Gln Ala Gly Glu Ala Val Gly
195 200 205

tat ggt ttc gat att acc gag agt gtc agc tca aat atc gac gac aag 1671
Tyr Gly Phe Asp Ile Thr Glu Ser Val Ser Ser Asn Ile Asp Asp Lys
210 215 220

acg ctt cac gag ctg tac ctt tgg ccc ttt gcg gat gct gtt cgc g 1717
Thr Leu His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg
225 230 235

gtaagcagtc cccccctcat aggtgattgt acatgtgtat ttctgactcg ctttcaaag 1776

ct ggc gtt ggt tcg ttc atg tgc tcc tac aac cag gtt aac aac agc 1823
Ala Gly Val Gly Ser Phe Met Cys Ser Tyr Asn Gln Val Asn Asn Ser
245 250 255

tac agc tgc tcg aac agc tac ctc cta aac aag ttg ctc aaa tcg gag 1871

ES 2 527 681 T3

Tyr	Ser	Cys	Ser	Asn 260	Ser	Tyr	Leu	Leu	Asn 265	Lys	Leu	Leu	Lys	Ser 270	Glu		
ctt	gat	ttt	cag	ggc	ttc	gtg	atg	agt	gac	tgg	gga	gcg	cac	cac	agc	1919	
Leu	Asp	Phe	Gln 275	Gly	Phe	Val	Met	Ser 280	Asp	Trp	Gly	Ala	His 285	His	Ser		
ggc	ggt	gga	gct	ggc	ctg	gct	ggc	ctt	gac	atg	tcg	atg	cca	gga	gac	1967	
Gly	Val	Gly 290	Ala	Ala	Leu	Ala	Gly 295	Leu	Asp	Met	Ser	Met 300	Pro	Gly	Asp		
acc	gcc	ttt	ggt	acc	ggc	aaa	tcc	ttc	tgg	gga	acc	aac	ctg	acc	atc	2015	
Thr	Ala 305	Phe	Gly	Thr	Gly	Lys 310	Ser	Phe	Trp	Gly	Thr 315	Asn	Leu	Thr	Ile		
gcc	ggt	ctc	aac	ggt	act	ggt	ccg	gaa	tgg	cgt	gtg	gat	gac	atg	gct	2063	
Ala	Val	Leu	Asn	Gly	Thr 325	Val	Pro	Glu	Trp	Arg 330	Val	Asp	Asp	Met	Ala 335		
ggt	cgc	atc	atg	gcg	ggc	ttt	tac	aag	ggt	ggt	cgc	gac	cgt	tac	cag	2111	
Val	Arg	Ile	Met	Ala 340	Ala	Phe	Tyr	Lys	Val 345	Gly	Arg	Asp	Arg	Tyr 350	Gln		
gtg	ccg	gtc	aac	ttc	gac	tcg	tgg	acg	aag	gat	gaa	tac	ggt	tac	gag	2159	
Val	Pro	Val	Asn 355	Phe	Asp	Ser	Trp	Thr 360	Lys	Asp	Glu	Tyr	Gly 365	Tyr	Glu		
cac	gca	ctg	ggt	ggc	cag	aac	tat	gtc	aag	gtc	aat	gac	aag	gtg	gat	2207	
His	Ala 370	Leu	Val	Gly	Gln	Asn	Tyr 375	Val	Lys	Val	Asn 380	Asp	Lys	Val	Asp		
ggt	cgt	gcc	gac	cat	gcg	gac	atc	atc	cgt	caa	att	ggg	tct	gct	agt	2255	
Val	Arg 385	Ala	Asp	His	Ala	Asp 390	Ile	Ile	Arg	Gln	Ile 395	Gly	Ser	Ala	Ser		
ggt	gtc	ctt	ctt	aag	aac	gat	gga	gga	ctc	cca	ttg	acc	ggc	tat	gaa	2303	
Val	Val	Leu	Leu	Lys	Asn 405	Asp	Gly	Gly	Leu	Pro 410	Leu	Thr	Gly	Tyr	Glu 415		
aag	ttc	acc	gga	ggt	ttt	gga	gag	gat	gcc	gga	tcg	aac	cgt	tgg	ggc	2351	
Lys	Phe	Thr	Gly 420	Val	Phe	Gly	Glu	Asp	Ala 425	Gly	Ser	Asn	Arg	Trp 430	Gly		
gct	gac	ggc	tgc	tct	gat	cgt	ggt	tgc	gac	aac	ggc	acg	ttg	gca	atg	2399	
Ala	Asp	Gly	Cys 435	Ser	Asp	Arg	Gly	Cys 440	Asp	Asn	Gly	Thr	Leu 445	Ala	Met		
ggt	tgg	ggc	agt	ggc	act	gct	gac	ttc	ccc	tac	ctt	gtc	act	ccc	gag	2447	
Gly	Trp	Gly 450	Ser	Gly	Thr	Ala	Asp 455	Phe	Pro	Tyr	Leu	Val 460	Thr	Pro	Glu		
cag	gca	atc	cag	aat	gaa	atc	ctt	tcc	aag	ggg	aag	ggg	tta			2489	
Gln	Ala 465	Ile	Gln	Asn	Glu	Ile 470	Leu	Ser	Lys	Gly	Lys 475	Gly	Leu				
gtgagtgtg	tcaccgacaa	tggtgccctt	gaccagatgg	aacaggttgc	gtctcaggcc											2549	
aggtattcct	tcctccgtat	ccctagcaat	cgaatctcca	ctgactttag	gac agc											2605	
					Asp	Ser											
ggt	tct	atc	ggt	ttc	gtc	aac	gcc	gac	tct	ggt	gaa	ggc	tac	atc	aac	2653	
Val	Ser	Ile	Val	Phe	Val 485	Asn	Ala	Asp	Ser	Gly 490	Glu	Gly	Tyr	Ile	Asn 495		
ggt	gat	ggc	aac	gaa	ggt	gat	cgg	aag	aac	ctc	acc	ctc	tgg	aaa	gga	2701	
Val	Asp	Gly	Asn	Glu 500	Gly	Asp	Arg	Lys	Asn 505	Leu	Thr	Leu	Trp	Lys 510	Gly		

ES 2 527 681 T3

ggc gag gag gtg atc aag act gtt gca gcc aac tgc aac aac acc att Gly Glu Glu Val Ile Lys Thr Val Ala Ala Asn Cys Asn Asn Thr Ile 515 520 525	2749
ggt gtg atg cac act gtg gga cct gtc ttg atc gat gag tgg tat gac Val Val Met His Thr Val Gly Pro Val Leu Ile Asp Glu Trp Tyr Asp 530 535 540	2797
aac ccc aac gtc acc gcc atc gtc tgg gcc ggt ctt cca ggc cag gag Asn Pro Asn Val Thr Ala Ile Val Trp Ala Gly Leu Pro Gly Gln Glu 545 550 555	2845
agc ggc aac agt ctc gtc gat gtg ctc tac ggc cgt gtc agc ccc gga Ser Gly Asn Ser Leu Val Asp Val Leu Tyr Gly Arg Val Ser Pro Gly 560 565 570 575	2893
gga aag acg ccg ttt acg tgg gga aag act cgc gag tcg tac ggc gct Gly Lys Thr Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr Gly Ala 580 585 590	2941
cct ctg ctc acc aaa ccc aac aac ggc aag ggt gct ccc cag gac gac Pro Leu Leu Thr Lys Pro Asn Asn Gly Lys Gly Ala Pro Gln Asp Asp 595 600 605	2989
ttc acc gag ggc gtc ttc atc gac tac aga agg ttc gac aag tac aac Phe Thr Glu Gly Val Phe Ile Asp Tyr Arg Arg Phe Asp Lys Tyr Asn 610 615 620	3037
gag acg ccc atc tat gag ttc ggg ttt ggt ctg agt tat act act ttt Glu Thr Pro Ile Tyr Glu Phe Gly Phe Gly Leu Ser Tyr Thr Thr Phe 625 630 635	3085
gaa tac tcg aac atc tac gtc cag ccc ctt aac gca cga cct tac acc Glu Tyr Ser Asn Ile Val Glu Gln Pro Leu Asn Ala Arg Pro Tyr Thr 640 645 650 655	3133
cca gcc tcc ggc agc acc aag gcg gct cct acc ttt ggg aat atc agc Pro Ala Ser Gly Ser Thr Lys Ala Ala Pro Thr Phe Gly Asn Ile Ser 660 665 670	3181
acg gac tat gca gat tac ttg tac cct gag gat ata cac aag gtc cca Thr Asp Tyr Ala Asp Tyr Leu Tyr Pro Glu Asp Ile His Lys Val Pro 675 680 685	3229
tta tac atc tat cct tgg ctt aac acg acg gac ccc gaa gaa gtc ctc Leu Tyr Ile Tyr Pro Trp Leu Asn Thr Thr Asp Pro Glu Glu Val Leu 690 695 700	3277
cgg cga tcc cga ctt acg gaa atg aag gcc gag gac tac atc cca tct Arg Arg Ser Arg Leu Thr Glu Met Lys Ala Glu Asp Tyr Ile Pro Ser 705 710 715	3325
ggc gcg act gat gga tct cct cag ccc atc ctt ccg gca ggc ggt gct Gly Ala Thr Asp Gly Ser 725 Pro Gln Pro Ile Leu Pro Ala Gly Gly Ala 720 730 735	3373
cct ggt ggc aac ccg ggt ctc tat gat gag atg tac agg gta tct gca Pro Gly Gly Asn Pro Gly Leu Tyr Asp Glu Met Tyr Arg Val Ser Ala 740 745 750	3421
atc atc acc aac acc ggt aac gtt gtt ggt gat gag gtt cct cag ctg Ile Ile Thr Asn Thr Gly Asn Val Val Gly Asp Glu Val Pro Gln Leu 755 760 765	3469
gtgagtttcg cagtctcatt gatatatgtc ttctcaggttg gtcactgacc cgcgatctat	3529
ag tat gtc tct ctt ggt ggt cca gat gac ccc aag gtc gtg ctc cgc	3576

ES 2 527 681 T3

Tyr	Val	Ser	Leu	Gly	Gly	Pro	Asp	Asp	Pro	Lys	Val	Val	Leu	Arg	
		770					775					780			
aac	ttt	gac	cg	atc	acg	ctc	cac	ccc	ggc	caa	cag	aca	atg	tgg	acc
Asn	Phe	Asp	Arg	Ile	Thr	Leu	His	Pro	Gly	Gln	Gln	Thr	Met	Trp	Thr
		785					790					795			
acg	aca	ttg	acg	cga	cg	gat	atc	tcg	aac	tgg	gac	cct	gcc	tcc	cag
Thr	Thr	Leu	Thr	Arg	Arg	Asp	Ile	Ser	Asn	Trp	Asp	Pro	Ala	Ser	Gln
		800				805					810				
aat	tgg	gtt	gtg	acc	aaa	tat	ccc	aag	aca	gtc	tac	atc	ggc	agc	tct
Asn	Trp	Val	Val	Thr	Lys	Tyr	Pro	Lys	Thr	Val	Tyr	Ile	Gly	Ser	Ser
		815			820					825					830
tcg	cgg	aaa	ctg	cac	ctg	cag	gca	ccg	ctt	ccc	cct	tac	tgaggtttta		
Ser	Arg	Lys	Leu	His	Leu	Gln	Ala	Pro	Leu	Pro	Pro	Tyr			
				835					840						
tccggaagga	ggaagtaaaa	acacaatggt	ttagttgtac	aggcgtcttt	cgtttgtgat										3829
tatccatagg	catatcaaga	ccactttggg	ttatatatat	atatatatat	ataagcggcc										3889
gaggaaaggc	aatgggtagc	atgggtcaag	gggaggaacc	gtcttgaaac	tactctcaat										3949
ttctttcagt	agatagtgc	ctccggttga	gtcccaaata	tagttttaat	aatggtaaat										4009
ggttcagaaa	aagaaaatgt	agaggtttca	aacacgctag	ttgaccctga	taggaattga										4069
gcatgaatgc	ctacacattc	caagtcgtgt	tagcagtcg	atagccgatg	aacctattcc										4129
gtaggttgag	gttcacccta	caaataagcc	aggatttaag	taaataacctg	ctcgtgaaat										4189
ctacaacgca	tcagatcaga	ggaaaattca	aatggcagaa	gtgcgagcac	ctcggtgaga										4249
agagatcgag	ctgtcgaagt	cggctggaac	acaggtaaaag	agaagtaata	caattcattg										4309
atttttacat	cgtttaacat	gtagaaggta	tctaaaatag	taagtccaga	tatgggcat										4369
ggagatcgcc	tcggcgatct	tcgggagtat	ctcgggagac	gcacatgacc	gcgcttaacc										4429
ctgtcgggtg	gacccgagtc	cgaccgacgt	catcagcgca	gcgcagggtca	ggctgcgcgc										4489
aacgtcaatg	ccagggggtg	ctgggacagt	tgcatatcaa	tcgatcagtc	aattaaagca										4549
tctgctttcc	acgttctttt	tttatcacct	ttcacttccc	ctgtcccaact	tgcttgagg										4609
ttgttgagcc	caaagaagaa	ggagaagaaa	atgggctcga	caccccgga	cgggtggtcg										4669
acgagcacat	catcagcagc	gtcttattat	caacattccc	aaccaccggc	cctcgttctc										4729
ctcgtctacc	cgctcactct	cctcctcggc	tccctgtaca	gagccatttc	ccccaccgcg										4789
cgggtgaggc	acgatgctgc	agaccctgct	ctggccccga	ccatagcgtc	cgacatcaac										4849
ctgtcccagt	catcccggta	ttcccattcc	catagcaaca	gcaacagccc	ggtcaattac										4909
ttcggccgca	aggacaacat	ctttaacgtc	tacttcgtca	agatcggtcg	gttctggacg										4969
accctgcct	tcctcacgtt	actcctcacc	cagcctgcct	acacaaacgc	cgggtcccctg										5029
cgcgcccgac	gcaccctcca	agccctgtcc	cgctacgcca	tcgtcacctc	actacctgga										5089
tcc															5092

<210> 22

<211> 843

<212> PRT

<213> Thermoascus aurantiacus

<400> 22

ES 2 527 681 T3

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

ES 2 527 681 T3

Gly Val Gly Ser Phe Met Cys Ser Tyr Asn Gln Val Asn Asn Ser Tyr
 245 250 255
 Ser Cys Ser Asn Ser Tyr Leu Leu Asn Lys Leu Leu Lys Ser Glu Leu
 260 265 270
 Asp Phe Gln Gly Phe Val Met Ser Asp Trp Gly Ala His His Ser Gly
 275 280 285
 Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Thr
 290 295 300
 Ala Phe Gly Thr Gly Lys Ser Phe Trp Gly Thr Asn Leu Thr Ile Ala
 305 310 315 320
 Val Leu Asn Gly Thr Val Pro Glu Trp Arg Val Asp Asp Met Ala Val
 325 330 335
 Arg Ile Met Ala Ala Phe Tyr Lys Val Gly Arg Asp Arg Tyr Gln Val
 340 345 350
 Pro Val Asn Phe Asp Ser Trp Thr Lys Asp Glu Tyr Gly Tyr Glu His
 355 360 365
 Ala Leu Val Gly Gln Asn Tyr Val Lys Val Asn Asp Lys Val Asp Val
 370 375 380
 Arg Ala Asp His Ala Asp Ile Ile Arg Gln Ile Gly Ser Ala Ser Val
 385 390 395 400
 Val Leu Leu Lys Asn Asp Gly Gly Leu Pro Leu Thr Gly Tyr Glu Lys
 405 410 415
 Phe Thr Gly Val Phe Gly Glu Asp Ala Gly Ser Asn Arg Trp Gly Ala
 420 425 430
 Asp Gly Cys Ser Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met Gly
 435 440 445
 Trp Gly Ser Gly Thr Ala Asp Phe Pro Tyr Leu Val Thr Pro Glu Gln
 450 455 460
 Ala Ile Gln Asn Glu Ile Leu Ser Lys Gly Lys Gly Leu Asp Ser Val
 465 470 475 480
 Ser Ile Val Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile Asn Val
 485 490 495
 Asp Gly Asn Glu Gly Asp Arg Lys Asn Leu Thr Leu Trp Lys Gly Gly
 500 505 510

