

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 521 940**

21 Número de solicitud: 201330678

51 Int. Cl.:

**C12N 9/42** (2006.01)

**C12P 19/02** (2006.01)

**C12P 7/10** (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

**10.05.2013**

43 Fecha de publicación de la solicitud:

**13.11.2014**

71 Solicitantes:

**ABENGOA BIOENERGÍA NUEVAS  
TECNOLOGÍAS, S.A. (100.0%)  
C/ Energía Solar, 1 Campus Palmas Altas  
41014 Sevilla ES**

72 Inventor/es:

**DÍEZ GARCÍA, Bruno;  
GÓMEZ RODRÍGUEZ, Ana;  
GIL MARTÍNEZ, Jorge;  
VALBUENA CRESPO, Noelia;  
MORENO PÉREZ, Antonio Javier;  
DUEÑAS SÁNCHEZ, Rafael;  
MUÑOZ GONZÁLEZ, Ana María;  
PÉREZ GÓMEZ, Dolores;  
GAVALDÁ MARTÍN, Sandra;  
SÁNCHEZ ZAMORANO, Laura;  
ÁLVAREZ NÚÑEZ, Consolación;  
BERMÚDEZ ALCANTARA, María De Los Ángeles;  
GUTIÉRREZ GÓMEZ, Pablo;  
ARJONA ANTOLÍN, Ricardo y  
MARTÍN PÉREZ, Lucía**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

54 Título: **Variantes de beta-glucosidasa con actividad de transglicosilación reducida**

57 Resumen:

Variantes de beta-glucosidasa con actividad de transglicosilación reducida.

La presente invención se refiere a variantes de beta-glucosidasa con actividad de transglicosilación reducida. La invención también se refiere a una construcción génica, a una célula huésped y a una composición enzimática que comprende dichas variantes. También se proporciona un procedimiento para producir dichas variantes, un procedimiento para producir azúcar fermentable y un procedimiento para producir un bioproducto, tal como bioetanol, a partir de material celulósico con las variantes de beta-glucosidasa, la célula huésped o la composición enzimática que comprende dichas variantes.

**ES 2 521 940 A1**

**VARIANTES DE BETA-GLUCOSIDASA CON ACTIVIDAD DE TRANSGLICOSILACIÓN**  
**REDUCIDA**

**DESCRIPCIÓN**

5

La invención pertenece al campo de las enzimas útiles para la producción de bioproductos y, más particularmente, a variantes de la beta-glucosidasa y a su uso en la producción de azúcares fermentables y de etanol a partir de material celulósico.

10 **TÉCNICA ANTERIOR**

La biomasa vegetal proporciona una fuente abundante de potencial energía en forma de azúcares y, por tanto, es una importante fuente renovable para la generación de azúcares fermentables. La fermentación de estos azúcares puede dar lugar a productos finales  
15 comercialmente valiosos, tales como el etanol también denominado bioetanol.

Aunque la fermentación de los azúcares a etanol es relativamente directa, la conversión eficiente de biomasa celulósica en azúcares fermentables, tales como la glucosa, supone un mayor desafío. La enorme energía potencial de las grandes cantidades de hidratos de  
20 carbono que integran la biomasa vegetal no está suficientemente utilizada porque los azúcares forman parte de polímeros complejos (polisacáridos, tales como celulosa y hemicelulosa) y, por tanto, no son fácilmente accesibles para la fermentación. Por tanto, la celulosa se puede tratar previamente, de forma mecánica, química, enzimática o de otros  
25 modos, para aumentar su susceptibilidad a la hidrólisis. Después, tras este proceso de pretratamiento tiene lugar una etapa de sacarificación, que es un proceso enzimático por el cual los hidratos de carbono complejos (como almidón o celulosa) se hidrolizan en sus componentes monosacáridos. El objetivo de cualquier tecnología de sacarificación es alterar o eliminar los impedimentos estructurales y de composición para la hidrólisis con el fin de  
30 mejorar la tasa de hidrólisis enzimática y aumentar los rendimientos de azúcares fermentables a partir de celulosa o de hemicelulosas (N. Mosier y *col.*, 2005, Bioresource Technology 96, 673–686). Después de esta etapa de sacarificación se realiza un proceso de fermentación.

La hidrólisis enzimática de los polisacáridos en azúcares solubles y, por último, en

monómeros tales como xilosa, glucosa y otras pentosas y hexosas, se cataliza mediante varias enzimas que en conjunto se denominan "celulasas". Las celulasas (1,4-beta-D-glucano-4-glucanohidrolasa, EC 3.2.1.4) son complejos multienzimáticos que comprenden tres componentes principales, endo- $\beta$ -glucanasa (EC 3.2.1.4), exo- $\beta$ -glucanasa o celobiohidrolasa (EC 3.2.1.9.1) y  $\beta$ -glucosidasa (EC 3.2.1.21), que se ha demostrado que actúan de forma sinérgica en la hidrólisis de la celulosa (Ekperigin, M.M., 2007, African Journal of Biotechnology, Vol.6 (1), pág. 028-033).

Específicamente, la beta-glucosidasa es una enzima glucosidasa que actúa sobre los enlaces  $\beta$ 1- $\rightarrow$ 4 que unen dos glucosas o moléculas sustituidas con glucosa (es decir, el disacárido celobiosa). Es una exocelulasa con especificidad por diversos sustratos de beta-D-glucósidos. Cataliza la hidrólisis de residuos terminales no reductores en los beta-D-glucósidos con liberación de glucosa. Por tanto, se usa ampliamente junto con otras celulasas en procesos para la conversión de la biomasa celulósica en azúcares fermentables. Su aplicación en la conversión de biomasa con un contenido elevado en celulosa en azúcares fermentables para la producción de etanol combustible es un área extensamente estudiada.

Por consiguiente, las celulasas microbianas se han convertido en biocatalizadores focales debido a su naturaleza compleja y a sus extensas aplicaciones industriales (Kuhad R. C. y *col.*, 2011, Enzyme Research, Article ID 280696). Hoy en día se ha prestado una considerable atención a los conocimientos actuales sobre la producción de celulasas y los retos en la investigación sobre celulasas, especialmente en la dirección de la mejora de la economía del proceso de varias industrias, con el fin de obtener celulasas con mayor actividad y mejores propiedades. En este sentido, se han diseñado beta-glucosidasas con mejor actividad hidrolítica o termoestabilidad (WO2011063308A2).

Las beta-glucosidasas (BGLs, EC 3.2.1.21) catalizan la transferencia de grupos glicosil entre nucleófilos de oxígeno sobre diferentes tipos de sustratos. Las BGLs con un mecanismo de retención (Rye y *col.*, 2000. Curr. Opin. Chem. Biol. 4), como es el caso de Bgl1, tienen actividad hidrolítica reducida a concentraciones elevadas de sustrato o de producto y esto depende al menos de que se produzcan dos fenómenos diferentes: inhibición y transglicosilación (Bohlin *et al.*, 2013, Appl. Microbiol. Biotechnol., 97(1):159-169). El mecanismo de retención forma inicialmente un complejo glicosil covalente con la enzima.

Este intermedio se descompone en la segunda etapa de la reacción, lo que se puede producir mediante una molécula de agua (actividad hidrolítica) o mediante un azúcar aceptor (actividad de transglicosilación). En la reacción de transglicosilación más sencilla, el aceptor es el propio sustrato. No obstante, la celobiosa se ha descrito como sustrato preferencial para retener a las BGLs en su reacción de transglicosilación.

Una beta glucosidasa con una actividad hidrolítica capaz de producir glucosa de forma eficiente pero con una actividad transglicosiladora reducida sería de utilidad, pudiendo disminuir la síntesis de otros compuestos, tales como etil-beta-D-glucopiranosido en presencia de etanol, de modo que aumentase la producción de azúcares fermentables y mejorara el rendimiento de una mezcla enzimática que contuviera BGL.

## **DESCRIPCIÓN DE LA INVENCIÓN**

La presente invención describe nuevas variantes de beta-glucosidasa con capacidad de transglicosilación reducida para la hidrólisis de material celulósico en azúcares fermentables, así como un procedimiento para producir dichas variantes de beta-glucosidasa, un procedimiento para producir azúcares fermentables y un procedimiento para producir bioproductos, tales como etanol, con dichas variantes.

Por tanto, la presente invención representa una solución a la necesidad de proporcionar variantes de beta-glucosidasa con la propiedad mejorada de una actividad de transglicosilación reducida para la optimización de la etapa de hidrólisis de material celulósico en azúcares fermentables.

La capacidad de transglicosilación de las beta-glucosidasas disminuye la concentración final de azúcares fermentables en etanol y, en consecuencia, el rendimiento del producto final del proceso. Sin embargo, los inventores han demostrado que las variantes de beta-glucosidasa de la presente invención tienen actividad de transglicosilación reducida y, por tanto, aumentan significativamente el rendimiento de la etapa de hidrólisis, lo que conduce a un aumento de la producción del bioproducto, preferiblemente etanol.

La presente invención también proporciona un procedimiento para la selección de variantes de beta-glucosidasa con actividad de transglicosilación reducida. Por tanto, un primer

aspecto de la presente invención se refiere a un procedimiento para la selección de variantes de beta-glucosidasa con actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa, en lo sucesivo "primer procedimiento de la invención", que comprende:

5

a) Incubar las variantes de beta-glucosidasa con p-nitro-fenil-glucopiranosido en presencia de celobiosa,

b) Medir la liberación de p-nitrofenol,

10

c) Comparar el valor del p-nitrofenol medido en la etapa (b) con un valor de referencia, y

d) Seleccionar aquellas variantes de beta-glucosidasa cuyos valores medidos en la etapa (b) son más elevados que el valor de referencia.

El término "beta-glucosidasa", como se usa aquí, se refiere a una enzima que cataliza la hidrólisis de un dímero de azúcar, incluyendo, pero sin limitarnos, celobiosa, con la liberación de un monómero de azúcar correspondiente, la cual se usa, pero sin limitación, para la síntesis de etanol. La enzima beta-glucosidasa actúa sobre los enlaces  $\beta 1 \rightarrow 4$  que unen dos glucosas o moléculas sustituidas con glucosa (es decir, el disacárido celobiosa). Es una exocelulasa con especificidad por una variedad de sustratos beta-D-glucósido. Cataliza la hidrólisis de residuos terminales no reductores en los beta-D-glucósidos con liberación de glucosa.

El término "variante", como se usa aquí, se refiere a una enzima que procede de una enzima nativa mediante una o más deleciones, inserciones y/o sustituciones de uno o más aminoácidos en los punto(s) dentro de su secuencia de aminoácidos y, por tanto, tiene una secuencia diferente a la de la enzima nativa. Como se usa aquí, la expresión "variante de beta-glucosidasa" significa un polipéptido que tiene actividad beta-glucosidasa producido por un organismo que expresa una secuencia de nucleótidos modificada que codifica una beta-glucosidasa nativa. Dicha secuencia de nucleótidos modificada se obtiene mediante intervención humana mediante modificación de la secuencia de nucleótidos que codifica una beta-glucosidasa nativa. El término "modificación" significa en el presente documento cualquier modificación química de la secuencia de aminoácidos o de ácido nucleico de una beta-glucosidasa nativa.

La expresión "beta-glucosidasa nativa" se refiere a una enzima beta-glucosidasa o su preproteína, expresada por un microorganismo tal como bacterias u hongos filamentosos. Preferentemente, la enzima beta-glucosidasa nativa a la que se hace referencia en la presente invención es expresada por un hongo filamentosos, más preferiblemente por un hongo que pertenece al género *Myceliophthora*, aún más preferiblemente por *Myceliophthora thermophila*, aún más preferiblemente la enzima beta-glucosidasa nativa es la enzima de SEQ ID NO: 1 (también denominada Bgl1) o SEQ ID NO: 2. La SEQ ID NO: 2 es la preproteína de la SEQ ID NO: 1 y consiste en un péptido señal correspondiente a los aminoácidos 1 a 19 de la SEQ ID NO: 2 unido a la SEQ ID NO: 1.

10

La "actividad de transglicosilación" de las beta-glucosidasas consiste en la formación inicial de un complejo glicosil covalente con la enzima y su posterior descomposición en la segunda etapa de la reacción mediante un azúcar aceptor, en lugar de producirse mediante una molécula de agua (actividad hidrolítica). En la reacción de transglicosilación más sencilla, el azúcar aceptor es el propio sustrato.

15

El sustrato p-nitro-fenil-glucopiranosido también se denomina pNGP. Este sustrato es hidrolizado por la beta-glucosidasa. Por tanto, este sustrato se incuba en condiciones que permiten que las variantes de beta-glucosidasa realicen el proceso hidrolítico en la etapa (a) del primer procedimiento de la invención. Adicionalmente, dicha incubación se realiza en presencia de celobiosa, ya que este disacárido se ha descrito como sustrato preferencial para retener beta-glucosidasas en su actividad de transglicosilación. Por tanto, en presencia de celobiosa, la actividad hidrolítica de beta-glucosidasa sobre pNGP muestra una reducción debido a la actividad de transglicosilación. Un porcentaje elevado de capacidad hidrolítica, medido mediante el p-nitrofenol liberado en presencia del sustrato celobiosa altamente susceptible a la transglicosilación, se usa como indicación de una baja capacidad de transglicosilación. Por tanto, cuanto más p-nitrofenol se mide en la etapa (b) del primer procedimiento de la invención, menor es la capacidad de transglicosilación de la variante de beta-glucosidasa analizada.

20

25

30

La expresión "valor de referencia" se refiere a la medición de p-nitrofenol liberado por una beta-glucosidasa nativa, preferiblemente la SEQ ID NO: 1 o la SEQ ID NO: 2, incubada con pNGP en presencia de celobiosa. Una beta-glucosidasa nativa incubada con pNGP en presencia de celobiosa tiene una producción baja de p-nitrofenol, dado que la enzima

utilizará la celobiosa como sustrato para realizar el proceso de transglicosilación en lugar de utilizar pNGP para llevar a cabo el proceso hidrolítico que genera p-nitrofenol. Por tanto, la medición del p-nitrofenol liberado tras la incubación en la etapa (a) y su comparación con la medición del p-nitrofenol liberado tras la incubación de una beta-glucosidasa nativa con pNGP en presencia de celobiosa permite la selección de las variantes de beta-glucosidasa con menor capacidad de transglicosilación que la de una beta-glucosidasa nativa. La comparación en la etapa (c) se puede realizar manualmente o por ordenador.

