

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 612 512**

51 Int. Cl.:

**C12N 9/42** (2006.01)

**C12P 7/10** (2006.01)

**C12P 7/06** (2006.01)

**C12P 19/14** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.12.2010 PCT/EP2010/070927**

87 Fecha y número de publicación internacional: **07.07.2011 WO11080317**

96 Fecha de presentación y número de la solicitud europea: **30.12.2010 E 10796442 (1)**

97 Fecha y número de publicación de la concesión europea: **02.11.2016 EP 2519630**

54 Título: **Procedimiento para tratar el material celulósico y enzimas CBHII/CEL6A útiles en el mismo**

30 Prioridad:

**30.12.2009 FI 20096412**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**17.05.2017**

73 Titular/es:

**ROAL OY (100.0%)  
Tykkimäentie 15b  
05200 Rajamäki, FI**

72 Inventor/es:

**PURANEN, TERHI;  
TOIKKA, SAULI;  
LANGFELDER, KIM y  
VEHMAANPERÄ, JARI**

74 Agente/Representante:

**CURELL AGUILÁ, Mireia**

**Observaciones :**

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

ES 2 612 512 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

Procedimiento para tratar el material celulósico y enzimas CBHII/CEL6A útiles en el mismo.

5 **Campo de la invención**

La presente invención se refiere a un procedimiento para el tratamiento de material celulósico con un enzima celobiohidrolasa CBHII/Cel6A fúngico o una preparación enzimática que comprende dicho enzima. El enzima resulta útil en diversas aplicaciones industriales, en particular en la producción de biocombustibles, en las que la producción de azúcares fermentables a partir de materia lignocelulósico en intervalos de temperatura moderada a elevada resulta ventajosa. La invención se refiere además a polipéptidos CBHII/Cel6A fúngicos y moléculas de ácidos nucleicos aisladas codificantes de dichos enzimas, a un vector recombinante, a células hospedadoras para producir dichos enzimas, a composiciones enzimáticas que comprenden dichos enzimas, así como a un procedimiento para la preparación de dichas composiciones. La invención se refiere además a diversos usos de dichos enzimas o composiciones enzimáticas, en los que se desea la conversión enzimática de material celulósico o lignocelulósico.

**Antecedentes de la invención**

La limitada disponibilidad de combustibles fósiles y las crecientes cantidades de CO<sub>2</sub> liberadas a partir de los mismos y que causan el fenómeno del efecto invernadero han hecho surgir la necesidad de utilizar biomasa como fuente renovable y limpia de energía. Una tecnología alternativa prometedora es la producción de biocombustibles, tales como etanol, butanol o propanol a partir de materiales celulósicos. En el sector del transporte los biocombustibles son hasta el momento la única opción que podría reducir las emisiones de CO<sub>2</sub> en un orden de magnitud. El etanol puede utilizarse en vehículos y sistemas de distribución existentes y, de esta manera, no requiere caras inversiones en infraestructuras. Los azúcares derivados de las materias primas renovables celulósicas y lignocelulósicas también pueden utilizarse como materias primas para una diversidad de productos químicos que pueden sustituir a los productos químicos procedentes del petróleo.

La mayoría de carbohidratos en las plantas se encuentran en forma de lignocelulosa, que consiste esencialmente de celulosa, hemicelulosa y lignina. En un procedimiento convencional de lignocelulosa a etanol, el material lignocelulósico en primer lugar se pretrata química o físicamente utilizando la hidrólisis ácida, la explosión de vapor, la expansión de fibras con amoníaco, la oxidación húmeda alcalina o el pretratamiento con ozono, para que la fracción de celulosa resulte más accesible a la hidrólisis enzimática. A continuación, la fracción de celulosa se hidroliza para obtener azúcares que pueden ser fermentadas por levaduras en etanol y destilarse para obtener etanol puro. Se obtiene lignina como coproducto principal que puede utilizarse como combustible sólido. En este procedimiento de hidrólisis y fermentación separados (HFS), la temperatura de la hidrólisis enzimática típicamente es más elevada que la de la fermentación. La utilización de enzimas termoestables en la hidrólisis ofrece beneficios potenciales, tales como tasas de reacción más altas a temperaturas elevadas, reducción de la carga enzimática debido a la actividad específica y vida útil más elevadas de los enzimas, mayor flexibilidad con respecto a la configuración del procedimiento y una mayor higiene.

Se sigue investigando para hacer que el procedimiento de producción de bioetanol resulte más económico. Una de las opciones es la simultaneidad de los procesos de sacarificación y fermentación (SFS). Los principales beneficios son la menor inhibición por parte del producto final de la hidrólisis enzimática y los menores costes de inversión. Los retos están en encontrar condiciones favorables, por ejemplo de temperatura y pH, para tanto la hidrólisis enzimática como la fermentación. En el bioprocedimiento consolidado (BPC), la cantidad de enzimas añadidos externamente puede reducirse significativamente mediante el aprovechamiento de un organismo fermentador o productor de etanol que sea capaz de producir un conjunto de enzimas lignocelulolíticos.

En los últimos años la ingeniería metabólica de los microorganismos utilizados en la producción de etanol ha mostrado avances significativos. Aparte de *Saccharomyces cerevisiae*, la atención se ha centrado en microorganismos tales como las especies bacterianas *Zymomonas* y *Escherichia coli*, y levaduras tales como *Pichia stipitis* y *Kluyveromyces fragilis* para la producción de etanol a partir de celulosa. En el procedimiento SFS, el inhibidor y la tolerancia de temperatura, así como la capacidad de utilizar múltiples azúcares, son propiedades importantes del microorganismo fermentador. Se han desarrollado levaduras manipuladas que son capaces de fermentar los azúcares pentosa xilosa y arabinosa además de la glucosa. Se han manipulado algunos microbios termofílicos, tales como *Thermoanaerobacterium saccharolyticum* o *Clostridium thermocellum* para fermentar azúcares, incluyendo xilosa, formando etanol a temperatura elevadas, de entre 50°C y 60°C (SFS termofílico o SFST). Estos organismos fermentadores también presentan potencial como BPC (Shaw et al., 2008).

La hidrólisis enzimática se considera la tecnología más prometedora para convertir la biomasa celulósica en azúcares fermentables. Sin embargo, la hidrólisis enzimática se utiliza únicamente en cantidad limitada a escala industrial y especialmente al utilizar material fuertemente lignificado, tal como madera o residuos agrícolas, la tecnología no resulta satisfactoria. Se han realizado esfuerzos por mejorar la eficiencia de la hidrólisis enzimática del material celulósico (Badger, 2002; Kurabi et al., 2005).

El documento WO 2001/060752 (Forskningcenter Risø, Dinamarca) describe un procedimiento continuo para convertir la biomasa lignocelulósica sólida en productos combustibles. Tras el pretratamiento mediante oxidación húmeda o explosión de vapor, la biomasa se separa parcialmente en celulosa, hemicelulosa y lignina y después se somete a hidrólisis parcial utilizando uno o más enzimas carbohidrasa (EC 3.2).

El documento WO 2002/024882 (Iogen Bio-Products Corp., CA) se refiere a un procedimiento de conversión de celulosa en glucosa mediante el tratamiento de un sustrato lignocelulósico pretratado con una mezcla de enzimas que comprende celulosa y una celobiohidrolasa I (CBHI) modificada que se ha obtenido mediante la inactivación de su dominio de unión a celulosa (DUC). El documento US2004/0005674 (Athenix Corp., EEUU) describe nuevas mezclas de enzimas que pueden utilizarse directamente sobre los sustratos de lignocelulosa. La mezcla sinérgica de enzimas contiene una celulosa y un enzima auxiliar, tal como celulasa, xilanas, ligninasa, amilasa, proteasa, lipidasa o glucuronidasa, o cualquier combinación de los mismos. La celulasa se considera que incluye los enzimas endoglucanasa (EG), beta-glucosidasa (BG) y celobiohidrolasa (CBH). El documento US2005/0164355 (Novozymes Biotech Inc., EEUU) describe un procedimiento para degradar el material lignocelulósico con uno o más enzimas celulolíticos seleccionados de entre EG, BG y CBH, y en presencia de por lo menos un surfactante. También pueden utilizarse enzimas adicionales tales como hemicelulasas, esterasa, peroxidasa, proteasa, lacasa o una mezcla de los mismos.

Los enzimas celulolíticos de origen fúngico mejor investigados y aplicados más ampliamente se han obtenido de *Trichoderma reesei* (el anamorfo de *Hypocrea jecorina*). También se han dado a conocer celulasas de hongos menos conocidos.

Hong et al. (2003a y 2003b) han caracterizado EG y CBHI de *Thermoascus aurantiacus* y la producción de los mismos en levaduras. Tuohy et al. (2002) describen tres formas de celobiohidrolasas, incluyendo CBHI y CBHII de *Talaromyces emersonii*.

Es conocida la utilización de la celobiohidrolasa I (CBHI), un miembro de la familia 7 de glucosil hidrolasa en la conversión enzimática de material celulósico, por ejemplo a partir del documento WO 03/000941 (Novozymes A/S, Dinamarca), que se refiere a enzimas CBHI obtenidos de diversos hongos. El documento WO 2005/074656 (Novozymes Inc., EEUU) da a conocer polipéptidos que presentan actividad celulolítica obtenidos de, por ejemplo, *Thermoascus aurantiacus*.

El documento WO 2007/071818 (Roal Oy, FI) describe la producción de hidrolizados de azúcares de material celulósico mediante conversión enzimática y preparaciones enzimáticas que comprenden dichos enzimas. Entre los enzimas útiles en el procedimiento se incluyen celobiohidrolasa, endoglucanasa, beta-glucosidasa y opcionalmente xilanas termoestables, obtenidas de *Thermoascus aurantiacus*, *Acremonium thermophilum* o *Chaetomium thermophilum*.

Las celobiohidrolasas II han sido dadas a conocer en varias solicitudes. El documento nº WO 2004/056981 (Novozymes A/S, Dinamarca) da a conocer polipéptidos que presentan actividad de celobiohidrolasas II y polinucleótidos codificantes de los polipéptidos, así como procedimientos para producir y utilizar los polipéptidos en aplicaciones, tales como en la producción de etanol. Se dan a conocer secuencias de ADN de longitud completa de *Aspergillus tubigenensis*, *Chaetomium thermophilum*, *Myceliophthora thermophila*, especies de *Thielavia*, *Acremonium thermophilum*, *Trichophaea saccata*, *Stibella anualata* y *Malbrancheae cinnamonea*. El documento EP 1578964 B1 (Novozymes A/S, Dinamarca) da a conocer la secuencia de aminoácidos de longitud completa de la CBHII de *C. thermophilum* y un polipéptido codificado por una secuencia de nucleótidos que se hibrida bajo condiciones restrictivas con un fragmento de la secuencia de nucleótidos codificante de dicho enzima.

El documento CN1757709 (Shandong Agricultural Univ., CN) da a conocer la secuencia de nucleótidos de un enzima CBHII termofílico de *Chaetomium thermophilum* CT2 y la expresión del mismo en la levadura *Pichia pastoris*. El enzima es capaz de convertir el material de fibra rechazado.

El documento WO 2006/074005 (Genencor Int., Inc., EEUU) da a conocer una variante del enzima CBHII/Cel6A de *Hypocrea jecorina* (*Trichoderma reesei*). El enzima variante resulta útil, por ejemplo, en la producción de bioetanol.

El documento WO 2007/094852 (Diversa Corp., EEUU; Verenium Corp., EEUU) da a conocer enzimas celulolíticos, ácidos nucleicos codificantes de los mismos y procedimientos para la producción y utilización de los mismos. El enzima puede ser una endoglucanasa, una celobiohidrolasa, una beta-glucosidasa, una xilanas, una mananasa, una beta-xilosidasa, una arabinofuranosidasa o una oligomerasa. El enzima y mezclas de enzimas resultan útiles, por ejemplo, en la producción de combustible o bioetanol.

El documento WO 2008/095033 (Syngenta, CH; Verenium Corp., EEUU) da a conocer enzimas que presentan actividad lignocelulolítica, incluyendo celobiohidrolasas útiles, por ejemplo, en la preparación de combustibles y el procesamiento de materiales de biomasa. El documento WO 2009/045627 (Verenium Corp., EEUU) da a conocer procedimientos para descomponer la hemicelulosa mediante la utilización de enzimas que presentan actividad de xilanas, mananasa y/o glucanasa y actividad y estabilidad incrementadas a pH y temperatura incrementados.

El documento WO 2009/089630 (Iogen Energy Corp., CA) da a conocer una variante de una celulasa de la familia 6 con inhibición reducida por la glucosa, que comprende una o más sustituciones de aminoácidos.

- 5 El documento WO 2009/059234 (Novozymes Inc., EEUU) da a conocer procedimientos de producción de material celulósico reducido en un metal redox-activo útil en la degradación o conversión de un material celulósico y en la producción de un producto de fermentación. La solicitud da a conocer, por ejemplo, un polipéptido CBHII de *Chaetomium thermophilum*.
- 10 El documento WO 2009/085868 (Novozymes A/S, Dinamarca) da a conocer polipéptidos, polinucleótidos codificantes del enzima y un procedimiento para producir un producto de fermentación, que comprende la sacarificación de un material celulósico con una composición de enzimas celulolíticas que comprende dicho polipéptido. Dichas celobiohidrolasas pueden obtenerse de *Trichoderma reesei*, *Hemicola insolens*, *Myceliophthora thermophila*, *Thielavia terrestris* y *Chaetomium thermophilum*. Los documentos WO 2006/074435 y US2006/218671
- 15 (Novozymes Inc., EEUU) dan a conocer secuencias de nucleótidos y de aminoácidos de la celobiohidrolasa Cel6A de *Thielavia terrestris*. El documento US 7220565 (Novozymes Inc., EEUU) da a conocer polipéptidos que presentan una actividad potenciadora de la celulolisis e identidad respecto a la secuencia de aminoácidos madura de *Myceliophthora thermophila* CBHII.
- 20 Los documentos US2007/0238155 y US2009/0280105 (Dyadic Int., Inc., EEUU) dan a conocer composiciones de enzimas que comprenden nuevos enzimas de *Chrysosporium lucknowense*, que comprenden, por ejemplo, los enzimas CBHIIa y CBHIIb asignados a la familia 6 de glucosil hidrolasas. Las composiciones de enzimas resultan eficaces en la hidrólisis del material lignocelulósico.
- 25 La secuencia genómica de *Neurospora crassa* OR74A se da a conocer en Galagan et al. (2003), incluyendo una secuencia del precursor de la exoglucanasa 2. Collins et al. (2003) dan a conocer la secuencia codificante de CBHII/Cel6A de *Talaromyces emersonii*.

30 Se espera que en el futuro próximo se incremente considerablemente el mercado de biocombustibles, tales como los combustibles para transporte renovables. Como resultado, existe un interés en rápido crecimiento en la utilización de materias primas alternativas para la producción de biocombustible. La fermentación de la biomasa celulósica presente en plantas y maderas o residuos municipales para producir etanol y otros alcoholes es una ruta atractiva a combustibles que complementen los combustibles fósiles. Una barrera a la producción de biocombustibles a partir de la biomasa celulósica y lignocelulósica es la robustez de las paredes celulares y la presencia de monómeros

35 sacáridos en forma de polímeros inaccesibles que requieren una gran cantidad de procesamiento para producir los sacáridos monoméricos disponibles para los microorganismos que se utilizan típicamente para producir alcohol mediante fermentación. De esta manera, existe una necesidad continua de nuevos procedimientos, así como de nuevos enzimas y mezclas de enzimas, que potencien la eficiencia de la degradación de los sustratos celulósicos y lignocelulósicos. En particular, se requieren enzimas y mezclas de enzimas que sean capaces de atacar los

40 diferentes enlaces glucosídicos del material celulósico cristalino y proporcionen de esta manera una hidrólisis prácticamente completa de los diversos materiales que deben tratarse. También existe una necesidad de enzimas que sean estables a las elevadas temperaturas del procedimiento, permitiendo de esta manera la utilización de una consistencia de elevada biomasa y que conduzca a concentraciones de azúcar y etanol elevadas. Este enfoque podría conducir a ahorros significativos de energía y de costes de inversión. La elevada temperatura también reduce

45 el riesgo de contaminación durante la hidrólisis. La presente invención presenta como objetivo satisfacer por lo menos partes de dichas necesidades.

### Sumario de la invención

50 El objetivo de la presente invención es proporcionar nuevos enzimas y composiciones enzimáticas para potenciar la eficiencia de la hidrólisis de la celulosa. Las celobiohidrolasas y en particular las celobiohidrolasas II obtenibles de *Acremonium thermophilum*, *Melanocarpus albomyces*, *Chaetomium thermophilum* o *Talaromyces emersonii* resultan útiles para hidrolizar y degradar el material celulósico. Los enzimas resultan cinéticamente muy eficaces en un intervalo amplio de temperaturas, y aunque presentan una actividad elevada a temperaturas elevadas, también son

55 muy eficientes a temperaturas de hidrólisis estándares. Lo anterior los convierte en extremadamente adecuados a la modificación de los procedimientos de hidrólisis de sustratos celulósicos que se llevan a cabo tanto a temperaturas convencionales como a temperaturas elevadas.

60 Todas las referencias a secuencias diferentes de SEC ID nº 11 y nº 12 o de las cepas de *A. thermophilum* depositadas diferentes de CBS 116240 o de las cepas de *E. coli* diferentes de DSM 22946 se proporcionan únicamente a título comparativo.

65 La presente invención se refiere a un procedimiento para el tratamiento de material celulósico con un polipéptido CBHII/Cel6A o con una preparación de enzimas que comprende dicho polipéptido o un microorganismo fermentador productor de dicho polipéptido, comprendiendo dicho procedimiento las etapas de: i) producción de un polipéptido CBHII/Cel6A de la invención o una preparación enzimática que comprende dicho polipéptido o un microorganismo

fermentador productor de dicho polipéptido; ii) hacer reaccionar el material celulósico con el polipéptido CBHII/Cel6A de la invención o la preparación enzimática que comprende dicho polipéptido o el microorganismo fermentador productor de dicho polipéptido, e iii) obtener material celulósico parcial o totalmente hidrolizado. Según una forma de realización, el polipéptido CBHII/Cel6A útil en dicho procedimiento presenta actividad de celobiohidrolasa y comprende una secuencia de aminoácidos que presenta una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 12, una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto al polipéptido de longitud completa de SEC ID nº 15 o una identidad de por lo menos 91% respecto al polipéptido de longitud completa de SEC ID nº 16. Según otra forma de realización, el polipéptido CBHII/Cel6A también puede ser un fragmento o variante que presenta propiedades similares, tal como una especificidad de sustrato similar y una dependencia o estabilidad frente al pH y la temperatura. La celobiohidrolasa CBHII/Cel6A útil en el procedimiento es un enzima celobiohidrolasa II de la familia 6 de glucosilhidrolasas y presenta tanto un plegamiento y estereoquímica conservados de la reacción de hidrólisis.

Las celobiohidrolasas CBHII/Cel6A pueden obtenerse de un género de *Acremonium*, *Melanocarpus*, *Chaetomium* o *Talaromyces*, más preferentemente de *A. thermophilum*, *M. albomyces*, *C. thermophilum* o *T. emersonii*, más preferentemente de las cepas depositadas *A. thermophilum* CBS 116240, *M. albomyces* CBS 685.95, *C. thermophilum* CBS 730.95 o *T. emersonii* DSM 2432.

La celobiohidrolasa CBHII/Cel6A aplicable en el procedimiento es capaz de hidrolizar diferentes materiales celulósicos a temperaturas moderadas a elevadas, en particular en combinación con otros enzimas utilizados en la hidrólisis de diversos materiales celulósicos o lignocelulósicos.

La celobiohidrolasa CBHII/Cel6A de la invención es aplicable en diversos usos, en particular en la producción de biocombustible.

También se dan a conocer nuevas celobiohidrolasas CBHII/Cel6A, que presentan actividad de celobiohidrolasa y comprenden una secuencia de aminoácidos que presenta una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 12, una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto al polipéptido de longitud completa de SEC ID nº 15 o una identidad de por lo menos 91% respecto al polipéptido de longitud completa de SEC ID nº 16. Según una forma de realización, el polipéptido CBHII/Cel6A también puede ser un fragmento o variante con propiedades similares, tal como una especificidad de sustrato similar y dependencia o estabilidad frente al pH y la temperatura. Dicha celobiohidrolasa CBHII/Cel6A es capaz de hidrolizar material celulósico a temperaturas moderadas a elevadas.

Dicho enzima se encuentra codificado por una molécula aislada de ácidos nucleicos, que comprende una secuencia polinucleotídica codificante de un polipéptido de la invención. Según una forma de realización, dicha molécula de ácidos nucleicos comprende la secuencia polinucleotídica definida en SEC ID nº 11, SEC ID nº 13, SEC ID nº 9 o SEC ID nº 10 o una subsecuencia de las mismas.

Según una forma de realización, la celobiohidrolasa CBHII/Cel6A puede encontrarse codificada por una molécula aislada de ácidos nucleicos que se hibrida bajo condiciones restrictivas con una secuencia polinucleotídica incluida en SEC ID nº 11, SEC ID nº 13, SEC ID nº 9, SEC ID nº 10, SEC ID nº 7, SEC ID nº 8 o una subsecuencia de las mismas.

Según una forma de realización, el enzima se encuentra codificado por un polinucleótido aislado incluido en el plásmido pALK2582 depositado en *Escherichia coli* bajo el número de acceso DSM 22946, el plásmido pALK2581 depositado en *E. coli* bajo el número de acceso DSM 22945, el plásmido pALK2904 depositado en *E. coli* bajo el número de acceso DSM 22947 o el plásmido pALK3006 depositado en *E. coli* bajo el número de acceso DSM 23185.

La celobiohidrolasa CBHII/Cel6A de la invención puede producirse a partir de un vector de expresión recombinante que comprende la molécula de ácidos nucleicos o secuencia de nucleótidos codificante de dicha celobiohidrolasa. Dicho polipéptido celobiohidrolasa puede producirse en un hospedador heterólogo, preferentemente en un hospedador microbiano.

Una molécula aislada de ácidos nucleicos que comprende una secuencia polinucleotídica codificante de un polipéptido fúngico CBHII/Cel6A seleccionado de entre el grupo que consiste en:

(a) una molécula de ácidos nucleicos o secuencia polinucleotídica codificante de un polipéptido que presenta actividad de celobiohidrolasa y que comprende la secuencia de aminoácidos de longitud completa ilustrada en SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 o SEC ID nº 16, o un fragmento o variante de la misma que presenta propiedades similares.

(b) una molécula de ácidos nucleicos o secuencia polinucleotídica codificante de un polipéptido que presenta

actividad de celobiohidrolasa y una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 12, una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 15 o una identidad de por lo menos 91% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 16, o un fragmento o variante de la misma que presenta propiedades similares,

(c) una molécula de ácidos nucleicos que comprende la secuencia codificante de la secuencia polinucleotídica ilustrada como SEC ID nº 11, SEC ID nº 13, SEC ID nº 9 o SEC ID nº 10,

(d) una molécula de ácidos nucleicos que comprende la secuencia codificante de la secuencia polinucleotídica contenida en DSM 22946, DSM 22945, DSM 22947 o DSM 23185,

(e) una molécula de ácidos nucleicos la secuencia codificante de la cual difiere de la secuencia codificante de una molécula de ácidos nucleicos según cualquiera de (c) a (d) debido a la degeneración del código genético, y

(f) se da a conocer además una molécula de ácidos nucleicos que se hibrida bajo condiciones restrictivas con una molécula de ácidos nucleicos contenida en DSM 22946, DSM 22945, DSM 22947 o DSM 23185, y codificante de un polipéptido que presenta actividad de celobiohidrolasa y una secuencia de aminoácidos que muestra una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa tal como se ilustra en SEC ID nº 12, una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 15 o una identidad de por lo menos 91% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 16, o un fragmento o variante de la misma que presenta propiedades similares.

La invención se refiere además a un vector de expresión recombinante que comprende la molécula de ácidos nucleicos o secuencia polinucleotídica de la invención ligada operablemente a secuencias reguladoras capaces de dirigir la expresión del gen codificante de la celobiohidrolasa CBHII/Cel6A de la invención y la producción de dicha celobiohidrolasa CBHII/Cel6A en un hospedador adecuado.

La invención se refiere además a una célula hospedadora que comprende el vector de expresión recombinante tal como se ha indicado anteriormente. Preferentemente, el hospedador es un hospedador microbiano, tal como un hospedador hongo filamentoso. Son hospedadores preferentes, *Trichoderma*, *Aspergillus*, *Fusarium*, *Humicola*, *Chrysosporium*, *Neurospora*, *Rhizopus*, *Penicillium* y *Mortierella*. Más preferentemente, el hospedador es *Trichoderma* o *Aspergillus*, más preferentemente el hongo filamentoso *T. reesei*.

La presente invención se refiere a un procedimiento de producción de un polipéptido de la invención que presenta actividad de celobiohidrolasa, comprendiendo dicho procedimiento las etapas de cultivar la célula hospedadora de la invención y recuperar el polipéptido. También comprendido dentro de la invención se encuentra un polipéptido que presenta actividad de celobiohidrolasa codificado por la molécula de ácidos nucleicos de la invención y que es obtenible mediante el procedimiento indicado anteriormente.

La invención se refiere además a un procedimiento para obtener una preparación de enzimas que comprende las etapas de cultivar una célula hospedadora de la invención y preparar el caldo de cultivo completo, o separar las células del medio de cultivo agotado y obtener el sobrenadante. También comprendida dentro de la invención se encuentra comprendida una preparación enzimática obtenible mediante el procedimiento indicado anteriormente. La invención se refiere además a una preparación enzimática que comprende la celobiohidrolasa CBHII/Cel6A de la invención.

La preparación enzimática puede comprender además otros enzimas seleccionados de entre el grupo: celobiohidrolasa, endoglucanasa, beta-glucosidasa, beta-glucanasa, xiloglucanasa, xilanasa, beta-xilosidasa, mananasa, beta-manosidasa,  $\alpha$ -glucuronidasa, acetil-xilán-esterasa,  $\alpha$ -arabinofuranosidasa,  $\alpha$ -galactosidasa, pectinasa, que implica endo- y exo- $\alpha$ -L-arabinasas,  $\alpha$ -galactosidasa, endo- y exo-galactoronasa, endopectiniliasa, pectato liasa y pectinesterasa, fenol-esterasa, ligninasa que implica lignina peroxidasa, peroxidasa dependiente de manganeso, enzima generador de H<sub>2</sub>O<sub>2</sub> y lacasa con o sin un mediador.

La preparación enzimática puede encontrarse en forma de un caldo de cultivo completo o medio de cultivo agotado. Puede encontrarse en forma de líquido, polvos o granulado.

También se encuentra comprendida dentro de la invención la utilización del polipéptido CBHII/Cel6A o preparación enzimática de la invención para la producción de biocombustible, para detergentes, para tratar fibras, para tratar alimentos o piensos, para pasta ("pulp") o papel, para bebidas o para cualesquiera aplicaciones que implican hidrólisis o la modificación del material celulósico.

En particular, el polipéptido CBHII/Cel6A o composición enzimática que comprende dicho polipéptido resulta útil para la producción de biocombustible.

### Breve descripción de los dibujos

5 La figura 1 representa esquemáticamente los casetes de expresión utilizados en la transformación de protoplastos de *Trichoderma reesei* para producir en exceso las proteínas CBHII/Cel6A recombinantes. Los genes *cbh2/cel6A* se encontraban bajo el control del promotor *cbh1/cel7A* de *T. reesei* (*p cbh1*) y la terminación de la transcripción se garantiza mediante la utilización de la secuencia terminadora *cbh1/cel7A* de *T. reesei* (*t cbh1*). Se incluyó el gen *amdS* como marcador de transformación.

15 La figura 2 representa la determinación de dependencia del pH para las composiciones enzimáticas (100 µg de proteína en la reacción) que comprenden las celobiohidrolasas CBHII/Cel6A recombinantes de *Acremonium thermophilum* CBS 116240, *Melanocarpus albomyces* CBS 685.95, *Chaetomium thermophilum* CBS 730.95 y *Talaromyces emersonii* DSM 2432. La hidrólisis se llevó a cabo en celulosa Avicel Ph 101 en un intervalo de pH de 3 a 10 a 50°C durante 21 horas. Se determinó la formación de azúcares reductores mediante el procedimiento de la hidrazida de ácido para-hidroxibenzoico (PAHBAH) (Lever, 1972) utilizando una curva estándar de celobiosa.

20 La figura 3 representa la determinación de la estabilidad térmica de las composiciones enzimáticas (100 µg de proteína en la reacción) que comprenden las celobiohidrolasas CBHII/Cel6A recombinantes de *Acremonium thermophilum* CBS 116240, *Melanocarpus albomyces* CBS 685.95, *Chaetomium thermophilum* CBS 730.95 y *Talaromyces emersonii* DSM 2432. La hidrólisis se llevó a cabo en celulosa Avicel Ph 101 en un intervalo de temperaturas de 40°C a 80°C a pH óptimo de las composiciones enzimáticas durante 21 horas. Se determinó la formación de azúcares reductores mediante el procedimiento de la hidrazida de ácido para-hidroxibenzoico (PAHBAH) (Lever, 1972) utilizando una curva estándar de celobiosa.

30 La figura 4 representa los resultados de la hidrólisis de material de madera dura mediante explosión con vapor realizada con mezclas enzimáticas que comprenden la celobiohidrolasa CBHII/Cel6A de la invención. El sustrato de madera dura se hidrolizó utilizando diferentes mezclas enzimáticas a una dosis de 5 mg de proteína por cada gramo de sólidos totales tanto a 55°C como a 37°C. La composición de la mezcla enzimática termofílica (MEZCLA 2) y las mezclas enzimáticas mesofílicas (MEZCLA ENZIMAS T. REESEI y MEZCLA ACC) que comprende *At\_ALKO4245\_Cel6A*, *Ma\_ALKO4237\_Cel6A* o *Ct\_ALKO4265\_Cel6A* se describen en mayor detalle en el Ejemplo 4. Las muestras de matraces de agitación por duplicado se extrajeron tras un tiempo de hidrólisis de 72 horas y se cuantificaron mediante HPLC, en la que se determinó la concentración de glucosa y xilosa. Los resultados de los blancos de sustrato, que contenían tampón en lugar de la muestra enzimática, se restaron de los resultados obtenidos con las mezclas enzimáticas. Se presenta la concentración agrupada de glucosa y xilosa.