ES 2 527 681 T3

Glu Glu Val Ile Lys Thr Val Ala Ala Asn Cys Asn Asn Thr Ile Val
 515 520 525
 Val Met His Thr Val Gly Pro Val Leu Ile Asp Glu Trp Tyr Asp Asn
 530 535 540
 Pro Asn Val Thr Ala Ile Val Trp Ala Gly Leu Pro Gly Gln Glu Ser
 545 550 555 560
 Gly Asn Ser Leu Val Asp Val Leu Tyr Gly Arg Val Ser Pro Gly Gly
 565 570 575
 Lys Thr Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr Gly Ala Pro
 580 585 590
 Leu Leu Thr Lys Pro Asn Asn Gly Lys Gly Ala Pro Gln Asp Asp Phe
 595 600 605
 Thr Glu Gly Val Phe Ile Asp Tyr Arg Arg Phe Asp Lys Tyr Asn Glu
 610 615 620
 Thr Pro Ile Tyr Glu Phe Gly Phe Gly Leu Ser Tyr Thr Thr Phe Glu
 625 630 635 640
 Tyr Ser Asn Ile Tyr Val Gln Pro Leu Asn Ala Arg Pro Tyr Thr Pro
 645 650 655
 Ala Ser Gly Ser Thr Lys Ala Ala Pro Thr Phe Gly Asn Ile Ser Thr
 660 665 670
 Asp Tyr Ala Asp Tyr Leu Tyr Pro Glu Asp Ile His Lys Val Pro Leu
 675 680 685
 Tyr Ile Tyr Pro Trp Leu Asn Thr Thr Asp Pro Glu Glu Val Leu Arg
 690 695 700
 Arg Ser Arg Leu Thr Glu Met Lys Ala Glu Asp Tyr Ile Pro Ser Gly
 705 710 715 720
 Ala Thr Asp Gly Ser Pro Gln Pro Ile Leu Pro Ala Gly Gly Ala Pro
 725 730 735
 Gly Gly Asn Pro Gly Leu Tyr Asp Glu Met Tyr Arg Val Ser Ala Ile
 740 745 750
 Ile Thr Asn Thr Gly Asn Val Val Gly Asp Glu Val Pro Gln Leu Tyr
 755 760 765
 Val Ser Leu Gly Gly Pro Asp Asp Pro Lys Val Val Leu Arg Asn Phe
 770 775 780
 Asp Arg Ile Thr Leu His Pro Gly Gln Gln Thr Met Trp Thr Thr Thr
 785 790 795 800
 Leu Thr Arg Arg Asp Ile Ser Asn Trp Asp Pro Ala Ser Gln Asn Trp
 805 810 815
 Val Val Thr Lys Tyr Pro Lys Thr Val Tyr Ile Gly Ser Ser Ser Arg
 820 825 830
 Lys Leu His Leu Gln Ala Pro Leu Pro Pro Tyr
 835 840

<210> 23
 <211> 3510
 <212> ADN
 <213> Acremonium thermophilum
 5
 <220>
 <221> CDS
 <222> (391)..(447)
 <223>
 10
 <220>
 <221> Intrón
 <222> (448)..(539)
 <223>
 15
 <220>
 <221> CDS
 <222> (540)..(685)
 <223>
 20
 <220>
 <221> Intrón
 <222> (686)..(759)
 <223>
 25
 <220>
 <221> CDS
 <222> (760)..(1148)
 <223>
 30
 <220>
 <221> Intrón
 <222> (1149)..(1217)
 <223>
 35
 <220>
 <221> CDS
 <222> (1218)..(3208)
 <223>
 40
 <400> 23

gcaggtagct acgacattcg acggtccacg cccagtggcg tctgctcggc cgtctgggaa	60
ccatgcacgc cgcctcttta ggtcgagcga ggtataacat actatctgca cggctaccta	120
tatattacgt cgatgtcacc cgcaggatgc gagcaccatt acttcgtgtc tcacccgccc	180
ttccgctccg catctcgtga acctaaaccc acgcgggcac actgcttctt gtgagagcct	240
ctacccgttc cacaagagcc atagctagag agagaagggc agccaaggga ccggtcaagc	300
ggcgctcttc atcgaccaa tctcgacaac ccggcagacg tcaccaccgg ctcccgccgc	360
acgacgtcac acgggactga ctacgaagac atg agg cag gcc ctt gtt tcg ctg	414
Met Arg Gln Ala Leu Val Ser Leu	
1 5	
gcc ttg ctg gcc agc agc cct gtt tcg gcg gcg gtgaccgcca gggacgccca	467
Ala Leu Leu Ala Ser Ser Pro Val Ser Ala Ala	
10 15	
ggtatgggcc caactgctct tcctccctgt ttctctctct accggtgctg acaacgacaa	527
tagctgcacc ag cga gaa ctc gcc act tcc gac cct ttc tat cct tcg cca	578
Arg Glu Leu Ala Thr Ser Asp Pro Phe Tyr Pro Ser Pro	
20 25 30	
tgg atg aac cct gaa gcc aat ggc tgg gag gac gcc tac gcc aag gcc	626
Trp Met Asn Pro Glu Ala Asn Gly Trp Glu Asp Ala Tyr Ala Lys Ala	
35 40 45	
aag gcg ttc gtt tcc cag ctg acg ctc ttg gaa aag gtc aac ctg acg	674
Lys Ala Phe Val Ser Gln Leu Thr Leu Leu Glu Lys Val Asn Leu Thr	
50 55 60	
act ggc atc gg gtgagtcttg ttctctcctg tagaaccgcc taccagaaga	725
Thr Gly Ile Gly	
65	

cattcaggaa	gtgctaata	tgggcggttg	acag	c	tgg	caa	gga	gga	caa	tgc		778
					Trp	Gln	Gly	Gly	Gln	Cys		
					70							
gtg	ggc	aac	gtc	ggt	tcc	gtc	ccg	cgt	ctc	ggc	ctt	826
Val	Gly	Asn	Val	Gly	Ser	Val	Pro	Arg	Leu	Gly	Leu	
75					80					85		
atg	cag	gac	tcc	ccc	gtg	ggt	atc	cgc	ttt	ggg	gac	874
Met	Gln	Asp	Ser	Pro	Val	Gly	Ile	Arg	Phe	Gly	Asp	
				95					100		Tyr	
											gtc	
											Ser	
											Val	
											105	
											gtc	
											Val	
ttc	ccc	tct	ggt	cag	acc	acg	gct	gcc	acc	ttc	gac	922
Phe	Pro	Ser	Gly	Gln	Thr	Thr	Ala	Ala	Thr	Phe	Asp	
			110					115			Aas	
											Lys	
											ggt	
											Gly	
											120	
											ctg	
											Leu	
											Met	
aac	cgt	cgc	ggc	aat	gcc	atg	ggc	cag	gag	cac	aaa	970
Asn	Arg	Arg	Gly	Asn	Ala	Met	Gly	Gln	Glu	His	Lys	
		125					130				gga	
											Gly	
											135	
											Lys	
											ggt	
											Gly	
											Val	
aac	gtc	ctg	ctc	ggc	ccg	gtc	gct	ggc	ccc	att	ggc	1018
Asn	Val	Leu	Leu	Gly	Pro	Val	Ala	Gly	Pro	Ile	Gly	
	140					145					150	
											cgt	
											Arg	
											Thr	
											ccc	
											Pro	
											gag	
											Glu	
ggg	gga	cga	aac	tgg	gag	ggc	ttc	tcc	ccc	gac	ccc	1066
Gly	Gly	Arg	Asn	Trp	Glu	Gly	Phe	Ser	Pro	Asp	Pro	
155					160					165		
											gtc	
											Val	
											cta	
											Leu	
											acg	
											Thr	
											ggt	
											Gly	
											170	
att	gcc	ttg	gcc	gaa	acg	atc	aag	gga	atc	cag	gat	1114
Ile	Ala	Leu	Ala	Glu	Thr	Ile	Lys	Gly	Ile	Gln	Asp	
				175					180		Ala	
											gct	
											Gly	
											Val	
											185	
											gtc	
											Ile	
											att	
											Ile	
gct	tgc	gcc	aag	cat	ttc	atc	gcg	aac	gaa	cag	g	1158
Ala	Cys	Ala	Lys	His	Phe	Ile	Ala	Asn	Glu	Gln	gtg	
			190					195			cgtgatg	
gaacgcggga	cgtgctctga	tgcaaaccga	cgagcactga	ccacgctttc	ctcgaacag							1217
aa	cac	ttc	cgc	cag	tcc	ggc	gag	gcc	cag	ggc	tac	1264
Glu	His	Phe	Arg	Gln	Ser	Gly	Glu	Ala	Gln	Gly	Tyr	
		200					205				ggc	
											Gly	
											210	
											ttt	
											Phe	
											gac	
											Asp	
											atc	
											Ile	
tcc	gag	tcg	ctg	tcg	tcc	aac	atc	gac	gac	aag	acc	1312
Ser	Glu	Ser	Leu	Ser	Ser	Asn	Ile	Asp	Asp	Lys	Thr	
						220					225	
											atg	
											Met	
											His	
											cag	
											Glu	
											gag	
											Glu	
											ctg	
											Leu	
tat	ctg	tgg	ccc	ttc	gcc	gac	ggc	gtg	cgt	gcc	ggc	1360
Tyr	Leu	Trp	Pro	Phe	Ala	Asp	Gly	Val	Arg	Ala	Gly	
230					235					240		
											gtc	
											Val	
											Gly	
											245	
											gcc	
											Ala	
											atc	
											Ile	
											245	
atg	tgc	tcg	tac	aac	cag	atc	aac	aac	tcg	tac	ggg	1408
Met	Cys	Ser	Tyr	Asn	Gln	Ile	Asn	Asn	Ser	Tyr	Gly	
				250					255		255	
											tgc	
											Cys	
											cag	
											Gln	
											aac	
											Asn	
											260	
											tcc	
											Ser	
aag	acc	ctg	aac	aac	ctg	ctc	aag	aac	gag	ctc	ggc	1456
Lys	Thr	Leu	Asn	Asn	Leu	Leu	Lys	Asn	Glu	Leu	Gly	
			265					270			ttc	
											Phe	
											cag	
											Gln	
											ggc	
											Gly	
											275	
											gtc	
											Val	
gtc	atg	agc	gac	tgg	cag	gcc	cag	cac	acc	ggc	gcg	1504
Val	Met	Ser	Asp	Trp	Gln	Ala	Gln	His	Thr	Gly	Ala	
		280					285				290	
											agg	
											Ala	
											ser	
											gcc	
											Ala	
											gtc	
											Val	
gcc	ggc	ctg	gac	atg	acc	atg	ccc	ggc	gac	acc	agc	1552
Ala	Gly	Leu	Asp	Met	Thr	Met	Pro	Gly	Asp	Thr	Ser	
						300					305	
											ttc	
											Phe	
											aac	
											Asn	
											acc	
											Thr	
											ggc	
											Gly	
ctc	agc	tac	tgg	ggc	acg	aac	ctc	acc	ctc	gcc	gtc	1600
											ctg	
											aac	
											ggc	
											acc	

ES 2 527 681 T3

Leu 310	Ser	Tyr	Trp	Gly	Thr 315	Asn	Leu	Thr	Leu	Ala 320	Val	Leu	Asn	Gly	Thr 325	
gtc Val	ccc Pro	gag Glu	tac Tyr	cgc Arg 330	atc Ile	gac Asp	gac Asp	atg Met	gtc Val 335	atg Met	cgc Arg	atc Ile	atg Met	gcc Ala 340	gcc Ala	1648
ttc Phe	ttc Phe	aag Lys	acc Thr 345	ggc Gly	cag Gln	acc Thr	ctg Leu	gac Asp 350	ctg Leu	ccg Pro	ccc Pro	atc Ile	aac Asn 355	ttc Phe	gac Asp	1696
tcg Ser	tgg Trp	acc Thr 360	acc Thr	gac Asp	acc Thr	ttc Phe	ggc Gly 365	ccg Pro	ctc Leu	cac His	ttc Phe	gcc Ala 370	gtc Val	aac Asn	gag Glu	1744
gac Asp	cgc Arg 375	cag Gln	cag Gln	atc Ile	aac Asn	tgg Trp 380	cac His	gtc Val	aac Asn	gtc Val	cag Gln 385	gac Asp	aac Asn	cat His	ggc Gly	1792
agc Ser 390	ctc Leu	atc Ile	cgc Arg	gag Glu	atc Ile 395	gcg Ala	gcc Ala	aag Lys	gga Gly	acc Thr 400	gtc Val	ctg Leu	ctg Leu	aag Lys	aac Asn 405	1840
acc Thr	ggg Gly	tcc Ser	ctc Leu	ccg Pro 410	ctc Leu	aac Asn	aag Lys	ccc Pro	aag Lys 415	ttc Phe	ctc Leu	gtc Val	gtg Val	gtc Val 420	ggc Gly	1888
gac Asp	gac Asp	gcg Ala	ggc Gly 425	ccc Pro	aac Asn	ccg Pro	gcg Ala	gga Gly 430	ccc Pro	aac Asn	gcc Ala	tgc Cys	ccc Pro 435	gac Asp	cgc Arg	1936
gga Gly	tgc Cys	gac Ala 440	gtc Val	ggc Gly	acc Thr	ctc Leu	ggc Gly 445	atg Met	gcc Ala	tgg Trp	ggc Gly	tcc Ser 450	ggc Gly	tcg Ser	gcc Ala	1984
aac Asn 455	ttc Phe	ccc Pro	tac Tyr	ctg Leu	atc Ile	acc Thr 460	ccg Pro	gac Asp	gcc Ala	gcg Ala	ctg Leu 465	cag Gln	gcg Ala	cag Gln	gcg Ala	2032
atc Ile 470	aag Lys	gac Asp	ggc Gly	acc Thr	cgc Arg 475	tac Tyr	gag Glu	agc Ser	gtg Val	ctg Leu 480	tcc Ser	aac Asn	tac Tyr	cag Gln	ctc Leu 485	2080
gac Asp	cag Gln	acc Thr	aag Lys	gcg Ala 490	ctg Leu	gtc Val	acc Thr	cag Gln	gcc Ala 495	aac Asn	gcc Ala	acg Thr	gcc Ala	atc Ile 500	gtc Val	2128
ttc Phe	gtc Val	aac Asn	gcc Ala 505	gac Asp	tcg Ser	ggc Gly	gag Glu	ggc Gly 510	tac Tyr	atc Ile	aac Asn	gtc Val	gac Asp 515	ggc Gly	aac Asn	2176
gag Glu	ggc Gly	gac Asp 520	cgc Arg	aag Lys	aac Asn	ctc Leu	acg Thr 525	ctc Leu	tgg Trp	cac His	gac Asp	ggc Gly 530	gac Asp	gcc Ala	ctg Leu	2224
atc Ile 535	aag Lys	agc Ser	gtg Val	gcc Ala	ggc Gly	tgg Trp 540	aac Asn	ccg Pro	aac Asn	acc Thr	atc Ile 545	gtc Val	gtc Val	atc Ile	cac His	2272
tcg Ser 550	acc Thr	ggc Gly	ccc Pro	gtc Val	ctc Leu 555	gtg Val	acc Thr	gac Asp	tgg Trp 560	tac Tyr	gac Asp	cac His	ccc Pro	aac Asn	atc Ile 565	2320
acc Thr	gcc Ala	atc Ile	ctg Leu	tgg Trp 570	gcc Ala	ggc Gly	gtg Val	ccc Pro	ggg Gly 575	cag Gln	gag Glu	tcc Ser	ggc Gly	aac Asn 580	gcc Ala	2368
atc	acc	gac	gtc	ctc	tac	gga	aaa	gtc	aac	ccg	tcg	ggc	cgc	agc	ccc	2416