5

También se puede reconocer una baja actividad de transglicosilación cuando se mide en la etapa (b) una capacidad hidrolítica en pNGP de entre 30-70%.

10

Las variantes de la beta-glucosidasa usadas en el primer procedimiento de la invención pueden derivar bien de una librería de mutantes ya conocidos en el estado de la técnica o disponibles comercialmente o bien se pueden diseñar por medio de cualquier procedimiento conocido por los expertos en la materia para generar una librería de mutantes de una enzima. Los mutantes que constituyen dicha librería pueden comprender sustituciones, deleciones y/o inserciones de uno o más aminoácidos en sus secuencias de aminoácidos, así como sustitución de una o más cadenas laterales de aminoácido. Por tanto, otro aspecto de la invención hace referencia a un procedimiento para producir variantes de beta-glucosidasa con actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa, en lo sucesivo "segundo procedimiento de la invención", que comprende el diseño de una librería de mutantes de enzimas beta-glucosidasa y la selección de variantes de beta-glucosidasa con actividad de transglicosilación reducida de acuerdo con las etapas (a) a (d) del primer procedimiento de la invención

15

20

25

Otro aspecto de la invención hace referencia a una variante de beta-glucosidasa aislada producida por el segundo procedimiento de la invención.

30

Otro aspecto de la invención hace referencia a una variante de beta-glucosidasa aislada que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos un 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% con la SEQ ID NO: 2 y comprende una sustitución de aminoácido en la posición Q211 correspondiente a las posiciones 1 a 871 de la SEQ ID NO: 2, donde la variante de beta-glucosidasa tiene una actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa, en lo

sucesivo denominada "variante de beta-glucosidasa de la invención". La actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa se mide, preferiblemente, por medio del primer procedimiento de la invención descrito arriba.

- 5 En una realización preferida, la sustitución de aminoácido de dicha variante es Q211H.

El término "identidad", como se usa aquí, en el contexto de describir dos o más secuencias polipeptídicas, hace referencia a un porcentaje especificado de coincidencias de residuos de aminoácidos en las posiciones desde una alineación de dos secuencias de aminoácidos.

- 10 Los procedimientos de alineación de secuencias para comparar son bien conocidos en la técnica. El grado de identidad se puede determinar mediante el método de Clustal (Higgins, 1989, CABIOS 5: 151-153), el método de Wilbur-Lipman (Wilbur y Lipman, 1983, *Proceedings of the National Academy of Science USA* 80: 726-730), el programa GAG, incluyendo GAP (Devereux *et al.* 1984, *Nucleic Acids Research* 12: 287 Genetics Computer  
15 Group University of Wisconsin, Madison, (WI)); BLAST o BLASTN, EMBOSS Needle y FASTA (Altschul *et al.* 1999, *J. Mol. Biol.* 215: 403-410). Además, se puede usar el algoritmo de Smith Waterman con el fin de determinar el grado de identidad entre dos secuencias.

- La variante de beta-glucosidasa de la invención puede exhibir cambios limitados en su  
20 secuencia de aminoácidos. Estos cambios permiten el mantenimiento de la función de la variante de beta-glucosidasa y la actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa. Estos cambios pueden ser sustituciones, deleciones o adiciones. Las sustituciones son por aminoácidos conservados que son aminoácidos con cadenas laterales y propiedades similares con respecto a, por ejemplo, propiedades  
25 hidrofóbicas o aromáticas. Estas sustituciones incluyen, pero no se limitan a, sustituciones entre Glu y Asp, Lys y Arg, Asn y Gln, Ser y Thr, y/o entre los aminoácidos incluidos en la lista siguiente: Ala, Leu, Val e Ile. Los cambios no conducen a modificaciones relevantes en las características o propiedades esenciales de la variante beta-glucosidasa de la invención.

- 30 Asimismo, la sustitución de aminoácido en la posición Q211, correspondiente a las posiciones 1 a 871 de la SEQ ID NO: 2, es por aminoácidos que tienen las mismas propiedades, sobre, por ejemplo, las propiedades hidrofóbicas o aromáticas, que el aminoácido H. Por tanto, dichas sustituciones permiten que las variantes de beta-glucosidasa de la invención mantengan la misma función que la variante preferida de SEQ



ID NO: 4, incluyendo la actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa.

5 En una realización más preferida, la variante de beta-glucosidasa de la invención comprende la secuencia de aminoácidos SEQ ID NO: 4. Un ejemplo de variante de beta-glucosidasa de la invención que comprende la secuencia de aminoácidos SEQ ID NO: 4 es el polipéptido de SEQ ID NO: 5, que es la preproteína de SEQ ID NO: 4, que consiste en un péptido señal correspondiente a los aminoácidos 1 a 19 de la SEQ ID NO: 5 unido a la SEQ ID NO: 4. Por tanto, en una realización más preferida, la variante de beta-glucosidasa de la invención consiste en la secuencia de aminoácidos SEQ ID NO: 5. Esta SEQ ID NO: 5  
10 corresponde a la beta-glucosidasa nativa de SEQ ID NO: 2, que comprende la sustitución de aminoácido Q211H. Como se muestra en los ejemplos más abajo, la sustitución Q211H reduce la actividad de transglicosilación de la enzima, lo que incrementa la concentración final de azúcares fermentables en un proceso hidrolítico a partir de material celulósico. Dicha  
15 secuencia SEQ ID NO: 5 también se denominará en lo sucesivo preproteína Bgl1Q211H.

En una realización aún más preferida, la variante de beta-glucosidasa de la invención consiste en la secuencia de aminoácidos SEQ ID NO: 4. Esta SEQ ID NO: 4 corresponde a la beta-glucosidasa madura de la SEQ ID NO: 5. Dicha secuencia SEQ ID NO: 4 también se  
20 denominará en lo sucesivo proteína madura de Bgl1Q211H.

La variante de beta-glucosidasa de la invención puede sintetizarse, por ejemplo, pero sin limitaciones, *in vitro*. Por ejemplo, por medio de la síntesis de péptidos en fase sólida o abordajes de ADN recombinante. La variante de beta-glucosidasa de la invención puede  
25 producirse de forma recombinante, incluida su producción como péptido maduro o como una preproteína que incluye un péptido señal.

La preparación de la variante de beta-glucosidasa de la invención se puede realizar por cualquier medio conocido en la técnica, tales como modificación de una secuencia de ADN que codifica una beta-glucosidasa nativa, tal como, por ejemplo, pero sin limitarnos, la SEQ ID NO: 3, que codifica la preproteína de SEQ ID NO: 2, transformación de la secuencia de ADN modificada en una célula huésped adecuada y la expresión de la secuencia de ADN modificada para formar la variante enzimática.  
30

Debido a la degeneración del código genético, varias secuencias de nucleótidos pueden codificar la misma secuencia de aminoácidos. Por tanto, en otro aspecto, la invención proporciona una secuencia de ácido nucleico aislada que codifica la variante de beta-glucosidasa de la invención o la variante beta-glucosidasa producida por el segundo procedimiento de la invención, en lo sucesivo "secuencia de ácido nucleico de la invención",  
5 y la secuencia de ácido nucleico complementaria a la misma.

De acuerdo con la presente invención, una "molécula de ácido nucleico aislada", "secuencia de nucleótidos", "secuencia de ácido nucleico" o "polinucleótido" es una molécula de ácido nucleico (polinucleótido) que se ha eliminado de su medio natural (es decir, que se ha sometido a manipulación humana) y puede incluir ADN, ARN o derivados de ADN o ARN, incluyendo ADNc. La secuencia de nucleótidos de la presente invención puede estar o no química o bioquímicamente modificada, y se puede obtener artificialmente por medio de clonación y procedimientos de selección o mediante secuenciación. La secuencia de ácido nucleico de la invención puede codificar el polipéptido maduro o una preproteína que consiste en un péptido señal unido a la enzima madura que tendrá que procesarse después.  
10  
15

La secuencia de nucleótidos de la presente invención también puede comprender otros elementos, tales como intrones, secuencias no codificantes en los extremos 3' y/o 5', sitios de unión al ribosoma, etc. Esta secuencia de nucleótidos también puede incluir secuencias codificantes para aminoácidos adicionales que son útiles para la purificación o estabilidad del péptido codificado.  
20

En una realización preferida, la secuencia de ácido nucleico de la invención es la SEQ ID NO: 6, que es la secuencia de ácido nucleico que codifica la secuencia de aminoácidos de SEQ ID NO: 5 (preproteína de SEQ ID NO: 4).  
25

La expresión, "secuencia de ácido nucleico complementaria" de una secuencia de ácido nucleico que codifica la variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida por el segundo procedimiento de la invención, hace referencia a la secuencia de ácido nucleico de la hebra complementaria a la que codifica la variante de beta-glucosidasa de la invención, o la variante de beta-glucosidasa producida por el segundo procedimiento de la invención. Se apreciará que un ADN bicatenario que codifica una secuencia de aminoácidos dada comprende un ADN monocatenario y su hebra  
30

complementaria, que tiene una secuencia que es complementaria al ADN monocatenario.

La Tabla 1 siguiente muestra una descripción detallada de algunas de las secuencias mencionadas a lo largo de la presente invención.

<b>Secuencia</b>	<b>DESCRIPCIÓN</b>
SEQ ID NO: 1	Proteína madura de la Bgl1 nativa
SEQ ID NO: 2	Preproteína de la Bgl1 nativa (SEQ ID NO: 1, incluyendo el péptido señal de 19 aminoácidos en el extremo N-terminal de la SEQ ID NO: 1)
SEQ ID NO: 3	Polinucleótido que codifica la SEQ ID NO: 2 (sin intrones)
SEQ ID NO: 4	Proteína madura de Bgl1Q211H
SEQ ID NO: 5	Preproteína Bgl1Q211H (SEQ ID NO: 4, incluyendo el péptido señal de 19 aminoácidos en el extremo N-terminal de la SEQ ID NO: 4)
SEQ ID NO: 6	Polinucleótido que codifica la SEQ ID NO: 5 (sin intrones)

5

Tabla 1. Descripción de algunas de las secuencias mencionadas en la presente invención.

La secuencia de ácido nucleico de la invención se puede incluir en una construcción genética, preferiblemente en un vector de expresión. Dicha construcción genética puede comprender además una o más secuencias reguladoras de la expresión génica, tales como promotores, terminadores etc. Por tanto, en otro aspecto, la invención proporciona una construcción genética que comprende la secuencia de ácido nucleico de la invención o la secuencia de ácido nucleico complementaria a la misma, en lo sucesivo "construcción genética de la invención". En una forma de realización preferida, dicha construcción genética es un vector de expresión.

10

15

La expresión "construcción genética" o "construcción de ácido nucleico" como se usa aquí hace referencia a una unidad funcional necesaria para la transferencia o la expresión de un gen de interés, en el presente documento, la secuencia de ácido nucleico de la invención

como se ha descrito, y secuencias reguladoras, incluyendo, por ejemplo, un promotor, operativamente unidas a la secuencia que codifica la proteína. Se refiere a una molécula de ácido nucleico, mono o bicatenaria, que se encuentra aislada de un gen natural o que se modifica para que contenga segmentos de ácidos nucleicos de un modo que, de otro modo, no existirían en la naturaleza. La expresión construcción de ácido nucleico es sinónima a la expresión "casete de expresión", cuando la construcción de ácido nucleico contiene las secuencias control requeridas para la expresión de la secuencia codificante.

La expresión "vector de expresión", también conocida como "construcción de expresión" o "plásmido", hace referencia a una molécula de ADN, lineal o circular, que comprende la secuencia de ácido nucleico de la invención y que está operativamente unida a segmentos adicionales que proporcionan la transcripción del péptido codificado. Generalmente, un plásmido se usa para introducir un gen específico en una célula diana. Una vez que el vector de expresión está en el interior de la célula, la proteína que está codificada por el gen es producida mediante los complejos ribosómicos de la maquinaria de transcripción y traducción celular. Con frecuencia el plásmido se somete a ingeniería para que contenga secuencias reguladoras que actúan como regiones potenciadoras y promotoras y que conducen a una transcripción eficiente del gen portado en el vector de expresión. El objetivo de un vector de expresión bien diseñado es la producción de grandes cantidades de ARN mensajero estable y, por tanto, de proteínas. Los vectores de expresión son herramientas básicas de biotecnología y de la producción de proteínas, tales como enzimas. El vector de expresión de la invención se introduce en una célula huésped de modo que el vector se mantiene como integrante cromosómico o como vector autoreplicante extracromosómico.

Ejemplos de vectores de expresión son fagos, cósmidos, fagémidos, cromosomas artificiales de levaduras (YAC), cromosomas artificiales bacterianos (BAC), cromosomas artificiales humanos (HAC) o vectores virales, tales como adenovirus, retrovirus o lentivirus.

Las construcciones génicas de la presente invención abarcan un vector de expresión, donde el vector de expresión se puede usar para transformar una célula huésped u hospedadora adecuada para que el huésped pueda expresar la variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida por el segundo procedimiento de la invención. Los procedimientos para la expresión recombinante de proteínas en hongos y otros organismos son bien conocidos en la técnica y se dispone de numerosos vectores de

expresión o se pueden construir usando procedimientos de rutina.

La expresión "secuencias control" se define aquí para incluir todos los componentes que son necesarios o ventajosos para la expresión de la secuencia de ácido nucleico de la presente  
5 invención. Dichas secuencias control incluyen, pero sin limitarse, un líder, una secuencia de poliadenilación, una secuencia propéptido, un promotor, una secuencia de péptido señal y un terminador de la transcripción. Como mínimo, las secuencias control incluyen un promotor y señales de terminación de la transcripción y de la traducción. Las secuencias control se pueden proporcionar con ligadores con el fin de introducir sitios de restricción  
10 específicos que facilitan la unión de las secuencias control con la región codificante de la secuencia de ácido nucleico de la presente invención. La expresión "operativamente unido" indica en el presente documento una configuración en la que una secuencia control se coloca en una posición adecuada respecto a la secuencia de ácido nucleico de la presente invención, de un modo tal que la secuencia control dirige la expresión de la secuencia de  
15 ácido nucleico de la presente invención.

El vector de expresión de la invención puede ser un vector de replicación autónoma, es decir un vector que existe como entidad extracromosómica, cuya replicación es independiente de la replicación del cromosoma, por ejemplo un plásmido, un elemento extracromosómico, un  
20 minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para garantizar la autoreplicación. Como alternativa, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica junto con el(los) cromosoma(s) en el(los) que se ha integrado.

