

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 501 944**

51 Int. Cl.:

C12P 19/02 (2006.01)

C12P 19/14 (2006.01)

C12P 7/10 (2006.01)

C12N 9/42 (2006.01)

C12N 15/56 (2006.01)

C12N 15/62 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.12.2006 E 06830936 (8)**

97 Fecha y número de publicación de la concesión europea: **13.08.2014 EP 1963498**

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:

02.10.2014

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
VIAKARI, LIISA**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 501 944 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 útiles p. ej. 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 xilanasas, y a preparaciones de enzimas y usos de las mismas. Se describen nuevos polipéptidos celulolíticos, los polinucleótidos que los codifican, y vectores y células anfitrionas que contienen los polinucleótidos. Adicionalmente, se describen otros usos de los polipéptidos y un método para prepararlos.

Antecedentes de la invención

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

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

La mayoría de los carbohidratos en las plantas están en la 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 más accesible a la hidrólisis la fracción de celulosa. La fracción de celulosa se hidroliza a continuación para obtener azúcares que pueden ser fermentados por levadura a etanol. La lignina se obtiene como un co-producto principal que se puede utilizar como carburante sólido.

Los costes de producción de bioetanol son altos y la producción de energía es baja, y existe una continua búsqueda para 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 solamente a una cantidad limitada a escala industrial, y especialmente cuando se utiliza material fuertemente lignificado tal como 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 eficacia de la hidrólisis enzimática del material celulósico (Badger 2002).

El documento US 2002/0192774 A1 describe un procedimiento continuo para la conversión de biomasa lignocelulósica sólida en productos carburantes combustibles. Después del pretratamiento mediante oxidación húmeda o explosión de vapor la biomasa se separa parcialmente en celulosa, hemicelulosa y lignina, y a continuación se somete a hidrólisis parcial utilizando una o más enzimas carbohidrasas (CE 3.2). Se proporciona como ejemplo Celluclast™, un producto comercial de Novo Nordisk A/S que contiene actividades celulasa y xilanasas.

El documento US 2004/0005674 A1 describe nuevas mezclas de enzimas que se pueden utilizar directamente sobre un sustrato de lignocelulosa, por medio de las cuales se pueden evitar los productos de desecho tóxicos formados durante los procedimientos de pretratamiento, y se puede ahorrar energía. La mezcla de enzimas sinérgicas contiene una celulasa y una enzima auxiliar tal como celulasa, xilanasas, 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 xilanasas y celulasa de *Trichoderma*.

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

ensayo dos preparaciones de celulasa comerciales de *Trichoderma reesei*, y dos preparaciones novedosas producidas por cepas mutantes de *Trichoderma* sp. y *Penicillium* sp. La preparación de *Trichoderma* sp. mostró un rendimiento significativamente mejor que las otras preparaciones. Se creyó que el mejor rendimiento se debía al menos en parte a una actividad beta-glucosidasa significativamente mayor, que alivia la inhibición por producto de celobiohidrolasa y endoglucanasa.

El documento US 2004/0053373 A1 se refiere un método de conversión de celulosa en glucosa mediante tratamiento de un sustrato lignocelulósico pretratado con una mezcla de enzimas que comprende celulasa y una celobiohidrolasa I (CBHI) modificada. La CBHI ha sido modificada inactivando su dominio de unión a celulosa (CBD). Las ventajas de la modificación de CBHI son, p. ej. 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 2005/0164355 A1 describe un método para degradar material lignocelulósico con una o más enzimas celulolíticas en presencia de al menos un tensioactivo. También se pueden usar enzimas adicionales tales como hemicelulasas, esterasa, peroxidasa, proteasa, lacasa o una mezcla de las mismas. La presencia de 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 lignocelulosa incluyendo CBH, EG y BG.

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

Se describen celobiohidrolasas (CBHS) p. ej., en el documento WO 03/000941, que se refiere a enzimas CBHI obtenidas a partir de diversos hongos. No se proporcionan las propiedades fisiológicas de las enzimas, ni ningún ejemplo de sus usos. Hong *et al.* (2003b) caracteriza CBHI de *Thermoascus aurantiacus* producida en levadura. No se describen las aplicaciones de la enzima. Tuohy *et al.* (2002) describen tres formas de celobiohidrolasas de *Talaromyces emersonii*.

Se describen las endoglucanasas de la familia cel5 (EG fam 5) p. ej., en el documento WO 03/062409, que se refiere a composiciones que comprenden al menos dos enzimas termoestables para su uso en aplicaciones para forraje. Hong *et al.* (2003a) describen la producción de endo- β -1,4-glucanasa termoestable a partir de *T. aurantiacus* en levadura. No se explican las aplicaciones. El documento WO 01/70998 se refiere a β -glucanasas de *Talaromyces*. También describe β -glucanasas de *Talaromyces emersonii*. Se comentan las aplicaciones para alimentos, forrajes, bebidas, elaboración de la cerveza, y detergentes. No se menciona la hidrólisis de lignocelulosa. El documento WO 98/06858 describe beta-1,4-endoglucanasa de *Aspergillus niger* y comenta las aplicaciones para forraje y alimento de la enzima. El documento WO 97/13853 describe métodos para el escrutinio de fragmentos de ADN que codifican enzimas en genotecas de ADNc. La genoteca de ADNc es de origen de levadura o fúngico, 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*.

Se describen endoglucanasas de la familia cel7 (EG fam 7) p. ej., en el documento US 5. 912. 157, que pertenece a la endoglucanasa de *Myceliphthora* y sus homólogos y las aplicaciones de la misma en detergentes, productos textiles, y pasta de celulosa. El documento US 6. 071. 735 describe celulasas que muestran una alta actividad endoglucanasa en condiciones alcalinas. Se comentan los usos como detergente, en aplicaciones de pasta de celulosa y papel, y textiles. 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 CBD.

Se describen endoglucanasas de la familia cel45 (EG fam 45) p. ej., en el documento US 6. 001. 639, que se refiere a las enzimas que tienen actividad endoglucanasa y que tienen dos secuencias de aminoácidos conservadas. Se comentan en general los usos en aplicaciones textiles, detergentes, y de pasta de celulosa y papel y se menciona el tratamiento de material lignocelulósico pero no se proporcionan ejemplos. El documento WO 2004/053039 se refiere a las aplicaciones detergentes de las endoglucanasas. El documento US 5. 958. 082 describe el uso de endoglucanasa, especialmente de *Thielavia terrestris* en aplicaciones textiles. 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 celulosa. No se menciona la hidrólisis de lignocelulosa.

Una pequeña cantidad de beta-glucosidasa (BG) mejora la hidrólisis de la biomasa a glucosa hidrolizando la celobiosa producida por celobiohidrolasas. La conversión de celobiosa en glucosa es normalmente la principal etapa limitante de la velocidad. Se describen beta-glucosidasas, p. ej. en el documento US 2005/0214920, que se refiere a BG de *Aspergillus fumigatus*. La enzima ha sido producida en *Aspergillus oryzae* y *Trichoderma reesei*. Se comenta en general el uso de la enzima en la degradación de biomasa o en aplicaciones detergentes, pero no se ilustra. El documento WO 02/095014 describe una enzima de *Aspergillus oryzae* que tiene actividad celobiasa. Se comenta en general el uso en la producción de etanol a partir de biomasa, pero no se ilustra. El documento WO 2005/074656 describe polipéptidos que tienen actividad potenciadora celulolítica derivada, p. ej., 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 6022725 describe la clonación y la amplificación del gen de la beta-glucosidasa de *Trichoderma reesei*, y el documento US 6103464 describe un método para detectar el ADN que codifica una beta-glucosidasa de un hongo filamentoso. No se proporcionan ejemplos de aplicación.

5 Se describen xilanasas p. ej., en el documento FR 2786784, que se refiere a una xilanasas termoestable útil, p. ej., en el tratamiento de forraje animal y en la fabricación de pan. La enzima deriva de un hongo termófilo, concretamente del género *Thermoascus*.

10 El documento US 6197564 describe enzimas que tienen actividad de 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 para forraje y panadería. El documento WO 01/42433 describe xilanasas termoestables de *Talaromyces emersonii* para uso en aplicaciones de alimentación y forraje.

15 Las enzimas celulolíticas mejor investigadas y más ampliamente aplicadas de origen fúngico derivan de *Trichoderma reesei* (el anamorfo de *Hypocrea jecorina*). Por consiguiente casi la mayor parte de las celulasas fúngicas disponibles en el mercado derivan de *Trichoderma reesei*. Sin embargo, la mayor parte de las celulasas de hongos menos conocidos no han sido aplicadas en los procedimientos de importancia práctica tales como la degradación de material celulósico, incluyendo lignocelulosa.

20 Existe una necesidad continua de nuevos métodos de degradación de sustratos celulósicos, en particular sustratos lignocelulósicos, y de nuevas enzimas y mezclas de enzimas, que mejoren la eficiencia de la degradación. También hay una necesidad de procesos y enzimas, que funcionen a altas temperaturas, permitiendo de ese modo el uso de una elevada consistencia de biomasa y conduciendo a elevadas concentraciones de azúcar y etanol. Este enfoque puede conducir a ahorros significativos en los costes de energía e inversiones. La elevada 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

25 Sorprendentemente, se ha descubierto ahora que las enzimas celulolíticas, y especialmente las celobiohidrolasas obtenibles a partir de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum* son particularmente útiles en la hidrólisis de material celulósico. Además de las 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 eficaces a las temperaturas de hidrólisis convencionales. Esto las convierte en extremadamente bien adaptadas para variar procesos de hidrólisis de sustratos celulósicos llevados a cabo tanto a temperaturas convencionales y a temperaturas elevadas.

La presente invención proporciona un método para el tratamiento de material celulósico con celobiohidrolasa, endoglucanasa y beta-glucosidasa, en donde dicha celobiohidrolasa es una proteína de fusión como se define en la reivindicación 1.

35 La invención proporciona adicionalmente una preparación enzimática que comprende celobiohidrolasa, endoglucanasa y beta-glucosidasa, en donde dicha celobiohidrolasa es una proteína de fusión como se define en la reivindicación 13.

Asimismo se proporciona el uso de dicha preparación de enzima para degradar material celulósico, así como el uso de dicho método en un procedimiento para preparar etanol a partir de material celulósico.

40 Se describen polipéptidos que comprenden un fragmento que tiene actividad celulolítica y que se seleccionan del grupo que consiste en:

45 a) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos 66% con el SEQ ID NO: 4, una identidad de 79% con el SEQ ID NO: 6, una identidad de 78% con el SEQ ID NO: 12, una identidad de 68% con el SEQ ID NO: 14, una identidad de 72% con el SEQ ID NO: 16, una identidad de 68% con el SEQ ID NO: 20, una identidad de 74% con el SEQ ID NO: 22 o 24, o una identidad de 78% con el SEQ ID NO: 26;

b) una variante de a) que comprende un fragmento que tiene actividad celulolítica;

y

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

También se describen polinucleótidos aislados seleccionados del grupo que consiste en:

50 a) una secuencia de nucleótidos del SEQ ID NO: 3, 5, 11, 13, 15, 19, 21, 23 o 25, o una secuencia que codifica dichos polipéptidos;

b) una hebra complementaria de a)

c) un fragmento de a) o b) que comprende al menos 20 nucleótidos; y

d) una secuencia que es degenerada como resultado del código genético con respecto a una cualquiera de las secuencias definidas en a), b) o c).

5 Se describen un vector, que comprende dicho polinucleótido como secuencia heteróloga, y una célula anfitriona que comprende dicho vector. Asimismo se describen cepas de *Escherichia coli* que tienen el número de acceso DSM 16728, DSM 16729, DSM 17324, DSM 17323, DSM 17729, DSM 16726, DSM 16725, DSM 17325 o DSM 17667 también se incluyen en la invención.

10 Se describen preparaciones de enzimas que comprenden al menos uno de los nuevos polipéptidos, y el uso de dicho polipéptido o preparación de enzima en la industria del combustible, textil, detergente, pasta de celulosa y papel, alimentos, forraje o bebidas.

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

15 a) un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos 66% con el SEQ ID NO: 4, una identidad de 79% con el SEQ ID NO: 6, una identidad de 78% con el SEQ ID NO: 12, una identidad de 68% con el SEQ ID NO: 14, una identidad de 72% con el SEQ ID NO: 16, una identidad de 68% con el SEQ ID NO: 20, una identidad de 74% con el SEQ ID NO: 22 o 24, o una identidad de 78% con el SEQ ID NO: 26;

b) una variante de a) que comprende un fragmento que tiene actividad celulolítica;

y

c) un fragmento de a) o b) que tiene actividad celulolítica,

20 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.

25 También se describe 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

30 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 pH del análisis fue de 5,0 (actividad MUL) o 4,8 (CMCasa o BGU). La actividad obtenida a 60°C se establece como la actividad relativa del 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.

35 Figura 2. Imagen esquemática de los casetes de expresión utilizados en la transformación de protoplastos de *Trichoderma reesei* para producir las proteínas fúngicas recombinantes. Los genes recombinantes estuvieron bajo el control del promotor de *cbh1* (prom *cbh1*) de *T. reesei* (*cel7A*) y la terminación de la transcripción se garantizó utilizando la secuencia terminadora de *CBH1* (term *CBH1*) de *T. reesei*. Se incluyó el gen *amdS* como marcador de la transformación.

40 Figura 3. A) pH óptimos de las preparaciones de proteínas recombinantes CBH/Cel7 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 la media (± DT) de tres mediciones separadas. B) estabilidad térmica de las preparaciones de proteínas recombinantes CBH/Cel7 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 proporcionan como la media (± DT) de tres mediciones separadas. Ambas reacciones contenían BSA (100 µg/ml) como estabilizador.

45 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 M. A) celobiohidrolasas que albergan un CBD, B) celobiohidrolasas (núcleo) sin un CBD.

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

Figura 7. A) Se determinó la dependencia del pH de la actividad de BG_101/Cel3A de *Acremonium*, BG_76/Cel3A de *Chaetomium*, y BG_81/Cel3A de *Thermoascus* producida de manera heteróloga con un sustrato de 4-nitrofenil-β-D-glucopiranosido en una reacción de 10 min a 50°C. B) Se determinó el óptimo de temperatura de βG_101/Cel3A de *Acremonium*, βG_76/Cel3A de *Chaetomium*, y βG_81/Cel3A de *Thermoascus* a pH 4,5, 5,5, y 4,5, respectivamente. La reacción que contenía 4-nitrofenil-β-D-glucopiranosido como sustrato se llevó a cabo durante 60 minutos, se añadió BSA (100 µg/ml) como estabilizador.

Figura 8. A) Se determinó la dependencia del pH de la actividad xilanasa XYN_30/Xyn10A *Thermoascus* producida de manera heteróloga con sustrato de xilano de abedul en una reacción de 10 min a 50°C. B) Se determinó el óptimo de temperatura de XYN_30/Xyn10A a pH 5,3 en una reacción de 60 min, se añadió BSA (100 µg/ml) como estabilizador.

Figura 9. Hidrólisis de fibra de píceas 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 dosificación de enzima se administró mediante FPU/g de materia seca de sustrato, FPU analizado a 50°C, pH 5. La hidrólisis se llevó a cabo durante 72 horas a pH 5, mezclando. Los resultados se proporcionan como la media (± DT) de tres mediciones separadas.

Figura 10. Hidrólisis 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 dosificación de enzima para la "MEZCLA 2" fue de 5 FPU/g de materia seca de sustrato y para las enzimas de *T. reesei* de 5 FPU/g de materia seca de Celluclast con un suplemento de 100 nkat/g de materia seca de Novozym 188 (la actividad en 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 dan como la media (± 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 rastrojo de maíz sometido a explosión de vapor de agua (10 mg/ml) con una mezcla de enzimas termófilas que contiene una nueva xilanasa termófila de *Thermoascus aurantiacus* (MEZCLA 3) y enzimas de *T. reesei* a 45, 55 y 60°C. La dosificación para la enzima fue para la "MEZCLA 3" de 5 FPU/g de materia seca de sustrato y para las enzimas de *T. reesei* de 5 FPU/g de materia seca de Celluclast con un suplemento de 100 nkat/g de materia seca de Novozym 188 (la actividad en 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 proporcionan como la media (± 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 fibra de píceas sometida a explosión de vapor de agua (10 mg/ml) con una mezcla de enzimas termófilas que contiene una nueva xilanasa termófila XYN_30/Xyn10A de *Thermoascus aurantiacus* (MEZCLA 3) y enzimas de *T. reesei* a 45, 55 y 60°C. La dosificación de enzima para la "MEZCLA 3" fue de 5 FPU/g de materia seca de sustrato y para las enzimas de *T. reesei* de 5 FPU/g de materia seca de Celluclast con un suplemento de 100 nkat/g de materia seca de Novozym 188 (la actividad en 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 proporcionan como la media (± DT) de tres mediciones separadas.

Figura 13. Efecto de la glucosa sobre la actividad de diferentes preparaciones de β-glucosidasa. El análisis convencional utilizando p-nitrofenil-β-D-glucopiranosido como sustrato se llevó a cabo en presencia de glucosa en la mezcla de análisis. La actividad se presenta como el 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 en condiciones convencionales (50°C, 1 h).

Figura 15. Actividad celulasa relativa de dos cepas de *T. reesei* diferentes cultivadas en medios que contienen Nutriosa sin tratar (N0) o Nutriosa pretratada con BG_81/Cel3A (NBG81) como fuente de carbono.

Descripción detallada de la invención

La celulosa es el principal componente estructural de las plantas superiores. Proporciona a las células vegetales una 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 glicosídicos β-1,4. La

celobiosa es la unidad repetitiva más pequeña de la celulosa. En las paredes celulares la celulosa está empaquetada en láminas orientadas de manera diversa, que están incluidas en una matriz de hemicelulosa y lignina. La hemicelulosa es un grupo heterogéneo de polímeros carbohidratados que contiene principalmente diferentes glucanos, xilanos y mananos. La hemicelulosa consiste en una cadena principal lineal con residuos unidos por β -1,4 sustituidos con cadenas laterales cortas que contienen normalmente acetilo, glucuronilo, arabinosilo y galactosilo. La hemicelulosa puede estar entrecruzada químicamente con lignina. La lignina es un polímero entrecruzado complejo de unidades de p-hidroxifenilpropano diversamente sustituido que proporciona resistencia a la pared celular para soportar 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 y e inaccesible y el material bioquímico más abundante en la biosfera. Los materiales que contienen lignocelulosa son, por ejemplo: virutas de madera dura y de madera blanda, pasta de madera, serrín y residuos industriales de forestales y de madera; biomasa agrícola tal como paja de cereales, pasta de remolacha azucarera, rastrojo y mazorcas de maíz, bagazo de caña de azúcar, tallos, hojas, cáscaras, cascarrillas, y similares; productos de desecho tales como residuos sólidos urbanos, papel de periódico y de oficina de desecho, residuos de la molienda de, p. ej. cereales; cultivos destinados a la producción de energía (p. ej., sauce, chopo, pasto varilla 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, p. ej. materiales vegetales tales como madera incluyendo madera blanda y dura, cultivos herbáceos, residuos agrícolas, residuos de pasta de celulosa y de papel, papel usado, residuos de la industria alimenticia y de forraje, etc. Las fibras textiles, como el algodón, las fibras derivadas de algodón, lino, cáñamo, yute y fibras celulósicas elaboradas por el hombre como modal, viscosa, liocel son ejemplos específicos de materiales celulósicos.

El material celulósico es degradado en la naturaleza por diversos organismos diferentes, 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 generalmente tres tipos de enzimas hidrolíticas: (1) Endoglucanasas que cortan los enlaces glicosídicos β -1,4 internos; (2) Exocelobiohidrolasas que cortan el disacárido celobiosa del extremo de la cadena polimérica de celulosa; (3) β -1,4-glucosidasas que hidrolizan la celobiosa y otros celo-oligosacáridos cortos a glucosa. En otras palabras, los tres principales grupos de celulasas son celobiohidrolasas (CBH), endoglucanasas (EG) y β -glucosidasas (BG).

La degradación de los sustratos que contienen celulosa más complejos requiere una amplia gama de enzimas diferentes. Por ejemplo la lignocelulosa es degradada por las hemicelulasas, como xilanasas y mananasas. La hemicelulasa es un enzima que hidroliza la 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. Celulasas según se utiliza en la presente memoria incluyen celobiohidrolasas, endoglucanasas y β -glucosidasas.

T. reesei tiene un sistema de celulasa bien conocido y eficaz, que contiene dos CBH, dos EG y BG principales y varias minoritarias. La CBHI de *T. reesei* (Cel7A) 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. La CBHII de *T. reesei* (Cel6A) corta el azúcar desde el extremo no reductor de la cadena de celulosa, tiene un dominio de unión a celulosa N-terminal y pueden constituir hasta 20% de la proteína secretada total. Las endoglucanasas EGI (Cel7B), y EGV (Cel45A) tienen un CBD en su extremo C-terminal, EGII (Cel5A) tiene un CBD N-terminal y EGIII (Cel12A) no tiene ningún dominio de unión a celulosa. La CBHI, CBHII, EGI y EGII también se denominan "celulasas principales de *Trichoderma*" que comprenden juntas 80-90% de las proteínas secretadas totales. Los expertos en la técnica saben que una enzima puede ser activa sobre diversos sustratos y las actividades enzimáticas se pueden medir utilizando diferentes sustratos, métodos y condiciones. La identificación de las diferentes actividades celulolíticas son comentadas, p. ej., por Van Tilbeurgh *et al.* 1988.

Además de un dominio/núcleo catalítico 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 localizados 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 de 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 enzimas que escinden la celulosa desde el extremo de la cadena de glucosa y producen principalmente celobiosa. También se denominan 1,4- β -D-glucanocelobiohidrolasas o celulosa 1,4- β -celobiosidasas. Hidrolizan los enlaces 1,4- β -D-glicosídicos desde los extremos reductores o no reductores de un polímero que contiene dichos enlaces, tales como celulosa, por

medio de lo cual se libera celobiosa. Se han aislado dos CBH diferentes de *Trichoderma reesei*, CBHI y CBHII. Estas tienen una estructura modular que consiste en un dominio catalítico conectado a un dominio de unión a celulosa (CBD). También existen celobiohidrolasas en la naturaleza que carecen de CBD.

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

10 "Beta-glucosidasa" o "BG" o "βG" se refiere a enzimas que degradan oligosacáridos solubles pequeños incluyendo celobiosa a glucosa. Se clasifican como EC 3.2.1.21. Son beta-D-glucósido glucohidrolasas, que típicamente catalizan la hidrólisis de residuos de 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 las celobiohidrolasas, por tanto la degradación de la celobiosa es importante para superar la inhibición por producto final de celobiohidrolasas.

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

20 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 glicosilhidrolasas son enzimas multifuncionales que contienen dominios catalíticos que pertenecen a diferentes familias de glicosilhidrolasas. La familia 3 consiste en beta-glucosidasas (EC 3.2.1.21) tales como BG_81 de Ta, BG_101 de At y BG_76 de Ct descritas en la presente memoria. La familia 5 (antes conocida como celA) consiste principalmente en endoglucanasas (EC 3.2.1.4) como EG_28 de Ta descrita en la presente memoria. La familia 7 (antes familia de celulasa CELC) contiene endoglucanasas (EC 3.2.1.4) y celobiohidrolasas (EC 3.2.1.91) tales como EG_54 de Ct, CBH de Ta, CBH_A de At, CBH_C de At y CBH de Ct descritas en la presente memoria. La familia 10 (antes CELF) consiste principalmente en xilanasas (EC 3.2.1.8) tales como XYN_30 de Ta y XYN_60 de At descritas en la presente memoria. La familia 45 (antes CELK) contiene endoglucanasas (EC 3.2.1.4), tales como la EG_40 de At y de tipo EG_40 de At descritas en la presente memoria.

35 Las enzimas celulolíticas útiles para hidrolizar material celulósico son obtenibles de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*. "Obtenible de" significa que se pueden obtener de dichas especies, pero no excluye la posibilidad de obtenerlas de otras fuentes. En otras palabras, se pueden originar a partir de cualquier organismo, incluidas plantas. Preferiblemente se originan a partir de microorganismos, por ejemplo bacterias u hongos. Las bacterias pueden ser, p. ej., 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*, *Humicola*, *Hypocrea*, *Lentinus*, *Melanocarpus*, *Myceliophthora*, *Myriococcum*, *Neurospora*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Pleurotus*, *Podospora*, *Polyporus*, *Rhizoctonia*, *Scytalidium*, *Pycnoporus*, *Trametes* y *Trichoderma*.

Las enzimas son obtenibles de *Thermoascus aurantiacus* cepa ALKO4242 depositada como CBS 116239, cepa ALKO4245 depositada como CBS 116240 actualmente clasificada como *Acremonium thermophilum*, o *Chaetomium thermophilum* cepa ALKO4265 depositada como CBS 730.95.

45 La celobiohidrolasa puede comprender una secuencia de aminoácidos que tiene una identidad de al menos el 80% con el SEQ ID NO: 2, 4, 6 u 8, o un fragmento enzimáticamente activo del mismo.

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

Celobiohidrolasa	Gen	Obtenible de	CBD	ácido nucleico SEQ ID NO:	amino ácido SEQ ID NO:
CBH de Ct	<i>cel7A</i> de Ct	<i>C. thermophilum</i>	+	7	8

Estas CBH tienen una inhibición de celobiosa constante ventajosa en comparación con la de la CBH de *Trichoderma reesei*, y muestran mejores resultados de hidrólisis cuando se someten a ensayo diversos sustratos celulósicos. Los SEQ ID NO: 2 y 4 no lo comprenden un CBD. Se pueden obtener resultados de hidrólisis particularmente mejorados cuando un dominio de unión a celulosa (CBD) se ancla una CBH que no tiene un CBD propio. El CBD se puede obtener p. ej. a partir de una especie de *Trichoderma* o *Chaetomium*, y se ancla a la CBH preferiblemente a través de un conector. La proteína de fusión resultante que contiene una región núcleo de CBH anclada a un CBD a través de un conector puede comprender una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 28 o 30. Los polinucleótidos que comprenden una secuencia del 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 el SEQ ID NO: 10, 12, 14 o 16, o un fragmento enzimáticamente activo del mismo. Estas endoglucanasas tienen buena termoestabilidad.

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

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

Beta-glucosidasa	Gen	Obtenible de	ácido nucleico SEQ ID NO:	aminoácido SEQ ID NO:
BG_81 de Ta	<i>ce3A</i> de Ta	<i>T. aurantiacus</i>	21	22
BG_101 de At	<i>ce3A</i> de At	<i>A. thermophilum</i>	23	24
BG_76 de Ct	<i>ce3A</i> de Ct	<i>C. thermophilum</i>	25	26

La xilanasas puede comprender una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 18 o 20, o un fragmento enzimáticamente activo del mismo.

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

Mediante el término "identidad" se quiere significar aquí la identidad global entre dos secuencias de aminoácidos comparadas entre sí desde el primer aminoácido codificado por el gen correspondiente hasta el último aminoácido. La identidad de las secuencias completas se mide utilizando el programa de alineamiento global de Needleman-Wunsch en el paquete de programas EMBOSS (European Molecular Biology Open Software Suite; Rice *et al.*, 2000), versión 3.0.0, con los siguientes parámetros: EMBL0SUM62, 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 en los que se utiliza el algoritmo de Needleman-Wunsch son comparables solamente cuando se alinean los dominios correspondientes de la secuencia. Por consiguiente, la comparación p. ej. de secuencias de celulasa que incluyen CBD o secuencias señal con secuencias que carecen de esos elementos no se puede realizar.

Se puede utilizar un polipéptido celulolítico que tiene una identidad al menos 80, 85, 90, 95 o 99% con el SEQ ID NO: 2, 4, 6, 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 quiere significar cualquier fragmento de una secuencia definida que tiene actividad celulolítica. En otras palabras un fragmento enzimáticamente activo puede ser la parte de la proteína madura de la secuencia definida, o puede ser solamente un fragmento de la parte de la proteína madura, siempre que tenga todavía actividad celobiohidrolasa, endoglucanasa, beta-glucosidasa o xilanasas.

Las enzimas celulolíticas son preferiblemente enzimas recombinantes, que pueden ser producidas de una manera generalmente conocida. 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 en 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 mediante tecnología recombinante en diferentes sistemas anfitriones son bien conocidos en la técnica (Sambrook *et al.*, 1989; Coen, 2001; Gellissen, 2005). Preferiblemente, las enzimas se producen como enzimas extracelulares que son secretadas al medio de cultivo, a partir del cual se puede recuperar y aislarse fácilmente. Se puede utilizar el medio de cultivo gastado del anfitrión de producción tal cual, o se pueden retirar de allí las células anfitrionas, 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 residuos celulares se han retirado del medio de cultivo que contiene el polipéptido. Convenientemente, los polipéptidos se aíslan, p. ej., añadiendo 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. A continuación el medio se filtra, utilizando un agente de filtración inorgánico y un filtro para retirar los precipitantes formados. Después de esto el producto filtrado se procesa adicionalmente utilizando una membrana semi-permeable para retirar el exceso de sales, azúcares y productos metabólicos.

El polinucleótido heterólogo puede comprender 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 filamentoso. Preferiblemente, el anfitrión recombinante se modifica para expresar y secretar enzimas celulolíticas como su actividad principal o una de sus actividades principales. Esto se puede realizar mediante la supresión de 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 o bien simultáneamente, p. ej., en forma de una mezcla de enzimas, o bien sucesivamente, o bien como una parte de la sacarificación y fermentación simultáneas (SSF). Se puede utilizar cualquier combinación de las celobiohidrolasas que comprenden una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 2, 4, 6 u 8 o con un fragmento enzimáticamente activo del mismo junto con cualquier combinación de endoglucanasas y beta-glucosidasas. Si el material celulósico comprende hemicelulosa, se utilizan adicionalmente para la degradación hemicelulasas, preferiblemente xilanasas. Las endoglucanasas, beta-glucosidasas y xilanasas se pueden seleccionar entre las descritas en la presente memoria, pero no se limitan a las mismas. Estas pueden ser también por ejemplo preparaciones de enzimas asequibles comercialmente. Además de celulasas y 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 de enzima" se refiere a 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. Ésta 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 de enzima puede estar en forma de líquido, polvo o granulado. Preferiblemente, la preparación de enzima es medio de cultivo gastado.

5 El "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.

La preparación de enzima puede comprender 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 80% con el SEQ ID NO: 2, 4, 6 u 8 o con un fragmento enzimáticamente activo del mismo, y se puede obtener a partir de

10 *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*, mientras que EG, BG y xilanasas pueden ser de cualquier origen, incluyendo a partir de dichos organismos. Otras enzimas que podrían estar presentes en la preparación son, p. ej. proteasas, amilasas, lacasas, lipasas, pectinasas, esterases y/o peroxidasas.

Se pueden utilizar mezclas y combinaciones de enzimas para adaptarse a diferentes condiciones del proceso. Por ejemplo, si el proceso de degradación se lleva 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 xilanasas de la familia 10 tuvo un excelente rendimiento de hidrólisis tanto a 45°C, como a temperaturas elevadas.

15 Las enzimas celulolíticas de *Trichoderma reesei* se utilizan convencionalmente a temperaturas en el intervalo de aproximadamente 40-50°C en la hidrólisis, y a 30-40°C en SSF. La CBH, EG, BG y Xyn obtenibles de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum* también son eficientes a estas temperaturas, pero además la mayoría de ellas también funciona 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, p. ej., entre 60°C y 65°C. Para tiempos de incubación cortos las mezclas de enzimas son funcionales incluso hasta 85°C, para la hidrólisis completa se utilizan normalmente temperaturas más bajas.

Las enzimas celulolíticas de *Trichoderma reesei* se utilizan convencionalmente a temperaturas en el intervalo de aproximadamente 40-50°C en la hidrólisis, y a 30-40°C en SSF. La CBH, EG, BG y Xyn obtenibles de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum* también son eficientes a estas temperaturas, pero además la mayoría de ellas también funciona 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, p. ej., entre 60°C y 65°C. Para tiempos de incubación cortos las mezclas de enzimas son funcionales incluso hasta 85°C, para la hidrólisis completa se utilizan normalmente temperaturas más bajas.

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

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

25 Se encontraron diversos nuevos polipéptidos celulolíticos en *Thermoascus aurantiacus*, *Acremonium thermophilum*, y *Chaetomium thermophilum*. Los nuevos polipéptidos pueden comprender un fragmento que tiene actividad celulolítica y se pueden seleccionar del grupo que consiste en un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos 66%, preferiblemente 70% o 75%, con el SEQ ID NO: 4, una identidad de 79% con el SEQ ID NO: 6, una identidad de 78% con el SEQ ID NO: 12, una identidad de 68%, preferiblemente 70% o 75%, identidad con el SEQ ID NO: 14, una identidad de 72%, preferiblemente 75%, con el SEQ ID NO: 16, una identidad de 68%, preferiblemente 70% o 75%, con el SEQ ID NO: 20, una identidad de 74% con el SEQ ID NO: 22 o 24, o una identidad de 78% con el SEQ ID NO: 26.