Ile Thr Asp Val Leu Tyr Gly Lys Val Asn Pro Ser Gly Arg Ser Pro	
585 590 595	
ttc acc tgg ggt ccg acc cgc gag agc tac ggc acc gac gtc ctc tac	2464
Phe Thr Trp Gly Pro Thr Arg Glu Ser Tyr Gly Thr Asp Val Leu Tyr	
600 605 610	
act ccc aac aac ggc aag ggc gcg ccg cag cag gcc ttc tcc gag ggc	2512
Thr Pro Asn Asn Gly Lys Gly Ala Pro Gln Gln Ala Phe Ser Glu Gly	
615 620 625	
gtc ttc atc gac tac cgc cac ttc gac cgc acc aac gcg tcc gtc atc	2560
Val Phe Ile Asp Tyr Arg His Phe Asp Arg Thr 640 Asn Ala Ser Val Ile	
630 635 645	
tac gag ttc ggc cac ggc ctc agc tac acg acg ttc cag tac agc aac	2608
Tyr Glu Phe Gly His 650 Gly Leu Ser Tyr Thr 655 Phe Gln Tyr Ser Asn	
660	
atc cag gtg gtc aag tcc aac gcc ggc gcg tac aag ccc acg acg ggc	2656
Ile Gln Val Val Lys Ser Asn Ala Gly Ala Tyr Lys Pro Thr Thr Gly	
665 670 675	
acg acc atc ccc gcg ccc acg ttt ggc agc ttc tcc aag gac ctc aag	2704
Thr Thr Ile Pro Ala Pro Thr Phe Gly Ser Phe Ser Lys Asp Leu Lys	
680 685 690	
gac tac ctc ttc ccg tcg gac cag ttc cgc tac atc acc cag tac atc	2752
Asp Tyr 695 Leu Phe Pro Ser Asp Gln Phe Arg Tyr Ile Thr Gln Tyr Ile	
700 705	
tac ccg tac ctc aac tcc acc gac ccg gcc aag gcg tcg ctc gac ccg	2800
Tyr Pro Tyr Leu Asn Ser 715 Thr Asp Pro Ala Lys 720 Ala Ser Leu Asp Pro	
710 715 725	
cac tac ggc aag acg gcg gcc gag ttt ctg ccg ccg cac gcg ctg gac	2848
His Tyr Gly Lys Thr 730 Ala Ala Glu Phe Leu Pro Pro His Ala Leu Asp	
735 740	
agc aac ccg cag ccg ctg ctg cgg tcg tcg ggc aag aac gag ccc ggc	2896
Ser Asn Pro Gln Pro Leu Leu Arg Ser 750 Ser Gly Lys Asn Glu Pro Gly	
745 755	
ggc aac cgc cag ctg tac gac atc ctg tac acg gtg acg gcg gac atc	2944
Gly Asn Arg Gln Leu Tyr Asp Ile Leu Tyr Thr Val Thr 770 Ala Asp Ile	
760 765 770	
acc aac acg ggc agc atc gtg ggt gcg gag gtg ccg cag ctg tac gtg	2992
Thr Asn Thr Gly Ser Ile Val 780 Gly Ala Glu Val Pro Gln Leu Tyr Val	
775 785	
tcg ctg ggc ggc ccc gac gac ccc aaa gtg gtc ctg cgc ggc ttc gac	3040
Ser Leu Gly Gly Pro Asp Asp Pro Lys Val Val Leu Arg Gly Phe Asp	
790 795 800 805	
cgc atc cgc atc gac ccg ggc aag acg gcg cag ttc cgc gtc acc ctg	3088
Arg Ile Arg Ile Asp Pro Gly Lys Thr Ala Gln Phe Arg Val Thr Leu	
810 815 820	
acc cgc cgg gat ctc agc aac tgg gac ccg gcg atc cag gac tgg gtc	3136
Thr Arg Arg Asp Leu Ser Asn Trp Asp Pro Ala Ile Gln Asp Trp Val	
825 830 835	
atc agc aag tac ccc aag aag gtg tac atc ggc cgg agc agc agg aag	3184
Ile Ser Lys Tyr Pro Lys Lys Val Tyr Ile Gly Arg Ser Ser Arg Lys	
840 845 850	
ctg gaa ctc tcc gcc gac ctc gcg tgatccggcg acggccaagt acgtatgtgg	3238
Leu Glu Leu Ser Ala Asp Leu Ala	
855 860	
actgccatcc gaacacctat actttttggc taggtagggg gagcagcaag gcctgagcat	3298
atactctctc cattgcacat ttctaattgta aatatatata tcattaattg ggagacccaa	3358
actcgaattt atgcatgcgt acaaagtgtg ttgaacaagt ttcggtccag cagatagtaa	3418
ccgtcttagt tcgtccatcc ctctctcgaa tgcgctgtat acacatgcgt atagacgt	3478
tgtatagggtg ccattgctag caatgcaagc tt	3510
<210> 24	

<211> 861
 <212> PRT
 <213> Acremonium thermophilum

5 <400> 24
 Met Arg Gln Ala Leu Val Ser Leu Ala Leu Leu Ala Ser Ser Pro Val
 1 5 10 15
 Ser Ala Ala Arg Glu Leu Ala Thr Ser Asp Pro Phe Tyr Pro Ser Pro
 20 25 30
 Trp Met Asn Pro Glu Ala Asn Gly Trp Glu Asp Ala Tyr Ala Lys Ala
 35 40 45
 Lys Ala Phe Val Ser Gln Leu Thr Leu Leu Glu Lys Val Asn Leu Thr
 50 55 60
 Thr Gly Ile Gly Trp Gln Gly Gly Gln Cys Val Gly Asn Val Gly Ser
 65 70 75 80
 Val Pro Arg Leu Gly Leu Arg Ser Leu Cys Met Gln Asp Ser Pro Val
 85 90 95
 Gly Ile Arg Phe Gly Asp Tyr Val Ser Val Phe Pro Ser Gly Gln Thr
 100 105 110
 Thr Ala Ala Thr Phe Asp Lys Gly Leu Met Asn Arg Arg Gly Asn Ala
 115 120 125
 Met Gly Gln Glu His Lys Gly Lys Gly Val Asn Val Leu Leu Gly Pro
 130 135 140
 Val Ala Gly Pro Ile Gly Arg Thr Pro Glu Gly Gly Arg Asn Trp Glu
 145 150 155 160
 Gly Phe Ser Pro Asp Pro Val Leu Thr Gly Ile Ala Leu Ala Glu Thr
 165 170 175

Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Cys Ala Lys His Phe
180 185 190

Ile Ala Asn Glu Gln Glu His Phe Arg Gln Ser Gly Glu Ala Gln Gly
195 200 205

Tyr Gly Phe Asp Ile Ser Glu Ser Leu Ser Ser Asn Ile Asp Asp Lys
210 215 220

Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Gly Val Arg Ala
225 230 235 240

Gly Val Gly Ala Ile Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr
245 250 255

Gly Cys Gln Asn Ser Lys Thr Leu Asn Asn Leu Leu Lys Asn Glu Leu
260 265 270

Gly Phe Gln Gly Phe Val Met Ser Asp Trp Gln Ala Gln His Thr Gly
275 280 285

Ala Ala Ser Ala Val Ala Gly Leu Asp Met Thr Met Pro Gly Asp Thr
290 295 300

Ser Phe Asn Thr Gly Leu Ser Tyr Trp Gly Thr Asn Leu Thr Leu Ala
305 310 315 320

Val Leu Asn Gly Thr Val Pro Glu Tyr Arg Ile Asp Asp Met Val Met
325 330 335

Arg Ile Met Ala Ala Phe Phe Lys Thr Gly Gln Thr Leu Asp Leu Pro
340 345 350

Pro Ile Asn Phe Asp Ser Trp Thr Thr Asp Thr Phe Gly Pro Leu His
355 360 365

Phe Ala Val Asn Glu Asp Arg Gln Gln Ile Asn Trp His Val Asn Val
370 375 380

Gln Asp Asn His Gly Ser Leu Ile Arg Glu Ile Ala Ala Lys Gly Thr
385 390 395 400

Val Leu Leu Lys Asn Thr Gly Ser Leu Pro Leu Asn Lys Pro Lys Phe
405 410 415

Leu Val Val Val Gly Asp Asp Ala Gly Pro Asn Pro Ala Gly Pro Asn
420 425 430

Ala Cys Pro Asp Arg Gly Cys Asp Val Gly Thr Leu Gly Met Ala Trp
435 440 445

ES 2 527 681 T3

Gly Ser Gly Ser Ala Asn Phe Pro Tyr Leu Ile Thr Pro Asp Ala Ala
 450 455 460
 Leu Gln Ala Gln Ala Ile Lys Asp Gly Thr Arg Tyr Glu Ser Val Leu
 465 470 475 480
 Ser Asn Tyr Gln Leu Asp Gln Thr Lys Ala Leu Val Thr Gln Ala Asn
 485 490 495
 Ala Thr Ala Ile Val Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile
 500 505 510
 Asn Val Asp Gly Asn Glu Gly Asp Arg Lys Asn Leu Thr Leu Trp His
 515 520 525
 Asp Gly Asp Ala Leu Ile Lys Ser Val Ala Gly Trp Asn Pro Asn Thr
 530 535 540
 Ile Val Val Ile His Ser Thr Gly Pro Val Leu Val Thr Asp Trp Tyr
 545 550 555 560
 Asp His Pro Asn Ile Thr Ala Ile Leu Trp Ala Gly Val Pro Gly Gln
 565 570 575
 Glu Ser Gly Asn Ala Ile Thr Asp Val Leu Tyr Gly Lys Val Asn Pro
 580 585 590
 Ser Gly Arg Ser Pro Phe Thr Trp Gly Pro Thr Arg Glu Ser Tyr Gly
 595 600 605
 Thr Asp Val Leu Tyr Thr Pro Asn Asn Gly Lys Gly Ala Pro Gln Gln
 610 615 620
 Ala Phe Ser Glu Gly Val Phe Ile Asp Tyr Arg His Phe Asp Arg Thr
 625 630 635 640
 Asn Ala Ser Val Ile Tyr Glu Phe Gly His Gly Leu Ser Tyr Thr Thr
 645 650 655
 Phe Gln Tyr Ser Asn Ile Gln Val Val Lys Ser Asn Ala Gly Ala Tyr
 660 665 670
 Lys Pro Thr Thr Gly Thr Thr Ile Pro Ala Pro Thr Phe Gly Ser Phe
 675 680 685
 Ser Lys Asp Leu Lys Asp Tyr Leu Phe Pro Ser Asp Gln Phe Arg Tyr
 690 695 700
 Ile Thr Gln Tyr Ile Tyr Pro Tyr Leu Asn Ser Thr Asp Pro Ala Lys
 705 710 715 720

ES 2 527 681 T3

Ala Ser Leu Asp Pro His Tyr Gly Lys Thr Ala Ala Glu Phe Leu Pro
725 730 735
Pro His Ala Leu Asp Ser Asn Pro Gln Pro Leu Leu Arg Ser Ser Gly
740 745 750
Lys Asn Glu Pro Gly Gly Asn Arg Gln Leu Tyr Asp Ile Leu Tyr Thr
755 760 765
Val Thr Ala Asp Ile Thr Asn Thr Gly Ser Ile Val Gly Ala Glu Val
770 775 780
Pro Gln Leu Tyr Val Ser Leu Gly Gly Pro Asp Asp Pro Lys Val Val
785 790 795 800
Leu Arg Gly Phe Asp Arg Ile Arg Ile Asp Pro Gly Lys Thr Ala Gln
805 810 815
Phe Arg Val Thr Leu Thr Arg Arg Asp Leu Ser Asn Trp Asp Pro Ala
820 825 830
Ile Gln Asp Trp Val Ile Ser Lys Tyr Pro Lys Lys Val Tyr Ile Gly
835 840 845
Arg Ser Ser Arg Lys Leu Glu Leu Ser Ala Asp Leu Ala
850 855 860

<210> 25
<211> 3392
5 <212> ADN
<213> Chaetomium thermophilum

<220>
<221> CDS
10 <222> (608)..(2405)
<223>

<220>
<221> Intrón
15 <222> (2406)..(2457)
<223>

<220>
<221> CDS
20 <222> (2458)..(2861)
<223>

<400> 25

tgcggggttg ctgcgactta attaataact ggcaaacg cccggagctc agctctgacc	60
tccgccacat ccgctcggca ccatgccagc gcgttgcaac ggcataaagc gctcaggttt	120
ttcttccgcc tgctccccac tgccgatggc catctgcacc ccagctcgtc acatttatct	180
cgcgcacagc gtcttccccac cagttgcctt gctcatgacg ctgttaaaga tggccctacc	240
tagccgctga gtccacaac gccgagatgt ttttgccct ttacaaggca cgccatggcc	300
gtccaaggctc tgttcatgag tgtgtttgtg gggccgaagg acacctcagt ggccacgaaa	360
tgccgccgag cgggccagca catgtcgaga gagacatgga catttatccc cgagatgctg	420
tattagggaa ccggtccttt tctcggagcc gtgatccgag agcgttcggg agtcgttgag	480
taaaagatgt cgagttgccg ttatatatcg cgggcctgta gctatgtgcc ctctattctc	540
acaggttcaa tcatcagtc tccgctgag acgtagcgcg ctgaactagc gctcgaatc	600
ttccgctc atg gct ctt cat gcc ttc ttg ttg ctg gca tca gca ttg ctg Met Ala Leu His Ala Phe Leu Leu Leu Ala Ser Ala Leu Leu 1 5 10	649
gcc cgg ggt gcc ctg agc caa cct gac aac gtc cgt cgc gct gct ccg Ala Arg Gly Ala Leu Ser Gln Pro Asp Asn Val Arg Arg Ala Ala Pro 15 20 25 30	697
acc ggg acg gcc gcc tgg gat gcc gcc cac tgc cag gct gcc gct gcc Thr Gly Thr Ala Ala Trp Asp Ala Ala His Ser Gln Ala Ala Ala 35 40 45	745
gtg tgc aga tta tca cag caa gac aag atc aac att gtc acc ggc gtt Val Ser Arg Leu Ser Gln Gln Asp Lys Ile Asn Ile Val Thr Gly Val 50 55 60	793
ggc tgg ggt aag ggt cct tgc gtc ggc aat acg aac cct gtc tac agc Gly Trp Gly Lys Gly Pro Cys Val Gly Asn Thr Asn Pro Val Tyr Ser 65 70 75	841
atc aac tac cca cag ctc tgc ctg cag gat ggc cca ctg ggt atc cgc Ile Asn Tyr Pro Gln Leu Cys Leu Gln Asp Gly Pro Leu Gly Ile Arg 80 85 90	889
tcc gcc acc agc gtc acg gcc ttc acg ccg ggc att caa gcc gcg tgc Ser Ala Thr Ser Val Thr Ala Phe Thr Pro Gly Ile Gln Ala Ala Ser 95 100 105 110	937
acc tgg gat gtg gag ttg atc cgg cag cgt ggt gtc tac cta gga cag Thr Trp Asp Val Glu Leu Ile Arg Gln Arg Gly Val Tyr Leu Gly Gln 115 120 125	985
gag gcc cgg gga act ggc gtg cat gtc ctg ctc ggc ccc gtg gcc ggt Glu Ala Arg Gly Thr Gly Val His Val Leu Leu Gly Pro Val Ala Gly 130 135 140	1033
gct ctt ggc aag atc ccg cac gga ggc cgt aac tgg gaa gcc ttc ggc	1081

ES 2 527 681 T3

Ala	Leu	Gly	Lys	Ile	Pro	His	Gly	Gly	Arg	Asn	Trp	Glu	Ala	Phe	Gly		
		145					150					155					
tcc	gac	ccc	tac	ttg	gcc	ggt	atc	gct	atg	tcc	gag	acc	atc	gag	ggc	1129	
Ser	Asp	Pro	Tyr	Leu	Ala	Gly	Ile	Ala	Met	Ser	Glu	Thr	Ile	Glu	Gly		
		160				165					170						
att	cag	tcg	gag	ggt	gtg	cag	gct	tgc	gcg	aag	cac	tac	atc	gcc	aat	1177	
Ile	Gln	Ser	Glu	Gly	Val	Gln	Ala	Cys	Ala	Lys	His	Tyr	Ile	Ala	Asn		
					180					185					190		
gag	cag	gaa	ctc	aac	cgc	gag	aca	atg	agc	agc	aac	gtc	gac	gac	cgc	1225	
Glu	Gln	Glu	Leu	Asn	Arg	Glu	Thr	Met	Ser	Ser	Asn	Val	Asp	Asp	Arg		
				195					200					205			
act	atg	cac	gag	cta	tac	ctc	tgg	ccg	ttc	gcc	gac	gcc	gtg	cat	tcc	1273	
Thr	Met	His	Glu	Leu	Tyr	Leu	Trp	Pro	Phe	Ala	Asp	Ala	Val	His	Ser		
				210				215					220				
aac	gtg	gcc	agc	gtc	atg	tgc	agc	tac	aac	aag	ctc	aac	ggc	acc	tgg	1321	
Asn	Val	Ala	Ser	Val	Met	Cys	Ser	Tyr	Asn	Lys	Leu	Asn	Gly	Thr	Trp		
		225					230					235					
ctc	tgc	gag	aac	gat	agg	gcc	caa	aac	cag	ctg	ctt	aag	agg	gag	ctc	1369	
Leu	Cys	Glu	Asn	Asp	Arg	Ala	Gln	Asn	Gln	Leu	Leu	Lys	Arg	Glu	Leu		
		240				245					250						
ggc	ttc	cg	ggc	tac	atc	gtg	agc	gac	tgg	aac	gcg	cag	cac	acc	acc	1417	
Gly	Phe	Arg	Gly	Tyr	Ile	Val	Ser	Asp	Trp	Asn	Ala	Gln	His	Thr	Thr		
					260					265					270		
gtg	ggc	tcg	gcc	aac	agt	ggc	atg	gac	atg	acc	atg	cct	ggc	agc	gac	1465	
Val	Gly	Ser	Ala	Asn	Ser	Gly	Met	Asp	Met	Thr	Met	Pro	Gly	Ser	Asp		
				275					280					285			
ttc	aac	ggc	tgg	aac	gtc	ctc	tgg	ggt	ccg	cag	ctc	aac	aac	gcc	gtc	1513	
Phe	Asn	Gly	Trp	Asn	Val	Leu	Trp	Gly	Pro	Gln	Leu	Asn	Asn	Ala	Val		
			290					295					300				
aac	agc	ggc	cag	gtc	tcg	cag	tcc	cg	ctc	aac	gac	atg	gtc	cag	cgc	1561	
Asn	Ser	Gly	Gln	Val	Ser	Gln	Ser	Arg	Leu	Asn	Asp	Met	Val	Gln	Arg		
		305					310					315					
att	ctt	gct	gcg	tgg	tac	ctc	ctc	ggc	cag	aac	tcc	gga	tac	ccg	tcc	1609	
Ile	Leu	Ala	Ala	Trp	Tyr	Leu	Leu	Gly	Gln	Asn	Ser	Gly	Tyr	Pro	Ser		
						325					330						
atc	aac	ctg	cgt	gcc	aac	gtc	caa	gcc	aac	cac	aag	gag	aat	gtg	cgt	1657	
Ile	Asn	Leu	Arg	Ala	Asn	Val	Gln	Ala	Asn	His	Lys	Glu	Asn	Val	Arg		
					340					345					350		
gcc	gta	gcc	cg	gat	ggc	atc	gtc	ctc	ctc	aag	aac	gac	ggc	att	ctg	1705	
Ala	Val	Ala	Arg	Asp	Gly	Ile	Val	Leu	Leu	Lys	Asn	Asp	Gly	Ile	Leu		
				355					360					365			
cct	ctt	cag	cgt	ccc	aat	aag	att	gct	ctt	gtc	ggc	tcc	gcc	gca	gtc	1753	
Pro	Leu	Gln	Arg	Pro	Asn	Lys	Ile	Ala	Leu	Val	Gly	Ser	Ala	Ala	Val		
				370				375					380				
gtc	aac	ccc	cgt	ggt	atg	aac	gca	tgc	gtg	gac	cgt	ggc	tgc	aac	gag	1801	
Val	Asn	Pro	Arg	Gly	Met	Asn	Ala	Cys	Val	Asp	Arg	Gly	Cys	Asn	Glu		
		385					390					395					
ggt	gcc	ctt	ggc	atg	ggc	tgg	ggc	tca	ggc	acg	gtc	gag	tat	ccc	tac	1849	
Gly	Ala	Leu	Gly	Met	Gly	Trp	Gly	Ser	Gly	Thr	Val	Glu	Tyr	Pro	Tyr		
						405					410						
ttt	gtt	gcg	ccg	tat	gat	gct	ctg	cgt	gag	cg	gca	cag	cgc	gat	ggc	1897	