25 Además se puede usar un único vector o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total que se va a introducir en el genoma de la célula huésped, o un transposon.

Los vectores usados en la presente invención contienen, preferiblemente, uno o más  
30 marcadores seleccionables que permitan la fácil selección de las células transformadas, transfectadas, transducidas o similares. Un marcador seleccionable es un producto génico que proporciona resistencia a un biocida o a un virus, a metales pesados, prototrofia a los auxótrofos y similares. Los marcadores seleccionables para usar en una célula huésped de un hongo filamentoso incluyen, pero sin limitarse, amdS (acetamidasa), argB (ornitina

carbamoiltransferasa), bar (fosfinotricina acetiltransferasa), hph (higromicina fosfotransferasa), niaD (nitrato reductasa), pyrG (orotidin-5'-fosfato descarboxilasa), cysC (sulfato adeniltransferasa), y trpC (antranilato sintasa), así como equivalentes de los mismos.

5

Los vectores usados en la presente invención contienen, preferiblemente, uno o más elementos que permiten la integración del vector en el genoma de la célula huésped o la replicación autónoma del vector en la célula con independencia del genoma. Para la integración en el genoma de la célula huésped, el vector puede depender de la secuencia de ácido nucleico de la presente invención o de cualquier otro elemento del vector para la integración en el genoma mediante recombinación homóloga o no homóloga. Como alternativa, el vector puede contener secuencias adicionales de nucleótidos para dirigir la integración mediante recombinación homóloga en el genoma de la célula huésped en una o más localización(es) precisas en el(los) cromosoma(s).

15

Para la replicación autónoma, el vector puede comprender además un origen de replicación que permite que el vector se replique de forma autónoma en la célula huésped en cuestión. El origen de replicación puede ser cualquier replicador plasmídico que participe en la replicación autónoma que funciona en una célula. La expresión "origen de replicación" o "replicador plasmídico" se define aquí como una secuencia de nucleótidos que permite que un plásmido o vector se replique *in vivo*. Ejemplos de orígenes de replicación útiles en una célula fúngica filamentosa son AMAI y ANSI.

En la célula huésped se puede insertar más de una copia de la secuencia de ácido nucleico de la presente invención para aumentar la producción del producto génico. Se puede obtener un incremento del número de copias del polinucleótido mediante integración de al menos una copia adicional de la secuencia en el genoma de la célula huésped o incluyendo un gen marcador seleccionable amplificable con el polinucleótido, donde las células que contienen copias amplificadas del gen marcador seleccionable y, por consiguiente, copias adicionales del polinucleótido, se pueden seleccionar cultivando las células en presencia del agente seleccionable adecuado. Los procedimientos usados para ligar los elementos descritos anteriormente para construir los vectores de expresión recombinante a los que se hace referencia en la presente invención son bien conocidos por un experto en la técnica.

En otro aspecto, la invención proporciona una célula huésped que comprende la construcción génica de la invención, en adelante denominada "célula huésped de la invención". Por tanto, dicha célula huésped expresa la variante de beta-glucosidasa de la invención. La "célula huésped", como se usa aquí, incluye cualquier tipo celular que es susceptible de transformación, transfección, transducción y similar con la construcción génica de la invención. La célula huésped puede ser eucariota, tal como una célula de mamífero, insecto, vegetal o fúngica. En una realización preferida, la célula huésped es una célula de hongo filamentoso. Los hongos filamentosos generalmente se caracterizan por una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. En una realización más preferida, la célula huésped de hongo filamentoso es una célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolyptocladium*, *Trametes*, o *Trichoderma*. En una realización más preferida, la célula huésped de hongo filamentoso es una célula de *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* o *Aspergillus oryzae*. En otra realización más preferida, la célula huésped de hongo filamentoso es una célula de *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium pseudograminearum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, o *Fusarium venenatum*. En otra realización más preferida, la célula huésped de hongo filamentoso es una célula de *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Gibberella zeae*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* o *Trichoderma viride*. En otra realización aún más preferida, la célula huésped de la invención es cualquier cepa de la especie *Myceliophthora thermophila*. En una realización aún más preferida, la célula

huésped de la invención es la cepa C1 de la especie *Myceliophthora thermophila*.

5 Se entenderá que, para las especies mencionadas anteriormente, la invención abarca estados tanto perfectos como imperfectos y otros equivalentes taxonómicos, por ejemplo anamorfos, con independencia del nombre de la especie por el que se conocen. Los expertos en la técnica reconocerán fácilmente la identidad de los equivalentes adecuados. Por ejemplo, *Myceliophthora thermophila* es equivalente a *Chrysosporium lucknowense*.

10 La variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida por el segundo procedimiento de la invención tiene una actividad de transglicosilación reducida, por lo que su uso en una composición enzimática para la etapa de hidrólisis del material celulósico en azúcares fermentables en los procesos para la producción de un bioproducto, preferiblemente etanol, es interesante para mejorar la actividad de toda la composición enzimática.

15 Por tanto, en otro aspecto de la invención se proporciona una composición enzimática que comprende la variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida por el segundo procedimiento de la invención, de ahora en adelante conocida como "composición enzimática de la invención". En una realización preferida, la composición enzimática de la invención comprende además otras enzimas celulolíticas.

25 Se debe entender que la variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida por el segundo procedimiento de la invención se pueden combinar con una o más de las enzimas celulolíticas descritas en el presente documento o con cualquier otra enzima disponible y adecuada para producir una composición multienzimática. Uno o más componentes de la composición multienzimática (aparte de las enzimas descritas en la presente invención) se pueden obtener o derivar de una fuente microbiana, vegetal o de otro tipo o combinación de las mismas, y contendrán enzimas capaces de degradar el material celulósico.

30 La expresión "enzimas celulolíticas", también conocidas como "celulasas" hace referencia a una categoría de enzimas capaces de hidrolizar la celulosa (enlaces  $\beta$ -1,4-glucano o  $\beta$ -D-glucosídico) en oligosacáridos más cortos, celobiosa y/o glucosa. Ejemplos de enzimas celulolíticas son, pero sin limitarnos, endoglucanasas, beta-glucosidasas, celobiohidrolasas



o beta-xilosidasas. Por tanto, en una realización más preferida, estas enzimas celulolíticas se seleccionan de la lista que consiste en: endoglucanasas, beta-glucosidasas, celobiohidrolasas, beta-xilosidasas o cualquier combinación de las mismas. Estas enzimas celulolíticas pueden derivar de la célula huésped de la invención u otros microorganismos  
5 productores de enzimas celulolíticas distintos a la célula huésped de la invención. Asimismo, se pueden producir de forma natural o recombinante.

El término "endoglucanasa" o "EG" hace referencia a un grupo de enzimas celulasas clasificadas como E.C. 3.2.1.4. Estas enzimas hidrolizan los enlaces  $\beta$ -1,4 glucosídicos de la  
10 celulosa.

El término "celobiohidrolasa" hace referencia a una proteína que cataliza la hidrólisis de la celulosa en celobiosa mediante una actividad exoglucanasa, liberando secuencialmente moléculas de celobiosa a partir de los extremos reductores o no reductores de la celulosa o  
15 los celooligosacáridos.

El término " $\beta$ -xilosidasa" se refiere a una proteína que hidroliza los 1,4- $\beta$ -D-xilooligómeros cortos en xilosa.

20 En una realización preferida, la composición enzimática de la invención comprende además la célula huésped de la invención.

La composición de la invención se puede preparar de acuerdo con procedimientos conocidos en la técnica y puede estar en forma de un líquido o una composición seca. Por  
25 ejemplo, la composición puede estar en forma de un granulado o un microgranulado. Las enzimas que se van a incluir en la composición se pueden estabilizar de acuerdo con procedimientos conocidos en la técnica.

La variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida  
30 mediante el segundo procedimiento de la invención, así como la célula huésped o la composición de la presente invención se pueden usar en la producción de monosacáridos, disacáridos y polisacáridos como reservas químicas o de fermentación a partir de material celulósico para la producción de etanol, plásticos u otros productos o intermedios.

La célula huésped de la presente invención se puede usar como fuente de la variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida mediante el segundo procedimiento de la invención y otras enzimas celulolíticas, en un proceso de fermentación a partir del material celulósico.

5

Por tanto, otro aspecto de la invención hace referencia a una composición enzimática obtenida mediante la célula huésped de la invención.

La degradación o hidrólisis del material celulósico en azúcares fermentables, proceso también conocido como "sacarificación", por medio de la variante de beta-glucosidasa de la invención, la variante de beta-glucosidasa producida mediante el segundo procedimiento de la invención, la célula huésped de la invención o la composición de la invención puede acompañarse después de un proceso de fermentación en el que los azúcares fermentables obtenidos se usan con el fin de obtener finalmente un bioproducto tal como bioetanol.

15

Por tanto, en otro aspecto, la presente invención se refiere a un procedimiento para producir azúcares fermentables, en lo sucesivo "tercer procedimiento de la invención", que comprende:

20

- a) Incubar el material celulósico con la variante de beta-glucosidasa de la invención, la variante de beta-glucosidasa producida mediante el segundo procedimiento de la invención, la célula huésped de la invención o la composición enzimática de la invención, y
- b) Recuperar el azúcar fermentable obtenido tras la incubación en la etapa (a).

25

La expresión "azúcar fermentable", como se usa en el presente documento, se refiere a azúcares simples, tales como glucosa, xilosa, arabinosa, galactosa, manosa, ramnosa, sacarosa o fructosa.

30

En otro aspecto, la presente invención se refiere a un procedimiento de producir un bioproducto, en lo sucesivo "cuarto procedimiento de la invención", que comprende:

- a) Incubar el material celulósico con la variante de beta-glucosidasa de la invención, la variante de beta-glucosidasa producida mediante el segundo procedimiento de la

invención, la célula huésped de la invención o la composición enzimática de la invención,

b) Fermentar los azúcares fermentables obtenidos tras la incubación de la etapa (a) con al menos un microorganismo fermentador, y

5 c) Recuperar el bioproducto obtenido tras la fermentación en la etapa (b).

La expresión "material celulósico" significa la fracción biodegradable de productos, residuos y desechos de origen biológico de la agricultura (incluidos vegetales, tales como residuos de cosechas y sustancias animales), forestales (como recursos madereros) e industrias  
10 relacionadas, incluidas pescaderías y acuicultura, así como la fracción biodegradable de residuos industriales y municipales, tales como residuos sólidos municipales o papeleras. En una realización preferida, el material celulósico es paja o fracción orgánica de residuos sólidos municipales. En una realización más preferida, el material celulósico es biomasa vegetal, más preferiblemente seleccionada de la lista que consiste en: biomasa rica en  
15 azúcares fermentables, tales como caña de azúcar, biomasa de almidón, por ejemplo, grano de trigo o paja de maíz. Aún más preferentemente, la biomasa vegetal es grano de cereales, tales como almidón, trigo, cebada o mezclas de los mismos.

En algunas realizaciones, el tercero y/o cuarto procedimiento de la invención comprende,  
20 preferiblemente, un proceso de pretratamiento antes de la etapa (a). En general, un proceso de pretratamiento dará como resultado que los componentes del material celulósico sean más accesibles para las etapas posteriores o sean más digeribles por las enzimas tras el tratamiento en ausencia de hidrólisis. El pretratamiento puede ser un pretratamiento químico, físico o biológico, o cualquier mezcla de los mismos.

25 El término "recuperación", como se usa aquí, se refiere a la recuperación de los azúcares fermentables obtenidos tras la incubación en la etapa (a) del tercer procedimiento de la invención o el bioproducto obtenido tras la fermentación de la etapa (b) del cuarto procedimiento de la invención. La recuperación se puede producir por cualquier  
30 procedimiento conocido en la técnica, incluidos mecánicos o manuales.

El término " fermentación", como se usa aquí, se refiere a un proceso de transformación biológica causada por la actividad de algunos microorganismos en el que azúcares tales como glucosa, fructosa y sacarosa se convierten en etanol. Por tanto, los microorganismos

usados son microorganismos fermentadores que tienen una capacidad de fermentación, tales como levaduras, preferiblemente *Saccharomyces cerevisiae*.

5 En otra realización preferida, las etapas (a) y (b) del cuarto procedimiento de la invención se pueden realizar simultáneamente.

El término "bioproducto" o "productos biológicos" se refiere a los materiales, químicos y derivados de energía de recursos biológicos renovables. Ejemplos de estos bioproductos son, pero sin limitarnos, compuestos de hidrocarburos en sus diferentes formas, tales como  
10 hidrocarburos alifáticos (saturados, insaturados, cíclicos) o aromáticos, como alcanos, alquenos, alquinos, formas cíclicas de estos compuestos o hidrocarburos aromáticos; sustancias oxigenadas como alcoholes, éteres, aldehídos, cetonas o ácidos carboxílicos; sustancias nitrogenadas como aminas, amidas, compuestos de nitrógeno o nitrilos; sustancias halogenadas como haluros. El término "bioproductos" también incluye cualquier  
15 combinación de los compuestos descritos arriba, compuestos que además derivan de los compuestos descritos arriba mediante cualquier tipo de tratamiento físico, químico o biológico, polímeros de los compuestos descritos arriba, compuestos descritos anteriormente sustituidos por cualquier grupo o elemento funcional en uno o más de sus enlaces y formas ramificadas de los compuestos descritos arriba.

20 El etanol se puede producir mediante degradación enzimática del material celulósico y la conversión de los sacáridos liberados en etanol. Este tipo de etanol a menudo se denomina bioetanol. Se puede usar como aditivo de combustible o como expansor en mezclas de menos de 1% y de hasta 100% (un sustituto del combustible).

25 Por tanto, en una forma de realización más preferida, el bioproducto es biocombustible. El término "biocombustible", como se usa aquí, hace referencia a un hidrocarburo, o una mezcla de los mismos, que se puede usar como combustible y se obtiene usando material celulósico fermentable como material de partida. Ejemplos de biocombustibles incluyen,  
30 pero sin limitarse, etanol o bioetanol y biodiésel. En una forma de realización preferida, el biocombustible es bioetanol.

El término "bioetanol" o "etanol" hace referencia a un alcohol realizado mediante fermentación, principalmente a partir de material celulósico fermentable tal como hidratos de

carbono producido mediante la variante de beta-glucosidasa de la invención o la variante de beta-glucosidasa producida mediante el segundo procedimiento de la invención, o cultivos de almidón tales como maíz o caña de azúcar.