Se encontraron diversos nuevos polipéptidos celulolíticos en *Thermoascus aurantiacus*, *Acremonium thermophilum*, y *Chaetomium thermophilum*. Los nuevos polipéptidos pueden comprender un fragmento que tiene actividad celulolítica y se pueden seleccionar del grupo que consiste en un polipéptido que comprende una secuencia de aminoácidos que tiene una identidad de al menos 66%, preferiblemente 70% o 75%, con el SEQ ID NO: 4, una identidad de 79% con el SEQ ID NO: 6, una identidad de 78% con el SEQ ID NO: 12, una identidad de 68%, preferiblemente 70% o 75%, identidad con el SEQ ID NO: 14, una identidad de 72%, preferiblemente 75%, con el SEQ ID NO: 16, una identidad de 68%, preferiblemente 70% o 75%, con el SEQ ID NO: 20, una identidad de 74% con el SEQ ID NO: 22 o 24, o una identidad de 78% con el SEQ ID NO: 26.

35 Los nuevos polipéptidos también pueden ser variantes de dichos 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 puede haber sido generado mediante mutagénesis. Puede comprender sustituciones, delecciones o inserciones de aminoácidos, pero todavía funciona de una manera sustancialmente similar a las enzimas definidas anteriormente, es decir, comprende un fragmento que tiene actividad celulolítica.

Los nuevos polipéptidos también pueden ser variantes de dichos 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 puede haber sido generado mediante mutagénesis. Puede comprender sustituciones, delecciones o inserciones de aminoácidos, pero todavía funciona de una manera sustancialmente similar a las enzimas definidas anteriormente, es decir, comprende un fragmento que tiene actividad celulolítica.

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

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

50 Los nuevos polipéptidos 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 solamente una parte enzimáticamente activa de la proteína madura. De acuerdo con una realización de la invención, el polipéptido tiene una secuencia de aminoácidos que tiene una identidad de al menos 80, 85, 90, 95, o 99% con el SEQ ID NO: 4, 6, 12, 14, 16, 20, 22, 24 o 26, o con un fragmento celulolíticamente activo del mismo. También puede ser una variante del mismo, o un fragmento del mismo que tiene actividad celobiohidrolasa, endoglucanasa,

Los nuevos polipéptidos 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 solamente una parte enzimáticamente activa de la proteína madura. De acuerdo con una realización de la invención, el polipéptido tiene una secuencia de aminoácidos que tiene una identidad de al menos 80, 85, 90, 95, o 99% con el SEQ ID NO: 4, 6, 12, 14, 16, 20, 22, 24 o 26, o con un fragmento celulolíticamente activo del mismo. También puede ser una variante del mismo, o un fragmento del mismo que tiene actividad celobiohidrolasa, endoglucanasa,

55 Los nuevos polipéptidos 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 solamente una parte enzimáticamente activa de la proteína madura. De acuerdo con una realización de la invención, el polipéptido tiene una secuencia de aminoácidos que tiene una identidad de al menos 80, 85, 90, 95, o 99% con el SEQ ID NO: 4, 6, 12, 14, 16, 20, 22, 24 o 26, o con un fragmento celulolíticamente activo del mismo. También puede ser una variante del mismo, o un fragmento del mismo que tiene actividad celobiohidrolasa, endoglucanasa,

xilanasa, o beta-glucosidasa. De acuerdo con otra realización de la invención, el polipéptido consiste esencialmente en un fragmento celolíticamente activo de una secuencia del SEQ ID NO: 4, 6, 12, 14, 16, 20, 22, 24 o 26.

Los nuevos polinucleótidos pueden comprender una secuencia de nucleótidos del 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 hebras complementarias del mismo. Polinucleótido según se utiliza en la presente memoria se refiere tanto a ARN y ADN, y puede ser de hebra sencilla o de doble hebra. 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. De acuerdo con una realización de la invención tiene al menos 100, 200 o 300 nucleótidos de longitud. Adicionalmente, el polinucleótido puede ser degenerado como resultado del código genético con respecto a una cualquiera de las secuencias definidas anteriormente. Esto significa que diferentes codones pueden codificar para el mismo aminoácido.

El polinucleótido puede estar "comprendido en" el 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. Alternativamente, el polinucleótido comprende 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 se pueden insertar en un vector, que es capaz de expresar el polipéptido codificado por la secuencia heteróloga, y el vector se puede insertar en una célula anfitriona capaz de expresar dicho polipéptido. La célula anfitriona es preferiblemente del género *Trichoderma* o *Aspergillus*.

Se ha introducido un gen heterólogo que codifica los nuevos polipéptidos en un plásmido en una cepa de *Escherichia coli* 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 enzimas pueden ser componentes de una preparación de enzima. La preparación de enzima 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. Puede comprender actividad celobiohidrolasa, endoglucanasa, beta-glucosidasa, y opcionalmente xilanasa y/o otras actividades enzimáticas. Puede comprender adicionalmente cualquier aditivo convencional.

Las nuevas enzimas se pueden aplicar en cualquier proceso que implique enzimas celolíticas, tal como en la industria de combustibles, textil, detergentes, pasta de celulosa y papel, alimentos, forrajes 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 celulosa y el papel se pueden utilizar para modificar la fibra celulósica, por ejemplo, en el tratamiento de pasta kraft, pasta mecánica, o papel reciclado.

Ejemplos

Ejemplo 1. Escrutinio para determinar las cepas que expresan la actividad celolítica 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 celolítica incluyendo beta-glucosidasas. Después del escrutinio preliminar se eligieron seis cepas para estudios posteriores. Estas fueron ALKO4239 y ALKO4242 de *Thermoascus aurantiacus*, ALKO4245 de *Acremonium thermophilum*, ALKO4246 de *Talaromyces thermophilus* y ALKO4261 y ALKO4265 de *Chaetomium thermophilum*.

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 g/litro: celulosa Solka Floc 18, bagazo de destilería 18, xilano de espelta 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 elementos traza 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 g/litro: MnSO₄ 1,6, ZnSO₄·7H₂O 3,45 y CoCl₂·6H₂O 2,0; la disolución de elementos traza Núm. 2 tiene g/litro: FeSO₄·7H₂O 5,0 con dos gotas de H₂SO₄ concentrado.

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

La cepa ALKO 4265 se cultivó en matraces oscilantes en el siguiente medio, g/l: celulosa Solka Floc 40, Pharmamedia™ (Traders Protein, Memphis, TN, USA) 10, polvo de aguas de infusión de maíz 5, (NH₄)₂SO₄ 5 y KH₂PO₄ 15; el pH se ajustó a 6,5. La cepa se cultivó a 45°C durante 7 d.

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

Para la estimación de la termoactividad de las enzimas, se realizaron análisis de las preparaciones de cultivo en matraz oscilante 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 análisis 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 qué pH era el más apropiado para el análisis de termoactividad.

- 5 Se analizaron todos los sobrenadantes de los matraces oscilantes para determinar las siguientes actividades:

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

- 10 Éstas 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 de beta-glucosidasa interferente. El 4-metilumbeliferilo liberado se midió a 370 nm. Las actividades de 'CBHI' y 'EGI' se distinguieron midiendo actividad en presencia y ausencia de celobiosa (5 mM). La actividad que no es inhibida por 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 abajo).

Actividad endoglucanasa (CMCasa):

- 15 Ésta se analizó 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 abajo).

Actividad beta-glucosidasa (BGU):

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

- 25 Para las purificaciones de las proteínas, también se cultivó ALKO4242 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, MgSO₄·7H₂O 0,5, CaCl₂·2H₂O 0,05, disolución de elementos traza Núm. 1 0,5, disolución de elementos traza Núm. 2 0,5. La aireación a fue de 1 vvm, control antiespumante con Struktol, agitación a 200-800 rpm y temperatura a 47°C. Se hicieron circular dos lotes, uno a pH 4,7 ± 0,2 (NH₃/H₂SO₄) y el otro con un 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.

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

- 35 Para la purificación de la enzima, se cultivó ALKO4261 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 espelta de avena 5, CaCO₃ 2, harina de soja 10, salvado de trigo 3,0, (NH₄)₂SO₄ 5, KH₂PO₄ 5, MgSO₄·7H₂O 0,5, NaCl 0,5, KNO₃ 0,3, solución 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), aireación a 1 vvm, agitación a 200-600 rpm, control antiespumante con Struktol y temperatura de 42°C. El tiempo de cultivo fue de 5 d. Se cultivó un segundo lote en condiciones similares excepto que se añadió Solka Floc a 40 g/l y bagazo a 15 g/l. 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 diez veces utilizando el sistema ultrafiltración Pellicon Mini (filtro NMWL de 10 kDa; Millipore, Billerica, MA, USA).

- 45 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), aireación a 0,6 vvm, agitación a 500 rpm, control antiespumante con Struktol y 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 (filtro NMWL de 10 kDa; Millipore, Billerica, MA, USA).

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

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

Los sobrenadantes del cultivo se tamponaron en primer lugar 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 hicieron eluir 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 de gel fue fosfato de sodio 50 mM de pH 7,0, que contenía cloruro de sodio 0,15 M.

Las celobiohidrolasas purificadas se analizaron mediante electroforesis en gel de poliacrilamida-SDS y se determinó que la masa molecular de ambas proteínas era de 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 se denominaron Cel7A de At y Cel7A de Ct, respectivamente, siguiendo el esquema en Henrissat *et al.* (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-celobiósido (MUG2) o 4-metilumbeliferil- β -D-celotriósido (MUG3) como sustrato (van Tilbeurgh *et al.*, 1988) en tampón de citrato sódico 0,05 M de pH 5 a 50°C durante 10 min. Las actividades endoglucanasa y xilanasa se determinaron mediante procedimientos normalizados (de acuerdo con la IUPAC, 1987) utilizando carboximetilcelulosa (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 horas 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 de enzima purificada se midió de acuerdo con Lowry *et al.*, 1951. Para caracterizar los productos finales de la hidrólisis, los azúcares solubles liberados en 24 h en el experimento de hidrólisis, como se ha descrito anteriormente, se analizaron mediante HPLC (Dionex). La celobiohidrolasa purificada I (CBHI/Cel7A) 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 Cel7A de At y Cel7A de Ct purificadas poseen actividades específicas superiores frente a sustratos sintéticos pequeños en comparación con CBHI/Cel7A de *T. reesei*. La actividad específica frente a Avicel fue claramente superior con las enzimas descritas en la presente memoria. Las actividades bajas de las preparaciones de enzima purificada frente a xilano y CMC pueden ser o debidas a las propiedades de las proteínas por sí mismas, o al menos parcialmente a las pequeñas cantidades restantes de enzimas contaminantes. El principal producto final de hidrólisis de la celulosa por todas las enzimas purificadas fue la celobiosa lo que es típico 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	ALKO4245 Cel7A de <i>A. thermophilum</i>	ALKO4265 Cel7A de <i>C. thermophilum</i>	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 min utilizando 4-metilumbeliferil- β -D-lactósido como sustrato. ALKO4265 CBH/Cel7A de *C. thermophilum* y ALKO4245 CBH/Cel7A de *A. thermophilum* fueron estables hasta 65° 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* ALKO4245

Se cultivó *Acremonium thermophilum* ALKO4245 como se ha descrito en el Ejemplo 1. El sobrenadante de cultivo se incubó a 70°C durante 24 horas después de lo cual se concentró mediante ultrafiltración. La endoglucanasa pura se obtuvo mediante purificación secuencial 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 (NH₄)₂SO₄ 1 M. Las proteínas unidas se hicieron eluir con el gradiente lineal desde el tampón anterior a fosfato de potasio 5 mM, pH 6,0. Las fracciones se recogieron y la actividad de endoglucanasa se determinó como se ha descrito anteriormente. La actividad endoglucanasa eluyó en una amplia zona 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 hicieron eluir con un gradiente lineal de NaCl de 0 a 0,25 M en el tampón de equilibrado. La actividad endoglucanasa que contenía proteína se hizo eluir en la zona de conductividad de 3-7 mS/cm. Se repitió la cromatografía de intercambio catiónico y el eluato de la proteína se concentró mediante liofilización.

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 se hizo eluir 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, designada como EG₄₀ de At, a 50°C era 450 de nkat/mg (Procedimiento de la IUPAC 1987, utilizando como sustrato CMC).

La estabilidad térmica de la endoglucanasa purificada se determinó a diferentes temperaturas. La reacción se realizó en presencia de 0,1 mg/ml de BSA a pH 5,0 durante 60 min utilizando carboximetilcelulosa como sustrato. *A. thermophilum* EG₄₀/Cel45A fue estable hasta 80°C. Las enzimas de referencia de *T. reesei* EGI (Cel7B) y EGII (Cel5A) 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 ha descrito en el Ejemplo 1. La endoglucanasa pura se obtuvo mediante purificación secuencial 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).

Se añadió sulfato amónico al sobrenadante de cultivo para alcanzar la misma conductividad que fosfato de potasio 20 mM de pH 6,0, que contenía (NH₄)₂SO₄ 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 (NH₄)₂SO₄ 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 hicieron eluir 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 contiene actividad endoglucanasa se hizo eluir al principio del gradiente de urea.

Las fracciones se combinaron, se equilibraron a Tris-HCl 16 mM, pH 7,5 (I = 1,4 mS/cm) mediante una columna 10DG (Bio-Rad) y se aplicó a una columna de intercambio aniónico HiTrap DEAE FF equilibrada con Tris-HCl 20 mM, pH 7,5. Las proteínas unidas se hicieron eluir 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 hizo eluir en el intervalo de 10-20 mS/cm.

La muestra se equilibró a 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 hicieron eluir con un gradiente lineal de acetato de sodio de 0 a 0,4 M, pH 4,5. La actividad endoglucanasa 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, Broad Range, Bio-Rad) era de 54 kDa. Se determinó que el pI de la proteína purificada, denominada EG₅₄ de Ct con PhastSystem (Pharmacia) era aprox. 5,5.

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

Se cultivó *Thermoascus aurantiacus* ALKO4242 como se ha descrito en el Ejemplo 1. La endoglucanasa pura se obtuvo mediante purificación secuencial 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 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 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 hicieron eluir 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 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 hicieron eluir con un gradiente lineal de NaCl 0 a 0,4 M en el tampón de equilibrado. La proteína que contiene actividad endoglucanasa se hizo eluir 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 hizo eluir 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 poli(acrilamida)-SDS y se consideró que el peso molecular era de 28 kDa. Se determinó que el pI de la proteína purificada, denominada EG_28 de Ta, en un gel IEF (PhastSystem, Pharmacia) era de aproximadamente 3,5. Se determinó que la actividad específica de EG_28 de Ta a 50°C era de 4290 nkat/mg (Procedimiento de la IUPAC 1987, utilizando como sustrato CMC).

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

Se cultivó *Acremonium thermophilum* ALKO4245 como se ha descrito en el Ejemplo 1. La β -glucosidasa pura se obtuvo mediante purificación secuencial con cromatografía de interacción hidrófoba y de 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 Phenyl Sepharose FF de interacción hidrófoba HiPrep 16/10 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 hicieron eluir con un gradiente lineal desde el tampón de equilibrado a fosfato de potasio 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 hicieron eluir con un gradiente lineal desde el 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 tampón de fosfato de potasio 4 mM de pH 7,2 mM. Las proteínas se hicieron eluir a la zona conductividad de 6-9 mS/cm. Las fracciones que contenían actividad de β -glucosidasa se recogieron, se combinaron y se concentraron.

El material activo de la cromatografía de intercambio aniónico se aplicó a una columna HR 26/60 Sephacryl S-300 equilibrada con fosfato de sodio 20 mM de pH 6,5, que contenía NaCl 0,15 M. La proteína con actividad β -glucosidasa se hizo eluir 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 poli(acrilamida)-SDS y se evaluó que el peso molecular era de 101 kDa. Se determinó que el pI de la proteína purificada, denominada A β G_101, en un gel IEF (PhastSystem, Pharmacia) estaba en la zona de 5,6-4,9. Se determinó que la actividad específica de β G_101 de At 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 0,1 mg/ml de BSA a pH 5,0 durante 60 min utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato. La β G_101 de *A. thermophilum* fue estable hasta 70°C. La enzima de referencia de *Aspergillus* (Novozym 188) conservó 100% de la actividad hasta 60°C.

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

Se cultivó *Chaetomium thermophilum* ALKO4261 como se ha descrito en el Ejemplo 1. La β -glucosidasa pura se obtuvo mediante purificación secuencial con cromatografía de interacción hidrófoba, de intercambio aniónico y catiónico 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 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 desde tampón de equilibrado a fosfato de potasio 3 mM, pH 6,0, seguido de elución con agua y urea 6 M. Las primeras fracciones con actividad β -glucosidasa se hicieron eluir en la zona de conductividad de 80-30 mS/cm. La segunda actividad β -glucosidasa se hizo eluir con urea 6 M. Las fracciones activas eluidas mediante urea se reunieron y se desalaron en columnas 10DG (Bio-Rad) equilibradas con Tris-HCl 10 pH 7,0.

- 5 Después de la eliminación de las sales, 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, sino que eluyó durante la introducción de la muestra. Esta fracción de flujo directo se desaló en columnas 10DG (Bio-Rad) equilibradas con acetato de Na 7 mM, pH 4,5.

- 10 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 hicieron eluir 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 que eluyeron 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.

- 15 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 hizo eluir 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 consideró que el peso molecular era de 76 kDa. La proteína se denominó β G₇₆ de Ct.

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

- 20 Se cultivó *Thermoascus aurantiacus* ALKO4242 como se ha descrito en el Ejemplo 1. La β -glucosidasa pura se obtuvo mediante purificación secuencial con cromatografía de interacción hidrófoba, de intercambio aniónico y catiónico 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 por BioRad Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

- 25 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 hicieron eluir 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 hizo eluir durante el gradiente en la zona conductividad de 28,0 a 1,1 mS/cm. Las fracciones se combinaron y se concentraron mediante ultrafiltración.

- 30 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 hizo eluir 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 una columna 10DG (Bio-Rad).

- 35 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 hizo eluir con un gradiente lineal de NaAc 10 mM a 400 mM y mediante elución retardada utilizando NaAc 400 mM de pH 4,5. Las proteínas con actividad β -glucosidasa se hicieron eluir ampliamente durante el gradiente lineal en la zona de conductividad de 5,0 a 11,3 mS/cm.

- 40 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 hizo eluir 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 poliacrilamida-SDS y se consideró que el peso molecular era de 81 kDa, que representaba muy probablemente la forma monomérica de la proteína. El enfoque isoeléctrico (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 de 4,55. Se determinó que la actividad específica de la proteína purificada, denominada β G₈₁ de Ta, a 50°C era de 600 nkat/mg utilizando 4-nitrofenil- β -D-glucopiranosido como sustrato (Bailey y Linko, 1990).

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

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

- 55 Se cultivó *Acremonium thermophilum* ALKO4245 como se ha descrito 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 secuencial con cromatografía de interacción hidrófoba y de intercambio catiónico

seguida de filtración en gel. La actividad xilanasa se determinó utilizando como sustrato xilano de abedul (Procedimiento de la IUPAC 1987). La proteína se analizó mediante BioRad Protein Assay Kit (Bio-Rad Laboratories) utilizando albúmina de suero bovino como patrón.

5 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 hicieron eluir con el gradiente lineal desde el tampón anterior a fosfato de potasio 5 mM, pH 6,0. La fracción de proteína se hizo eluir en una amplia zona de conductividad de 120 a 15 mS/cm.

10 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, sino que eluyó en el flujo directo durante la introducción de la muestra. Este eluato se concentró mediante ultrafiltración. La cromatografía hidrófoba se repitió como se ha descrito anteriormente. Las proteínas no unidas se recogieron y se liofilizaron.

15 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 poliacrilamida-SDS. Se consideró que la masa molecular de la proteína purificada sobre la base de los patrones de masa molecular (patrones SDS-PAGE teñidos previamente, Intervalo Amplio, Bio-Rad) era de 60 kDa. Se determinó que la actividad específica de la proteína, denominada XYN_60 de At, 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 20 70°C (10 min, pH 5,0) en comparación con 50°C. Las actividades específicas contra MUG2 (4-metilumbeliferil- β -D-celobiosido), MUL (4-metilumbeliferil-beta-D-lactósido) y MUG3 (4-metilumbeliferil- β -D-celotriósido) fueron de 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).

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

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

30 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 $(\text{NH}_4)_2\text{SO}_4$ 0,7 M. Las proteínas unidas se hicieron eluir con un protocolo de elución de dos etapas. La elución se llevó a cabo dejando caer primero la concentración de sal a $(\text{NH}_4)_2\text{SO}_4$ 0,2 M y después de eso se aplicó un gradiente lineal de fosfato de potasio 20 mM de pH 6,0, que contenía $(\text{NH}_4)_2\text{SO}_4$ 0,2 M a fosfato de potasio 5 mM de pH 6,0. La proteína se hizo eluir con $(\text{NH}_4)_2\text{SO}_4$ 0,2 M ($I = 39$ mS/cm).

35 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 sino que eluyó en el flujo directo. La conductividad de la muestra se ajustó para que correspondiera a la de acetato de sodio 20 mM, pH 4,5 mediante adición de agua y el pH se ajustó a 4,5 durante la concentración mediante ultrafiltración.

40 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 hicieron eluir con un gradiente lineal desde el tampón de equilibrado al mismo tampón que contenía NaCl 1 M. La enzima se hizo eluir en la zona de conductividad de 1-7 mS/cm. La muestra se liofilizó y después de eso se disolvió en agua.

45 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 hizo eluir 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 teñidos previamente, Intervalo Amplio, Bio-Rad). Se determinó que el pI de la proteína purificada, denominada XYN_30 de Ta con PhastSystem (Pharmacia) era de aprox. 6,8. Se determinó que la actividad específica de XYN_30 de Ta a 50°C era de 4800 50 nkat/mg (Procedimiento de la IUPAC 1987, utilizando como sustrato xilano de abedul).

Ejemplo 11. Secuenciación de aminoácidos internos

55 Los péptidos internos se secuenciaron mediante ionización por electropulverización combinada con espectrometría de masas en tándem (ESI-MS/MS) utilizando un aparato Q-TOF1 (Micromass). La proteína se alquiló primero y se digirió en péptidos tripticos. Los péptidos generados se desalaron y se separaron parcialmente mediante nanocromatografía líquida (de fase inversa) antes de aplicarlos al aparato Q-TOF1. Las secuencias de los 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 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 de 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 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

5 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 la lisina difiere solo en 0,036 Da y no se pueden distinguir en el análisis ESI-MS/MS.

10 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.

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

Proteína	Péptido	Secuencia ^(a)
EG_40 de At	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
EG_54 de Ct	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
	Péptido 7	A L N D D F C R
EG_28 de Ta	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
	Péptido 3	P F M M E R
	Péptido 4	Q Y A V V D P H N Y G R

15 (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 la lisina difiere solo en 0,036 Da y no se pueden distinguir en el análisis ESI-MS/MS.

Tabla 4. Secuencias de péptidos internos determinados a partir de las beta-glucosidasas β G_101 de *Acremonium thermophilum* ALK04245, β G_76 de *Chaetomium thermophilum* ALK04261 y β G_81 de *Thermoascus aurantiacus* ALK04242.

Proteína	Péptido	Secuencia ^(a)
β G_101 de At	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
β G_76 de Ct	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
β G_81 de Ta	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)

- 5 Tabla 5. Secuencias de péptidos internos determinados a partir de las xilanasas XYN_60 de *Acremonium thermophilum* ALK04245 y XYN_30 de *Thermoascus aurantiacus* ALK04242.

Proteína	Péptido	Secuencia
XYN_60 de At	Péptido 1	Y N D Y N L E Y N Q K
	Péptido 2	F G Q V T P E N
	Péptido 3	V D G D A T Y M S Y V N N K
	Péptido 4	K P A W T S V S S V L A A K
	Péptido 5	S Q G D I V P R A K
XYN_30 de Ta	Péptido 1	V Y F G V A T D Q N R
	Péptido 2	N A A I I Q A D F G Q V T P E N S M K
	Péptido 3	G H T L V W H S Q L P S W V S S I T D K
	Péptido 4	N H I T T L M T R
	Péptido 5	A W D V V N E A F N E D G S L R
	Péptido 6	L Y I N D Y N L D S A S Y P K
	Péptido 7	A S T T P L L F D G N F N P K P A Y N A I V Q D L Q Q
	Péptido 8	Q T V F L N V I G E D Y I P I A F Q T A R

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

Se prepararon las bibliotecas genómicas de *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 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 *Sau* 3A. Los ADN digeridos también se fraccionaron por tamaños y los fragmentos del tamaño seleccionado ($\approx$ 5-23 kb) se desfosforilaron y se ligaron a los brazos del vector lambda digerido con *Bam*HI. 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 de $3,6 \times 10^6$ pfu/ml y $3,7 \times 10^5$ pfu/ml y los de las bibliotecas amplificadas fueron de $6,5 \times 10^{10}$ pfu/ml y de $4,2 \times 10^8$ pfu/ml, respectivamente.

Se utilizó Lambda FIX[®] II/Xho I 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 mediante el método de Raeder y Broda (1985), se digirieron parcialmente con *Sau*3A. Los ADN digeridos se fraccionaron por tamaños y los fragmentos del tamaño seleccionado ($\approx$ 6-23 kb) se rellenaron y se ligaron a los brazos del vector Lambda FIX II digerido 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 de $0,2 \times 10^6$ y $0,3 \times 10^6$ pfu/ml y los de las bibliotecas amplificadas fueron de $1,8 \times 10^9$ y $3,8 \times 10^9$ pfu/ml, respectivamente.

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

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

Las sondas para el escrutinio de las bibliotecas genómicas que se habían construido como se ha descrito 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. Varios cebadores sometidos a ensayo en reacciones de PCR se diseñaron 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, cada cebador 5 mM y 1 unidad de ADN polimerasa Dynazyme EXT (Finzymes, Finlandia) y $\approx$ 0,5-1 μ g del 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, o bien 1 min de recocido a 62°C (gradiente de $\pm$ 8°C) para los moldes de *Thermoascus* ALKO4242 y *Chaetomium* ALKO4265 o 1 min de recocido a 58°C (gradiente de $\pm$ 6°C) para el molde de *Acremonium* ALKO4245, 2 min de extensión a 72°C y una extensión final a 72°C durante 10 min.

Los productos de ADN de los tamaños esperados (calculados a partir de las 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 en el vector pCR[®] Blunt-TOPO[®] (Invitrogen, USA). Los insertos se caracterizaron mediante secuenciación y llevando a cabo hibridaciones de transferencia Southern para los ADN genómicos digeridos con varias enzimas de restricción. Los fragmentos de PCR, que se eligieron para ser utilizados como sondas para el escrutinio 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 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 ADN molde genómico y el nombre del plásmido que contiene el fragmento sonda.

Gen	Cebador directo	Cebador inverso	DNA molde	Fragmento (kb)	Plásmido
<i>cbh</i> de Ta	TCEL11 atcggaactggcgttggtgctc	TCEL12 gaattggagctagtgctgacg	<i>Thermoascus</i> ALKO4242	0,8 kb	pALK1633
<i>cbh</i> de Ct	TCEL7 cgatgccaactggcgctggac	TCEL8 ttcttggtggtgctgacggtc	<i>Chaetomium</i> ALKO4265	0,8 kb	pALK1632
<i>cbh</i> de At	TCEL13 agctcgaccaactgctacacg	TCEL4 accgtgaactcttgctggtg	<i>Acremonium</i> ALKO4245	0,7 kb	pALK1634

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

Los insertos de los plásmidos enumerados en la Tabla 6 se marcaron 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 sonda marcados. 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 a partir 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 otro u otros genes de celobiohidrolasa podrían estar presentes en el genoma, causando una reacción cruzada. Se purificaron de cuatro a cinco placas que hibridaron fuertemente a partir 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 fagos se aislaron y se caracterizaron mediante hibridaciones de transferencia Southern. Los fragmentos de restricción seleccionados que hibridaron con la sonda se subclonaron con 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, estos tenían los códigos At_cbh_C At_cbh_A y, a continuación, se denominaron At cel7 A y cel7B de At, respectivamente). La Tabla 7 resume la información sobre las sondas utilizadas para los escrutinios de los genes, los clones de fagos a partir de los cuales se aislaron los genes, los fragmentos de restricción seleccionados que contenían los genes completos con sus regiones promotoras y terminadoras, los nombres de los plásmidos, y los números de depósito DSM para las cepas de *E. coli* que portaban estos plásmidos.

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

Gen	Sonda utilizada en el escrutinio	Clon de fago	Fragmento subclonado en pBluescript II	Número de plásmido	Núm. de depósito de <i>E. coli</i>
Cel7A de Ta	pALK1633	F12	XbaI de 3,2 kb	pALK1635	DSM 16723
Cel7A de Ct	pALK1632	F36	PvuI - HindIII de 2,3 kb	pALK1642	DSM 16727
cel7B de At	pALK1634	F6	EcoRI de 3,1 kb	pALK1646	DSM 16728
Cel7A de At	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 de péptidos 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 Cel7A de Ct y Cel7A de At. Por lo tanto, se pudo concluir que los genes que codificaban las proteínas CBH/Cel7 purificadas de *Chaetomium thermophilum* y *Acremonium thermophilum* se habían clonado.

Tabla 8. Resumen 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	Longitudes de los intrones (pb)	SEQ ID NO:
Cel7A de Ta	1439	1371	1	65	1
Cel7A de Ct	1663	1596	1	64	7
cel7B de At	1722	1377	3	134, 122, 87	3

Gen Cbh	Longitud con intrones (pb) ^(a)	Región codificante (pb) ^(b)	Núm. de intrones	Longitudes de los intrones (pb)	SEQ ID NO:
Cel7A de At	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* ALK4265 y *Acremonium thermophilum* ALK04245. ss, secuencia señal.

Proteína CBH	Núm. de aa	Longitud de ss NN/HMM ^(a)	CBD C-terminal ^(b)	PM pronosticado (Da, no incluida ss) ^(c)	pl pronosticado (no incluida ss)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
Cel7A de Ta	457	17/17	NO	46873	4,44	2	2
Cel7A de Ct	532	18/18	SI, T497 a L532	54564	5,05	3	8
Cel7B de At	459	21/21	NO	47073	4,83	2	4
Cel7A de At	523	17/17	SI, Q488 a L523	53696	4,67	4	6

5 (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) Se indican el dominio de unión a celulosa (CBD), los aminoácidos de la región CBD C-terminal (M1 (se incluye en la numeración la Met Núm. 1))

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

(d) Número de secuencias NXS/T.

15 Las secuencias de aminoácidos deducidas de Cel7A de *Thermoascus aurantiacus* y Cel7A de *Acremonium thermophilum* (núcleo, sin el CBD) eran muy homólogas entre sí (analizadas mediante el alineamiento global de Needleman-Wunsch, EMBOSS 3.0.0 Needle, con Matrix 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 menor con la Cel7A de *Chaetomium thermophilum* deducida. La Cel7B de *Acremonium thermophilum* fue muy distinta de las secuencias de CBH/Cel7 descritas aquí.

20 La secuencia de Cel7A de *Chaetomium* deducida poseía las identidades más altas (analizadas mediante alineamiento global de Needleman-Wunsch, EMBOSS Needle, véase más arriba) con los 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 relacionada con el polipéptido de CBHI de *Oryza sativa*. Los mayores homólogos de la

25 secuencia de Cel7A de *Acremonium thermophilum* deducida fueron con 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* ALKO4245 codifican las proteínas CBH que tienen alta homología con los polipéptidos de la familia 7 de glicosido hidrolasas, por lo tanto, éstas se denominaron Cel7A o Cel7B (Henrissat *et al.* 1998).

30 La comparación de las secuencias de aminoácidos deducidas de los genes *CBH/cel 7* de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245 y *Thielavia* entre sí, y adicionalmente con las secuencias encontradas en las bases de datos, se muestran en la Tabla 10.

35 Tabla 10. Secuencias con una homología más alta con las secuencias de aminoácidos deducidas de los genes *CBH/cel 7* de *Thermoascus aurantiacus* ALKO4242, *Chaetomium thermophilum* ALKO4265 y *Acremonium thermophilum* ALKO4245. El alineamiento se realizó utilizando el alineamiento global Needleman-Wunsch (EMBLOSUM62, Penalización por hueco 10,0, Penalización por extensión 0,5). * indica una secuencia de

aminoácidos derivada de uno de los genes celobiohidrolasa clonados en este trabajo. 'Central' indica el alineamiento sin el CBD.

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

Organismo, enzima y número de acceso	Identidad, (%)
<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* (Cel7A de Ta), *Chaetomium thermophilum* (Cel7A de Ct) y *Acremonium thermophilum* (Cel7A de At, Cel7B de At; en la fase inicial del trabajo estas proteínas tenían los códigos temporales CBH_C de At y CBH_A de At, respectivamente). Los plásmidos de expresión contruidos se enumeran en la Tabla 11. Los genes *cbh1cel7* recombinantes, incluyendo sus propias secuencias señal, se fusionaron exactamente con el promotor de *cbh1* (cel7A) de *T. reesei* mediante PCR. La terminación de la transcripción se garantizó por medio del terminador de 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). 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 A96 y A98 de *T. reesei* (ambas cepas tienen suprimidos los genes que codifican las cuatro celulasas principales CBHI/Cel7A, CBHII/Cel6A, EGI/Cel7B y EGII/Cel5A). Las transformaciones se realizaron como en Penttilä *et al.* (1987) con las modificaciones descritas en Karhunen *et al.* (1993), seleccionando con acetamida como única fuente de nitrógeno. Los transformantes se purificaron sobre placas de selección a través de conidios individuales antes de su esporulación sobre PD.