Phe 415	Val	Ala	Pro	Tyr	Asp 420	Ala	Leu	Arg	Glu	Arg 425	Ala	Gln	Arg	Asp	Gly 430	
acg Thr	cag Gln	atc Ile	agt Ser	ctg Leu 435	cat His	gca Ala	tcg Ser	gac Asp	aat Asn 440	aca Thr	aac Asn	ggg Gly	gtt Val	aac Asn 445	aac Asn	1945
gcc Ala	gtg Val	cag Gln	ggc Gly 450	gct Ala	gac Asp	gcg Ala	gcg Ala	ttt Phe 455	gtg Val	ttc Phe	atc Ile	act Thr	gct Ala 460	gac Asp	tcc Ser	1993
ggc Gly	gaa Glu	ggg Gly 465	tac Tyr	att Ile	acc Thr	gtt Val	gag Glu 470	ggc Gly	cat His	gct Ala	ggc Gly	gac Asp 475	cgg Arg	aat Asn	cat His	2041
ctg Leu	gat Asp 480	cct Pro	tgg Trp	cat His	aat Asn	ggt Gly 485	aac Asn	cag Gln	ctt Leu	gtg Val	cag Gln 490	gct Ala	gtt Val	gcg Ala	cag Gln	2089
gca Ala 495	aat Asn	aag Lys	aac Asn	gtc Val	att Ile 500	gtg Val	gtt Val	gtg Val	cac His	agc Ser 505	gtt Val	ggg Gly	ccg Pro	gtt Val	att Ile 510	2137
ctg Leu	gag Glu	acg Thr	atc Ile	ctc Leu 515	aat Asn	acg Thr	ccc Pro	ggt Gly	gtg Val 520	agg Arg	gct Ala	gtt Val	gtt Val	tgg Trp 525	gct Ala	2185
ggc Gly	ttg Leu	ccg Pro	agc Ser 530	cag Gln	gag Glu	agc Ser	ggt Gly	aac Asn 535	gcg Ala	ctg Leu	gtt Val	gat Asp	gtg Val 540	ctg Leu	tac Tyr	2233
ggc Gly	ctt Leu	gtt Val 545	tcg Ser	ccg Pro	tcg Ser	ggc Gly	aag Lys 550	ctt Leu	gtc Val	tac Tyr	acc Thr	att Ile 555	gcg Ala	aag Lys	agc Ser	2281
ccg Pro	agc Ser 560	gac Asp	tac Tyr	ccg Pro	act Thr	agc Ser 565	att Ile	gtc Val	cgt Arg	ggc Gly	gat Asp 570	gat Asp	aac Asn	ttc Phe	cgc Arg	2329
gag Glu 575	ggt Gly	ctg Leu	ttc Phe	atc Ile	gac Asp 580	tac Tyr	agg Arg	cac His	ttc Phe	gat Asp 585	aac Asn	gcc Ala	cgg Arg	atc Ile	gag Glu 590	2377
ccc Pro	cgt Arg	ttc Phe	gag Glu	ttt Phe 595	ggc Gly	ttc Phe	ggt Gly	ctc Leu	t	gtaagtctct	taccactccg					2425
ttttgtaaca	accgcattct	aacatcccc	ag	ca	tac	acc	aac	ttc	agc	tat						2477
ser	tyr	thr	asn	phe	ser	tyr										605
tcc Ser	aac Asn	ctg Leu	ggc Gly 610	atc Ile	tcc Ser	tcg Ser	tcc Ser	gca Ala 615	acc Thr	gcc Ala	ggc Gly	cca Pro	gcc Ala 620	acg Thr	ggc Gly	2525
ccc Pro	acc Thr	gtc Val 625	ccc Pro	ggc Gly	ggc Gly	ccg Pro	gcc Ala 630	gac Asp	ctc Leu	tgg Trp	aac Asn	tat Tyr 635	gtc Val	gcg Ala	acc Thr	2573
gtc Val	acg Thr 640	gcg Ala	acc Thr	gtt Val	acc Thr	aac Asn 645	acc Thr	ggc Gly	ggc Gly	gtg Val	gaa Glu 650	ggt Gly	gcc Ala	gag Glu	gtc Val	2621
gct Ala 655	cag Gln	ctg Leu	tac Tyr	atc Ile	tct Ser 660	ttg Leu	cca Pro	tct Ser	tcg Ser	gct Ala 665	cct Pro	gca Ala	tcg Ser	cca Pro	ccg Pro 670	2669
aag	cag	ctt	cgt	ggc	ttt	gtc	aag	ctt	aag	ttg	gcg	cct	ggt	caa	agc	2717

Lys Gln Leu Arg Gly Phe Val Lys Leu Lys Leu Ala Pro Gly Gln Ser
675 680 685

ggg acg gca acg ttt aga cta agg aag agg gat ttg gct tat tgg gat 2765
Gly Thr Ala Thr Phe Arg Leu Arg Lys Arg Asp Leu Ala Tyr Trp Asp
690 695 700

gtg ggg agg cag aat tgg gtt gtt cct tcg ggg agg ttt ggc gtg ctt 2813
Val Gly Arg Gln Asn Trp Val Val Pro Ser Gly Arg Phe Gly Val Leu
705 710 715

gtg ggg gct agt tcg agg gat att agg ttg cag ggg gag att gtt gtt 2861
Val Gly Ala Ser Ser Arg Asp Ile Arg Leu Gln Gly Glu Ile Val Val
720 725 730

taggggggta tggtcagcac ctagtgggg aattgatgtg taagttggag taggggtttt 2921

cgtgtacata cataccattt ggtcaatgtt acgacattta gtttatgaag tttcctggtg 2981

gctaccgctg atgagccctc gtatgatacc cacaatctat atgttttact cttctctttc 3041

cttttttctc ttccttttcc tttattactt cattccttgt gtactttctg tgaacctcca 3101

gtcgaccatc cgaccaatt cgaaagtctt tcctgacctg gttcagggtg gcatattctc 3161

gaaaggatgt cgaccttctt gaccttactg ggctaccggg aaagccctag gatggctgat 3221

ggacagatct ggtgatcaac tatgggaaca ctccggagat ggtgactaat atgcgatggt 3281

catttaaaga gcaccgcttc cagcgatctc cccagttgct cctcaacgat tgacacggcc 3341

aatttatcca gattccggga ttctctgagt gagctgtccc ttttttctag a 3392

<210> 26

<211> 734

<212> PRT

<213> Chaetomium thermophilum

<400> 26

Met Ala Leu His Ala Phe Leu Leu Leu Ala Ser Ala Leu Leu Ala Arg
1 5 10 15

Gly Ala Leu Ser Gln Pro Asp Asn Val Arg Arg Ala Ala Pro Thr Gly
20 25 30

Thr Ala Ala Trp Asp Ala Ala His Ser Gln Ala Ala Ala Val Ser
35 40 45

Arg Leu Ser Gln Gln Asp Lys Ile Asn Ile Val Thr Gly Val Gly Trp
50 55 60

Gly Lys Gly Pro Cys Val Gly Asn Thr Asn Pro Val Tyr Ser Ile Asn
65 70 75 80

Tyr Pro Gln Leu Cys Leu Gln Asp Gly Pro Leu Gly Ile Arg Ser Ala
85 90 95

ES 2 527 681 T3

Thr Ser Val Thr Ala Phe Thr Pro Gly Ile Gln Ala Ala Ser Thr Trp
 100 105 110
 Asp Val Glu Leu Ile Arg Gln Arg Gly Val Tyr Leu Gly Gln Glu Ala
 115 120 125
 Arg Gly Thr Gly Val His Val Leu Leu Gly Pro Val Ala Gly Ala Leu
 130 135 140
 Gly Lys Ile Pro His Gly Gly Arg Asn Trp Glu Ala Phe Gly Ser Asp
 145 150 155 160
 Pro Tyr Leu Ala Gly Ile Ala Met Ser Glu Thr Ile Glu Gly Ile Gln
 165 170 175
 Ser Glu Gly Val Gln Ala Cys Ala Lys His Tyr Ile Ala Asn Glu Gln
 180 185 190
 Glu Leu Asn Arg Glu Thr Met Ser Ser Asn Val Asp Asp Arg Thr Met
 195 200 205
 His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val His Ser Asn Val
 210 215 220
 Ala Ser Val Met Cys Ser Tyr Asn Lys Leu Asn Gly Thr Trp Leu Cys
 225 230 235 240
 Glu Asn Asp Arg Ala Gln Asn Gln Leu Leu Lys Arg Glu Leu Gly Phe
 245 250 255
 Arg Gly Tyr Ile Val Ser Asp Trp Asn Ala Gln His Thr Thr Val Gly
 260 265 270
 Ser Ala Asn Ser Gly Met Asp Met Thr Met Pro Gly Ser Asp Phe Asn
 275 280 285
 Gly Trp Asn Val Leu Trp Gly Pro Gln Leu Asn Asn Ala Val Asn Ser
 290 295 300
 Gly Gln Val Ser Gln Ser Arg Leu Asn Asp Met Val Gln Arg Ile Leu
 305 310 315 320
 Ala Ala Trp Tyr Leu Leu Gly Gln Asn Ser Gly Tyr Pro Ser Ile Asn
 325 330 335
 Leu Arg Ala Asn Val Gln Ala Asn His Lys Glu Asn Val Arg Ala Val
 340 345 350
 Ala Arg Asp Gly Ile Val Leu Leu Lys Asn Asp Gly Ile Leu Pro Leu
 355 360 365

ES 2 527 681 T3

Gln Arg Pro Asn Lys Ile Ala Leu Val Gly Ser Ala Ala Val Val Asn
 370 375 380
 Pro Arg Gly Met Asn Ala Cys Val Asp Arg Gly Cys Asn Glu Gly Ala
 385 390 395 400
 Leu Gly Met Gly Trp Gly Ser Gly Thr Val Glu Tyr Pro Tyr Phe Val
 405 410 415
 Ala Pro Tyr Asp Ala Leu Arg Glu Arg Ala Gln Arg Asp Gly Thr Gln
 420 425 430
 Ile Ser Leu His Ala Ser Asp Asn Thr Asn Gly Val Asn Asn Ala Val
 435 440 445
 Gln Gly Ala Asp Ala Ala Phe Val Phe Ile Thr Ala Asp Ser Gly Glu
 450 455 460
 Gly Tyr Ile Thr Val Glu Gly His Ala Gly Asp Arg Asn His Leu Asp
 465 470 475 480
 Pro Trp His Asn Gly Asn Gln Leu Val Gln Ala Val Ala Gln Ala Asn
 485 490 495
 Lys Asn Val Ile Val Val Val His Ser Val Gly Pro Val Ile Leu Glu
 500 505 510
 Thr Ile Leu Asn Thr Pro Gly Val Arg Ala Val Val Trp Ala Gly Leu
 515 520 525
 Pro Ser Gln Glu Ser Gly Asn Ala Leu Val Asp Val Leu Tyr Gly Leu
 530 535 540
 Val Ser Pro Ser Gly Lys Leu Val Tyr Thr Ile Ala Lys Ser Pro Ser
 545 550 555 560
 Asp Tyr Pro Thr Ser Ile Val Arg Gly Asp Asp Asn Phe Arg Glu Gly
 565 570 575
 Leu Phe Ile Asp Tyr Arg His Phe Asp Asn Ala Arg Ile Glu Pro Arg
 580 585 590
 Phe Glu Phe Gly Phe Gly Leu Ser Tyr Thr Asn Phe Ser Tyr Ser Asn
 595 600 605
 Leu Gly Ile Ser Ser Ser Ala Thr Ala Gly Pro Ala Thr Gly Pro Thr
 610 615 620
 Val Pro Gly Gly Pro Ala Asp Leu Trp Asn Tyr Val Ala Thr Val Thr
 625 630 635 640

ES 2 527 681 T3

Ala Thr Val Thr Asn Thr Gly Gly Val Glu Gly Ala Glu Val Ala Gln
645 650 655

Leu Tyr Ile Ser Leu Pro Ser Ser Ala Pro Ala Ser Pro Pro Lys Gln
660 665 670

Leu Arg Gly Phe Val Lys Leu Lys Leu Ala Pro Gly Gln Ser Gly Thr
675 680 685

Ala Thr Phe Arg Leu Arg Lys Arg Asp Leu Ala Tyr Trp Asp Val Gly
690 695 700

Arg Gln Asn Trp Val Val Pro Ser Gly Arg Phe Gly Val Leu Val Gly
705 710 715 720

Ala Ser Ser Arg Asp Ile Arg Leu Gln Gly Glu Ile Val Val
725 730

<210> 27

<211> 1631

<212> ADN

5 <213> Thermoascus aurantiacus

<220>

<221> Intrón

<222> (610)..(674)

10 <223>

<220>

<221> CDS

<222> (675)..(1628)

15 <223>

<220>

<221> CDS

<222> (1)..(609)