5 A menos que se defina otra cosa, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el que les daría un experto en la técnica a la que esta invención pertenece. En la práctica de la presente invención se pueden usar procedimientos y materiales similares o equivalentes a los descritos en el presente documento. A lo largo de la descripción y las reivindicaciones, con la palabra “comprende” y  
 10 sus variaciones no se pretende excluir otras características técnicas, aditivos, componentes o etapas. Otros objetos, ventajas y características adicionales de la invención serán obvias para los expertos en la técnica a partir del análisis de la descripción o se pueden aprender mediante la práctica de la invención. Los ejemplos y dibujos siguientes y el listado de secuencias se proporcionan a modo de ilustración y no se pretende que sean limitantes de  
 15 la presente invención.

#### **BREVE DESCRIPCIÓN DE LAS FIGURAS**

**Fig. 1. Vector de expresión con *Tcbh1* como secuencia de terminación y *pyr5* como  
 20 marcador de selección.** *SacI* y *NotI* fueron los sitios de restricción elegidos para la clonación del fragmento *Pcbh1-bgl1*.

**Fig. 2. Plásmido de expresión del ADNc de *bgl1* de *Myceliophthora thermophila* C1.**

25 **Fig. 3. Actividad relativa beta-glucosidasa de 31L2D y otros transformantes** analizada usando pNGP como sustrato en presencia y en ausencia de celobiosa.

**Fig. 4. Plásmido de expresión de *bgl1Q211H* amplificado de la cepa 31L2D.**

30 **Fig. 5. SDS-PAGE de las enzimas beta-glucosidasa purificadas.** Se cargaron 20 µg de proteína en cada carril. Carril 1: Marcador; Carril 2: Bgl1; Carril 3: proteína madura de Bgl1Q211H.

**Fig. 6. Ensayo de transglicosilación indirecta de la enzima Bgl1 y la proteína madura**

de **Bgl1Q211H**. Todas las mediciones se analizaron por triplicado y las barras de error corresponden a la desviación estándar.

**Fig. 7. Estudios térmicos usando la enzima Bgl1 y la proteína madura de Bgl1Q211H.**

5 A. Determinación de la temperatura de desnaturalización. B. Resistencia térmica en condiciones de hidrólisis de la biomasa. Todas las mediciones se analizaron por triplicado y las barras de error corresponden a la desviación estándar.

**EJEMPLOS**

10

**Ejemplo 1. Mutagénesis de *bgl1*. Construcción de un vector de expresión, mutagénesis, amplificación de los bancos con mutaciones en *bgl1*.**

15 Se ha descrito a *M. thermophila* C1 como un sistema de transformación de buena calidad para expresar y secretar proteínas y polipéptidos heterólogos. El gen *bgl1* de la beta-glucosidasa de *M. thermophila* C1 fue el elegido para mejorar su calidad enzimática en la presente invención. Esta enzima muestra un elevado perfil de transglicosilación, de modo que el objetivo de la mutagénesis era reducir esta actividad de transglicosilación sin afectar a la actividad hidrolítica de la beta-glucosidasa *per se*.

20

La secuencia de ADNc de *bgl1* se sintetizó *in vitro* tras optimización, para eliminar sitios de reconocimiento para las enzimas de restricción más habituales sin alterar la secuencia de aminoácidos. La secuencia de nucleótidos de ADNc de *bgl1* y la secuencia de aminoácidos deducida se muestran en la SEQ ID NO: 3 y la SEQ ID NO: 2, respectivamente. La  
25 secuencia codificante tiene una longitud de 2616 incluido el codón de terminación. La proteína codificada por esta secuencia tiene una longitud de 871 aminoácidos con una masa molecular prevista de 95 KDa y un punto isoeléctrico de 5.10. Usando el programa Signal IP (Petersen *et al.*, 2011, *Nature Methods*, 8:785-786), se predijo un péptido señal de 19 residuos. La proteína madura prevista (SEQ ID NO: 1) contiene 852 aminoácidos con una  
30 masa molecular prevista de 93 KDa y un punto isoeléctrico de 5.04.

El gen *bgl1* se sintetizó *in vitro* junto con el promotor del gen de la celobiohidrolasa 1 (*Pcbh1*) de *M. thermophila* C1. La secuencia de *Pcbh1* incluye una región de 1796 pb aguas arriba del gen de la celobiohidrolasa 1 de *M. thermophila* C1 (*cbh1*, número de registro en

NCBI XP\_003660789.1). Este fragmento (*Pcbh1-bgl1*) se sintetizó *in vitro*, incluyendo la secuencia de las enzimas de restricción *SacI* y *NotI* en los extremos (*SacI* en el extremo 5' y *NotI* en el extremo 3') para clonarse en un vector de expresión denominado pBASE1. El vector de expresión pBASE1 también contenía la secuencia terminadora del gen de la celobiohidrolasa 1 de *Myceliophthora thermophila* C1 (*Tcbh1*, correspondiente a una región de 1014 pb aguas abajo de *cbh1*) y el gen *pyr5* (número de acceso en NCBI XP\_003660657.1) de la misma cepa como marcador de selección. El gen *pyr5* codifica una orotato-fosforibosil transferasa funcional y su vector de expresión permite la complementación de la auxotrofia de uridina en la correspondiente cepa huésped auxotrófica, *M. thermophila* C1 *pyr5*-. El vector de expresión pBASE1 se muestra en la figura 1.

El fragmento *Pcbh1-bgl1* se digirió con las enzimas de restricción *SacI* y *NotI* y se clonó en pBASE1, digerido previamente con las mismas enzimas de restricción. El vector de expresión pBASE1 y el casete *Pcbh1-bgl1* se ligaron y el producto de la unión se transformó en células electrocompetentes de *Escherichia coli* XL1Blue MRF siguiendo el protocolo proporcionado por el fabricante (Agilent Technologies Inc.). El plásmido recombinante obtenido se denominó pABC334 y se muestra en la Figura 2.

El gen *bgl1* clonado en pABC344 se sometió a mutagénesis al azar mediante amplificación con PCR usando el kit de mutagénesis GeneMorpholl EZClone Domain Mutagenesis Kit (Agilent Technologies Inc.). La amplificación mutagénica se realizó usando los oligonucleótidos 1 y 2 que se muestran a continuación. El péptido señal y el codón de terminación se excluyeron de la amplificación mutagénica. Por tanto, estos oligonucleótidos amplifican un fragmento de 2553 pb correspondiente a la secuencia de *bgl1* sin péptido señal ni codón de terminación.

Oligonucleótido 1 (SEQ ID NO: 7):

5'-GCTGACAATCATCGTCAGGTTCCACCAGAAGCCCCTCGCGA-3'

30

Oligonucleótido 2 (SEQ ID NO: 8):

5'-AGGAAGCTCAATCTTGAGATCCAACCTCCGGCTGCTCCTC-3'

El sistema GeneMorpholl EZClone Domain Mutagenesis permite diferentes tasas de

mutación dependiendo de la cantidad de ADN diana y los ciclos de amplificación usados durante el proceso. Con estas premisas se generaron dos bancos mutantes diferentes: el primero con una frecuencia de mutación entre 0-1 mutaciones/kb y el segundo con una frecuencia de mutación entre 1-4,5 mutaciones/kb. La cantidad diana inicial fue 2 µg y 0,8 µg de pABC344 respectivamente para cada banco. Las condiciones para la reacción de amplificación fueron un ciclo de 95 °C durante 2 minutos; y 30 ciclos cada uno a 95 °C durante 30 segundos, 57 °C durante 30 segundos y 72 °C durante 3 minutos. El bloque de calor se mantuvo a 72 °C durante 10 minutos, seguido de un ciclo de retención a 12 °C. Los productos de PCR correspondiente a versiones mutadas de *bgl1* se purificaron en gel de agarosa con un kit de extracción de gel QIAquick (Qiagen) y se usaron como megacebadores en una segunda PCR para amplificar el pABC344 completo usando las condiciones siguientes: un ciclo a 95°C durante 1 minuto y 25 ciclos de 95 °C durante 50 segundos, 60 °C durante 50 segundos y 68 °C durante 10 minutos. Se realizó una digestión de las reacciones de amplificación con *DpnI* (10U/µL) durante 2 horas a 37 °C para eliminar el plásmido de expresión parental pABC344 usado como diana, ya que *DpnI* solo reconoce ADN metilado. Por tanto, solo los plásmidos amplificados durante esta segunda reacción de PCR permanecen tras la digestión con *DpnI*.

Ambos bancos de mutaciones se transformaron en células ultracompetentes *Escherichia coli* XL-10 Gold siguiendo el protocolo proporcionado por el fabricante (Agilent Technologies Inc.) y el ADN del plásmido a partir de un total de 320.000 colonias transformadas con ambos bancos de mutantes se purificó usando el kit Plasmid Maxi (Omega bio-tek, Inc).

**Ejemplo 2. Transformación de los bancos de mutantes de *bgl1* en *Myceliophthora thermophila* C1 y selección de una versión mejorada de *bgl1*.**

El ADN plasmídico de los bancos que contenían las diferentes versiones mutadas de *bgl1* se transformaron en la cepa huésped auxótrofa *M. thermophila* C1 *pyr5*<sup>-</sup> (Verdoes *et al.*, 2007, *Ind. Biotechnol.*, 3 (1)), usada previamente en otras selecciones de alto rendimiento en *M. thermophila*. El ADN se introdujo en la cepa huésped usando un método de transformación en protoplastos (US7399627B2). Los transformantes se sembraron en placas de agar sin suplemento de uridina. Tras 5 días de incubación a 35 °C se analizaron los transformantes prototróficos resultantes (que expresan el gen *pyr5*). Los transformantes obtenidos se inocularon en cultivos de placas de microtitulación (PMT) de 96 pocillos para llevar a cabo la



detección selectiva de alto rendimiento o *high throughput screening* (US7794962B2).

El objetivo de la selección o *screening* era identificar las versiones mutadas de *bgl1* con baja actividad de transglicosilación. Por tanto, se estableció un ensayo indirecto para estimar la capacidad de transglicosilación de una enzima BGL (beta-glucosidasa). La actividad hidrolítica en pNGP (p-nitro-fenil-glucopiranosido) se midió en presencia y ausencia de celobiosa (5,5, mM). La celobiosa se ha descrito como sustrato preferencial de las BGL con mecanismo de retención o “retainining” en su reacción de transglicosilación. Por tanto, en presencia de celobiosa, la actividad de BGL sobre pNGP muestra una reducción debido a la actividad de transglicosilación. El porcentaje de capacidad hidrolítica, medido mediante el p-nitrofenol liberado (y el consiguiente incremento de la  $A_{410}$ ) en presencia del sustrato celobiosa altamente susceptible a la transglicosilación, se usó como indicación de una baja capacidad de transglicosilación. La actividad beta-glucosidasa se mide en unidades por litro de cultivo (U/l). Una unidad de actividad de hidrólisis de pNGP se definió como la cantidad de enzima equivalente a la liberación de 1  $\mu\text{mol}$  de p-nitrofenol por minuto.

La actividad de beta-glucosidasa de 5  $\mu\text{l}$  de sobrenadantes de cada transformante se analizó con 100 mg/l de pNGP (Sigma N7006) durante 10 minutos a 50 °C en un volumen final de 100  $\mu\text{l}$ . La reacción se detuvo añadiendo a las mezclas de reacción 100  $\mu\text{l}$  de carbonato sódico 1M. Cada reacción también se llevó a cabo en presencia de celobiosa (5,5 mM), La capacidad hidrolítica se midió mediante la liberación de p-nitrofenol en presencia y en ausencia de celobiosa como indicación de baja capacidad de transglicosilación.

Para seleccionar los mutantes de *bgl1* que muestran actividad de transglicosilación reducida, se estableció que una baja actividad de transglicosilación correspondía a la disminución de su capacidad hidrolítica sobre pNGP entre 30-70% cuando había celobiosa en la reacción. De entre los 6.900 transformantes analizados, uno de ellos, denominado 31L2D, mostró una capacidad hidrolítica sobre pNGP del 62% en presencia de celobiosa. Todos los transformantes con la actividad deseada se confirmaron en una segunda ronda de ensayos en PMT. La producción de proteínas de los mejores transformantes, incluido 31L2D, también se realizó a escala de producción en matraz (Verdoes et al., 2007, *Ind. Biotechnol.* 3(1); Visser et al., 2011, *Ind. Biotechnol.*, 7 (3)) y se confirmó que el transformante 31L2D mostraba el mejor perfil de acuerdo con las condiciones de *screening* (62% de actividad beta-glucosidasa hidrolítica en presencia de celobiosa). La actividad de

beta-glucosidasa de la cepa 31L2D seleccionada y otros transformantes analizados en presencia y en ausencia de celobiosa se muestra en la Figura 3.

Para determinar la secuencia del gen *bgl1* expresado en 31L2D, se amplificó un fragmento de ADN que contenía *Pcbh1-bgl1* a partir del ADN genómico usando los oligonucleótidos 3 y 4.

Oligonucleótido 3 (SEQ ID NO: 9): El sitio de restricción *SacI* está subrayado.

5'-CGAGGAGCTCCTTACAAAAAAGGTATCC-3'

10

Oligonucleótido 4 (SEQ ID NO: 10): Los sitios de restricción *Bam*HI, *Sma*I y *Pst*I están subrayados. El codón de terminación está en recuadro.

5'TTCCTGCAGCCCGGGGGATCC[TCAGGAAGCTCAATCTTGAGATCCAACCTTCC-3'

15 El oligonucleótido 3 incluye el sitio de restricción *SacI* e hibrida en el extremo 3' de *Pcbh1*. El oligonucleótido 4 hibrida en el extremo 3' de *bgl1* e incluye el codón de terminación de *bgl1* y los sitios de restricción *Bam*HI, *Sma*I y *Pst*I para clonar en el plásmido de expresión pBASE1. Los oligonucleótidos 3 y 4 se usaron para amplificar el fragmento *Pcbh1-bgl1* usando ADN genómico de la cepa 31L2D como diana (obtenido usando el kit DNeasy Plant Mini Kit de Qiagen) con la ADN polimerasa iProof High-Fidelity (BioRad) y se programaron durante un ciclo a 98 °C durante 2 minutos y 30 ciclos de 98 °C durante 10 segundos, 72 °C durante 90 segundos y 72 °C durante 10 minutos. El fragmento de ADN amplificado se digirió con las enzimas de restricción *SacI* y *Bam*HI y se clonó en pABC344 digerido previamente con las mismas enzimas de restricción. La mezcla de ligación se transformó en células electrocompetentes de *Escherichia coli* XL1Blue MRF siguiendo el protocolo proporcionado por el fabricante (Stratagene). El plásmido recombinante se denominó pABC410 y se muestra en la Figura 4.