40 La figura 4A representa los resultados de la hidrólisis de madera dura mediante explosión con vapor realizada a 55°C con una mezcla enzimática termofílica (MEZCLA 2) complementada con *At\_ALKO4245\_Cel6A* (MEZCLA 2\_AT), *Ma\_ALKO4237\_Cel6A* (MEZCLA 2\_MA) o *Ct\_ALKO4265\_Cel6A* (MEZCLA 2\_CT).

45 La figura 4B representa los resultados de la hidrólisis de madera dura mediante explosión con vapor realizada a 37°C con una mezcla enzimática mesofílica (MEZCLA ENZIMAS T. REESEI) complementada con *At\_ALKO4245\_Cel6A* (MEZCLA TR\_AT), *Ma\_ALKO4237\_Cel6A* (MEZCLA TR\_MA) o *Ct\_ALKO4265\_Cel6A* (MEZCLA TR\_CT).

50 La figura 4C representa los resultados de la hidrólisis de madera dura mediante explosión con vapor realizada a 37°C con una mezcla enzimática mesofílica (MEZCLA ACC) complementada con *At\_ALKO4245\_Cel6A* (MEZCLA ACC\_AT), *Ma\_ALKO4237\_Cel6A* (MEZCLA ACC\_MA) o *Ct\_ALKO4265\_Cel6A* (MEZCLA ACC\_CT).

55 La figura 5 representa los resultados de la hidrólisis de mazorcas de maíz mediante explosión con vapor realizada con mezclas enzimáticas que comprenden una celobiohidrolasa CBHII/Cel6A de la invención. El sustrato de mazorcas de maíz se hidrolizó utilizando diferentes mezclas enzimáticas a una dosis de 5 mg de proteína por cada gramo de sólidos totales a 55°C. La composición de la mezcla enzimática termofílica (MEZCLA 2) que comprendía el polipéptido *At\_ALKO4245\_Cel6A* de la invención (MEZCLA 2\_AT) se describe con mayor detalle en el Ejemplo 5. Las muestras de matraces de agitación por duplicado se extrajeron tras un tiempo de hidrólisis de 72 horas y se cuantificó mediante HPLC, en la que se determinó la concentración de glucosa y xilosa. Los resultados de los blancos de sustrato, que contenían tampón en lugar de la muestra de enzima, se restaron de los resultados obtenidos con las mezclas enzimáticas. Se presenta la concentración agrupada de glucosa y xilosa.

### Listado de secuencias

SEC ID nº 1: secuencia del cebador oligonucleótido CBH\_1S.

65 SEC ID nº 2: secuencia del cebador oligonucleótido CBH\_1AS.

SEC ID nº 3: secuencia del cebador oligonucleótido CBH\_8.

SEC ID nº 4: secuencia del cebador oligonucleótido CBH\_9.

5 SEC ID nº 5: secuencia del cebador oligonucleótido Te\_CBH\_A.

SEC ID nº 6: secuencia del cebador oligonucleótido Te\_CBH\_B.

10 SEC ID nº 7: secuencia del fragmento de PCR obtenido de *Acremonium thermophilum* AKO4245 (CBS 116240) utilizando los cebadores CBH\_1S y CBH\_1AS.

SEC ID nº 8: secuencia del fragmento de PCR obtenido de *Melanocarpus albomyces* ALKO4237 (CBS 685.95) utilizando los cebadores CBH\_1S y CBH\_1AS.

15 SEC ID nº 9: secuencia de nucleótidos del gen *cbh2/cel6A* de *Chaetomium thermophilum* ALKO4265 (CBS 730.95).

SEC ID nº 10: secuencia de nucleótidos del gen *cbh2/cel6A* de *Talaromyces emersonii* RF8069 (DSM 2432).

20 SEC ID nº 11: secuencia de nucleótidos del gen *cbh2/cel6A* de *Acremonium thermophilum* ALKO4245 (CBS 116240).

SEC ID nº 12: secuencia de aminoácidos deducida de *Acremonium thermophilum* ALKO4245 (CBS 116240) CBHII/Cel6A.

25 SEC ID nº 13: secuencia de nucleótidos del gen *cbh2/cel6A* de *Melanocarpus albomyces* ALKO4237 (CBS 685.95).

SEC ID nº 14: secuencia de aminoácidos deducida de *Melanocarpus albomyces* ALKO4237 (CBS 685.95) CBHII/Cel6A.

30 SEC ID nº 15: secuencia de aminoácidos deducida de *Chaetomium thermophilum* ALKO4265 (CBS 730.95) CBHII/Cel6A.

SEC ID nº 16: secuencia de aminoácidos deducida de *Talaromyces emersonii* RF8069 (DSM 2432) CBHII/Cel6A.

### 35 Depósitos

Se depositó *Acremonium thermophilum* ALKO4245 en el Centraalbureau Voor Schimmelcultures, Uppsalalaan 8, 3584 CT, Utrecht, Países Bajos, el 20 de septiembre de 2004 y número de acceso asignado CBS 116240.

40 Se depositó *Chaetomium thermophilum* ALKO4265 en el Centraalbureau Voor Schimmelcultures, Oosterstraat 1, 3742 SK BAARN, Países Bajos, el 8 de noviembre de 1995 y número de acceso asignado CBS 730.95. Tras la terminación del periodo de depósito actual, las muestras se almacenaron bajo contrato de manera que la cepa se encuentre disponible más allá del tiempo de la patente.

45 Se depositó *Melanocarpus albomyces* ALKO4237 en el Centraalbureau Voor Schimmelcultures, Oosterstraat 1, 3740 AG BAARN, Países Bajos, el 11 de octubre de 1995 y número de acceso asignado CBS 685.95. Tras la terminación del periodo de depósito actual, las muestras se almacenaron de manera que la cepa se encuentre disponible más allá del tiempo de la patente.

50 La cepa de *Escherichia coli* RF8175, incluyendo el plásmido pALK2582, se depositó en el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Inhoffenstrasse 7B, D-38124 Braunschweig, Alemania, el 9 de septiembre de 2009 y número de acceso asignado DSM 22946.

55 La cepa de *E. coli* RF8174, incluyendo el plásmido pALK2581, se depositó en el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Inhoffenstrasse 7B, D-38124 Braunschweig, Alemania, el 9 de septiembre de 2009 y número de acceso asignado DSM 22945.

60 La cepa de *E. coli* RF8214, incluyendo el plásmido pALK2904, se depositó en el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Inhoffenstrasse 7B, D-38124 Braunschweig, Alemania, el 9 de septiembre de 2009 y número de acceso asignado DSM 22947.

La cepa de *E. coli* RF8333, incluyendo el plásmido pALK3006, se depositó en el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Inhoffenstrasse 7B, D-38124 Braunschweig, Alemania, el 10 de diciembre de 2009 y número de acceso asignado DSM 23185.

65



### Descripción detallada de la invención

La presente invención proporciona un procedimiento para el tratamiento de material celulósico con un polipéptido celobiohidrolasa CBHII/Cel6A o una composición enzimática que comprende dicho polipéptido. La invención proporciona además polipéptidos celobiohidrolasa CBHII/Cel6A fúngicos y composiciones enzimáticas que comprenden dichos polipéptidos CBHII/Cel6A, los cuales muestran una amplia especificidad de sustrato y son estables en intervalos de pH y temperatura amplios. Presentan un buen rendimiento a temperaturas tanto moderadas como elevadas. Según una forma de realización, los polipéptidos celobiohidrolasa CBHII/Cel6A son estables durante como mínimo 21 horas hasta 70°C, preferentemente en un intervalo de temperaturas de 40°C a 60°C. Los polipéptidos y composiciones enzimáticas que comprenden dichos polipéptidos resultan ideales en aplicaciones diferentes que requieren una hidrólisis eficiente de materiales celulósicos o lignocelulósicos complejos, en los que la celulosa es uno de los componentes principales. Los polipéptidos y composiciones enzimáticas resultan útiles, por ejemplo, en la hidrólisis de material celulósico con el fin de producir monómeros sacáridos a partir de materia prima polimérica, que después puede ser fermentada por microorganismos en la producción de biocombustible. De esta manera, la presente invención proporciona celobiohidrolasas CBHII/Cel6A alternativas para la utilización en biocombustible y otras aplicaciones. Las celobiohidrolasas CBHII/Cel6A fúngicas pueden producirse en hospedadores fúngicos de alto rendimiento con o sin procesamiento posterior, por ejemplo la separación de caldo de fermentación y micelios es fácil de llevar a cabo, o en organismos fermentadores.

"Celulosa" o "material celulósico" tal como se utilizan en la presente memoria se refieren a cualquier material que comprende celulosa como componente significativo. La celulosa es el componente estructural principal de las plantas superiores. Proporciona a las células vegetales una resistencia ténsil elevada, ayudándolas a resistir el estrés mecánico y la presión osmótica. La celulosa es un  $\beta$ -1,4-glucano compuesto de cadenas lineales de residuos de glucosa unidos mediante enlaces  $\beta$ -1,4-glucosídicos. La celobiosa es la unidad repetitiva más pequeña de la celulosa. Entre los ejemplos de material celulósico se incluyen fibras textiles derivadas de, por ejemplo, algodón, lino, cáñamo, yute y las fibras celulósicas artificiales, tales como Modal, viscosa y Lyocell.

El término "celulosa" o "material celulósico" se refiere también a "lignocelulosa" o "material lignocelulósico", que comprende celulosa como componente significativo.

En las paredes celulares, la celulosa se encuentra empaquetada en láminas de diversa orientación, las cuales se encuentran incluidas en una matriz de hemicelulosa, pectina y/o polímeros de unidades de fenol-propanol, tales como lignina, para formar "material lignocelulósico" o "lignocelulosa". La lignocelulosa es físicamente dura, densa e inaccesible. Entre los materiales que contienen lignocelulosa se incluyen, por ejemplo, materiales vegetales tales como madera, incluyendo madera dura y madera blanda, virutas de madera dura y de madera blanda, pulpa de madera, serrín y residuos industriales de silvicultura y de madera, cultivos herbáceos, biomasa agrícola tal como paja de cereales, pulpa de remolacha azucarera, fibra de maíz, rastrojo y mazorcas de maíz, beta-glucanos de cereales, bagazo de caña azucarera, tallos, hojas, vainas, cáscaras y similares; productos residuales, tales como residuos urbanos sólidos, residuos de papel de periódicos y de papel de oficina, fibras de desecho, lodos de papel, residuos triturados de, por ejemplo, cereales; cultivos energéticos específicos (por ejemplo sauce, chopo, pastos altos ('switchgrass') o alpiste, y similares).

El material celulósico y lignocelulósico es degradado naturalmente por varios organismos, entre ellos bacterias y hongos que producen enzimas capaces de hidrolizar polímeros de carbohidratos. La degradación habitualmente requiere la acción de muchos enzimas que típicamente actúan secuencial o simultáneamente. La conversión biológica de la celulosa en glucosa generalmente requiere tres tipos principales de enzimas hidrolíticas: (1) endoglucanasas (EG), que cortan enlaces beta-1,4-glucosídicos internos principalmente en las regiones amorfas de la celulosa, (2) exoglucanasas o celobiohidrolasas (CBH) que cortan el disacárido celobiosa desde el extremo reductor o no reductor de la cadena polimérica de la celulosa cristalina, (3) beta-1,4-glucosidasas (BG), que hidrolizan la celobiosa y otros celo-oligosacáridos cortos en glucosa. La glucosa y la celobiosa actúan como inhibidores de producto final de la reacción de hidrólisis.

"Celulasa" o "enzima celulolítica" es un enzima que presenta "actividad de celulasa" o "actividad celulolítica", que se refiere a que es capaz de hidrolizar material celulósico. Uno de los sistemas enzimáticos celulolíticos más estudiados es el del hongo filamentoso "*Trichoderma reesei*", que es conocido que muestra por lo menos dos CBH, 8 EG y 5 BG.

Los "enzimas lignocelulolíticos" son enzimas que presentan "actividad de lignocelulasa" o "actividad lignocelulolítica", lo que significa que son capaces de hidrolizar material lignocelulósico, tal como celulosas, hemicelulosas o derivados de los mismos en sacáridos más pequeños.

Los "beta-glucanos" son beta-D-glucanos de enlace mixto (1 $\rightarrow$ 3),(1 $\rightarrow$ 4). Son frecuentes en, por ejemplo, cereales.

La "beta-glucanasa" tal como se utiliza en la presente memoria se refiere a enzimas que pueden cortar por lo menos parcialmente los enlaces en el beta-glucano.

La "celobiohidrolasa" o "CBH" tal como se utiliza en la presente memoria se refiere a enzimas que cortan los enlaces 1,4-beta-D-glucosídicos de polímeros tales como la celulosa desde el extremo reductor o no reductor y producen principalmente celobiosa. También se denominan exoglucanasas o 1,4-beta-D-glucano celobiohidrolasas o celulosa 1,4-beta-celobiosidasas. Las CBH presentan una estructura modular que consiste en dominios diferentes, tales como un dominio catalítico y un dominio de unión a celulosa N- o C-terminal (DUC). También existen celebiohidrolasas en la naturaleza que no presentan DUC. Las CBH también pueden presentar dominios adicionales de función desconocida. Los diferentes dominios habitualmente se unen entre sí mediante un conector O-glucosilado o péptido bisagra rico en glicina, prolina, serina o treonina.

El término "celobiohidrolasa II" o "celobiohidrolasa CBHII" o "celobiohidrolasa CBHII/Cel6A" o "celulasa CBHII/Cel6A" o "polipéptido CBHII/Cel6A" o "enzima CBHII/Cel6A" se refiere, en relación a la presente invención, a un enzima 1,4-β-D-glucano celobiohidrolasa clasificado como EC 3.2.1.91 por la nomenclatura de la unión internacional de bioquímica y biología molecular (IUBMB, por sus siglas en inglés; ver <http://www.iubmb.org>). Basándose en sus similitudes estructurales, las celobiohidrolasas II o las celobiohidrolasas CBHII/Cel6A de la presente invención se clasifican en la familia 6 de glucosil hidrolasas (GH6), las cuales presentan secuencias de aminoácidos y estructuras tridimensionales similares (Henrissat, 1991; Henrissat y Bairoch, 1993, 1996; Henrissat et al., 1998; ver también <http://www.cazy.org/fam/GH6.html>). Debido a que existe una relación directa entre la secuencia y las similitudes de pliegue, dicha clasificación se cree que refleja las características estructurales de los enzimas mejor que únicamente su especificidad de sustrato, ayuda a revelar las relaciones evolutivas entre los enzimas y proporciona una cómoda herramienta para obtener información sobre los mecanismos.

La expresión "actividad de celobiohidrolasa" o "actividad de CBH" tal como se utiliza en la invención se refiere a la actividad hidrolítica que actúa sobre los enlaces 1,4-beta-D-glucosídicos en la celulosa, celotriosa, celotetraosa o en cualquier polímero de glucosa unido mediante enlaces beta-1,4 que libera principalmente celobiosa del extremo reductor o no reductor de la cadena polimérica. La rotura enzimática de un enlace glucosídico es un proceso estereoselectivo, en el que la configuración en entorno al centro anomérico (carbono C1) puede invertirse o conservarse. Ambos mecanismos contienen una pareja de residuos ácido carboxílico dispuestos en cada extremo del enlace que debe cortarse. Los enzimas inversores utilizan un mecanismo de único desplazamiento, mientras que los enzimas conservadores implican una reacción de doble desplazamiento (Sinnott, 1990; Withers y Aebersold, 1995). Habitualmente se determina el curso estereoquímico de la hidrólisis mediante RMN de protones, en la que los protones α- y β-anoméricos proporcionan diferentes desplazamientos químicos (Withers et al., 1986).

La expresión "actividad de celobiohidrolasa II" o "actividad de CBHII/Cel6A" tal como se utiliza en la invención se refiere a actividad hidrolítica que actúa sobre los enlaces 1,4-beta-D-glucosídicos en la celulosa, celotriosa, celotetraosa o en cualquier polímero de glucosas unidas mediante enlace beta-1,4, liberando principalmente celobiosa del extremo no reductor de la cadena polimérica. El mecanismo de reacción de las celobiohidrolasas CBHII/Cel6A es la inversión.

Los procedimientos de análisis de la actividad de celulasa son bien conocidos en la literatura y se hace referencia a los mismos en, por ejemplo, Ghose (1987), Tomme et al. (1988) y van Tilbeurgh et al. (1988). La actividad global de celulasa se mide comúnmente como actividad degradadora de papel de filtro (UPF). La actividad de celobiohidrolasa puede analizarse utilizando celodextrinas solubles pequeñas y sus glucósidos cromogénicos, tales como 4-metilumbeliferil-beta-D-glucósidos. Las celobiohidrolasas CBHII/Cel6A pueden identificarse basándose en el resultado de la reacción de hidrólisis; los enzimas CBHII/Cel6A no pueden cortar el enlace heterosídico de los oligosacáridos cromogénicos pequeños (van Tilbeurgh et al., 1988). La actividad de celobiohidrolasa también puede analizarse en celulosa microcristalina, tal como Avicel Ph 101, tal como se utiliza en el Ejemplo 3. La formación de azúcares reductores solubles tras la hidrólisis puede determinarse mediante el procedimiento de la hidrazida de ácido para-hidroxibenzoico (PAHBAH) (Lever, 1972) utilizando una curva estándar de celobiosa, el procedimiento de Somogyi-Nelson (Somogyi, 1952), el procedimiento de ferricianuro alcalino (Robyt y Whelan, 1972), el procedimiento de 2,2'-bichinoninato (Waffenschmidt y Jaenicke, 1987) o el procedimiento del ácido dinitrosalicílico (DNS) de Miller (1959). Entre otros sustratos celulósicos se incluyen, por ejemplo, la celulosa Solka-Floc o la celulosa hinchada con ácido fosfórico (Karlsson et al., 2001).

La celobiohidrolasa II también puede identificarse en un ensayo de transferencia western o un ensayo ELISA utilizando anticuerpos policlonales o monoclonales generados contra el polipéptido CBHII/Cel6A purificado.

De esta manera, la expresión "celobiohidrolasa CBHII/Cel6A" se refiere a enzimas celobiohidrolasa II que son elementos de la familia 6 de glucosil-hidrolasas, que presentan tanto un pliegue conservado como la estereoquímica de la reacción de hidrólisis.

La expresión "temperatura moderada" o "temperatura convencional" en el contexto de la presente invención se refiere a temperaturas utilizadas comúnmente en la hidrólisis de la celulosa y correspondientes a las temperaturas óptimas o estabildades térmicas de los enzimas utilizados en dichos procedimientos. De esta manera, las expresiones se refieren a intervalos de temperatura de entre 30°C y 45°C.

La expresión "temperatura elevada" o "temperatura alta" se refiere a intervalos de temperatura de entre 45°C y 70°C.

En procedimientos de hidrólisis a corto plazo, los enzimas pueden resultar eficaces incluso hasta 80°C. Los enzimas activos o estables en dichos intervalos de temperatura elevada también se denominan enzimas "termoestables" o "termofílicos".

5 Las cepas de microorganismo capaces de producir polipéptido celobiohidrolasa CBHII/Cel6A o actividad de celobiohidrolasa CBHII/Cel6A pueden cribarse en diferentes sustratos. En primer lugar, las cepas seleccionadas se cultivan en un medio adecuado. Tras producir una cantidad suficiente de una celobiohidrolasa interesante, el enzima puede aislarse o purificarse y sus propiedades pueden caracterizarse más completamente. Alternativamente, pueden aislarse genes codificantes de celobiohidrolasas o celobiohidrolasas CBHII/Cel6A y compararse la  
10 secuencia de aminoácidos codificada por los genes, con las secuencias de aminoácidos de las celobiohidrolasas CBHII/Cel6A aislarse y caracterizarse en el Ejemplo 1.

Los enzimas celobiohidrolasa producidos, en particular los enzimas celobiohidrolasa CBHII/Cel6A, pueden purificarse mediante la utilización de procedimientos convencionales de química enzimática, tal como la preparación de sales, la ultrafiltración, la cromatografía de intercambio iónico, la cromatografía de afinidad, la filtración en gel y la cromatografía de intercambio hidrofóbica. La purificación puede monitorizarse a partir de la determinación de la proteína, ensayos de actividad enzimática y mediante electroforesis en gel de SDS-poliacrilamida. Puede determinarse la actividad enzimática y estabilidad del enzima purificado a diversos valores de temperatura y pH, así como la masa molecular y el punto isoeléctrico. Alternativamente, pueden identificarse las propiedades de las celobiohidrolasas CBHII/Cel6A de la invención mediante la producción de los enzimas en un hospedador recombinante y la purificación y caracterización del enzima celobiohidrolasa CBHII/Cel6A recombinante. Además, las propiedades de la preparación de enzima recombinante que comprende la celobiohidrolasa CBHII/Cel6A de la invención como uno de los componentes enzimáticos principales puede caracterizarse tal como se indica en el  
15 Ejemplo 3.

La masa molecular de la celobiohidrolasa CBHII/Cel6A purificada puede determinarse a partir de la espectrometría de masas o en SDS-PAGE según Laemmli (1970). La masa molecular también puede predecirse a partir de la secuencia de aminoácidos del enzima utilizando la herramienta pI/MW en el servidor ExPASy ([http://expasy.org/tools/pi\\_tool.html](http://expasy.org/tools/pi_tool.html); Gasteiger et al., 2003).  
20

La dependencia de la temperatura o la termoestabilidad del enzima celobiohidrolasa CBHII/Cel6A puede determinarse en un tampón adecuado a diferentes temperaturas mediante la utilización de, por ejemplo, celulosa Avicel como sustrato, tal como se indica en el Ejemplo 3, o mediante la utilización de otros sustratos y sistemas de tampón descritos en la literatura. La determinación de la dependencia del pH y de la estabilidad frente al pH puede llevarse a cabo en un tampón adecuado a diferentes valores de pH siguiendo la actividad sobre un sustrato  
25 celulósico.

El pI puede determinarse mediante enfoque isoeléctrico sobre un gel de gradiente de pH inmovilizado compuesto de poli(acrilamida), almidón o agarosa o mediante la estimación del pI a partir de la secuencia de aminoácidos, por ejemplo mediante la utilización de la herramienta pI/MW en el servidor ExPASy ([http://expasy.org/tools/pi\\_tool.html](http://expasy.org/tools/pi_tool.html); Gasteiger et al., 2003).  
30

El extremo N-terminal del enzima CBHII/Cel6A recombinante purificado, así como los péptidos internos, pueden secuenciarse según la reacción de degradación de Edman (Edman y Begg, 1967) o mediante la predicción del sitio de corte de la secuencia de señal de secreción a partir de la secuencia de aminoácidos, por ejemplo mediante la utilización del programa SignalP V3.0 (Nielsen et al., 1997; Nielsen y Krogh, 1998; Bendtsen et al., 2004) tal como se indica en el Ejemplo 1 o en otros procedimientos descritos en la literatura.  
35

La expresión "de longitud completa" se refiere a la forma del enzima traducido a partir de la secuencia de ADN codificante que se inicia con el codón de inicio ATG, que codifica la primera metionina en la secuencia de aminoácidos y que finaliza en el codón de parada TGA, TAG o TAA.  
40

El término "maduro" se refiere a la forma del enzima que después de la eliminación de la secuencia de señal (péptido de señal de secreción o prepéptido) comprende los aminoácidos esenciales para la actividad enzimática o catalítica. En los hongos filamentosos es la forma secretada al medio de cultivo como resultado de, por ejemplo, el procesamiento N-terminal de la secuencia de señal y otro procesamiento N-terminal y la glucosilación post-traducciona. Además, la forma madura se refiere a un enzima que ha sido cortado de su proteína portadora en las construcciones de fusión.  
45

Muchas de las celobiohidrolasas bacterianas y fúngicas se producen como enzimas modulares (Srisodsuk et al., 1993; Suurnäkki et al., 2000). Además de un dominio catalítico o nuclear que expresa actividad celulolítica, estos enzimas pueden comprender uno o más dominios de unión a celulosa (DUC), también denominados dominios/módulos de unión a carbohidrato (DUC/MUC), los cuales pueden encontrarse situados en el extremo N-terminal o C-terminal del dominio catalítico. Los DUC presentan actividad de unión a carbohidrato y median en la unión del enzima celulasa a la celulosa cristalina pero presentan poco o ningún efecto sobre la actividad hidrolítica de la celulasa del enzima sobre los sustratos solubles. Estos dos dominios típicamente se conectan mediante un  
50

65

conector flexible y altamente glucosilado o región bisagra, tal como se pone de manifiesto a partir de las secuencias de aminoácidos de SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 y SEC ID nº 16.

El término "fragmento" tal como se da a conocer se refiere a un polipéptido que no presenta uno o más residuos aminoácidos desde el extremo N-terminal y/o C-terminal del polipéptido de longitud completa de SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 o SEC ID nº 16, tal como la forma madura del polipéptido u otra parte enzimáticamente activa del polipéptido. El fragmento todavía presenta la actividad catalítica esencial o actividad de celobiohidrolasa de la celobiohidrolasa CBHII/Cel6A de longitud completa y propiedades sustancialmente similares, tales como la dependencia y estabilidad frente al pH y la temperatura y la especificidad de sustrato. Los polipéptidos dados a conocer en SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 y SEC ID nº 16 contienen naturalmente un DUC N-terminal y un conector. Estas regiones conectoras o DUC nativas pueden sustituirse por, por ejemplo, un DUC de una especie de *Trichoderma* o de *Chaetomium*. Los enzimas CBHII/Cel6A de la invención pueden utilizarse en las aplicaciones también sin una secuencia de señal y/o DUC o la secuencia de señal y/o el DUC puede obtenerse a partir de diferentes enzimas de los microorganismos anteriormente indicados o de microorganismos diferentes o puede incorporarse sintética o recombinantemente al dominio catalítico de los enzimas anteriormente indicados.

El término "identidad" tal como se utiliza en la presente memoria se refiere a la identidad entre dos secuencias de aminoácidos comparados entre sí dentro de la región de secuencia correspondiente que presenta aproximadamente la misma cantidad de aminoácidos. Por ejemplo, puede compararse la identidad de una secuencia de longitud completa o de una secuencia madura de las dos secuencias de aminoácidos. Además, puede compararse la identidad de una secuencia de longitud completa o de una secuencia madura que no presenta el DUC N-terminal de las dos secuencias de aminoácidos. De esta manera, la comparación de, por ejemplo, las secuencias de CBHII/Cel6A, incluyendo el DUC y/o las secuencias de señal con secuencias que no presentan dichos elementos, no se encuentra dentro del contexto de la invención. La identidad de las secuencias de longitud completa puede medirse mediante la utilización de, por ejemplo, la alineación de ClustalW (por ejemplo en [www.ebi.ac.uk/Tools/Clustalw](http://www.ebi.ac.uk/Tools/Clustalw)) con una matriz tal como la siguiente: BLOSUM, apertura de hueco: 10, extensión de hueco: 0,5, o utilizando un programa de Clone Manager (versión 9) (Scientific and Educational Software, Care, EEUU), incluyendo las funciones "Comparar dos secuencias/Global/Comparar secuencias" como aminoácidos/matriz de puntuación BLOSUM62, tal como se indica en el Ejemplo 1.

Las secuencias de aminoácidos de las dos moléculas que deben compararse pueden diferir en una o más posiciones, que, sin embargo, no alteran la función biológica o estructura de las moléculas. Dichas "variantes" pueden producirse naturalmente debido a diferentes organismos hospedadores o mutaciones en la secuencia de aminoácidos, por ejemplo como variante alélica dentro de la misma cepa, especie o género, o pueden conseguirse mediante mutagénesis específica. Según una forma de realización, pueden comprender sustituciones, deleciones, combinaciones o inserciones de aminoácidos de una o más posiciones en la secuencia de aminoácidos, aunque todavía funcionen de manera sustancialmente similar a los enzimas definidos en SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 o SEC ID nº 16, es decir, comprenden una variante que presenta actividad celulolítica.

Según una forma de realización, en un procedimiento de tratamiento de material celulósico, que incluye también material lignocelulósico con un polipéptido CBHII/Cel6A o una composición enzimática que comprende dicho polipéptido o tal como en un bioprocedimiento consolidado, el material celulósico puede tratarse con un microorganismo fermentador que produce dicho polipéptido, en el que el polipéptido CBHII/Cel6A muestra actividad de celobiohidrolasa y comprende una secuencia de aminoácidos que presenta una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 12, una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto al polipéptido de longitud completa de SEC ID nº 15 o una identidad de por lo menos 91% respecto al polipéptido de longitud completa de SEC ID nº 16, o un fragmento o variante de los mismos con propiedades similares. Dicho enzima es capaz de hidrolizar material celulósico, incluyendo lignocelulosa a temperaturas moderadas a elevadas. Dicho procedimiento comprende las etapas siguientes: i) producción de un polipéptido CBHII/Cel6A de la invención o una composición enzimática que comprende dicho polipéptido o un microorganismo fermentador que produce dicho polipéptido, ii) reaccionar del material celulósico con el polipéptido CBHII/Cel6A de la invención o la composición enzimática que comprende dicho polipéptido o el microorganismo fermentador productor de dicho polipéptido, y iii) obtención de material celulósico parcial o totalmente hidrolizado, incluyendo material lignocelulósico hidrolizado.