Tabla 11. Casetes de expresión contruidos para producir las proteínas CBH/Cel7 de *Thermoascus aurantiacus* ALK04242 (Cel7A de Ta), *Chaetomium thermophilum* ALKO4265 (Cel7A de Ct), y *Acremonium thermophilum* ALKO4245 (Cel7A de At, Cel7B de At) en *Trichoderma reesei*. La estructura general de los casetes de expresión fue la descrita en la Fig. 2. Los genes *cbh1cel7* clonados se fusionaron exactamente con el promotor *cbh1/cel7A* de *T. reesei*.

CBH/Cel7	Plásmido de expresión	Tamaño del casete de expresión ^(a)	terminador de cel7A ^(b)
Cel7A de Ta	pALK1851	9,0 kb	245 pb (<i>XbaI</i>)
Cel7A de Ct	pALK1857	9,2 kb	240 pb (<i>HindIII</i>)
Cel7B de At	pALK1860	9,4 kb	361 pb (<i>EcoRI</i>)
Cel7A de At	pALK1865	9,5 kb	427 pb (<i>EcoRV</i>)

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

(b) Número de nucleótidos de la región terminadora de *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.

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 a base de 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 utilizando transferencias Southern en las que se incluyeron varios productos digeridos genómicos y el casete de expresión respectivo se utilizó como sonda. Se analizó la expresión heteróloga de las proteínas Cel7A de Ta, Cel7A de Ct, Cel7A de At y Cel7B de At mediante SDS-PAGE con la posterior tinción de Coomassie. Las conclusiones de que no se pudiera detectar ninguna actividad celobiohidrolasa o la producción de proteína heteróloga en SDS-PAGE para los transformantes de Cel7B de At que contenía el casete de expresión integrado, sugieren que Cel7B de At es producida por debajo de los niveles de detección en *Trichoderma* utilizando el diseño experimental descrito.

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

Los transformantes de CBH/Cel7 seleccionados se cultivaron biorreactores de laboratorio a 28°C en el medio indicado anteriormente durante 3-4 días con un pH de control de $4,4 \pm 0,2$ ($\text{NH}_3/\text{H}_3\text{PO}_4$) 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).

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

Se fusionó Cel7A de *Thermoascus aurantiacus* (AF478686, Hong *et al.*, 2003b; SEQ ID NO: 1) con el conector y CBD de CBH/Cel7A de *Trichoderma reesei* (AR088330, Srisodsuk *et al.* 1993) (= CBD de Tr), seguido de la producción de la proteína de fusión (SEQ ID NO: 28 que corresponde al ácido nucleico del SEQ ID NO: 27) en *T. reesei* como se describe en FI20055205/US 11/119526; presentado el 29 de Abril 2005. Además, Cel7A de *Thermoascus aurantiacus* se fusionó con el conector y el CBD de Cel7A de *Chaetomium thermophilum* (SEQ ID NO: 7) (CBD de Ct). Para ello, se sintetizaron la secuencia codificante del conector y el CBD de Cel7A de *Chaetomium thermophilum* mediante PCR utilizando los siguientes cebadores:

5'-TTAAACATATGTTATCTACTCCAACATCAAGGTCGGACCCATCGGCTC-GACCGTCCCTGGCCTTGAC-3'
(secuencia directa)

y

5'-TATATGCGGCCGCAAGCTTTACCATCAAGTTACTCCAGCAAATCAGGG-AACTG-3' (secuencia inversa).

La mezcla de reacción de PCR contenía 1 x tampón de reacción de DyNAzyme™ EXT (Finnzymes, Finlandia), Mg^{2+} 15 mM, dNTP 0,2 mM, 2 μM de cada cebador, 0,6 unidades de ADN polimerasa DyNAzyme™ EXT (Finnzymes, Finlandia), y aproximadamente 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 inicial a 98°C, seguido de 30 ciclos de 30 seg. a 98°C, 30 seg. de recocido a 68°C (gradiente de $\pm 4^\circ\text{C}$), 30 segundos de extensión a 72°C y una extensión final a 72°C durante 10 min. El fragmento de ADN específico en la reacción de PCR se obtuvo a un intervalo de temperatura de recocido entre 64°C a 68,5°C. El fragmento CBD sintetizado de *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 del plásmido se confirmó mediante secuenciación (SEQ ID NO: 29). El gen fusión *cel7A* construido se fusionó exactamente al promotor de *cbh1* (*cel7A*) de *T. reesei*. La terminación de la transcripción se aseguró por medio del terminador de *cel7A* de *T. reesei* y se utilizó el gen marcador *amdS* de *A. nidulans* para la selección de los transformantes como se describe en 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 A96 de *T. reesei*. Las transformaciones se realizaron como en Penttilä *et al.* (1987) con las modificaciones descritas en Karhunen *et al.* (1993), con la selección de acetamida como única fuente de nitrógeno. Los transformantes se purificaron en placas de selección a través de conidios individuales antes de la su esporulación sobre PD.

Se analizó la producción Cel7A + CBD (SEQ ID NO. : 28 y 30) de *Thermoascus aurantiacus* de los transformantes 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 KH_2PO_4 a pH 5,5. La actividad de 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 utilizando transferencias Southern en las que se habían incluido 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* recombinante se habían producido en forma de proteínas de fusión estables en *T. reesei*.

El transformante seleccionado que producía la proteína de fusión Cel7A de Ta + CBD de Tr (SEQ ID NO: 28) también se cultivó en un biorreactor 2 litros a 28°C en el medio indicado anteriormente durante 3-4 días con control de pH a $4,4 \pm 0,2$ ($\text{NH}_3/\text{H}_3\text{PO}_4$) 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).

50 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 la celobiosa se determinaron a partir de las celobiohidrolasas producidas de manera heteróloga en *T. reesei* (Ejemplos 14 y 15). Las enzimas se purificaron como se describe en el Ejemplo 2. Las concentraciones de proteína de las enzimas purificadas se midieron mediante su absorción a 280 nm utilizando un coeficiente de extinción molar teórico, que se evaluó a partir de las secuencias de aminoácidos (Gill y von Hippel, 1989).

Las constantes cinéticas (valores K_m y k_{cat}) y la constante de inhibición de celobiosa (K_i) para CBHI/Cel7A de Tr, CBH/Cel7A de Ta, CBH/Cel7A de At y CBH/Cel7A de Ct, se midieron utilizando CNPLac (2-Cloro-4-nitrofenil- β -D-lactósido) como sustrato a la temperatura ambiente (22°C) en tampón de fosfato de sodio 50 mM, pH 5,7. Para la determinación de la constante de inhibición (K_i), se utilizaron ocho concentraciones de sustrato diferentes (31 a 4000 μ M) en presencia de un intervalo de cinco concentraciones de inhibidor (0-100 μ M o 0-400 μ M), que agrupan el valor de K_i . Todos los experimentos se realizaron en placas de microtitulación y el volumen de reacción total fue de 200 μ l. Las tasas iniciales se midieron en cada caso mediante control continuo de la liberación del anión de 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 de la curva patrón de CNP (de 0 a 100 μ M). Las concentraciones de enzima utilizadas fueron: CBHI/Cel7A de Tr 2,46 μ M, CBH/Cel7A de Ta 1,58 μ M, CBH/Cel7A de Ct 0,79 μ M y CBH/Cel7A de At 3 μ M. Las constantes K_m y k_{cat} se calcularon a partir del ajuste de la ecuación de Michaelis-Menten utilizando el programa de Origen. Se utilizaron diagramas de Lineweaver-Burk, regráficos "replots" (pendiente de LWB frente a [glc2; celobiosa]) y diagramas de Hanes para distinguir entre la inhibición de tipo competitivo y mixta y para determinar las constantes de inhibición (K_i).

Los resultados de las mediciones cinéticas se muestran en la Tabla 12 y la Tabla 13. Como se puede observar, CBH/Cel7A de Ct tiene claramente el número de recambio más alto (k_{cat}) sobre CNPLac y también la constante de especificidad (k_{cat}/K_m) es mayor en comparación con CBHI/Cel7A de *T. reesei*. La celobiosa (glc2) es un inhibidor competitivo para todas las celulasas medidas, y la CBHI/Cel7A de Tr (utilizada como control) tiene la inhibición más fuerte (es decir, el valor de K_i más bajo) por celobiosa. La CBH/Cel7A de At tuvo una constante de inhibición 7 veces mayor en comparación con el de CBHI/Cel7A de Tr. Estos resultados indican que las tres nuevas celobiohidrolasas podrían funcionar mejor en la hidrólisis de celulosa debido a la disminución de la inhibición por celobiosa en comparación con la celobiohidrolasa I de Cel7A de *Trichoderma reesei*.

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

Enzima	K_i (μ M)	Tipo de inhibición
Cel7A de Ct	39	competitiva
Cel7A de Ta	107	competitiva
Cel7A de At	141	competitiva
Cel7A de Tr	19	competitiva

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

Enzima	k_{cat} (min^{-1})	K_m (μ M)	k_{cat}/K_m ($\text{min}^{-1} \text{M}^{-1}$)
Cel7A de Ct	18,8	1960	9,5 103
Cel7A de Tr	2,6	520	5,0 103

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

Las celobiohidrolasas recombinantes purificadas Cel7A de Ct, Cel7A de Ta, Cel7A de Ta+CBD de Tr, Cel7A de Ta+CBD de Ct, Cel7A de At así como la versión núcleo de Cel7A de Ct (véase más adelante) se sometieron a ensayo en cantidades equimolares en la hidrólisis de celulosa cristalina a dos temperaturas, 45°C y 70°C; la Cel7A de Tr de *T. reesei* purificada y su versión núcleo (véase más adelante) se utilizaron como comparación. Los análisis de hidrólisis de la celulosa cristalina (Ph 101, Avicel; Fluka, Bucsh, Suiza) se llevaron a cabo en acetato de sodio 50 mM en un tubo con escala de 1,5 ml, pH 5,0. Se aplicó movimiento oscilante a Avicel a 45°C o a 70°C, con la disolución de enzima (1,4 μ M), y el volumen final de la mezcla de reacción fue de 325 μ l. La hidrólisis siguió hasta 24 horas tomando muestras en seis momentos diferentes y deteniendo la reacción mediante la adición de 163 μ l 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 Millex GV13 de 0,22 μ m (Millipore, Billerica, MA, USA). La formación de azúcares reductores solubles en el sobrenadante se determinó mediante el método de la hidrazida de ácido para-hidroxibenzoico (PAHBAH) (Lever, 1972) utilizando una curva patrón de celobiosa (celobiosa de 50 a 1600 μ M). Se añadió una disolución de PAHBAH 0,1 M (Sigma-Aldrich, St. Louis, MO, USA) de nueva aportación en NaOH 0,5 M (100 μ l) a 150 μ l de la muestra filtrada y se hirvió durante 10 minutos después de lo cual la disolución se enfrió sobre hielo. Se midió la absorbancia de las muestras a 405 nm.

Las versiones núcleo de las celobiohidrolasas que albergan un CBD en su forma nativa se obtuvieron como sigue: Cel7A de Ct y Cel7A de Tr fueron expuestas 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 cantidad de una quinta parte o una décima parte de papaína con respecto a la cantidad total de la Cel7A en la mezcla de reacción). La proteína núcleo resultante se purificó con una columna de intercambio aniónico DEAE Sepharose FF (Pharmacia, Uppsala, Suecia) como se ha descrito anteriormente. El producto se analizó en SDS-PAGE.

Los resultados de 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 termoestable 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 de Cel7A de *Thermoascus* está conectado al CBD de Cel7A de *T. reesei* (Cel7A de Ta + CBD de Tr). Resultó sorprendente que las celobiohidrolasas aisladas y clonadas en este trabajo sean superiores, cuando albergan un CBD, en la velocidad y la formación del producto en la hidrólisis de celulosa cristalina también a la temperatura de hidrólisis convencional de 45°C cuando se comparó con la celobiohidrolasa del estado de la técnica de Cel7A (CBHI) de *T. reesei* a la misma concentración de la enzima. Los resultados también están de acuerdo con las preparaciones de enzima (Cel7A de At y Cel7A de Ct), que se purificaron a partir de los anfitriones originales y se sometieron 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 convencionales de biología molecular como se ha descrito 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 purificadas de *Acremonium* y *Chaetomium* comparten homología con varias endoglucanasas de la familia 45 de glicosilhidrolasas tales como la endoglucanasa Cel45A de *Melanocarpus albomyces* (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 la endoglucanasa EG1 de *Thermoascus aurantiacus* (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 los péptidos. El orden de los péptidos en la secuencia de la proteína y la naturaleza efectora o antisentido correspondiente de los cebadores se dedujeron a partir de la comparación con las endoglucanasas publicadas homólogas. Las secuencias de los cebadores y los péptidos correspondientes se enumeran en la Tabla 14. Debido a una identidad de casi 100% de los péptidos de *Thermoascus* con la secuencia publicada, el gen de la 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* a partir de las bibliotecas genómicas correspondientes.

Proteína	Péptido	Localización del cebador ^(a)	Secuencia del cebador ^(b)
EG_40 de At	Péptido 5	1—6	TAYTGGGAYTGYTGAAARCC
	WFQNADN ^(c)		RTTRTCNGCRTTYTGAAACCA
EG_54 de Ct	Péptido 7	3—7	GCAAGCTTCGRCAARTCRTRCTT ^(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 EG_40 purificada de *Acremonium*, pero se origina a partir de la secuencia Cel45A de *M. albomyces* (AJ515703) homóloga a EG_40.

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

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

Se amplificó la sonda específica del gen *cel45A* de *Acremonium thermophilum* para escrutar la biblioteca genómica con los cebadores directo (TAYTGGGAYTGYTGAAARCC) e inverso (RTTRTCNGCRTTYTGAAACCA) 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, 1 min de recocido a 50-60°C, 2 min de extensión a 72°C y una final extensión a 72°C durante 10 min. Para la amplificación de la sonda específica de gen *cel7B* de *Chaetomium thermophilum* (que codifica EG_54 de Ct), se utilizaron un cebador directo (GGAATTCGAYCARACNGARCARTA) y un cebador inverso (GCAAGCTTCGRCARAARTCRCTT). 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 unidades 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 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 por secuenciación y realizando hibridaciones mediante 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 los dos supuestos genes de endoglucanasa podrían ser escrutados a partir de la biblioteca genómica de *Acremonium*. Las secuencias de aminoácidos deducidas de los dos productos de PCR tienen homología con varias secuencias publicadas de endoglucanasa de la familia 45 de glicosilhidrolasa (programa BLAST del Nacional 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 homóloga de *Humicola insolens*, A35275) 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 7 glicosilhidrolasa (programa BLAST, versión 2.2.9 en el NCBI, Nacional Center for Biotechnology Information; Altschul *et al.*, 1990).

El inserto de los plásmidos pALK1710, pALK1711, pALK2005 y 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 amplificada de *Acremonium* o *Chaetomium*. La temperatura de hibridación 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%. Se obtuvieron varias placas positivas, de las cuales de cinco a seis placas que hibridación fuertemente se purificaron a partir de cada escrutinio. Los ADN de los fagos se aislaron y se analizaron mediante hibridación por transferencia Southern. Los fragmentos de restricción que hibridaban con la sonda se subclonaron en el vector pBluescript II KS + (Stratagene, USA) y se secuenciaron las porciones relevantes. En todos los casos el fragmento de fago subclonado contiene 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 a partir de los cuales 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 del fago subclonado, y los números de depósito de la colección de cultivos Deutsche Sammlung von und Mikroorganismen Zellkulturen GmbH (DSM) de cepas de *E. coli* que portaban estos plásmidos.

Tabla 15. Sondas utilizadas para la clonación del gen de la endoglucanasa, clon del fago y subclón seleccionados, 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>
<i>cel45A</i> de <i>At</i>	<i>A. thermophilum</i>	pALK1710	P24	<i>Sma</i> I de 5,5 kb	pALK1908	DSM 17324
	ALKO4245					
<i>cel45B</i> de <i>At</i>	<i>A. thermophilum</i>	pALK1711	P41	<i>Xho</i> I de 6,0 kb	pALK1904	DSM 17323
	ALKO4245					
<i>cel7B</i> de <i>Ct</i>	<i>C. thermophilum</i>	pALK2005	P55	<i>Bam</i> HI de 5,1 kb	pALK2010	DSM 17729
	ALKO4261					

El gen *cel5A* de *Thermoascus aurantiacus* (que codificaba EG_28) (SEQ ID NO: 9) se amplificó directamente a partir del ADN genómico aislado mediante reacción de PCR. los cebadores directo (ATTAACCGCGACTGCGCATCATGAAGCTCGGCTCTCTCGTGCTC) e inverso (AACTGAGGCATAGAACTGACGTCATATT) que se utilizaron para la amplificación se diseñaron basándose en el gen *eg1* de *T. aurantiacus* publicado (AF487830). Las mezclas de reacción PCR contenían 1 x tampón Phusion HF,

dNTP 0,3 mM, 0,5 M 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, 30 s de recocido a 57-67°C, 2,5 minutos de extensión 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 (del codón de INICIO al de 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 peptídicas de las endoglucanasas EG_40 de *Acremonium* (gen *cel45A* de At), EG_54 de *Chaetomium* (gen *cel7B* de Ct), y EG_28 de *Thermoascus* (gen *cel5A* de Ta) se encontraron en las correspondientes secuencias de aminoácidos deducidas de los genes clonados confirmando que se clonaron los genes apropiados.

Tabla 16 Resumen de los genes de endoglucanasa aislados 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 intrones (pb)	SEQ ID NO:
<i>cel45A</i> de At	1076	891	2	59, 123	11
<i>cel45B</i> de At	1013	753	2	155, 102	13
<i>cel7B</i> de Ct	1278	1275	-	-	15
<i>cel5A</i> de Ta	1317	1005	5	55, 60, 59, 74, 61	9

(a) El codón de PARADA está incluido.

(b) El codón de PARADA no está incluido.

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

Proteína endoglucanasa	Núm. de aa	Longitud de la ss de NN/HMM ^(a)	CBD ^(b)	PM pronosticado (Da, ss no incluida) ^(c)	pl pronosticado (ss no incluida)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
EG_40 de At	297	21/21	Si, K265 a L297	28625	4,79	2	12
De tipo EG_40 de At	251	20/20	No	23972	6,11	2	14
EG_54 de Ct	425	17/17	No	45358	5,44	1	16
EG_28 de Ta	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) Se indican la presencia de un dominio de unión a celulosa en la proteína, 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 en el servidor ExPASy (Gasteiger *et al.*, 2003).

(d) Los supuestos sitios de N-glicosilación N-X-S/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 (*Cel45A* de At) y de tipo EG_40 (*Cel45B* de At) de *Acremonium*, EG_54 (*Cel7B* de Ct) de *Chaetomium*, y EG_28 (*Cel5A* de Ta) de *Thermoascus* comparten homología con celulasas de la familia 45 (*Acremonium*), la familia 7 (*Chaetomium*), y la familia 5 (*Thermoascus*) de glicosilhidrolasa, identificando de este modo 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 endoglucanasas de *Thielavia terrestris* (CQ827970, identidad del 77,3%) y *Myceliophthora thermophila* (AR094305, identidad de 66,9%), respectivamente (Tabla 18). Las dos endoglucanasas aisladas de la familia 45 de *Acremonium*

comparten una identidad de solo 53,7% entre sí. De estas enzimas sólo EG_40/Cel45A contiene un dominio de unión a celulosa (CBD).

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

La secuencia de la proteína de la endoglucanasa aislada *Thermoascus aurantiacus* es completamente idéntica a la de la EGI publicada de *T. aurantiacus* (AF487830, Tabla 18). La homología más próxima se encontró 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* con sus contrapartes homólogas. Este alineamiento se realizó utilizando el programa de 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> , hipotético 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 hipotético, 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
<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*

Se construyeron plásmidos de expresión para la producción de las proteínas recombinantes EG_40/Cel45A, de tipo EG_40/Cel45B de *Acremonium*, y EG_28/Cel5A de *Thermoascus* como se ha descrito en el Ejemplo 14. Los casetes de expresión lineales (Tabla 19) se aislaron 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 ha descrito en el Ejemplo 14.

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 del casete de expresión ^(a)	Terminador heterólogo ^(b)
EG_40 de At	pALK1920	<i>NotI</i> 10,9 kb	156 pb (<i>HindIII</i>)
De tipo EG_40 de At	pALK1921	<i>EcoRI</i> 8,6 kb	282 pb (<i>SspI</i>)
EG_28 de Ta	pALK1930	<i>NotI</i> 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 a partir de carboximetilcelulosa (2% (p/v) de CMC) a 50°C en tampón de citrato 50 mM de 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. Los anticuerpos policlonales específicos de EG_40 de *Acremonium* se produjeron 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 AP para transferencia Western ProtoBlot (Promega). Los genotipos de los transformantes seleccionados se analizaron mediante transferencia Southern utilizando el casete de expresión como sonda.

El pH óptimo de las endoglucanasas producidas de manera heteróloga se determinó en tampón universal de McIlvaine 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. El pH óptimo para las otras endoglucanasas producidas de manera heteróloga son pH 5,0-5,5 y 6,0 para el 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 ha descrito anteriormente. Se determinó que la mayor actividad de la enzima 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 ha descrito en el Ejemplo 14, en un biorreactor de 2 litros durante cuatro días (28°C, pH 4,2) para obtener material para las pruebas de aplicación.

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

Se utilizaron métodos convencionales de biología molecular como se ha descrito 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 purificadas de *Acremonium*, *Chaetomium*, y *Thermoascus* comparten homología con varias β -glucosidasas de la familia 3 de glicosil hidrolasa 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 los 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 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* de las bibliotecas genómicas correspondientes.

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

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

- 5 (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, sino que se origina a partir de la secuencia de β -glucosidasa (BD168028) de *A. cellulolyticus* homóloga a β G_101.

- 10 Las sondas para el escrutinio de las bibliotecas genómicas construidas se amplificaron 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 μ g de cada cebador, 1 unidad de ADN polimerasa Dynazyme EXT (Finnzymes, Finlandia) y aproximadamente 0,5 μ g 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, 1 min de hibridación 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), 2-3 min de extensión a 72°C y una extensión final a 72°C durante 5-10 min.

- 15 Se aislaron productos de PCR específicos del tamaño esperado (estimado a partir de las secuencias de β -glucosidasa homólogas BD168028, AY072918, y AY368687) 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 mediante secuenciación y realizando hibridaciones de transferencia Southern del 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).

- 20 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* amplificadas como se ha descrito en el Ejemplo 18. Se obtuvieron varias placas positivas, de las cuales de cinco a seis placas que hibridaron fuertemente se purificaron a partir de cada escrutinio. Los ADN de los fagos 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 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 colección de cultivos Deutsche Sammlung von und Mikroorganismen Zellkulturen GmbH (DSMZ) de cepas de *E. coli* que portan estos plásmidos.

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

Gen	Biblioteca genómica	Sonda utilizada en el escrutinio	Clon del fago	Fragmento subclonado	Plásmido	Núm. de depósito de <i>E. coli</i>
ce3A de At	<i>A. thermophilum</i>	pALK1924	P44	6,0 kb	pALK1925	DSM
	ALKO4245			<i>HindIII</i>		17325
ce3A de Ct	<i>C. thermophilum</i>	pALK1935	P51	7,0 kb	pALK2001	DSM
	ALKO4261			<i>XbaI</i>		17667
ce3A de Ta	<i>T. aurantiacus</i>	pALK1713	P21	5,3 kb	pALK1723	DSM
	ALKO4242			<i>BamHI</i>		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 de péptidos de las proteínas β G_101 (At Cel3A) *Acremonium*, β G_76 (Ct Cel3A) de *Chaetomium*, y β G_81 (Ta Cel3A) de *Thermoascus* purificadas se encontraron en las correspondientes secuencias de aminoácidos deducidas de los genes clonados confirmando que se habían clonado los genes apropiados.

Tabla 22. Resumen de los genes de β -glucosidasa aislados 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:
ce3A de At	2821	2583	3	92, 74, 69	23
ce3A de Ct	2257	2202	1	52	25
ce3A de Ta	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 de β -glucosidasa deducidas de *Acremonium thermophilum*, *Chaetomium thermophilum*, y *Thermoascus aurantiacus*. ss, secuencia señal.

Proteína β -glucosidasa	Núm. de aa	Longitud de la ss NN/HMM ^(a)	CBD ^(b)	PM pronosticado (Da, no incluida ss) ^(c)	pl pronosticado no incluida ss)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
β G_101 de At	861	19/18	No	91434	5,46	8	24
β G_76 de Ct	734	20/20	No	76457	6,3	2	26
β G_81 de Ta	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 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 el Compute pI/MW en el 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).

Las secuencias de proteínas deducidas de las β -glucosidasas β G_101/Cel3A de *Acremonium*, β G_76/Cel3A de *Chaetomium*, y β G_81/Cel3A de *Thermoascus* comparten homología con las enzimas de la familia 3 de glicosil hidrolasa, identificando de este modo 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%) era la de β G_76/Cel3A de *C. thermophilum* para la proteína hipotética de *N. crassa* indicando que se habían clonado genes de nuevas enzimas.

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. La alineación se ha realizado mediante 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 (%)
*Acremonium thermophilum 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
*Thermoascus aurantiacus βG_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
*Chaetomium thermophilum βG_76	40,9
Chaetomium thermophilum 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
*Acremonium thermophilum βG_101	40,9
<i>Trichoderma reesei</i> Cel3B, AY281374	40,0
*Thermoascus aurantiacus βG_81	39,9
Thermoascus aurantiacus 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
*Acremonium thermophilum βG_101	62,2
<i>Trichoderma reesei</i> Cel3B, AY281374	57,9
*Chaetomium thermophilum βG_76	39,9

5

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

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

10

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.

15

β -glucosidasa	Plásmido de expresión	Tamaño del casete de expresión ^(a)	Terminador heterólogo ^(b)
β G_101 de At	pALK1933	<i>NotI</i> de 10,5 kb	300 pb (<i>HindIII</i>)

β -glucosidasa	Plásmido de expresión	Tamaño del casete de expresión ^(a)	Terminador heterólogo ^(b)
β G_76 de Ct	pALK2004	EcoRI de 10,1 kb	528 pb (<i>Xba</i> I)
β G_81 de Ta	pALK1914	EcoRI de 10,9 kb	452 pb (<i>Apo</i> I)

(a) El casete de expresión para la transformación en *T. reesei* se aisló de la cadena principal del vector mediante digestión con *Eco* RI o *Not*I.

(b) Se indica el número de nucleótidos después del codón de PARADA del gen clonado que está incluido 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.

Se analizó La producción de beta-glucosidasa de los transformantes a partir de los sobrenadantes de cultivo de 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. También se detectó la producción de proteínas recombinantes a partir de los sobrenadantes de cultivo mediante electroforesis en gel de poliacrilamida-SDS. Además, se verificó la expresión de β G_81 de *Thermoascus* mediante análisis de transferencia de Western con anticuerpos anti- β G_81 como se ha descrito en el Ejemplo 19. Los genotipos de los transformantes seleccionados se analizaron mediante transferencia Southern utilizando el casete de expresión como sonda.

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

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

Ejemplo 22. Clonación de los genes de xilanasas de *Acremonium thermophilum* ALKO4245 y *Thermoascus aurantiacus* ALKO4242

Se utilizaron métodos convencionales de la biología molecular como se ha descrito 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 *Acremonium* purificada comparten homología con las xilanasas de la familia 10 de glicosil hidrolasa tales como *Humicola grisea* XYNI (AB001030). Todos los péptidos derivados de xilanasas de *Thermoascus* fueron completamente idénticos a la secuencia XynA publicada de *Thermoascus aurantiacus* (AJ132635) identificando de este modo 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 basándose en 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 dedujeron a partir de la comparación con secuencia XYNI de *Humicola insolens* homóloga (AB001030). La secuencia del cebador efector (GAYGGYGAYGCSACYTAYATG) se basa en el Péptido 3 (aminoácidos 2-8) y el cebador anti-sentido (YTTYTGRTCTAYTCSAGRTTTRTA) en el Péptido 1 (aminoácidos 4-11).

Se obtuvo un producto de PCR del tamaño esperado (estimado a partir del secuencia XYNI de *Humicola insolens* homólogo AB001030) de la reacción. 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 hidrolasa (programa BLAST del 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 escrutaron aproximadamente $1-2 \times 10^5$ placas de la biblioteca genómica de *Acremonium* amplificada como se ha descrito en el Ejemplo 18. Se obtuvieron varias placas positivas, de las cuales se purificaron cinco placas que hibridaban fuertemente. Los ADN de fagos se aislaron y se analizaron mediante hibridación de transferencia Southern. Un fragmento de restricción *Xba*I 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 *xyn10A* de *Acremonium thermophilum* completo (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 *xyn10A* de *Thermoascus aurantiacus* (SEQ ID NO: 17) se amplificó directamente a partir del ADN genómico aislado mediante reacción de PCR. Los cebadores directo (TTATACCGCGGGAAGCCATGGTTCGACCAACGATCCTAC) e inverso (TTATAGGATCCACCGGTCTATACTCACTGCTGCAGGTCCTG) que se utilizaron en la amplificación del gen se diseñaron basándose en el gen *XynA* de *T. aurantiacus* publicado (AJ132635). 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 0,3 mM, cada cebador 1 µM, 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, 1 min de recocido a 60-66°C, 3 min de extensión 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 (desde el codón 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 contenía 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 proteínas deducidas (SEQ ID NO: 17-20) se resumen en la Tabla 26 y la Tabla 27, respectivamente. Las secuencias peptídicas de las proteínas purificados XYN_60 de *Acremonium* y XYN_30 de *Thermoascus* se encontraron en las correspondientes secuencias de aminoácidos deducidas de los genes clonados (*At xyn10A* y *Ta xyn10A*, respectivamente) confirmando que los se habían clonado los genes apropiados.

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

Gen de xilanasa	Longitud con intrones (pb) ^(a)	Región codificante (pb) ^(b)	Núm. de intrones	Longitud de los intrones (pb)	SEQ ID NO:
<i>At xyn10A</i>	1471	1248	2	135, 85	19
<i>Ta xyn10A</i>	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 de xilanasa deducidas DE *Acremonium thermophilum* y *Thermoascus aurantiacus*. ss, secuencia señal.

Proteína xilanasa	Núm. de aa	Longitud de la ss NN/HMM ^(a)	CBD ^(b)	PM pronosticado (Da, ss no incluida) ^(c)	pl pronosticado (ss no incluida)	Supuestos sitios de N-glicosilación ^(d)	SEQ ID NO:
XYN_60 de <i>At</i>	416	19/19	Si, W385 a L416	42533	6,32	1-2	20
XYN_30 de <i>Ta</i>	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 el valor HMM utilizando modelos ocultos de Markov.

(b) Se indican la presencia de un dominio de unión a carbohidratos CBD, los aminoácidos del CBD C-terminal (numeración de acuerdo con el polipéptido completo)

(c) No está incluida la secuencia pronosticada. La predicción se realizó utilizando la herramienta Compute pI/MW en el 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 Lo Leggio *et al.*, 1999

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

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 de 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
<i>Humicola grisea</i> XYL I, AB001030	67,1
<i>Magnaporthe grisea</i> 70-15, hipotética XM_364947	63,8
<i>Aspergillus aculeatus</i> xilanasas, AR149839	53,7
<i>Talaromyces emersonii</i> xilanasas, AX403831	51,8
<i>Gibberella zeae</i> xilanasas, AY575962	51,4
<i>Magnaporthe grisea</i> XYL5, AY144348	48,5
<i>Talaromyces emersonii</i> , AX172287	46,9

5

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

Se construyeron plásmidos de expresión para la producción de las proteínas recombinantes XYN_60/Xyn10A de *Acremonium* y XYN_30/Xyn10A de *Thermoascus* como se ha descrito en el Ejemplo 14. Los casetes de expresión lineales (Tabla 29) se aislaron 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 ha descrito en el Ejemplo 14.

10

Tabla 29. Casetes de expresión construidos para la producción de las xilanasas XYN_60/Xyn10A de *Acremonium thermophilum* y XYN_30/Xyn10A de *Thermoascus aurantiacus* en *Trichoderma reesei*. Las estructuras esquemáticas de los casetes de expresión se describen en la Figura 2.

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

15

(a) El casete de expresión para la transformación en *T. reesei* se aisló de 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 está incluido 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.

20

Se analizó la producción de xilanasas de los transformantes 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 xilano de abedul (1% p/v) a 50°C en tampón citrato 50 mM de pH 5,3 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 ha descrito en el Ejemplo 19. Los genotipos de los transformantes seleccionados se analizaron mediante transferencia Southern utilizando el casete de expresión como sonda.

25

Se produjo XYN₃₀/Xyn10A de *Thermoascus* en *T. reesei* y se determinó el pH óptimo de la proteína producida de manera heteróloga en tampón de universal McIlvaine en un intervalo de pH de 3,0-8,0 utilizando como sustrato xilano de abedul (Figura 8A). Se determinó que el pH óptimo era de 4,5. Se determinó que la temperatura óptima para la actividad enzimática de XYN₃₀ era de 75°C (Figura 8B).