20

25

El gen *bgl1* de pABC410 se secuenció. El *bgl1* mutado mostró una mutación: la guanina de la posición 633 de la secuencia de nucleótidos de *bgl1* nativo SEQ ID NO: 3 se había mutado a timina, dando, por tanto, una secuencia de nucleótidos de SEQ ID NO: 6, que codificaba una preproteína (SEQ ID NO: 5, preproteína Bgl1Q211H), en la que la glutamina (Q) en el residuo 211 de la preproteína nativa SEQ ID NO: 2 se había intercambiado por histidina (H), de modo que 31L2D expresaba una versión madura mutada (SEQ ID NO: 4)

30

de Bgl1 madura en la que la glutamina (Q) en el residuo 192 de la proteína nativa madura de SEQ ID NO: 1 se había intercambiado por histidina (H), denominada proteína madura de Bgl1Q211H. La secuencia de nucleótidos que codifica la preproteína Bgl1Q211H y la secuencia de aminoácidos de la proteína madura se muestran en la SEQ ID NO: 6 y la SEQ ID NO: 4, respectivamente. La secuencia codificante tiene una longitud de 2616 incluido el codón de terminación. La preproteína predicha codificada tiene 871 aminoácidos (SEQ ID NO: 5) con una masa molecular prevista de 95 KDa y un punto isoeléctrico de 5.14. El plásmido pABC410 se transformó en *M. thermophila* C1 con el fin de confirmar el fenotipo de baja actividad de transglicosilación observado en la cepa 31L2D. La transformación, detección selectiva y medición de la actividad BGL se llevó a cabo como se ha descrito anteriormente en la presente invención. Todos los transformantes que expresaban Bgl1Q211H de pABC410 mostraron la misma actividad de transglicosilación baja que la cepa 31L2D.

### 15 **Ejemplo 3. Análisis comparativo de Bgl1 de la cepa C1 de *Myceliophthora thermophila* y la Bgl1Q211H mutante.**

#### Producción de las enzimas Bgl1 y Bgl1Q211H.

20 La producción de las mezclas enzimáticas se realizó como se describe en Verdoes *et al.*, 2007, Ind. Biotechnol. 3 (1), y Visser *et al.*, 2011, Ind. Biotechnol., 7 (3), usando la plataforma de expresión para enzimas industriales basada en *M. thermophila* C1 desarrollada por Dyadic Netherlands.

#### 25 Purificación de las enzimas Bgl1 y Bgl1Q211H.

Tanto la enzima nativa Bgl1 (SEQ ID NO: 1) como la proteína madura de Bgl1Q211H (SEQ ID NO: 4) se purificaron usando una cromatografía de intercambio iónico. Las muestras se prepararon centrifugando a 4.000xg durante 10 minutos producciones en matraz. Los sedimentos se descartaron y los sobrenadantes se filtraron a través de un filtro de celulosa en jeringa estéril de 0.45 µm (VWR). Las preparaciones enzimáticas resultantes (muestras de 15 ml) se desalaron después en una columna HiPrep Desalting Sephadex G-25 (GE Healthcare) en tampón Tris-HCl 100 mM, a pH 7,0 a un caudal de 15 ml min<sup>-1</sup>.

Posteriormente se aplicó una muestra que contenía 150 mg de proteína total en una columna HiLoad 26/10 Q-Sepharose HP (GE Helthcare) equilibrada con el mismo tampón. La columna se lavó con el tampón de partida y las proteínas unidas a la misma se eluyeron con un gradiente de NaCl a un caudal de  $5 \text{ mL min}^{-1}$  usando un perfil de elución lineal de 0 a 30 %. En total se realizaron nueve ciclos de purificación para cada enzima. De este modo se obtuvieron puras 109 y 116 mg de las enzimas Bgl1 nativa y mutante Q211H, respectivamente.

Las BGLs purificadas se cargaron en geles de electroforesis SDS-PAGE con una concentración del 7,5% de acrilamida, con el fin de comprobar la homogeneidad de las muestras (Fig. 5). De acuerdo con la secuencia de aminoácidos de Bgl1, cabe esperar que esta enzima tenga un tamaño molecular de 93kDa. No obstante, las glicósido hidrolasas fúngicas y otras enzimas pertenecientes a diferentes familias de proteínas a menudo están glicosiladas, portando tanto glicanos O-unidos como N-unidos. De hecho, la glicosilación es la modificación postraducciona l más frecuente en estas proteínas. Se detectó una única banda de 116 kDa para ambas muestras, lo que indica que se han obtenido enzimas electroforéticamente puras y que las versiones glicosiladas de las enzimas se han purificado con éxito.

#### Caracterización cinética de la actividad hidrolítica de la proteína Bgl1 nativa y madura mutante de Bgl1Q211H.

Las beta-glucosidasas (BGLs, EC 3.2.1.21) catalizan la transferencia de grupos glicosil entre nucleófilos de oxígeno sobre diferentes tipos de sustratos. Las BGLs con un mecanismo de retención, como es el caso de Bgl1, tienen actividad hidrolítica reducida a concentraciones elevadas de sustrato o de producto y esto depende al menos de la existencia de dos fenómenos diferentes: inhibición y transglicosilación. La caracterización cinética de Bgl1 y de la proteína madura mutante de Bgl1Q211H se realizó en base a su capacidad hidrolítica.

La actividad hidrolítica de BGL se determinó usando p-nitrofenil-beta-D-glucopiranosido (pNGP, Sigma) como sustrato. Para este ensayo de pNGP, las mezclas de la reacción enzimática (1 ml volumen final) que contienen 100  $\mu\text{mol}$  de tampón acetato sódico (pH 5,0), 100  $\mu\text{g}$  de pNGP (0,33  $\mu\text{mol}$ ) y una cantidad adecuada de la enzima purificada, se incubaron a 50 °C durante 10 minutos. La cantidad de p-nitrofenol liberada se midió a  $A_{410}$  ( $\epsilon_{410} = 15,2$

mM<sup>-1</sup> cm<sup>-1</sup>) tras la adición de 100 µg de carbonato sódico a las mezclas de reacción. Una unidad de actividad de hidrólisis sobre pNGP se definió como la cantidad de enzima equivalente a la liberación de 1 µmol de p-nitrofenol por minuto. Cuando se estudia el efecto de un compuesto soluble (glucosa o xilosa) sobre la actividad de Bgl, a la mezcla de reacción enzimática se añaden cantidades adecuadas del azúcar correspondiente.

Las BGLs purificadas se caracterizaron en términos de actividad enzimática ( $V_{max}$ ), afinidad por el sustrato ( $K_m$ ), resistencia térmica y efecto sobre la actividad enzimática de la glucosa, que es el compuesto más abundante al final de los procesos de hidrólisis enzimática de biomasa.  $V_{max}$  y  $K_m$  se calcularon usando un gráfico de Hanes-Woolf. La cinética de inhibición se calculó usando un gráfico de Eadie-Hofstee. Los resultados se muestran en la tabla 2.

Enzima	$K_m$ (mM)	$V_{max}$ (U mg <sup>-1</sup> )	$K_i$ , Glucosa (mM)
Bgl1	0,66	20,02	3,83
Proteína madura de Bgl1Q211H	0,25	25,36	1,38

Tabla 2. Caracterización cinética de las beta-glucosidasas purificadas.

No se obtuvieron indicios de un impacto directo de la afinidad por el sustrato ( $K_m$ ) sobre el rendimiento de hidrólisis enzimática de la correspondiente mezcla enzimática que contiene BGL. No obstante, se observó un impacto razonable de la actividad enzimática ( $V_{max}$ ) sobre la reducción de la dosis. La Bgl1Q211H mutante se seleccionó como una enzima más activa sobre pNGP y esto se confirmó usando enzimas purificadas: proteína madura de Bgl1Q211H mostró un incremento del 25% de la actividad específica de beta-glucosidasa en comparación con la Bgl1 purificada (25,36 Vs. 20,02 U mg prot<sup>-1</sup>).

La glucosa es un potente inhibidor competitivo de la mayoría de las BGLs. Esto también fue cierto para la enzima nativa y para la proteína madura de Bgl1Q211H purificada. Los ensayos en presencia de diferentes concentraciones de glucosa se ajustaban a una inhibición de patrón competitivo por la glucosa usando un gráfico de Eadie-Hofstee. Las  $K_i$  calculadas fueron de 3,83 y 1,38 mM<sup>-1</sup> para la enzima Bgl1 y la proteína madura de

Bgl1Q211H, indicando una potente inhibición competitiva por la glucosa para ambas enzimas.

5 Estudios de transglicosilación usando la enzima Bgl1 nativa y la proteína madura de Bgl1Q211H.

10 En Bgl1, el mecanismo de retención forma inicialmente un complejo glicosil covalente con la enzima. Este intermedio se descompone en la segunda etapa de la reacción, lo que se puede producir mediante una molécula de agua (actividad hidrolítica) o mediante un azúcar aceptor (actividad de transglicosilación). En la reacción de transglicosilación más sencilla, el aceptor es el propio sustrato. No obstante, la celobiosa se ha descrito como sustrato preferencial de las BGLs con mecanismo de retención en su reacción de transglicosilación.

15 Por tanto, se estableció un ensayo indirecto para estimar la capacidad de transglicosilación de una enzima BGL. La actividad hidrolítica sobre el pNGP se midió en presencia de concentraciones crecientes de celobiosa. El porcentaje de capacidad hidrolítica, medido mediante el p-nitrofenol liberado (y el consiguiente incremento de la  $A_{410}$ ) en presencia del sustrato celobiosa altamente susceptible a la transglicosilación, se usó como indicación de una baja capacidad de transglicosilación. Este ensayo se llevó a cabo para la Bgl1 nativa y  
20 para la proteína madura mutante de Bgl1Q211H. Los resultados se muestran en la Figura 6. Se determinó una mayor capacidad de transglicosilación para la enzima Bgl1 nativa.

25 Con el fin de comprobar si estos ensayos indirectos se correlacionaban con la capacidad de transglicosilación real, se llevaron a cabo ensayos de medición de la formación real de oligómeros. Las mezclas de reacción enzimáticas (volumen final de 1 ml) que contienen celobiosa 138 mM (50 g/L), tampón de acetato sódico 100 mM (pH 5,0) y 0,1 U (actividad de la celobiosa) de la Bgl1 nativa o la proteína madura mutante de Bgl1Q211H se incubaron a 50 °C durante 24 horas. Las reacciones se detuvieron tras 24 horas de incubación mediante la adición de 100 µg de carbonato a las mezclas de reacción. Después, las mezclas se  
30 analizaron mediante HPLC (Agilent Technologies, 1200 Series) usando un detector del índice de refracción (DIR) y una columna Aminex HPX-87 H. Los tiempos de retención y el pico de las áreas obtenidas debido a la formación de cada azúcar se muestran en la Tabla 3.

Azúcar	Tiempo de retención (min)	Área relativa (%)	
		Bgl1	Proteína madura de Bgl1Q211H
Glucosa	9,23	59,13	78,81
Celobiosa	7,49	38,36	21,18
Celotriosa	6,79	2,50	0,00

Tabla 3. Tiempos de retención y áreas relativas para los azúcares detectados tras el ensayo de transglicosilación. El área relativa representa el porcentaje de cada azúcar en la concentración total en la mezcla de reacción tras el ensayo de transglicosilación.

5

En este ensayo de transglicosilación, la glucosa es el producto de la capacidad hidrolítica de las BGLs sobre el sustrato real, celobiosa. La celotriosa solo se puede formar dentro de una reacción de transglicosilación por la cual una molécula de celobiosa actúa como aceptor para la glucosa retenida en el sitio activo de la enzima BGL. La celobiosa puede ser el resultado de (1) sustrato no hidrolizado que permanece en la mezcla de reacción o (2) el producto de una reacción de transglicosilación, siendo la glucosa tanto el aceptor como el donante. Dado que tras la incubación de la proteína madura mutante de Bgl1Q211H con celobiosa no se determinó celotriosa y se determinó una concentración significativamente menor de celobiosa, se puede concluir que esta enzima mutante tiene una capacidad de transglicosilación reducida y una actividad hidrolítica mejorada a concentraciones elevadas de celobiosa.

15

Considerando estos resultados, se ha demostrado que la enzima madura mutante de Bgl1Q211H tiene una capacidad de transglicosilación menor en comparación con Bgl1 *de M. thermophila* C1. En el ensayo acoplado a HPLC directo, se demostró la anulación completa de la formación de oligómeros para la proteína madura de Bgl1Q211H, y fue detectable para Bgl1. Considerando el método indirecto para la estimación de la transglicosilación se estimó un intervalo de reducción del 75% de la capacidad de transglicosilación.

20

La reducción de la capacidad de transglicosilación de la proteína madura de la enzima Bgl1Q211H parece ser responsable de una mejora de esta enzima en términos de eficiencia

25

en la hidrólisis enzimática de la biomasa, proceso en el cual hay una elevada concentración de sustratos de transglicosilación que perjudica la hidrólisis de dichos sustratos en azúcares fermentables.

5 Estudios térmicos usando la enzima Bgl1 y la proteína madura mutante de Bgl1Q211H.

La enzima Bgl1 y la proteína madura de Bgl1Q211H purificadas se analizaron en términos de la estabilidad térmica en base al ensayo de pNGP. Se llevaron a cabo dos tipos de ensayos: determinación de la temperatura de desnaturalización y estudios de estabilidad en  
10 condiciones de hidrólisis de la biomasa.

Para la determinación de la temperatura de desnaturalización, las enzimas purificadas se incubaron durante 10 minutos a temperaturas que variaban entre 30 y 80 °C, después se realizaron ensayos convencionales de pNGP usando enzimas tratadas y la actividad se  
15 representó como porcentaje relativo en comparación con la ausencia de tratamiento. En estos estudios se puede calcular un parámetro característico por el cual se determina la temperatura responsable de un 50% de pérdida de actividad enzimática (T1/2). Los resultados se muestran en la Figura 7A. En el estudio de estabilidad en condiciones de hidrólisis de la biomasa, las mezclas de reacción (1 ml de volumen final) conteniendo 0,1 mg  
20 de enzima purificada y 100 µmol de tampón acetato sódico (pH 5,0) se incubaron a 50 °C durante 72 horas y en condiciones de agitación (300 rpm, Thermomixer Confort, Eppendorf). Se tomaron muestras a 24, 48 y 72 horas y se realizaron ensayos de pNGP convencionales. Los resultados se muestran en la Figura 7B.