Los enzimas celobiohidrolasa CBHII/Cel6A útiles para tratar o hidrolizar material celulósico son "obtenibles de" cualquier organismo, incluyendo plantas. Preferentemente, los enzimas CBHII/Cel6A se originan a partir de microorganismos, por ejemplo bacterias u hongos. Según una forma de realización, las bacterias pueden ser, por ejemplo, de un género seleccionado de entre *Bacillus*, *Azospirillum*, *Streptomyces* y *Pseudomonas*. Según una forma de realización, el enzima se origina a partir de hongos (incluyendo hongos filamentosos y levaduras), por ejemplo de un género seleccionado de entre el grupo que consiste en las levaduras *Schizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Candida* y *Yarrowia*, o de los hongos filamentosos *Achaetomium*, *Acremonium*, *Aspergillus*, *Aureobasidium*, *Botrytis*, *Chaetomium*, *Chrysosporium*, *Cryptococcus*, *Collybia*, *Fomes*, *Fusarium*, *Humicola*, *Hypocrea*, *Lentinus*, *Magnaporthea*, *Melanocarpus*, *Mucor*, *Myceliophthora*, *Myriococcum*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Pleurotus*, *Podospora*, *Polyporus*, *Pycnoporus*, *Rhizoctonia*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Trametes* y

*Trichoderma*. Preferentemente, los enzimas CBHII/Cel6A se obtienen de *Acremonium thermophilum*, *Melanocarpus albomyces*, *Chaetomium thermophilum* o *Talaromyces emersonii*.

5 Se dan a conocer además los enzimas obtenibles de la cepa de hongo filamentoso ALKO4245 depositada como CBS 116240 y que se clasifica actualmente como *Acremonium thermophilum*, *Melanocarpus albomyces* cepa ALKO4237, depositada como CBS 685.95, *Talaromyces emersonii* cepa DSM 2432 (en la colección de cultivos del solicitante bajo el número RF8069) o *Chaetomium thermophilum* cepa ALKO4265, depositada como CBS 730.95.

10 Las celobiohidrolasas CBHII/Cel6A dadas a conocer se comercializan como *At*\_ALKO4245\_Cel6A, *Ma*\_ALKO4237\_Cel6A, *Ct*\_ALKO4265\_Cel6A y *Te*\_RF8069\_Cel6A, siendo las celobiohidrolasas CBHII/Cel6A originadas de las cepas *Acremonium thermophilum* CBS 116240, *Melanocarpus albomyces* CBS 685.95, *Chaetomium thermophilum* CBS 730.95 o *Talaromyces emersonii* DSM 2432 y elementos de la familia 6 de glucósido hidrolasas.

15 Tabla 1: celobiohidrolasas CBHII/Cel6A de la invención

Celobiohidrolasa II	Gen	Obtenible de	Ácidos nucleicos SEC ID nº	Aminoácidos SEC ID nº
<i>At</i> _ALKO4245_Cel6A	<i>At</i> _ALKO4245_cel6A	<i>Acremonium thermophilum</i>	11	12
<i>Ma</i> _ALKO4237_Cel6A	<i>Ma</i> _ALKO4237_cel6A	<i>Melanocarpus albomyces</i>	13	14
<i>Ct</i> _ALKO4265_Cel6A	<i>Ct</i> _ALKO4265_cel6A	<i>Chaetomium thermophilum</i>	9	15
<i>Te</i> _RF8069_Cel6A	<i>Te</i> _RF8069_cel6A	<i>Talaromyces emersonii</i>	10	16

20 Según una forma de realización, el enzima celobiohidrolasa CBHII/Cel6A fúngico útil en el procedimiento es un polipéptido que presenta actividad de celobiohidrolasa y que comprende el enzima *At*\_ALKO4245\_Cel6A que presenta la secuencia de aminoácidos de longitud completa de SEC ID nº 12 o una secuencia de aminoácidos que presenta una identidad de por lo menos 76% respecto a la secuencia de aminoácidos SEC ID nº 12. Los enzimas preferentes muestran una identidad de por lo menos 78%, 80% o 82%, preferentemente de por lo menos 84%, 86% o 88%, más preferentemente de por lo menos 90%, todavía más preferentemente de por lo menos 92%. Todavía más preferentemente, las secuencias de aminoácidos muestran una identidad de por lo menos 94% o de por lo menos 96% o 97%, más preferentemente de por lo menos 98%, todavía más preferentemente de 99% respecto a la secuencia de aminoácidos de SEC ID nº 12. Las identidades de los dos enzimas se comparan dentro de las regiones de secuencia correspondientes, es decir, dentro de la región completa de la celobiohidrolasa CBHII/Cel6A.

30 Se da a conocer además un enzima celobiohidrolasa CBHII/Cel6A fúngica útil en el procedimiento que presenta una actividad de celobiohidrolasa y que comprende el enzima *Ma*\_ALKO4237\_Cel6A que presenta la secuencia de aminoácidos de longitud completa de SEC ID nº 14 o una secuencia de aminoácidos que presenta una identidad de por lo menos 76% respecto a la secuencia de aminoácidos SEC ID nº 14. según una forma de realización, los enzimas muestran una identidad de por lo menos 78%, 80% o 82%, preferentemente de por lo menos 84%, 86% o 88%, más preferentemente de por lo menos 90%, todavía más preferentemente de por lo menos 92%. Según una forma de realización, las secuencias de aminoácidos muestran una identidad de por lo menos 94% o de por lo menos 96% o 97%, más preferentemente de por lo menos 98%, todavía más preferentemente de 99% respecto a la secuencia de aminoácidos de SEC ID nº 14. Se comparan las identidades de los dos enzimas dentro de las regiones de secuencia correspondientes, es decir, dentro de la región de longitud completa de la celobiohidrolasa CBHII/Cel6A.

40 Según una forma de realización es un enzima celobiohidrolasa CBHII/Cel6A fúngico útil en el procedimiento que presenta actividad de celobiohidrolasa y que comprende el enzima *Ct*\_ALKO4265\_Cel6A que presenta la secuencia de aminoácidos de longitud completa de SEC ID nº 15 o una secuencia de aminoácidos que presenta una identidad de por lo menos 76% respecto a la secuencia de aminoácidos SEC ID nº 15. Los enzimas preferentes muestran una identidad de por lo menos 76%, preferentemente de por lo menos 97%, más preferentemente de por lo menos 98%, todavía más preferentemente de por lo menos 99% respecto a la secuencia de aminoácidos de SEC ID nº 15. Las identidades de los dos enzimas se comparan dentro de las regiones de secuencia correspondientes, es decir, dentro de la región de longitud completa de la celobiohidrolasa CEBHII/Cel6A.

50 Se da a conocer además un enzima celobiohidrolasa CBHII/Cel6A fúngico útil en el procedimiento que presenta actividad de celobiohidrolasa y que comprende el enzima *Te*\_RF8069\_Cel6A que presenta la secuencia de aminoácidos de longitud completa de SEC ID nº 16 o una secuencia de aminoácidos que presenta una identidad de por lo menos 91% respecto a la secuencia de aminoácidos SEC ID nº 16. Según una forma de realización, los enzimas muestran una identidad de por lo menos 92%, más preferentemente de por lo menos 93%, todavía más

preferentemente de por lo menos 94%. Todavía más preferentemente, las secuencias de aminoácidos muestran una identidad de por lo menos 95% o de por lo menos 96% o 97%, más preferentemente de por lo menos 98%, todavía más preferentemente de 99% respecto a la secuencia de aminoácidos de SEC ID nº 16. Se comparan las identidades de los dos enzimas dentro de las regiones de secuencia correspondientes, es decir, dentro de la región de longitud completa de la celobiohidrolasa CBHII/Cel6A.

Las celobiohidrolasas CBHII/Cel6A fúngicas tal como se dan a conocer son activas o estables en un amplio intervalo de pH, de por lo menos pH 3 a pH 10, y más preferentemente en un intervalo de pH de por lo menos pH 3 a pH 7 al someterlas a ensayo a 50°C durante 21 horas utilizando celulosa Avicel como sustrato, tal como se indica en el Ejemplo 3.

Según una forma de realización, *At\_ALKO4245\_Cel6A* se encuentra activa entre pH 3 y pH 7, preferentemente entre pH 4 y pH 6, y más preferentemente entre pH 4 y pH 5. La actividad máxima de *At\_ALKO4245\_Cel6A* se mide a pH 5, al someterla a ensayo a 50°C durante 21 horas utilizando celulosa Avicel como sustrato.

Según una forma de realización, *Ma\_ALKO4237\_Cel6A* se encuentra activo en un intervalo de pH de entre pH 3 y pH 10, preferentemente de entre pH 4 y pH 8, más preferentemente de entre pH 4 y pH 7, todavía más preferentemente de entre pH 4 y pH 6, y todavía más preferentemente de entre pH 4 y pH 5. La actividad máxima de *Ma\_ALKO4237\_Cel6A* se mide a pH 4 al someterla a ensayo a 50°C durante 21 horas utilizando celulosa Avicel como sustrato.

Según una forma de realización, *Ct\_ALKO4265\_Cel6A* se encuentra activo en un intervalo de pH de entre pH 3 y pH 10, preferentemente de entre pH 3 y pH 8, más preferentemente de entre pH 3 y pH 7, todavía más preferentemente de entre pH 4 y pH 7, y todavía más preferentemente de entre pH 4 y pH 6. La actividad máxima de *Ct\_ALKO4265\_Cel6A* se mide a pH 5 al someterla a ensayo a 50°C durante 21 horas utilizando celulosa Avicel como sustrato.

Según una forma de realización, *Te\_RF8089\_Cel6A* es activo en un intervalo de pH de entre pH 3 y pH 7, preferentemente de entre pH 3 y pH 6, y más preferentemente de entre pH 4 y pH 6. La actividad máxima de *Te\_RF8089\_Cel6A* se mide a pH 4 al someterla a ensayo a 50°C durante 21 horas utilizando celulosa Avicel como sustrato.

Los enzimas de la invención resultan eficaces en la degradación de material celulósico y lignocelulósico en un amplio intervalo de temperaturas. Las celobiohidrolasas CBHII/Cel6A de la invención son activas o estables durante un máximo de 21 horas en un intervalo de temperaturas de entre 40°C y 70°C al someterla a ensayo al pH óptimo de los enzimas utilizando celulosa Avicel como sustrato, tal como se indica en el Ejemplo 3.

La celobiohidrolasa *At\_ALKO4245\_Cel6A* se encuentra activa entre 40°C y 70°C, preferentemente entre 40°C y 60°C, más preferentemente entre 50°C y 60°C. Al pH óptimo la actividad máxima de *At\_ALKO4245\_Cel6A* se mide a 60°C al utilizar un tiempo de incubación de 21 horas y celulosa Avicel como sustrato.

Según una forma de realización, la celobiohidrolasa *Ma\_ALKO4237\_Cel6A* se encuentra activa en un intervalo de temperaturas de entre 40°C y 60°C. El enzima muestra una actividad máxima a 50°C al incubarla durante 21 horas al pH óptimo utilizando celulosa Avicel como sustrato. La celobiohidrolasa *Ct\_ALKO4265\_Cel6A* se encuentra activa en un intervalo de temperaturas de entre 40°C y 70°C, preferentemente en el intervalo de entre 50°C y 60°C. Al pH óptimo, la actividad máxima se mide a 60°C al utilizar un tiempo de incubación de 21 horas y celulosa Avicel como sustrato. La celobiohidrolasa *Te\_RF8089\_Cel6A* se encuentra activa entre 40°C y 70°C, preferentemente entre 40°C y 60°C, más preferentemente entre 50°C y 60°C. La actividad máxima se mide a 60°C al incubar durante 21 horas al pH óptimo utilizando celulosa Avicel como sustrato.

Los enzimas celulolíticos actualmente utilizados en la hidrólisis de material celulósico y la producción de azúcares fermentables para, por ejemplo, aplicaciones de bioetanol se obtienen principalmente de microorganismos bien estudiados, tales como el hongo filamentoso *Trichoderma reesei* (por ejemplo la línea de productos Accellerase®, Genencor Int., Inc., EEUU). Los enzimas celulolíticos se utilizan convencionalmente a temperaturas comprendidas en el intervalo de entre 30°C y 45°C. Las celobiohidrolasas CBHII/Cel6A de la presente invención son eficientes también a estas temperaturas, aunque además funcionan extremadamente bien incluso a temperaturas de hasta 70°C, tal como de entre 40°C y 70°C, por ejemplo de entre 40°C y 70°C o de entre 40°C y 60°C o de entre 50°C y 60°C, tal como se muestra en el Ejemplo 3. En los tiempos de hidrólisis cortos las composiciones enzimáticas pueden resultar funcionales hasta 80°C. En tiempos de hidrólisis cortos se requieren tiempos de incubación más prolongados y, por lo tanto, se utilizan normalmente temperaturas más bajas. Lo anterior provoca que las celobiohidrolasas CBHII/Cel6A de la invención resulten extremadamente adecuadas a la modificación de los procedimientos de hidrólisis de sustratos celulósicos llevados a cabo a temperaturas convencionales o moderadas y a temperaturas elevadas.

En los Ejemplos 4 y 5, se describen experimentos realizados con diversos materiales celulósicos, tales como madera dura y mazorcas de maíz. A partir de la figura 4 resulta evidente que el rendimiento de las mezclas

enzimáticas complementadas con las celobiohidrolasas CBHII/Cel6A fúngicas *At\_ALKO4245\_Cel6A*, *Ma\_ALKO42237\_Cel6A* o *Ct\_ALKO4265\_Cel6A* en la hidrólisis de madera dura sometida a explosión de vapor es mucho mejor que el rendimiento de las mezclas enzimáticas sin dicha complementación. En los experimentos realizados a 55°C (figura 4A), se encontró que la cantidad de azúcares liberada del sustrato de madera dura se incrementaba en 12%, 14% y 26% mediante la complementación con *Ma\_ALKO4237\_Cel6A*, *Ct\_ALKO4265\_Cel6A* o *At\_ALKO4245\_Cel6A* en la MEZCLA 2, respectivamente. Se encontró que el enzima ALKO4245 de *Acremonium thermophilum* era el enzima CBHII/Cel6A de mejor rendimiento de entre los estudiados en la presente memoria. El enzima *At\_ALKO4245\_Cel6A* mostró una hidrólisis incrementada también a 37°C, añadido en la mezcla de *Trichoderma* del estado de la técnica (MEZCLA ENZIMAS T.REESEI) (figura 4B9 o en el producto comercial (MEZCLA ACC) (figura 4C).

También se obtuvieron resultados similares en las mazorcas de maíz sometidas a explosión con vapor (figura 5). La MEZCLA 2 de enzima termofílico se complementó con el enzima *At\_ALKO4245\_Cel6A* de la invención. La cantidad agrupada de glucosa y xilosa tras 72 horas de hidrólisis era notablemente más elevada que sin dicha complementación.

Según una forma de realización, el procedimiento de tratamiento del material celulósico y lignocelulósico, tal como material lignocelulósico que incluía hemicelulosa, pectina y lignina, implica la utilización de una o más de las celobiohidrolasas CBHII/Cel6A de la invención como "composición enzimática" o "preparación enzimática" que comprende por lo menos un enzima adicional capaz de hidrolizar dicho material. Los enzimas adicionales pueden seleccionarse de entre el grupo de celobiohidrolasa, endoglucanasa, beta-glucosidasa, beta-glucanasa, xiloglucanasa, xilanasa, beta-xilosidasa, mananasa, beta-manosidasa,  $\alpha$ -glucuronidasa, acetil-xilán-esterasa,  $\alpha$ -arabinofuranosidasa,  $\alpha$ -galactosidasa, pectinasa, que implica endo- y exo- $\alpha$ -L-arabinasas,  $\alpha$ -galactosidasa, endo- y exo-galactoronasa, endopectiniliasa, pectinesterasa, pectato liasa, fenol-esterasa, ligninasa que implica lignina peroxidasa, peroxidasa dependiente de manganeso, enzima generador de H<sub>2</sub>O<sub>2</sub> y lacasa con o sin un mediador.

Según una forma de realización, la preparación o composición enzimática comprende por lo menos uno de los enzimas indicados anteriormente. Puede contener enzimas en forma por lo menos parcialmente purificada y aislada. Incluso puede consistir esencialmente del enzima o enzimas deseados. Alternativamente, la preparación puede ser un medio de cultivo o filtrado agotado que contiene uno o más de los enzimas deseados. Preferentemente, la preparación enzimática es medio de cultivo agotado. La expresión "medio de cultivo agotado" se refiere al medio de cultivo del hospedador que comprende los enzimas producidos. Preferentemente, las células hospedadoras se separan de dicho medio después de la producción. La preparación o composición enzimática también puede ser un "caldo de cultivo completo" obtenido opcionalmente después de eliminar el/los hospedador/hospedadores o microorganismo/microorganismos de producción sin ningún procesamiento o producción posterior del enzima o enzimas celulolíticos deseados. En el "bioprocedimiento consolidado", la composición enzimática o por lo menos algunos de los enzimas de la composición enzimática puede ser producido por el microorganismo fermentador.

La expresión "polipéptido aislado" en el presente contexto puede referirse simplemente a que las células y residuos celulares han sido eliminados del medio de cultivo que contiene el polipéptido. Convenientemente, los polipéptidos son aislados, por ejemplo mediante la adición de polímeros aniónicos y/o catiónicos, al medio de cultivo agotado con el fin de potenciar la precipitación de células, residuos celulares y algunos enzimas que presentan actividades secundarios no deseadas. A continuación, el medio se filtra utilizando un agente de filtración inorgánico y un filtro para eliminar los sedimentos formados. Después, el filtrado se procesa adicionalmente utilizando una membrana semipermeable para eliminar el exceso de sales, azúcares y productos metabólicos.

Además de la actividad enzimática, la preparación puede contener aditivos, tales como mediadores, estabilizadores, tampones, conservantes, surfactantes y/o componentes del medio de cultivo. Los aditivos preferentes son los que se utilizan comúnmente en preparaciones enzimáticas destinadas a una aplicación particular. La preparación enzimática puede encontrarse en forma de líquido, polvos o granulado.

Según una forma de realización de la invención, la preparación enzimática comprende una mezcla de CBH, EG y BG, opcionalmente junto con enzimas degradadores de la hemicelulosa en combinación con las celobiohidrolasas CBHII/Cel6A de la invención. Pueden utilizarse diferentes mezclas y combinaciones enzimáticas para adecuarse a diferentes materiales de sustrato y condiciones del procedimiento. Por ejemplo, en el caso de que el procedimiento de degradación deba llevarse a cabo a una temperatura elevada, se seleccionan enzimas termoestables.

La "hemicelulosa" es un grupo heterogéneo de polímeros de carbohidrato que contiene principalmente diferentes glucanos, xilanos y mananos. La hemicelulosa consiste en un esqueleto lineal con residuos unidos mediante enlace  $\beta$ -1,4, sustituidos con cadenas laterales cortas que habitualmente contienen grupos ácido acetilo, ácido 4-O-glucurónico, L-arabinosa y galactosilo. La hemicelulosa puede entrecruzarse químicamente con lignina y celulosa. Entre los "enzimas degradadores de xilano", o "xilanasas", se incluyen enzimas tanto exohidrolíticos como endohidrolíticos, tales como endo-1,4-beta-D-xilanasa (EC 3.2.1.8) o exo-1,4-beta-D-xilosidasa (EC 3.2.1.37), que descomponen la hemicelulosa en xilosas. Los gluco- y galacto-mananos son hidrolizados por las endo-1,4-beta-mananasas (EC 3.2.1.78) y la beta-manosidasa (EC 3.2.1.25), rindiendo beta-D-manosa. Entre los enzimas capaces de eliminar los sustituyentes de cadena lateral se incluyen  $\alpha$ -glucuronidasas, acetil-xilán-esterasas,  $\alpha$ -

arabinofuranosidasas y  $\alpha$ -galactosidasas, que actúan cooperativamente con los enzimas degradadores del esqueleto (Biely et al., 1997; Sundberg y Poutanen, 1991; Stålbrand et al., 1995).

5 Los enzimas que participan en la degradación de la "pectina" implican endo- y exo- $\alpha$ -L-arabinasas y  $\alpha$ -galactosidasa, endo- y exo-galactoronasas, endopectinilinasas y pectinesterasas (Del Cañizo et al., 1994).

10 La "lignina" es un polímero entrecruzado complejo de unidades p-hidroxifenilpropano de diversa sustitución. Su hidrólisis implica peroxidasas de lignina, fenol-esterasas, peroxidasas dependientes de manganeso, enzimas generadores de H<sub>2</sub>O<sub>2</sub> y lacasas (Cullen y Kersten, 2004).

15 Los enzimas de la composición enzimática pueden añadirse al material lignocelulósico de manera simultánea o secuencial.

20 Según una forma de realización, el procedimiento es aplicable a diversos materiales que contienen celulosa, lignocelulosa y beta-glucano, tales como materiales vegetales, por ejemplo madera, incluyendo madera dura y madera blanda, virutas de madera dura y de madera blanda, pulpa de madera, serrín y residuos industriales silvícolas y de madera, cultivos herbáceos, biomasa agrícola, tal como paja de cereales, pulpa de remolacha azucarera, fibra de maíz, rastrojo y mazorcas de maíz, beta-glucanos de cereales, bagazo de caña azucarera, tallos, hojas, vainas y cáscaras, y similares; productos de desecho, tales como residuos sólidos urbanos, papel de periódico y residuos de papel de oficina, fibra de desecho, lodos de papel, residuos triturados de, por ejemplo, cereales; cultivos energéticos específicos (por ejemplo sauce, chopo, pastos altos ('switchgrass') o alpiste, y similares).

25 Según una forma de realización de la invención, el procedimiento resulta aplicable a la producción de biocombustibles, tales como etanol, propanol y butanol y similares a partir de material celulósico.

30 Según una forma de realización se encuentra un procedimiento en el que el polipéptido CBHII/Cel6A utilizado en la hidrólisis del material celulósico se obtiene de la cepa CBS 116240 de *Acremonium thermophilum* y presenta la secuencia de aminoácidos de SEC ID nº 12 o una identidad de por lo menos 78%, 80%, 82%, 84%, 86%, 88%, 90%, 92%, 94%, 96%, 97%, 98% o 99% respecto a la secuencia de aminoácidos de SEC ID nº 12, o un fragmento o variante de la misma que presenta actividad de celobiohidrolasa.

35 También se da a conocer un polipéptido CBHII/Cel6A fúngico que presenta actividad de celobiohidrolasa y que comprende una secuencia de aminoácidos que presenta una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 12, una identidad de por lo menos 76% respecto al polipéptido de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto al polipéptido de longitud completa de SEC ID nº 15, o una identidad de por lo menos 91% respecto al polipéptido de longitud completa de SEC ID nº 16, o un fragmento o variante de los mismos que presentan propiedades similares.

40 El polipéptido CBHII/Cel6A puede obtenerse a partir de los géneros de hongo filamentoso *Acremonium*, *Melanocarpus*, *Chaetomium* o *Talaromyces*. Entre las especies dadas a conocer se incluyen *Acremonium thermophilum*, *Melanocarpus albomyces*, *Chaetomium thermophilum* o *Talaromyces emersonii*.

45 Se dan a conocer además enzimas obtenibles a partir de una cepa de hongo filamentoso ALKO4245 depositada como CBS 116240 actualmente clasificada como *Acremonium thermophilum*, *Melanocarpus albomyces* cepa ALKO4237, depositada como CBS 685.95, *Chaetomium thermophilum* cepa ALKO4265, depositada como CBS 730.95 o *Talaromyces emersonii* cepa DSM 2432 (en la colección de cultivos del solicitante bajo el número RF8069).

50 Según una forma de realización preferente de la invención, el enzima celobiohidrolasa CBHII/Cel6A fúngico es un polipéptido que presenta actividad de celobiohidrolasa y que comprende el enzima At\_ALKO4245\_Cel6A que presenta la secuencia de aminoácidos de longitud completa de SEC ID nº 12.

55 Se da a conocer además un enzima celobiohidrolasa CBHII/Cel6A fúngico que presenta actividad de celobiohidrolasa y que comprende el enzima Ma\_ALKO4237\_Cel6A con la secuencia de aminoácidos de longitud completa de SEC ID nº 14.

60 Se da a conocer además un enzima celobiohidrolasa CBHII/Cel6A fúngico que presenta actividad de celobiohidrolasa y que comprende el enzima Ct\_ALKO4265\_Cel6A con la secuencia de aminoácidos de longitud completa de SEC ID nº 15.

65 Se da a conocer además un enzima celobiohidrolasa CBHII/Cel6A fúngico que presenta actividad de celobiohidrolasa y que comprende el enzima Te\_RF8069\_Cel6A con la secuencia de aminoácidos de longitud completa de SEC ID nº 16.

Según una forma de realización preferida de la invención, el enzima celobiohidrolasa CBHII/Cel6A es capaz de hidrolizar material celulósico a temperaturas moderadas a elevadas. Las celobiohidrolasas CBHII/Cel6A de la



invención son activas o estables incluso a temperaturas de hasta 70°C, tal como de entre 40°C y 70°C, por ejemplo de entre 40°C y 60°C o de entre 50°C y 60°C, al someterlas a ensayo al pH óptimo de los enzimas utilizando celulosa Avicel Ph101 como sustrato, tal como se indica en el Ejemplo 3. Para tiempos de hidrólisis cortos, las composiciones enzimáticas pueden ser funcionales hasta una temperatura de 80°C. Para la hidrólisis total, resultan necesarios tiempos de incubación más prolongados y, por lo tanto, normalmente se utilizan temperaturas más bajas. Lo anterior convierte a las celobiohidrolasas CBHII/Cel6A de la invención en extremadamente adecuadas para modificar los procesos de hidrólisis del sustrato celulósico realizados a temperaturas tanto convencionales como moderadas y a temperaturas elevadas que requieren la termoestabilidad de los enzimas.

Según una forma de realización, el enzima CBHII/Cel6A fúngico se encuentra codificado por una molécula aislada de ácidos nucleicos que comprende una secuencia polinucleotídica, que codifica un polipéptido que comprende la secuencia de aminoácidos caracterizada en SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 o SEC ID nº 16. De esta manera, se da a conocer el enzima celobiohidrolasa CBHII/Cel6A o el polipéptido que comprende la secuencia de aminoácidos de la forma de longitud completa del enzima *At\_ALKO4245\_Cel6A* caracterizado en SEC ID nº 12, el enzima *Ma\_ALKO4237\_Cel6A* caracterizado en SEC ID nº 14, *Ct\_ALKO4265* caracterizado en SEC ID nº 15 o *Te\_RF8069\_Cel6A* caracterizado en SEC ID nº 16.

Además, se dan a conocer los polipéptidos codificados por las moléculas de ácidos nucleicos codificantes de un polipéptido CBHII/Cel6A que presentan una actividad de celobiohidrolasa y una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 12, una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 15 o de por lo menos 91% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 16. Las identidades de los dos enzimas se comparan dentro de las regiones de secuencia correspondientes, es decir, dentro de la región de longitud completa del polipéptido CBHII/Cel6A.

Se da a conocer además una secuencia polipeptídica que se encuentra codificada por una molécula de ácidos nucleicos codificante de un fragmento del polipéptido, en el que el fragmento de polipéptido no presenta uno o más residuos aminoácidos del extremo N-terminal y/o C-terminal del polipéptido de longitud completa de SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 o SEC ID nº 16. El fragmento todavía presenta la actividad catalítica esencial o la actividad de celobiohidrolasa de la celobiohidrolasa CBHII/Cel6A de longitud completa y propiedades sustancialmente similares, tales como dependencia del pH y de la temperatura y especificidad de sustrato. El fragmento puede ser, por ejemplo, un enzima que no presenta la secuencia de señal de secreción o el dominio de unión a carbohidrato o DUC.

Se dan a conocer además variantes naturales o sintéticas de las celobiohidrolasas CBHII/Cel6A que presentan propiedades similares a los polipéptidos definidos en SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 o SEC ID nº 16. La variación puede resultar de la delección, sustitución, inserción, adición o combinación de una o más posiciones en la secuencia de aminoácidos.

Una forma de realización es una celobiohidrolasa CBHII/Cel6A que se encuentra codificada por una molécula aislada de ácidos nucleicos que comprende una secuencia polinucleotídica incluida en SEC ID nº 11, SEC ID nº 13, SEC ID nº 9 o SEC ID nº 10 o una subsecuencia de las mismas.

Se da a conocer además una celobiohidrolasa CBHII/Cel6A que se encuentra codificada por una molécula de ácidos nucleicos o secuencia polinucleotídica que se hibrida bajo condiciones restrictivas con una secuencia polinucleotídica o una subsecuencia de la misma incluida en SEC ID nº 11, SEC ID nº 13, SEC ID nº 9 o SEC ID nº 10. Además, una forma de realización preferente es un polipéptido codificado por una molécula de ácidos nucleicos o secuencia polinucleotídica que se hibrida con una sonda preparada utilizando la PCR, tal como el fragmento de PCR incluido en SEC ID nº 7 o SEC ID nº 8.

Pueden utilizarse procedimientos estándares de biología molecular en el aislamiento del ADNc o un ADN genómico del organismo hospedador, por ejemplo, los procedimientos descritos en los manuales de biología molecular, tal como Sambrook y Russell, 2001. El ADNc o un gen genómico codificante de la celobiohidrolasa CBHII/Cel6A de la invención puede aislarse utilizando sondas de ADN, que han sido preparadas basándose en la secuencia de aminoácidos de péptidos N-terminales o trípticos del enzima CBHII/Cel6A purificado. Alternativamente, la sonda puede diseñarse basándose en las secuencias de nucleótidos o de aminoácidos conocidas de las celobiohidrolasas homólogas. Los clones de CBHII/Cel6a también pueden cribarse basándose en la actividad sobre placas que contienen un sustrato específico del enzima o mediante la utilización de anticuerpos específicos para una celobiohidrolasa CBHII/Cel6A.

La hibridación con una sonda de ADN, tal como, por ejemplo, SEC ID nº 7, que consiste en más de 100 a 200 nucleótidos, habitualmente se lleva a cabo bajo condiciones "altamente restrictivas", es decir, la hibridación a una temperatura que es 20°C a 25°C inferior a la temperatura de fusión ( $T_f$ ) calculada de un híbrido perfecto, la  $T_f$  calculada según Bolton y McCarthy (1962) y lavados post-hibridación a baja concentración salina. Habitualmente la prehibridación y la hibridación se llevan a cabo por lo menos a 65°C en 6xSSC (o 6xSSPE), reactivo de Denhardt 5x,

SDS al 0,5% (p/v), 100 µg/ml de ADN de esperma de salmón desnaturalizado y fragmentado. La adición de formamida al 50% reduce las temperaturas de prehibridación e hibridación a 42°C. Los lavados de alta astringencia se llevaron a cabo a baja concentración salina, por ejemplo en 2xSSC-SDS al 0,1% (p/v) a TA, y finalmente en 0,1xSSC-SDS al 0,1% (p/v) por lo menos a 65°C, por ejemplo a 68°C.