20 <223>

<400> 27

atg tat cag cgc gct ctt ctc ttc tct ttc ttc ctc gcc gcc gcc cgc

48

25

ES 2 527 681 T3

Met 1	Tyr	Gln	Arg	Ala 5	Leu	Leu	Phe	Ser	Phe 10	Phe	Leu	Ala	Ala	Ala 15	Arg		
gcg Ala	cag Gln	cag Gln	gcc Ala 20	ggt Gly	acc Thr	gta Val	acc Thr 25	gca Ala	gag Glu	aat Asn	cac His	cct Pro	tcc Ser 30	ctg Leu	acc Thr	96	
tgg Trp	cag Gln	caa Gln 35	tgc Cys	tcc Ser	agc Ser	ggc Gly	ggt Gly 40	agt Ser	tgt Cys	acc Thr	acg Thr	cag Gln 45	aat Asn	gga Gly	aaa Lys	144	
gtc Val	gtt Val 50	atc Ile	gat Asp	gcg Ala	aac Asn	tgg Trp 55	cgt Arg	tgg Trp	gtc Val	cat His	acc Thr 60	acc Thr	tct Ser	gga Gly	tac Tyr	192	
acc Thr 65	aac Asn	tgc Cys	tac Tyr	acg Thr	ggc Gly 70	aat Asn	acg Thr	tgg Trp	gac Asp	acc Thr 75	agt Ser	atc Ile	tgt Cys	ccc Pro	gac Asp 80	240	
gac Asp	gtg Val	acc Thr	tgc Cys	gct Ala 85	cag Gln	aat Asn	tgt Cys	gcc Ala	ttg Leu 90	gat Asp	gga Gly	gcg Ala	gat Asp	tac Tyr 95	agt Ser	288	
ggc Gly	acc Thr	tat Tyr	ggt Gly 100	gtt Val	acg Thr	acc Thr	agt Ser	ggc Gly 105	aac Asn	gcc Ala	ctg Leu	aga Arg	ctg Leu 110	aac Asn	ttt Phe	336	
gtc Val	acc Thr	caa Gln 115	agc Ser	tca Ser	ggg Gly	aag Lys	aac Asn 120	att Ile	ggc Gly	tcg Ser	cgc Arg	ctg Leu 125	tac Tyr	ctg Leu	ctg Leu	384	
cag Gln 130	gac Asp	gac Asp	acc Thr	act Thr	tat Tyr	cag Gln 135	atc Ile	ttc Phe	aag Lys	ctg Leu	ctg Leu 140	ggt Gly	cag Gln	gag Glu	ttt Phe	432	
acc Thr 145	ttc Phe	gat Asp	gtc Val	gac Asp	gtc Val 150	tcc Ser	aat Asn	ctc Leu	cct Pro	tgc Cys 155	ggg Gly	ctg Leu	aac Asn	ggc Gly	gcc Ala 160	480	
ctc Leu	tac Tyr	ttt Phe	gtg Val 165	gcc Ala	atg Met	gac Asp	gcc Ala	gac Asp	ggc Gly 170	gga Gly	ttg Leu	tcc Ser	aaa Lys	tac Tyr 175	cct Pro	528	
ggc Gly	aac Asn	aag Lys	gca Ala 180	ggc Gly	gct Ala	aag Lys	tat Tyr	ggc Gly 185	act Thr	ggt Gly	tac Tyr	tgc Cys	gac Asp 190	tct Ser	cag Gln	576	
tgc Cys	cct Pro	cgg Arg 195	gat Asp	ctc Leu	aag Lys	ttc Phe	atc Ile 200	aac Asn	ggt Gly	cag Gln	gtacg	tcaga	agtgataact			629	
agccagcaga	gcccataaat	cattaactaa	cgctgtcaaa	tacag	gcc Ala 205	aac Asn	ggt Val	gaa								686	
ggc Gly	tgg Trp	cag Gln 210	ccg Pro	tct Ser	gcc Ala	aac Asn	gac Asp 215	cca Pro	aat Asn	gcc Ala	ggc Gly	ggt Val 220	ggt Gly	aac Asn	cac His	734	
ggt Gly	tcc Ser	tgc Cys	tgc Cys	gct Ala	gag Glu	atg Met 230	gat Asp	gtc Val	tgg Trp	gaa Glu	gcc Ala 235	aac Asn	agc Ser	atc Ile	tct Ser	782	
act Thr 240	gcg Ala	gtg Val	acg Thr	cct Pro	cac His 245	cca Pro	tgc Cys	gac Asp	acc Thr	ccc Pro 250	ggc Gly	cag Gln	acc Thr	atg Met	tgc Cys 255	830	
cag	gga	gac	gac	tgt	ggt	gga	acc	tac	tcc	tcc	act	cga	tat	gct	ggt	878	

Gln	Gly	Asp	Asp	Cys 260	Gly	Gly	Thr	Tyr	Ser 265	Ser	Thr	Arg	Tyr	Ala 270	Gly		
acc	tgc	gac	cct	gat	ggc	tgc	gac	ttc	aat	cct	tac	cgc	cag	ggc	aac	926	
Thr	Cys	Asp	Pro 275	Asp	Gly	Cys	Asp	Phe 280	Asn	Pro	Tyr	Arg	Gln 285	Gly	Asn		
cac	tgc	ttc	tac	ggc	ccc	ggg	cag	atc	gtc	gac	acc	agc	tcc	aaa	ttc	974	
His	Ser	Phe 290	Tyr	Gly	Pro	Gly	Gln 295	Ile	Val	Asp	Thr	Ser 300	Ser	Lys	Phe		
acc	gtc	gtc	acc	cag	ttc	atc	acc	gac	gac	ggg	acc	ccc	tcc	ggc	acc	1022	
Thr	Val 305	Val	Thr	Gln	Phe	Ile 310	Thr	Asp	Asp	Gly	Thr 315	Pro	Ser	Gly	Thr		
ctg	acg	gag	atc	aaa	cgc	ttc	tac	gtc	cag	aac	ggc	aag	gta	atc	ccc	1070	
Leu 320	Thr	Glu	Ile	Lys	Arg 325	Phe	Tyr	Val	Gln	Asn 330	Gly	Lys	Val	Ile	Pro 335		
cag	tgc	gag	tgc	acg	atc	agc	ggc	gtc	acc	ggc	aac	tca	atc	acc	acc	1118	
Gln	Ser	Glu	Ser	Thr 340	Ile	Ser	Gly	Val	Thr 345	Gly	Asn	Ser	Ile	Thr 350	Thr		
gag	tat	tgc	acg	ggc	cag	aag	gcc	gcc	ttc	ggc	gac	aac	acc	ggc	ttc	1166	
Glu	Tyr	Cys	Thr 355	Ala	Gln	Lys	Ala	Ala 360	Phe	Gly	Asp	Asn	Thr 365	Gly	Phe		
ttc	acg	cac	ggc	ggg	ctt	cag	aag	atc	agt	cag	gct	ctg	gct	cag	ggc	1214	
Phe	Thr	His 370	Gly	Gly	Leu	Gln	Lys 375	Ile	Ser	Gln	Ala	Leu 380	Ala	Gln	Gly		
atg	gtc	ctc	gtc	atg	agc	ctg	tgg	gac	gat	cac	gcc	gcc	aac	atg	ctc	1262	
Met	Val 385	Leu	Val	Met	Ser	Leu 390	Trp	Asp	Asp	His	Ala 395	Ala	Asn	Met	Leu		
tgg	ctg	gac	agc	acc	tac	ccg	act	gat	gcg	gac	ccg	gac	acc	cct	ggc	1310	
Trp 400	Leu	Asp	Ser	Thr	Tyr 405	Pro	Thr	Asp	Ala	Asp 410	Pro	Asp	Thr	Pro	Gly 415		
gtc	gcg	cgc	ggt	acc	tgc	ccc	acg	acc	tcc	ggc	gtc	ccg	gcc	gac	ggt	1358	
Val	Ala	Arg	Gly	Thr 420	Cys	Pro	Thr	Thr	Ser 425	Gly	Val	Pro	Ala	Asp 430	Val		
gag	tgc	cag	tac	ccc	aat	tca	tat	ggt	atc	tac	tcc	aac	atc	aag	gtc	1406	
Glu	Ser	Gln 435	Tyr	Pro	Asn	Ser	Tyr	Val 440	Ile	Tyr	Ser	Asn	Ile 445	Lys	Val		
gga	ccc	att	ggc	agc	acc	ggc	aac	cct	agc	ggc	ggc	aac	cct	ccc	ggc	1454	
Gly	Pro	Ile 450	Gly	Ser	Thr	Gly	Asn 455	Pro	Ser	Gly	Gly	Asn 460	Pro	Pro	Gly		
gga	aac	ccg	cct	ggc	acc	acc	acc	acc	cgc	cgc	cca	gcc	act	acc	act	1502	
Gly	Asn 465	Pro	Pro	Gly	Thr	Thr	Thr	Thr	Arg	Arg	Pro	Ala	Thr	Thr	Thr		
gga	agc	tct	ccc	gga	cct	acc	cag	tct	cac	tac	ggc	cag	tgc	ggc	ggt	1550	
Gly 480	Ser	Ser	Pro	Gly	Pro 485	Thr	Gln	Ser	His	Tyr 490	Gly	Gln	Cys	Gly	Gly 495		
att	ggc	tac	agc	ggc	ccc	acg	gtc	tgc	gcc	agc	ggc	aca	act	tgc	cag	1598	
Ile	Gly	Tyr	Ser	Gly 500	Pro	Thr	Val	Cys	Ala 505	Ser	Gly	Thr	Thr	Cys 510	Gln		
gtc	ctg	aac	cct	tac	tac	tct	cag	tgc	ctg	taa						1631	
Val	Leu	Asn 515	Pro	Tyr	Tyr	Ser	Gln	Cys 520	Leu								

<210> 28

<211> 521

<212> PRT

<213> Thermoascus aurantiacus

<400> 28

ES 2 527 681 T3

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

ES 2 527 681 T3

Ser Cys Cys Ala Glu Met Asp Val Trp Glu Ala Asn Ser Ile Ser Thr
 225 230 235 240
 Ala Val Thr Pro His Pro Cys Asp Thr Pro Gly Gln Thr Met Cys Gln
 245 250 255
 Gly Asp Asp Cys Gly Gly Thr Tyr Ser Ser Thr Arg Tyr Ala Gly Thr
 260 265 270
 Cys Asp Pro Asp Gly Cys Asp Phe Asn Pro Tyr Arg Gln Gly Asn His
 275 280 285
 Ser Phe Tyr Gly Pro Gly Gln Ile Val Asp Thr Ser Ser Lys Phe Thr
 290 295 300
 Val Val Thr Gln Phe Ile Thr Asp Asp Gly Thr Pro Ser Gly Thr Leu
 305 310 315 320
 Thr Glu Ile Lys Arg Phe Tyr Val Gln Asn Gly Lys Val Ile Pro Gln
 325 330 335
 Ser Glu Ser Thr Ile Ser Gly Val Thr Gly Asn Ser Ile Thr Thr Glu
 340 345 350
 Tyr Cys Thr Ala Gln Lys Ala Ala Phe Gly Asp Asn Thr Gly Phe Phe
 355 360 365
 Thr His Gly Gly Leu Gln Lys Ile Ser Gln Ala Leu Ala Gln Gly Met
 370 375 380
 Val Leu Val Met Ser Leu Trp Asp Asp His Ala Ala Asn Met Leu Trp
 385 390 395 400
 Leu Asp Ser Thr Tyr Pro Thr Asp Ala Asp Pro Asp Thr Pro Gly Val
 405 410 415
 Ala Arg Gly Thr Cys Pro Thr Thr Ser Gly Val Pro Ala Asp Val Glu
 420 425 430
 Ser Gln Tyr Pro Asn Ser Tyr Val Ile Tyr Ser Asn Ile Lys Val Gly
 435 440 445
 Pro Ile Gly Ser Thr Gly Asn Pro Ser Gly Gly Asn Pro Pro Gly Gly
 450 455 460
 Asn Pro Pro Gly Thr Thr Thr Thr Arg Arg Pro Ala Thr Thr Thr Gly
 465 470 475 480
 Ser Ser Pro Gly Pro Thr Gln Ser His Tyr Gly Gln Cys Gly Gly Ile
 485 490 495
 Gly Tyr Ser Gly Pro Thr Val Cys Ala Ser Gly Thr Thr Cys Gln Val
 500 505 510
 Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
 515 520

<210> 29

<211> 1734

<212> ADN

<213> Thermoascus aurantiacus

<220>
 <221> Intrón
 <222> (610)..(674)
 <223>
 5
 <220>
 <221> CDS
 <222> (1726)..(1731)
 <223>
 10
 <220>
 <221> CDS
 <222> (675)..(1661)
 <223>
 15
 <220>
 <221> CDS
 <222> (1)..(609)
 <223>
 20
 <220>
 <221> Intrón
 <222> (1662)..(1725)
 <223>
 25
 <400> 29

ES 2 527 681 T3

atg	tat	cag	cgc	gct	ctt	ctc	ttc	tct	ttc	ttc	ctc	gcc	gcc	gcc	cgc	48
Met	Tyr	Gln	Arg	Ala	Leu	Leu	Phe	Ser	Phe	Phe	Leu	Ala	Ala	Ala	Arg	
1				5				10						15		
gcg	cag	cag	gcc	ggg	acc	gta	acc	gca	gag	aat	cac	cct	tcc	ctg	acc	96
Ala	Gln	Gln	Ala	Gly	Thr	Val	Thr	Ala	Glu	Asn	His	Pro	Ser	Leu	Thr	
			20					25					30			
tgg	cag	caa	tgc	tcc	agc	ggc	ggg	agt	tgt	acc	acg	cag	aat	gga	aaa	144
Trp	Gln	Gln	Cys	Ser	Ser	Gly	Gly	Ser	Cys	Thr	Thr	Gln	Asn	Gly	Lys	
			35				40					45				
gtc	gtt	atc	gat	gcg	aac	tgg	cgt	tgg	gtc	cat	acc	acc	tct	gga	tac	192
Val	Val	Ile	Asp	Ala	Asn	Trp	Arg	Trp	Val	His	Thr	Thr	Ser	Gly	Tyr	
			50			55					60					
acc	aac	tgc	tac	acg	ggc	aat	acg	tgg	gac	acc	agt	atc	tgt	ccc	gac	240
Thr	Asn	Cys	Tyr	Thr	Gly	Asn	Thr	Trp	Asp	Thr	Ser	Ile	Cys	Pro	Asp	
					70					75					80	
gac	gtg	acc	tgc	gct	cag	aat	tgt	gcc	ttg	gat	gga	gcg	gat	tac	agt	288
Asp	Val	Thr	Cys	Ala	Gln	Asn	Cys	Ala	Leu	Asp	Gly	Ala	Asp	Tyr	Ser	
				85				90						95		
ggc	acc	tat	ggg	gtt	acg	acc	agt	ggc	aac	gcc	ctg	aga	ctg	aac	ttt	336
Gly	Thr	Tyr	Gly	Val	Thr	Thr	Ser	Gly	Asn	Ala	Leu	Arg	Leu	Asn	Phe	
			100					105					110			
gtc	acc	caa	agc	tca	ggg	aag	aac	att	ggc	tcg	cgc	ctg	tac	ctg	ctg	384
Val	Thr	Gln	Ser	Ser	Gly	Lys	Asn	Ile	Gly	Ser	Arg	Leu	Tyr	Leu	Leu	
			115				120					125				
cag	gac	gac	acc	act	tat	cag	atc	ttc	aag	ctg	ctg	ggg	cag	gag	ttt	432
Gln	Asp	Asp	Thr	Thr	Tyr	Gln	Ile	Phe	Lys	Leu	Leu	Gly	Gln	Glu	Phe	
			130			135					140					
acc	ttc	gat	gtc	gac	gtc	tcc	aat	ctc	cct	tgc	ggg	ctg	aac	ggc	gcc	480
Thr	Phe	Asp	Val	Asp	Val	Ser	Asn	Leu	Pro	Cys	Gly	Leu	Asn	Gly	Ala	
					150					155					160	
ctc	tac	ttt	gtg	gcc	atg	gac	gcc	gac	ggc	gga	ttg	tcc	aaa	tac	cct	528
Leu	Tyr	Phe	Val	Ala	Met	Asp	Ala	Asp	Gly	Gly	Leu	Ser	Lys	Tyr	Pro	
				165					170					175		
ggc	aac	aag	gca	ggc	gct	aag	tat	ggc	act	ggg	tac	tgc	gac	tct	cag	576
Gly	Asn	Lys	Ala	Gly	Ala	Lys	Tyr	Gly	Thr	Gly	Tyr	Cys	Asp	Ser	Gln	
			180					185					190			
tgc	cct	cgg	gat	ctc	aag	ttc	atc	aac	ggg	cag	gtacgtcaga	agtgataact				629
Cys	Pro	Arg	Asp	Leu	Lys	Phe	Ile	Asn	Gly	Gln						
			195				200									
agccagcaga	gcccataaat	cattaactaa	cgctgtcaaaa	tacag	gcc	aat	gtt	gaa								686
					Ala	Asn	Val	Glu								
							205									
ggc	tgg	cag	ccg	tct	gcc	aac	gac	cca	aat	gcc	ggc	gtt	ggg	aac	cac	734
Gly	Trp	Gln	Pro	Ser	Ala	Asn	Asp	Pro	Asn	Ala	Gly	Val	Gly	Asn	His	
			210				215					220				
ggg	tcc	tgc	tgc	gct	gag	atg	gat	gtc	tgg	gaa	gcc	aac	agc	atc	tct	782
Gly	Ser	Cys	Cys	Ala	Glu	Met	Asp	Val	Trp	Glu	Ala	Asn	Ser	Ile	Ser	
			225			230					235					
act	gcg	gtg	acg	cct	cac	cca	tgc	gac	acc	ccc	ggc	cag	acc	atg	tgc	830
Thr	Ala	Val	Thr	Pro	His	Pro	Cys	Asp	Thr	Pro	Gly	Gln	Thr	Met	Cys	
					245					250					255	