25 No se observaron diferencias significativas en términos de temperatura de desnaturalización entre Bgl1 y la proteína madura de Bgl1Q211H. T1/2 se calculó que era 65 y 62 °C para la enzima Bgl1 y la proteína madura de Bgl1Q211H, respectivamente, ambos por encima de la temperatura de hidrólisis de la biomasa (50 °C). Respecto a la estabilidad térmica en las condiciones de hidrólisis de biomasa ambas enzimas permanecieron activas tras 72 h a 50  
30 °C. Además, para la enzima mutante se determinó un 20% de activación.



**REIVINDICACIONES**

1. Una variante de beta-glucosidasa que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos un 80% con la SEQ ID NO: 2 y  
5 comprende una sustitución de aminoácido en la posición Q211 correspondiente a las posiciones 1 a 871 de la SEQ ID NO: 2, donde la variante de beta-glucosidasa tiene una actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa.
2. La variante de beta-glucosidasa de acuerdo con la reivindicación 1, donde la  
10 sustitución del aminoácido es Q211H.
3. La variante de beta-glucosidasa de acuerdo con cualquiera de las reivindicaciones 1 ó 2, que comprende la secuencia de aminoácidos SEQ ID NO: 4.
- 15 4. La variante de beta-glucosidasa de acuerdo con la reivindicación 3, que consiste en la secuencia de aminoácidos SEQ ID NO: 4.
5. La variante de beta-glucosidasa de acuerdo con la reivindicación 3, que consiste en la secuencia de aminoácidos SEQ ID NO: 5.  
20
6. Una secuencia de ácido nucleico aislada que codifica la variante de beta-glucosidasa de acuerdo con cualquiera de las reivindicaciones 1 a 5.
7. Una secuencia de ácido nucleico aislada complementaria a la secuencia de ácido  
25 nucleico de acuerdo con la reivindicación 6.
8. Una construcción génica que comprende la secuencia de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 6 ó 7.
- 30 9. La construcción génica de la reivindicación 8, donde la construcción génica es un vector de expresión.
10. Una célula huésped que comprende la construcción génica de acuerdo con cualquiera de las reivindicaciones 8 ó 9.

11. La célula huésped de la reivindicación 10, donde dicha célula es *Myceliophthora thermophila* C1.
12. Una composición enzimática que comprende la variante de beta-glucosidasa de acuerdo con cualquiera de las reivindicaciones 1 a 5.
13. La composición enzimática de la reivindicación 12, que además comprende otras enzimas celulolíticas.
14. La composición enzimática de acuerdo con la reivindicación 13, donde las otras enzimas celulolíticas se seleccionan de la lista que consiste en: endoglucanasas, beta-glucosidasas, celobiohidrolasas, beta-xilosidasas o cualquier combinación de las mismas.
15. La composición enzimática de acuerdo con cualquiera de las reivindicaciones 12 a 14 que además comprende la célula de acuerdo con cualquiera de las reivindicaciones 10 u 11.
16. Un procedimiento para la selección de la variante de beta-glucosidasa con actividad de transglicosilación reducida en comparación con la beta-glucosidasa nativa, según cualquiera de las reivindicaciones 1 a 5, que comprende:
- Incubar variantes de beta-glucosidasa que comprenden una secuencia de aminoácidos que tiene una identidad de secuencia de al menos un 80% con la SEQ ID NO: 2 y comprende una sustitución de aminoácido en la posición Q211 correspondiente a las posiciones 1 a 871 de la SEQ ID NO: 2, con p-nitro-fenil-glucopiranosido en presencia de celobiosa,
  - Medir la liberación de p-nitrofenol,
  - Comparar el valor del p-nitrofenol medido en la etapa (b) con un valor de referencia, y
  - Seleccionar aquellas variantes de beta-glucosidasa cuyos valores medidos en la etapa (b) son más elevados que el valor de referencia.
17. Un procedimiento para producir la variante de beta-glucosidasa según cualquiera de las reivindicaciones 1 a 5, que comprende el diseño de una librería de mutantes de enzimas beta-glucosidasa que presentan una secuencia de aminoácidos que tiene una identidad de secuencia de al menos un 80% con la SEQ ID NO: 2 y comprenden una sustitución de

aminoácido en la posición Q211 correspondiente a las posiciones 1 a 871 de la SEQ ID NO: 2, y la selección de la variante de beta-glucosidasa con actividad de transglicosilación reducida, de acuerdo con las etapas (a) a (d) del procedimiento de la reivindicación 16.

- 5 18. Un procedimiento para producir azúcar fermentable, que comprende:
- a. Incubar material celulósico con la variante de beta-glucosidasa de acuerdo con cualquiera de las reivindicaciones 1 a 5, con la célula huésped de acuerdo con cualquiera de las reivindicaciones 10 u 11 o con la composición enzimática de acuerdo con cualquiera de las reivindicaciones 12 a 15 y
- 10 b. Recuperar el azúcar fermentable obtenido tras la incubación en la etapa (a).
19. Un procedimiento para producir un bioproducto que comprende:
- a. Incubar material celulósico con la variante de beta-glucosidasa de acuerdo con cualquiera de las reivindicaciones 1 a 5, con la célula huésped de acuerdo con cualquiera de
- 15 las reivindicaciones 10 u 11 o con la composición enzimática de acuerdo con cualquiera de las reivindicaciones 12 a 15,
- b. Fermentar los azúcares fermentables obtenidos tras la incubación de la etapa (a) con al menos un microorganismo fermentador, y
- c. Recuperar el bioproducto obtenido tras la fermentación en la etapa (b).
- 20
20. El procedimiento de acuerdo con la reivindicación 19, donde el bioproducto es etanol.

Figura 1

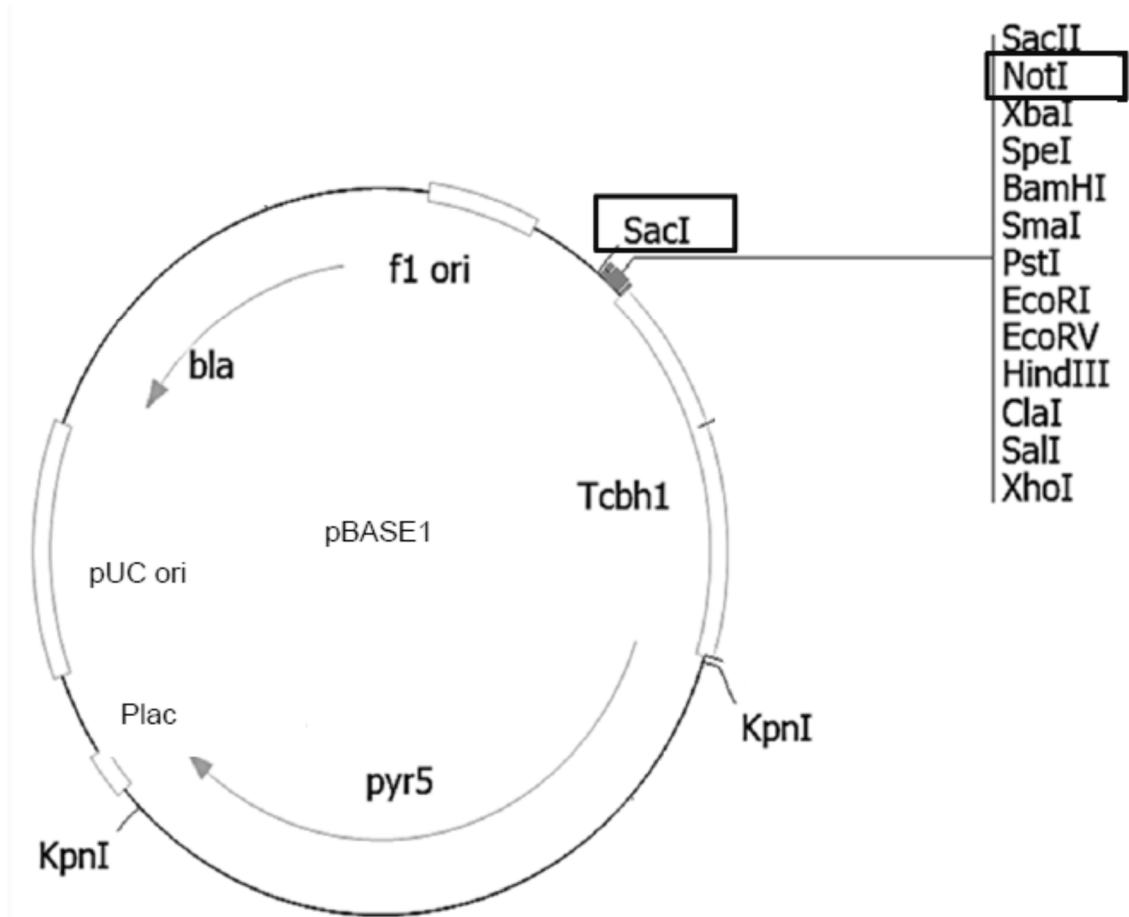
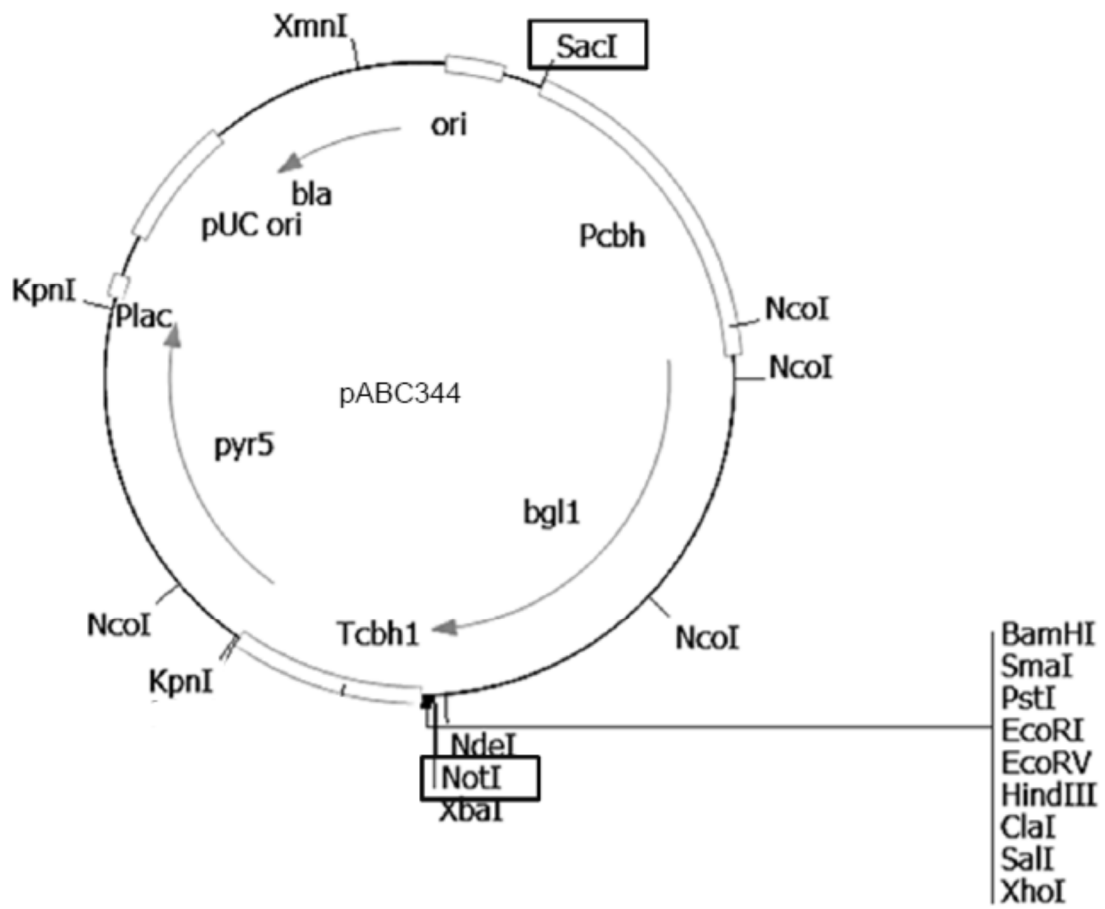