Según una forma de realización, los genes *At\_ALKO4245\_Cel6A*, *Ma\_ALKO4237\_Cel6A*, *Ct\_ALKO4265\_Cel6A* y *Te\_RF8069\_Cel6A* se aislaron con una sonda preparada mediante PCR utilizando condiciones de hibridación restrictiva tal como se indica en el Ejemplo 1. Los genes *cbh2/cel6A* genómicos se aislaron mediante la utilización de cebadores oligonucleótidos diseñados basándose en las secuencias de nucleótidos publicadas o utilizando cebadores oligonucleótidos degenerados planificados basándose en la alineación de las secuencias de aminoácidos previamente conocidas de las proteínas CBHII/Cel6A.

Según una forma de realización preferida de la invención, el enzima celobiohidrolasa CBHII/Cel6A se encuentra codificado por una molécula aislada de ácidos nucleicos que comprende la secuencia de nucleótidos de SEC ID nº 11 codificante de la forma de longitud completa del enzima *At\_ALKO4245\_Cel6A* de SEC ID nº 12. Se da a conocer además una celobiohidrolasa CBHII/Cel6A fúngica codificada por una molécula aislada de ácidos nucleicos que comprende la secuencia de nucleótidos de SEC ID nº 13, que codifica la forma de longitud completa del enzima *Ma\_ALKO4237\_Cel6A* que presenta la secuencia de aminoácidos de SEC ID nº 14. Se da a conocer además una celobiohidrolasa CBHII/Cel6A fúngica codificada por una molécula aislada de ácidos nucleicos que comprende la secuencia de nucleótidos de SEC ID nº 9 que codifica la forma de longitud completa del enzima *Ct\_ALKO4265\_Cel6A* que presenta la secuencia de aminoácidos de SEC ID nº 15. Se da a conocer además una celobiohidrolasa CBHII/Cel6A fúngica codificada por una molécula aislada de ácidos nucleicos que comprende la secuencia de nucleótidos de SEC ID nº 10, que codifica la forma de longitud completa del enzima *Te\_RF8069\_Cel6A* que presenta la secuencia de aminoácidos de SEC ID nº 16.

Según una forma de realización, la celobiohidrolasa CBHII/Cel6A fúngica se encuentra codificada por la secuencia polinucleotídica incluida en el plásmido pALK2582 depositada en *Escherichia coli* RF8175 bajo el número de acceso DSM 22945, el plásmido pALK2581, depositado en *Escherichia coli* RF8174 bajo el número de acceso DSM 22945, el plásmido pALK2904, depositado en *Escherichia coli* RF8214 bajo el número de acceso DSM 22947, o el plásmido pALK3006, depositado en *Escherichia coli* RF8333 bajo el número de acceso DSM 23185.

Una forma de realización de la invención es la celobiohidrolasa CBHII/Cel6A producida a partir de un vector de expresión recombinante que comprende la molécula de ácidos nucleicos, que codifica la celobiohidrolasa CBHII/Cel6A fúngica tal como se ha caracterizado anteriormente, operablemente ligada a secuencias reguladoras capaces de dirigir la expresión de un gen codificante de dicho enzima celobiohidrolasa CBHII/Cel6A en un hospedador adecuado. La construcción de dicho vector de expresión recombinante y la utilización de dicho vector se describen en mayor detalle en el Ejemplo 2.

Los hospedadores adecuados para la producción de la celobiohidrolasa CBHII/Cel6A fúngica son hospedadores homólogos o heterólogos, tal como los hospedadores microbianos, incluyendo bacterias, levaduras y hongos. Son hospedadores de producción preferentes los hongos filamentosos, tales como *Trichoderma*, *Aspergillus*, *Fusarium*, *Hemicola*, *Chrysosporium*, *Neurospora*, *Rhizopus*, *Penicillium* y *Mortierella*, debido a la facilidad de procesamiento y recuperación posteriores del producto enzimático. Entre los hospedadores adecuados se incluyen especies tales como las cepas de tipo *T. reesei*, *A. niger*, *A. oryzae*, *A. sojae*, *A. awamori* o *A. japonicus*, *F. venenatum* o *F. oxysporum*, *H. insolens* o *H. lanuginosa*, *N. crassa* y *C. lucknowense*, algunos de los cuales se listan como organismos hospedadores productores de enzimas en, por ejemplo, la lista AMFEP 2007 de enzimas comerciales (<http://www.amfep.org/list.html>). Más preferentemente, el enzima es producido en un hospedador fúngico filamentoso del género *Trichoderma* o *Aspergillus*, tal como *T. reesei* o *A. niger*, *A. oryzae* o *A. awamori*. Según la forma de realización más preferente de la invención, el enzima celobiohidrolasa CBHII/Cel6A fúngico es producido en *T. reesei*.

Según la forma de realización preferida de la invención, el polipéptido CBHII/Cel6A es *At\_ALKO4245\_Cel6A* obtenido de *Acremonium thermophilus* CBS 116240 y que presenta la secuencia de aminoácidos de SEC ID nº 12.

Se da a conocer además una molécula aislada de ácidos nucleicos que comprende una secuencia polinucleotídica codificante de la celobiohidrolasa CBHII/Cel6A fúngica seleccionada de entre el grupo que consiste en:

- (a) una molécula de ácidos nucleicos o secuencia polinucleotídica codificante de un polipéptido que presenta actividad de celobiohidrolasa y que comprende la secuencia de aminoácidos de longitud completa tal como se ilustra en SEC ID nº 12, SEC ID nº 14, SEC ID nº 15 o SEC ID nº 16, o un fragmento o variante de los mismos que presenta propiedades similares,
- (b) una molécula de ácidos nucleicos o secuencia polinucleotídica codificante de un polipéptido que presenta actividad de celobiohidrolasa y una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 12, una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto a la

secuencia de aminoácidos de longitud completa de SEC ID nº 15, o una identidad de por lo menos 91% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 16, o un fragmento o variante de los mismos que presenta propiedades similares,

- 5 (c) una molécula de ácidos nucleicos que comprende la secuencia codificante de la secuencia de nucleótidos tal como se ilustra en SEC ID nº 11, SEC ID nº 13, SEC ID nº 9 o SEC ID nº 10,
- (d) una molécula de ácidos nucleicos que comprende la secuencia codificante de la secuencia polinucleotídica contenida en DSM 22946, DSM 22945, DSM 22947 o DSM 23185,
- 10 (e) una molécula de ácidos nucleicos la secuencia codificante de la cual difiere de la secuencia codificante de una molécula de ácidos nucleicos de cualquiera de entre (c) y (d) debido a la degeneración del código genético, y
- 15 (f) una molécula de ácidos nucleicos hibridante bajo condiciones restrictivas con una molécula de ácidos nucleicos contenida en DSM 22946, DSM 22945, DSM 22947 o DSM 23185 o una subsecuencia de los mismos, y codificante de un polipéptido que presenta actividad de celobiohidrolasa y una secuencia de aminoácidos que muestra una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa tal como se ilustra en SEC ID nº 12, una identidad de por lo menos 76% respecto a la
- 20 secuencia de aminoácidos de longitud completa tal como se ilustra en SEC ID nº 14, una identidad de por lo menos 95% respecto a la secuencia de aminoácidos de longitud completa tal como se ilustra en SEC ID nº 15 o una identidad de por lo menos 91% respecto a la secuencia de aminoácidos de longitud completa tal como se ilustra en SEC ID nº 16.

25 La molécula de ácidos nucleicos de la invención puede ser de ARN o de ADN, en la que el ADN puede estar constituido de ADN genómico o ADNc.

Pueden utilizarse procedimientos estándares de biología molecular aisladamente y tratamientos enzimáticos de la secuencia polinucleotídica codificante de la celobiohidrolasa CBHII/Cel6A fúngica de la invención, incluyendo el aislamiento de ADN genómico y plasmídico, la digestión de ADN para producir fragmentos de ADN, la secuenciación, las transformaciones de *E. coli*, etc. Los procedimientos básicos se describen en los manuales estándares de biología molecular, por ejemplo Sambrook y Russell, 2001.

30

El aislamiento de los genes *At\_ALKO4245\_cel6A*, *Ma\_ALKO4237\_cel6A*, *Ct\_ALKO4265\_cel6A* y *Te\_RF8069\_cel6A* codificantes de los polipéptidos *At\_ALKO4245\_Cel6A*, *Ma\_ALKO4237\_Cel6A*, *Ct\_ALKO4265\_Cel6A* y *Te\_RF8069\_Cel6A* se describe en el Ejemplo 1. Brevemente, se utilizó el fragmento de PCR de 1.032 pb (SEC ID nº 7) y el fragmento de PCR de 831 pb (SEC ID nº 8) obtenidos mediante la utilización de las secuencias de los cebadores oligonucleótidos degenerados (SEC ID nº 1 y SEC ID nº 2) para aislar el gen *At\_ALKO4245\_cel6A* de *Acremonium thermophilum* ALKO4245 y el gen *Ma\_ALKO4237\_cel6A* de *Melanocarpus albomyces* ALKO42237 en vector pCR@4Blunt-TOPO®. Se incluyó el gen *cel6A* de *A. thermophilum* de longitud completa en el plásmido pALK2582 depositado en *E. coli*, en la colección de cultivos del DSMZ bajo el número de acceso DSM 22946. Se incluyó el gen *cel6A* de *M. albomyces* en el plásmido pALK2581 depositado en *E. coli*, en la colección de cultivos del DSMZ bajo el número de acceso DSM 22945.

35

40

El gen *Ct\_ALKO4265\_cel6A* se aisló utilizando las parejas de cebadores de SEC ID nº 3 y SEC ID nº 4 y el gen *Te\_RF8069\_cel6A* se aisló utilizando las parejas de cebadores de SEC ID nº 5 y SEC ID nº 6 tal como se indica en el Ejemplo 1. Los fragmentos de PCR que contenían los genes *cel6A* de longitud completa de *Chaetomium thermophilum* ALKO4265 y de *Talaromyces emersonii* RF8069 se incluyeron en los plásmidos pALK2904 y pALK3006 depositados en *E. coli*, en la colección de cultivos del DSMZ bajo los números de acceso DSM 22947 y DSM 23185, respectivamente.

45

50

Las secuencias de aminoácidos deducidas de las celobiohidrolasas CBHII/Cel6A se analizaron a partir de las secuencias de ADN tal como se indica en el Ejemplo 1.

La longitud del gen *At\_ALKO4245\_cel6A* (SEC ID nº 11), codificante de la celobiohidrolasa CBHII/Cel6A de *Acremonium thermophilum* (SEC ID nº 12), es de 1.830 pb (incluyendo el codón de parada). Se encontraron cuatro intrones putativos que presentaban longitudes de 79, 72, 117 y 126 pb. De esta manera, la región codificante de *At\_ALKO4245\_cel6A* es de 1.434 pb (codón de parada no incluido) y la secuencia de proteína deducida consistía de 478 aminoácidos, incluyendo una secuencia de señal predicha de 18 aminoácidos (SignalP V3.0; Nielsen, 1997; Nielsen y Krogh, 1998 y Bendtsen et al., 2004). La secuencia de aminoácidos deducida presentaba homología respecto a las secuencias de CBHII/Cel6A publicadas al analizarlas utilizando el programa BLAST, versión 2.2.9 en el NCBI, National Center for Biotechnology Information, Altschul et al., 1990). La masa molecular predicha del enzima maduro, excluyendo la secuencia de señal, era de 48.918 Da y el pI predicho era de 4,82. Estas predicciones se realizaron utilizando la herramienta Compute pI/MW en el servidor ExpPASy (Gasteiger et al., 2003). Los valores de identidad de la secuencia de *At\_ALKO4245\_Cel6A* de longitud completa respecto a las regiones correspondientes de las secuencias homólogas se obtuvieron mediante la utilización de un programa Clone Manager

55

60

65

(versión 9) que incluía las funciones "Comparar dos secuencias/Global/comparar secuencias como aminoácidos/matriz de puntuaciones BLOSUM62". Los valores (%) que muestran identidad respecto a las otras celobiohidrolasas CBHII/Cel6A de la presente invención se muestran en la Tabla 6. La identidad respecto a las secuencias de aminoácidos de CBHII/Cel6A publicadas se presenta en las Tablas 7 y 8.

5 El *At\_ALKO4245\_Cel6A* de la invención mostraba la homología más alta respecto a los polipéptidos de longitud completa de *Thielavia terrestris* NRRL 8126 (SEC ID nº 2 en la patente US nº 7.220.565, Novozymes Inc., EEUU, y SEC ID nº 49 en el documento WO2009/085868, Novozymes A/S, Dinamarca), el producto proteína sin nombre de *Podospora anserina* DSM 980 (EMBL nº de acceso XP\_001903170) y el precursor de endoglucanasa 2 deducido de *Neurospora crassa* OR74A (XM\_955677). La identidad respecto a la proteína de *T. terrestris* era, dentro del polipéptido de longitud completa, de 75%. Las identidades respecto a los polipéptidos de *P. anserina* y de *N. crassa* eran de 69%.

15 La longitud del gen *Ma\_ALKO4237\_cel6A* (SEC ID nº 13), codificante de la celobiohidrolasa CBHII/Cel6A de *Melanocarpus albomyces* (SEC ID nº 14) era de 1.607 pb (incluyendo el codón de parada). Se encontraron dos intrones putativos que presentaban una longitud de 93 y 95 pb. De esta manera, la región codificante de *Ma\_ALKO4237\_cel6A* era de 1.416 pb (codón de parada no incluido) y la secuencia de proteína deducida consistía de 472 aminoácidos, incluyendo una secuencia de señal predicha de 17 aminoácidos (SignalP V3.0; Nielsen et al., 1997; Nielsen y Krogh, 1998, y Bendtsen et al., 2004). La masa molecular predicha del enzima maduro excluyendo la secuencia de señal era de 48.627 Da y el pl predicho era de 4,50. Estas predicciones se realizaron utilizando la herramienta Compute pI/MW en el servidor ExpASY (Gasteiger et al., 2003). El *Ma\_ALKO4237\_Cel6A* de la invención mostraba la homología más alta respecto al polipéptido de longitud completa de un organismo no caracterizado, dado a conocer como la secuencia de aminoácidos SEC ID nº 413 en el documento WO 2008/095033 (Syngenta Inc., EEUU) (72%) respecto al polipéptido CBHII de *Thielavia terrestris* NRRL 8126 (SEC ID nº 49 en el documento nº WO2009/085868, Novozymes A/S, Dinamarca) (71%) y la proteína hipotética CHGG\_10762 de *Chaetomium globosum* CBS 148.151 (EMBL nº de acceso XP001226029) (identidad de 75%).

30 La longitud del gen *Ct\_ALKO4265\_cel6A* (SEC ID nº 9), codificante de la celobiohidrolasa CBHII/Cel6A de *Chaetomium thermophilum* (SEC ID nº 15) era de 1.757 pb (incluyendo el codón de parada). Se encontraron tres intrones putativos que presentaban longitudes de 77, 196 y 56 pb. De esta manera, la región codificante de *Ct\_ALKO4265\_cel6A* era de 1.425 pb (codón de parada no incluido) y la secuencia de proteína deducida consiste en 475 aminoácidos, incluyendo una secuencia de señal predicha de 17 aminoácidos (SignalP V3.0; Nielsen et al., 1997; Nielsen y Krogh, 1998 y Bendtsen et al., 2004). La masa molecular predicha del enzima maduro, excluyendo la secuencia de señal, era de 49.408 Da y el pl predicho era de 5,31. El *Ct\_ALKO4265\_Cel6A* de la invención mostraba la homología más alta respecto a la celobiohidrolasa deducida de la familia 6 de *Chaetomium thermophilum* CT2 (EMB nº de acceso AY861348; CN 1757709, Shandong Agricultural University, CN), respecto al polipéptido de *C. thermophilum* CGMCC0859 que presenta actividad de celobiohidrolasa II (SEC ID nº 2 en la patente nº EP1578964B1, Novozymes A/S, Dinamarca) y respecto a CBHII de *Chaetomium thermophilum* de SEC ID nº 36 (la secuencia de aminoácidos del listado de secuencias) o de SEC ID nº 46 (la secuencia de aminoácidos de la descripción) o de SEC ID nº 45 (la secuencia de nucleótidos de la descripción) en el documento WO2009/059234, Novozymes Inc., EEUU. Las identidades dentro de los polipéptidos de longitud completa eran de 94%.

45 La longitud del gen *Te\_RF8069\_cel6A* (SEC ID nº 10), codificante de la celobiohidrolasa CBHII/Cel6A de *Talaromyces emersonii* (SEC ID nº 16), es de 1.754 pb (incluyendo el codón de parada). Se encontraron siete intrones putativos de longitudes 50, 44, 52, 56, 53, 59 y 60 pb. De esta manera, la región codificante de *Te\_RF8069\_cel6A* es de 1.377 pb (codón de parada no incluido) y la secuencia de proteína deducida consiste en 459 aminoácidos, incluyendo una secuencia de señal predicha de 19 aminoácidos (SignalP V3.0; Nielsen et al., 1997; Nielsen y Krogh, 1998, y Bendtsen et al., 2004). La masa molecular predicha del enzima maduro, excluyendo la secuencia de señal, era de 46.618 Da y el pl predicho, de 4,27. El enzima *Te\_RF8069\_Cel6A* de la invención mostraba una homología más elevada respecto al polipéptido de *Talaromyces emersonii* (Q8N1B5 en la figura 3A-C del documento WO 2006/074005, Novozymes Inc., EEUU) y la celobiohidrolasa II de *Talaromyces emersonii* (AY075018). Las identidades dentro de las secuencias de aminoácidos de longitud completa eran de 90%.

55 De esta manera, según una forma de realización, es una secuencia aislada de polipéptido o molécula aislada de ácidos nucleicos, que codifica un enzima celobiohidrolasa CBHII/Cel6A o polipéptido que comprende la secuencia de aminoácidos de la forma de longitud completa del enzima *At\_ALKO4245\_Cel6A* caracterizado en la SEC ID nº 12, el enzima *Ma\_ALKO4237\_Cel6A* caracterizado en la SEC ID nº 14, *Ct\_ALKO4265* caracterizado en la SEC ID nº 15 o el enzima *Te\_RF8069\_Cel6A* caracterizado en la SEC ID nº 16.

60 Además, también se dan a conocer moléculas de ácidos nucleicos o secuencias polinucleotídicas que codifican un polipéptido CBHII/Cel6A que presentan actividad de celobiohidrolasa y una identidad de por lo menos 76%, preferentemente por lo menos 78%, 80%, 82%, 84%, 86%, 88%, 90%, 92%, 94%, 96% o 97%, más preferentemente por lo menos de 98%, todavía más preferentemente de 99% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 12. Las identidades de los dos enzimas se comparan dentro de las regiones de secuencia correspondientes, es decir, dentro de la región de longitud completa del polipéptido CBHII/Cel6A.

Se da a conocer además es un enzima celobiohidrolasa CBHII/Cel6A fúngico que muestra una identidad de por lo menos 76%, preferentemente de por lo menos 78%, 80%, 82%, 84%, 86%, 88%, 90%, 92%, 94%, 96% o 97%, más preferentemente de por lo menos 98%, todavía más preferentemente de 99% respecto a la longitud completa de la secuencia de aminoácidos de SEC ID nº 14.

Se dan a conocer además celobiohidrolasas CBHII/Cel6A que muestran una identidad de por lo menos 95%, preferentemente de por lo menos 96% o de por lo menos 97%, todavía más preferentemente de por lo menos 98%, todavía más preferentemente de por lo menos 99% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 15.

Se dan a conocer además polipéptidos CBHII/Cel6A que muestran una identidad de por lo menos 91%, preferentemente de por lo menos 92%, 93%, 94%, 95%, 96% o 97%, más preferentemente de por lo menos 98%, y todavía más preferentemente de 99% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 16.

Según una forma de realización es una molécula aislada de ácidos nucleicos que comprende una secuencia polinucleotídica codificante de un polipéptido, que presenta la secuencia de aminoácidos de la celobiohidrolasa CBHII/Cel6A de longitud completa de la invención, así como fragmentos o variantes naturales o sintéticos de los polipéptidos de la invención que presentan actividad de celobiohidrolasa o propiedades similares a los polipéptidos de longitud completa. Dicho fragmento puede ser, por ejemplo, un enzima que no presenta la secuencia de señal de secreción o dominio de unión a carbohidrato o DUC.

Una forma de realización es una molécula aislada de ácidos nucleicos que comprende una "secuencia codificante" de la secuencia polinucleotídica incluida en SEC ID nº 11, SEC ID nº 13, SEC ID nº 9 o SEC ID nº 10 o una subsecuencia de los mismos. Según una forma de realización preferente de la invención, el polipéptido se encuentra codificado por la molécula de ácidos nucleicos que presenta la secuencia de nucleótidos SEC ID nº 11 que comprende la secuencia codificante del enzima *At\_ALKO4245\_Cel6A*. La expresión "secuencia codificante" se refiere a la secuencia de nucleótidos que se inicia desde el codón de inicio de traducción (ATG) y se para en el codón de parada de traducción (TAA, TAG o TGA) y puede comprender regiones de intrón. El polipéptido de longitud completa traducido habitualmente se inicia con metionina.

Según otra forma de realización, la molécula aislada de ácidos nucleicos comprende una secuencia codificante de una secuencia polinucleotídica incluida en el plásmido pALK2582 depositado en *Escherichia coli* RF8175 bajo el número de acceso DSM 22946, el plásmido pALK2581 depositado en *Escherichia coli* RF8174 bajo el número de acceso DSM 22945, el plásmido pALK2904 depositado en *Escherichia coli* RF8214 bajo el número de acceso DSM 22947, o el plásmido pALK3006 depositado en *Escherichia coli* RF8333 bajo el número de acceso DSM 23185, codificando dichas secuencias polinucleotídicas las celobiohidrolasas *At\_ALKO4245\_Cel6A*, *Ma\_ALKO4237\_Cel6A*, *Ct\_ALKO4265\_Cel6A* y *Te\_RF8069\_Cel6A*, respectivamente.

La molécula de ácidos nucleicos de la invención también puede ser un análogo de la secuencia de nucleótidos caracterizada anteriormente. La "degeneración" se refiere a análogos de la secuencia de nucleótidos que difieren en uno o más nucleótidos o codones pero que codifican el polipéptido CBHII/Cel6A recombinante de la invención.

Según una forma de realización, la molécula de ácidos nucleicos también puede ser una molécula de ácidos nucleicos o secuencia polinucleotídica hibridante bajo condiciones restrictivas con una secuencia polinucleotídica contenida en los plásmidos pALK2582, pALK2581, pALK2904 o pALK3006 depositados en *E. coli* bajo los números de acceso DSM 22946, DSM 22945, DSM 22947 o DSM 23185 o una subsecuencia de los mismos, y codificante de un polipéptido que presenta actividad de celobiohidrolasa y una secuencia de aminoácidos que dentro de la región de secuencia correspondiente muestra una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa tla como se ilustra en SEC ID nº 12, una identidad de por lo menos 76% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 14, una identidad de por lo menos 95% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 15 o de por lo menos 91% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 16. El ADN hibridante puede originarse a partir de un hongo perteneciente al género *Acremonium*, *Melanocarpus*, *Chaetomium* o *Talaromyces* o puede originarse a partir de otra especie fúngica.

Según una forma de realización preferente de la invención, el enzima celobiohidrolasa CBHII/Cel6A fúngico se encuentra codificado por una molécula aislada de ácidos nucleicos que comprende la secuencia de nucleótidos de SEC ID nº 11 codificante de la forma de longitud completa del enzima *At\_ALKO4245\_Cel6A* de SEC ID nº 12.

La presente invención se refiere además a un vector de expresión recombinante o constructo de expresión recombinante, que puede utilizarse para propagar o expresar la molécula de ácidos nucleicos o secuencia polinucleotídica codificante de la celobiohidrolasa CBHII/Cel6A seleccionada en un hospedador procariótico o eucariótico adecuado. El vector de expresión recombinante comprende secuencias de ADN o de ácidos nucleicos que facilitan o dirigen la expresión y la secreción de la secuencia codificante o gen del polipéptido CBHII/Cel6A en

- un hospedador adecuado, tal como promotores, intensificadores, terminadores (incluyendo señales de terminación de transcripción y traducción) y secuencias de señal operablemente ligadas a la secuencia polinucleotídica codificante de dicho polipéptido. El vector de expresión puede comprender además genes marcadores para la selección de las cepas transformantes o el marcador de selección puede introducirse en el hospedador en otro constructo de vector mediante cotransformación. Dichas secuencias reguladoras pueden ser homólogas o heterólogas respecto al organismo de producción o pueden originarse a partir del organismo, a partir del cual se aísla el gen codificante del polipéptido CBHII/Cel6A.
- Son ejemplos de promotores para la expresión del polipéptido CBHII/Cel6A de la invención en hospedadores hongos filamentosos, los promotores de la TAKA amilasa, la proteasa alcalina ALP y la triosa fosfato isomerasa de *A. oryzae*, la lipasa de *Rhizopus miehei*, la glucoamilasa (*glaA*) de *Aspergillus niger* o *A. awamori*, la proteasa de tipo tripsina de *Fusarium oxysporum*, el promotor de la celobiohidrolasa 1 de *Chrysosporium lucknowense*, la celobiohidrolasa I (Cel7A) de *Trichoderma reesei*, etc.
- En levaduras, pueden utilizarse por ejemplo promotores de enolasa (ENO-1), galactoquinasa (GAL1), alcohol deshidrogenasa (ADH2) y 3-fosfoglicerato quinasa de *S. cerevisiae* para proporcionar la expresión.
- Son ejemplos de secuencias de promotor para dirigir la transcripción del polipéptido CBHII/Cel6A de la invención en un hospedador bacteriano, el promotor del operón *lac* de *Escherichia coli*, el promotor de *dagA* de la agarasa de *Streptomyces coelicolor*, el promotor del gen alfa-amilasa (*amyL*) de *B. licheniformis*, el promotor del gen amilasa maltogénica (*amyM*) de *B. stearothermophilus*, los promotores de los genes *xylA* y *xylB* de *B. subtilis*, etc.
- Entre los terminadores adecuados se incluyen los de los genes anteriormente mencionados o cualesquiera otras secuencias de terminador caracterizadas.
- Entre los marcadores de transformación o selección adecuados se incluyen aquellos que complementan un defecto o añaden una nueva característica, por ejemplo una actividad enzimática, en el hospedador, por ejemplo los genes *dal* de *B. subtilis* o *B. licheniformis* o *amdS* y *niaD* de *Aspergillus*. La selección también puede basarse en un marcador que confiera resistencia a un antibiótico, tal como ampicilina, canamicina, cloranfenicol, tetraciclina, fleomicina o higromicina.
- Resulta preferida la producción extracelular del polipéptido CBHII/Cel6A de la invención. De esta manera, el vector recombinante comprende secuencias que facilitan la secreción en el hospedador seleccionado. La secuencia de señal de la celobiohidrolasa CBHII/Cel6A de la invención o la presecuencia o prepéptido puede incluirse en el vector de expresión recombinante o la secuencia de señal natural puede sustituirse por otra secuencia de señal capaz de facilitar la expresión en el hospedador seleccionado. De esta manera, la secuencia de señal seleccionada puede ser homóloga o heteróloga respecto al hospedador de expresión.
- Son ejemplos de secuencias de señal adecuadas aquellas de los organismos fúngicos o de levadura, por ejemplo secuencias de señal de genes bien expresados. Dichas secuencias de señal son bien conocidas de la literatura.
- El vector recombinante puede comprender además secuencias que faciliten la integración del vector en el ADN cromosómico del hospedador a fin de obtener una expresión estable. El vector también puede ser un constructo de fusión y comprende secuencias codificantes de un polipéptido portador, que se encuentra genéticamente fusionado con la misma secuencia codificante que la secuencia codificante de la proteína de interés y que mejora la secreción del polipéptido en un organismo hospedador heterólogo o que facilita la purificación de la proteína tras la producción. Entre dichos portadores se incluyen, por ejemplo, las proteínas producidas en grandes cantidades por el hospedador nativo, tal como las celulasas de *T. reesei* o las glucoamilasas de las especies de *Aspergillus*. El polipéptido portador también puede ser un dominio intacto de la proteína secretable, tal como el DUC. La proteína portadora y la proteína de interés puede mantenerse como una fusión después de la secreción o separarse química o bioquímicamente después de la secreción al sobrenadante.
- Según una forma de realización, las celobiohidrolasas *At\_ALKO4245\_Cel6A*, *Ma\_ALKO4237\_Cel6A*, *Ct\_ALKO4265\_Cel6A* y *Te\_RF8069\_Cel6A* se expresaron con su propia secuencia de señal a partir del promotor de *cbh1* (*cel7A*) de *Trichoderma reesei*, tal como se indica en el Ejemplo 2. El constructo de expresión utilizado para transformar el hospedador *T. reesei* incluía también el terminador *cbh1* y el marcador *amdS* para la selección de los transformantes separándolos de las células no transformadas.
- La presente invención se refiere además a células hospedadoras que comprenden el vector de expresión recombinante tal como se ha indicado anteriormente. Los hospedadores adecuados para la producción de la celobiohidrolasa CBHII/Cel6A fúngica son hospedadores homólogos o heterólogos, tales como los hospedadores microbianos, incluyendo bacterias, levaduras y hongos. Los sistemas de producción en células vegetales o de mamífero también son posibles.
- Los hongos filamentosos, tales como *Trichoderma*, *Aspergillus*, *Fusarium*, *Humicola*, *Chrysosporium*, *Neurospora*, *Rhizopus*, *Penicillium* y *Mortierella* son hospedadores de producción preferentes debido a la facilidad del

procesamiento y recuperación posteriores del producto enzimático. Son sistemas de producción y expresión del hospedador adecuados, por ejemplo, el sistema de producción desarrollado para el hospedador hongo filamentoso *Trichoderma reesei* (documento EP 244234), o los sistemas de producción de *Aspergillus*, tales como las cepas de tipo *A. oryzae* o *A. niger* (documento WO 97/08325, patentes US nº 5.843.745 y nº 5.770.418), *A. awamori*, *A. sojae* y *A. japonicus*, o el sistema de producción desarrollado para *Fusarium*, tal como *F. oxysporum* (Malardier et al., 1989) o *F. venenatum*, y para *Neurospora crassa*, *Rhizopus miehei*, *Mortierella alpinis*, *H. lanuginosa* o *H. insolens*, o para *Chrysosporium lucknowense* (patente US nº 6.573.086). Los sistemas de producción adecuados desarrollados para levaduras son sistemas desarrollados para *Saccharomyces*, *Schizosaccharomyces* o *Pichia pastoris*. Los sistemas de producción adecuados desarrollados para bacterias son un sistema de producción desarrollado para *Bacillus*, por ejemplo para *B. subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, para *E. coli* o para el actinomiceto *Streptomyces*. Preferentemente, la celobiohidrolasa CBHII/Cel6A de la invención se produce en un hospedador hongo filamentoso del género *Trichoderma* o *Aspergillus*, tal como cepas de tipo *T. reesei* o *A. niger*, *A. oryzae*, *A. sojae*, *A. awamori* o *A. japonicus*. Según la forma de realización más preferente de la invención, la celobiohidrolasa CBHII/Cel6A fúngica se produce en *T. reesei*.