- 5 Los transformantes seleccionados se cultivaron, como se ha descrito en el Ejemplo 14, en un biorreactor de 2 litros durante cuatro días (28°C, pH 4,2) para obtener material para las pruebas de aplicación.

Ejemplo 24. Rendimiento de las celobiohidrolasas recombinantes en la hidrólisis

- 10 El rendimiento de las celobiohidrolasas recombinantes purificadas se evaluó en los estudios de hidrólisis con las enzimas de *T. reesei* purificadas. La hidrólisis se llevó a cabo con mezclas controladas de enzimas purificadas sobre varios sustratos pretratados. Se obtuvieron productos filtrados de cultivo de *T. reesei*, que contenían diferentes enzimas CBH/Cel7 clonadas como se ha descrito en los Ejemplos 14 y 15, y las enzimas CBH se purificaron mediante cromatografía de afinidad como se ha descrito 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:

- 15 CBH de *Thermoascus aurantiacus* ALKO4242 (Cel7A de Ta)

CBH de *Thermoascus aurantiacus* ALKO4242 (Cel7A de Ta) con CBD anclado genéticamente de *Trichoderma reesei* (Cel7A de Ta + CBD de Tr)

CBH de *Thermoascus aurantiacus* ALKO4242 (Cel7A de Ta) con el CBD anclado genéticamente de *Chaetomium thermophilum* (Cel7A de Ta + CBD de Ct)

- 20 CBH de *Acremonium thermophilum* ALKO4245 (Cel7A de At)

CBH de *Chaetomium thermophilum* ALKO4265 (Cel7A de Ct).

- 25 Cada CBH/Cel7 que se iba a someter a ensayo (dosificación 14,5 mg/g de materia seca de sustrato) se utilizó o bien con EGII/Cel5A de *T. reesei* (3,6 mg/g) o bien con una mezcla que contenía EGI/Cel7B (1,8 mg/g), EGII/Cel5A (1,8 mg/g) de *T. reesei*, 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 de peso seco). Se incubaron 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 en la mezcla mediante agitación magnética a 45°C durante 48 h. También se prepararon las muestras de referencia con las enzimas inactivadas y los sustratos correspondientes. Se midió la liberación de los productos de hidrólisis como azúcares reductores con el método DNS utilizando glucosa como patrón (Tabla 30).

Se utilizaron los siguientes sustratos en el experimento:

Celulosa cristalina (Avicel)

Fibra de píceas 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), materia seca de 25,9% (PICEA).

- 35 Fibra de rastrojo de maíz oxidada en mojado lavada ("WOCS en sus siglas en inglés").

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

- 40 Tabla 30. Productos de hidrólisis con enzimas CBH (45°C, pH 5,0). 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	PICEA	WOCS	SAUCE
Cel7A de Ta	EGII, bG	2,0	2,0	2,8	2,0
Cel7A de Ta +CBD de Tr	EGII, bG	5,8	4,0	4,4	4,0
Cel7A de Ta +CBD de Ct	EGII, bG	4,9	3,7	4,6	3,7
Cel7A de At	EGII, bG	5,3	3,3	4,5	3,3
Cel7A de Ct	EGII, bG	6,0	2,6	3,4	2,6

Enzimas		Sustratos			
CBH	Enzimas adicionales	Avicel	PICEA	WOCS	SAUCE
Cel7A de <i>T. reesei</i>	EGII, bG	4,7	2,9	2,9	2,9
Cel7A de Ta	EGII, EGI, XYL, AXE, bG	nd	nd	4,3	2,8
Cel7A de Ta +CBD de Tr	EGII, EGI, XYL, AXE, bG	nd	nd	7,2	5,9
Cel7A de Ta +CBD de Ct	EGII, EGI, XYL, AXE, bG	nd	nd	7,2	5,6
Cel7A de At	EGII, EGI, XYL, AXE, bG	nd	nd	6,4	5,4
Cel7A de Ct	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 dosificación de proteína en la hidrólisis. Los resultados demuestran que sobre los sustratos celulósicos (Avicel y fibra de píceas) Cel7A de *Thermoascus aurantiacus* con el CBD unido genéticamente mostraba una hidrólisis claramente superior a la hidrólisis que CBHI/Cel7A de *T. reesei*. Sin CBD, Cel7A de *T. aurantiacus* fue menos eficaz sobre estos sustratos. El rendimiento de las celobiohidrolasas de *Acremonium thermophilum* y *Chaetomium thermophilum* también fue mejor que el de CBHI/Cel7A de *T. reesei* sobre varios sustratos; en particular, Cel7A de *C. thermophilum* mostró una alta eficacia sobre la celulosa pura (Avicel).

En el caso de los sustratos que contenían cantidades notables de hemicelulosa (sauce y el rastrojo de maíz), las enzimas CBH/Cel7 necesitaban claramente, tanto hemicelulasas como endoglucanasas adicionales para funcionar eficazmente. Si no se encontraban presentes hemicelulasas adicionales, Cel7A de *T. aurantiacus* con CBD anclado genéticamente mostró de nuevo una hidrólisis claramente máxima. Con las enzimas degradantes de hemicelulosa más importantes (xilanasa, acetil xilano esterasa y EGI), Cel7A de *T. aurantiacus* con CBD anclado genéticamente funcionaba 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 rendimiento hidrolítico de Cel7A de *T. aurantiacus*, incluso aunque la diferencia fuera bastante pequeña.

Se puede concluir que cuando se reemplaza 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 trasposiciones experimentales resultó claramente mejorada 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 de temperatura de las celobiohidrolasas producidas en la presente memoria, los resultados indican que el rendimiento de las mezclas de enzimas celulósicas a temperaturas más altas de 45°C se puede mejorar claramente utilizando las celobiohidrolasas producidas en la presente memoria.

Ejemplo 25. Rendimiento de las endoglucanasas recombinantes en la hidrólisis

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

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

Endoglucanasa de tipo EG_40 de At/Cel45B de *Acremonium thermophilum* ALKO4245 (de tipo EG_40 de ALKO4245)

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

Se utilizaron los siguientes sustratos en el experimento:

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

Fibra de rastrojo de maíz sometida a explosión de vapor de agua (pretratamiento con vapor de agua a 210°C durante 5 min), materia seca 31,0% ("SECS en sus siglas en inglés").

Las endoglucanasas que se iban a estudiar (dosificación de 840 nkat/g de materia seca, basándose en la actividad endoglucanasa contra HEC de acuerdo con la IUPAC, 1987) se utilizaron o bien con celobiohidrolasas de *T. reesei* (CBHI/Cel7A, 8,1 mg/g de materia seca y CBHII/Cel6A, 2,0 mg/g de materia seca) o bien con Cel7A de *Thermoascus aurantiacus* con el CBD de *T. reesei* anclado genéticamente (10,1 mg/g de materia seca). 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 de peso seco, 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 muestra de referencia con enzimas inactivadas y los sustratos correspondientes. Se midió la liberación de los productos de la hidrólisis como 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) como 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	PICEA	SECS
EG no añadida	CBHI y CBHII de <i>T. reesei</i>	2,4	3,2
EGI	CBHI y CBHII de <i>T. reesei</i>	3,5	4,6
EGII	CBHI y CBHII de <i>T. reesei</i>	3,8	3,5
EG_40 de At	CBHI y CBHII de <i>T. reesei</i>	4,9	4,3
At de tipo EG_40	CBHI y CBHII de <i>T. reesei</i>	4,5	4,8
EG_28 de Ta	CBHI y CBHII de <i>T. reesei</i>	3,0	3,9
sin EG añadida	Cel7A de <i>T. aurantiacus</i> + CBD de Tr	1,8	2,1
EG I	Cel7A de <i>T. aurantiacus</i> + CBD de Tr	nd.	4,2
EG II	Cel7A de <i>T. aurantiacus</i> + CBD de Tr	3,2	nd.
EG_40 de At	Cel7A de <i>T. aurantiacus</i> + CBD de Tr	4,8	4,0
EG_28 de Ta	Cel7A de <i>T. aurantiacus</i> + CBD de Tr	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) utilizadas en el análisis y subestimar la eficacia de las enzimas con alta actividad específica contra hidroxietilcelulosa. En cualquier caso, los resultados demuestran que las endoglucanasas de *Acremonium thermophilum* funcionan muy bien en la hidrólisis cuando influyen juntas en 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 sobre el sustrato rastrojo de maíz que contiene hemicelulosa debido a su fuerte actividad xilanasa secundaria. La endoglucanasa Cel5A de *T. aurantiacus* (EG_28 de ALKO4242) mostró una hidrólisis inferior que las enzimas de *T. reesei*.

Se puede concluir que las endoglucanasas de *A. thermophilum* funcionan con una eficacia comparable o mayor en comparación con las correspondientes enzimas de *Trichoderma* en la hidrólisis a juzgar por esta disposición experimental. Considerando también la estabilidad de temperatura de las endoglucanasas descritas en la presente memoria, los resultados indican que el rendimiento de las 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.

Ejemplo 26. La hidrólisis de píceas pre-tratada con vapor de agua a altas temperaturas

Se suspendió fibra de píceas 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 25,9% de materia seca en 5 ml de tampón de acetato de sodio 0,05 M con una 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 muestreo, se retiraron de la hidrólisis tubos de ensayo por triplicado, 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 la reacción de hidrólisis. Los blancos que contenían

sustrato solo (solamente tampón añadido en lugar de enzimas) también se incubaron en las condiciones correspondientes.

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

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

- 10 Preparación de endoglucanasa termófila que contiene endoglucanasa At EG_40/Cel45A de *Acremonium thermophilum* ALKO4245. La proteína se produjo en *T. reesei* como se ha descrito 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 422 nkat/ml de actividad endoglucanasa (de acuerdo con la IUPAC, 1987).

- 15 Preparación de β -glucosidasa termófila preparada como se ha descrito en el Ejemplo 21 que contenía β -glucosidasa β G_81 de Ta/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 actividad β -glucosidasa de 6270 nkat/ml (de acuerdo con Bailey y Linko, 1990).

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

- 20 Como comparación y 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 Novozym 188 (de Novozymes A/S). Ésta se denominó "ENZIMAS de *T. REESEI*".

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

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

- 30 Los resultados muestran claramente un mejor rendimiento 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 sobre 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 seco de píceas. Debido a la dosificación relativamente baja de enzima, los rendimientos finales de azúcar eran claramente inferiores. Para las 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
- 35 *Trichoderma* del estado de la técnica éste fue solo de 31% y 11% a 55°C y 60°C, respectivamente.

Ejemplo 27. Hidrólisis de rastrojo de maíz pre-tratado con vapor de agua a altas temperaturas

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

- 40 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 de CBHI/Cel7A de *T. reesei* anclado genéticamente (Cel7A de Ta + CBD de Tr, Ejemplo 15). El contenido de proteína de la preparación fue 31 mg/ml.

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

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

Producto xilanasa termófila que contenía una xilanasa AM24 procedente de *Nonomuraea flexuosa* DSM43186. El producto se preparó mediante el uso de una cepa recombinante de *Trichoderma reesei* que había sido transformada con el casete de expresión de pALK1502, como se describe en el documento WO2005/100557. El producto sólido se

disolvió en agua para preparar una disolución al 10% y se obtuvo una preparación de enzima con actividad xilanasa (analizada de acuerdo con Bailey *et al.*, 1992) de 208000 nkat/ml.

Estas preparaciones de enzimas se combinaron como sigue (por 10 ml de mezcla): preparación de CBH/Cel7 7,79 ml, preparación de endoglucanasa 0,96 ml, preparación de β -glucosidasa 1,14 ml y preparación de xilanasa 0,31 ml. Esta mezcla se utilizó como "MEZCLA 2" de las enzimas termófilas.

Como comparación y 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). Ésta 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 la hidrólisis se cuantificaron utilizando el análisis para 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 los productos de hidrólisis como azúcares reductores se presentan en la Figura 10.

Los resultados muestran claramente un mejor rendimiento de las enzimas descritas aquí 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. Basándose en el análisis de HPLC el máximo rendimiento de azúcares (incluyendo azúcares solubles libres en el sustrato no lavado que se había utilizado) del sustrato sería de 5,73 mg por 10 mg de sustrato seco. De este modo, la hidrólisis por las enzimas de la mezcla 2 fue casi completa en el plazo de 48 horas. A 55°C y 57,5°C las enzimas termófilas descritas la presente memoria también mostraron un rendimiento claramente mejor en la hidrólisis, en comparación con las enzimas de *Trichoderma* del estado de la técnica.

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

Se repitió el procedimiento explicado en el Ejemplo 27 excepto que el producto de xilanasa XT 02026A3 se reemplazó por preparación de xilanasa termófila que contenía xilanasa XYN_30 de Ta/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).

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

Como comparación y 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). Ésta 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 la hidrólisis se cuantificaron utilizando el análisis para 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 los productos de hidrólisis como azúcares reductores se presenta en la Figura 11.

Los resultados muestran claramente un mejor rendimiento 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 de *T. reesei*. A 55°C y 60°C las enzimas termófilas descritas en la presente memoria mostraron un rendimiento claramente mejor en la hidrólisis, en comparación con las enzimas de *Trichoderma* del estado de la técnica. El rendimiento de la nueva mezcla de enzimas a 60°C estuvo al mismo nivel que el rendimiento de las enzimas del estado de la técnica a 45°C.

Ejemplo 29. Hidrólisis de píceas pre-tratada a altas temperaturas utilizando una mezcla con una xilanasa termoestable

Se repitió el procedimiento descrito en el Ejemplo 28 con fibra de píceas 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 25,9% de materia seca) 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 la hidrólisis se cuantificaron utilizando el análisis para 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 los productos de hidrólisis como azúcares reductores se presenta en la Figura 12.

Los resultados muestran claramente un mejor rendimiento 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 eficaz 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 todavía estuvo al mismo nivel que a 45°C, mientras que la mezcla de estado de la técnica fue 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 estuvo 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. Como comparación, también se incluyeron en el experimento las enzimas comerciales Celluclast 1,5 L (β -glucosidasa de 534 nkat/ml) y Novozym 188 (β -glucosidasa de 5840 nkat/ml).

Con el fin de evaluar la sensibilidad de las diferentes β -glucosidasas a la inhibición por glucosa, se realizó un procedimiento de análisis de la actividad patrón 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 la glucosa de manera que la concentración de glucosa en la mezcla de análisis se ajustó a valores entre 0 y 0,5 M. Excepto esta adición de glucosa el análisis se realizó utilizando el procedimiento convencional (Bailey y Linko, 1990). Las actividades en presencia de concentraciones variables de glucosa como porcentaje de la actividad sin glucosa se presentan en la Figura 13.

Los resultados demuestran que las β -glucosidasas de *C. thermophilum* y *T. aurantiacus* se vieron 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 1,5 L. La enzima de *A. thermophilum* mostró un comportamiento comparable a la enzima de *T. reesei* de Celluclast. Especialmente la enzima de *C. thermophilum* fue claramente menos afectada por la alta concentración de glucosa. Por lo tanto, estos resultados indican que teniendo en cuenta la inhibición por glucosa el uso de las nuevas β -glucosidasas, 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 en papel de filtro de mezclas de enzimas a altas temperaturas

La actividad en papel de filtro de las preparaciones de enzimas se midió de acuerdo con el método de la IUPAC (1987) como se ha descrito en el procedimiento, excepto 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% de sustrato de papel de filtro en 1 h en condiciones experimentales. Se considera que la actividad FPU representa la actividad celulasa global total de una preparación de enzima.

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

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

"ENZIMAS A de T. REESEI" preparada como se describe la preparación de "ENZIMAS de T. REESEI" descrita en el Ejemplo 26.

"ENZIMAS B de T. REESEI" se construyó 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 de enzimas a diferentes temperaturas se presentan en la Figura 14 como porcentajes de la actividad en condiciones convencionales (IUPAC, 1987) (a 50°C).

Los resultados demuestran claramente que las mezclas descritas aquí muestran una actividad celulasa global superior a temperaturas elevadas (60-70°) 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

Una mezcla de producto hidrolizado de almidón de alta concentración (Nutriosa 74/968, Roquette) se trató con preparación enzimática enriquecida pG_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.

- 5 La preparación enzimática enriquecida en β G₈₁/Cel3A de Ta se añadió a la disolución de Nutriosa al 70% (p/p) 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 una fuente de carbono en un medio en matraces oscilantes para dos cepas de *Trichoderma* diferentes (A47 y Rut-C30). El efecto del tratamiento de la enzima se midió como la actividad endoglucanasa formada durante un cultivo en matraz oscilante durante 7 días. Como referencia se realizaron cultivos en las mismas condiciones con Nutriosa no tratada como fuente de carbono. Se obtuvo un incremento de más del doble en las actividades en los cultivos en matraz oscilante realizados en medio de Nutriosa pretratado con β G₈₁/Cel3A de Ta con las cepas sometidas a ensayo. Los resultados se muestran en la Figura 15.

Lista de depósito de organismos

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 May 2005	DSM 17323
<i>Escherichia coli</i>	pALK1908	DSMZ	13 May 2005	DSM 17324
<i>Escherichia coli</i>	pALK1925	DSMZ	13 May 2005	DSM 17325
<i>Escherichia coli</i>	pALK1926	DSMZ	13 May 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

- 10 ⁽¹⁾ Centraalbureau Voor Schimmelcultures en Uppsalalaan 8, 3584 CT, Utrecht, Países Bajos

⁽²⁾ Centraalbureau Voor Schimmelcultures en Oosterstraat 1, 3742 SK BAARN, Países Bajos

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

- 15 ⁽⁴⁾ [Después de la terminación del período de depósito actual, las muestras se almacenarán bajo los acuerdos para dejar la cepa disponible más allá del tiempo aplicable 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.
- 20 Badger, P.C. (2002) Ethanol from cellulose: a general review. In Trends in new crops and new uses. J. Janick y A. Whipkey (eds.). ASHS Press, Alexandria, VA, USA, pp. 17-21.
- Bailey M. J. y 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.
- 25 Bailey, M.J., Biely, P. y Poutanen, K. (1992) Interlaboratory testing for assay of xylanase activity. J Biotechnol. 23:257-270.
- 30 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.
- 5 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.
- 10 Bernfeld, B. (1955) Amylases, α and β . In: *Methods in Enzymology*, vol. 1, Eds. Colowick, S.P. y Kaplan, N.O. Academic Press, New York, págs. 149-158.
- 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.
- 15 Coen, D.M. (2001) The polymerase chain reaction. In: 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.
- 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.
- 20 Gellissen, G. (ed.) (2005) *Production of recombinant proteins. Novel microbial and eukaryotic expression systems*. Wiley-VCH Verlag GmbH&Co. Weinheim, Germany.
- 25 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/
- 30 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.
- 35 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
- 45 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.
- 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.
- 50 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.
- 55 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.
- 60 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.
- 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 Biotechn.* Vol 121-124: 219-229.
- 65

- 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.
- 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.
- 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.
- Parry N., Beever D., Owen E., Nerinckx W. Claeysens 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.
- 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.
- 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 acetylxylan esterases of *Trichoderma reesei*. Biotechnol. Appl. Biochem. 13: 1-11.
- 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.
- Tomme, P. McRae, S., Wood, T. y Claeysens, M. (1988) Chromatographic separation of cellulolytic enzymes. Methods in Enzymol. 160: 187-192.
- Tuohy M., Walsh J., Murray P., Claeysens 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, págs. 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.
- 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

5	<110> Roal		
	<120> Tratamiento de material celulósico y enzimas útiles en el mismo		
	<130> 2051999		
10	<160> 30		
	<170> PatentIn versión 3.1		
15	<210> 1		
	<211> 3192		
	<212> ADN		
	<213> Thermoascus aurantiacus		
20	<220>		
	<221> CDS		
	<222> (1514)..(2122)		
	<223>		
25	<220>		
	<221> Intrón		
	<222> (2123)..(2187)		
30	<223>		
	<220>		
	<221> CDS		
35	<222> (2188)..(2949)		
	<223>		
	<400> 1		
	ctagaccttt atcctttcat ccgaccagac ttcccccttg accttggcgc cctgttgact	60	
	acctacctac ctaggtaggt aacgtcgtcg accctcttga atgatcctcg tcacactgca	120	
	aacatccgaa acatacggca aaagatgatt gggcatggat gcaggagaca tcgaatgagg	180	
	gcttagaagg aaatgaaaac ctgggaccag gacgctaggt acgatgaaat ccgccaatgg	240	
	tgaaacttta agtcgtgcct acagcacagg ctctgtgaag attgcgctgt tcagacttaa	300	
	tcttctcadc acagtccaag tctttatgaa aaggaaaaga gagagaagag cgctatttcg	360	
	agctgtcggc ctcataggga gacagtcgag cataccagcg gtatcgacgt tagactcaac	420	
	caagaataat gacgagaata aacacagaag tcaaccttga actgtatatc agggttccag	480	
	cagcagatag ttacttgcat aaagacaact ccccgagggc tctctgcata caccaggatg	540	
	ttccggaatt attcactgct cgtttccgac gtggcgctcag tgatccgtct ccacagaacc	600	
	tctacctggg gaataaccca ggggaggaat ctgcaagtaa gaacttaata ccaatccccg	660	
	gggctgccgg ggtgaatcaa atctcccgcg ggaaattaaa cccatacgat gtttttgcac	720	
40	cacatgcatg cttggcacga tttctccgca agggagtcac agagaaagac atatttcgca	780	

tactactgtg actctgcaga gttacatatc actcaggata cattgcagat cattgtccga	840
gcatcaaaca tggacctgca ggatcaacgg cccgacaaaa cacaagtggc taaagctggg	900
ggatgcccga acccgctgcg caatatcatt gatggatggt cccccacatt tttaaaacat	960
cgacggatcg gcccgcatatc taatcctttt atcaacaaaa agttccactc gactagagaa	1020
aaaaaggcca aggccactaa ttgcagtcgg atactgggtct ttctgcgctc caacaccttc	1080
atccatgatc cccttagcca ccaatgcccc acataataca tgttgacata ggtacgtagc	1140
tctgttatcc aatcgcatcc gaacctcttt aacggacccc tctacacac cttatcctaa	1200
cttcaggaga ctgttgccca ttggggattg aggaggtccg gggtgcagga tgcgttctag	1260
gctaaattct cggccggtag ccatctcgaa tctctcgtga agccttcac tgaacggttg	1320
gcggcccgtc aagccgatga ccatgggttc ctgatagagc ttgtgcctga ccggccttgg	1380
cggcatagac gagctgaaca catcaggtat gaacagatca gatataaagt cggattgagt	1440
cctagtacga agcaatccgc caccacaaa tcaagcaacg agcgacagca ataacaatat	1500
caatcgaatc gca atg tat cag cgc gct ctt ctc ttc tct 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
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	

ES 2 501 944 T3

ggt cag gag ttt acc ttc gat gtc gac gtc tcc aat ctc cct tgc ggg Gly Gln Glu Phe Thr Phe Asp Val Asp Val Ser Asn Leu Pro Cys Gly 145 150 155	1981
ctg aac ggc gcc ctc tac ttt'gtg gcc atg gac gcc gac ggc gga ttg Leu Asn Gly Ala Leu Tyr Phe Val Ala Met Asp Ala Asp Gly Gly Leu 160 165 170	2029
tcc aaa tac cct ggc aac aag gca ggc gct aag tat ggc act ggt tac Ser Lys Tyr Pro Gly Asn Lys Ala Gly Ala Lys Tyr Gly Thr Gly Tyr 175 180 185	2077
tgc gac tct cag tgc cct cgg gat ctc aag ttc atc aac ggt cag Cys Asp Ser Gln Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln 190 195 200	2122
gtacgtcaga agtgataact agccagcaga gcccatgaat cattaactaa cgctgtcaaa	2182
tacag gcc aat gtt gaa ggc tgg cag ccg tct gcc aac gac cca aat gcc Ala Asn Val Glu Gly Trp Gln Pro Ser Ala Asn Asp Pro Asn Ala 205 210 215	2232
ggc gtt ggt aac cac ggt tcc tgc tgc gct gag atg gat gtc tgg gaa Gly Val Gly Asn His Gly Ser Cys Cys Ala Glu Met Asp Val Trp Glu 220 225 230	2280
gcc aac agc atc tct act gcg gtg acg cct cac cca tgc gac acc ccc Ala Asn Ser Ile Ser Thr Ala Val Thr Pro His Pro Cys Asp Thr Pro 235 240 245 250	2328
ggc cag acc atg tgc cag gga gac gac tgt ggt gga acc tac tcc tcc Gly Gln Thr Met Cys Gln Gly Asp Asp Cys Gly Gly Thr Tyr Ser Ser 255 260 265	2376
act cga tat gct ggt acc tgc gac cct gat ggc tgc gac ttc aat cct Thr Arg Tyr Ala Gly Thr Cys Asp Pro Asp Gly Cys Asp Phe Asn Pro 270 275 280	2424
tac cgc cag ggc aac cac tcg ttc tac ggc ccc ggg cag atc gtc gac Tyr Arg Gln Gly Asn His Ser Phe Tyr Gly Pro Gly Gln Ile Val Asp 285 290 295	2472
acc agc tcc aaa ttc acc gtc gtc acc cag ttc atc acc gac gac ggg Thr Ser Ser Lys Phe Thr Val Val Thr Gln Phe Ile Thr Asp Asp Gly 300 305 310	2520
acc ccc tcc ggc acc ctg acg gag atc aaa cgc ttc tac gtc cag aac Thr Pro Ser Gly Thr Leu Thr Glu Ile Lys Arg Phe Tyr Val Gln Asn 315 320 325 330	2568
ggc aag gta atc ccc cag tcg gag tcg acg atc agc ggc gtc acc ggc Gly Lys Val Ile Pro Gln Ser Glu Ser Thr Ile Ser Gly Val Thr Gly 335 340 345	2616
aac tca atc acc acc gag tat tgc acg gcc cag aag gcc gcc ttc ggc Asn Ser Ile Thr Thr Glu Tyr Cys Thr Ala Gln Lys Ala Ala Phe Gly 350 355 360	2664
gac aac acc ggc ttc ttc acg cac ggc ggg ctt cag aag atc agt cag Asp Asn Thr Gly Phe Phe Thr His Gly Gly Leu Gln Lys Ile Ser Gln 365 370 375	2712

ES 2 501 944 T3

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 gtggttgagg 3009

gattgtcgtg tacaggggac atgcgatgct caaaaatcta catcagtttg ccaattgaac 3069

catgaaaaaa aggggggagat caaagaagtc tgtcaaaaga ggggggctgt ggcagcttaa 3129

gccttgttgt agatcgagtc gacgccctat agtgagtcgt attagagctc gcggccgcga 3189

gct 3192

<210> 2
5 <211> 457
<212> PRT
<213> *Thermoascus aurantiacus*

10 <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

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

ES 2 501 944 T3

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 2 501 944 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
385 390 395 400

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 Asn Ser Thr Phe Thr Ala Asn
450 455

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

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

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

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

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

30 <220>
<221> CDS

<222> (2413)..(2540)
<223>

<220>
5 <221> Intrón
<222> (2541)..(2627)
<223>

<220>
10 <221> CDS
<222> (2628)..(2691)
<223>

<400> 3
gaattcggat cacaccgaga gcttcgcgat ggccagctgt ctcagcttgt acccgtctac 60
caacgttccg catcttcggt accttgatag ctgcggtttg ctggactgct ttgtgagggg 120
actgtgccac gcctgggaga cgggtgccgt accatcggtt actgcgcaga ctgagaaccg 180
tcgttgccga aacagccagg caggaagcct gtccaccttc atgtatcttc atatggaccc 240
cagcgcgccc ctctctttct cctcatttct tgcccaccac gatggacacc atgccaatct 300
atttcttgat cccttgactc ctcagcccc cagcagtcgg acaatgtaca gtgatgggca 360
tctctttctg tacatacgtc cctctcgcg gtgtccacgc gcggccgggg atgcctggga 420
cggagtgcc acccgcaggga acgagacttg gctgatgggg tcggtgcat ggtggcacia 480
gagatccagg ccccccgatc tcgttctcgc acgtatcctt ccccccgcgg cgatgcccac 540
gtgggaagtc ttcggagcgg caccagggc catcttgccg atgcccggca cggctctggc 600
ggttgccctc atctatcgtg gctgcacatc cgccgtgcc ccattgggaa agcaggcttt 660
gttcttcccc tctgtcgate gtctcccacc taccctccct cctcgcaagg gcttaccctg 720
gcccctcact gctgcttcac ctactgctg cttccccgca atgcccctc gcccccccc 780
ccccctctc ctttgagta cagatctaca taatatcgag acgcccccca agctgtttct 840
ctggcacagc cctctcgcgc gtggtgcaag agcaagtcag agtatcaatt ccccatctc 900
tcattctcagc cttctgccc tgggccacc gacattctgg gccgtagcc aagaccgatc 960
cgctctcac c atg cac aag cgg gcg gcc acc ctg tcc gcc ctg gtc gtc 1010
Met His Lys Arg Ala Ala Thr Leu Ser Ala Leu Val Val
1 5 10
gcc gcc gcc ggc ttc gcc cgc ggc cag ggc gtg gcc acg cag cag acg 1058
Ala Ala Ala Gly Phe Ala Arg Gly Gln Gly Val Gly Thr Gln Gln Thr
15 20 25
15 gag acg cac ccc aag ctg acc ttc cag aag tgc tcc gcc gcc ggc agc 1106

ES 2 501 944 T3

Glu Thr His Pro Lys Leu Thr Phe Gln Lys Cys Ser Ala Ala Gly Ser	
30 35 40 45	
tgc acg acc cag aac ggc gag gtg gtc atc gac gcc aac tgg cgc tgg	1154
Cys Thr Thr Gln Asn Gly Glu Val Val Ile Asp Ala Asn Trp Arg Trp	
50 55 60	
gtg cac gac aag aac ggc tac acc aac tgc tac acg ggc aac gag tgg	1202
Val His Asp Lys Asn Gly Tyr Thr Asn Cys Tyr Thr Gly Asn Glu Trp	
65 70 75	
aac acc acc atc tgc gcc gac gcc gcc tcg tgc gcc agc aac tgc gtc	1250
Asn Thr Thr Ile Cys Ala Asp Ala Ala Ser Cys Ala Ser Asn Cys Val	
80 85 90	
gtc gac ggc gcc gac tac cag ggc acc tac ggc gcc tcc acc tcc ggc	1298
Val Asp Gly Ala Asp Tyr Gln Gly Thr Tyr Gly Ala Ser Thr Ser Gly	
95 100 105	
aac gcc ctg acc ctc aag ttc gtc acc aag ggc agc tac gcc acc aac	1346
Asn Ala Leu Thr Leu Lys Phe Val Thr Lys Gly Ser Tyr Ala Thr Asn	
110 115 120 125	
atc ggc tcg cgc atg tac ctg atg gcc agc ccc acc aag tac gcc atg	1394
Ile Gly Ser Arg Met Tyr Leu Met Ala Ser Pro Thr Lys Tyr Ala Met	
130 135 140	
ttc acc ctg ctg ggc cac gag ttc gcc ttc gac gtc gac ctg agc aag	1442
Phe Thr Leu Leu Gly His Glu Phe Ala Phe Asp Val Asp Leu Ser Lys	
145 150 155	
ctg ccc tgc ggc ctc aac ggc gcc gtc tac ttc gtc agc atg gac gag	1490
Leu Pro Cys Gly Leu Asn Gly Ala Val Tyr Phe Val Ser Met Asp Glu	
160 165 170	
gac ggc ggc acc agc aag tac ccc tcc aac aag gcc ggc gcc aag tac	1538
Asp Gly Gly Thr Ser Lys Tyr Pro Ser Asn Lys Ala Gly Ala Lys Tyr	
175 180 185	
ggc acg ggc tac tgc gac tcg cag tgt ccg cgc gac ctc aag ttt atc	1586
Gly Thr Gly Tyr Cys Asp Ser Gln Cys Pro Arg Asp Leu Lys Phe Ile	
190 195 200 205	
gac ggc aag gtgagaaccc gcactagcgt cccgccttcc gtgtccctcc	1635
Asp Gly Lys	
ttttgccttc ttogaccgcc ctcttccctg cgggccaggg tcgctggggt gctgtcctcc	1695
tttctggtgg gcagcgggtgc tgatcccgcg ccag gcc aac tcg gcc agc tgg cag	1750
Ala Asn Ser Ala Ser Trp Gln	
210 215	
ccc tcg tcc aac gac cag aac gcc ggc gtg ggc ggc atg ggc tcg tgc	1798
Pro Ser Ser Asn Asp Gln Asn Ala Gly Val Gly Gly Met Gly Ser Cys	
220 225 230	
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	
235 240 245	