Figura 2



Transformantes con el banco de mutantes *bgI1*

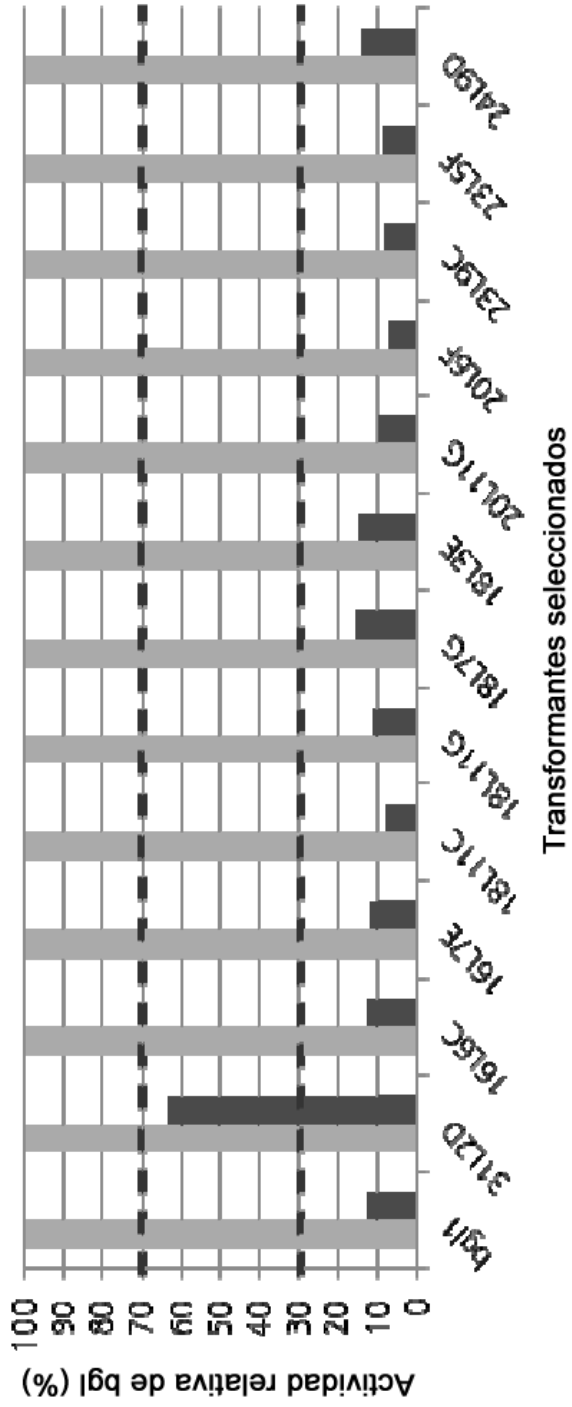


Figura 3

Figura 4

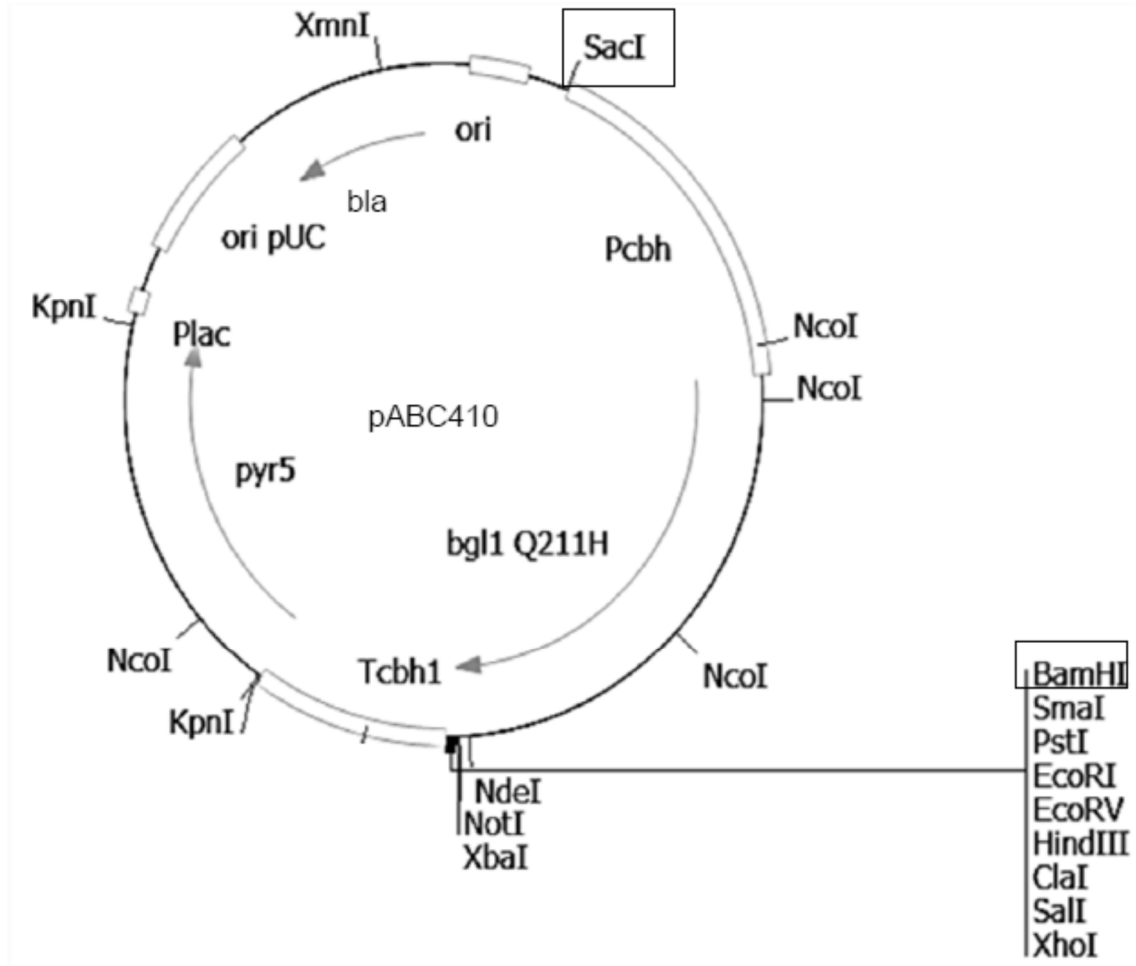


Figura 5

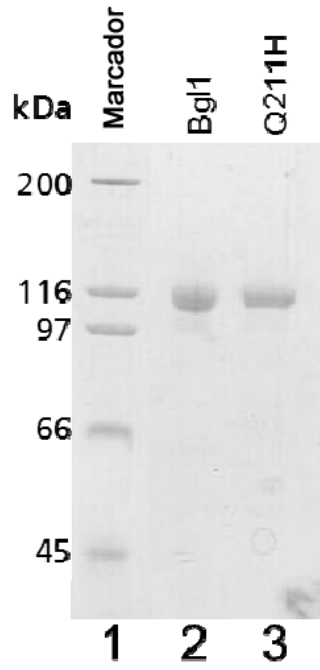


Figura 6

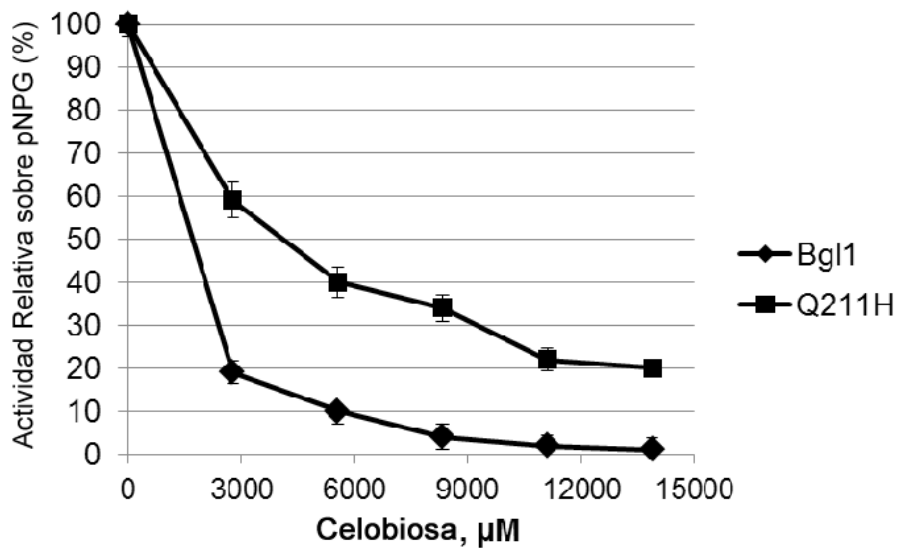
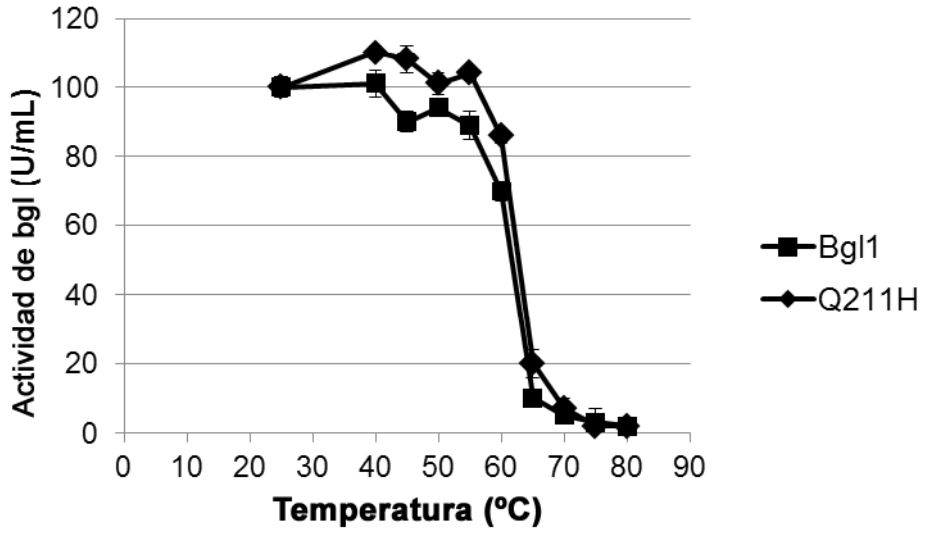


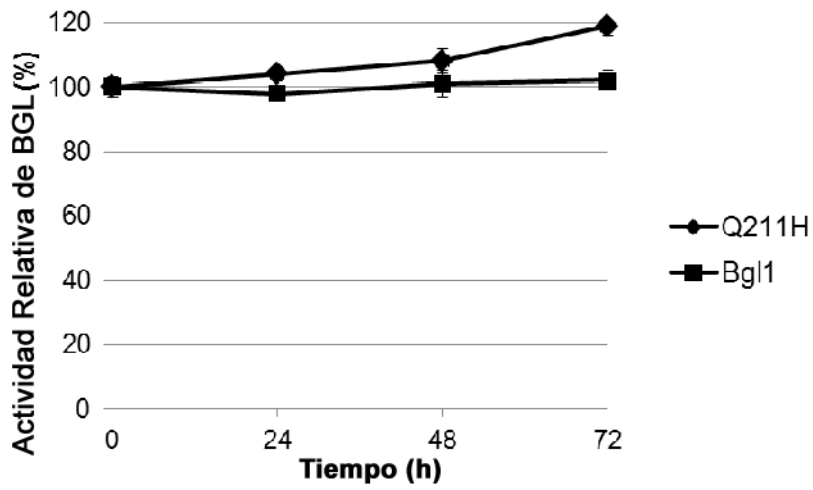


Figura 7

A



B



# ES 2 521 940 A1

## LISTADO DE SECUENCIAS

<110> ABENGOA BIOENERGÍA NUEVAS TECNOLOGÍAS, S.A.  
<120> VARIANTES DE BETA-GLUCOSIDASA CON ACTIVIDAD DE TRANSGLICOSILACIÓN REDUCIDA  
<130> ES1861.32  
<160> 10  
<170> PatentIn version 3.5  
<210> 1  
<211> 852  
<212> PRT  
<213> Secuencia Artificial  
<220>  
<223> Enzima beta-glucosidasa nativa Bgl1  
<400> 1  
Ala Asp Asn His Arg Gln Val His Gln Lys Pro Leu Ala Arg Ser Glu  
1 5 10 15  
Pro Phe Tyr Pro Ser Pro Trp Met Asn Pro Asn Ala Asp Gly Trp Ala  
20 25 30  
Glu Ala Tyr Ala Gln Ala Lys Ser Phe Val Ser Gln Met Thr Leu Leu  
35 40 45  
Glu Lys Val Asn Leu Thr Thr Gly Val Gly Trp Gly Ala Glu Gln Cys  
50 55 60  
Val Gly Gln Val Gly Ala Ile Pro Arg Leu Gly Leu Arg Ser Leu Cys  
65 70 75 80  
Met His Asp Ser Pro Leu Gly Ile Arg Gly Ala Asp Tyr Asn Ser Ala  
85 90 95  
Phe Pro Ser Gly Gln Thr Val Ala Ala Thr Trp Asp Arg Gly Leu Met  
100 105 110  
Tyr Arg Arg Gly Tyr Ala Met Gly Gln Glu Ala Lys Gly Lys Gly Ile  
115 120 125  
Asn Val Leu Leu Gly Pro Val Ala Gly Pro Leu Gly Arg Met Pro Glu  
130 135 140  
Gly Gly Arg Asn Trp Glu Gly Phe Ala Pro Asp Pro Val Leu Thr Gly  
145 150 155 160  
Ile Gly Met Ser Glu Thr Ile Lys Gly Ile Gln Asp Ala Gly Val Ile  
165 170 175  
Ala Cys Ala Lys His Phe Ile Gly Asn Glu Gln Glu His Phe Arg Gln

ES 2 521 940 A1

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

ES 2 521 940 A1

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

ES 2 521 940 A1

725 730 735

Pro Leu Leu Arg Ser Ser Gly Gly Asn Ser Pro Gly Gly Asn Arg Gln  
740 745 750

Leu Tyr Asp Ile Val Tyr Thr Ile Thr Ala Asp Ile Thr Asn Thr Gly  
755 760 765

Ser Val Val Gly Glu Glu Val Pro Gln Leu Tyr Val Ser Leu Gly Gly  
770 775 780

Pro Glu Asp Pro Lys Val Gln Leu Arg Asp Phe Asp Arg Met Arg Ile  
785 790 795 800

Glu Pro Gly Glu Thr Arg Gln Phe Thr Gly Arg Leu Thr Arg Arg Asp  
805 810 815

Leu Ser Asn Trp Asp Val Thr Val Gln Asp Trp Val Ile Ser Arg Tyr  
820 825 830

Pro Lys Thr Ala Tyr Val Gly Arg Ser Ser Arg Lys Leu Asp Leu Lys  
835 840 845

Ile Glu Leu Pro  
850

<210> 2  
 <211> 871  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Preproteína de la SEQ ID NO: 1

<220>  
 <221> SIGNAL  
 <222> (1)..(19)