cag gga gac gac tgt ggt gga acc tac tcc tcc act cga tat gct ggt 878
 Gln Gly Asp Asp Cys 260 Gly Gly Thr Tyr Ser Ser Thr Arg Tyr Ala Gly
 acc tgc gac cct gat ggc tgc gac ttc aat cct tac cgc cag ggc aac 926
 Thr Cys Asp Pro Asp Gly Cys Asp Phe Asn Pro Tyr Arg Gln Gly Asn
 cac tcg ttc tac ggc ccc ggg cag atc gtc gac acc agc tcc aaa ttc 974
 His Ser Phe Tyr Gly Pro Gly Gln Ile Val Asp Thr Ser Ser Lys Phe
 acc gtc gtc acc cag ttc atc acc gac gac ggc acc ccc tcc ggc acc 1022
 Thr Val Val Thr Gln Phe Ile Thr Asp Asp Gly Thr Pro Ser Gly Thr
 ctg acg gag atc aaa cgc ttc tac gtc cag aac ggc aag gta atc ccc 1070
 Leu Thr Glu Ile Lys Arg Phe Tyr Val Gln Asn Gly Lys Val Ile Pro
 cag tcg gag tcg acg atc agc ggc gtc acc ggc aac tca atc acc acc 1118
 Gln Ser Glu Ser Thr Ile Ser Gly Val Thr Gly Asn Ser Ile Thr Thr
 gag tat tgc acg gcc cag aag gcc gcc ttc ggc gac aac acc ggc ttc 1166
 Glu Tyr Cys Thr Ala Gln Lys Ala Ala Phe Gly Asp Asn Thr Gly Phe
 ttc acg cac ggc ggc ctt cag aag atc agt cag gct ctg gct cag ggc 1214
 Phe Thr His Gly Gly Leu Gln Lys Ile Ser Gln Ala Leu Ala Gln Gly
 atg gtc ctc gtc atg agc ctg tgg gac gat cac gcc gcc aac atg ctc 1262
 Met Val Leu Val Met Ser 390 Trp Asp Asp His Ala Ala Asn Met Leu
 tgg ctg gac agc acc tac ccg act gat gcg gac ccg gac acc cct ggc 1310
 Trp Leu Asp Ser Thr Tyr 405 Pro Thr Asp Ala Asp Pro Asp Thr Pro Gly
 gtc gcg cgc ggt acc tgc ccc acg acc tcc ggc gtc ccg gcc gac gtt 1358
 Val Ala Arg Gly Thr Cys Pro Thr Thr Ser Gly Val Pro Ala Asp Val
 gag tcg cag tac ccc aat tca tat gtt atc tac tcc aac atc aag gtc 1406
 Glu Ser Gln Tyr Pro Asn Ser Tyr Val Ile Tyr Ser Asn Ile Lys Val
 gga ccc atc ggc tcg acc gtc cct ggc ctt gac ggc agc aac ccc ggc 1454
 Gly Pro Ile Gly Ser Thr Val Pro Gly Leu Asp Gly Ser Asn Pro Gly
 aac ccg acc acc acc acc gtc gtt cct ccc gct tct acc tcc acc tcc cgt 1502
 Asn Pro Thr Thr Thr Val Val Pro Pro Ala Ser Thr Ser Thr Ser Arg
 ccg acc agc agc act agc tct ccc gtt tcg acc ccg act ggc cag ccc 1550
 Pro Thr Ser Ser Thr Ser 485 Ser Pro Val Ser Thr Pro Thr Gly Gln Pro
 ggc ggc tgc acc acc cag aag tgg ggc cag tgc ggc ggt atc ggc tac 1598
 Gly Gly Cys Thr Thr Gln Lys Trp Gly Gln Cys Gly Gly Ile Gly Tyr
 acc ggc tgc act aac tgc gtt gct ggc acc acc tgc act cag ctc aac 1646
 Thr Gly Cys Thr Asn Cys Val Ala Gly Thr Thr Cys Thr Gln Leu Asn
 ccc tgg tac agc cag gtatgtttct cttccccctt ctagactcgc ttggatttga 1701
 Pro Trp Tyr Ser Gln
 cagttgctaa catctgctca acag tgc ctg taa 1734
 Cys Leu

<210> 30

<211> 534

<212> PRT

<213> Thermoascus aurantiacus

<400> 30

ES 2 527 681 T3

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

ES 2 527 681 T3

Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly
 195 200 205
 Trp Gln Pro Ser Ala Asn Asp Pro Asn Ala Gly Val Gly Asn His Gly
 210 215 220
 Ser Cys Cys Ala Glu Met Asp Val Trp Glu Ala Asn Ser Ile Ser Thr
 225 230 235 240
 Ala Val Thr Pro His Pro Cys Asp Thr Pro Gly Gln Thr Met Cys Gln
 245 250 255
 Gly Asp Asp Cys Gly Gly Thr Tyr Ser Ser Thr Arg Tyr Ala Gly Thr
 260 265 270
 Cys Asp Pro Asp Gly Cys Asp Phe Asn Pro Tyr Arg Gln Gly Asn His
 275 280 285
 Ser Phe Tyr Gly Pro Gly Gln Ile Val Asp Thr Ser Ser Lys Phe Thr
 290 295 300
 Val Val Thr Gln Phe Ile Thr Asp Asp Gly Thr Pro Ser Gly Thr Leu
 305 310 315 320
 Thr Glu Ile Lys Arg Phe Tyr Val Gln Asn Gly Lys Val Ile Pro Gln
 325 330 335
 Ser Glu Ser Thr Ile Ser Gly Val Thr Gly Asn Ser Ile Thr Thr Glu
 340 345 350
 Tyr Cys Thr Ala Gln Lys Ala Ala Phe Gly Asp Asn Thr Gly Phe Phe
 355 360 365
 Thr His Gly Gly Leu Gln Lys Ile Ser Gln Ala Leu Ala Gln Gly Met
 370 375 380
 Val Leu Val Met Ser Leu Trp Asp Asp His Ala Ala Asn Met Leu Trp
 385 390 395 400
 Leu Asp Ser Thr Tyr Pro Thr Asp Ala Asp Pro Asp Thr Pro Gly Val
 405 410 415
 Ala Arg Gly Thr Cys Pro Thr Thr Ser Gly Val Pro Ala Asp Val Glu
 420 425 430
 Ser Gln Tyr Pro Asn Ser Tyr Val Ile Tyr Ser Asn Ile Lys Val Gly
 435 440 445
 Pro Ile Gly Ser Thr Val Pro Gly Leu Asp Gly Ser Asn Pro Gly Asn
 450 455 460

ES 2 527 681 T3

Pro Thr Thr Thr Val Val Pro Pro Ala Ser Thr Ser Thr Ser Arg Pro
 465 470 475 480
 Thr Ser Ser Thr Ser Ser Pro Val Ser Thr Pro Thr Gly Gln Pro Gly
 485 490 495
 Gly Cys Thr Thr Gln Lys Trp Gly Gln Cys Gly Gly Ile Gly Tyr Thr
 500 505 510
 Gly Cys Thr Asn Cys Val Ala Gly Thr Thr Cys Thr Gln Leu Asn Pro
 515 520 525
 Trp Tyr Ser Gln Cys Leu
 530

REIVINDICACIONES

1. Un polipéptido que comprende un fragmento que tiene actividad celulolítica y se selecciona del grupo que consiste en:
 - a) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos 90% con SEQ ID NO: 6; y
 - b) un fragmento a) que tiene actividad celulolítica.
2. El polipéptido de la reivindicación 1, que comprende una secuencia de aminoácidos que tiene una identidad del al menos 95% o 99% con SEQ ID NO: 6, o un fragmento del mismo que tiene actividad celulolítica.
3. Un polinucleótido aislado seleccionado del grupo que consiste en:
 - a) una secuencia de nucleótidos de SEQ ID NO: 5, o una secuencia que codifica un polipéptido de la reivindicación 1;
 - b) una cadena complementaria de a) y
 - c) una secuencia que es degenerada como resultado del código genético para una cualquiera de las secuencias definidas en a) o b).
4. Un vector, que comprende como secuencia heteróloga un polinucleótido de la reivindicación 3.
5. Una célula anfitriona que comprende el vector de la reivindicación 4.
6. Una cepa de *Escherichia coli* que tiene el número de acceso DSM 16729.
7. Una preparación enzimática que comprende un polipéptido de la reivindicación 1.
8. La preparación enzimática de la reivindicación 7, que comprende adicionalmente endoglucanasa y beta-glucosidasa, en donde la endoglucanasa comprende preferiblemente una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 10, 12, 14 o 16, o un fragmento de los mismos que tiene actividad celulolítica, y la beta-glucosidasa comprende preferiblemente una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 22, 24 o 26, o un fragmento de los mismos que tiene actividad celulolítica.
9. La preparación enzimática de la reivindicación 8, que comprende al menos una enzima adicional, preferiblemente una xilanasa, que comprende preferiblemente una secuencia de aminoácidos que tiene una identidad de al menos 80% con SEQ ID NO: 18 o 20, o un fragmento de los mismos que tiene actividad celulolítica.
10. El uso de un polipéptido de acuerdo con la reivindicación 1, o una preparación enzimática según la reivindicación 7 en la industria de combustibles, textil, detergentes, pasta de papel y papel, alimentos, piensos o bebidas, o en el tratamiento de pasta kraft, pasta mecánica, o papel reciclado.
11. El uso de acuerdo con la reivindicación 10, en donde la preparación enzimática es medio de cultivo gastado.
12. El uso de una preparación enzimática de acuerdo con la reivindicación 8 o 9 para degradar material celulósico.
13. Un método para preparar un polipéptido que comprende un fragmento que tiene actividad celulolítica y que se selecciona del grupo que consiste en:
 - a) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos 90% con SEQ ID NO: 6; y
 - b) un fragmento de a) que tiene actividad celulolítica,
 comprendiendo dicho método transformar una célula anfitriona con un vector que codifica dicho polipéptido, y cultivar dicha célula anfitriona en condiciones que permiten la expresión de dicho polipéptido, y opcionalmente recuperar y purificar el polipéptido producido.
14. Un método de tratamiento de material celulósico con un medio de cultivo gastado de al menos un microorganismo capaz de producir un polipéptido que comprende un fragmento que tiene actividad celulolítica y que se selecciona del grupo que consiste en:
 - a) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos 90% con SEQ ID NO: 6; y
 - b) un fragmento de a) que tiene actividad celulolítica,

comprendiendo dicho método hacer reaccionar el material celulósico con el medio de cultivo gastado para obtener material celulósico hidrolizado.

- 5 15. Un método para el tratamiento de material celulósico con celobiohidrolasa, endoglucanasa y beta-glucosidasa, por medio del cual dicha celobiohidrolasa comprende una secuencia de aminoácidos que tiene una identidad de al menos 90% con el SEQ ID NO: 6, o un fragmento del mismo que tiene actividad celulolítica.

16. El uso del método de acuerdo con la reivindicación 15 en un procedimiento para la preparación de etanol a partir de material celulósico.

Fig. 1A.

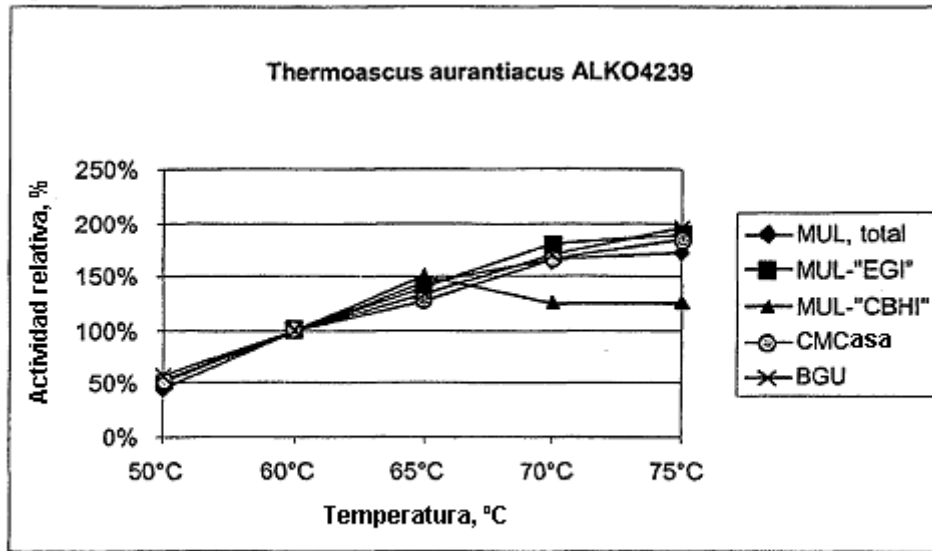


Fig. 1B.

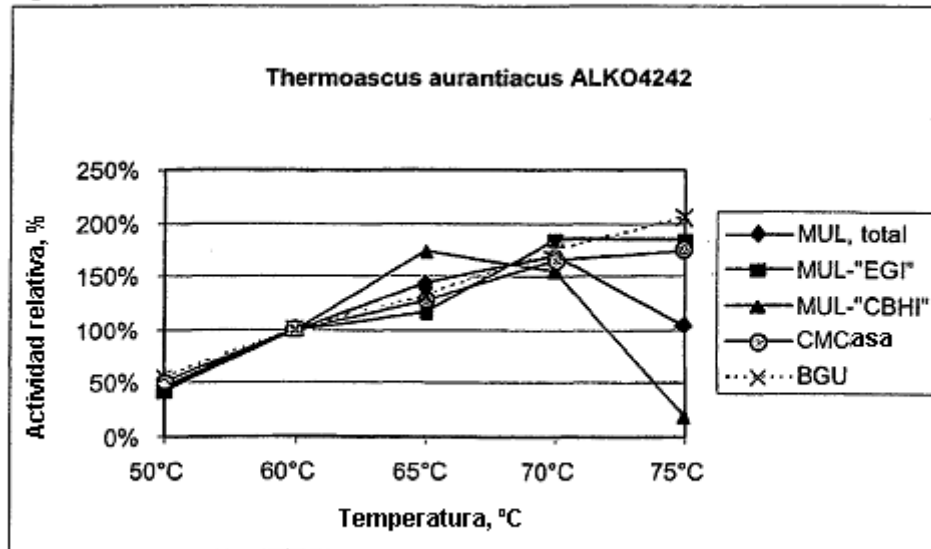


Fig. 1C.

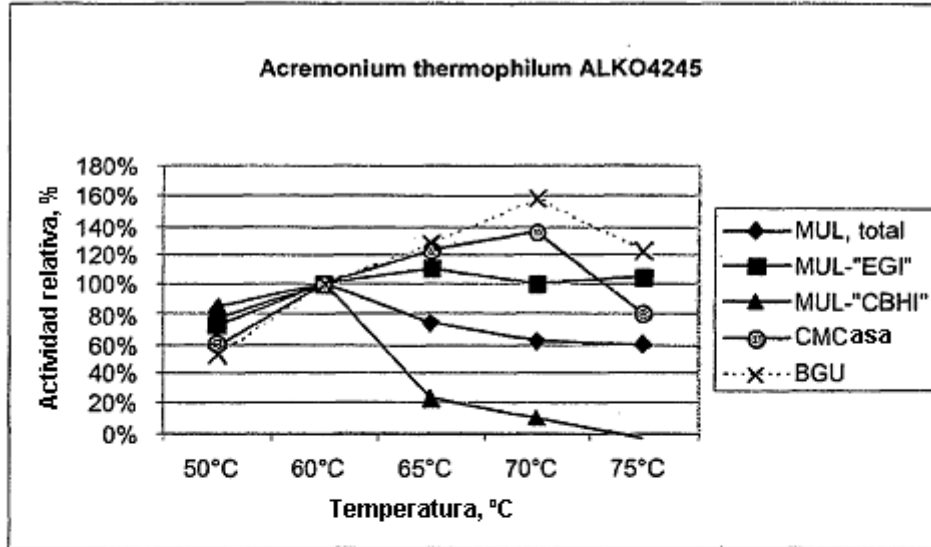


Fig. 1D.

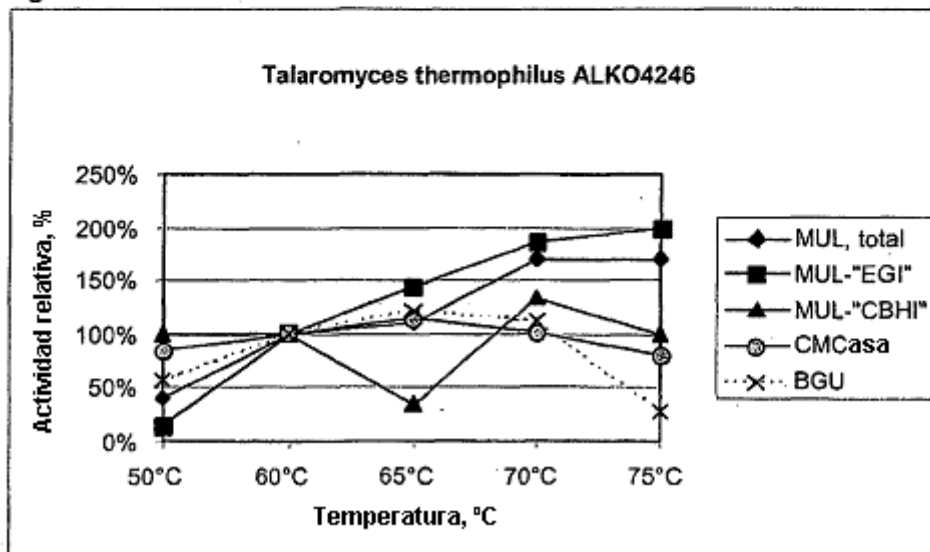


Fig. 1E.