ES 2 501 944 T3

acg ccg cac ccg tgc cag aac tac cag cag cac agc tgc agc ggc gac Thr Pro His Pro Cys Gln Asn Tyr Gln Gln His Ser Cys Ser Gly Asp 250 255 260	1894
gac tgc ggc ggc acc tac tcg gcc acc cgc ttc gcc ggc gac tgc gac Asp Cys Gly Gly Thr Tyr Ser Ala Thr Arg Phe Ala Gly Asp Cys Asp 265 270 275	1942
ccg gac ggc tgc gac tgg aac gcc tac cgc atg ggc gtg cac gac ttc Pro Asp Gly Cys Asp Trp Asn Ala Tyr Arg Met Gly Val His Asp Phe 280 285 290 295	1990
tac ggc aac ggc aag acc gtc gac acc ggc aag aag ttc tcc atc gtc Tyr Gly Asn Gly Lys Thr Val Asp Thr Gly Lys Lys Phe Ser Ile Val 300 305 310	2038
acc cag ttc aag ggc tcc ggc tcc acc ctg acc gag atc aag cag ttc Thr Gln Phe Lys Gly Ser Gly Ser Thr Leu Thr Glu Ile Lys Gln Phe 315 320 325	2086
tac gtc cag gac ggc agg aag atc gag aac ccc aac gcc acc tgg ccc Tyr Val Gln Asp Gly Arg Lys Ile Glu Asn Pro Asn Ala Thr Trp Pro 330 335 340	2134
ggc ctc gag ccc ttc aac tcc atc acc ccg gac ttc tgc aag gcc cag Gly Leu Glu Pro Phe Asn Ser Ile Thr Pro Asp Phe Cys Lys Ala Gln 345 350 355	2182
aag cag gtc ttc ggc gac ccc gac cgc ttc aac gac atg ggc ggc ttc Lys Gln Val Phe Gly Asp Pro Asp Arg Phe Asn Asp Met Gly Gly Phe 360 365 370 375	2230
acc aac atg gcc aag gcc ctg gcc aac ccc atg gtc ctg gtg ctg tcg Thr Asn Met Ala Lys Ala Leu Ala Asn Pro Met Val Leu Val Leu Ser 380 385 390	2278
ctg tgg gac gac gtgagccatt ttgcattct ctctgactc tcctccgctg Leu Trp Asp Asp 395	2330
ccatcaccac ctctccacc accgccacga ggggtgtagct tgatctccgc tgactgacgt	2390
gtgcccacac ccccgtttct ag cac tac tcc aac atg ctg tgg ctc gac tct His Tyr Ser Asn Met Leu Trp Leu Asp Ser 400 405	2442
acc tac ccg acc gac gcc gat ccc agc gcg ccc ggc aag gga cgt ggc Thr Tyr Pro Thr Asp Ala Asp Pro Ser Ala Pro Gly Lys Gly Arg Gly 410 415 420	2490
acc tgc gac acc agc agc ggc gtg cca agc gac gtg gag tcg aag aat Thr Cys Asp Thr Ser Ser Gly Val Pro Ser Asp Val Glu Ser Lys Asn 425 430 435	2538
gg gtgagtcgga tcttctgcat gcggcccgtt ttccgagcat tgcttgggg Gly	2590
cctccctcag gctgacacac gcgcgccttc gatacag c gat gcg acc gtc atc Asp Ala Thr Val Ile 440	2643

ES 2 501 944 T3

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 ccccgacct 2811

gcagcgtcgg cgggcatcgt atagtctggt gtaactgtat atagtctctgt gcgtgtgaat 2871

cgaacgagca ccgacgaaat gtggtgtttc atgctatcgt acatgctctt gcgagatctg 2931

aagtcgtcaa ttagacattg ccaccatcca acttggcgac tgtccaccg gtccatttgt 2991

atcactggct cttccgagac ccggtctctc tcacaccgta atcactgcaa gcagagttga 3051

attc 3055

<210> 4

<211> 459

5 <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

Arg Met Tyr Leu Met Ala Ser Pro Thr Lys Tyr Ala Met Phe Thr Leu
130 135 140

10

ES 2 501 944 T3

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

ES 2 501 944 T3

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
5 <212> ADN
<213> Acremonium thermophilum

<220>
<221> CDS
10 <222> (891)..(1299)
<223>

<220>
<221> Intrón
15 <222> (1300)..(1387)
<223>

<220>
<221> CDS
20 <222> (1388)..(1442)
<223>

<220>
<221> Intrón
25 <222> (1443)..(1495)
<223>

<220>
<221> CDS
30 <222> (1496)..(1643)
<223>

<220>
<221> Intrón
35 <222> (1644)..(1697)
<223>

<220>
<221> CDS
40 <222> (1698)..(1928)
<223>

<220>
<221> Intrón
45 <222> (1929).. (2014)

<223>

<220>

<221> CDS

5 <222> (2015)..(2740)

<223>

<400> 5

```

ctcgagtttc cctggtcggc cactctctgc tcatctcgct ctgcgcccct ggatgtgccg      60
tgtgtccagt cgtgtatctc ttgactgcac gacgtgttcc tcgcgactcg tctcgcgccg      120
gtggatgccc gtccactcat ttgtccgtct actgggtcag cctctcgtct cgaacgagct      180
tccacggccc actccccgga caacctcggc tctggatggc cctcctcccc ctccgtgtct      240
cccctcctgc ggggtccgtc gtgccctggc tgcattgtcc acatcgcttg atcacgctgc      300
gagccaccgc agagcccat ctccaaagcg accgtggcag cactacctct gtttctggga      360
tggggcccac gtcgatggcc tggcatccct tgccaccctc ctccatcccc ctgacctcac      420
tcccaaccga taggagaagt ggtcatgggc acgaccccg tgcacgtcttg gactcgacga      480
gcttgatcgg gccggaagcc gtcaacgacg ggggagccgt gtcttgccac gcgtggccgt      540
ccttcgacag tggacagcga gaaaactggt ggggaagagg gctgctacag tcttgcttg      600
cgaggcccgga cgctcctagt ccgagaacca cctacgtgtt tctcgcgaag acggggccag      660
cttagcggcc aaatttgccc cccgggccta gggcttagcg atggggatga tgaactgggtg      720
tcgacgatgt ctatataacg acggcgatct cctgtctctg agatcccatc ctttcatctc      780
caaccactt catcccttcc tctctctctc cccctccctt ctctgacata ccgagtcctc      840
agaagcctcg tccgtcgtca cctattctca cttccccgcg aactccggcc atg tat      896
                                   Met Tyr
                                   1

acc aag ttc gcc gcc ctc gcc gcc ctc gtg gcc acc gtc cgc ggc cag      944
Thr Lys Phe Ala Ala Leu Ala Ala Leu Val Ala Thr Val Arg Gly Gln
      5              10              15

gcc gcc tgc tgc ctc acc gcc gag acc cac ccg tgc ctg cag tgg cag      992
Ala Ala Cys Ser Leu Thr Ala Glu Thr His Pro Ser Leu Gln Trp Gln
      20              25              30

aag tgc acc gcg ccc ggc agc tgc acc acc gtc agc ggc cag gtc acc      1040

```

10

ES 2 501 944 T3

Lys Cys Thr Ala Pro Gly Ser Cys Thr Thr Val Ser Gly Gln Val Thr	
35 40 45 50	
atc gac gcc aac tgg cgc tgg ctg cac cag acc aac agc agc acc aac	1088
Ile Asp Ala Asn Trp Arg Trp Leu His Gln Thr Asn Ser Ser Thr Asn	
55 60 65	
tgc tac acc ggc aac gag tgg gac acc agc atc tgc agc tcc gac acc	1136
Cys Tyr Thr Gly Asn Glu Trp Asp Thr Ser Ile Cys Ser Ser Asp Thr	
70 75 80	
gac tgc gcc acc aag tgc tgc ctc gac ggc gcc gac tac acc ggc acc	1184
Asp Cys Ala Thr Lys Cys Cys Leu Asp Gly Ala Asp Tyr Thr Gly Thr	
85 90 95	
tac ggc gtc acc gcc agc ggc aac tcg ctc aac ctc aag ttc gtc acc	1232
Tyr Gly Val Thr Ala Ser Gly Asn Ser Leu Asn Leu Lys Phe Val Thr	
100 105 110	
cag ggg ccc tac tcc aag aac atc ggc tcg cgc atg tac ctc atg gag	1280
Gln Gly Pro Tyr Ser Lys Asn Ile Gly Ser Arg Met Tyr Leu Met Glu	
115 120 125 130	
tcg gag tcc aag tac cag g gtgagcatat agatcacatc tttcgctact	1329
Ser Glu Ser Lys Tyr Gln	
135	
tgcgtccggtt tcgcacggca agcgggtccag acgctaacgg gacgggttctc ttctctag	1387
gc ttc act ctc ctc ggt cag gag ttt acc ttt gac gtg gac gtc tcc	1434
Gly Phe Thr Leu Leu Gly Gln Glu Phe Thr Phe Asp Val Asp Val Ser	
140 145 150	
aac ctc gg gtaggtgatg acttctcccg catgagaaga gctctgctaa	1482
Asn Leu Gly	
ccgtgttgtc cag c tgc ggt ctg aac gga gcg ctc tac ttc gtg tcc atg	1532
Cys Gly Leu Asn Gly Ala Leu Tyr Phe Val Ser Met	
160 165	
gac ctc gac ggc ggc gtg tcc aag tac acc acc aac aag gcc ggc gcc	1580
Asp Leu Asp Gly Gly Val Ser Lys Tyr Thr Thr Asn Lys Ala Gly Ala	
170 175 180	
aag tac ggc acc ggc tac tgc gac tcc cag tgc ccg cgg gat ctc aag	1628
Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys Pro Arg Asp Leu Lys	
185 190 195	
ttc atc aac ggc cag gtgggtcgag agaccctctt cccctctcag tgaacgatgt	1683
Phe Ile Asn Gly Gln	
200	
ctgaccctct ctag gcc 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 gcc 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	

ES 2 501 944 T3

atc tgg gag gcc aac aag gtc tcc gcc gcc tac acg ccg cac ccc tgc Ile Trp Glu Ala Asn Lys Val Ser Ala Ala Tyr Thr Pro His Pro Cys 235 240 245	1829
acc acc atc ggc cag acc atg tgc acc ggc gac gac tgc ggc ggc acc Thr Thr Ile Gly Gln Thr Met Cys Thr Gly Asp Asp Cys Gly Gly Thr 250 255 260	1877
tat tcg tcg gac cgc tat gcc ggc atc tgc gac ccc gac ggt tgc gat Tyr Ser Ser Asp Arg Tyr Ala Gly Ile Cys Asp Pro Asp Gly Cys Asp 265 270 275 280	1925
ttt gtaggttctt tctctcgccg ctccctgacg acctatatgt gtgaagggac Phe	1978
gcacagaaaa gacaagggtca aagctgacca gagcag aac tcg tac cgc atg ggc Asn Ser Tyr Arg Met Gly 285	2032
gac acc agc ttc tac ggc ccc ggc aag acg gtc gac acc ggc tcc aag Asp Thr Ser Phe Tyr Gly Pro Gly Lys Thr Val Asp Thr Gly Ser Lys 290 295 300	2080
ttc acc gtc gtg acc cag ttc ctc acg ggc tcc gac ggc aac ctc agc Phe Thr Val Val Thr Gln Phe Leu Thr Gly Ser Asp Gly Asn Leu Ser 305 310 315	2128
gag atc aag cgc ttc tac gtg cag aac ggc aag gtc atc ccc aac tcc Glu Ile Lys Arg Phe Tyr Val Gln Asn Gly Lys Val Ile Pro Asn Ser 320 325 330 335	2176
gag tcc aag atc gcc ggc gtc tcc ggc aac tcc atc acc acc gac ttc Glu Ser Lys Ile Ala Gly Val Ser Gly Asn Ser Ile Thr Thr Asp Phe 340 345 350	2224
tgc acc gcc cag aag acc gcc ttc ggc gac acc aac gtc ttc gag gag Cys Thr Ala Gln Lys Thr Ala Phe Gly Asp Thr Asn Val Phe Glu Glu 355 360 365	2272
cgc ggc ggc ctc gcc cag atg ggc aag gcc ctg gcc gag ccc atg gtc Arg Gly Gly Leu Ala Gln Met Gly Lys Ala Leu Ala Glu Pro Met Val 370 375 380	2320
ctg gtc ctg tcc gtc tgg gac gac cac gcc gtc aac atg ctc tgg ctc Leu Val Leu Ser Val Trp Asp Asp His Ala Val Asn Met Leu Trp Leu 385 390 395	2368
gac tcc acc tac ccc acc gac agc acc aag ccc ggc gcc gcc cgc ggc Asp Ser Thr Tyr Pro Thr Asp Ser Thr Lys Pro Gly Ala Ala Arg Gly 400 405 410 415	2416
gac tgc ccc atc acc tcc ggc gtg ccc gcc gac gtc gag tcc cag gcg Asp Cys Pro Ile Thr Ser Gly Val Pro Ala Asp Val Glu Ser Gln Ala 420 425 430	2464
ccc aac tcc aac gtc atc tac tcc aac atc cgc ttc ggc ccc atc aac Pro Asn Ser Asn Val Ile Tyr Ser Asn Ile Arg Phe Gly Pro Ile Asn 435 440 445	2512
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 tcg 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 tcg cag tgc ctg taagccataa 2750
Cys Lys Tyr Ser Asn Asp Trp Tyr Ser Gln Cys Leu
515 520

gccccctgta cggttcggaag acggtggcaa cagacaaacc cctccccga gcacaccccc 2810

cagggatcta aggggggttgt ggttaagaca taagaatgcg ccgtggcttg gcctacgcca 2870

cggtcatgaa agtgcagtga aaatgggggc aagagtcgga aaaagtgagt ttgcttgcaa 2930

gggagagagg atgtcgagag gtgatgactt cgtttgtaca tagttggctc ttcgtgattg 2990

ggaacgggag gagtgtcggg gggagccctc cagactcctt ggcctctccg ctcgttccat 3050

ctttctcagt acatatacat ctgcattttc atccacgtct ctggcgtctc tggatgtgaa 3110

cgaatccgac aactggtggg 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

5 <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

Trp Gln Lys Cys Thr Ala Pro Gly Ser Cys Thr Thr Val Ser Gly Gln
35 40 45

ES 2 501 944 T3

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

ES 2 501 944 T3

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		

ES 2 501 944 T3

<210> 7
 <211> 3649
 <212> ADN
 <213> Chaetomium thermophilum

5

<220>
 <221> CDS
 <222> (1290)..(2879)
 <223>

10

<220>
 <221> Intrón
 <222> (2880)..(2943)
 <223>

15

<220>
 <221> CDS
 <222> (2944)..(2949)
 <223>

20

<400> 7
 tctagagctg tgcacgcggc cgcgtaatac gactcactat agggcgaaga attcggatcg 60
 gactagagct cgtcacgggc tcgcgccgac gaggcgatga ggacgaaggc ccgacataat 120
 ccgtacttta cgctacatga cgactctcga aaattgtaaa gggccggcat ttcggagcga 180
 gtgctgcgag ggcgcattcg cggcgtacct ggaattcctg gaatggtaag caatggccag 240
 caatgggcca ggtatggacc agcttgaatc ctgggtgcgg cgtcaccagg ccagcatgg 300
 tgcccagaat ggcccaccgt ggcccatcgt cctaagaaac aagctgcgtc ccgcgatcca 360
 aaaacgtcgt cttcggcgca cgtcctccgt ggtcccccg gctggacacc ctggctggcc 420
 ctccaatgag cggcatttgc ccctgtcgag cgtgtcggca accttaatcg actccatctc 480
 tcggctccac gccgtccatc ctgtcctcga cctcgtcatc tgtgctcccc ttgccctccc 540
 ttgcccttcc ttgcctccgc caccgacgtc cacaatgtga ccctgctgcc cggagcgccc 600
 agcgccatgc accgtttggg cttgtcgccg tgtcgccagt ctccatcgag cgattcgacc 660
 gtgtgcctct ctccaccagc gttccccgcg ctctccatag tccatgctac ttcgagccgt 720
 tgccctcaca gctgccagcg gcatggctct gtcggtctcg cctctccttt tcccgtgaag 780
 cgctgccata caattctccg tctgccccag tccttgaggc gccgctattc ccaatcggcc 840
 atggcactgg ccagcccgat ccatgttcga tcgagcttcg acgggccgtg agccgtctgc 900
 acggaggagc ttgcgagcct gcgaacctgg cggacctgga gaagcctggc ccctctccct 960
 ggatggagat actgggtgcg ctagcaccac ggcgtgccac ggccaagctc cggccgaccc 1020
 ggaggcggga agagggttgc gttgctgtct tcggcggctg tcagggcaaa gggtaatcgt 1080
 caatgtggga aaaggggctc atctccatga gattcatgac tcggacatcg tctatataag 1140
 tcgagtcccc cactctccaa cagccgattc tgcctcctcat cccatcacca ccctcgtcca 1200

ES 2 501 944 T3

caaccacgca gttgtgtaca tcaaaacaag ttcgctcctt ttacatcttc accacaacaa	1260
cagcacatcc tctccttttcg gctttcaag atg atg tat aag aag ttc gcc gct	1313
Met Met Tyr Lys Lys Phe Ala Ala	
1 5	
ctc gcc gcc ctc gtg gct ggc gcc tcc gcc cag cag gct tgc tcc ctc	1361
Leu Ala Ala Leu Val Ala Gly Ala Ser Ala Gln Gln Ala Cys Ser Leu	
10 15 20	
acc gct gag aac cac cct agc ctc acc tgg aag cgc tgc acc tct ggc	1409
Thr Ala Glu Asn His Pro Ser Leu Thr Trp Lys Arg Cys Thr Ser Gly	
25 30 35 40	
ggc agc tgc tcg acc gtg aac ggc gcc gtc acc atc gat gcc aac tgg	1457
Gly Ser Cys Ser Thr Val Asn Gly Ala Val Thr Ile Asp Ala Asn Trp	
45 50 55	
cgc tgg act cac acc gtc tcc ggc tcg acc aac tgc tac acc ggc aac	1505
Arg Trp Thr His Thr Val Ser Gly Ser Thr Asn Cys Tyr Thr Gly Asn	
60 65 70	
cag tgg gat acc tcc ctc tgc act gat ggc aag agc tgc gcc cag acc	1553
Gln Trp Asp Thr Ser Leu Cys Thr Asp Gly Lys Ser Cys Ala Gln Thr	
75 80 85	
tgc tgc gtc gat ggc gct gac tac tct tcg acc tat ggt atc acc acc	1601
Cys Cys Val Asp Gly Ala Asp Tyr Ser Ser Thr Tyr Gly Ile Thr Thr	
90 95 100	
agc ggt gac tcc ctg aac ctc aag ttc gtc acc aag cac cag tac ggc	1649
Ser Gly Asp Ser Leu Asn Leu Lys Phe Val Thr Lys His Gln Tyr Gly	
105 110 115 120	
acc aac gtc ggc tcc cgt gtc tat ctg atg gag aac gac acc aag tac	1697
Thr Asn Val Gly Ser Arg Val Tyr Leu Met Glu Asn Asp Thr Lys Tyr	
125 130 135	
cag atg ttc gag ctc ctc ggc aac gag ttc acc ttc gat gtc gat gtc	1745
Gln Met Phe Glu Leu Leu Gly Asn Glu Phe Thr Phe Asp Val Asp Val	
140 145 150	
tcc aac ctg ggc tgc ggt ctc aac ggc gcc ctc tac ttc gtt tcc atg	1793
Ser Asn Leu Gly Cys Gly Leu Asn Gly Ala Leu Tyr Phe Val Ser Met	
155 160 165	
gat gct gat ggt ggc atg agc aaa tac tct ggc aac aag gct ggc gcc	1841
Asp Ala Asp Gly Gly Met Ser Lys Tyr Ser Gly Asn Lys Ala Gly Ala	
170 175 180	
aag tac ggt acc ggc tac tgc gat gct cag tgc ccg cgc gac ctc aag	1889
Lys Tyr Gly Thr Gly Tyr Cys Asp Ala Gln Cys Pro Arg Asp Leu Lys	
185 190 195 200	
ttc atc aac ggc gag gcc aac gtt ggg aac tgg acc ccc tcg acc aac	1937
Phe Ile Asn Gly Glu Ala Asn Val Gly Asn Trp Thr Pro Ser Thr Asn	
205 210 215	
gat gcc aac gcc ggc ttc ggc cgc tat ggc agc tgc tgc tct gag atg	1985
Asp Ala Asn Ala Gly Phe Gly Arg Tyr Gly Ser Cys Cys Ser Glu Met	

ES 2 501 944 T3

220	225	230	
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 235 240 245			2033
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 250 255 260			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 265 270 275 280			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 285 290 295			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 300 305 310			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 315 320 325			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 330 335 340			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 345 350 355 360			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 365 370 375			2417
aag gcc ctc gca ggc ccc atg gtc ctg gtc atg tcc gtc tgg gat gac Lys Ala Leu Ala Gly Pro Met Val Leu Val Met Ser Val Trp Asp Asp 380 385 390			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 395 400 405			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 410 415 420			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 425 430 435 440			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 445 450 455			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 Thr Val Val Pro Pro Ala Ser 460 465 470			2705

ES 2 501 944 T3

acc tcc acc tcc cgt ccg acc agc agc act agc tct ccc gtt tgc acc 2753
 Thr Ser Thr Ser Arg Pro Thr Ser Ser Thr Ser Ser Pro Val Ser Thr
 475 480 485

ccg act ggc cag ccc ggc ggc tgc acc acc cag aag tgg ggc cag tgc 2801
 Pro Thr Gly Gln Pro Gly Gly Cys Thr Thr Gln Lys Trp Gly Gln Cys
 490 495 500

ggc ggt atc ggc tac acc ggc tgc act aac tgc gtt gct ggc acc acc 2849
 Gly Gly Ile Gly Tyr Thr Gly Cys Thr Asn Cys Val Ala Gly Thr Thr
 505 510 515 520

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 ttctgtacata aatttacgca 3129

ttgagctttt cccccccaa aaacagttcc ctgatttgcg ggagtaactt gatggtaaag 3189

cttggtcata agctcttcaa tggaaaaaac gatacagtca tgccttgaca catcctccca 3249

aagtcttcgt ccatgacatc acggtcgcgc cttaagcaca agttcaataa ccccatgtgg 3309

cgttgcccttg tcttgaaaca cagatgagat cttcagccca gccgcacgcg ccacttcctt 3369

gaactgagcc aacgagcgtt ccttcccgcc gattgagagc atcgcatagt ccttgaaggc 3429

tgcatagaga ggaatagggg gcttggttcc ggtagttggg ctgccggaac tcggatctgt 3489

tggcgcaagg gggtcagggt tgatctgctc ggcatgagc acgcgtccat cgggggttgt 3549

tagtgcacga gcgacattgc gcaggatggt gactgccaca gggtcggagt aatcgcgag 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

ES 2 501 944 T3

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
 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

ES 2 501 944 T3

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	
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

5 <210> 9
<211> 1339
<212> ADN
<213> Thermoascus aurantiacus

10 <220>
<221> CDS
<222> (17)..(122)
<223>

15 <220>
<221> Intrón
<222> (123)..(177)
<223>

20 <220>
<221> CDS
<222> (178)..(236)
<223>

25 <220>
<221> Intrón
<222> (237)..(296)
<223>

30 <220>
<221> CDS
<222> (297)..(449)
<223>

35 <220>
<221> Intrón
<222> (450)..(508)
<223>

40 <220>
<221> CDS
<222> (509)... (73)
<223>

45 <220>
<221> Intrón
<222> (574)..(647)
<223>

50 <220>
<221> CDS
<222> (648)..(745)
<223>

55 <220>
<221> Intrón
<222> (746)..(806)
<223>

60 <220>
<221> CDS
<222> (807)..(1330)
<223>

<400> 9

ES 2 501 944 T3

ccgcgggaactg cgcatac 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 tgc 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 gttagtaaca 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 attccaccat gtttgagact gtcttcggtg tgctgacatt taatggtag	508
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	

ES 2 501 944 T3

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

ES 2 501 944 T3

<400> 10

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			

ES 2 501 944 T3

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

5 <212> ADN

<213> Acremonium thermophilum

<220>

<221> característica miscelánea

10 <222> (13)..(13)

<223> N = no identificada

<220>

<221> CDS

15 <222> (715)..(797)

<223>

<220>

<221> Intrón

20 <222> (798)..(856)

<223>

<220>

<221> CDS

25 <222> (857)..(1105)

<223>

<220>

<221> Intrón

30 <222> (1106)..(1228)

<223>

<220>

<221> CDS

ES 2 501 944 T3

<222> (1229)..(1787)

<223>

<400> 11

tctgtctctt gntcagaac agatctcctg gcggcctgct ttgccgggtcc 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 gcccatccca tccggcaaca cgcatctgct ttctgttccg gcattccgat	360
acctgggttc tggagtgcct accgagcctc gcttcctggg atcgggaggt gcaccccgcc	420
aaaccctatg ccccaaacgg 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 cccatcgta gcatttcact 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
	Tyr Trp Asp Cys Cys
	30
aag ccg tcc tgc ggc tgg ccg gga aag gcc tcg 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 tcg tgc tcg 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	
tcg ggc tgc gac gga ggc tcc gcc tac tcg tgc gcc gac cag acg ccc	1016
Ser Gly Cys Asp Gly Gly Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro	

5

ES 2 501 944 T3

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 ttcccacccc cgctttctgt gcaggccgct tccccctac			1165
ccacccactt cccccccccc gcctctgtga tcgggcatcc gagctaagtt gcgtgtcgtc			1225
cag a ctc acc ttc aac tgc ggc ccc gtc gcg ggc aag acc atg gtg 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 gcc 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			
175	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 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			
255	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		

```

gtccggtttcc ctagggtctga ggctgacgtg aactgggtcc tcttgccgc cccatcacgg 1867
gttcgtattc gcgcgcttag ggagaggagg atgcagtttg agggggccac attttgaggg 1927
ggacgcagtc tggggtcgaa gcttgctcgt tagggctgcc gtgacgtggt agagcagatg 1987
ggaccaagtg cggagctagg caggtgggtg gttgtggtgg tggcttacct tctgtaacgc 2047
aatggcatct catctcactc gcctgctccc tgattggtgg ctctgttcgg cctggcgctt 2107
tttgggaccg ctggctggaa tggattgctc cggaacgcca ggttgagctg ggctggcgcg 2167
agtagattgg ccgctccgag ctgcaaccat aataaaattt tcggaccctg taagccgcac 2227
ccgaccaggt ctccattggc ggacatgcac gacgtccttc gcaggcacgg cctgcccgcc 2287
tctgatcacc cgcagttttc gtaccgtcag accagataca agccccg 2334

```

<210> 12

<211> 297

5 <212> PRT

<213> Acremonium thermophilum

<220>

<221> característica miscelánea

10 <222> (13)..(13)

<223> N = no identificada

<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

```

15

ES 2 501 944 T3

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
 <212> ADN
 <213> Acremonium thermophilum

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

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

<223>

<220>

<221> CDS

5 <222> (858)..(888)

<223>

<220>

<221> Intrón

10 <222> (889)..(990)

<223>

<220>

<221> CDS

15 <222> (991)..(1268)

<223>

<400> 13

ctcgaggaga ggaaccgagt ttgaaagatg ctatatatcg atagactacc ggcgtcgcct 60

cgccctgtcc gctctcttgc attccccctg ttgatgagac gagacaaaat tcctggtag 120

aaaagatccg tcgccgagat ttcaccagtg gtaagtcccg agaattggtc attcgacgtt 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

ES 2 501 944 T3

```

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 gtt ggt cag tcc aat g      888
      Ile Pro Gly Gly Gly Val Gly Gln Ser Asn
                      150                      155

gtaggttcct tccctgaagt 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
195                      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 ccaggacat ttgtcgcgac      1348

ctgatcatag gcgccagggg agttgaaagg ggttgccgta cgagaagaca ttttgcgcc      1408

gtcttactcc cagccacttc tgtacatatt caatgacatt acatagcccg caaatatggt      1468

catatatcgt ggccgccc aa accgccccgg tttgcttagg ctggagctga agtggctcgc      1528

cgatggctgt caaaggcagt cggaatattc ctcggttgctt cggcaacacg gtagctgctt      1588

gaaccgtacc cagcattaga acaccccccg ccgaggggctt gctacgtcaa tggcgggggtc      1648

tccaaccctt gcgcggcaca aaaccaacca cgccctcgtc ttttatgatg tcctcgtcga      1708

```

aacgtcccgt gaagacactc cgctcatggt ctggtcctct gatgtagaag gggtaggtca 1768
gccgatggtc gtcaccgtcg tcaatgcttc cctcaagctt cttgcggcct ttatcctcca 1828
actcttccca catgagaact ccattctttcc gccttttcac aaagccactg cctccttgt 1888
caagggccaa aaaccaacgc cgctgatgaa tgcttccgat cgtgtttgac gcgcccgggg 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

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

5

10

	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
 <213> Chaetomium thermophilum

<220>
 <221> característica miscelánea
 <222> (2786)..(2786)
 <223> N = no identificada

<220>
 <221> CDS
 <222> (768)..(2042)
 <223>

<400> 15
 ggatccaaga ccatcccca ggattctcgg attatgtttg catctcacc tccgaaaccg 60
 catgaaaaat tgaaatgggc aactgtcgct gtgtttaatg ctttgcacat catgggatca 120
 tgttcacccg ctctaattct tcatactcca gatcctatct atcctccgca tctagccggc 180
 ttcttgcttg tgatccaaag cctgatccc acgcggcttc tagacgcttt agaaattaca 240
 ccgaatctcc ccatgccctt cttgcaatat cttcccgacc aggaacttcg ggtgctcaac 300
 atccgcgagc ttgaagacga ccttctttgg ccggcttgge atgcgactct gttcgggact 360
 caatgcaact ctggggccctt caatgccgag catgaccgtt actgaggctt agccgcccc 420
 atcgcttggc acggtacctt gcagacggaa tcccgggccc gttgtccgat ctgctttgg 480
 tccggtagag aagcctcgga ggaagagaca cacggacaca acgattgcgg gccccaatgc 540
 gctgctccta attgaggctc cgaggctcgtg tgccgtgtgg agaggccgag actgggtctg 600

ES 2 501 944 T3

gggtgcggag gattgcggag atgaagataa tctgggtgca accgtggata cataaaagg	660
agtagttctc ccctctgtga aaccttcttc cccaggattc tcctcgcctc taagagtcca	720
aagtcattca agacatccta cagcgggggtc 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 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 Tyr Tyr Gly Thr Gly Tyr Cys Asp Ala Gln Cys Phe Val Thr Pro Phe 180 185 190 195	1352
atc aat ggc atc ggc aac atc gag ggc aag ggc tcg tgc tgc aac gag Ile Asn Gly Ile Gly Asn Ile Glu Gly Lys Gly Ser Cys Cys Asn Glu 200 205 210	1400
atg gac att tgg gag gcc aac tcg cgt agt cag tcc att gct ccg cac	1448

ES 2 501 944 T3

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