La célula hospedadora de producción puede ser homóloga o heteróloga respecto a la celobiohidrolasa CBHII/Cel6A de la invención. Preferentemente, el hospedador recombinante se modifica para expresar y secretar enzimas celulolíticas como actividad principal o como una de sus actividades principales. Lo anterior puede llevar a cabo delecionando genes secretados homólogos principales, por ejemplo las cuatro celulasas principales de *Trichoderma* y dirigiendo genes heterólogos a un locus que ha sido modificado para garantizar niveles elevados de expresión y producción. Por ejemplo, el hospedador puede encontrarse libre de celobiohidrolasas homogéneas debido a la eliminación de dichas celobiohidrolasas mediante inactivación o eliminación de una o más celobiohidrolasas del hospedador, por ejemplo mediante delección del gen o genes codificantes de dicha celobiohidrolasa o celobiohidrolasas homogéneas u homólogas.

La presente invención se refiere además a un procedimiento para la producción de un polipéptido CBHII/Cel6A que presenta actividad de celobiohidrolasa, comprendiendo dicho procedimiento las etapas de cultivar la célula hospedadora natural o recombinante portadora del vector de expresión recombinante para una celobiohidrolasa CBHII/Cel6A de la invención bajo condiciones adecuadas y opcionalmente aislar dicho enzima. El medio de producción puede ser un medio adecuado para cultivar el organismo hospedador y que contiene inductores para la expresión eficiente. Los medios adecuados son bien conocidos de la literatura.

La invención se refiere a un polipéptido que presenta actividad de celobiohidrolasa, estando codificado dicho polipéptido por la molécula de ácidos nucleicos de la invención y que puede obtenerse mediante el procedimiento indicado anteriormente.

La invención se refiere además a un procedimiento para la obtención de una preparación enzimática que comprende un polipéptido CBHII/Cel6A que presenta actividad de celobiohidrolasa, comprendiendo dicho procedimiento las etapas de cultivar una célula hospedadora portadora del vector de expresión de la invención y preparar el caldo de cultivo completo, o separar las células del medio de cultivo agotado y obtener el sobrenadante que presenta actividad de celobiohidrolasa.

La presente invención se refiere además a una preparación enzimática que comprende el enzima CBHII/Cel6A caracterizado anteriormente. La preparación o composición enzimática presenta actividad de celobiohidrolasa y puede obtenerse mediante el procedimiento según la invención.

Se encuentra comprendida dentro de la invención una preparación enzimática que comprende la celobiohidrolasa CBHII/Cel6A fúngica de la invención.

Dicha preparación enzimática puede comprender a demás diferentes tipos de enzimas además de la celobiohidrolasa CBHII/Cel6A de la presente invención, por ejemplo otra celulasas, incluyendo una celobiohidrolasa, endoglucanasa y beta-glucosidasa, beta-glucanasa, una amilasa, una lipasa, cutinasa, una proteasa, xilanasas, beta-xilosidasa, mananasa, beta-manosidasa,  $\alpha$ -glucuronidasa, acetil-xilán-esterasa,  $\alpha$ -arabinofurnosidasa,  $\alpha$ -galactosidasa, pectinasa, que implica endo- y exo- $\alpha$ -L-arabinasas,  $\alpha$ -galactosidasa, endo- y exo-galactoronasa, xiloglucanasa, endopectinliasa, pectato liasa y pectinesterasa, fenol esterasa, ligninasa, que implica lignina peroxidasa, peroxidasa dependiente de manganeso, enzima generador de H<sub>2</sub>O<sub>2</sub> y/o una oxidasa, tal como una lacasa o peroxidasa con o sin un mediador. Se espera que estos enzimas potencien el rendimiento del enzima CBHII/Cel6A de la invención mediante la eliminación de los carbohidratos, proteínas y aceites o grasas presentes en el material que debe manipularse. Según una forma de realización, dichos enzimas pueden ser enzimas naturales o recombinantes producidos por la cepa hospedadora o pueden añadirse al sobrenadante de cultivo tras el procedimiento de producción. Las composiciones enzimáticas o preparaciones enzimáticas pueden contener cualquier combinación de dichos enzimas. Según una forma de realización, las celobiohidrolasas CBHII/Cel6A también pueden utilizarse en combinación con preparaciones enzimáticas disponibles comercialmente.

Según una forma de realización, los enzimas necesarios para la hidrólisis del material celulósico según la invención pueden añadirse en una cantidad enzimáticamente eficaz simultáneamente, por ejemplo en forma de una mezcla de

enzimas, o secuencialmente o como parte del procedimiento simultáneo de sacarificación y fermentación (PSSF) o pueden ser producidos por el microorganismo fermentador en el bioprocedimiento consolidado.

5 Según una forma de realización, la preparación o composición enzimática puede contener los enzimas en forma por lo menos parcialmente purificada y aislada. Incluso puede consistir esencialmente del enzima o enzimas deseados. Alternativamente, la preparación puede ser un caldo de cultivo completo o medio de cultivo agotado o filtrado que contiene uno o más enzimas deseados. En un bioprocedimiento consolidado, los enzimas pueden ser producidos por el microorganismo fermentador utilizado en el procedimiento. Además de la actividad enzimática, la preparación puede contener aditivos, tales como mediadores, estabilizadores, tampones, conservantes, surfactantes y/o componentes del medio de cultivo.

15 Según una forma de realización, los aditivos son los utilizados comúnmente en preparaciones enzimáticas destinadas a una aplicación particular. Los surfactantes resultan útiles para emulsionar grasas y humectar superficies. Según una forma de realización, el surfactante puede ser no iónico, incluyendo semipolar y/o aniónico y/o catiónico y/o zwitteriónico. Pueden añadirse tampones a la preparación enzimática para modificar el pH o afectar al rendimiento o estabilidad de otros ingredientes. Según una forma de realización, entre los estabilizadores adecuados se incluyen polioles, tales como propilenglicol o glicerol, un azúcar o alcohol de azúcar, ácido láctico, ácido bórico o derivados del ácido bórico, péptidos, etc. Se utiliza agente blanqueador para oxidar y degradar los compuestos orgánicos. Son ejemplos de sistemas de blanqueamiento químico adecuados, las fuentes de H<sub>2</sub>O<sub>2</sub>, tales como perborato o percarbonato con o sin activadores de blanqueamiento formadores de perácido, tales como tetraacetiletildiamina, o alternativamente, peroxiácidos, por ejemplo de tipo amida, imida o sulfona. Según una forma de realización, pueden sustituirse los oxidantes químicos parcial o totalmente por la utilización de enzimas oxidantes, tales como lacasas o peroxidasas. Muchas lacasas no funcionan eficazmente en ausencia de mediadores. Entre los adyuvantes o agentes acomplejantes se incluyen sustancias tales como zeolita, difosfato, trifosfato, carbonato, citrato, etc. La preparación enzimática puede comprender además uno o más polímeros, tales como carboximetilcelulosa, poli(etilenglicol), poli(alcohol vinílico), poli(vinilpirrolidona), etc. Además, según una forma de realización, pueden añadirse suavizantes, agentes cáusticos, conservantes para evitar el deterioro de otros ingredientes, abrasivos y sustancias modificadoras de las propiedades espumantes y de viscosidad.

30 Según una forma de realización preferida de la invención, dicha preparación enzimática se encuentra en forma de medio de cultivo agotado, líquido, polvos o granulado. Preferentemente, la preparación enzimática es medio de cultivo agotado.

35 La presente invención se refiere además a diversos usos de la celobiohidrolasa CBHII/Cel6A fúngica de la invención, en la que se desea la hidrólisis o modificación de material celulósico. Entre dichos usos se incluyen cualquier aplicación en la que se utilizan convencionalmente enzimas celulolíticas, tal como en las industrias de los combustibles, textiles, detergentes, pulpa y papel, alimentos, piensos o bebidas. La adición de la celobiohidrolasa CBHII/Cel6A fúngica de la invención a una composición enzimática que comprende otros enzimas degradadores de la celulosa, tal como la celobiohidrolasa I, endoglucanasas y beta-glucosidasas, potencia en gran medida la hidrólisis y conduce a la hidrólisis prácticamente total del esqueleto polimérico de la celulosa en monómeros de glucosa. El producto principal de la acción de CBHII/Cel6A es la celobiosa, compuesta de dos unidades de glucosa.

45 Una forma de realización preferida es la utilización del procedimiento de la invención en aplicaciones que requieren estabilidad/rendimiento de la celobiohidrolasa CBHII/Cel6A a temperaturas moderadas/convencionales o elevadas, es decir, que requiere enzimas termofílicos o termoestables. Es conocido que las temperaturas elevadas incrementan la hidrólisis de la celulosa cristalina presente en los materiales celulósicos o lignocelulósicos, reduciendo de esta manera la cantidad total de enzimas necesaria en la hidrólisis o reduciendo el tiempo de hidrólisis requerido. Además, debido a que a temperaturas elevadas se reduce la viscosidad del sustrato lignocelulósico, los enzimas termoestables permiten trabajar a cargas de sólido más elevadas y un ahorro en los costes de inversión.

50 Otra forma de realización preferida es la utilización del polipéptido CBHII/Cel6A de la invención o de la preparación enzimática que comprende dicho polipéptido en la hidrólisis de material celulósico para la producción de biocombustible que comprende etanol, propanol, butanol y similares. En la producción de biocombustible, las celobiohidrolasas CBHII/Cel6A de la invención son como otros enzimas celulolíticos especialmente adecuados para producir monómeros de glucosa a partir de material celulósico polimérico que seguidamente puede ser fermentado por cepas de levadura produciendo etanol y utilizarse como combustible.

60 Según una forma de realización, el material lignocelulósico puede pretratarse antes de la hidrólisis enzimática para romper la estructura de fibras de los sustratos celulósicos y permitir que la fracción celulósica resulte más accesible a los enzimas celulolíticos. Entre los pretratamientos actuales se incluyen los procedimientos mecánicos, químicos o térmicos y las combinaciones de los mismos. El material puede, por ejemplo, pretratarse mediante explosión con vapor o hidrólisis ácida. El procedimiento de sacarificación, es decir, la producción de monómeros azúcar, y la fermentación por parte de cepas de levaduras pueden llevarse a cabo separada o simultáneamente en el mismo reactor. La celobiohidrolasa CBHII/Cel6A de la invención presenta una gran ventaja respecto a los enzimas comerciales que se comercializan actualmente en el aspecto de que también es estable a temperaturas elevadas, es



decir, es termoestable. La hidrólisis de materiales celulósicos o lignocelulósicos es conocido que se potencia a temperaturas elevadas y rinde monómeros de azúcar más eficientemente.

5 Según una forma de realización, la utilización de los polipéptidos CBHII/Cel6A de la invención permite la utilización de elevadas consistencias de la biomasa y conduce a concentraciones elevadas de azúcares y etanol. Este enfoque puede conducir a ahorros significativos de energía y de costes de inversión. La temperatura elevada también reduce el riesgo de contaminación durante la hidrólisis.

10 Los hidrolizados de azúcares también pueden servir como materia prima para otros procedimientos no microbianos, por ejemplo para el enriquecimiento, aislamiento y purificación de azúcares altamente valiosos o para diversos procedimientos de polimerización.

15 Los monómeros de glucosa también pueden utilizarse como intermediarios o materias primas para la producción de diversos bloques químicos o constructivos para los procedimientos de la industria química, por ejemplo en la denominada biorefinería.

En las industrias de la pasta y del papel, los polipéptidos pueden utilizarse para modificar la fibra celulósica, por ejemplo en el tratamiento de la pasta Kraft, la pasta mecánica o el papel reciclado.

20 En la industria textil, las celobiohidrolasas CBHII/Cel6A encuentran aplicaciones en la suavización y/o mejora del tacto de los tejidos de algodón y en la eliminación de colorantes índigo en sustitución del lavado a la piedra.

En la industria de los piensos, las celobiohidrolasas CBHII/Cel6A pueden utilizarse en la degradación de los materiales celulósicos o hemicelulósicos presentes en las materias primas.

25 En la industria de los detergentes, las celobiohidrolasas CBHII/Cel6A pueden utilizarse en composiciones de lavandería manual o automática o de lavavajillas para la eliminación de manchas celulósicas.

30 La invención se ilustra mediante los ejemplos no limitativos siguientes. A partir de los resultados experimentales puede concluirse que las celobiohidrolasas CBHII/Cel6A fúngicas de la invención son capaces de satisfacer las muy variables demandas de las diferentes industrias que requieren una hidrólisis eficiente del material celulósico presente en diversas materias primas.

### 35 Ejemplo 1. Clonación de los genes celobiohidrolasa 2 (*cbh2*)

Se utilizaron procedimientos estándar de biología molecular en el aislamiento y tratamientos enzimáticos de ADN (por ejemplo el aislamiento de ADN plasmídico, la digestión del ADN para producir fragmentos de ADN), en transformaciones de *E. coli*, en la secuenciación, etc. Los procedimientos básicos utilizados fueron los descritos por el fabricante del enzima, reactivo o kit o los descritos en manuales estándares de biología molecular, por ejemplo Sambrook y Russell (2001). El aislamiento del ADN genómico se llevó a cabo tal como indican en detalle Raeder y Broda (1985).

45 Se seleccionaron cuatro cepas fúngicas termófilas de la colección de cultivos de Roal Oy para la clonación, basándose en los resultados anteriores de que las cepas producen celulasas termoestables (documento WO 2007/071818; Maheshwari et al., 2000; Murray et al., 2003; Miettinen-Oinonen et al., 2004). Las sondas para la clonación de los genes *cbh2* de *Acremonium thermophilum* ALKO4245 y *Melanocarpus albomyces* ALKO4237 se sintetizaron mediante PCR. Los oligos degenerados se planificaron basándose en la alineación de las secuencias de aminoácidos anteriormente publicadas de las proteínas celobiohidrolasa II (CBHII). Las secuencias de los cebadores homólogos para la clonación de los genes *cbh2* de las cepas *Chaetomium thermophilum* ALK4265 y *Talaromyces emersonii* RF8069 se obtuvieron de las secuencias de nucleótidos publicadas (AY861348, DQ020255, CQ838150; Murray et al., 2003; AY075018, AF439936). Las secuencias de los cebadores se muestran en la Tabla 2 (SEC ID nº 1 a nº 6).

55 Tabla 2. Oligonucleótidos utilizados como cebadores de PCR para amplificar sondas para el cribado de genes *cbh2* de *Acremonium thermophilum* ALKO4245 y *Melanocarpus albomyces* ALKO4237 y para amplificar genes *cbh2* de longitud completa de *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069.

ADN genómico molde de	Oligonucleótido	Longitud (pb)	Secuencia <sup>(a)</sup>	SEC ID nº
ALKO4245	CBH_1S	17	TGGGGNCARTGYGGNGG (s)	1
	CBH_1AS	17	GCNGGCCANCCNARCCA (as)	2
ALKO4237	CBH_1S	17	TGGGGNCARTGYGGNGG (s)	1
	CBH_1AS	17	GCNGGCCANCCNARCCA (as)	2
ALKO4265	CBH_8	21	ATGGCTAAGCAGCTGCTGCTC (s)	3

	CBH_9	20	TCAGARCGGAGGGTTGGCAT(as)	4
RF8069	Te_CBH_A	44	TATTATCCCGCGGACTGCGCATCATGCGGAATCTT CTTGCTCTTG (s)	5
	Te_CBH_B	38	AATTTGGATCCTCAGAACAGCGGGTTAGCATTTCG TGAG (as)	6

<sup>(a)</sup> R=A o G, N=A o G o T o C, Y=T o C; "s" entre paréntesis=cadena de sentido, "as" entre paréntesis=cadena antisentido

Se amplificaron las sondas mediante PCR con los cebadores indicados en la Tabla 2 utilizando el ADN genómico como molde en las reacciones. Las mezclas de PCR de *Acremonium thermophilum* ALKO4245 y *Melanocarpus albomyces* ALKO4237 contenían en Tris-HCl 10 mM, pH 8,8, KCl 50 mM, Triton X-100 al 0,1%, MgCl<sub>2</sub> 1,5 mM, dNTP 0,2 mM, 5 µM de cada cebador y 1 a 2 unidades de ADN polimerasa Dynazyme EXT (Finnzymes, Finlandia) y 0,5 a 1 µg del ADN genómico correspondiente. Las condiciones para las reacciones de PCR eran 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 60°C (±5°C, gradiente), 2 min. de extensión a 72°C y una extensión final a 72°C durante 10 min. Para la clonación de PCR de los genes *cbh2* de las cepas *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069, las reacciones contenían los tampones 1x Phusion GC y 1x Phusion HF, respectivamente, con 0,2 mM de los dNTP, 5 µM de cada cebador y 1 a 2 unidades de ADN polimerasa Phusion (Finnzymes, Finlandia) y 0,5 a 1 µg del ADN genómico correspondiente. Las condiciones para las reacciones de PCR fueron las siguientes: 3 min. de desnaturalización inicial a 98°C, seguido de 30 ciclos de 30 s a 98°C, 30 s de hibridación a 55°C (±5°C, gradiente), 1 a 2 min. de extensión a 72°C y una extensión final a 72°C durante 10 min.

Las combinaciones de cebadores indicadas en la Tabla 2 produjeron un producto de ADN específico que presentaba el tamaño esperado (según los cálculos basados en las secuencias de *cbh2* publicadas). Los productos de ADN se aislaron y se purificaron a partir de las mezclas de reacción de PCR y se clonaron en el vector pCR®4Blunt-TOPO® siguiendo las instrucciones del fabricante (Invitrogen, EEUU). Las inserciones se caracterizaron mediante secuenciación y mediante la realización de hibridaciones de transferencia southern con los ADN genómicos digeridos con varios enzimas de restricción. Los fragmentos de PCR, que habían sido seleccionados para la utilización como sondas para la clonación génica de las cepas *Acremonium thermophilum* ALKO4245 y *Melanocarpus albomyces* AKO4237, se presentan en la Tabla 3. Además, la tabla indica los fragmentos de PCR que contienen los genes *cbh2* de longitud completa de las cepas *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069.

**Tabla 3.** Cebadores utilizados en las reacciones de PCR, sondas seleccionadas para el cribado de los genes *cbh2* de *Acremonium thermophilum* ALKO4245 y *Melanocarpus albomyces* ALKO4237 y fragmentos de ADN que contienen los genes *cbh2* de longitud completa de *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069. Se muestra el ADN genómico molde y el nombre del plásmido que contiene el fragmento de sonda.

Gen	Cebador directo	Cebador inverso	ADN genómico utilizado como molde en la reacción de PCR	Fragmento obtenido (kb)	Inserción en el plásmido	SEC ID n°
ALKO4245_cel6A	CBH_1S	CBH_1AS	ALKO4245	1,0 kb	pALK2580	7
ALKO4237_cel6A	CBH_1S	CBH_1AS	ALKO4237	0,8 kb	pALK2576	8
ALKO4265_cel6A	CBH_8	CBH_9	ALKO4265	1,8 kb	pALK2904	9
RF8069_cel6A	Te_CBH_A	Te_CBH_B	RF8069	1,8 kb	pALK3006	10

Las secuencias de aminoácidos deducidas de la totalidad de dichos fragmentos de PCR presentó homología respecto a las secuencias de CBHII/Cel6A publicadas (programa BLAST, versión 2.2.9 en el NCBI, National Center for Biotechnology Information; Altschul et al., 1990). El fragmento de PCR de 1.757 pb en el plásmido pALK2904 (SEC ID n° 9) y el fragmento de PCR de 1.788 pb en el plásmido pALK3006 (SEC ID n° 10) contenían los genes *cbh2/cel6A* de longitud completa de *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069, respectivamente. El gen codificante de *Chaetomium thermophilum* ALKO4265 se denominó Ct\_ALKO4265\_cel6A y el gen de *Talaromyces emersonii* RF8069 se denominó Te\_RF8069\_cel6A. Las cepas de *E. coli* RF8214 (=pALK2904) y RF8333 (=pALK3006) se depositaron en la colección del DSM bajo los números de acceso DSM 22947 y DSM 23185, respectivamente.

Los ADN genómicos de *Acremonium thermophilum* ALKO4245 y de *Melanocarpus albomyces* ALKO4237 se digirieron con varios enzimas de restricción para el análisis de transferencia southern. Las sondas para la hibridación eran los fragmentos de EcoRI de 1.032 pb (SEC ID n° 7) y de 831 pb (SEC ID n° 8), cortados de los plásmidos pALK2580 y pALK2576, respectivamente. Las sondas anteriormente indicadas se marcaron mediante la utilización de digoxigenina siguiendo las instrucciones del proveedor (Roche, Alemania). Las hibridaciones se llevaron a cabo

durante la noche a 68°C. Tras la hibridación, los filtros se lavaron 2x5 min. a TA utilizando 2xSSC-SDS al 0,1% seguido de 2x15 min. a 68°C utilizando 0,1xSSC-SDS al 0,1%.

A partir del ADN genómico de *Acremonium thermophilum* ALKO4245, se hibridó el fragmento digerido con XbaI de aproximadamente 8 kb utilizando el fragmento de EcoRI de 1.032 pb marcado con digoxigenina de pALK2580 a modo de sonda. Correspondientemente, se hibridó el fragmento digerido con BamHI de aproximadamente 4,5 kb con el fragmento de EcoRI de 831 pb marcado con digoxigenina de pALK2576 del ADN genómico de *Melanocarpus albomyces* ALKO4237. Los fragmentos de ADN genómico hibridantes se aislaron del pool de fragmentos genómicos digeridos basándose en su tamaño. Los fragmentos genómicos se aislaron del gel de agarosa y se clonaron en vectores pBluescript II KS+ (Stratagene, EEUU) con XbaI (*Acremonium thermophilum* ALKO4245) o BamHI (*Melanocarpus albomyces* ALKO4237). Las mezclas de ligación se transformaron en células de *Escherichia coli* XL10-Gold (Stratagene) y se sembraron en placas de LB (Luria-Bertani) que contenían 50 a 100 µg/ml de ampicilina. Se cribaron las colonias de *E. coli* para los clones positivos utilizando la hibridación colonial con las inserciones pALK2580 y pALK2576 como sondas bajo las condiciones de hibridación correspondientes a las indicadas anteriormente para los análisis de transferencia southern, excepto en que se utilizó una temperatura de hibridación de 65°C en lugar de 68°C. Se recolectaron varios clones positivos de las placas. Se demostró mediante digestión de restricción que contenían inserciones de los tamaños esperados. Se secuenció el gen de longitud completa codificante de CBHII/Cel6A de *Acremonium thermophilum* ALKO4246 (*At*\_ALKO4245\_cel6A, SEC ID nº 11) a partir de la inserción XbaI de 7 kb y el plásmido que contenía dicha inserción se denominó pALK2582. La cepa *E. coli* RF8175 que incluía el plásmido pALK2582 se depositó en la colección del DSM bajo el número de acceso DSM 22946. El gen codificante de la proteína de *Acremonium thermophilum* ALKO4245 se denominó *At*\_ALKO4245\_cel6A. Correspondientemente, el gen de longitud completa codificante de CBHII/Cel6A de *Melanocarpus albomyces* ALKO4237 (*Ma*\_ALKO4237\_cel6A, SEC ID nº 12) se secuenció a partir de la inserción BamHI de 5 kb y el plásmido que contenía dicha inserción se denominó pALK2581. La cepa *E. coli* RF8174 que incluía el plásmido pALK2581 se depositó en la colección del DSM bajo el número de acceso DSM 22945. El gen codificante de *Melanocarpus albomyces* ALKO4237 se denominó *Ma*\_ALKO4237\_cel6A. La información relevante sobre los genes y las secuencias de proteína deducidas (SEC ID nº 9 a nº 16) se resume en las Tablas 4 y 5, respectivamente.

Tabla 4. Resumen de los genes *cbh2/cel6A* aislados de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069.

Gen	Longitud incluyendo intrones (pb) <sup>(a)</sup>	Región codificante (pb) <sup>(b)</sup>	Nº de intrones putativos	Longitudes de intrones putativos (pb)	SEC ID nº
<i>At</i> _ALKO4245_cel6A	1830	1434	4	79, 72, 117, 126	11
<i>Ma</i> _ALKO4237_cel6A	1607	1416	2	93, 95	13
<i>Ct</i> _ALKO4265_cel6A	1757	1425	3	77, 196, 56	9
<i>Te</i> _RF8069_cel6A	1754	1377	7	50, 44, 52, 56, 53, 59, 60	10

<sup>(a)</sup> El codón de PARADA se encuentra incluido.

<sup>(b)</sup> El codón de PARADA no se encuentra incluido.

Tabla 5. Resumen de las secuencias de aminoácidos deducidas a partir de las secuencias de los genes *cbh2/cel6A* de *Acremonium thermophilum* ALKO4246, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069.

Proteína CBHII/Cel6A	Nº de aminoácidos	Longitud de la ss NN/HMM <sup>(a)</sup>	DUC <sup>(b)</sup>	PM predicho (Da), ss no incl. <sup>(c)</sup>	pl predicho, ss no incl.	SEC ID nº
<i>At</i> _ALKO4245_Cel6A	478	18	Q26 a L63	48918	4,82	12
<i>Ma</i> _ALKO4237_Cel6A	472	17	Q25 a L62	48627	4,50	14
<i>Ct</i> _ALKO4265_Cel6A	475	17	Q25 a I63	49408	5,31	15
<i>Te</i> _RF8069_Cel6A	459	19	Q20 a V55	46618	4,27	16

<sup>(a)</sup> La predicción de la secuencia de señal se realizó utilizando el programa SignalP V3.0 (Nielsen et al., 1997; Nielsen y Krogh, 1998; Bendtsen et al., 2004); el valor de NN se obtuvo utilizando redes neuronales.

<sup>(b)</sup> Se indica el dominio de unión a celulosa (DUC), los aminoácidos de la región del DUC [M1 (met nº 1) incluida en la numeración].

<sup>(c)</sup> No se incluye la secuencia de señal predicha. Se realizó la predicción utilizando la herramienta Compute pI/MW en el servidor ExPASy (Gasteiger et al., 2003).

La comparación entre las secuencias de CBHII/Cel6A deducidas de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* DSM 2432 (RF8069) se presenta en la Tabla 6. Se utilizó un programa del Clone Manager (versión 9) que incluía las funciones "Comparar dos secuencias/Global/Comparar secuencias como aminoácidos/matriz de puntuaciones BLOSUM62"

para determinar el grado de identidad.

5 Tabla 6. Valores de identidad (%) obtenidos de la alineación de las secuencias de aminoácidos deducidas de CBHII/Cel6A de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069. Se alinearon las secuencias de aminoácidos de longitud completa, incluyendo las secuencia de señal. Se utilizó un programa de Clone Manager 9 (Comparar dos secuencias/Global/Comparar secuencias como aminoácidos/matriz de puntuaciones BLOSUM62) para determinar el grado de identidad.

	ALKO4245_ Cel6A	ALKO4237_ Cel6A	ALKO4265_ Cel6A	RF8069_ Cel6A
ALKO4245_Cel6A	100	67	68	66
ALKO4237_Cel6A		100	71	60
ALKO4265_Cel6A			100	61
RF8069_Cel6A				100

10 La comparación entre las secuencias de CBHII/Cel6A deducidas de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069, y las secuencias encontradas en las bases de datos se muestra en las Tablas 7 y 8.

15 Tabla 7. Secuencias de identidad más alta respecto a las secuencias de aminoácidos deducidas de CBHII/Cel6A de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069. Se alinearon las secuencias de aminoácidos de longitud completa, incluyendo las secuencias de señal. Se llevó a cabo la búsqueda en la base de datos utilizando BLAST (tblastn, base de datos nr/nt) y se utilizó el programa Clone Manager 9 (Comparar dos secuencias/Global/Comparar secuencias como aminoácidos/matriz de puntuaciones BLOSUM62) para determinar el grado de identidad.

Organismo y número de acceso	Identidad (%)
<i>At</i> _ALKO4245_Cel6A	100
<i>Podospora anserina</i> , XP_001903170	69
<i>Neurospora crassa</i> , XM_955677	69
<i>Ma</i> _ALKO4237_Cel6A	100
<i>Chaetomium globosum</i> , XP_001226029	75
<i>Ct</i> _ALKO4265_Cel6A	100
<i>Chaetomium thermophilum</i> , AY861348	94
<i>Te</i> _RF8069_Cel6A	100
<i>Talaromyces emersonii</i> , AY075018	90

25 Tabla 8. Secuencias de identidad más alta respecto a las secuencias de aminoácidos deducidas de CBHII/Cel6A de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069. Se alinearon las secuencias de aminoácidos de longitud completa, incluyendo las secuencias de señal. Se llevaron a cabo las búsquedas en las bases de datos del sistema de registro del Chemical Abstracts Service (CAS) y de Patented Protein Sequences del NCBI utilizando BLAST y el programa Clone Manager 9 (Comparar dos secuencias/Global/Comparar secuencias como aminoácidos/matriz de puntuaciones BLOSUM62) para determinar el grado de identidad.