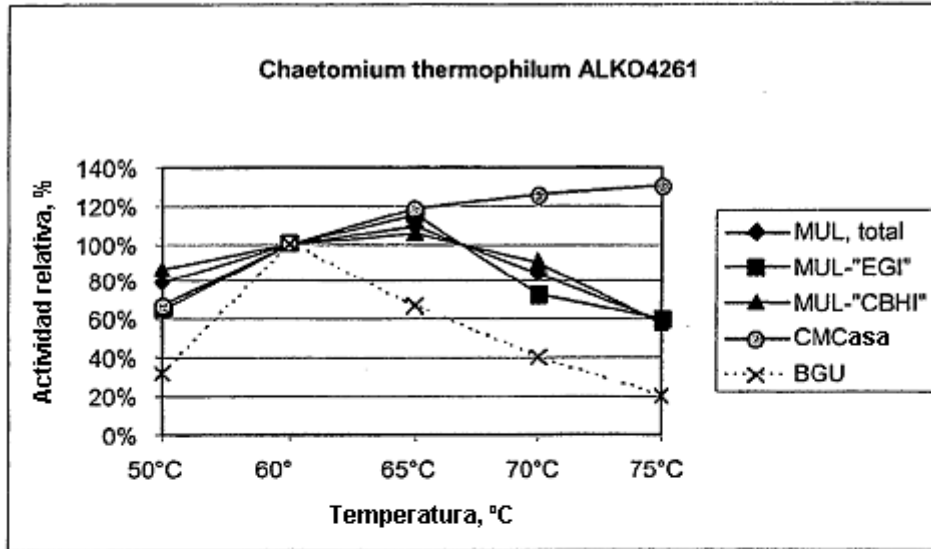


Fig. 1F.

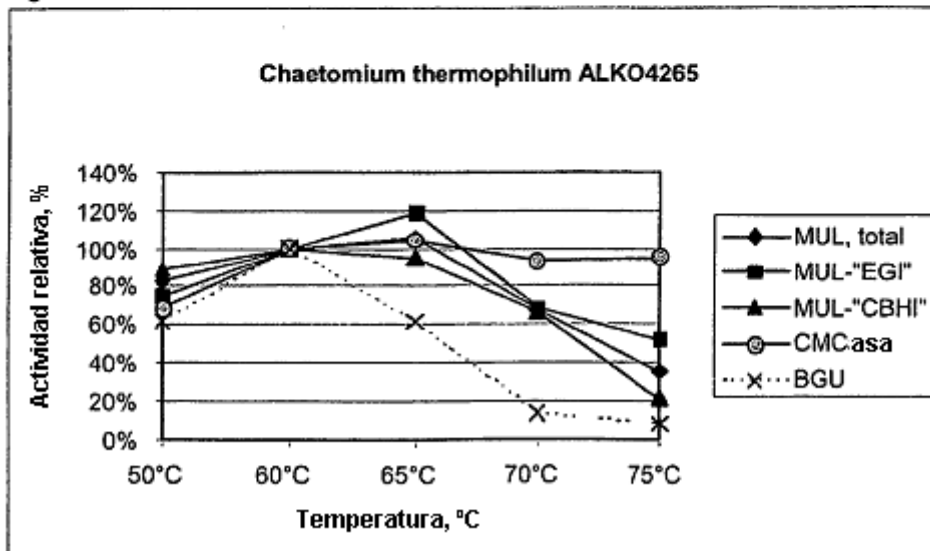


Fig. 2.

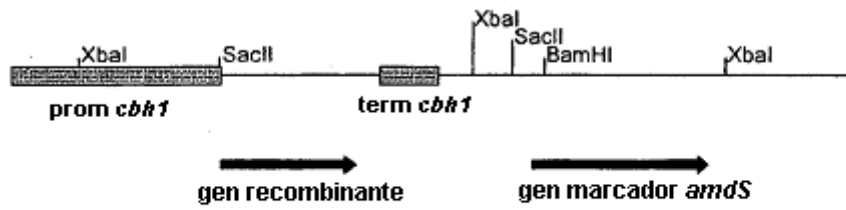


Fig. 3A.

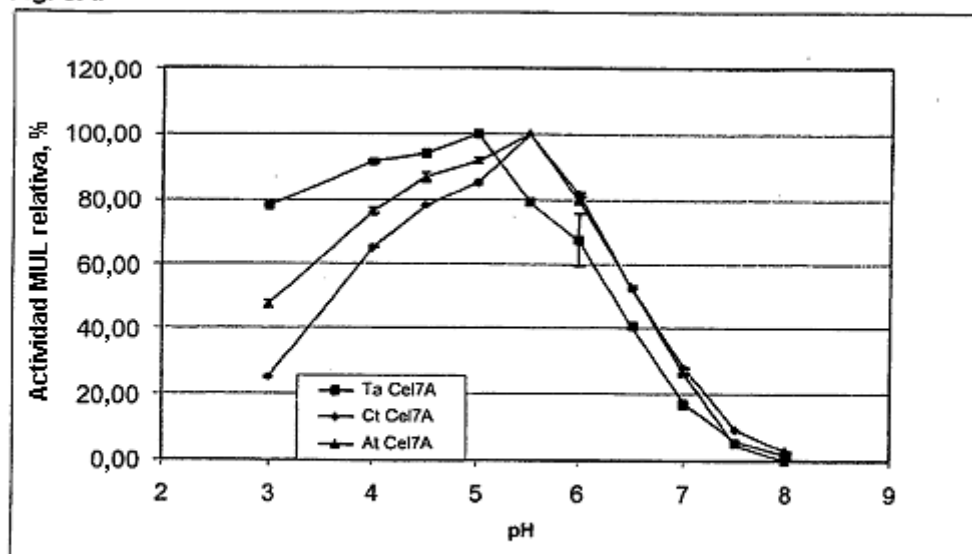


Fig. 3B.

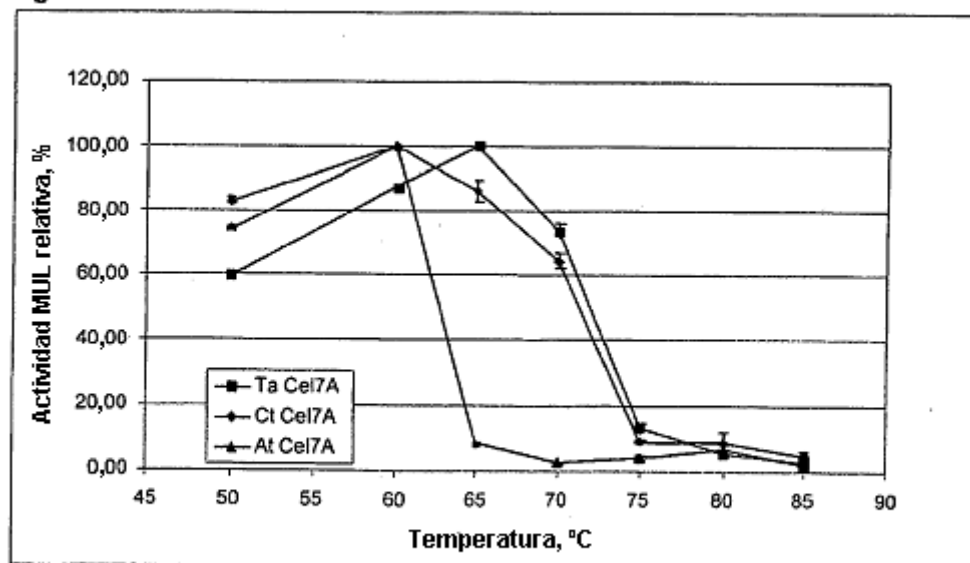


Fig. 4A.

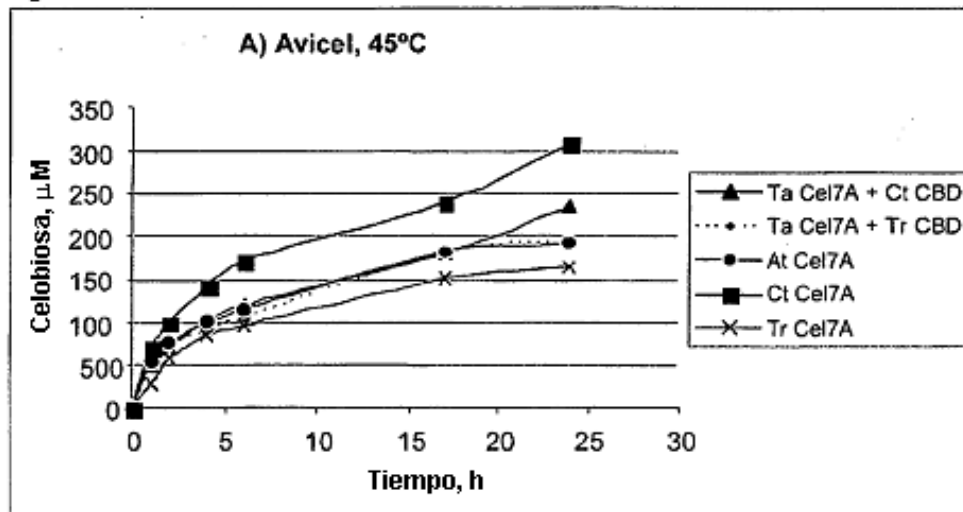


Fig. 4B.

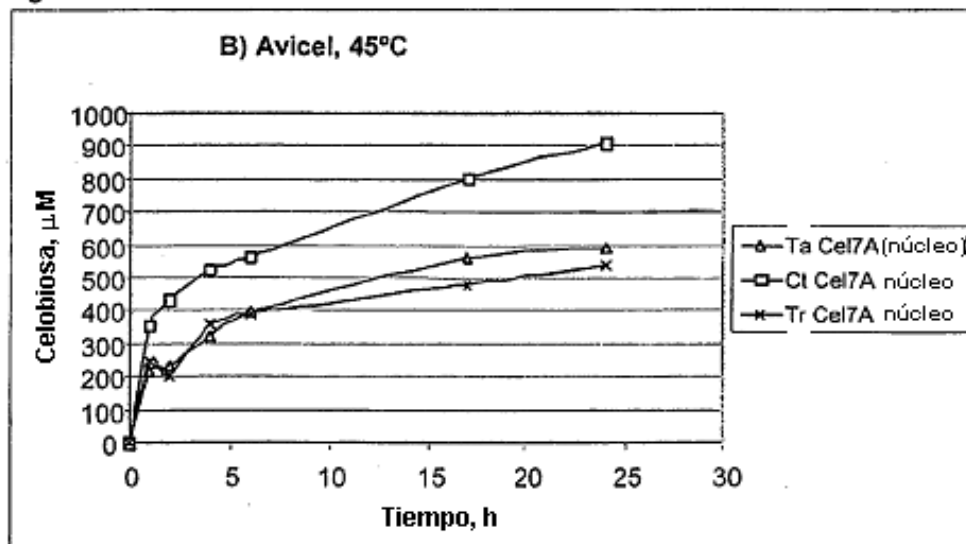


Fig. 5A.

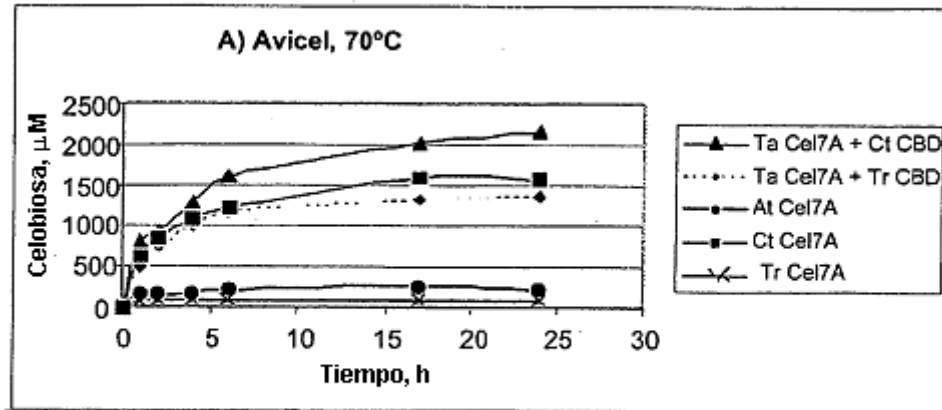


Fig. 5B.

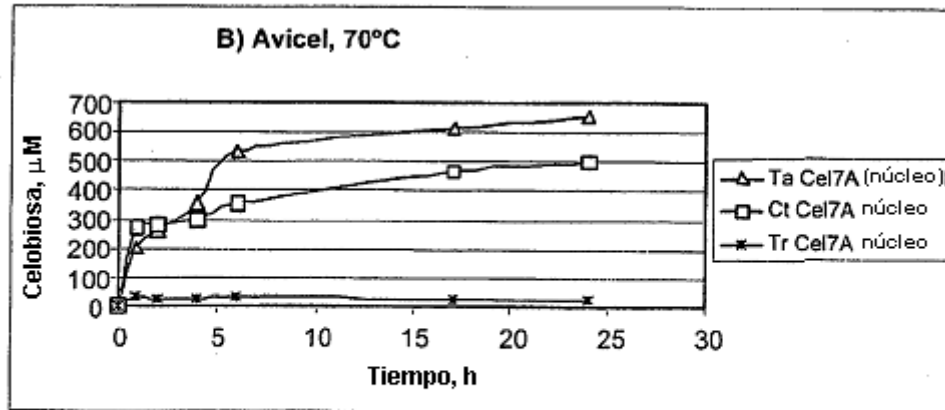


Fig. 6A.

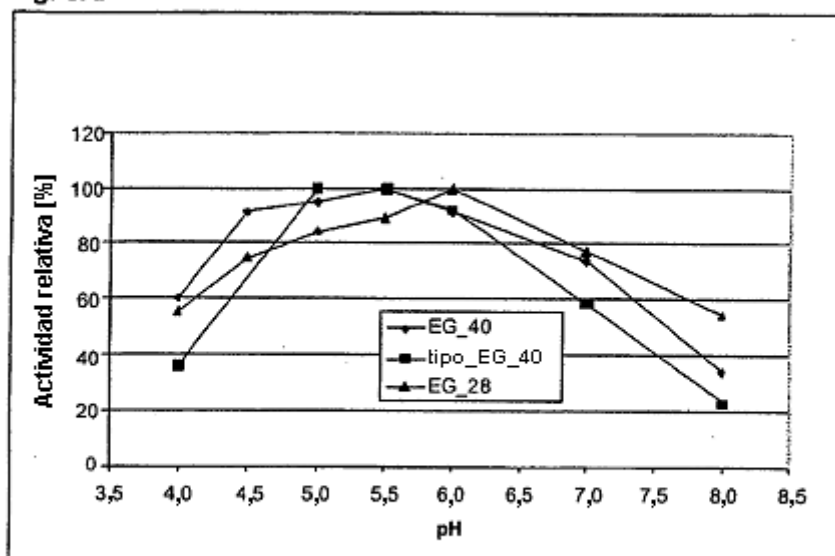


Fig. 6B.

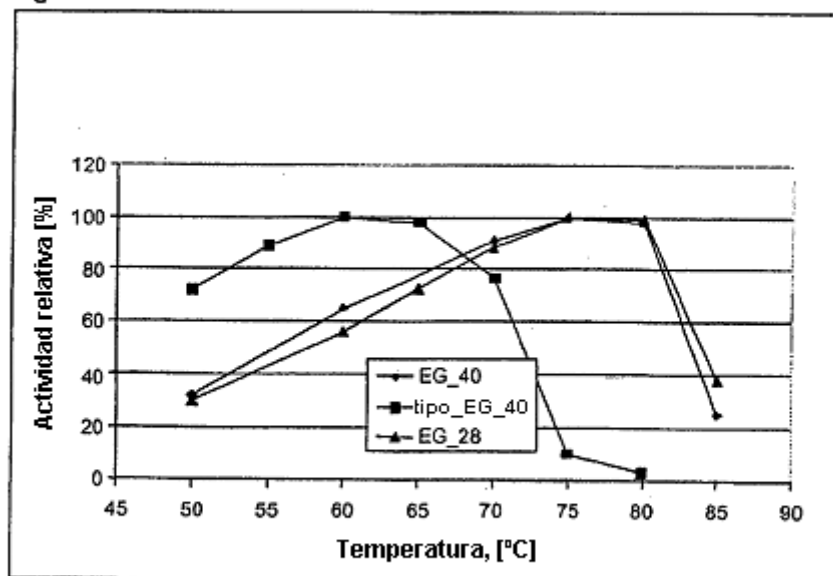


Fig. 7A.