```

cagtgggtcat gggtaaataa tctacaggct gtgaatggcg ttgcgtcagc ctcattaact 2312
taaacgattg gactccccctt ttcctaataca tcgccgttgc cgtgtaactc tcctagatct 2372
cttgtttgtat atggcttcaa ctcgaagtga agaaaaatgg atacggcgac ctctttgtgc 2432
caattttctt gctgttcttc cggtattgac cctcggcaag acaactatgg ccaatattct 2492
gttatagtcg gcagttagtg ttgtgtcgta caagtcgtgc gggagcaata ctcaacagcc 2552
gcccttaata tggttattta cgccacgacg cacttcatta cacggctttg gggggtatat 2612
attccgttca actctatccc tcattcggtg tgattgaacg tctccaacag tgaaagtata 2672
agtctgacaa aaatgcccaa ccgccatgcc actgatgatc ctgttgagat gctcgtggtc 2732
tataacatcc tgtctaagtg ttacctccct aatggttagcc ccagttctgc tctncttgtc 2792
tcgacagc 2800

```

<210> 16
 <211> 425
 <212> PRT
 <213> Chaetomium thermophilum

<220>
 <221> característica miscelánea
 <222> (2786)..(2786)
 <223> N = no identificada

<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 2 501 944 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

ES 2 501 944 T3

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
5 <212> ADN
<213> Thermoascus aurantiacus

<220>
10 <221> CDS
<222> (13)..(256)
<223>

<220>
15 <221> Intrón
<222> (257)..(329)
<223>

<220>
20 <221> CDS
<222> (330)..(370)
<223>

<220>
25 <221> Intrón
<222> (371)..(444)
<223>

<220>
30 <221> CDS
<222> (445)..(493)
<223>

<220>
35 <221> Intrón
<222> (494)..(561)
<223>

<220>
40 <221> CDS
<222> (562)..(683)
<223>

<220>
45 <221> Intrón
<222> (684)..(786)

<223>
 <220>
 <221> CDS
 5 <222> (787)..(932)
 <223>
 <220>
 <221> Intrón
 10 <222> (933)..(1001)
 <223>
 <220>
 <221> CDS
 15 <222> (1002)..(1090)
 <223>
 <220>
 <221> Intrón
 20 <222> (1091)..(1155)
 <223>
 <220>
 <221> CDS
 25 <222> (1156)..(1174)
 <223>
 <220>
 <221> Intrón
 30 <222> (1175)..(1267)
 <223>
 <220>
 <221> CDS
 35 <222> (1268)..(1295)
 <223>
 <220>
 <221> Intrón
 40 <222> (1296)..(1361)
 <223>
 <220>
 <221> CDS
 45 <222> (1362)..(1451)
 <223>
 <220>
 <221> Intrón
 50 <222> (1452)..(1551)
 <223>
 <220>
 <221> CDS
 55 <222> (1552)..(1617)
 <223>
 <220>
 <221> Intrón
 60 <222> (1618)..(1829)
 <223>
 <220>
 <221> CDS
 65 <222> (1830)..(1922)
 <223>

ES 2 501 944 T3

<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	

ES 2 501 944 T3

gcacgcaact gcgattgagg agaagataact aatcttcttg aaag gtc aat tgg gcc Val Asn Trp Ala	456
cag caa aat gga aag ctg atc cgt ggc cat act ctt g gttagtagaa Gln Gln Asn Gly Lys Leu Ile Arg Gly His Thr Leu 100 105 110	503
cgccaacctg cttccctaac ttactgaaga aggaaaaccg aattgaccgt cccccaag	561
ta tgg cac tcg cag ctg ccc tcg tgg gtg agc tcc atc acc gac aag Val Trp His Ser Gln Leu Pro Ser Trp Val Ser Ser Ile Thr Asp Lys 115 120 125	608
aat acg ctg acc aac gtg atg aaa aat cac atc acc acc ttg atg acc Asn Thr Leu Thr Asn Val Met Lys Asn His Ile Thr Thr Leu Met Thr 130 135 140	656
cgg tac aag ggc aag atc cgt gca tgg gtcagtcac ctaccctaag Arg Tyr Lys Gly Lys Ile Arg Ala Trp 145 150	703
ctgcggtttca atgaagagac aaataagaac acacgtatctt gcccgggcgt ttcagaatca	763
gaactgacag aatcactgaa tag gac gtg gtg aac gag gca ttc aac gag gat Asp Val Val Asn Glu Ala Phe Asn Glu Asp 155 160	816
ggc tcc ctc cgc cag act gtc ttc ctc aac gtc atc ggg gag gat tac Gly Ser Leu Arg Gln Thr Val Phe Leu Asn Val Ile Gly Glu Asp Tyr 165 170 175	864
atc ccg att gct ttc cag acc gcc cgc gcc gct gac ccg aat gcc aag Ile Pro Ile Ala Phe Gln Thr Ala Arg Ala Ala Asp Pro Asn Ala Lys 180 185 190	912
ctg tac atc aac gat tac aa gtaagattta aggctcagt atattccatt Leu Tyr Ile Asn Asp Tyr Asn 195 200	962
tagtgtgaga agcattgctt atgagcatct gtattacag c ctc gac agt gcc tcg Leu Asp Ser Ala Ser 205	1017
tac ccc aag acg cag gcc att gtc aac cgc gtc aag caa tgg cgt gca Tyr Pro Lys Thr Gln Ala Ile Val Asn Arg Val Lys Gln Trp Arg Ala 210 215 220	1065
gct gga gtc ccg att gac ggc ata g gtatgtctct ctttctgttt Ala Gly Val Pro Ile Asp Gly Ile 225 230	1110
gtgatgtgac cgatttgaaa ccagtctaac gttagctggg tctag ga tcg caa acg Gly Ser Gln Thr	1166
cac ctc ag gtaaataatc gggaatgcct cggagaataa aagagaaaaa His Leu Ser 235	1214
aaatgattgt cttatcagat cgtatcgact gactcatggc ttgtccaaaa tag c gct	1271

Ala

ggt cag gga gcc ggt gtt cta caa taagtgcgcc 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
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 cgttgtccct atcccccttg 1481
Ala Ser Pro Thr Asp Tyr Val Asn
270 275

gatactttgt ataattatta tcttcccggga 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 cgatcatgtct tagtcgtctt cagtcaggcc 1697

aaggccaaga cacaggacct gaaacggggca ggcagcagct gctagcagcc 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

5 <212> PRT

<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

10

ES 2 501 944 T3

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

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

ES 2 501 944 T3

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
<212> ADN .
<213> Acremonium thermophilum

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

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

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

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

<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 gcccggtggct 180
gaattttcct tcccgttctc tacagtcctt ccctcagcga cacatccgca gttttgacag 240

cggaatcgt caggatgctc cgccttctct cgcaacctga gtgccaggc gtctcgcca	300
ccgtctctta tataatggccg ctgggtccgc ctttcgatcg gttttcgatt tggctctctcc	360
tagttccctc agctgacccg ggatatcgct tgtggctccg aaacctcacc atcccagacg	420
agcaagttct ccgcagtcca cctcagctca tccggccctt ggtagcatcg cagcgacccc	480
agacgaaggc accaaagaag catactatat attaggctaa atcgagcccc acgtggaata	540
tttgccatcg aggaggggtg gttgggcttc ttgtcctcgc aggtgctgcg cctgtacctc	600
cctgggtgctc cagctggtgc tcccgtggt gctgttcag tcgccgtctg gcccgaatgc	660
tctgtatctc ggttcgtccc gcaactcctt cgccaagcgc taccaatgct ttgacgaacc	720
cggtaaattt gcagtggacc tgcagctggg caaacccgca gtgggaacca cagacctggt	780
tcgttcgaca cactccaatc gcaacccgc ccgcgcaaac cttgcaccac atgtcgcccc	840
tttcccagtt gggtcctga agacacggag ccacttccgt gatcgtcggc tccccagcc	900
gacagtcgga cgctgcaata ggatgccagc acccgtggat ccaagggcca gtgaccccaa	960
ctctttcgcg gtattctggc cctcccaaag gtatgccagg acttcctgt ctttctacc	1020
accagctctc ctccacggcg gaacggatac gccgtctcgc cggtctctgc tcgacaacat	1080
gcgagggggc gcgaaggcta ggttgtagcg atgcgacggt gcgatgtcac catttggcag	1140
tgatgttttc cgttgtcccc ttctccacc tcgcgcgttt cctcaaagac gcccacaacca	1200
taaatacgat gcgacgcaa cttcatgtg ttcgtggcat cttgcctgac cagtctcagc	1260
aagaaacctg tggcggcgcg attgtcttga cttctgatt gaaaacggat ctgcgtctc	1320
ctcgatagcc gacc atg cgc gcc aag caa ctc ctg gcg gcc ggc ctg ctg	1370
Met Arg Ala Lys Gln Leu Leu Ala Ala Gly Leu Leu	
1 5 10	
gcc ccc gcg tcc gtc tcg gcc cag ctc aac agc ctc gcc gtg gcg gct	1418
Ala Pro Ala Ser Val Ser Ala Gln Leu Asn Ser Leu Ala Val Ala Ala	
15 20 25	
ggc ctc aag tac ttc ggc acg gcc gtg cgg gag gcc aac gtc aac ggc	1466
Gly Leu Lys Tyr Phe Gly Thr Ala Val Arg Glu Ala Asn Val Asn Gly	
30 35 40	
gac gcc acc tac atg tcg tac gtc aac aac aag tcc gag ttc ggc cag	1514
Asp Ala Thr Tyr Met Ser Tyr Val Asn Asn Lys Ser Glu Phe Gly Gln	
45 50 55 60	
gtg acg ccc gag aac ggc cag aag tgg gat tcc acc gag ccc agc cag	1562
Val Thr Pro Glu Asn Gly Gln Lys Trp Asp Ser Thr Glu Pro Ser Gln	
65 70 75	
ggc cag ttc agc tac agc cag ggc gac atc gtc ccc ggc gtc gcg aag	1610
Gly Gln Phe Ser Tyr Ser Gln Gly Asp Ile Val Pro Gly Val Ala Lys	
80 85 90	
aag aac ggc cag gtg ctg cgc tgc cac acc ctg gtg tgg tac agc cag	1658

ES 2 501 944 T3

Lys	Asn	Gly	Gln	Val	Leu	Arg	Cys	His	Thr	Leu	Val	Trp	Tyr	Ser	Gln			
95					100					105								
ctc	ccc	agc	tgg	g	gtc	agt	gact	ctct	ctttt	ct	ctgt	cttt	ctct	ttgt	ct	1711		
Leu	Pro	Ser	Trp															
110																		
ttct	ctctt	ctct	ctct	ctct	ctct	ctct	ctct	ctct	ctct	ctct	ctcc	ctcc	agc	atc	g	1771		
act	gct	gat	ttg	ctg	acca	gaag	ctc	gtg	tgc	ag	tca	tcc	gga	agt	tgg	1823		
										Val	Ser	Ser	Gly	Ser	Trp			
115																		
acc	cgc	gcg	acg	ctt	cag	tcc	gtc	atc	gag	acg	cac	atc	tcg	aac	gtg	1871		
Thr	Arg	Ala	Thr	Leu	Gln	Ser	Val	Ile	Glu	Thr	His	Ile	Ser	Asn	Val			
120					125					130								
atg	ggc	cac	tac	aag	ggc	cag	tgc	tac	gcc	tgg	gac	gtg	gtc	aac	gag	1919		
Met	Gly	His	Tyr	Lys	Gly	Gln	Cys	Tyr	Ala	Trp	Asp	Val	Val	Asn	Glu			
135					140					145					150			
gcc	atc	aac	gac	gac	ggc	acg	tgg	cgg	acc	agc	gtc	ttc	tac	aac	acc	1967		
Ala	Ile	Asn	Asp	Asp	Gly	Thr	Trp	Arg	Thr	Ser	Val	Phe	Tyr	Asn	Thr			
155					160					165								
ttc	aac	acc	gac	tac	ctg	gcc	att	gcc	ttc	aac	gcc	gcg	aag	aag	gcc	2015		
Phe	Asn	Thr	Asp	Tyr	Leu	Ala	Ile	Ala	Phe	Asn	Ala	Ala	Lys	Lys	Ala			
170					175					180								
gat	gcg	ggc	gcg	aag	ct	gtag	gtg	tcg	gcct	ttac	gt	gcgc	gcgc			2062		
Asp	Ala	Gly	Ala	Lys	Leu													
185																		
cac	ctcc	gcg	acat	gag	ccc	cag	agc	gcgt	ggct	aat	agt	tcct	cac	gca	cg	2118		
tac	tac	aac	gac	tac	aat	ctc	gag	tac	aac	ggc	gcc	aag	acc	aac	acg	2166		
Tyr	Tyr	Asn	Asp	Tyr	Asn	Leu	Glu	Tyr	Asn	Gly	Ala	Lys	Thr	Asn	Thr			
190					195					200								
gcc	gtg	cag	ctg	gtg	cag	atc	gtg	cag	cag	gcc	ggc	gcg	ccc	atc	gac	2214		
Ala	Val	Gln	Leu	Val	Gln	Ile	Val	Gln	Gln	Ala	Gly	Ala	Pro	Ile	Asp			
205					210					215					220			
ggg	gtg	ggc	ttc	cag	ggc	cac	ctg	atc	gtg	ggg	tca	acg	ccg	tcg	cgc	2262		
Gly	Val	Gly	Phe	Gln	Gly	His	Leu	Ile	Val	Gly	Ser	Thr	Pro	Ser	Arg			
225					230					235								
agc	tcc	ctg	gcc	acg	gcg	ctg	aag	cgc	ttc	acg	gcg	ctt	ggc	ctg	gag	2310		
Ser	Ser	Leu	Ala	Thr	Ala	Leu	Lys	Arg	Phe	Thr	Ala	Leu	Gly	Leu	Glu			
240					245					250								
gtg	gcg	tac	acg	gag	ctg	gac	atc	cgg	cac	tcg	agc	ctg	ccg	ccg	tcg	2358		
Val	Ala	Tyr	Thr	Glu	Leu	Asp	Ile	Arg	His	Ser	Ser	Leu	Pro	Pro	Ser			
255					260					265								
tcg	gcg	gcg	ctg	gcg	acg	cag	ggc	aac	gac	ttc	gcc	agc	gtg	gtg	ggc	2406		
Ser	Ala	Ala	Leu	Ala	Thr	Gln	Gly	Asn	Asp	Phe	Ala	Ser	Val	Val	Gly			
270					275					280								
tcg	tgc	ctc	gac	gtg	gcg	ggc	tgc	gtg										

ES 2 501 944 T3

285	290	295	300	
acg gac aag tac	agc tgg gtg ccc gac	acg ttc ccc ggc tcg	ggc ggc ggc	2502
Thr Asp Lys Tyr	Ser Trp Val Pro Asp	Thr Phe Pro Gly	Ser Gly Ala	
	305	310	315	
gcg ctg ctg tac	gac gcg aac tac	agc aag aag ccg	gcg tgg acg tcg	2550
Ala Leu Leu Tyr	Asp Ala Asn Tyr	Ser Lys Lys Pro	Ala Trp Thr Ser	
	320	325	330	
gtc tcg tcg gtg	ctg gcg gcc aag	gcg acg aac ccg	ccc ggc ggc ggg	2598
Val Ser Ser Val	Leu Ala Ala Lys	Ala Thr Asn Pro	Pro Gly Gly Gly	
	335	340	345	
aac cca ccc ccc	gtc acc acc acg	acc acc acg acc	acg tcg aag	2646
Asn Pro Pro Pro	Val Thr Thr Thr	Thr Thr Thr Thr	Thr Thr Ser Lys	
	350	355	360	
ccg tcg cag ccc	acc acc acg acc	acc acc acc acg	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 ggg atc ggc	tgg acg ggg 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
ttttgacett ttctccgcag	acgtatatta tcaattatag	ttggatttct aga		2955
<210> 20				
<211> 416				
<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	

ES 2 501 944 T3

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

ES 2 501 944 T3

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
5 <212> ADN
<213> Thermoascus aurantiacus

<220>
<221> CDS
10 <222> (669)..(728)
<223>

<220>
<221> Intrón
15 <222> (729)..(872)
<223>

<220>
<221> CDS
20 <222> (873)..(1015)
<223>

<220>
<221> Intrón
25 <222> (1016)..(1082)
<223>

<220>
<221> CDS
30 <222> (1083)..(1127)
<223>

<220>

<221> Intrón
 <222> (1128)..(1183)
 <223>
 5 <220>
 <221> CDS
 <222> (1184)..(1236)
 <223>
 10 <220>
 <221> Intrón
 <222> (1237)..(1300)
 <223>
 15 <220>
 <221> CDS
 <222> (1301)..(1717)
 <223>
 20 <220>
 <221> Intrón
 <222> (1718)..(1776)
 <223>
 25 <220>
 <221> CDS
 <222> (1777)..(2489)
 <223>
 30 <220>
 <221> Intrón
 <222> (2490)..(2599)
 <223>
 35 <220>
 <221> CDS
 <222> (2600)..(3469)
 <223>
 40 <220>
 <221> Intrón
 <222> (3470)..(3531)
 <223>
 45 <220>
 <221> CDS
 <222> (3532)..(3759)
 <223>
 50 <400> 21

ggatccgtcc gcggacacag gcagagagac ggcacgggga ctcgacctga tcctcccagg 60
 gcgggggtgtt gtttgtggcg agggagcgat gctgatgttc ttccagctcc gttgctacct 120
 tcccacggcc atttagccgg cggacggcat gtaacatgtc aaacatgtgg gctcggcagt 180
 gggggcgtga gacgcagcac ctgacccggc ggcgcggcgc ttgcagggtc cagggacagc 240
 cggccgtggt cgtttgcggg gaaggcgaca cagacgactt ggcgcggccc gccggaaggc 300
 gaggaatcat gagtgcgacg gagacatggc aagaccacgg ccttcctggc gaagaagaag 360
 atgaataatc gcaggggcag tgtggcatgg accgcacggc cgccaggggac ctgccccgtg 420
 aggtttctcg ggtgtttcca ctggttccat cgctgggggc gatcccgagc ccgtgtgccc 480
 gtgtaactat tattgacgat caacatgcc a tggccagcca gcttctataa taatcatata 540
 taacaccccc cgttctcccg ctgccttgct ccgtggtctt cctggtcctg cttgagggtc 600
 acgagtctcc ttgcatggtc aactcgtcct ctgcttcac cgtgcttga 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 ttccagcaagt gcgcgccggc cgcgacatcc gccgccgctg ccctcaccga 818
 cgcagcaccc 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
 gtgagtccat 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

ES 2 501 944 T3

Leu Gly Phe Arg Gly Leu Cys Leu Gln Asp Ser Pro Leu Gly Val Arg	
85 90 95	
ttt g gtaggtcttt caacagagaa caagggctgt cgcgggagag atgctgacg	1286
Phe	
100	
atacctactt 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	
ggg 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
Tyr Ser Cys Ser Asn Ser Tyr Leu Leu Asn Lys Leu Leu Lys Ser Glu	
260 265 270	
ctt gat ttt cag ggc ttc gtg atg agt gac tgg gga gcg cac cac agc	1919
Leu Asp Phe Gln Gly Phe Val Met Ser Asp Trp Gly Ala His His Ser	
275 280 285	
ggc gtt gga gct gcc ctg gct ggc ctt gac atg tcg atg cca gga gac	1967
Gly Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp	
290 295 300	

ES 2 501 944 T3

acc gcc ttt ggt acc ggc aaa tcc ttc tgg gga acc aac ctg acc atc Thr Ala Phe Gly Thr Gly Lys Ser Phe Trp Gly Thr Asn Leu Thr Ile 305 310 315	2015
gcc gtt ctc aac ggt act gtt ccg gaa tgg cgt gtg gat gac atg gct Ala Val Leu Asn Gly Thr Val Pro Glu Trp Arg Val Asp Asp Met Ala 320 325 330 335	2063
gtt cgc atc atg gcg gcc ttt tac aag gtt ggt cgc gac cgt tac cag Val Arg Ile Met Ala Ala Phe Tyr Lys Val Gly Arg Asp Arg Tyr Gln 340 345 350	2111
gtg ccg gtc aac ttc gac tcg tgg acg aag gat gaa tac ggt tac gag Val Pro Val Asn Phe Asp Ser Trp Thr Lys Asp Glu Tyr Gly Tyr Glu 355 360 365	2159
cac gca ctg gtt ggc cag aac tat gtc aag gtc aat gac aag gtg gat His Ala Leu Val Gly Gln Asn Tyr Val Lys Val Asn Asp Lys Val Asp 370 375 380	2207
gtt cgt gcc gac cat gcg gac atc atc cgt caa att ggg tct gct agt Val Arg Ala Asp His Ala Asp Ile Ile Arg Gln Ile Gly Ser Ala Ser 385 390 395	2255
gtt gtc ctt ctt aag aac gat gga gga ctc cca ttg acc ggc tat gaa Val Val Leu Leu Lys Asn Asp Gly Gly Leu Pro Leu Thr Gly Tyr Glu 400 405 410 415	2303
aag ttc acc gga gtt ttt gga gag gat gcc gga tcg aac cgt tgg ggc Lys Phe Thr Gly Val Phe Gly Glu Asp Ala Gly Ser Asn Arg Trp Gly 420 425 430	2351
gct gac ggc tgc tct gat cgt ggt tgc gac aac ggc acg ttg gca atg Ala Asp Gly Cys Ser Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met 435 440 445	2399
ggt tgg ggc agt ggc act gct gac ttc ccc tac ctt gtc act ccc gag Gly Trp Gly Ser Gly Thr Ala Asp Phe Pro Tyr Leu Val Thr Pro Glu 450 455 460	2447
cag gca atc cag aat gaa atc ctt tcc aag ggg aag ggg tta Gln Ala Ile Gln Asn Glu Ile Leu Ser Lys Gly Lys Gly Leu 465 470 475	2489
gtgagtgtg tcaccgacaa tggtagccctt gaccagatgg aacaggttgc gtctcaggcc	2549
aggtattcct tcctccgtat cccatgaat cgaatctcca ctgacttttag gac agc Asp Ser	2605
gtt tct atc gtt ttc gtc aac gcc gac tct ggt gaa ggc tac atc aac Val Ser Ile Val Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile Asn 480 485 490 495	2653
gtt gat ggc aac gaa ggt gat cgg aag aac ctc acc ctc tgg aaa gga Val Asp Gly Asn Glu Gly Asp Arg Lys Asn Leu Thr Leu Trp Lys Gly 500 505 510	2701
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

ES 2 501 944 T3

gtt 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 Tyr Val 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 Pro Gln Pro Ile Leu Pro Ala Gly Gly Ala 720 725 730 735	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

ES 2 501 944 T3

gtgagtttcg cagtctcatt gatatatgtc tttcgagttg gtcactgacc cgcgatctat	3529
ag tat gtc tct ctt ggt ggt cca gat gac ccc aag gtc gtg ctc cgc Tyr Val Ser Leu Gly Gly Pro Asp Asp Pro Lys Val Val Leu Arg 770 775 780	3576
aac ttt gac cgc 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	3624
acg aca ttg acg cga cgc 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	3672
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	3720
tgc 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	3769
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 tagcgagtcg atagccgatg aacctattcc	4129
gtaggttgag gttcaccta caaataagcc aggatttaag taaataacctg ctctgtaa	4189
ctacaacgca tcagatcaga ggaaaattca aatggcagaa gtgcgagcac ctcggtgaga	4249
agagatcgag ctgtcgaagt cggctggaac acaggtaaag agaagtaata caattcattg	4309
atttttacat cgtttaacat gtagaaggta tctaaaatag taagtccaga tatgggccat	4369
ggagatcgcc tcggcgatct tcgggagtat ctcgggagac gcacatgacc gcgcttaacc	4429
ctgtcggttg gacccgagtc cgaccgacgt catcagcgca gcgcaggtca ggctgcgcgc	4489
aacgtcaatg ccagggggtg ctgggacagt tgcatatcaa tcgatcagtc aattaaagca	4549
tctgctttcc acgttctttt tttatcacct ttcacttccc ctgtcccaact tgccttgga	4609
ttgttgagcc caaagaagaa ggagaagaaa atgggctcga caccgccgaa cgggtggtcg	4669
acgagcacat catcagcagc gtcttattat caacattccc aaccaccggc cctcgttctc	4729
ctcgtctacc cgtcactct cctcctcggc tccctgtaca gagccatttc cccaccgcg	4789
cgggtgaggc acgatgctgc agaccctgct ctggccccga ccatagcgtc cgacatcaac	4849
ctgtcccagt catcccggtg ttcccattcc catagcaaca gcaacagccc ggtcaattac	4909
ttcgccccga aggacaacat ctttaacgtc tacttcgtca agatcggctg gttctggacg	4969

ES 2 501 944 T3

accctgcct tcttcacgtt actcctcacc cagcctgcct acacaaacgc cggccccctg 5029

cgcgcccgac gcaccctcca agccctgtcc cgctacgccca tcgtcacccct actacctgga 5089

tcc 5092

<210> 22

<211> 843

<212> PRT

<213> *Thermoascus aurantiacus*

<400> 22

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

ES 2 501 944 T3

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

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

ES 2 501 944 T3

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

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

ES 2 501 944 T3

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
5 <212> ADN
<213> Acremonium thermophilum

<220>
<221> CDS
10 <222> (391)..(447)
<223>

<220>
<221> Intrón
15 <222w (448)..(539)
<223>

<220>
 <221> CDS
 <222> (540)..(685)
 5 <223>

<220>
 <221> Intrón
 <222> (686)..(759)
 10 <223>

<220>
 <221> CDS
 <222> (760)..(1148)
 15 <223>

<220>
 <221> Intrón
 <222> (1149)..(1217)
 20 <223>

<220>
 <221> CDS
 <222> (1218)..(3208)
 25 <223>

<400> 23
 gcaggtagct acgacattcg acggtccacg cccagtggcg tctgctcggc cgtctgggaa 60
 ccatgcacgc ccgcctctta 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 gggacgcca 467
 Ala Leu Leu Ala Ser Ser Pro Val Ser Ala Ala
 10 15
 ggtatgggtcc caactgctct tcctccctgt ttcctcctct accggtgctg acaacgacaa 527
 tagctgcacc ag cga gaa ctc gcc act tcc gag 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

ES 2 501 944 T3

aag gcg ttc gtt tcc cag ctg acg ctc ttg gaa aag gtc aac ctg acg Lys Ala Phe Val Ser Gln Leu Thr Leu Leu Glu Lys Val Asn Leu Thr 50 55 60	674
act ggc atc gg gtgagttcttg ttctctcctg tagaaccgcc taccagaaga Thr Gly Ile Gly 65	725
cattcaggaa gtgctaataga tgggcgggttg acag c tgg caa gga gga caa tgc Trp Gln Gly Gly Gln Cys 70	778
gtg ggc aac gtc ggt tcc gtc ccg cgt ctc ggc ctt cgc agc ctg tgc Val Gly Asn Val Gly Ser Val Pro Arg Leu Gly Leu Arg Ser Leu Cys 75 80 85 90	826
atg cag gac tcc ccc gtg ggt atc cgc ttt ggg gac tac gtc tcc gtc Met Gln Asp Ser Pro Val Gly Ile Arg Phe Gly Asp Tyr Val Ser Val 95 100 105	874
ttc ccc tct ggt cag acc acg gct gcc acc ttc gac aag ggt ctg atg Phe Pro Ser Gly Gln Thr Thr Ala Ala Thr Phe Asp Lys Gly Leu Met 110 115 120	922
aac cgt cgc ggc aat gcc atg ggc cag gag cac aaa gga aag ggt gtc Asn Arg Arg Gly Asn Ala Met Gly Gln Glu His Lys Gly Lys Gly Val 125 130 135	970
aac gtc ctg ctc ggc ccg gtc gct ggc ccc att ggc cgt acg ccc gag Asn Val Leu Leu Gly Pro Val Ala Gly Pro Ile Gly Arg Thr Pro Glu 140 145 150	1018
ggg gga cga aac tgg gag ggc ttc tcc ccc gac ccc gtc cta acg ggt Gly Gly Arg Asn Trp Glu Gly Phe Ser Pro Asp Pro Val Leu Thr Gly 155 160 165 170	1066
att gcc ttg gcc gaa acg atc aag gga atc cag gat gct ggt gtc att Ile Ala Leu Ala Glu Thr Ile Lys Gly Ile Gln Asp Ala Gly Val Ile 175 180 185	1114
gct tgc gcc aag cat ttc atc gcg aac gaa cag g gtgcgtgatg Ala Cys Ala Lys His Phe Ile Ala Asn Glu Gln 190 195	1158
gaacgcggga cgtgctctga tgcaaaccga cgagcactga ccacgctttc ctggaacag	1217
aa cac ttc cgc cag tcc ggc gag gcc cag ggc tac ggc ttt gac atc Glu His Phe Arg Gln Ser Gly Glu Ala Gln Gly Tyr Gly Phe Asp Ile 200 205 210	1264
tcc gag tcg ctg tcg tcc aac atc gac gac aag acc atg cac gag ctg Ser Glu Ser Leu Ser Ser Asn Ile Asp Asp Lys Thr Met His Glu Leu 215 220 225	1312
tat ctg tgg ccc ttc gcc gac ggc gtg cgt gcc ggc gtc ggc gcc atc Tyr Leu Trp Pro Phe Ala Asp Gly Val Arg Ala Gly Val Gly Ala Ile 230 235 240 245	1360
atg tgc tcg tac aac cag atc aac aac tcg tac ggg tgc cag aac tcc Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr Gly Cys Gln Asn Ser 250 255 260	1408

ES 2 501 944 T3

aag acc ctg aac aac ctg ctc aag aac gag ctc ggc ttc cag ggc ttc Lys Thr Leu Asn Asn Leu Leu Lys Asn Glu Leu Gly Phe Gln Gly Phe 265 270 275	1456
gtc atg agc gac tgg cag gcc cag cac acc ggc gcg gcc agc gcc gtc Val Met Ser Asp Trp Gln Ala Gln His Thr Gly Ala Ala Ser Ala Val 280 285 290	1504
gcc ggc ctg gac atg acc atg ccc ggc gac acc agc ttc aac acc ggc Ala Gly Leu Asp Met Thr Met Pro Gly Asp Thr Ser Phe Asn Thr Gly 295 300 305	1552
ctc agc tac tgg ggc acg aac ctc acc ctc gcc gtc ctg aac ggc acc Leu Ser Tyr Trp Gly Thr Asn Leu Thr Leu Ala Val Leu Asn Gly Thr 310 315 320 325	1600
gtc ccc gag tac cgc atc gac gac atg gtc atg cgc atc atg gcc gcc Val Pro Glu Tyr Arg Ile Asp Asp Met Val Met Arg Ile Met Ala Ala 330 335 340	1648
ttc ttc aag acc ggc cag acc ctg gac ctg ccg ccc atc aac ttc gac Phe Phe Lys Thr Gly Gln Thr Leu Asp Leu Pro Pro Ile Asn Phe Asp 345 350 355	1696
tcg tgg acc acc gac acc ttc ggc ccg ctc cac ttc gcc gtc aac gag Ser Trp Thr Thr Asp Thr Phe Gly Pro Leu His Phe Ala Val Asn Glu 360 365 370	1744
gac cgc cag cag atc aac tgg cac gtc aac gtc cag gac aac cat ggc Asp Arg Gln Gln Ile Asn Trp His Val Asn Val Gln Asp Asn His Gly 375 380 385	1792
agc ctc atc cgc gag atc gcg gcc aag gga acc gtc ctg ctg aag aac Ser Leu Ile Arg Glu Ile Ala Ala Lys Gly Thr Val Leu Leu Lys Asn 390 395 400 405	1840
acc ggg tcc ctc ccg ctc aac aag ccc aag ttc ctc gtc gtg gtc ggc Thr Gly Ser Leu Pro Leu Asn Lys Pro Lys Phe Leu Val Val Val Gly 410 415 420	1888
gac gac gcg ggc ccc aac ccg gcg gga ccc aac gcc tgc ccc gac cgc Asp Asp Ala Gly Pro Asn Pro Ala Gly Pro Asn Ala Cys Pro Asp Arg 425 430 435	1936
gga tgc gac gtc ggc acc ctc ggc atg gcc tgg ggc tcc ggc tcg gcc Gly Cys Asp Val Gly Thr Leu Gly Met Ala Trp Gly Ser Gly Ser Ala 440 445 450	1984
aac ttc ccc tac ctg atc acc ccg gac gcc gcg ctg cag gcg cag gcg Asn Phe Pro Tyr Leu Ile Thr Pro Asp Ala Ala Leu Gln Ala Gln Ala 455 460 465	2032
atc aag gac ggc acc cgc tac gag agc gtg ctg tcc aac tac cag ctc Ile Lys Asp Gly Thr Arg Tyr Glu Ser Val Leu Ser Asn Tyr Gln Leu 470 475 480 485	2080
gac cag acc aag gcg ctg gtc acc cag gcc aac gcc acg gcc atc gtc Asp Gln Thr Lys Ala Leu Val Thr Gln Ala Asn Ala Thr Ala Ile Val 490 495 500	2128

ES 2 501 944 T3

ttc gtc aac gcc gac tcg ggc gag ggc tac atc aac gtc gac ggc aac Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile Asn Val Asp Gly Asn 505 510 515	2176
gag ggc gac cgc aag aac ctc acg ctc tgg cac gac ggc gac gcc ctg Glu Gly Asp Arg Lys Asn Leu Thr Leu Trp His Asp Gly Asp Ala Leu 520 525 530	2224
atc aag agc gtg gcc ggc tgg aac ccg aac acc atc gtc gtc atc cac Ile Lys Ser Val Ala Gly Trp Asn Pro Asn Thr Ile Val Val Ile His 535 540 545	2272
tcg acc ggc ccc gtc ctc gtg acc gac tgg tac gac cac ccc aac atc Ser Thr Gly Pro Val Leu Val Thr Asp Trp Tyr Asp His Pro Asn Ile 550 555 560 565	2320
acc gcc atc ctg tgg gcc ggc gtg ccc ggg cag gag tcc ggc aac gcc Thr Ala Ile Leu Trp Ala Gly Val Pro Gly Gln Glu Ser Gly Asn Ala 570 575 580	2368
atc acc gac gtc ctc tac gga aaa gtc aac ccg tcg ggc cgc agc ccc Ile Thr Asp Val Leu Tyr Gly Lys Val Asn Pro Ser Gly Arg Ser Pro 585 590 595	2416
ttc acc tgg ggt ccg acc cgc gag agc tac ggc acc gac gtc ctc tac Phe Thr Trp Gly Pro Thr Arg Glu Ser Tyr Gly Thr Asp Val Leu Tyr 600 605 610	2464
act ccc aac aac ggc aag ggc gcg ccg cag cag gcc ttc tcc gag ggc Thr Pro Asn Asn Gly Lys Gly Ala Pro Gln Gln Ala Phe Ser Glu Gly 615 620 625	2512
gtc ttc atc gac tac cgc cac ttc gac cgc acc aac gcg tcc gtc atc Val Phe Ile Asp Tyr Arg His Phe Asp Arg Thr Asn Ala Ser Val Ile 630 635 640 645	2560
tac gag ttc ggc cac ggc ctc agc tac acg acg ttc cag tac agc aac Tyr Glu Phe Gly His Gly Leu Ser Tyr Thr Thr Phe Gln Tyr Ser Asn 650 655 660	2608
atc cag gtg gtc aag tcc aac gcc ggc gcg tac aag ccc acg acg ggc Ile Gln Val Val Lys Ser Asn Ala Gly Ala Tyr Lys Pro Thr Thr Gly 665 670 675	2656
acg acc atc ccc gcg ccc acg ttt ggc agc ttc tcc aag gac ctc aag Thr Thr Ile Pro Ala Pro Thr Phe Gly Ser Phe Ser Lys Asp Leu Lys 680 685 690	2704
gac tac ctc ttc ccg tcg gac cag ttc cgc tac atc acc cag tac atc Asp Tyr Leu Phe Pro Ser Asp Gln Phe Arg Tyr Ile Thr Gln Tyr Ile 695 700 705	2752
tac ccg tac ctc aac tcc acc gac ccg gcc aag gcg tcg ctc gac ccg Tyr Pro Tyr Leu Asn Ser Thr Asp Pro Ala Lys Ala Ser Leu Asp Pro 710 715 720 725	2800
cac tac ggc aag acg gcg gcc gag ttt ctg ccg ccg cac gcg ctg gac His Tyr Gly Lys Thr Ala Ala Glu Phe Leu Pro Pro His Ala Leu Asp 730 735 740	2848
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 Ser Gly Lys Asn Glu Pro Gly	
745 750 755	
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 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 Gly Ala Glu Val Pro Gln Leu Tyr Val	
775 780 785	
tcg ctg ggc ggg ccc gac gac ccc aaa gtg gtc ctg cgc ggg 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 ctg 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 ctg tcc gcc gac ctg 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 ttcgggtccag cagatagtaa	3418
ccgtcttagt tcgtccatcc ctctctcgaa tgcgctgtat acacatgcgt atatagacgt	3478
tgtataggtg ccattgctag caatgcaagc tt	3510

<210> 24
 <211> 861
 <212> PRT
 <213> Acremonium thermophilum

<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

5

10

ES 2 501 944 T3

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

ES 2 501 944 T3

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		
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		

ES 2 501 944 T3

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

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

ES 2 501 944 T3

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

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

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

<400> 25
tgccggggttg ctgcgactta attaataact ggcaaaacgg cccggagctc agctctgacc 60
tccgccacat ccgctcggca ccatgccagc gcgttgcaac ggcataaagc gctcagggtt 120
ttcttccgcc tgctccccac tgccgatggc catctgcacc ccagctcgtc acatttatct 180
cgcgacacagc gtcttccac cagttgcctt gctcatgacg ctgttaaaga tggccctacc 240
tagccgctga gtcccacaac gccgagatgt ctttggccct ttacaaggca cgccatggcc 300
gtccaaggctc tggtcatgag tgtgtttgtg gggccgaagg acacctcagt ggccacgaaa 360
tgccgccgag cgggccagca catgtcgaga gagacatgga catttatccc cgagatgctg 420
tattagggaa ccggtccttt tctcggagcc gtgatccgag agcgttcggg agtcggtgag 480

25

taaaagatgt cgagttgccg ttatatatcg cgggcctgta gctatgtgcc ctctattctc	540
acaggttcaa tcatcagtcg tcgccgtgag acgtagcgcg ctgaactagc gctcgatctc	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 tcg cag gct gcc gct gcc Thr Gly Thr Ala Ala Trp Asp Ala Ala His Ser Gln Ala Ala Ala Ala 35 40 45	745
gtg tcg 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 tcg 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 Ala Leu Gly Lys Ile Pro His Gly Gly Arg Asn Trp Glu Ala Phe Gly 145 150 155	1081
tcc gac ccc tac ttg gcc ggt atc gct atg tcc gag acc atc gag ggc Ser Asp Pro Tyr Leu Ala Gly Ile Ala Met Ser Glu Thr Ile Glu Gly 160 165 170	1129
att cag tcg gag ggt gtg cag gct tgc gcg aag cac tac atc gcc aat Ile Gln Ser Glu Gly Val Gln Ala Cys Ala Lys His Tyr Ile Ala Asn 175 180 185 190	1177
gag cag gaa ctc aac cgc gag aca atg agc agc aac gtc gac gac cgc Glu Gln Glu Leu Asn Arg Glu Thr Met Ser Ser Asn Val Asp Asp Arg 195 200 205	1225
act atg cac gag cta tac ctc tgg ccg ttc gcc gac gcc gtg cat tcc Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val His Ser 210 215 220	1273
aac gtg gcc agc gtc atg tgc agc tac aac aag ctc aac ggc acc tgg	1321

ES 2 501 944 T3

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	cgc	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	
255					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	ggc	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	cgc	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	
	320					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	
335				340				345						350		
gcc	gta	gcc	cgc	gat	ggc	atc	gtc	ctc	ctc	aag	aac	gac	ggc	att	ctg	1705
Ala	Val	Ala	Arg	Asp	Gly	Ile	Val	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	ggc	atg	aac	gcc	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				
ggc	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	
	400					405					410					
ttt	gtt	gcg	ccg	tat	gat	gct	ctg	cgt	gag	cgg	gca	cag	cgc	gat	ggc	1897
Phe	Val	Ala	Pro	Tyr	Asp	Ala	Leu	Arg	Glu	Arg	Ala	Gln	Arg	Asp	Gly	
415				420				425						430		
acg	cag	atc	agt	ctg	cat	gca	tcg	gac	aat	aca	aac	ggg	gtt	aac	aac	1945
Thr	Gln	Ile	Ser	Leu	His	Ala	Ser	Asp	Asn	Thr	Asn	Gly	Val	Asn	Asn	
			435					440					445			
gcc	gtg	cag	ggc	gct	gac	gcg	gcg	ttt	gtg	ttc	atc	act	gct	gac	tcc	1993
Ala	Val	Gln	Gly	Ala	Asp	Ala	Ala	Phe	Val	Phe	Ile	Thr	Ala	Asp	Ser	
		450						455				460				
ggc	gaa	ggg	tac	att	acc	gtt	gag	ggc	cat	gct	ggc	gac	cgg	aat	cat	2041
Gly	Glu	Gly	Tyr	Ile	Thr	Val	Glu	Gly	His	Ala	Gly	Asp	Arg	Asn	His	

ES 2 501 944 T3

465	470	475	
ctg gat cct tgg cat aat ggt aac cag ctt gtg cag gct gtt gcg cag Leu Asp Pro Trp His Asn Gly Asn Gln Leu Val Gln Ala Val Ala Gln 480 485 490			2089
gca aat aag aac gtc att gtg gtt gtg cac agc gtt ggg ccg gtt att Ala Asn Lys Asn Val Ile Val Val Val His Ser Val Gly Pro Val Ile 495 500 505 510			2137
ctg gag acg atc ctc aat acg ccc ggt gtg agg gct gtt gtt tgg gct Leu Glu Thr Ile Leu Asn Thr Pro Gly Val Arg Ala Val Val Trp Ala 515 520 525			2185
ggc ttg ccg agc cag gag agc ggt aac gcg ctg gtt gat gtg ctg tac Gly Leu Pro Ser Gln Glu Ser Gly Asn Ala Leu Val Asp Val Leu Tyr 530 535 540			2233
ggc ctt gtt tgc ccg tgc ggc aag ctt gtc tac acc att gcg aag agc Gly Leu Val Ser Pro Ser Gly Lys Leu Val Tyr Thr Ile Ala Lys Ser 545 550 555			2281
ccg agc gac tac ccg act agc att gtc cgt ggc gat gat aac ttc cgc Pro Ser Asp Tyr Pro Thr Ser Ile Val Arg Gly Asp Asp Asn Phe Arg 560 565 570			2329
gag ggt ctg ttc atc gac tac agg cac ttc gat aac gcc cgg atc gag Glu Gly Leu Phe Ile Asp Tyr Arg His Phe Asp Asn Ala Arg Ile Glu 575 580 585 590			2377
ccc cgt ttc gag ttt ggc ttc ggt ctc t gtaagtctct taccactccg Pro Arg Phe Glu Phe Gly Phe Gly Leu 595			2425
ttttgtaaca acccgattct aacatccccc ag ca tac acc aac ttc agc tat Ser Tyr Thr Asn Phe Ser Tyr 605			2477
tcc aac ctg ggc atc tcc tgc tcc gca acc gcc ggc cca gcc acg ggc Ser Asn Leu Gly Ile Ser Ser Ser Ala Thr Ala Gly Pro Ala Thr Gly 610 615 620			2525
ccc acc gtc ccc ggc ggc ccg gcc gac ctc tgg aac tat gtc gcg acc Pro Thr Val Pro Gly Gly Pro Ala Asp Leu Trp Asn Tyr Val Ala Thr 625 630 635			2573
gtc acg gcg acc gtt acc aac acc ggc ggc gtg gaa ggt gcc gag gtc Val Thr Ala Thr Val Thr Asn Thr Gly Gly Val Glu Gly Ala Glu Val 640 645 650			2621
gct cag ctg tac atc tct ttg cca tct tgc gct cct gca tgc cca ccg Ala Gln Leu Tyr Ile Ser Leu Pro Ser Ser Ala Pro Ala Ser Pro Pro 655 660 665 670			2669
aag cag ctt cgt ggc ttt gtc aag ctt aag ttg gcg cct ggt caa agc Lys Gln Leu Arg Gly Phe Val Lys Leu Lys Leu Ala Pro Gly Gln Ser 675 680 685			2717
ggg acg gca acg ttt aga cta agg aag agg gat ttg gct tat tgg gat Gly Thr Ala Thr Phe Arg Leu Arg Lys Arg Asp Leu Ala Tyr Trp Asp 690 695 700			2765

ES 2 501 944 T3

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	
taggggggtta tggtcagcac ctagttgggg 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 cgaccttcct gaccctactg 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

5 <212> PRT

<213> Chaetomium thermophilum

<400> 26

ES 2 501 944 T3

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

Thr Ser Val Thr Ala Phe Thr Pro Gly Ile Gln Ala Ala Ser Thr Trp
100 105 110

ES 2 501 944 T3

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

ES 2 501 944 T3

Ala Arg Asp Gly Ile Val Leu Leu Lys Asn Asp Gly Ile Leu Pro Leu
355 360 365

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

ES 2 501 944 T3

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

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
5 <212> ADN
<213> Thermoascus aurantiacus

<220>
<221> Intrón
10 <222> (610)..(674)
<223>

<220>
<221> CDS
15 <222> (675)..(1628)
<223>

<220>
<221> CDS
20 <222> (1)..(609)
<223>

<400> 27
atg tat cag cgc gct ctt ctc ttc tct ttc ttc ctc gcc gcc gcc cgc 48

25

ES 2 501 944 T3

Met	Tyr	Gln	Arg	Ala	Leu	Leu	Phe	Ser	Phe	Phe	Leu	Ala	Ala	Ala	Arg	
1				5					10					15		
gcg	cag	cag	gcc	ggc	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	ggc	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	
65					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	ggc	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	ggc	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	
145					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	ggc	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	ggc	cag	gtacgtcaga	agtgataact				629
Cys	Pro	Arg	Asp	Leu	Lys	Phe	Ile	Asn	Gly	Gln						
		195				200										
agccagcaga	gccccatgaat	cattaactaa	cgctgtcaaa	tacag	gcc	aat	gtt	gaa								686
					Ala	Asn	Val	Glu								
					205											
ggc	tgg	cag	ccg	tct	gcc	aac	gac	cca	aat	gcc	ggc	gtt	ggc	aac	cac	734
Gly	Trp	Gln	Pro	Ser	Ala	Asn	Asp	Pro	Asn	Ala	Gly	Val	Gly	Asn	His	
		210					215					220				
ggc	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	

ES 2 501 944 T3

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			
240	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 Gly Gly Thr Tyr Ser Ser Thr Arg Tyr Ala Gly			
	260	265	270
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			
	275	280	285
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			
	290	295	300
acc gtc gtc acc cag ttc atc acc gac gac ggg acc ccc tcc ggc acc			1022
Thr Val Val Thr Gln Phe Ile Thr Asp Asp Gly Thr Pro Ser Gly Thr			
	305	310	315
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			
320	325	330	335
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			
	340	345	350
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			
	355	360	365
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			
	370	375	380
atg gtc ctc gtc atg agc ctg tgg gac gat cac gcc gcc aac atg ctc			1262
Met Val Leu Val Met Ser Leu Trp Asp Asp His Ala Ala Asn Met Leu			
	385	390	395
tgg ctg gac agc acc tac ccg act gat gcg gac ccg gac acc cct ggc			1310
Trp Leu Asp Ser Thr Tyr Pro Thr Asp Ala Asp Pro Asp Thr Pro Gly			
400	405	410	415
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			
	420	425	430
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			
	435	440	445
gga ccc att ggc agc acc ggc aac cct agc ggc ggc aac cct ccc ggc			1454
Gly Pro Ile Gly Ser Thr Gly Asn Pro Ser Gly Gly Asn Pro Pro Gly			
	450	455	460
gga aac ccg cct ggc acc acc acc acc cgc cgc cca gcc act acc act			1502
Gly Asn Pro Pro Gly Thr Thr Thr Thr Arg Arg Pro Ala Thr Thr Thr			
	465	470	475

ES 2 501 944 T3

gga agc tct ccc gga cct acc cag tct cac tac ggc cag tgc ggc ggt 1550
Gly Ser Ser Pro Gly Pro Thr Gln Ser His Tyr Gly Gln Cys Gly Gly
480 485 490 495

att ggc tac agc ggc ccc acg gtc tgc gcc agc ggc aca act tgc cag 1598
Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala Ser Gly Thr Thr Cys Gln
500 505 510

gtc ctg aac cct tac tac tct cag tgc ctg taa 1631
Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
515 520

<210> 28

<211> 521

<212> PRT

<213> *Thermoascus aurantiacus*

<400> 28

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

ES 2 501 944 T3

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
 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

ES 2 501 944 T3

				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
 5 <212> ADN
 <213> Thermoascus aurantiacus

<220>
 <221> Intrón
 10 <222> (610)..(674)
 <223>

<220>
 <221> CDS
 15 <222> (1726)..(1731)
 <223>

<220>
 <221> CDS
 20 <222> (675)..(1661)
 <223>

<220>
 <221> CDS
 25 <222> (1)..(609)
 <223>

<220>
 <221> Intrón
 30 <222> (1662)..(1725)
 <223>