Organismo y número de acceso	Identidad (%)
_ALKO4245_Cel6A	100
US7220565 B2, SEC ID: 2	75
WO2009085868 A1, SEC ID: 49	75
<i>Ma</i> _ALKO4237_Cel6A	100
WO2008095033 A2, SEC ID: 413	72
WO2009085868 A1, SEC ID: 49	71

Ct_ALKO4265_Cel6A	100
EP1578964 B1, SEC ID: 2	94
WO2009059234 A2, SEC ID: 45*)	94
CN1757709A	94
Te_RF8069_Cel6A	100
WO2006074005 A2, figura 3A-C	90

\*) Las descripciones del documento nº WO 2009/59234 se refieren a la SEC ID nº 45 (ADN) o a la SEC ID nº 46 (proteína). El listado de secuencias del documento nº WO 2009/59234 se refiere a la SEC ID nº 35 (ADN) o a la SEC ID nº 36 (proteína).

5

## Ejemplo 2. Producción de proteínas CBHII/Cel6A recombinantes en *Trichoderma reesei*

Se construyeron plásmidos de expresión para la producción de proteínas CBHII/Cel6A recombinantes de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069 en *Trichoderma reesei*. Los plásmidos de expresión construidos se listan en la Tabla 9. Los genes *cbh2/cel6A* recombinantes, incluyendo sus propias secuencias de señal, se fusionaron exactamente con el promotor *cbh1/cel7A* de *T. reesei* mediante PCR. Se garantizó la terminación de la transcripción con el terminador *cbh1/cel7A* de *T. reesei* y se utilizó el gen marcador *amdS* de *A. nidulans* para la selección de los transformantes tal como se indica en Paloheimo et al. (2003). Se aislaron los casetes de expresión lineales (figura 1) a partir de los esqueletos de vector tras la digestión con EcoRI o EcoRI-Spel y se transformaron en protoplastos de *T. reesei*. La cepa hospedadora utilizada no produce ninguna de las cuatro principales celulasas de *T. reesei* (CBHI, CBHII, EGI, EGII). Las transformaciones se llevaron a cabo tal como en Penttilä et al. (1987) con las modificaciones indicadas en Karhunen et al. (1993), seleccionando la acetamida como única fuente de nitrógeno (gen marcador *amdS*). Los transformantes se purificaron en placas de selección mediante conidios individuales antes de esporularlos sobre PD.

Tabla 9. Casetes de expresión construidos para producir proteínas recombinantes CBHII/Cel6A de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069 en *Trichoderma reesei*. La estructura global de los casetes de expresión fue la indicada en la figura 1. Los genes *cbh2/cel6A* clonados se fusionaron exactamente con el promotor *cbh1/cel7A* de *T. reesei*.

Proteína celobiohidrolasa II	Plásmido de expresión	Casete de expresión <sup>(a)</sup>	Terminador <sup>(b)</sup>
At_ALKO4245_Cel6A	pALK2906	9,4 kb EcoRI	125 pb ( <i>Xba</i> I)
Ma_ALKO4237_Cel6A	pALK2901	9,3 kb EcoRI-Spel	258 pb ( <i>Dr</i> alI)
Ct_ALKO4265_Cel6A	pALK2903	9,2 kb EcoRI	
Te_RF8069_Cel6A	pALK3010	7,9 kb EcoRI	

<sup>(a)</sup> Se aisló el casete de expresión para la transformación de *T. reesei* a partir del esqueleto del vector mediante la utilización de la digestión con EcoRI o EcoRI-Spel.

30

<sup>(b)</sup> Número de nucleótidos después del codón de PARADA del gen recombinante clonado que se incluyó en el casete de expresión. El sitio de restricción en el extremo 3' del fragmento génico genómico que se utilizó en la construcción del casete de expresión se indica entre paréntesis. Se extrajo el fragmento génico Ct\_ALKO4265\_cel6A del extremo 3' con EcoRI (un sitio presente en el vector pCR@4Blunt-TOPO®). Correspondientemente, se extrajo el fragmento génico Te\_RF8069\_cel6A de su extremo 3' con BamHI (un sitio creado después del codón de parada en la PCR). Lo anterior no deja ningún terminador original Ct\_ALKO4265\_cel6A o Te\_RF8069\_cel6A en los constructos antes de la secuencia de terminador *cbh1*.

Se analizó la producción de CBHII/Cel6A de los transformantes a partir de los sobrenadantes de cultivo de los cultivos en matraz de agitación. Se inocularon los transformantes desde cultivos sólidos en tubo con PD a matraces de agitación que contenían 50 ml de medio complejo inductor de celulasa basado en lactosa (Joutsjoki et al., 1993) tamponado con KH<sub>2</sub>PO<sub>4</sub> al 5%. Se analizó la producción de proteína CBHII/Cel6A de los transformantes a partir de los sobrenadantes de cultivo tras cultivarlos durante 7 días a 30°C, 250 rpm. Se analizó la producción heteróloga de proteínas recombinantes mediante SDS-PAGE con posterior tinción de Coomassie. Los genotipos de los transformantes seleccionados se confirmaron mediante la utilización de análisis de transferencia southern en los que se incluyeron varias digestiones genómicas y se utilizó el casete de expresión respectivo a modo de sonda.

Se seleccionaron los transformantes mejores productores para el cultivo en biorreactores a escala de laboratorio. Los transformantes se cultivaron en biorreactores de laboratorio a 28°C en el medio complejo inductor de celulasa durante 3 a 4 días con control del pH a 4,4±0,2 (NH<sub>3</sub>/H<sub>3</sub>PO<sub>4</sub>) con el fin de obtener material para los ensayos de aplicación. Se recuperaron los sobrenadantes mediante centrifugación y filtración a través de filtros Seitz-K 150 y EK (Pall SeitzSchenk Filtersystems GmbH, Bad Kreuznach, Alemania).

50

**Ejemplo 3. Hidrólisis de celulosa cristalina (Avicel) utilizando enzimas CBHII/Cel6A recombinantes**

Se caracterizaron las preparaciones de enzima CBHII/Cel6A recombinante en términos de dependencia del pH y estabilidad térmica utilizando celulosa cristalina (Avicel) como sustrato. Se determinó la dependencia del pH y la estabilidad térmica de las proteínas CBHII/Cel6A recombinantes de *Acremonium thermophilum* ALKO4245, *Melanocarpus albomyces* ALKO4237, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069 dentro de un intervalo de pH de 3,0 a 10,0 y un intervalo de temperatura de 40°C a 80°C, respectivamente. Los ensayos de hidrólisis de celulosa cristalina (Ph 101, Avicel; Fluka, Bucsh, Suiza) se llevaron a cabo a escala de tubo de 2,0 ml en acetato sódico 50 mM, pH 6,0. Para la determinación de los óptimos de pH, se sometieron a agitación soluciones de sustrato (Avicel, 50 mg/ml en acetato sódico, pH 6,0) de intervalo de pH de 3 a 10 con las preparaciones de enzima (100 µg de proteína en la reacción) a 50°C y volumen final de la mezcla de reacción de 650 µl. Se continuó la hidrólisis durante 21 horas y se detuvo mediante la adición de 326 µl de reactivo de parada que contenía 9 vol. de etanol al 94% y 1 vol. de glicina 1 M (pH 11). La solución se filtró a través de una unidad de filtración de 0,22 µm Millex GV13 (Millipore, Billerica, MA, EEUU). La formación de azúcares reductores solubles en el sobrenadante se determinó mediante el procedimiento de hidrazida de ácido para-hidroxibenzoico (PAHBAH) (Lever, 1972) utilizando una curva estándar de celobiosa (200 µM a 1.200 µM de celobiosa). Se añadió PAHBAH 0,1 M recién preparado (Sigma-Aldrich, St. Louis, MO, EEUU) en solución de NaOH 0,5 M (200 µl) a 300 µl de la muestra filtrada y sometida a ebullición durante 10 minutos, seguido del enfriamiento de la solución hasta la temperatura ambiente. Se midió la absorbancia a 405 nm de muestras por duplicado con un aparato Multiskan EX (Thermo Labsystems, Franklin, MA, EEUU). Correspondientemente, se determinó la estabilidad térmica de las proteínas CBHII/Cel6A recombinantes en soluciones de sustrato Avicel en un intervalo de temperatura de 40°C a 80°C al pH óptimo durante un tiempo de reacción de 21 horas.

Los resultados indican que el óptimo de pH de CBHII/Cel6A de *Talaromyces emersonii* RF8069 y de CBHII/Cel6A de *Melanocarpus albomyces* ALKO4237 es de 4,0, mientras que los enzimas de *Chaetomium thermophilum* ALKO4265 y de *Acremonium thermophilum* ALKO4245 es de 5,0 (figura 2). Se encontró que la estabilidad térmica era sustancialmente más alta para la CBHII/Cel6A de *Acremonium thermophilum* ALKO4245, *Chaetomium thermophilum* ALKO4265 y *Talaromyces emersonii* RF8069 que para la proteína de *Melanocarpus albomyces* ALKO4237 (figura 3).

**Ejemplo 4. Hidrólisis de sustrato de madera dura con preparaciones enzimáticas que comprendían una celobiohidrolasa CBHII/Cel6A recombinante**

Se suspendió madera dura sometida a explosión con vapor en tampón de acetato sódico 0,05 M, pH 4,8. El peso final de la mezcla de hidrólisis era de 20 g con una concentración de sólidos totales de 2% (p/p). Se hidrolizó el sustrato utilizando diferentes mezclas enzimáticas a una dosis de e5 mg de proteína por cada gramo de sólidos totales en matraces de agitación de 50 ml. Se determinó el contenido de proteínas de los componentes enzimáticos y de las mezclas utilizando el kit de ensayo Pierce BCA (Thermo Scientific, número de producto 23227) con albúmina de suero bovino (Thermo Scientific, número de producto 23209) como estándar. Los matraces de agitación de agitaban en un baño de agua de agitación lineal ajustado a diferentes temperaturas. Para cada punto de muestras, se extrajo una muestra de 1 ml de matraces de agitación por duplicado y se sometieron a ebullición durante 10 minutos para terminar la hidrólisis enzimática, se centrifugaron y el sobrenadante se analizó para los productos de reacción de la hidrólisis. Se prepararon blancos que contenían solo sustrato (adición de únicamente tampón en lugar de enzimas) de manera idéntica a las demás muestras.

Se prepararon tres combinaciones de mezcla separadas (una MEZCLA 2 termofílica, una MEZCLA ACC mesofílica y una MEZCLA ENZIMAS T. REESEI mesofílica) con diferentes sustituciones de Cel6A/CBHII.

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

Preparación de Cel7A/CBHII termofílica que contenía Cel7A de *Thermoascus aurantiacus* ALKO4242 con DUC unido genéticamente de *T. reesei* CBHI/Cel7A (documento nº WO2007/071818).

Preparación de endoglucanasa termofílica que contiene endoglucanasa Cel45A de *Acremonium thermophilum* ALKO4245 (At EG\_40/Cel45A, documento nº WO 2007/071818).

Preparación de β-glucosidasa termofílica que contiene β-glucosidasa de *Thermoascus aurantiacus* ALKO4242 (Ta βG\_81/Cel3A, documento nº WO 2007/071818).

Preparación de xilanasas termofílicas que contienen *Nonomurea flexuosa* Xyn11A (AM24, documento nº WO 2005/100557, AB Enzymes Oy, FI).

Todas las celulasas se produjeron heterológicamente como monoccomponentes en la cepa hospedadora *Trichoderma reesei* que presenta un fondo libre de celulasa (se deletaron los genes codificantes de las cuatro celulasas principales Cel7A/CBHI, Cel6A/CBHII, Cel7B/EGI y Cel5A/EGII). Se utilizaron sobrenadantes de cultivo en bruto en

la mezcla. Se combinaron los componentes enzimáticos de la manera siguiente (por cada 10 ml de mezcla): preparación de CBHI/Cel7A 330 mg (71,2%), preparación de endoglucanasa 105 mg (22,7%), preparación de  $\beta$ -glucosidasa 7,5 (1,6%) y preparación de xilanasas 21 mg (4,5%). El volumen se enrasó a 10 ml con agua corriente. La concentración de proteína final de la mezcla era de 46,35 mg/ml. Dicha mezcla de enzima se denominó MEZCLA 2.

Para someter a ensayo el rendimiento de Cel6A/CBHII en la hidrólisis con MEZCLA 2, se sustituyó 15% (49,5 mg) del componente CBHI/Cel7A de la MEZCLA 2 por *Ma*\_ALKO4237\_Cel6A (MEZCLA 2\_MA), *Ct*\_ALKO4265\_Cel6A (MEZCLA 2\_CT) o *At*\_ALKO4245\_Cel6A (MEZCLA 2\_AT), respectivamente.

Se preparó una mezcla del estado de la técnica mediante la combinación de los componentes siguientes (por cada 10 ml de la mezcla): ECONASE® CE (Roal Oy, un producto enzimático de *T. reesei* clásico) 470 mg (94%), preparación de  $\beta$ -glucosidasa (*At*  $\beta$ G\_101/Cel3A, documento nº WO 2007/071818) 20 mg (4%) y preparación de xilanasas (*Ta* XYN\_30, documento nº WO 2007/071818) 10 mg (2%). El volumen se enrasó a 10 ml con agua corriente. La concentración de proteína final en dicha mezcla era de 50 mg/ml. Dicha mezcla de enzima se denominó MEZCLA ENZIMAS T. REESEI. Los enzimas *At*  $\beta$ G\_101/Cel3A y *Ta* XYN\_30 se produjeron heterológicamente como monocómpoentes en la cepa hospedadora *Trichoderma reesei* que presentaba fondo libre de celulasa.

Para someter a ensayo el rendimiento de Cel6A/CBHII en la hidrólisis con MEZCLA ENZIMAS T. REESEI, el 15% (70,5 mg) del componente ECONASE® CE de MEZCLA ENZIMAS T. REESEI se sustituyó por *Ma*\_ALKO4237\_Cel6A (MEZCLA TR\_MA), *Ct*\_ALKO4265\_Cel6A (MEZCLA TR\_CT) o *At*\_ALKO4245\_Cel6A (MEZCLA TR\_AT), respectivamente.

Se preparó la MEZCLA ACC a partir del producto comercial Accellerase® 1000 (de Genencor International/Danisco A/S) (por cada 10 ml): 400 mg de proteína Accellerase® (100%). El volumen se enrasó a 10 ml con agua corriente. La concentración final de proteína en dicha mezcla era de 40 mg/ml.

Para someter a ensayo el rendimiento de Cel6A/CBHII en la hidrólisis con MEZCLA ACC, 15% (60 mg) del componente Accellerase® 1000 de la MEZCLA ACC se sustituyó por *Ma*\_ALKO4237\_Cel6A (MEZCLA ACC\_MA), *Ct*\_ALKO4265\_Cel6A (MEZCLA AC\_CT) o *At*\_ALKO4245\_Cel6A (MEZCLA ACC\_AT), respectivamente.

Para las combinaciones de MEZCLA 2, la hidrólisis se llevó a cabo a 55°C, mientras que la temperatura de hidrólisis para los experimentos con MEZCLA ENZIMAS T. REESEI y con MEZCLA ACC era de 37°C. Se extrajeron muestras de la hidrólisis tras 72 h, se cuantificaron mediante HPLC y se determinaron las concentraciones de glucosa y xilosa. Los resultados de los blancos de sustrato se restaron de las muestras con enzimas y la concentración de glucosa y xilosa agrupados se muestran en la figura 4A-C.

Los resultados muestran claramente un mejor rendimiento de la MEZCLA 2 con los enzimas Cel6A/CBHII termoestables a 55°C. Se encontró que la cantidad de azúcares liberados del sustrato de madera dura se incrementaba en 12%, 14% y 26% mediante la complementación con los enzimas *Ma*\_ALKO4237\_Cel6A, *Ct*\_ALKO4265\_Cel6A o *At*\_ALKO4245\_Cel6A en la MEZCLA 2, respectivamente. Se encontró que el enzima *Acremonium thermophilum* ALKO4245 era la proteína Cel6A/CBHII de mejor rendimiento estudiada en la presente memoria (figura 4A). *At*\_ALKO4245\_Cel6A mostraba una hidrólisis incrementada también a 37°C añadida en la mezcla de *Trichoderma* del estado de la técnica (MEZCLA ENZIMAS T. REESEI) (figura 4B) o en el producto comercial (MEZCLA ACC) (figura 4C).

#### **Ejemplo 5. Hidrólisis de mazorcas de maíz con preparaciones enzimáticas que comprenden una celobiohidrolasa CBHI/Cel6A recombinante**

Se suspendieron mazorcas de maíz sometidas a explosión con vapor en tampón de acetato sódico 0,05 M, pH 4,8. El peso final de la mezcla de hidrólisis era de 20 g, con una concentración de sólidos totales de 2% (p/p). El sustrato se hidrolizó utilizando diferentes mezclas enzimáticas a una dosis de 5 mg de proteína por cada gramo de sólidos totales en matraces de agitación de 50 ml. Se determinó el contenido de proteínas de los componentes enzimáticos y de las mezclas utilizando el kit de ensayo Pierce BCA (Thermo Scientific, número de producto 23227) con albúmina de suero bovino (Thermo Scientific, número de producto 23209) como estándar. Los matraces de agitación se agitaban en un baño de agua de agitación lineal ajustado a 55°C. Para cada punto de muestra, se extrajo una muestra de 1 ml de matraces de agitación por duplicado y se sometieron a ebullición durante 10 minutos para terminar la hidrólisis enzimática, se centrifugaron y se analizó el sobrenadante para los productos de reacción de la hidrólisis. Los blancos que contenían solo sustrato (adición de sólo tampón en lugar de enzimas) se prepararon de manera idéntica a la de las demás muestras.

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

Preparación de Cel7A/CBHI termofílico que contiene Cel7A de *Thermoascus aurantiacus* ALKO4242 con DUC unido genéticamente de *T. reesei* CBHI/Cel7A (documento WO 2007/071818).

Preparación de endoglucanasa termofílica que contiene endoglucanasa Cel45A de *Acremonium thermophilum* ALKO4245 (At EG\_40/Cel45A, documento nº WO 2007/071818).

5 Preparación de  $\beta$ -glucosidasa termofílica que contiene  $\beta$ -glucosidasa de *Thermoascus aurantiacus* ALKO4242 (Ta  $\beta$ G\_81/Cel3A, documento WO 2007/071818).

Preparación de xilanasa termofílica que contiene *Nomomurea flexuosa* Xyn11A (AM24, documento WO 2005/100557).

10 Todas las celulasas se produjeron heterológicamente como monoccomponentes en la cepa hospedadora *Trichoderma reesei* que presenta un fondo libre de celulasa (se deleccionaron los genes codificantes de las cuatro celulasas principales Cel7A/CBHI, Cel6A/CBHII, Cel7B/EGI y Cel5A/EGII). Se utilizaron en la mezcla sobrenadantes de cultivo en bruto. Los componentes enzimáticos se combinaron de la manera siguiente (por cada 10 ml de mezcla):  
15 preparación de CBHI/Cel7A 330 mg (71,2%), preparación de endoglucanasa 105 mg (22,7%), preparación de  $\beta$ -glucosidasa 7,5 mg (1,6%) y preparación de xilanasa 21 mg (4,5%). Se enrasó el volumen a 10 ml con agua corriente. La concentración de proteína final de la mezcla era de 46,35 mg/l. Dicha mezcla de enzima se denominó MEZCLA 2.

20 Para someter a ensayo el rendimiento de At\_ALKO4245\_Cel6A en la hidrólisis con MEZCLA 2, el 15% (49,5 mg) del componente CBHII/Cel7A de la MEZCLA 2 se substituyó por At\_ALKO4245\_Cel6A (MEZCLA 2\_AT), respectivamente.

25 Se extrajeron muestras de la hidrólisis tras 72 h, se cuantificaron mediante HPLC y se determinaron las concentraciones de glucosa y xilosa. Los resultados de los blancos de sustrato se restaron de las muestras con enzimas y la concentración de glucosa y xilosa agrupadas se muestran en la figura 5.

30 De manera similar a la indicada en el Ejemplo 4, los resultados en la presente memoria muestran claramente el mejor rendimiento de la MEZCLA 2 con el enzima At\_ALKO4245\_Cel6A a 55°C para el sustrato de mazorcas de maíz.

## Referencias

35 Altschul SF, W Gish, W Miller, EW Myers y DJ Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

AMFEP, 2007. Association of Manufacturers and Formulators of Enzyme products, List of commercial enzymes at <http://www.amfep.org/list.html> (actualizado el 30 de noviembre de 2007).

40 Badger, PC. 2002. Ethanol from cellulose: a general review. In Trends in new crops and new uses. J Janick and A Whipkey (eds.). ASHS Press, Alexandria, VA, USA, pp. 17-21.

45 Bendtsen JD, H Nielsen, G von Heijne and S Brunak. 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol.Biol. 340:783-795.

Biely P, M Vrsanska, M Tenkanen and D Kluepfel. 1997. Endo-beta-1,4-xylanase families: differences in catalytic properties. Journal of Biotechnology 57: 151-166.

50 Bolton ET y BJ McCarthy. 1962. A general method for the isolation of RNA complementary to DNA. Proc. Nat. Acad. Sci.USA 48:1390-1397.

Collins CM, PG Murray, S Denman, A Grassick, TT Teeri, L Byrnes and MG Tuohy. 2003. Molecular cloning of the cellobiohydrolase genes of *Talaromyces emersonii*. Biochem Biophys Res. Comm. 301:280-286.

55 Cullen D y PJ Kersten. 2004. Enzymology and molecular biology of lignin degradation. In: The Mycota III. Biochemistry and molecular biology. R Brambl and GA Marzluf (eds.). 2nd edition, Springer-Verlag, Berlin-Heidelberg, pages 249-273.

60 Del Cañizo AN, RA Hours, MV Miranda, O Cascone. 1994. Fractionation of fungal pectic enzymes by immobilised metal ion affinity chromatography. J. Sci. Food Agric. 64:527-531.

Edman P and G Begg. 1967. A protein sequenator. Eur. J. Biochem. 1:80-91.

65 Galagan JE, SE Calvo, KA Borkovich, EU Selker, ND Read, D Jaffe, W FitzHugh, LJ Ma, S Smirnov, S Purcell, B Rehman, T Elkins, R Engels, S Wang, CB Nielsen, J Butler, M Endrizzi, D Qui, P Ianakiev, D Bell-Pedersen, MA Nelson, M Werner-Washburne, CP Selitrennikoff, JA Kinsey, EL Braun, A Zelter, U Schulte, GO Kothe, G Jedd, W



- 5 Mewes, C Staben, E Marcotte, D Greenberg, A Roy, K Foley, J Naylor, N Stange-Thomann, R Barrett, S Gnerre, M Kamal, M Kamvysselis, E Mauceli, C Bielke, S Rudd, D Frishman, S Krystofova, C Rasmussen, RL Metznerberg, DD Perkins, S Kroken, C Cogoni, G Macino, D Catcheside, W Li, RJ Pratt, SA Osmani, CP DeSouza, L Glass, MJ Orbach, JA Berglund, R Voelker, O Yarden, M Plamann, S Seiler, J Dunlap, A Radford, R Aramayo, DO Natvig, LA Alex, G Mannhaupt, DJ Ebbolle, M Freitag, I Paulsen, MS Sachs, ES Lander, C Nusbaum, B Birren. 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422:859-868.
- 10 Gasteiger E, A Gattiker, C Hoogland, I Ivanyi, RD Appel y A Bairoch. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31: 3784-3788.
- 15 Ghose TK. 1987. International Union of Pure and Applied Chemistry. Measurement of cellulase activities. *Pure and Appl. Chem.* 59: 257-268.
- Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280: 309-316.
- Henrissat B y A Bairoch. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 293: 781-788.
- 20 Henrissat B y A Bairoch. 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316: 695-696
- Henrissat B, TT Teeri y RAJ Warren. 1998. A scheme for designating enzymes that hydrolyse the polysaccharides in the cell wall of plants. *FEBS Letters* 425: 352-354.
- 25 Hong J, H Tamaki, K Yamamoto y H Kumagai. 2003a. Cloning of a gene encoding a thermo-stabile endo- $\beta$ -1,4-glucanase from *Thermococcus aurantiacus* and its expression in yeast. *Biotechnol. Letters* 25: 657-661.
- 30 Hong J, H Tamaki, K Yamamoto y H Kumagai. 2003b. Cloning of a gene encoding thermostable cellobiohydrolase from *Thermoascus aurantiacus* and its expression in yeast. *Appl. Microbiol. Biotechnol.* 63: 42-50.
- Joutsjoki VV, TK Torkkeli y KMH Nevalainen. 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.
- 35 Karhunen T, A Mäntylä, KMH Nevalainen y PL Suominen. 1993. High frequency one-step gene replacement in *Trichoderma reesei*. I. Endoglucanase I overproduction. *Mol. Gen. Genet.* 241:515-522.
- 40 Karlsson J, M Saloheimo, M Siika-aho, M Tenkanen, M Penttilä y F Tjerneld. 2001. Homologous expression and characterization of Cel61A (EGIV) of *Trichoderma reesei*. *Eur. J. Biochem.* 268:6498-6507.
- Kurabi A, A Berlin, N Gilkes, D Kilburn, A Markov, A Skomarovsky, A Gusakov, O Okunev, A Sinitsyn, D Gregg, D Xie y J Saddler. 2005. Enzymatic hydrolysis of steam-exploded and ethanol organosolv-pretreated Douglas-Fir by novel and commercial fungal cellulases. *Appl. Biochem and Biotechnol.* Vol 121-124: 219-229.
- 45 Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Lever M. 1972. A new reaction for colorimetric determination of carbohydrates. *Anal. Biochem.*, 47: 276-279.
- 50 Maheshwari R, G Bharadwaj and MK Bhat. 2000. Thermophilic fungi: their physiology and enzymes. *Microbiol. Mol. Biol. Rev.* 64:461-488.
- Malardier L, MJ Daboussi, J Julien, F Roussel, C Scazzocchio y Y Brygoo. 1989. Cloning of the nitrate reductase gene (*niaD*) of *Aspergillus nidulans* and its use for transformation of *Fusarium oxysporum*. *Gene* 78:147-156.
- 55 Miettinen-Oinonen A, J Londesborough, V Joutsjoki, R Lantto, J Vehmaanperä. 2004. Three cellulases from *Melanocarpus albomyces* with application in textile industry. *Enzyme Microb. Technol.* 34:332-341.
- 60 Miller GL. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugars. *Anal. Chem.* 31:426-428.
- Murray PG, CM Collins, A Grassick y MG Tuohy. 2003. Molecular cloning, transcriptional, and expression analysis of the first cellulase gene (*cbh2*), encoding cellobiohydrolase II, from the moderately thermophilic fungus *Talaromyces emersonii* and structure prediction of the gene product. *Biochem. Biophys. Res. Commun.* 301:280-286.
- 65 Nielsen H, J Engelbrecht, S Brunak y G von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides

and prediction of their cleavage sites. *Protein Engineering* 10:1-6.

Nielsen H y A Krogh. 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. In: *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6)*, AAAI Press, Menlo Park, California, pp. 122-130.

Paloheimo M, A Mäntylä, J Kallio, y P Suominen. 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.

Penttilä M, H Nevalainen, M Rättö, E Salminen y J Knowles. 1987. A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. *Gene* 61:155-164.

Raeder U y P Broda. 1985. Rapid preparation of DNA from filamentous fungi. *Lett. Appl. Microbiol.* 1:17-20.

Robyt JF y WJ Whelan. 1972. Reducing value methods for maltodextrins: I. Chain length dependence of alkaline 3,5-dinitrosalisylate and chain-length independence of alkaline copper. *Anal. Biochem.* 45:510-516.

Sambrook J y DW Russell. 2001. *Molecular cloning, a laboratory manual*. Cold Spring Harbor Laboratory, New York, US.

Shaw AJ, KK Podkaminer, SG Desai, JS Bardsley, SR Rogers, PG Thorne, DA Hogsett y LR Lynd. 2008. Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. *Proc. Natl. Acad. Sci. USA* 105: 13769-13774.

Sinnott ML. 1990. Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* 90: 1171-1202.

Somogyi M. 1952. Notes on sugar determination. *J. Biol. Chem.* 195:19-23.

Srisodsuk M, T Reinikainen, M Penttilä y T Teeri. 1993. Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. *J. Biol. Chem.* 268(28): 20756-61.

Stålbrand H, A Saloheimo, J Vehmaanperä, B Henrissat y M Penttilä. 1995. Cloning and expression in *Saccharomyces cerevisiae* of a *Trichoderma reesei*  $\beta$ -mannanase gene containing a cellulose binding domain. *Appl. Environ. Microbiol.* 61:1090-1097.

Sundberg M y K Poutanen. 1991. Purification and properties of two acetylxylyl esterases of *Trichoderma reesei*. *Biotechnol. Appl. Biochem.* 13: 1-11.

Suurnäkki A, M Tenkanen, M Siika-aho, M-L Niku-Paavola, L Viikari y J Buchert. 2000. *Trichoderma reesei* cellulases and their core domains in the hydrolysis and modification of chemical pulp. *Cellulose* 7: 189-209.

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

Tomme P, S McRae, T Wood y M Claeysens. 1988. Chromatographic separation of cellulolytic enzymes. *Methods in Enzymol.* 160: 187-192.

Tuohy M, J Walsh, P Murray, M Claeysens, M Cuffe, A Savage y M Coughan. 2002. Kinetic parameters and mode of action of cellobiohydrolases produced by *Talaromyces emersonii*. *Biochem. Biophys. Acta* 1596: 366-380.

Waffenschmidt S and L Jaenicke. 1987. Assay of reducing sugars in the nanomole range with 2,2'-bicinchoninate. *Anal. Biochem.* 165:337-340.

Withers SG, D Dombroski, LA Berven, DG Kilburn, RC Miller Jr, RAJ Warren y NR Gilkes. 1986. Direct  $^1\text{H}$  NMR determination of the stereochemical course of hydrolases catalysed by glucanase components of the cellulose complex. *Biochem. Biophys. Res. Commun.* 139:487-494.