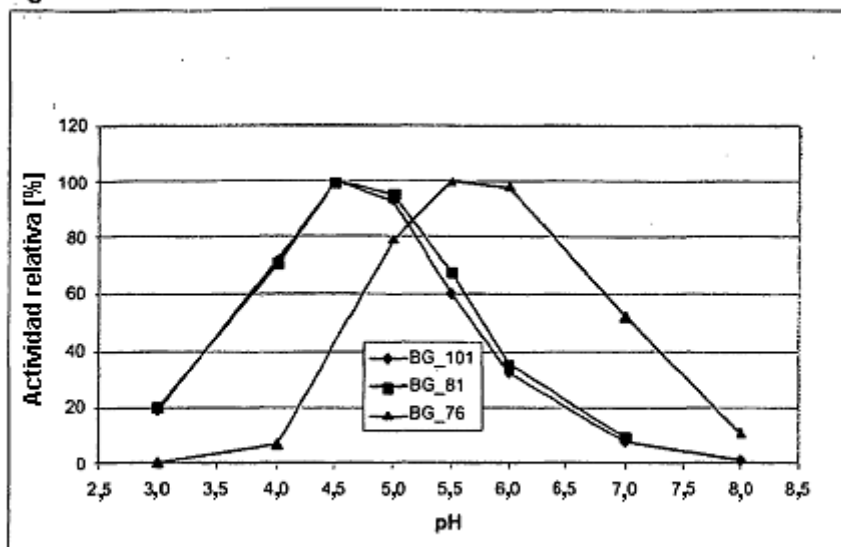


Fig. 7B.

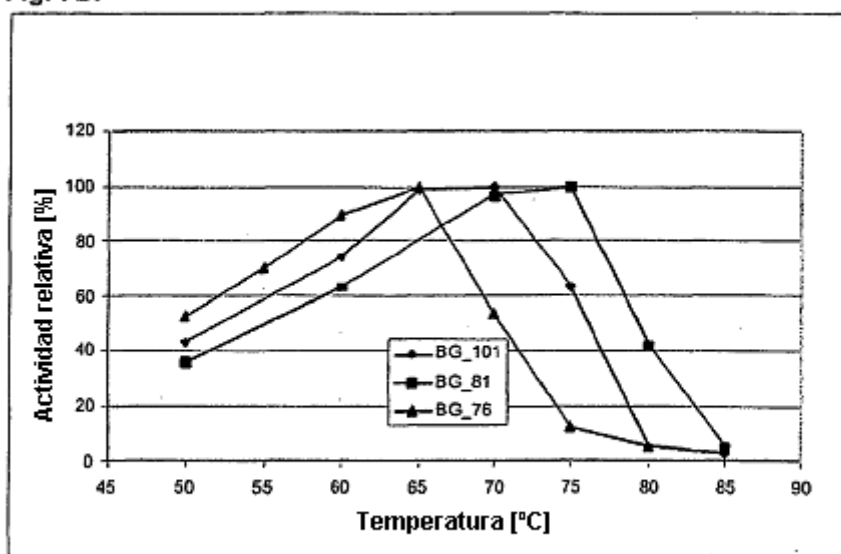


Fig. 8A.

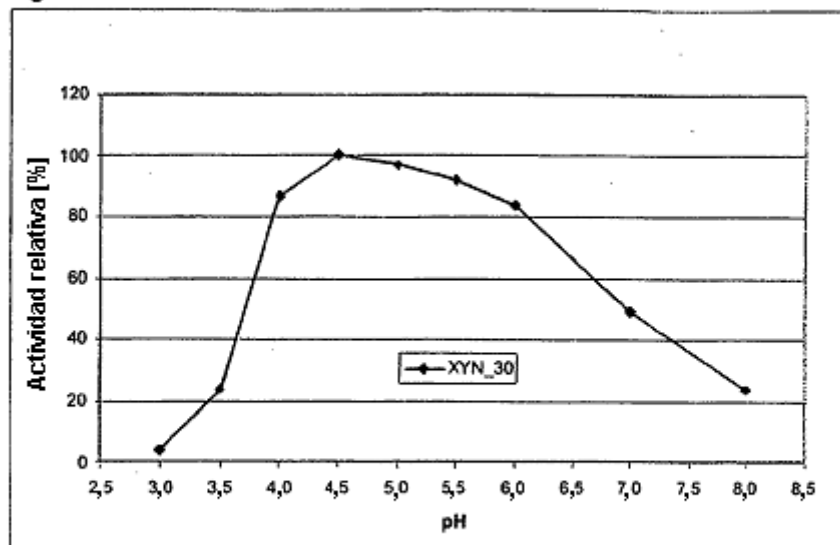


Fig 8B.

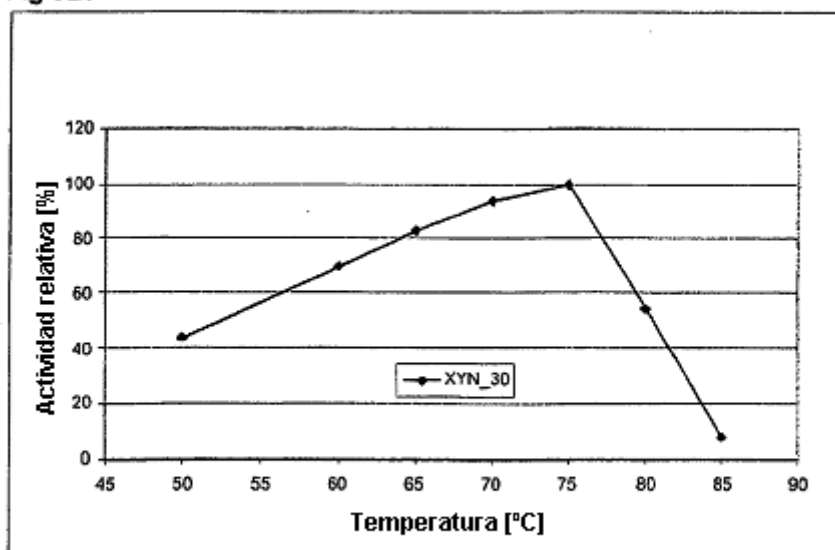


Fig. 9.

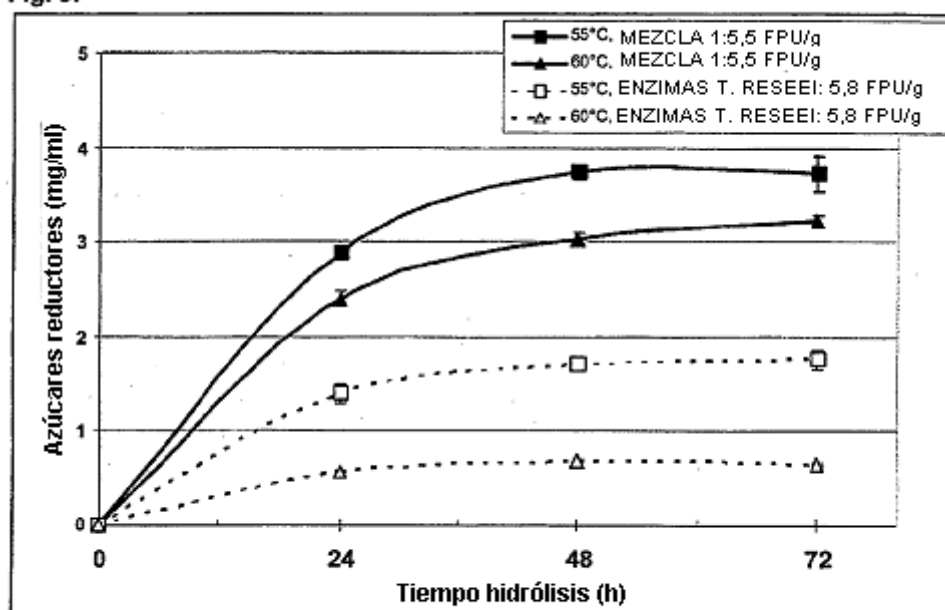


Fig. 10.

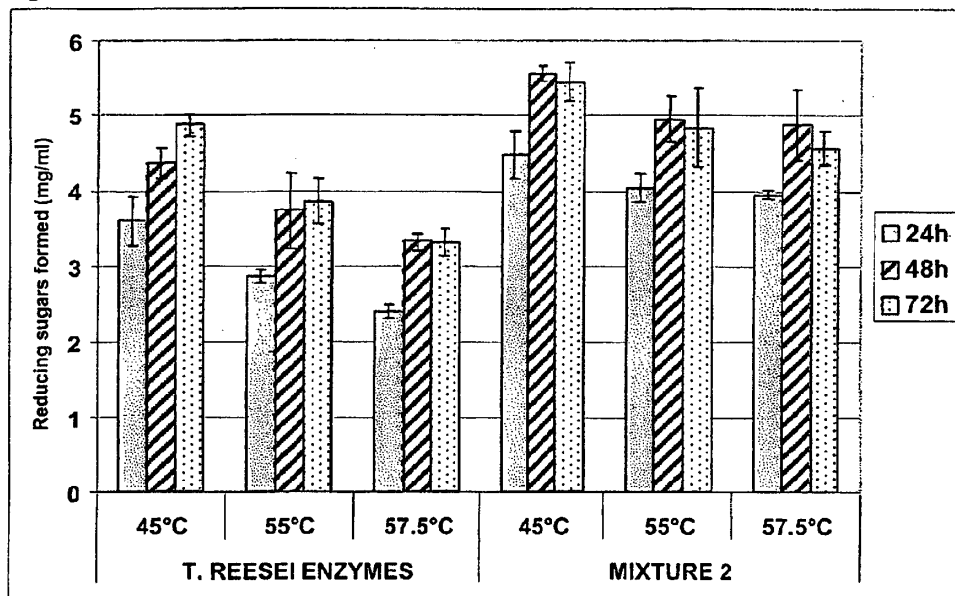


Fig. 11.

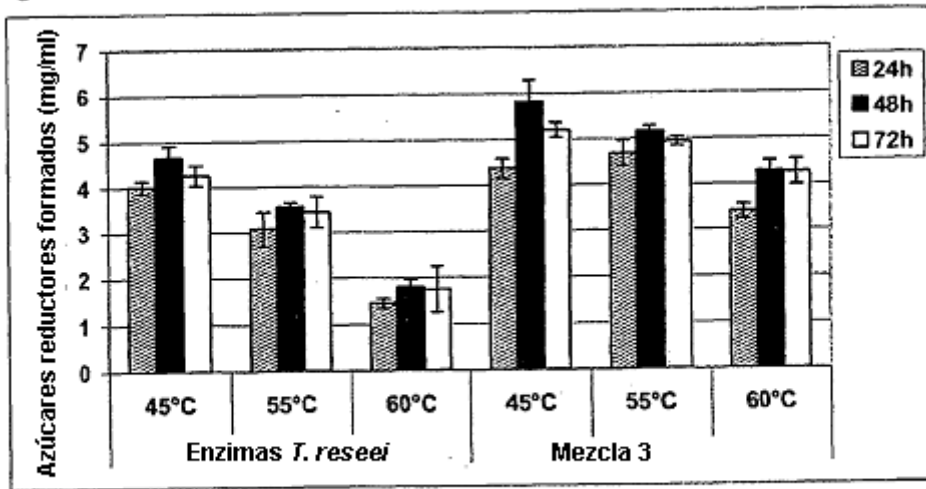


Fig. 12.

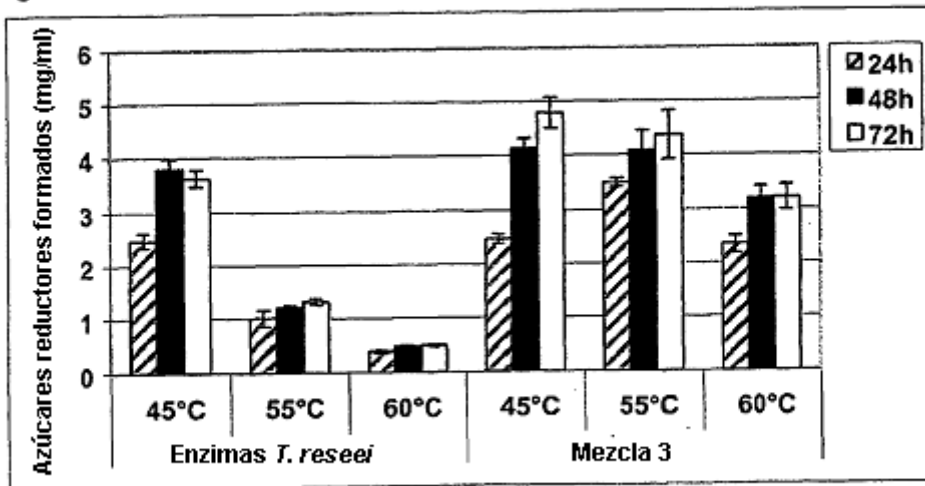


Fig. 13.

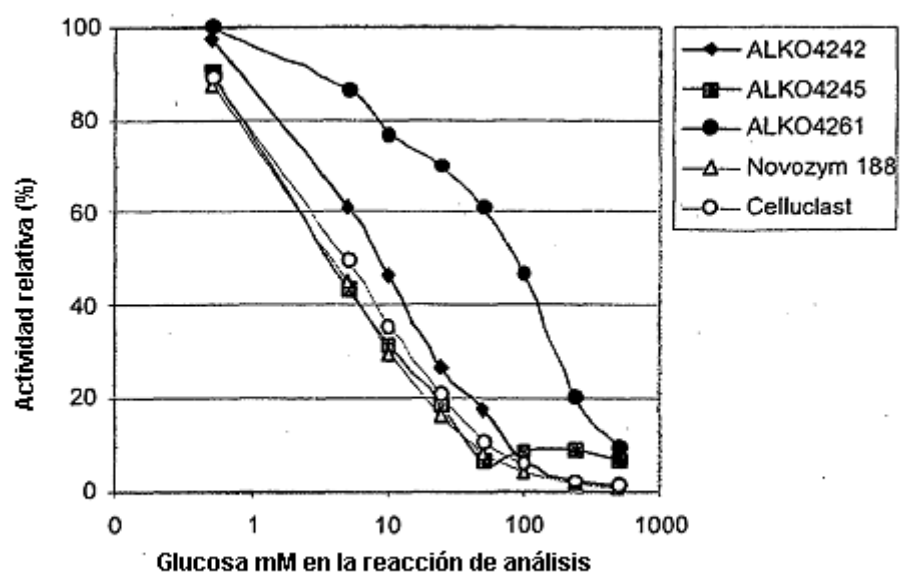


Fig. 14.

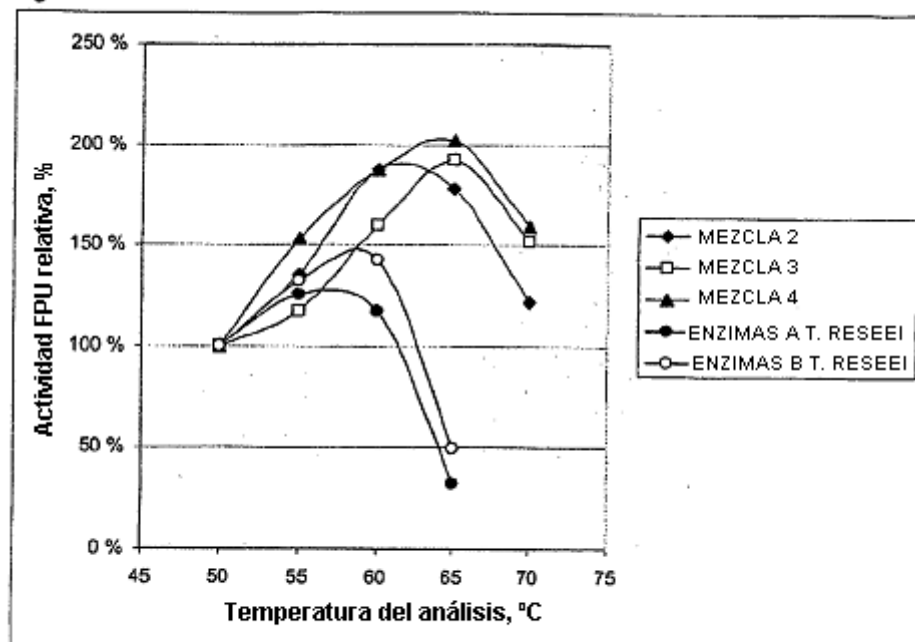


Fig. 15.