```

<400> 29
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 ggt 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 ggt 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
65          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 ggt 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 ggt 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
145          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 ggt 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 ggt cag gtacgtcaga agtgataact      629
Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln
          195          200

agccagcaga gcccatgaat cattaactaa cgctgtcaaa tacag gcc aat gtt gaa      686
Ala Asn Val Glu
          205

```

ES 2 501 944 T3

ggc tgg cag ccg tct gcc aac gac cca aat gcc ggc gtt ggt aac cac Gly Trp Gln Pro Ser Ala Asn Asp Pro Asn Ala Gly Val Gly Asn His 210 215 220	734
ggt tcc tgc tgc gct gag atg gat gtc tgg gaa gcc aac agc atc tct Gly Ser Cys Cys Ala Glu Met Asp Val Trp Glu Ala Asn Ser Ile Ser 225 230 235	782
act gcg gtg acg cct cac cca tgc gac acc ccc ggc cag acc atg tgc Thr Ala Val Thr Pro His Pro Cys Asp Thr Pro Gly Gln Thr Met Cys 240 245 250 255	830
cag gga gac gac tgt ggt gga acc tac tcc tcc act cga tat gct ggt Gln Gly Asp Asp Cys Gly Gly Thr Tyr Ser Ser Thr Arg Tyr Ala Gly 260 265 270	878
acc tgc gac cct gat ggc tgc gac ttc aat cct tac cgc cag ggc aac Thr Cys Asp Pro Asp Gly Cys Asp Phe Asn Pro Tyr Arg Gln Gly Asn 275 280 285	926
cac tcg ttc tac ggc ccc ggg cag atc gtc gac acc agc tcc aaa ttc His Ser Phe Tyr Gly Pro Gly Gln Ile Val Asp Thr Ser Ser Lys Phe 290 295 300	974
acc gtc gtc acc cag ttc atc acc gac gac ggg acc ccc tcc ggc acc Thr Val Val Thr Gln Phe Ile Thr Asp Asp Gly Thr Pro Ser Gly Thr 305 310 315	1022
ctg acg gag atc aaa cgc ttc tac gtc cag aac ggc aag gta atc ccc Leu Thr Glu Ile Lys Arg Phe Tyr Val Gln Asn Gly Lys Val Ile Pro 320 325 330 335	1070
cag tcg gag tcg acg atc agc ggc gtc acc ggc aac tca atc acc acc Gln Ser Glu Ser Thr Ile Ser Gly Val Thr Gly Asn Ser Ile Thr Thr 340 345 350	1118
gag tat tgc acg gcc cag aag gcc gcc ttc ggc gac aac acc ggc ttc Glu Tyr Cys Thr Ala Gln Lys Ala Ala Phe Gly Asp Asn Thr Gly Phe 355 360 365	1166
ttc acg cac ggc ggg ctt cag aag atc agt cag gct ctg gct cag ggc Phe Thr His Gly Gly Leu Gln Lys Ile Ser Gln Ala Leu Ala Gln Gly 370 375 380	1214
atg gtc ctc gtc atg agc ctg tgg gac gat cac gcc gcc aac atg ctc Met Val Leu Val Met Ser Leu Trp Asp Asp His Ala Ala Asn Met Leu 385 390 395	1262
tgg ctg gac agc acc tac ccg act gat gcg gac ccg gac acc cct ggc Trp Leu Asp Ser Thr Tyr Pro Thr Asp Ala Asp Pro Asp Thr Pro Gly 400 405 410 415	1310
gtc gcg cgc ggt acc tgc ccc acg acc tcc ggc gtc ccg gcc gac gtt Val Ala Arg Gly Thr Cys Pro Thr Thr Ser Gly Val Pro Ala Asp Val 420 425 430	1358
gag tcg cag tac ccc aat tca tat gtt atc tac tcc aac atc aag gtc Glu Ser Gln Tyr Pro Asn Ser Tyr Val Ile Tyr Ser Asn Ile Lys Val 435 440 445	1406

ES 2 501 944 T3

gga ccc atc ggc tgc 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
450 455 460

aac ccg 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
465 470 475

ccg acc agc agc act agc tct ccc gtt tgc acc ccg act ggc cag ccc 1550
Pro Thr Ser Ser Thr Ser Ser Pro Val Ser Thr Pro Thr Gly Gln Pro
480 485 490 495

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
500 505 510

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
515 520 525

ccc tgg tac agc cag gtatgtttct cttccccctt ctagactcgc ttggatttga 1701
Pro Trp Tyr Ser Gln
530

cagttgctaa catctgctca acag tgc ctg taa 1734
Cys Leu

<210> 30

<211> 534

<212> PRT

<213> Thermoascus aurantiacus

<400> 30

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

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

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

- 5 1. Un método para tratar material celulósico con celobiohidrolasa, endoglucanasa, y beta-glucosidasa, por medio del cual dicha celobiohidrolasa es una proteína de fusión que comprende una región núcleo de celobiohidrolasa anclada a un dominio de unión a celulosa, y comprende una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 28 o 30, y en donde dicha región núcleo tiene una identidad de al menos 95% con el SEQ ID NO: 2.
2. El método de la reivindicación 1, en donde la endoglucanasa comprende una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 10, 12, 14 o 16, o con un fragmento del mismo que tiene actividad endoglucanasa.
- 10 3. El método de la reivindicación 2, en donde la endoglucanasa es obtenible de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*.
4. El método de la reivindicación 3, en donde la endoglucanasa es obtenible de *Thermoascus aurantiacus* CBS 116239, *Acremonium thermophilum* CBS 116240, o *Chaetomium thermophilum* CBS 730.95.
- 15 5. El método de la reivindicación 1, en donde la beta-glucosidasa comprende una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 22, 24 o 26, o con un fragmento del mismo que tiene actividad beta-glucosidasa.
6. El método de la reivindicación 5, en donde la beta-glucosidasa es obtenible de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*, preferiblemente de *Thermoascus aurantiacus* CBS 116239, *Acremonium thermophilum* CBS 116240, o *Chaetomium thermophilum* CBS 730.95.
- 20 7. El método de una cualquiera de las reivindicaciones anteriores, en donde el material celulósico es material lignocelulósico.
8. El método de una cualquiera de las reivindicaciones anteriores, que comprende tratar material lignocelulósico con al menos una enzima adicional, preferiblemente una xilanasa que comprende preferiblemente una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 18 o 20, o con un fragmento del mismo que tiene actividad xilanasa.