Withers SG y Aebersold R. 1995. Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci.* 4:361-372.

## Listado de secuencias

<110> Roal Oy

<120> Procedimiento de tratamiento de material celulósico y enzimas CBHII/Cel6A útiles en el mismo

<130> A9693PC  
 <150> FI 20096412  
 5 <151> 2009-12-30  
 <160> 16  
 <170> PatentIn versión 3.5  
 10 <210> 1  
 <211> 17  
 <212> ADN  
 <213> Secuencia artificial  
 15 <220>  
 <223> Secuencia del cebador oligonucleótido CBH\_1S  
 <220>  
 20 <221> misc\_feature  
 <222> (6)..(6)  
 <223> N es A, G, T o C  
 <220>  
 25 <221> misc\_feature  
 <222> (9)..(9)  
 <223> R es A o G  
 <220>  
 30 <221> misc\_feature  
 <222> (12)..(12)  
 <223> Y es T o C  
 <220>  
 35 <221> misc\_feature  
 <222> (15)..(15)  
 <223> N es A, G, T o C  
 <400> 1  
 40 tggggncart gyggngg 17  
 <210> 2  
 <211> 17  
 <212> ADN  
 45 <213> Secuencia artificial  
 <220>  
 <223> Secuencia del cebador oligonucleótido CBH\_1AS  
 <220>  
 50 <221> misc\_feature  
 <222> (3)..(3)  
 <223> N es A, G, T o C  
 <220>  
 55 <221> misc\_feature  
 <222> (9)..(9)  
 <223> N es A, G, T o C  
 <220>  
 60 <221> misc\_feature  
 <222> (12)..(12)  
 <223> N es A, G, T o C  
 <220>  
 65 <221> misc\_feature

# ES 2 612 512 T3

<222> (14)..(14)  
 <223> R es A o G

5 <400> 2  
 gcnggccanc cnarcca 17

10 <210> 3  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia del cebador oligonucleótido CBH\_8

15 <400> 3  
 atggctaagc agctgctgct c 21

20 <210> 4  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia del cebador oligonucleótido CBH\_9

25 <220>  
 <221> misc\_feature  
 <222> (6)..(6)  
 <223> R es A o G

30 <400> 4  
 tcagarcgga gggttgcat 20

35 <210> 5  
 <211> 44  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia del cebador oligonucleótido Te\_CBH\_A

<400> 5  
 tattatccgc ggactcgcga tcatgcggaa tcttctgct cttg 44

45 <210> 6  
 <211> 38  
 <212> ADN  
 <213> Secuencia artificial

50 <220>  
 <223> Secuencia del cebador oligonucleótido Te\_CBH\_B

55 <400> 6  
 aattggatc ctcaaacag cgggtagca ttcgtgag 38

<210> 7  
 <211> 1014  
 <212> ADN  
 <213> Secuencia artificial

60 <220>  
 <223> Secuencia del fragmento de PCR obtenido de *Acremonium thermophilum* ALKO4245 (CBS 116240) utilizando los cebadores CBH\_1S y CBH\_1AS

ES 2 612 512 T3

<400> 7  
 tggggacaat gcggcgga cggcttctcg ggaccgacct gttgcacgtc gggcaacacc 60  
 tgcgtcaagc tcaatgactg gtactcccag tgcctgcca acagccaggt acgctgcgcc 120  
 cagccgctct agctggactt ctcttttac gagatgtctt ggctgacctc cttccaaagg 180  
 tgaccaccac cacaagctcg accaccacca ccaccaccac gaccctcac ggtcccacca 240  
 ccgccacaac caccaccacc aagccacctc ccaccaccac caccacgacc acgacgacga 300  
 agcctcctgg caccgcctcg ggaccgtgt cctacaccgg caacccttc tctggcgtgc 360  
 agctttgggc caactcccac tacgcctcgg agatctcggc ctccgccatc cccagcctga 420  
 cgggcgccat ggccaccaag gccgccggc tgcccaagggt gccagcttc cagtggctgt 480  
 acgtctcctt tctcctctct ctccctcttc tctcctctc tttttttcc tttttgtgtg 540  
 tgtgtgtgtg tgttcacaca catcttaagc caagccactg acagttttct tccagtgaca 600  
 ccgcctcaa ggtctccctg atggccgaca ccctcagcga catccgccag gccaacgcgc 660  
 ccggcgccaa cccgccctac gccggccagt tcgttgteta cgacctgcc gaccgcgact 720  
 gctccgcccgc cgctccaac ggcgagtaca gcatcgccga caacggcgtc gccactaca 780  
 aggcctacat cgacagcatc cgcgagcagc tggtcgcta ctccgacgtg cgcgtcctgc 840  
 tcgtcgtcga gcccgactcg ctggccaacc tggtcaccaa cctcaacgtg gccaaagtct 900  
 ccaacgccca gagcgctac ctcgagtga ccaatacgcc ctacccagc tcaacctgcc 960  
 caacgtcgcc atgtacctcg acgccggcca cgccggctgg ctccgctggc ctgc 1014

- 5 <210> 8
- <211> 813
- <212> ADN
- <213> Secuencia artificial

- 10 <220>
- <223> Secuencia del fragmento de PCR obtenido de *Melanocarpus albomyces* ALKO4237 (CBS 685.95) utilizando los cebadores CBH\_1S y CBH\_1AS

<400> 8  
 tggggccaat gcggcgga tggctggacc gggccgacct gctgcgagtc gggcagcagc 60

ES 2 612 512 T3

tgcgtggccc agaacgacta ctactctgca gtgcctgccg ggcggcgcga ccaccacgtc 120  
 gggcccgggc acgccgtcca ccaccacgtc ggtgacgacc cccaccagcc agggctcgcc 180  
 cacgtcgccg ccgccgacga cgccgaccac gacgatcccc ggcggcgcgt cgacgacggc 240  
 cagctacacg ggcaaccctg tgcggggcat gcagatgtgg gccaacagct actacgcctc 300  
 cgaggtctcg tcgctggcca tccccagcat gacgggcccc atggccacca aggcggccga 360  
 ggtggccaag gtgcccagct tccagtggct cgaccgcaac gtgacggctg acacgctctt 420  
 cacgcagacg ctggccgaga cccgggcggc caacgagggc ggcgcccaacc cgcccaacat 480  
 gggcatcttc gtcgtctatg acctgcccga ccgcgactgc gccgccgccg cgtccaacgg 540  
 cgagtgggcc atcgccgacg gcggcgtggc caactacaag gcctacatcg accgcatccg 600  
 caagcacatc atcgctact cggacatccg catggccatc gtgctcgagc ccgactcgct 660  
 cgccaacatg gtgaccaaca tggacgtgcc caagtgcgcc aacgcggccg acacgtataa 720  
 ggagctcacc atctacgccg tccagcagct cgacctgccc aacgtggcca tctacctgga 780  
 cgccggccac gccggctggc tcggctggcc agc 813

<210> 9

<211> 1757

5 <212> ADN

<213> *Chaetomium thermophilum*

<220>

<221> misc\_feature

10 <223> La secuencia de nucleótidos del gen cbh2/cel6A de *Chaetomium thermophilum* ALKO4265 (CBS 730.95)

<400> 9

atggctaagc agctgctgct cactgccgct cttgcccga cttcgctggc tgcccctcta 60  
 cttgaggagc gccagagctg ctccctggtc tggatgtct caaccccgga tctgagtgtt 120  
 cgggatctgc ttgatcatgg atgccgatgc tgacatgttt gtgtttcagg ggtcaatgag 180  
 gtggcatcaa ctacaacggc ccgacttgct gcccgctccg cagtgtttgc acatacctga 240  
 atgactggta cagccagtgc attcccggcc aggctcagcc cggcacgact agcaccacgg 300  
 cccggaccac cagcaccagc accagcactt cgtccgtccg cccgaccacc acctccaata 360  
 cccctgtgac gactgctcct ccggcgacca ccatcccggg cggcgccctg agcacggcca 420  
 gctacaacgg caacccttc tcgggtgtcc agctctgggc caacacctac tactcgtccg 480  
 aggtgcacac tctggccatc cccagcctgt ctccctgagct ggctgccaaag gccgccaaag 540  
 ttgctgaggt tcccagcttc cagtggctcg accgcaacgt gactgtcgac acctcttcg 600  
 tcggcaccct caatgacatt cgtggtgcca accagcgcgg tgccaaccgg ccttatggta 660  
 agtgcaatgc ctgccccaca tcagacccta cgccccaaagc ttggccactc tccagagtgg 720  
 agcagtgggt ccgacacaac cctaaccctg acaattccct tacctcaacc ccttttcaac 780

ES 2 612 512 T3

ctcgtccatc gcatcatcac aacctgtttt cttaacgtct atcatgatac tggtgctaaa 840  
ctccgtaaag cggcccaatt tgtcgtttat gaccttccgg accgtgattg cgctgccgct 900  
gcttcgaacg gcgagtgggc tatcgccaac aatggtgcca acaactacaa gcgttacatc 960  
gaccggatcc gcgagatcct tatccagtac tctgatatcc gcaactattct ggtcattgag 1020  
cctgattccc tggccaacat ggtcaccaac atgaacgtcc agaagtgcgc gaacgccgca 1080  
accacctaca aggagcttac catctatgcc ctcaagcagc tcaaccttcc tcatgtcgcc 1140  
atgtacatgg atgctggtca cgccggctgg cttggctggc ccgccaacat tcagcctgct 1200  
gctgagctct ttggtcagct ctaccgtgac gctggcaagc ccgcttccgt ccgcggctctc 1260  
gcgtaagaa agctcctgag acctcgactc ctggaacaac agttactgac ataacctaga 1320  
ccaacgttgc caactacaat gcttggtcga tcgccagcgc tccgtcgtat acttctccta 1380  
accctaacta cgacgagaag cactacattg aggcccttgc tcctcttctc cgcaaccagg 1440  
gcttcgatgc caagttcatc gtcgacaccg gccgtaacgg caagcagccc accggccagc 1500  
tccaatgggg cgattggtgc aatgtcaagg gaactggctt cgggtgtcgt cccacttcta 1560  
aactgggca tgagcttggt gatgctttcg tgtgggtcaa gcccggtggt gagtccgacg 1620  
gcaccagcga caccagcgt gctcgttacg actatcactg cggccttcc gacgcattga 1680  
ctccagcgc tgaggctggc caatggttcc aggcttattt cgaacagctg ctcattaatg 1740  
ccaaccctcc gttctga 1757

<210> 10

<211> 1754

5 <212> ADN

<213> *Talaromyces emersonii*

<220>

<221> misc\_feature

10 <223> La secuencia de nucleótidos del gen *cbh2/cel6A* de *Talaromyces emersonii* DSM 2432 (RF8069)

<400> 10

atgcggaatc ttcttgctct tgcaccggca gcgctgcttc tcggcgcagc ggacgcgcaa 60  
caatccctct ggggacaatg tgagcagctc ctogaacgtc tgtctgatga attggtctga 120  
cagcttcagg cggaggggaat tcgtggactg gagcgcagga ttgtgctgca ggagcgacgt 180  
gcagcaccat caattcttgt acgtctgttc tctactgctg tcgagctcag ctaactgcgt 240  
agactacgca caatgcgtcc ctgcaacggc cactcctacc acgttgacga caacgacaaa 300  
gccctcgtcg actgcgcaa cgaccctcc tccgacgtca gcgacgacca caggcactgg 360  
atcggcgaca tcgccctcca tcaccgcgtc tgcgtccggc aaccgcttg tcggatacca 420  
gctctacgcc aaccgctact atgcctctga ggtgattagc ctggccatcc cgtcgtctaa 480  
cagcgagctg gttcccaagg cgagcgaggt ggccaagggt ccgtcttttg tctggctgtg 540

ES 2 612 512 T3

ggtgatattc caccctgttt gggacttgca gagactgata ttgtcacagc gatcaagcgg 600  
 ccaaagtgcc caacatgggc gagtatctga aagacatcca gtcccagaat gcggccggcg 660  
 cagaccctcc gattgcaggc atcttcgtcg tttacgacct acctgaccgc gactgcgcgg 720  
 cggcagcggc caatggcgag ttctccatcg ccaacaacgg cgttgccctg tacaagcaat 780  
 acatcgactc gatccgcgag cagctgacga cgtattcggg tgtgcacacc atcctgatca 840  
 ttggttagta cgctagtgat atttatcaat tttttttttg tcaataactga ctgccgcaga 900  
 acccgacagc ctggccaacc tggtcaccaa cctgaacgtg gcgaaatgcg cgaatgccc 960  
 gggcgcctat ctgcaatgca tcaactacgc catcacgcag ctcaacctgc cgaatgtggc 1020  
 catgtatctt gatgctggtg agcttccctc acataccagg gaataaaaga cagaactgat 1080  
 tgtctttcag gacacgccgg atggctaggc tggtcagcaa acctccaacc cgctgcgcag 1140  
 ctgtttgcag aggtctacaa gaacgcctcg tcgccggcct cggcgcgcgg tctcgcgacc 1200  
 aacgtcgcca actacaacgc ctggacgatc agcccgtgcc cgtcgtacac gcagggcgac 1260  
 cccaactgcg acgaggagga ctatgtgaat gcccttgccg cgctgcttca gagccagggg 1320  
 ttaaatgctt actttatcac tgatacatgt gagtctcacc acaccagac ctcgactgga 1380  
 cgtaccctaa tctgacctgt tctgcagccc gcaacggcgt ccaaccacc aagcagaacc 1440  
 aatggggcga ctggtgcaac gtcacgcgca ccgggttcgg cgtccgcccg acgactgaca 1500  
 ctggcaaccc tctcggaggc gccttcgtct gggcgaagcc gggcggcgag agcgatggca 1560  
 catctaacac gacctctccg cgatacgact accactgcgg gctgagcgat gcgctgcagc 1620  
 cggctccgga ggcgggaact tggttccagg taagttgcaa gcagagatgt actgtacatt 1680  
 ggagcgtatg ctaattatgt gtgttacagg cgtactttga gcagctgctt acgaatgcta 1740  
 acccgctgtt ctga 1754

<210> 11  
 <211> 1830  
 <212> ADN  
 <213> *Acremonium thermophilum*

<220>  
 <221> misc\_feature

<223> La secuencia de nucleótidos del gen *cbh2/cel6A* de *Acremonium thermophilum* ALKO4245 (CBS 116240)

<400> 11  
 atggctgcca gacgtctcct gctcgtgcc tccctgtcgg cagtggctct cgccgcgccc 60  
 gtggtcgaag agcgcagaa ctgcgcctct acctgtagg caccactccc cagctgtttt 120  
 ttctcatctc tgagccaaag caacggtctg actctcttcg cttcttcaa acaggagcca 180  
 atgcggcggc aacggcttct cgggaccgac ctggtgcacg tcgggcaaca cctgcgtcaa 240  
 gctcaatgac tggactccc agtgcctgcc caacagccag gtacgctgcg ccagccgct 300



ES 2 612 512 T3

ctagctggac ttctcctttt acgagatgtc ttggctgacc tccttccaaa ggtgaccacc 360  
accacaagct cgaccaccac caccaccacc acgaccacct acggtcccac caccaccaca 420  
accaccacca ccaagccacc tcccaccacc accaccacga ccacgacgac gaagcctcct 480  
ggcaccgcct cgggcaccgt gtcctacacc ggcaaccctt tctctggcgt gcagctttgg 540  
gccaaactcct actacgcctc cgagatctcg gcctccgcca tcccagcct gacgggcgcc 600  
atggccacca aggccgccgc cgtcgccaag gtgcccagct tccagtggct gtacgtctcc 660  
tttctcctct ctctccctct tctctccttc tctttttttt cttttttgtg tgtgtgtgtg 720  
tgtgttcaca cacatcttaa gccaaagccac tgacagtttt cttccagtga caccgcctcc 780  
aaggtctccc tgatggccga caccctcagc gacatccgcc aggccaaccg cgccggcgcc 840  
aaccgcctcct acgcccggcca gttcgttgtc tacgacctgc ccgaccgcga ctgctccgcc 900  
gccgcctcca acggcgagta cagcatcgcc gacaacggcg tcgcccacta caaggcctac 960  
atcgacagca tccgcgagca gctggtcgcc tactccgacg tgcgcgtcct gctcgtcgtc 1020  
gagcccgact cgctggccaa cctggtcacc aacctcaacg tggccaagtg ctccaacgcc 1080  
cagagcgcct acctcgagtg caccaactac gccctcacc agctcaacct gcccaacgtc 1140  
gccatgtacc tcgacgccgg ccacgccggc tggctgggct ggcccgccaa cctgcagccc 1200  
gccgccacc tgttcgccaa ggtctacaac gacgccaaca agcccgtgc cgtgcgcggc 1260  
ctcgccacca acgtcgccaa ctacaacggc tggaacctga cctcgccgcc ctcgtaacc 1320  
caaggtatgc ccttcacgta tcccctcccc ttcataccga gcctcgcag ggactttccg 1380  
gcctctcttt tgtctccgcc ccactctcc gtatctccag ttgggaaagc aatactgata 1440  
cctcgggcag gcaacaacaa ctacgacgag atccactacg tccaggccat cgccccctc 1500  
ctcaagtctg ccggcttcga cgcccacttc atcaccgaca ccggccgcaa cggcaagcag 1560  
cccaccggcc agcagcaatg gggcgactgg tgcaacgtca tcggcaccgg cttcggcgtg 1620  
cgccccacca ccaaacacggg ccttgagctc gaggacgcct tcgtctgggt gaagcccggc 1680  
ggcgagtgcg acggcaccag cgacaccagc gccgcccgt acgactacca ctgcggtctg 1740  
tccgatgccc tgcagccccg gcccgaggcc ggcacctggt tcgaggccta tttcgagcag 1800  
ctgctcacca acgccaacc gtcgttctga 1830

<210> 12  
<211> 478  
5 <212> PRT  
<213> *Acremonium thermophilum*

<220>  
<221> misc\_feature

10 <223> La secuencia de aminoácidos deducida de CBHII/Cel6A de *Acremonium thermophilum* ALKO4245 (CBS 116240)

ES 2 612 512 T3

<400> 12

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

Leu Ala Ala Pro Val Val Glu Glu Arg Gln Asn Cys Ala Ser Thr Trp  
 20 25 30

Ser Gln Cys Gly Gly Asn Gly Phe Ser Gly Pro Thr Cys Cys Thr Ser  
 35 40 45

Gly Asn Thr Cys Val Lys Leu Asn Asp Trp Tyr Ser Gln Cys Leu Pro  
 50 55 60

Asn Ser Gln Val Thr Thr Thr Thr Ser Ser Thr Thr Thr Thr Thr Thr  
 65 70 75 80

Thr Thr Pro His Gly Pro Thr Thr Thr Thr Thr Thr Thr Thr Lys Pro  
 85 90 95

Pro Pro Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Lys Pro Pro Gly Thr  
 100 105 110

Ala Ser Gly Thr Val Ser Tyr Thr Gly Asn Pro Phe Ser Gly Val Gln  
 115 120 125

Leu Trp Ala Asn Ser Tyr Tyr Ala Ser Glu Ile Ser Ala Ser Ala Ile  
 130 135 140

Pro Ser Leu Thr Gly Ala Met Ala Thr Lys Ala Ala Ala Val Ala Lys  
 145 150 155 160

Val Pro Ser Phe Gln Trp Leu Asp Thr Ala Ser Lys Val Ser Leu Met  
 165 170 175

Ala Asp Thr Leu Ser Asp Ile Arg Gln Ala Asn Arg Ala Gly Ala Asn  
 180 185 190

Pro Pro Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp  
 195 200 205

Cys Ser Ala Ala Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Asn Gly  
 210 215 220

Val Ala His Tyr Lys Ala Tyr Ile Asp Ser Ile Arg Glu Gln Leu Val  
 225 230 235 240

Ala Tyr Ser Asp Val Arg Val Leu Leu Val Val Glu Pro Asp Ser Leu  
 245 250 255

ES 2 612 512 T3

Ala Asn Leu Val Thr Asn Leu Asn Val Ala Lys Cys Ser Asn Ala Gln  
260 265 270

Ser Ala Tyr Leu Glu Cys Thr Asn Tyr Ala Leu Thr Gln Leu Asn Leu  
275 280 285

Pro Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly  
290 295 300

Trp Pro Ala Asn Leu Gln Pro Ala Ala Thr Leu Phe Ala Lys Val Tyr  
305 310 315 320

Asn Asp Ala Asn Lys Pro Ala Ala Val Arg Gly Leu Ala Thr Asn Val  
325 330 335

Ala Asn Tyr Asn Gly Trp Asn Leu Thr Ser Pro Pro Ser Tyr Thr Gln  
340 345 350

Gly Asn Asn Asn Tyr Asp Glu Ile His Tyr Val Gln Ala Ile Ala Pro  
355 360 365

Leu Leu Lys Ser Ala Gly Phe Asp Ala His Phe Ile Thr Asp Thr Gly  
370 375 380

Arg Asn Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys  
385 390 395 400

Asn Val Ile Gly Thr Gly Phe Gly Val Arg Pro Thr Thr Asn Thr Gly  
405 410 415

Leu Glu Leu Glu Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu Cys  
420 425 430

Asp Gly Thr Ser Asp Thr Ser Ala Ala Arg Tyr Asp Tyr His Cys Gly  
435 440 445

Leu Ser Asp Ala Leu Gln Pro Ala Pro Glu Ala Gly Thr Trp Phe Glu  
450 455 460

Ala Tyr Phe Glu Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe  
465 470 475

<210> 13

<211> 1607

5 <212> ADN

<213> *Melanocarpus albomyces*

<220>

<221> misc\_feature

10 <223> La secuencia de nucleótidos del gen *cbh2/cel6A* de *Melanocarpus albomyces* ALKO4237 (CBS 685.95)

ES 2 612 512 T3

<400> 13  
atggtcaaga aactcctgct caccaccgcc ctggcggccg cctcgctggc ggctcccgtc 60  
atcgaggagc gccagagctg cgcctctctc tggtgagcga gcccgctccga gcccggtcccc 120  
gtccatccga gtgtcgagtg tgactgacac gaacgggatt cgtcccgtcc cgtcccgttc 180  
cacaggggcc aatgcggcgg caatggctgg accggggccga cctgctgcga gtcgggcagc 240  
acgtgctggtg cccagaacga ctactactcg cagtgcctgc cgggcggcgc gaccaccacg 300  
tcgggcccgg gcacgccgtc caccaccacg tcggtgacga cctccaccag ccagggtcgtg 360  
cccacgtcgc cgccgccgac gacgccgacc acgacgatcc ccggcggcgc gtcgacgacg 420  
gccagctaca cgggcaaccc gttcgcgggc atgcagatgt gggccaacag ctactacgcc 480  
tccgaggtct cgtcgctggc catccccagc atgacgggcc ccatggccac caaggcggcc 540  
gagtgggcca aggtgccag cttccagtgg ctcgaccgca acgtgacggt cgacacgctc 600  
ttcacgcaga cgctggccga gatccgggcg gccaacgagg cgggcgcca cccgccaac 660  
atgggcatct tcgtcgtcta tgacctgcc gaccgcgact gcgccgccgc cgcgtccaac 720  
ggcgagtggg ccatcgccga cggcggcgtg gccaaactaca aggctacat cgaccgcatc 780  
cgcgagcaca tcatcgcta ctcgacatc cgcattggcca tcgtgctcga gcccgactcg 840  
ctcgccaaca tggtgaccaa catggacgtg cccaagtgcg ccaacgcggc cgacacgtac 900  
aaggagctca ccatctacgc cgtccagcag ctcgacctgc ccaacgtggc catctacctg 960  
gacgccggcc acgccggctg gctcggctgg cccgccaaacc tgcagcccgc cgccgacctc 1020  
ttcgccggca tctaccgca cgccggccgc ccccgcgccc tgcgcggcct cgccaccaac 1080  
gtggccaact acaacgcctg gagcctgagc tcgccgcccc cgtacacgtc gccaaccccc 1140  
aactacgagc agctgcgctt catccaggcc ttccgcccgc tcctcgaggc caacggctgg 1200  
tccgcccagt tcatcaccga ccagggccgc tccggcaagc agccgactgg tacgtttcca 1260  
ccgttttttt ttttcccttt cttttgttc tttcttcgtc acggcaacgt ctggctgacc 1320  
tttgtttggtg tgtgtctatt acaggccagg aggaatgggg ccactggtgc aaccaggctcg 1380  
gcaccggctt cggcatgcgc ccgacggccg acacgggcta cgacttccag gacgccatcg 1440  
tctgggtcaa gcccggcggc gagagcgacg gcaccagcga cacctccgcc gagcgctacg 1500  
accaccactg cggcctgtcc gacgccctca agcccgtcc ggaggccggc cagtggttcc 1560  
aggcctactt tgagcagctg ctcgagaacg cgaacccgcc gttctaa 1607

5 <210> 14  
<211> 472  
<212> PRT  
<213> *Melanocarpus albomyces*

10 <220>  
<221> misc\_feature

ES 2 612 512 T3

<223> La secuencia de aminoácidos deducida de CBHII/Cel6A de *Melanocarpus albomyces* ALKO4237 (CBS 685.95)

<400> 14

```

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

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

Gln Cys Gly Gly Asn Gly Trp Thr Gly Pro Thr Cys Cys Glu Ser Gly
          35          40          45

Ser Thr Cys Val Ala Gln Asn Asp Tyr Tyr Ser Gln Cys Leu Pro Gly
          50          55          60

Gly Ala Thr Thr Thr Ser Gly Pro Gly Thr Pro Ser Thr Thr Thr Ser
65          70          75          80

Val Thr Thr Ser Thr Ser Gln Gly Ser Pro Thr Ser Pro Pro Pro Thr
          85          90          95

Thr Pro Thr Thr Thr Ile Pro Gly Gly Ala Ser Thr Thr Ala Ser Tyr
          100          105          110

Thr Gly Asn Pro Phe Ala Gly Met Gln Met Trp Ala Asn Ser Tyr Tyr
          115          120          125

Ala Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Met Thr Gly Pro Met
130          135          140

Ala Thr Lys Ala Ala Glu Val Ala Lys Val Pro Ser Phe Gln Trp Leu
145          150          155          160

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

Ile Arg Ala Ala Asn Glu Ala Gly Ala Asn Pro Pro Asn Met Gly Ile
          180          185          190

Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Ala Ala Ser
          195          200          205

Asn Gly Glu Trp Ala Ile Ala Asp Gly Gly Val Ala Asn Tyr Lys Ala
210          215          220

Tyr Ile Asp Arg Ile Arg Glu His Ile Ile Ala Tyr Ser Asp Ile Arg
225          230          235          240

```

5

ES 2 612 512 T3

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

Met Asp Val Pro Lys Cys Ala Asn Ala Ala Asp Thr Tyr Lys Glu Leu  
 260 265 270

Thr Ile Tyr Ala Val Gln Gln Leu Asp Leu Pro Asn Val Ala Ile Tyr  
 275 280 285

Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Leu Gln  
 290 295 300

Pro Ala Ala Asp Leu Phe Ala Gly Ile Tyr Arg Asp Ala Gly Arg Pro  
 305 310 315 320

Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Ala Trp  
 325 330 335

Ser Leu Ser Ser Pro Pro Pro Tyr Thr Ser Pro Asn Pro Asn Tyr Asp  
 340 345 350

Glu Leu Arg Phe Ile Gln Ala Phe Arg Pro Leu Leu Glu Ala Asn Gly  
 355 360 365

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

Thr Gly Gln Glu Glu Trp Gly His Trp Cys Asn Gln Val Gly Thr Gly  
 385 390 395 400

Phe Gly Met Arg Pro Thr Ala Asp Thr Gly Tyr Asp Phe Gln Asp Ala  
 405 410 415

Ile Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr Ser Asp Thr  
 420 425 430

Ser Ala Glu Arg Tyr Asp His His Cys Gly Leu Ser Asp Ala Leu Lys  
 435 440 445

Pro Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala Tyr Phe Glu Gln Leu  
 450 455 460

Leu Glu Asn Ala Asn Pro Pro Phe  
 465 470

<210> 15

<211> 475

<212> PRT

<213> *Chaetomium thermophilum*

5

ES 2 612 512 T3

<220>

<221> misc\_feature

<223> La secuencia de aminoácidos deducida de CBHII/Cel6A de *Chaetomium thermophilum* ALKO4265 (CBS 730.95)

5

<400> 15

Met Ala Lys Gln Leu Leu Leu Thr Ala Ala Leu Ala Ala Thr Ser Leu  
1 5 10 15

Ala Ala Pro Leu Leu Glu Glu Arg Gln Ser Cys Ser Ser Val Trp Gly  
20 25 30

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

Ser Val Cys Thr Tyr Leu Asn Asp Trp Tyr Ser Gln Cys Ile Pro Gly  
50 55 60

Gln Ala Gln Pro Gly Thr Thr Ser Thr Thr Ala Arg Thr Thr Ser Thr  
65 70 75 80

Ser Thr Ser Thr Ser Ser Val Arg Pro Thr Thr Thr Ser Asn Thr Pro  
85 90 95

Val Thr Thr Ala Pro Pro Ala Thr Thr Ile Pro Gly Gly Ala Ser Ser  
100 105 110

Thr Ala Ser Tyr Asn Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala  
115 120 125

Asn Thr Tyr Tyr Ser Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu  
130 135 140

Ser Pro Glu Leu Ala Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser  
145 150 155 160

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

Thr Leu Asn Asp Ile Arg Gly Ala Asn Gln Arg Gly Ala Asn Pro Pro  
180 185 190

Tyr Ala Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala  
195 200 205

Ala Ala Ser Asn Gly Glu Trp Ala Ile Ala Asn Asn Gly Ala Asn Asn  
210 215 220

ES 2 612 512 T3

Tyr Lys Arg Tyr Ile Asp Arg Ile Arg Glu Ile Leu Ile Gln Tyr Ser  
 225 230 235 240

Asp Ile Arg Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn Met  
 245 250 255

Val Thr Asn Met Asn Val Gln Lys Cys Ala Asn Ala Ala Thr Thr Tyr  
 260 265 270

Lys Glu Leu Thr Ile Tyr Ala Leu Lys Gln Leu Asn Leu Pro His Val  
 275 280 285

Ala Met Tyr Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro Ala  
 290 295 300

Asn Ile Gln Pro Ala Ala Glu Leu Phe Gly Gln Leu Tyr Arg Asp Ala  
 305 310 315 320

Gly Lys Pro Ala Ser Val Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr  
 325 330 335

Asn Ala Trp Ser Ile Ala Ser Ala Pro Ser Tyr Thr Ser Pro Asn Pro  
 340 345 350

Asn Tyr Asp Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu Arg  
 355 360 365

Asn Gln Gly Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn Gly  
 370 375 380

Lys Gln Pro Thr Gly Gln Leu Gln Trp Gly Asp Trp Cys Asn Val Lys  
 385 390 395 400

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

Val Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr  
 420 425 430

Ser Asp Thr Ser Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu Ser Asp  
 435 440 445

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

Glu Gln Leu Leu Ile Asn Ala Asn Pro Pro Phe  
 465 470 475

<210> 16



ES 2 612 512 T3

<211> 459  
 <212> PRT  
 <213> *Talaromyces emersonii*

5 <220>  
 <221> misc\_feature  
 <223> La secuencia de aminoácidos deducida de CBHIII/Cel6A de *Talaromyces emersonii* DSM 2432 (RF8069)

<400> 16  
 Met Arg Asn Leu Leu Ala Leu Ala Pro Ala Ala Leu Leu Leu Gly Ala  
 1 5 10 15  
 Ala Asp Ala Gln Gln Ser Leu Trp Gly Gln Cys Gly Gly Asn Ser Trp  
 20 25 30  
 Thr Gly Ala Thr Asp Cys Ala Ala Gly Ala Thr Cys Ser Thr Ile Asn  
 35 40 45  
 Ser Tyr Tyr Ala Gln Cys Val Pro Ala Thr Ala Thr Pro Thr Thr Leu  
 50 55 60  
 Thr Thr Thr Thr Lys Pro Ser Ser Thr Ala Pro Thr Thr Pro Pro Pro  
 65 70 75 80  
 Thr Ser Ala Thr Thr Thr Gly Thr Gly Ser Ala Thr Ser Pro Ser Ile  
 85 90 95  
 Thr Ala Ser Ala Ser Gly Asn Pro Phe Val Gly Tyr Gln Leu Tyr Ala  
 100 105 110  
 Asn Pro Tyr Tyr Ala Ser Glu Val Ile Ser Leu Ala Ile Pro Ser Leu  
 115 120 125  
 Ser Ser Glu Leu Val Pro Lys Ala Ser Glu Val Ala Lys Val Pro Ser  
 130 135 140  
 Phe Val Trp Leu Asp Gln Ala Ala Lys Val Pro Asn Met Gly Glu Tyr  
 145 150 155 160  
 Leu Lys Asp Ile Gln Ser Gln Asn Ala Ala Gly Ala Asp Pro Pro Ile  
 165 170 175  
 Ala Gly Ile Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala  
 180 185 190  
 Ala Ala Ser Asn Gly Glu Phe Ser Ile Ala Asn Asn Gly Val Ala Leu  
 195 200 205  
 10 Tyr Lys Gln Tyr Ile Asp Ser Ile Arg Glu Gln Leu Thr Thr Tyr Ser

ES 2 612 512 T3

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

## REIVINDICACIONES

1. Procedimiento para tratar material celulósico con un polipéptido CBHII/Cel6A o una preparación enzimática que comprende dicho polipéptido, en el que dicho polipéptido CBHII/Cel6A presenta una actividad de celobiohidrolasa y comprende una secuencia de aminoácidos que presenta una identidad de por lo menos 78% respecto al polipéptido de longitud completa de SEC ID nº 12, o un fragmento del mismo que presenta una actividad de celobiohidrolasa, y en el que dicho procedimiento comprende las etapas siguientes:
- 5
- 10 i) producir dicho polipéptido CBHII/Cel6A o una preparación enzimática que comprende dicho polipéptido o un microorganismo fermentativo que produce dicho polipéptido;
- ii) hacer reaccionar el material celulósico con dicho polipéptido CBHII/Cel6A o la preparación enzimática que comprende dicho polipéptido o el microorganismo fermentativo que produce dicho polipéptido; y
- 15 iii) obtener un material celulósico parcial o completamente hidrolizado;
- en el que el polipéptido CBHII/Cel6A es activo a una temperatura de entre 50°C y 70°C.
2. Procedimiento según la reivindicación 1, en el que el polipéptido CBHII/Cel6A puede obtenerse a partir de un género de *Acremonium*, más preferentemente a partir de *A. thermophilum*, todavía más preferentemente a partir de la cepa depositada *A. thermophilum* CBS 116240.
- 20
3. Procedimiento según la reivindicación 1 o 2, en el que el polipéptido CBHII/Cel6A presenta la secuencia de aminoácidos de SEC ID nº 12.
- 25
4. Procedimiento según cualquiera de las reivindicaciones 1 a 3, en el que el polipéptido CBHII/Cel6A puede hidrolizar material celulósico a una temperatura de entre 40°C y 70°C.
5. Procedimiento según cualquiera de las reivindicaciones 1 a 4, en el que el material celulósico se trata con una composición enzimática que comprende dicho polipéptido CBHII/Cel6A en combinación con por lo menos un enzima adicional que puede hidrolizar dicho material celulósico seleccionado de entre un grupo de celobiohidrolasa, endoglucanasa, beta-glucosidasa, beta-glucanasa, xiloglucanasa, xilanasa, beta-xilosidasa, mananasa, beta-manosidasa,  $\alpha$ -glucuronidasa, acetil xilán esterasa,  $\alpha$ -arabinofuranosidasa,  $\alpha$ -galactosidasa, pectinasa, endo- $\alpha$ -L-arabinasas y exo- $\alpha$ -L-arabinasas,  $\alpha$ -galactosidasa, endo- y exo-galactoronasa, endopectiniliasa, pectato liasa y pectinesterasa, fenol esterasa, ligninasa, lignina peroxidasa, peroxidasa dependiente de manganeso, enzima generador de H<sub>2</sub>O<sub>2</sub> y lacasa con o sin mediadores.
- 30
- 35
6. Procedimiento según la reivindicación 5, en el que los enzimas se añaden al material celulósico de manera simultánea o secuencial.
- 40
7. Procedimiento según cualquiera de las reivindicaciones 1 a 6, en el que el material celulósico se selecciona de entre el grupo que consiste en materiales vegetales, biomasa agrícola, productos de residuo y cultivos energéticos especializados.
- 45
8. Procedimiento según cualquiera de las reivindicaciones 1 a 7, en el que dicho polipéptido CBHII/Cel6A es aplicable en la producción de biocombustible.
9. Procedimiento según cualquiera de las reivindicaciones 1 a 8, en el que dicho polipéptido CBHII/Cel6A se obtiene a partir de *A. thermophilum* CBS 116240 y presenta la secuencia de aminoácidos de SEC ID nº 12.
- 50
10. Polipéptido CBHII/Cel6A fúngico, en el que dicho polipéptido presenta una actividad de celobiohidrolasa y comprende una secuencia de aminoácidos que presenta una identidad de por lo menos 78% respecto al polipéptido de longitud completa de SEC ID nº 12, o un fragmento del mismo que presenta una actividad de celobiohidrolasa, en el que el polipéptido CBHII/Cel6A es activo a una temperatura de entre 50°C y 70°C.
- 55
11. Polipéptido CBHII/Cel6A según la reivindicación 10, en el que dicho polipéptido puede obtenerse a partir de un género de *Acremonium*, más preferentemente a partir de la especie *A. thermophilum*, todavía más preferentemente a partir de *A. thermophilum* CBS 116240.
- 60
12. Polipéptido CBHII/Cel6A según la reivindicación 10 u 11, en el que dicho polipéptido presenta la secuencia de aminoácidos de SEC ID nº 12.
13. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 12, en el que dicho polipéptido puede hidrolizar material celulósico a una temperatura de entre 40°C y 70°C.
- 65
14. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 13, en el que dicho polipéptido se

encuentra codificado por una molécula de ácidos nucleicos aislada que comprende la SEC ID nº 11.

15. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 14, en el que dicho polipéptido se encuentra codificado por una molécula de ácidos nucleicos aislada que hibrida bajo condiciones restrictivas con la secuencia polinucleotídica de SEC ID nº 11.

16. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 15, en el que dicho polipéptido se encuentra codificado por la secuencia polinucleotídica incluida en el plásmido pALK2582 depositada en *Escherichia coli* bajo el número de acceso DSM 22946.

17. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 16, en el que dicho polipéptido se produce a partir de un vector de expresión recombinante que comprende la molécula de ácidos nucleicos, que comprende una secuencia polinucleotídica que codifica un enzima CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 16 ligada funcionalmente a unas secuencias reguladoras que pueden dirigir la expresión de la molécula de ácidos nucleicos codificante de CBHII/Cel6A en un hospedador adecuado.

18. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 17, en el que dicho polipéptido se produce en un hospedador heterólogo, preferentemente en un hospedador microbiano.

19. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 18, en el que dicho polipéptido se produce en un hospedador de género *Trichoderma*, *Aspergillus*, *Fusarium*, *Humicola*, *Chrysosporium*, *Neurospora*, *Rhizopus*, *Penicillium* o *Mortierella*.

20. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 19, en el que dicho polipéptido es producido en *Trichoderma* o *Aspergillus*, preferentemente en *Trichoderma reesei*.

21. Polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 20, en el que dicho polipéptido CBHII/Cel6A se obtiene a partir de *A. thermophilum* CBS 116240 y presenta la secuencia de aminoácidos de SEC ID nº 12.

22. Molécula de ácidos nucleicos aislada, que comprende una secuencia polinucleotídica que codifica un polipéptido CBHII/Cel6A fúngico seleccionado de entre el grupo que consiste en:

(a) una molécula de ácidos nucleicos, que comprende una secuencia polinucleotídica que codifica un polipéptido que presenta una actividad de celobiohidrolasa y que comprende la secuencia de aminoácidos de longitud completa tal como se representa en la SEC ID nº 12, o un fragmento del mismo que presenta una actividad de celobiohidrolasa;

(b) una molécula de ácidos nucleicos, que comprende una secuencia polinucleotídica que codifica un polipéptido que presenta una actividad de celobiohidrolasa y una identidad de por lo menos 78% respecto a la secuencia de aminoácidos de longitud completa de SEC ID nº 12 o un fragmento del mismo que presenta una actividad de celobiohidrolasa;

(c) una molécula de ácidos nucleicos que comprende la secuencia codificante de la secuencia polinucleotídica representada como SEC ID nº 11;

(d) una molécula de ácidos nucleicos que comprende la secuencia codificante de la secuencia polinucleotídica contenida en DSM 22946;

(e) una molécula de ácidos nucleicos cuya secuencia codificante difiere de la secuencia codificante de una molécula de ácidos nucleicos de cualquiera de (c) a (d) debido a la degeneración del código genético; y

(f) una molécula de ácidos nucleicos que hibrida bajo condiciones restrictivas con una molécula de ácidos nucleicos contenida en DSM 22946 y que codifica un polipéptido que presenta una actividad de celobiohidrolasa y una secuencia de aminoácidos que muestra una identidad de por lo menos 78% respecto a la secuencia de aminoácidos de longitud completa tal como se representa en SEC ID nº 12, o un fragmento del mismo que presenta actividad de celobiohidrolasa;

en la que el polipéptido CBHII/Cel6A es activo a una temperatura de entre 50°C y 70°C.

23. Molécula de ácidos nucleicos según la reivindicación 22, en la que dicha molécula de ácidos nucleicos que codifica CBHII/Cel6A se obtiene a partir de *A. thermophilum* CBS 116240 y presenta la secuencia polinucleotídica de SEC ID nº 11.

24. Vector de expresión recombinante que comprende la molécula de ácidos nucleicos según la reivindicación 22 o 23 ligada funcionalmente a unas secuencias reguladoras que pueden dirigir la expresión del gen de celobiohidrolasa

CBHII/Cel6A y la producción de dicha celobiohidrolasa CBHII/Cel6A en un hospedador adecuado.

25. Célula hospedadora que comprende el vector de expresión recombinante según la reivindicación 24.

5 26. Célula hospedadora según la reivindicación 25, caracterizada por que dicho hospedador es un hospedador microbiano.

27. Célula hospedadora según la reivindicación 25 o 26, caracterizada por que dicho hospedador es un hongo filamentoso.

10 28. Célula hospedadora según cualquiera de las reivindicaciones 25 a 27, caracterizada por que dicho hospedador es de un género *Trichoderma*, *Aspergillus*, *Fusarium*, *Humicola*, *Chrysosporium*, *Neurospora*, *Rhizopus*, *Penicillium* o *Mortierella*.

15 29. Célula hospedadora según la reivindicación 28, caracterizada por que dicho hospedador es *Trichoderma* o *Aspergillus*, preferentemente *Trichoderma reesei*.

30. Célula hospedadora según cualquiera de las reivindicaciones 25 a 29, caracterizada por que dicho hospedador es un microorganismo fermentador.

20 31. Procedimiento de producción de un polipéptido que presenta una actividad de celobiohidrolasa, comprendiendo dicho procedimiento las etapas de cultivar la célula hospedadora según cualquiera de las reivindicaciones 25 a 30 y recuperar el polipéptido.

25 32. Procedimiento para obtener una preparación enzimática que comprende las etapas de cultivar una célula hospedadora según cualquiera de las reivindicaciones 25 a 30 y preparar el caldo de cultivo completo, o de separar las células del medio de cultivo agotado y obtener el sobrenadante.

30 33. Preparación enzimática, que comprende el enzima CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 21.

34. Preparación enzimática según la reivindicación 33, en la que la celobiohidrolasa CBHII/Cel6A se obtiene a partir de *Acremonium thermophilum* y presenta la secuencia de aminoácidos de SEC ID nº 12.

35 35. Preparación enzimática según la reivindicación 33 o 34, caracterizada por que dicha preparación comprende otros enzimas seleccionados de entre el grupo de celobiohidrolasa, endoglucanasa, beta-glucosidasa, beta-glucanasa, xiloglucanasa, xilanasas, beta-xilosidasa, mananasa, beta-manosidasa,  $\alpha$ -glucuronidasa, acetil-xilán-esterasa,  $\alpha$ -arabinofuranosidasa,  $\alpha$ -galactosidasa, pectinasa, endo- $\alpha$ -L-arabinasas y exo- $\alpha$ -L-arabinasas,  $\alpha$ -galactosidasa, endo- y exo-galactoronasa, endopectinliasa, pectato liasa y pectinesterasa, fenol esterasa, ligninasa lignina peroxidasa, peroxidasa dependiente de manganeso, enzima generador de H<sub>2</sub>O<sub>2</sub> y lacasa con o sin un mediador.

40 36. Preparación enzimática según cualquiera de las reivindicaciones 33 a 35, caracterizada por que dicha preparación enzimática se encuentra en forma de caldo de cultivo completo, medio de cultivo agotado, líquido, polvo o granulado.

45 37. Utilización del polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 21 o la preparación enzimática según cualquiera de las reivindicaciones 33 a 36 para la producción de biocombustible, para detergentes, para el tratamiento de fibras, para el tratamiento de alimento o pienso, para pasta y papel, para bebidas o para cualesquiera aplicaciones que impliquen la hidrólisis o la modificación de material celulósico.

50 38. Utilización según la reivindicación 37, en la que el polipéptido CBHII/Cel6A según cualquiera de las reivindicaciones 10 a 21 o la preparación enzimática según cualquiera de las reivindicaciones 33 a 36 se utiliza para la producción de biocombustible.

55

Fig. 1

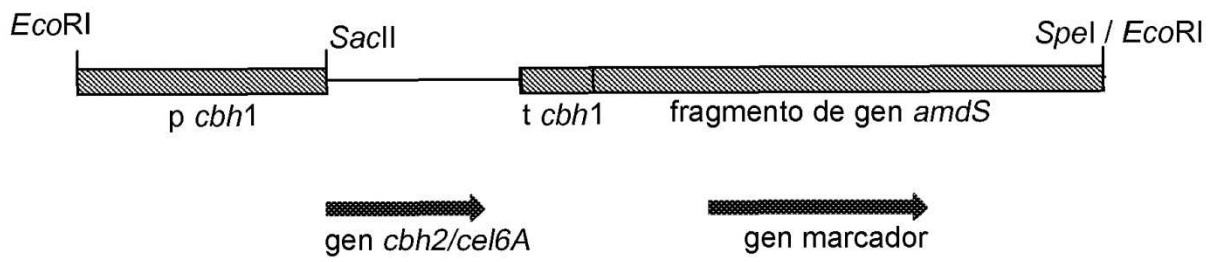


Fig. 2

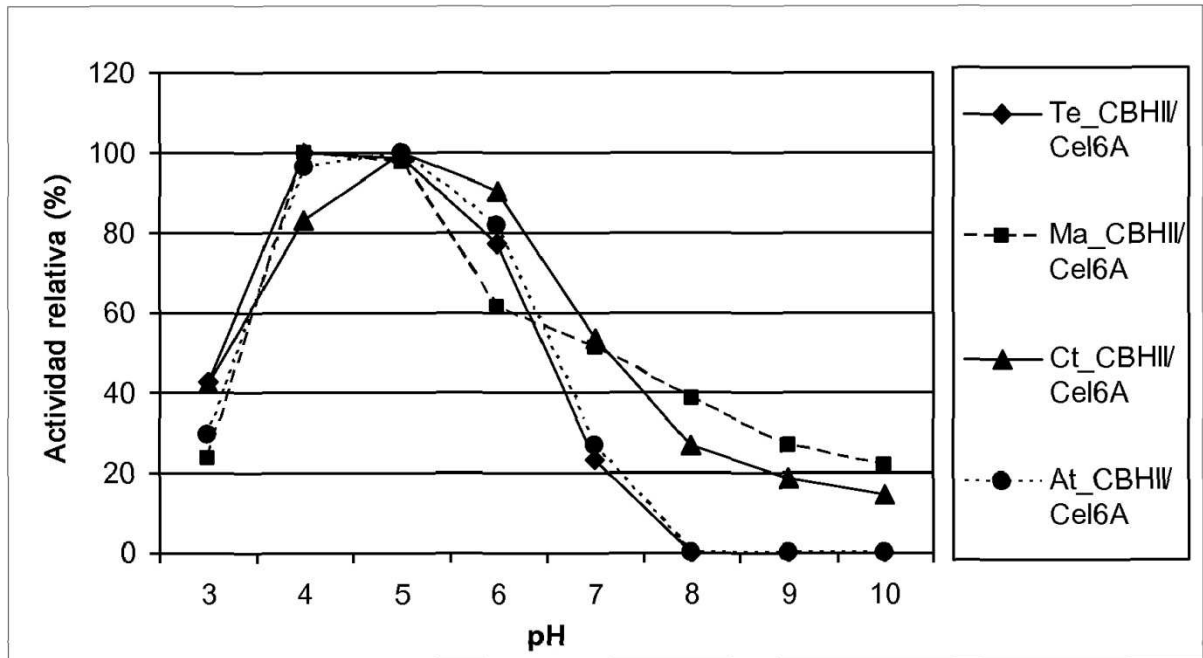


Fig. 3

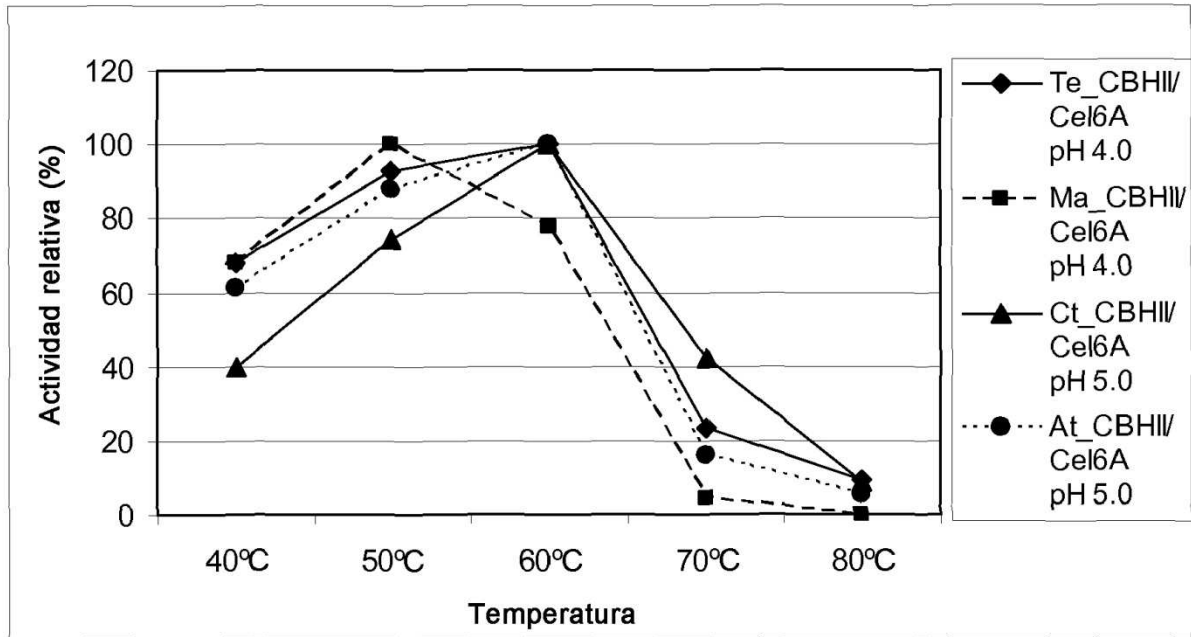




Fig. 4A

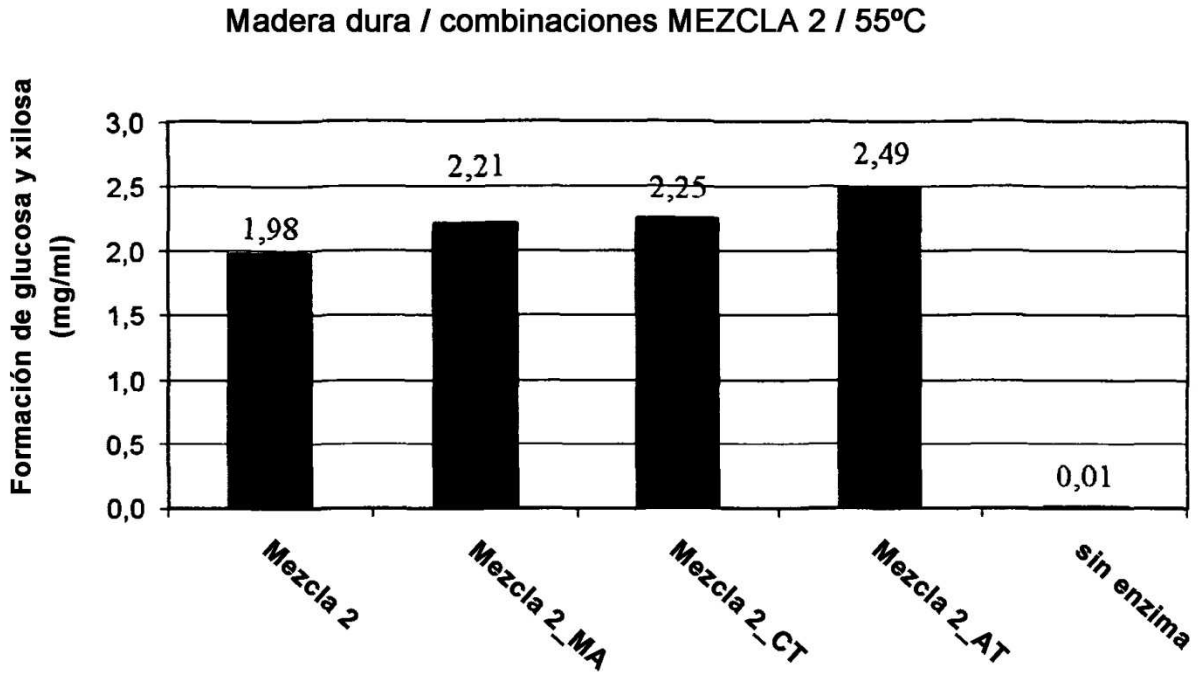


Fig. 4B

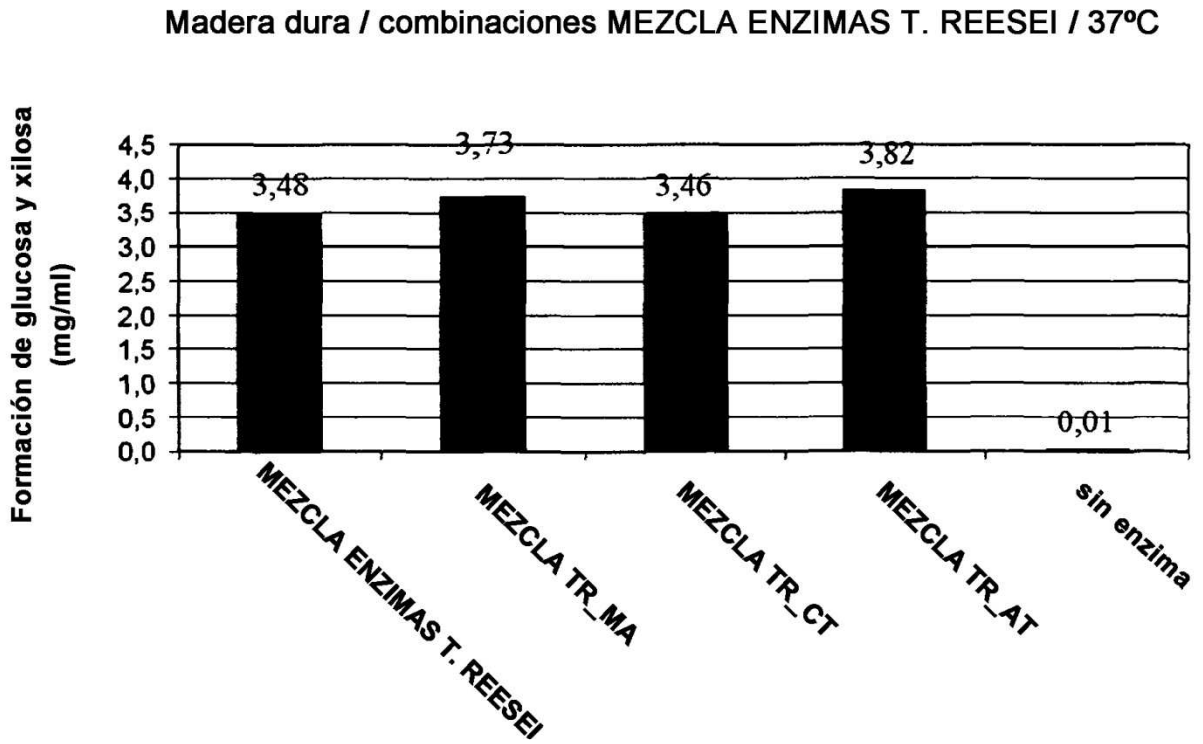


Fig. 4C

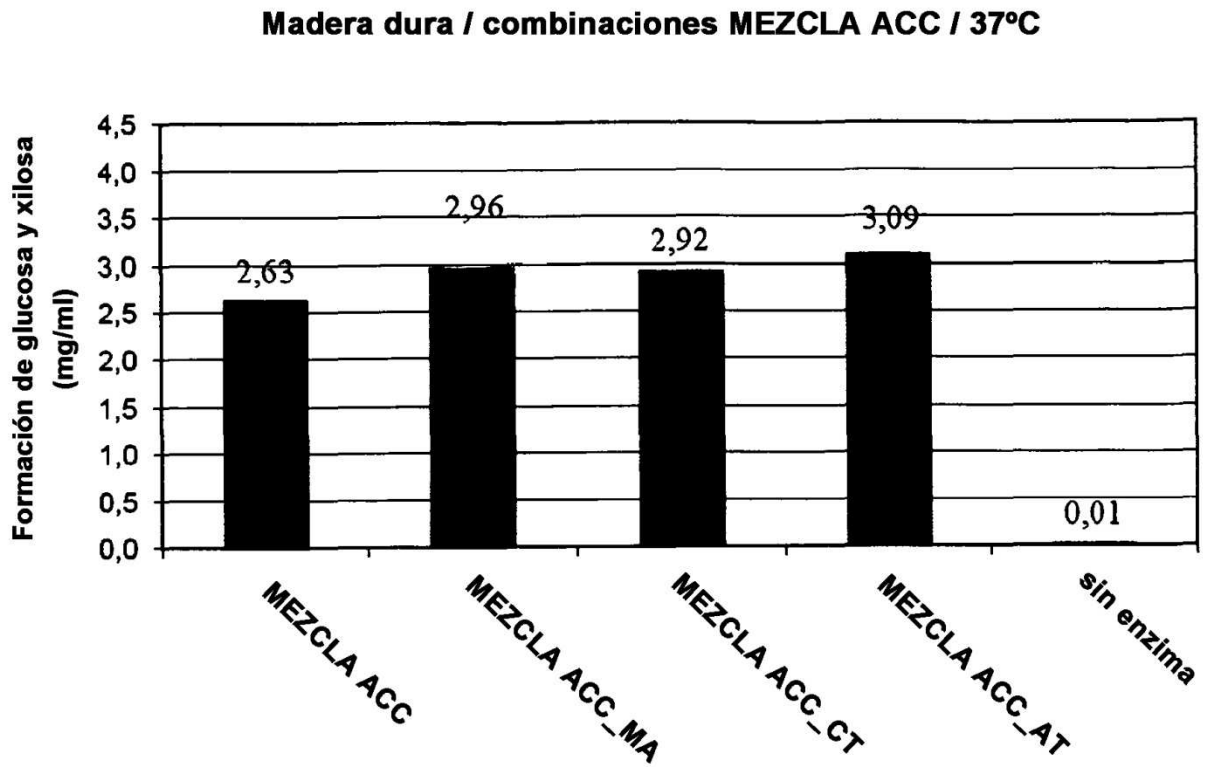


Fig. 5