- 25 9. El método de la reivindicación 8, en donde la xilanasa es obtenible de *Thermoascus aurantiacus* o *Acremonium thermophilum*, preferiblemente de *Thermoascus aurantiacus* CBS 116239, o *Acremonium thermophilum* CBS 116240.
10. El método de una cualquiera de las reivindicaciones anteriores, en donde las enzimas se añaden al material celulósico o bien simultáneamente o bien sucesivamente.
- 30 11. El método de una cualquiera de las reivindicaciones anteriores, en donde al menos una de las enzimas está codificada por un gen similar al incluido en un microorganismo que tiene el número de acceso DSM 17326, DSM 17324, DSM 17323, DSM 17729, DSM 16724, DSM 16726, DSM 16725, DSM 17325 o DSM 17667.
- 35 12. El método de la reivindicación 1, en donde el material celulósico se selecciona del grupo que consiste en rastrojo de maíz, pasto varilla, paja de cereales, bagazo de caña de azúcar y materiales derivados de madera.
- 40 13. Una preparación de enzima que comprende celobiohidrolasa, endoglucanasa, y beta-glucosidasa, en donde dicha celobiohidrolasa es una proteína de fusión que comprende una región núcleo de celobiohidrolasa anclada a un dominio de unión a celulosa, y comprende una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 28 o 30, y en donde dicha región núcleo tiene una identidad de al menos 95% con el SEQ ID NO: 2.
14. La preparación de enzima de la reivindicación 13, en donde la endoglucanasa comprende una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 10, 12, 14 o 16, o con un fragmento del mismo que tiene actividad endoglucanasa.
- 45 15. La preparación de enzima de la reivindicación 14, en donde la endoglucanasa es obtenible de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*, preferiblemente de *Thermoascus aurantiacus* CBS 116239, *Acremonium thermophilum* CBS 116240, o *Chaetomium thermophilum* CBS 730.95.
16. La preparación de enzima de la reivindicación 13, en donde la beta-glucosidasa comprende una secuencia de aminoácidos que tiene una identidad de al menos 80% con el SEQ ID NO: 22, 24 o 26, o con un fragmento del mismo que tiene actividad beta-glucosidasa.
- 50 17. La preparación de enzima de la reivindicación 16, en donde la beta-glucosidasa es obtenible de *Thermoascus aurantiacus*, *Acremonium thermophilum*, o *Chaetomium thermophilum*, preferiblemente de *Thermoascus aurantiacus* CBS 116239, *Acremonium thermophilum* CBS 116240, o *Chaetomium thermophilum* CBS 730.95.

18. La preparación de enzima de una cualquiera de las reivindicaciones 13-17, 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 el SEQ ID NO: 18 o 20, o con un fragmento del mismo que tiene actividad xilanasa.
- 5 19. La preparación de enzima de la reivindicación 18, en donde la xilanasa es obtenible de *Thermoascus aurantiacus* o *Acremonium thermophilum*, preferiblemente de *Thermoascus aurantiacus* CBS 116239 o *Acremonium thermophilum* CBS 116240.
20. La preparación de enzima de una cualquiera de las reivindicaciones 13-19, en donde al menos una de las enzimas está codificada por un gen similar al incluido en un microorganismo que tiene el número de acceso DSM 17326, DSM 17324, DSM 17323, DSM 17729, DSM 16724, DSM 16726, DSM 16725, DSM 17325 o DSM 17667.
- 10 21. La preparación de enzima de la reivindicación 13, que está en forma de medio de cultivo gastado, o polvos, gránulos, o líquido.
22. El método de una cualquiera de las reivindicaciones 1-12, o la preparación de enzima de una cualquiera de las reivindicaciones 13-21, en donde la proteína de fusión comprende una secuencia de aminoácidos que tiene el SEQ ID NO: 28 o el SEQ ID NO: 30.
- 15 23. El uso de una preparación de enzima de una cualquiera de las reivindicaciones 13-22 para la degradación de material celulósico.
24. El uso de la reivindicación 23, en donde el material celulósico se selecciona del grupo que consiste en rastrojo de maíz, pasto varilla, paja de cereales, bagazo de caña de azúcar y materiales derivados de madera.
- 20 25. El uso del método de acuerdo con una cualquiera de las reivindicaciones 1-12 o 22 en un procedimiento para preparar etanol a partir de material celulósico.

Fig. 1A.

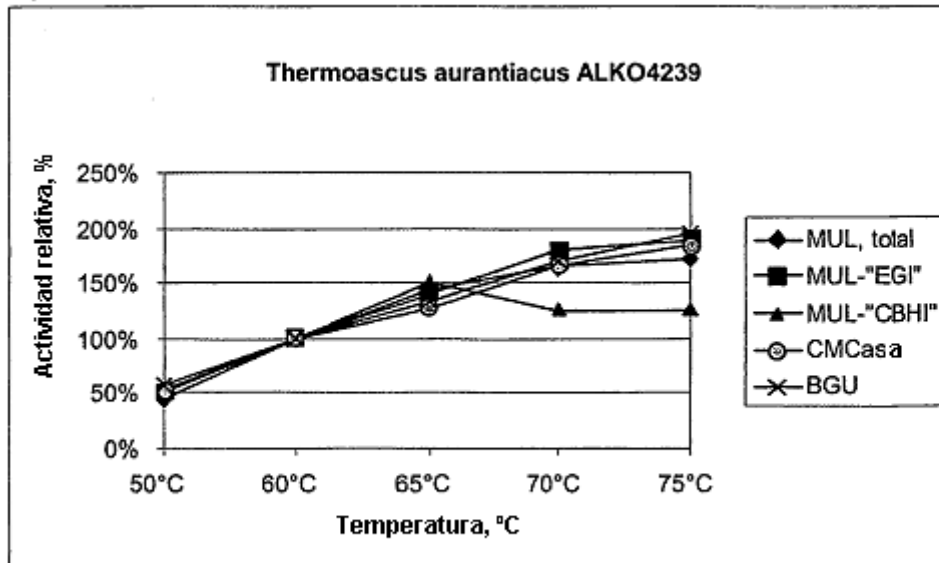


Fig. 1B.

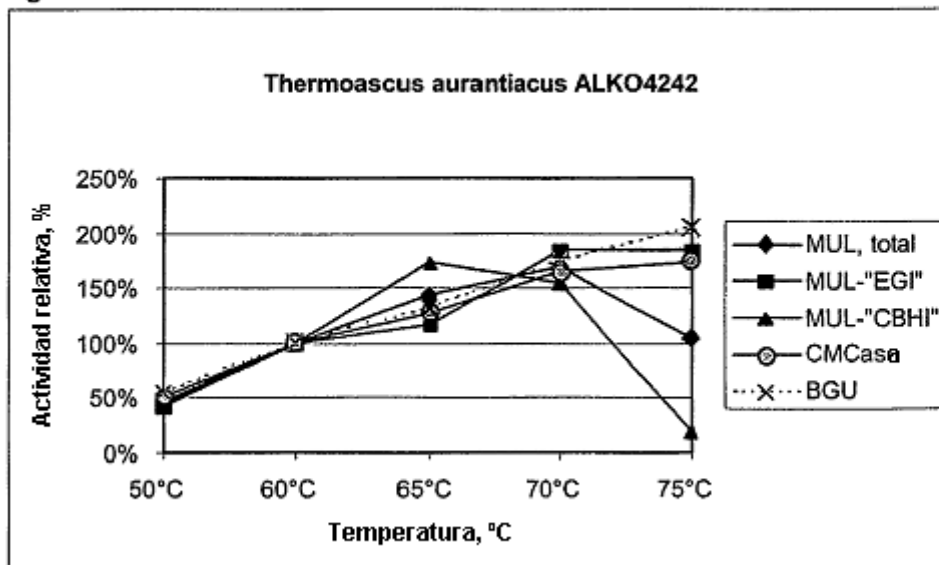


Fig. 1C.

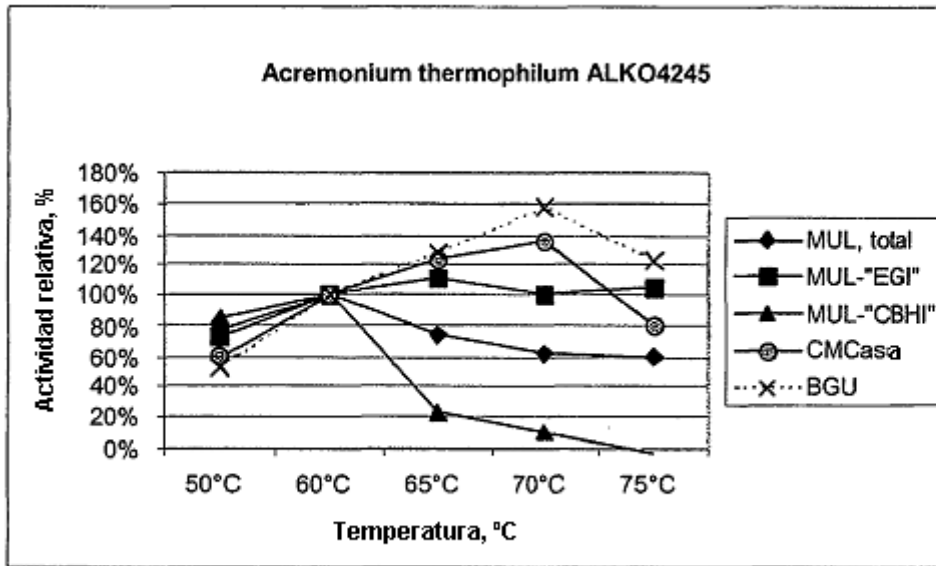


Fig. 1D.

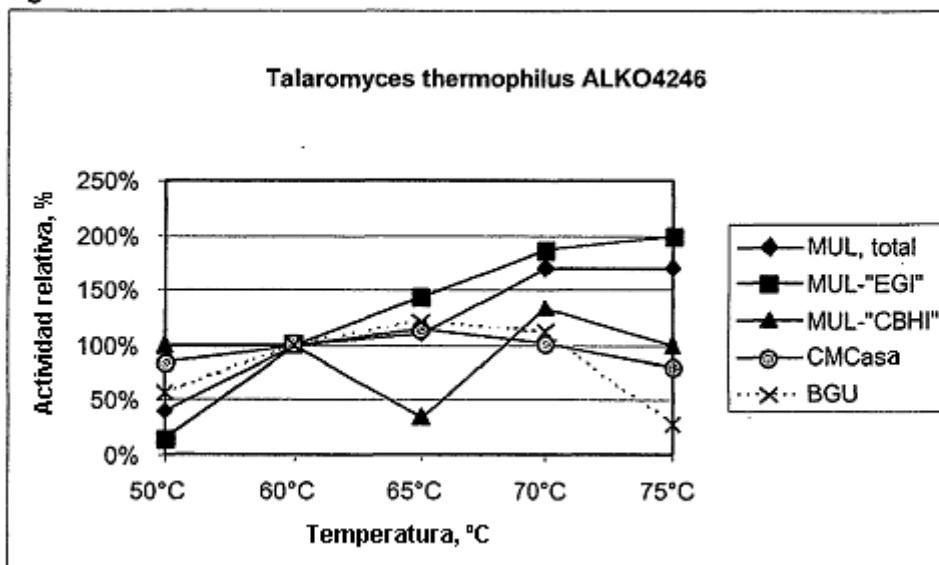


Fig. 1E.

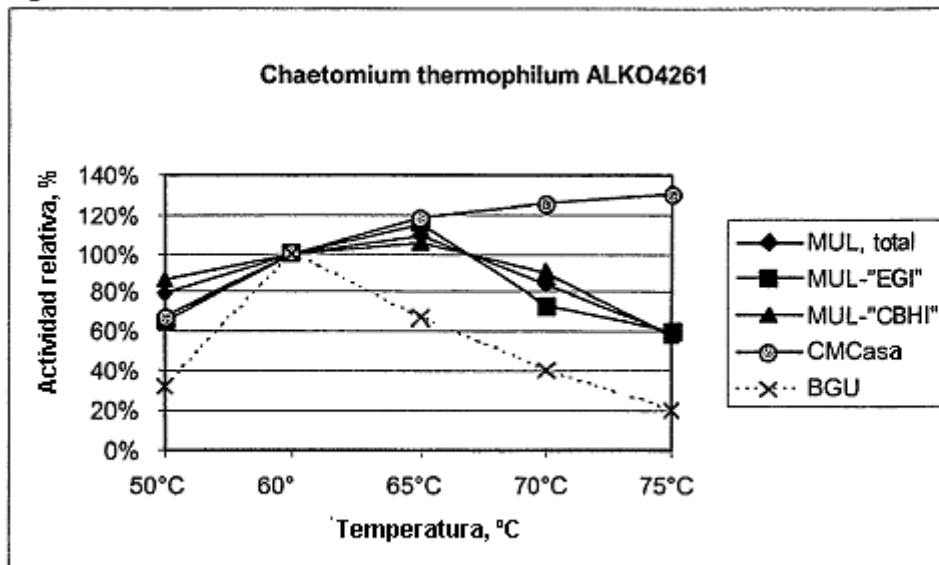


Fig. 1F.

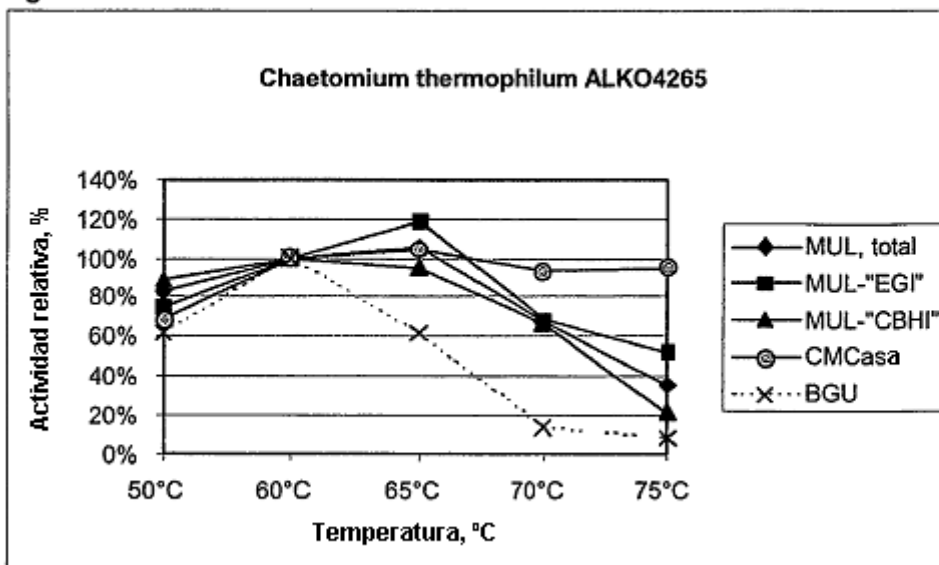


Fig. 2.

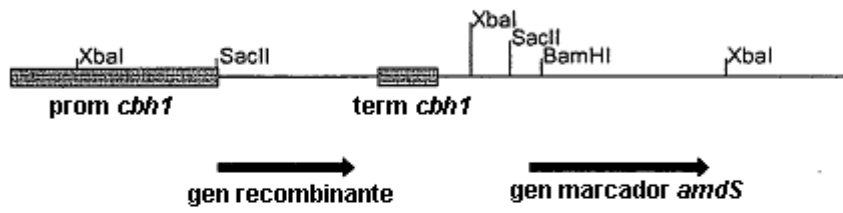


Fig. 3A.

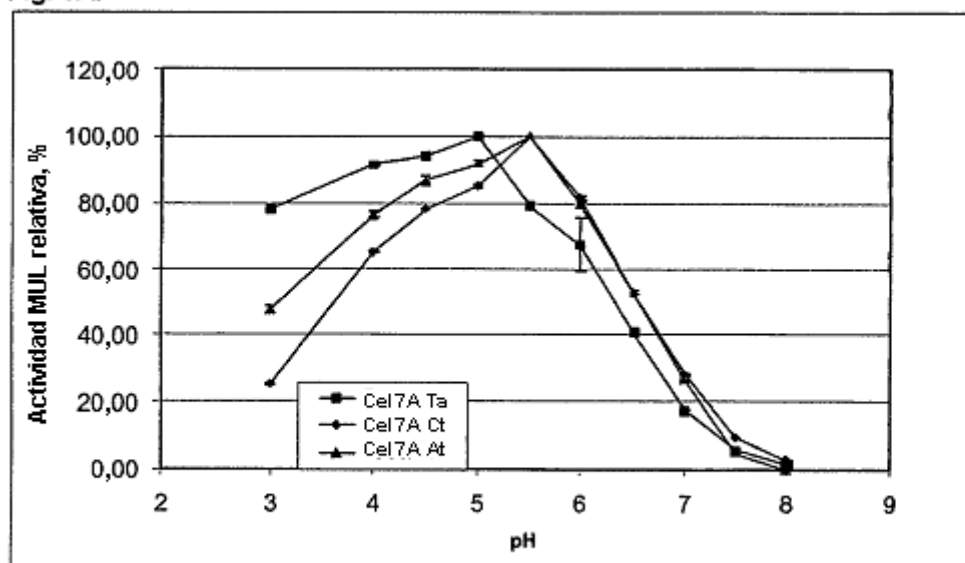


Fig. 3B.

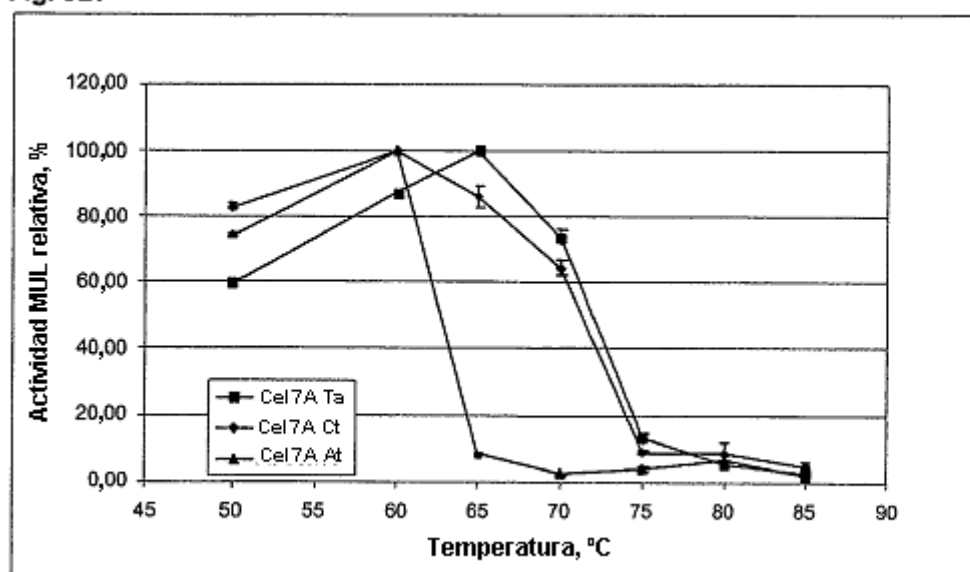


Fig. 4A.

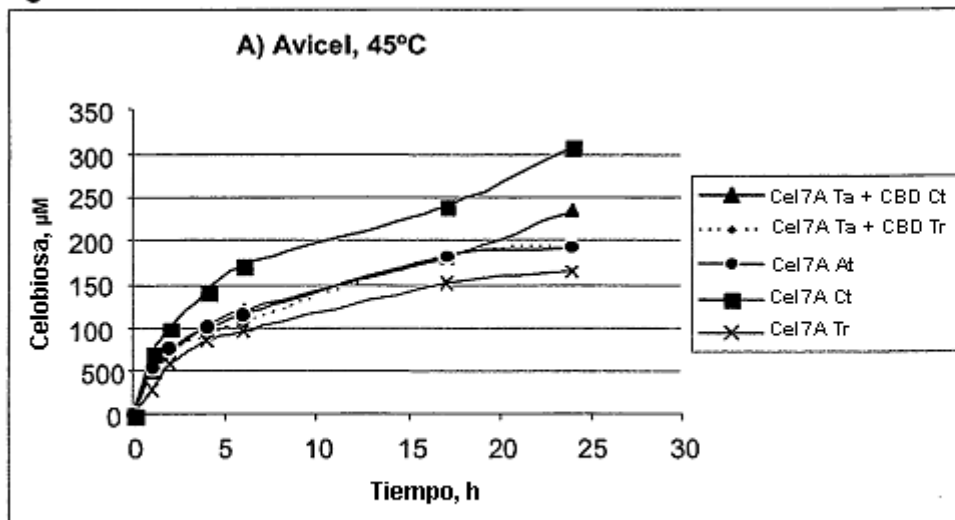


Fig. 4B.

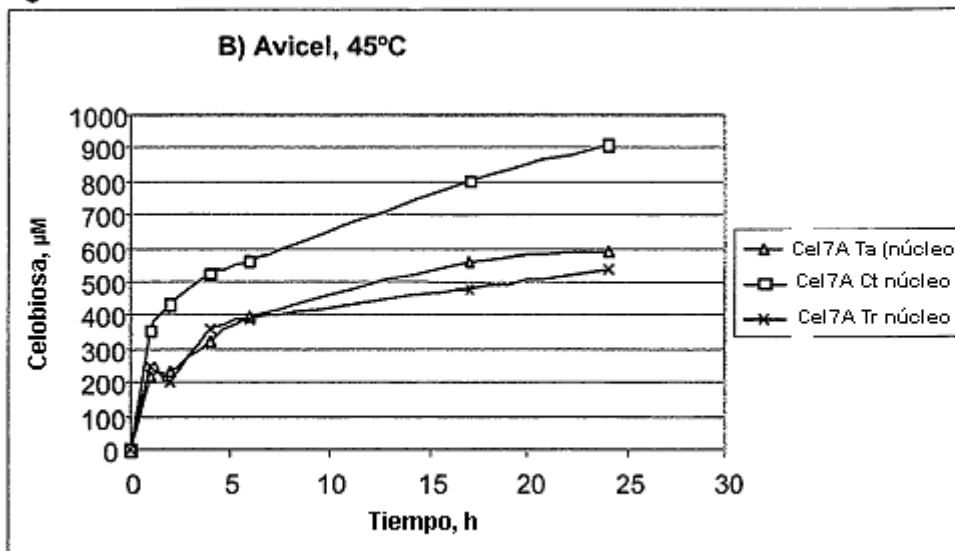


Fig. 5A.

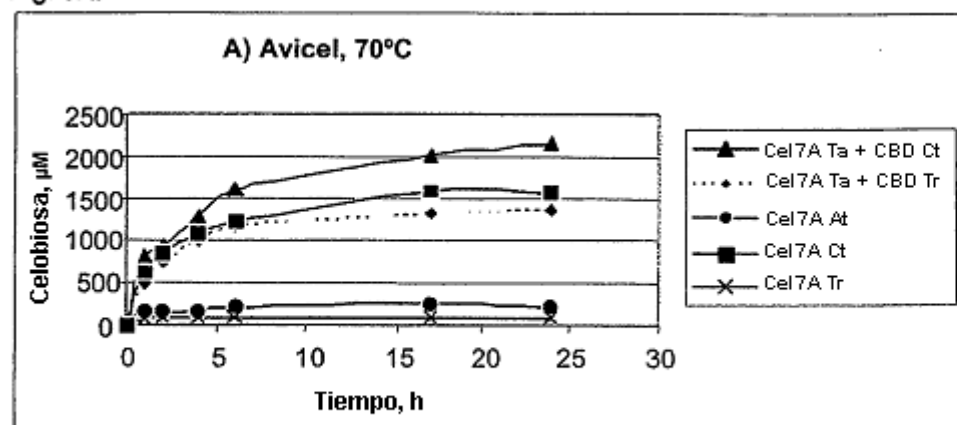


Fig. 5B.

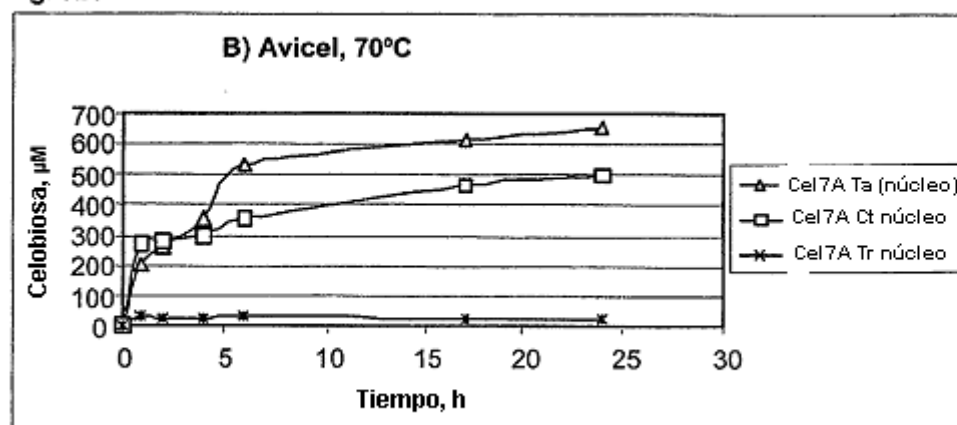


Fig. 6A.

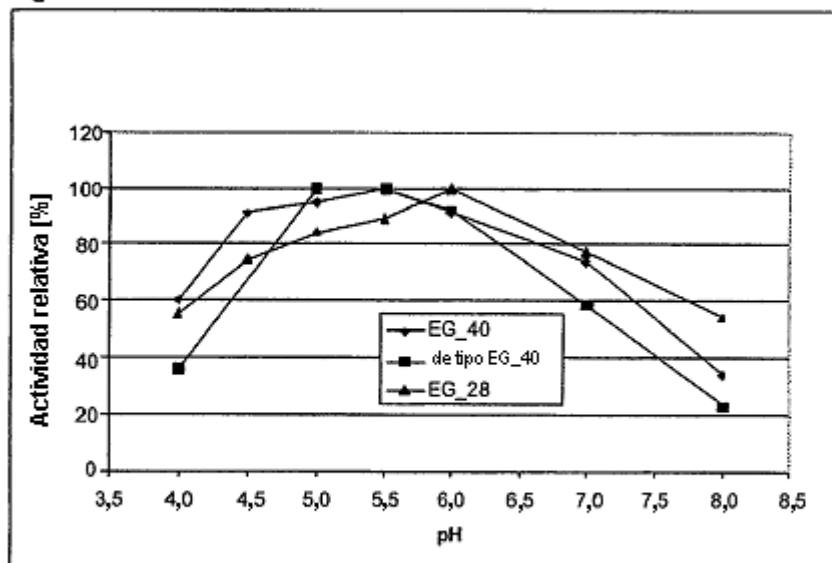


Fig. 6B.

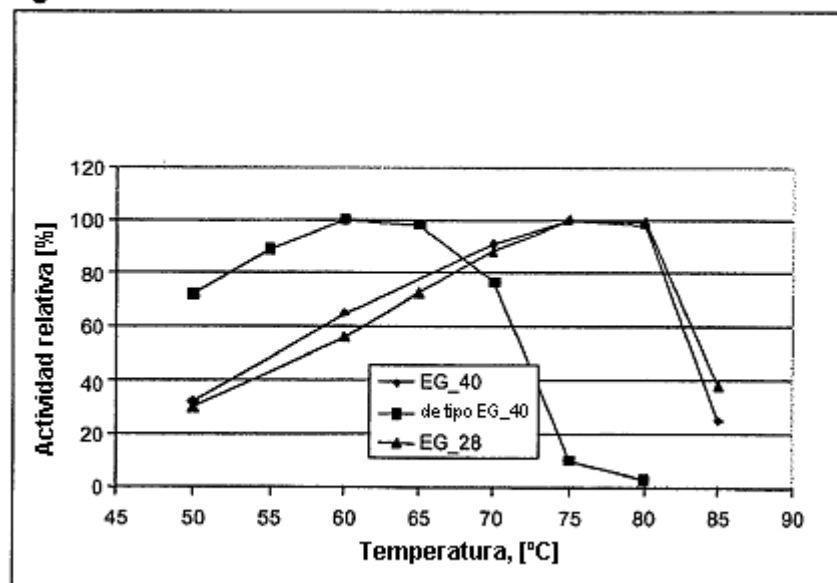


Fig. 7A.

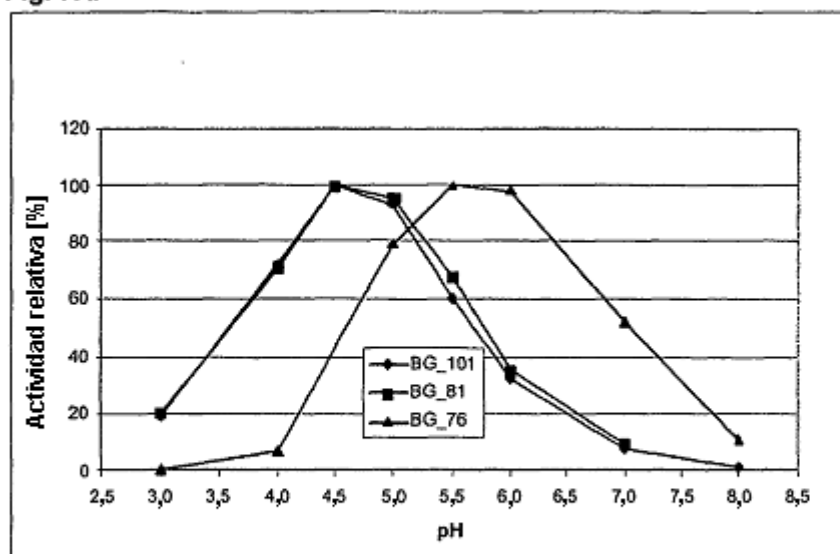


Fig. 7B.

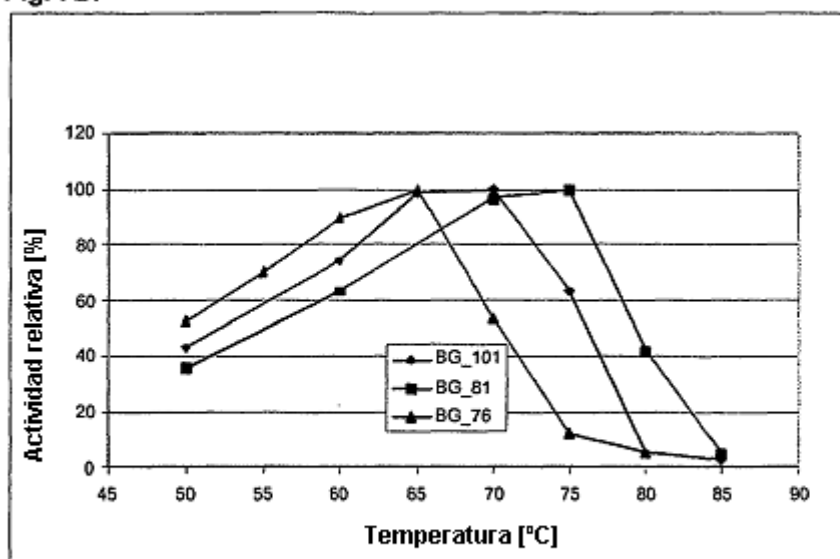


Fig. 8A.

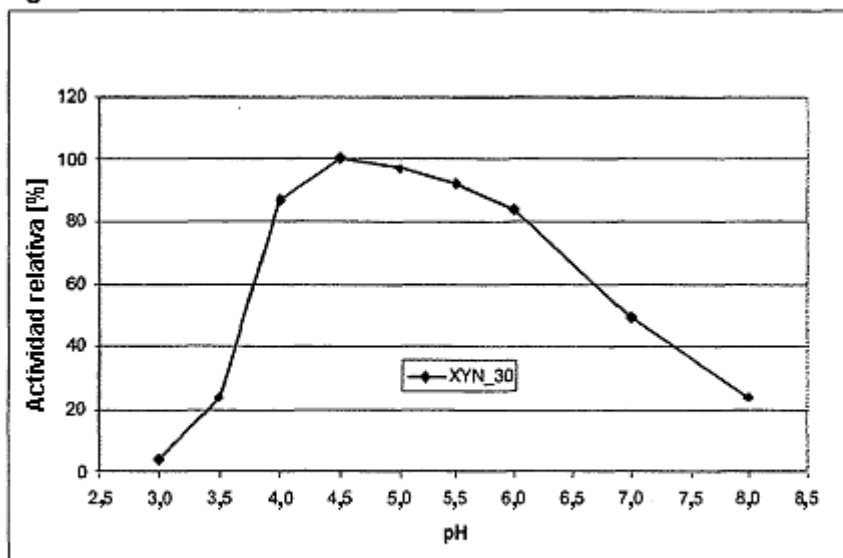


Fig 8B.

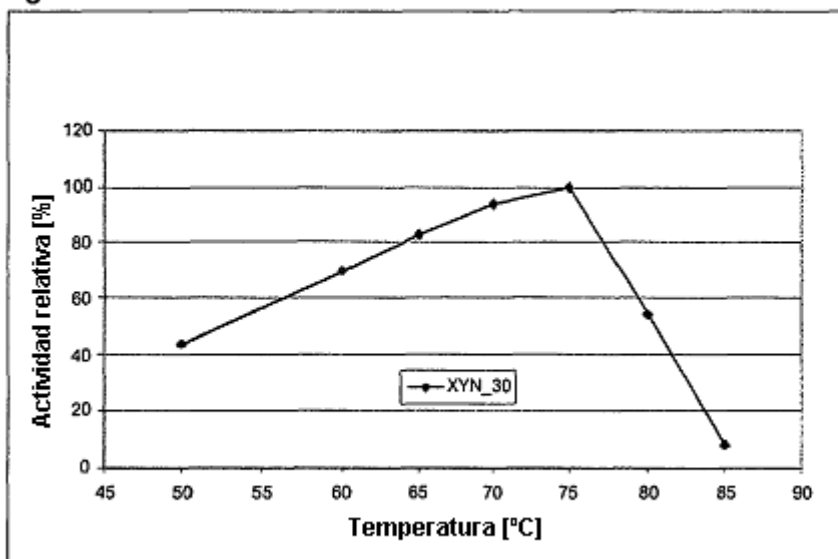


Fig. 9.

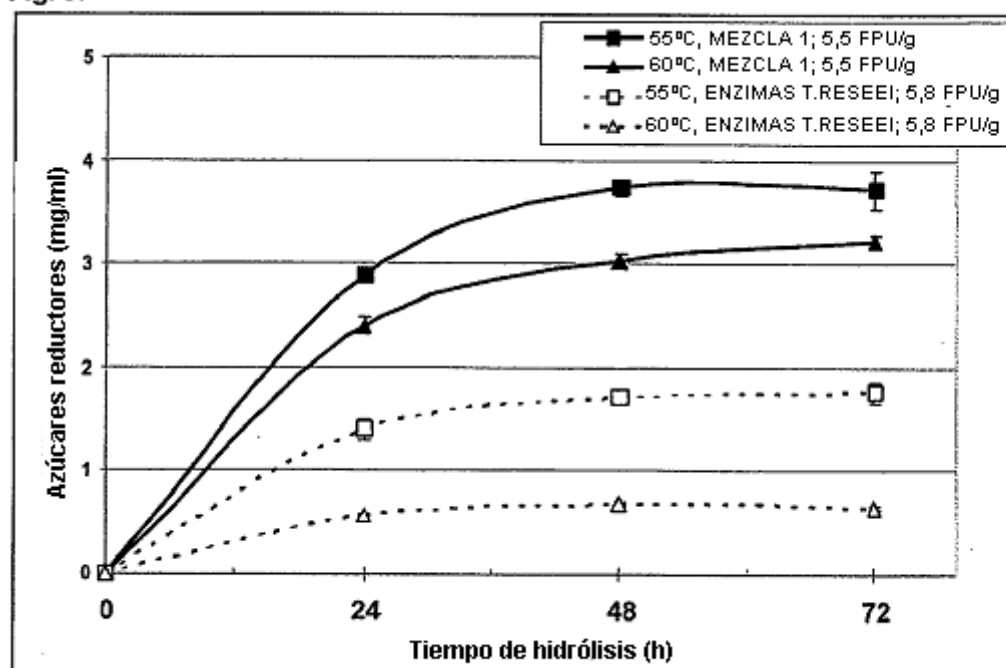


Fig. 10.

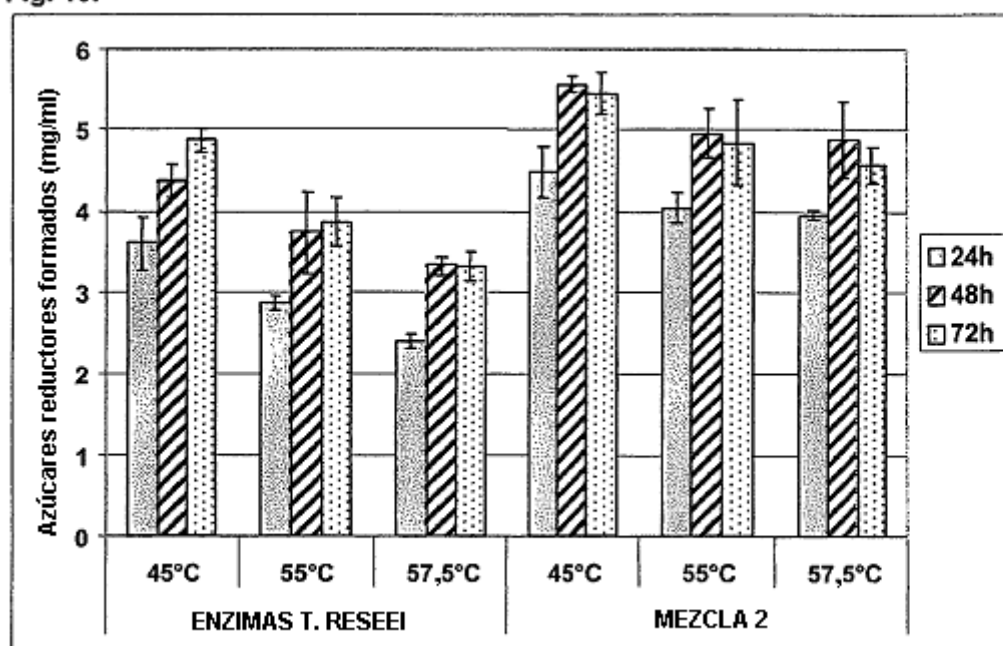


Fig. 11.

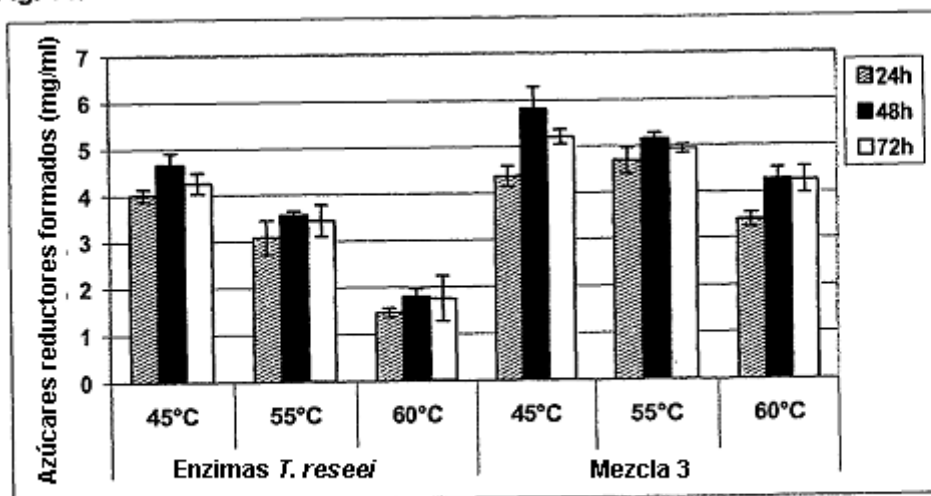


Fig. 12.

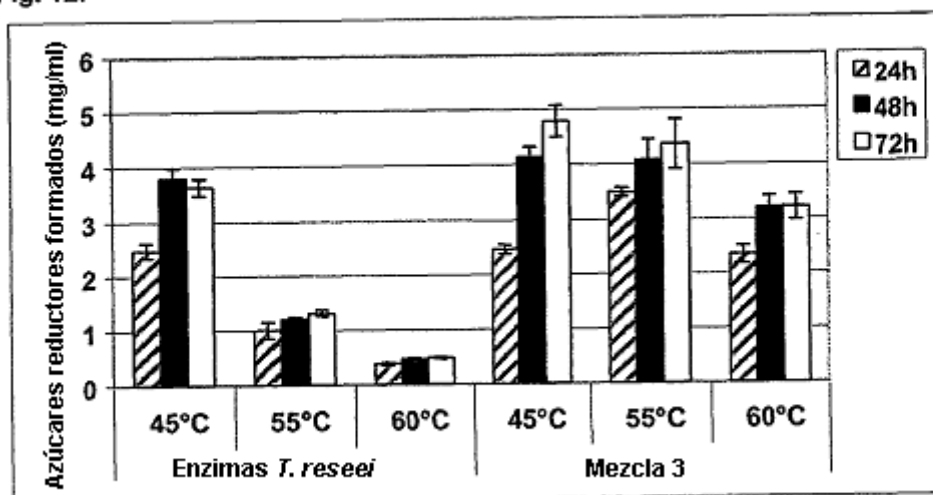


Fig. 13.

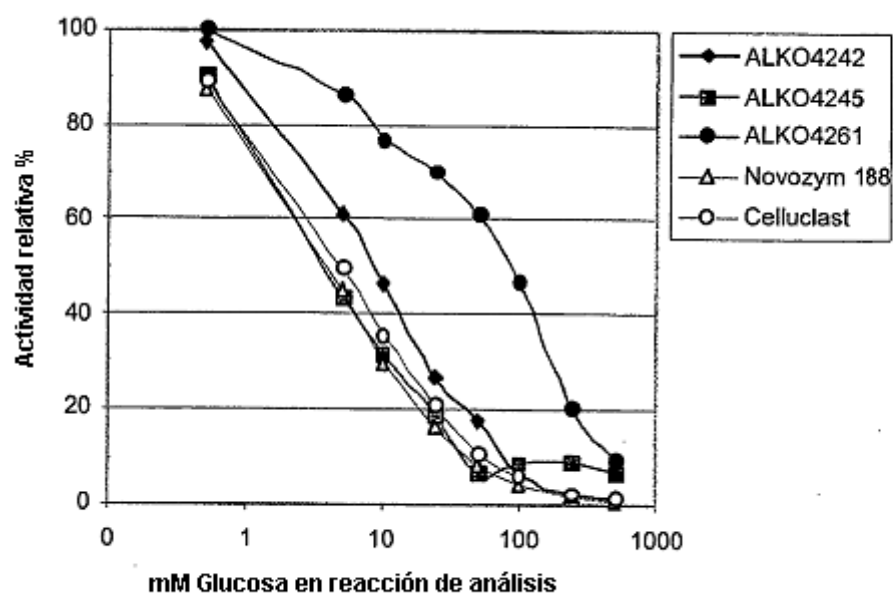


Fig. 14.

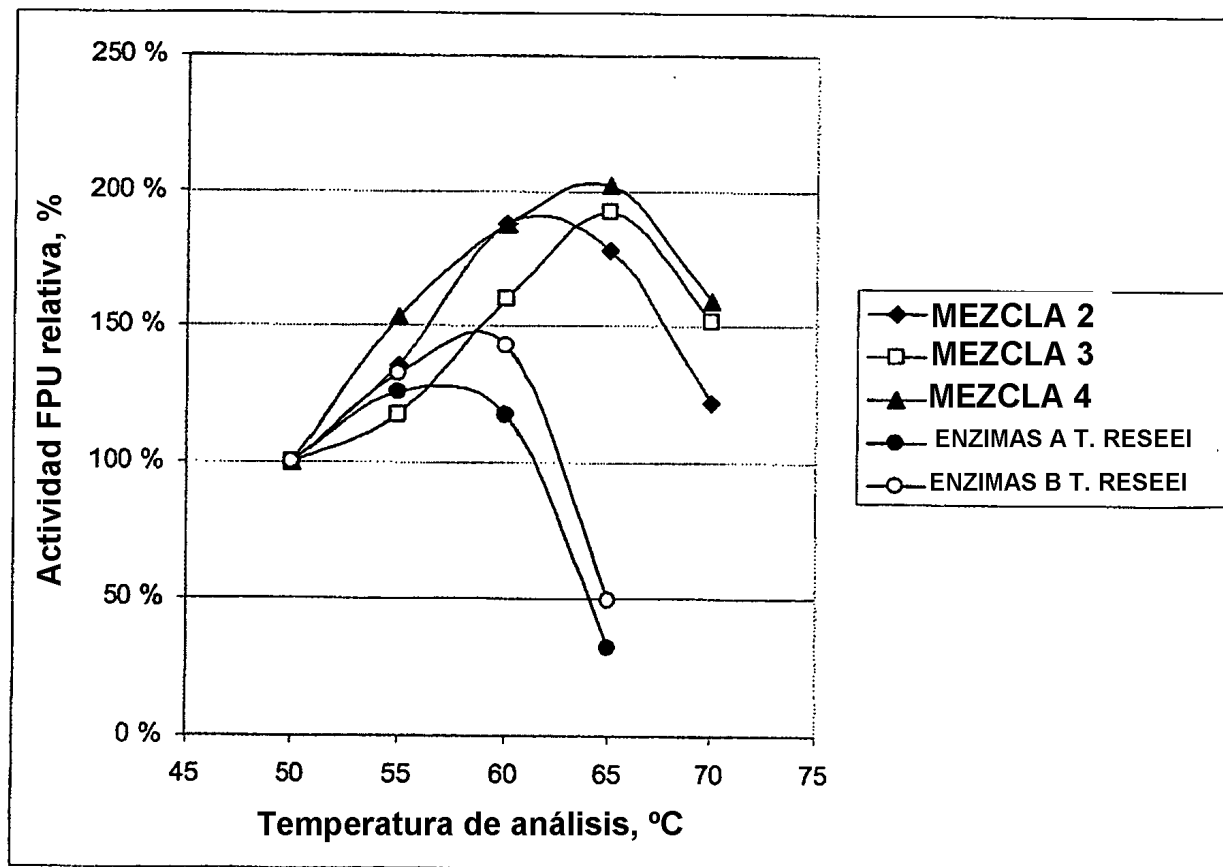


Fig. 15.