<400> 2

Met Gln Leu Pro Ala Ala Ala Gln Trp Leu Leu Thr Leu Pro Ala Lys  
1 5 10 15

Ala Ser Leu Ala Asp Asn His Arg Gln Val His Gln Lys Pro Leu Ala  
20 25 30

Arg Ser Glu Pro Phe Tyr Pro Ser Pro Trp Met Asn Pro Asn Ala Asp  
35 40 45

Gly Trp Ala Glu Ala Tyr Ala Gln Ala Lys Ser Phe Val Ser Gln Met  
50 55 60

Thr Leu Leu Glu Lys Val Asn Leu Thr Thr Gly Val Gly Trp Gly Ala  
65 70 75 80

ES 2 521 940 A1

Glu Gln Cys Val Gly Gln Val Gly Ala Ile Pro Arg Leu Gly Leu Arg  
 85 90 95  
 Ser Leu Cys Met His Asp Ser Pro Leu Gly Ile Arg Gly Ala Asp Tyr  
 100 105 110  
 Asn Ser Ala Phe Pro Ser Gly Gln Thr Val Ala Ala Thr Trp Asp Arg  
 115 120 125  
 Gly Leu Met Tyr Arg Arg Gly Tyr Ala Met Gly Gln Glu Ala Lys Gly  
 130 135 140  
 Lys Gly Ile Asn Val Leu Leu Gly Pro Val Ala Gly Pro Leu Gly Arg  
 145 150 155 160  
 Met Pro Glu Gly Gly Arg Asn Trp Glu Gly Phe Ala Pro Asp Pro Val  
 165 170 175  
 Leu Thr Gly Ile Gly Met Ser Glu Thr Ile Lys Gly Ile Gln Asp Ala  
 180 185 190  
 Gly Val Ile Ala Cys Ala Lys His Phe Ile Gly Asn Glu Gln Glu His  
 195 200 205  
 Phe Arg Gln Val Pro Glu Ala Gln Gly Tyr Gly Tyr Asn Ile Ser Glu  
 210 215 220  
 Thr Leu Ser Ser Asn Ile Asp Asp Lys Thr Met His Glu Leu Tyr Leu  
 225 230 235 240  
 Trp Pro Phe Ala Asp Ala Val Arg Ala Gly Val Gly Ser Val Met Cys  
 245 250 255  
 Ser Tyr Gln Gln Val Asn Asn Ser Tyr Ala Cys Gln Asn Ser Lys Leu  
 260 265 270  
 Leu Asn Asp Leu Leu Lys Asn Glu Leu Gly Phe Gln Gly Phe Val Met  
 275 280 285  
 Ser Asp Trp Gln Ala Gln His Thr Gly Ala Ala Ser Ala Val Ala Gly  
 290 295 300  
 Leu Asp Met Ser Met Pro Gly Asp Thr Gln Phe Asn Thr Gly Val Ser  
 305 310 315 320  
 Phe Trp Gly Ala Asn Leu Thr Leu Ala Val Leu Asn Gly Thr Val Pro  
 325 330 335  
 Ala Tyr Arg Leu Asp Asp Met Ala Met Arg Ile Met Ala Ala Leu Phe  
 340 345 350

ES 2 521 940 A1

Lys Val Thr Lys Thr Thr His Leu Glu Pro Ile Asn Phe Ser Phe Trp  
 355 360 365  
 Thr Asp Asp Thr Tyr Gly Pro Ile His Trp Ala Ala Lys His Gly Tyr  
 370 375 380  
 Gln Lys Ile Asn Ser His Val Asp Val Arg Ala Asp His Gly Asn Leu  
 385 390 395  
 Ile Arg Glu Ile Ala Ala Lys Gly Thr Val Leu Leu Lys Asn Thr Gly  
 405 410 415  
 Ser Leu Pro Leu Asn Lys Pro Lys Phe Val Ala Val Ile Gly Glu Asp  
 420 425 430  
 Ala Gly Ser Ser Pro Asn Gly Pro Asn Gly Cys Ser Asp Arg Gly Cys  
 435 440 445  
 Asn Glu Gly Thr Leu Ala Met Gly Trp Gly Ser Gly Thr Ala Asn Tyr  
 450 455 460  
 Pro Tyr Leu Val Ser Pro Asp Ala Ala Leu Gln Ala Arg Ala Ile Gln  
 465 470 475 480  
 Asp Gly Thr Arg Tyr Glu Ser Val Leu Ser Asn Tyr Ala Glu Glu Lys  
 485 490 495  
 Thr Lys Ala Leu Val Ser Gln Ala Asn Ala Thr Ala Ile Val Phe Val  
 500 505 510  
 Asn Ala Asp Ser Gly Glu Gly Tyr Ile Asn Val Asp Gly Asn Glu Gly  
 515 520 525  
 Asp Arg Lys Asn Leu Thr Leu Trp Asn Asn Gly Asp Thr Leu Val Lys  
 530 535 540  
 Asn Val Ser Ser Trp Cys Ser Asn Thr Ile Val Val Ile His Ser Val  
 545 550 555 560  
 Gly Pro Val Leu Leu Thr Asp Trp Tyr Asp Asn Pro Asn Ile Thr Ala  
 565 570 575  
 Ile Leu Trp Ala Gly Leu Pro Gly Gln Glu Ser Gly Asn Ser Ile Thr  
 580 585 590  
 Asp Val Leu Tyr Gly Lys Val Asn Pro Ala Ala Arg Ser Pro Phe Thr  
 595 600 605  
 Trp Gly Lys Thr Arg Glu Ser Tyr Gly Ala Asp Val Leu Tyr Lys Pro  
 610 615 620

ES 2 521 940 A1

Asn 625 Asn Gly 625 Asn Gly 630 Ala 630 Pro Gln Gln Asp 635 Phe Thr Glu Gly Val Phe 640  
Ile Asp Tyr Arg Tyr 645 Phe Asp Lys Val Asp 650 Asp Asp Ser Val Ile Tyr 655  
Glu Phe Gly His 660 Gly Leu Ser Tyr Thr 665 Thr Phe Glu Tyr Ser Asn Ile 670  
Arg Val Val 675 Lys Ser Asn Val Ser 680 Glu Tyr Arg Pro Thr Thr Gly Thr 685  
Thr Ala Gln Ala Pro Thr Phe 695 Gly Asn Phe Ser Thr Asp Leu Glu Asp 700  
Tyr 705 Leu Phe Pro Lys Asp 710 Glu Phe Pro Tyr Ile 715 Tyr Gln Tyr Ile Tyr 720  
Pro Tyr Leu Asn Thr 725 Thr Asp Pro Arg Arg 730 Ala Ser Ala Asp Pro His 735  
Tyr Gly Gln Thr 740 Ala Glu Glu Phe Leu 745 Pro Pro His Ala Thr Asp Asp 750  
Asp Pro Gln 755 Pro Leu Leu Arg Ser 760 Ser Gly Gly Asn Ser 765 Pro Gly Gly  
Asn Arg Gln Leu Tyr Asp Ile 775 Val Tyr Thr Ile Thr 780 Ala Asp Ile Thr  
Asn Thr Gly Ser Val Val 790 Gly Glu Glu Val Pro Gln Leu Tyr Val Ser 800  
Leu Gly Gly Pro Glu 805 Asp Pro Lys Val Gln 810 Leu Arg Asp Phe Asp Arg 815  
Met Arg Ile Glu Pro Gly Glu Thr Arg 825 Gln Phe Thr Gly Arg 830 Leu Thr  
Arg Arg Asp 835 Leu Ser Asn Trp Asp 840 Val Thr Val Gln Asp 845 Trp Val Ile  
Ser Arg Tyr Pro Lys Thr Ala 855 Tyr Val Gly Arg Ser 860 Ser Arg Lys Leu  
Asp 865 Leu Lys Ile Glu Leu Pro 870

<210> 3  
<211> 2616



ES 2 521 940 A1

<212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Secuencia nucleotídica codificante para la preprotenía SEQ ID NO: 2

<400> 3  
 atgcaacttc cagccgcagc ccaatggctg ctcacgcttc ccgcgaaagc ctcaacttgc 60  
 gacaatcatc gtcaggttca ccagaagccc ctcgcgagat ctgaaccttt ttaccctgcg 120  
 ccatggatga atcccaacgc cgacggctgg gcggaggcct atgcccaggc caagtccttt 180  
 gtctccaaa tgactctgct agagaaggtc aacttgacca cgggagtcgg gtggggggct 240  
 gagcagtgcg tcggccaagt gggcgcgac cctcgccttg gacttcgcag tctgtgcatg 300  
 catgactccc ctctcggcat ccgaggagcc gactacaact cagcgttccc ctctggccag 360  
 accgttgcg ctacctggga tcgcggtctg atgtaccgtc gcggctacgc aatgggccag 420  
 gaggccaaag gcaagggcat caatgtcctt ctcggaccag tcgccggccc cttggccgc 480  
 atgcccagg gcggtcgtaa ctgggaaggc ttcgctccc atcccgtcct taccggcatc 540  
 ggcatgtccg agacgatcaa gggcattcag gatgctggcg tcatcgcttg tgcgaagcac 600  
 ttatttgaa acgagcagga gcacttcaga caggtgccag aagcccaggg atacggttac 660  
 aacatcagcg aaaccctctc ctccaacatt gacgacaaga ccatgcacga gctgtacctt 720  
 tggccgtttg ccgatgccgt ccgggcccgc gtcggctctg tcatgtgctc gtaccagcag 780  
 gtcaacaact cgtacgcctg ccagaactcg aagctgctga acgacctcct caagaacgag 840  
 cttgggtttc agggcttcgt catgagcgac tggcaggcac agcacttg cgcagcaagc 900  
 gccgtggctg gtctcgatat gtccatgccg ggcgacacc agttcaacac tggcgtcagt 960  
 ttctggggcg ccaatctcac cctcgcctg ctcaacggca cagtccttgc ctaccgtctc 1020  
 gacgacatgg ccatgcgcat catggccgcc ctcttcaagg tcaccaagac caccacctg 1080  
 gaacctatca acttctcctt ctggaccgac gacacttatg gcccgatcca ctgggcccgc 1140  
 aagcatggct accagaagat taattcccac gttgacgtcc gcgccgacca cggcaacctc 1200  
 atccgggaga ttgccgcaa gggtaggtg ctgctgaaga ataccggctc tctaccctg 1260  
 aacaagccaa agttcgtggc cgtcatcggc gaggatgctg ggtcgagccc caacgggccc 1320  
 aacggctgta gcgaccgcg ctgtaacgaa ggcagctcg ccatgggctg gggttccggc 1380  
 acagccaact atccgtacct cgtttcccc gacgccgcg tccaggcacg ggccatccag 1440  
 gacggcacga ggtacgagag cgtcctgtcc aactacgcc aggaaaagac aaaggctctg 1500  
 gtctcgagg ccaatgcaac cgccatcgtc ttcgtcaatg ccgactcagg cgagggtac 1560  
 atcaacgtgg acggtaacga gggcgaccgt aagaacctga ctctctggaa caacgggtgat 1620  
 actctggctc agaacgtgtc gagctgggtc agcaacacca tcgtcgtcat cactcggctc 1680  
 ggcccgtcc tcctgaccga ttggtacgac aacccaaca tcacggccat tctctgggct 1740  
 ggtcttccg gccaggagtc gggcaactcc atcaccgacg tgctttacgg caaggtaac 1800  
 cccgccgcc gctcgcctt cacttggggc aagaccgcg aaagctatgg cgcggacgct 1860

ES 2 521 940 A1

ctgtacaagc cgaataatgg caatggtgcg cccaacagg acttcaccga gggcgtcttc 1920  
atcgactacc gctacttcga caaggttgac gatgactcgg tcactctacga gttcggccac 1980  
ggcctgagct acaccacctt cgagtacagc aacatccgcy tcgtcaagtc caacgtcagc 2040  
gagtaccggc ccacgacggg caccacggcc caggccccga cgtttggcaa cttctccacc 2100  
gacctggagg actatctctt cccaaggac gagttccctt acatctacca gtacatctac 2160  
ccgtacctca acacgaccga cccccggagg gcctcggccg atccccacta cggccagacc 2220  
gccgaggagt tcctcccgcc ccacgccacc gatgacgacc cccagccgct cctccggtcc 2280  
tcggggcggaa actccccgg cggcaaccgc cagctgtacg acattgtcta cacaatcacg 2340  
gccgacatca cgaatacggg ctccgttgta ggcgaggagg ttccgcagct ctacgtctcg 2400  
ctgggcggtc ccgaggacc caaggtgcag ctgcgcgact ttgacaggat gcggatcgaa 2460  
cccggcgaga cgaggcagtt caccggccgc ctgacgcgca gagatctgag caactgggac 2520  
gtcacggtgc aggactgggt catcagcagg tatccaaga cggcatatgt tgggaggagc 2580  
agccggaagt tggatctcaa gattgagctt ccttga 2616

<210> 4  
<211> 852  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Proteína madura de Bgl1Q211H

<400> 4

Ala Asp Asn His Arg Gln Val His Gln Lys Pro Leu Ala Arg Ser Glu  
1 5 10 15

Pro Phe Tyr Pro Ser Pro Trp Met Asn Pro Asn Ala Asp Gly Trp Ala  
20 25 30

Glu Ala Tyr Ala Gln Ala Lys Ser Phe Val Ser Gln Met Thr Leu Leu  
35 40 45

Glu Lys Val Asn Leu Thr Thr Gly Val Gly Trp Gly Ala Glu Gln Cys  
50 55 60

Val Gly Gln Val Gly Ala Ile Pro Arg Leu Gly Leu Arg Ser Leu Cys  
65 70 75 80

Met His Asp Ser Pro Leu Gly Ile Arg Gly Ala Asp Tyr Asn Ser Ala  
85 90 95

Phe Pro Ser Gly Gln Thr Val Ala Ala Thr Trp Asp Arg Gly Leu Met  
100 105 110

Tyr Arg Arg Gly Tyr Ala Met Gly Gln Glu Ala Lys Gly Lys Gly Ile  
115 120 125

ES 2 521 940 A1

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

ES 2 521 940 A1

Leu Asn Lys Pro Lys Phe Val Ala Val Ile Gly Glu Asp Ala Gly Ser  
 405 410 415  
 Ser Pro Asn Gly Pro Asn Gly Cys Ser Asp Arg Gly Cys Asn Glu Gly  
 420 425 430  
 Thr Leu Ala Met Gly Trp Gly Ser Gly Thr Ala Asn Tyr Pro Tyr Leu  
 435 440 445  
 Val Ser Pro Asp Ala Ala Leu Gln Ala Arg Ala Ile Gln Asp Gly Thr  
 450 455 460  
 Arg Tyr Glu Ser Val Leu Ser Asn Tyr Ala Glu Glu Lys Thr Lys Ala  
 465 470 475 480  
 Leu Val Ser Gln Ala Asn Ala Thr Ala Ile Val Phe Val Asn Ala Asp  
 485 490 495  
 Ser Gly Glu Gly Tyr Ile Asn Val Asp Gly Asn Glu Gly Asp Arg Lys  
 500 505 510  
 Asn Leu Thr Leu Trp Asn Asn Gly Asp Thr Leu Val Lys Asn Val Ser  
 515 520 525  
 Ser Trp Cys Ser Asn Thr Ile Val Val Ile His Ser Val Gly Pro Val  
 530 535 540  
 Leu Leu Thr Asp Trp Tyr Asp Asn Pro Asn Ile Thr Ala Ile Leu Trp  
 545 550 555 560  
 Ala Gly Leu Pro Gly Gln Glu Ser Gly Asn Ser Ile Thr Asp Val Leu  
 565 570 575  
 Tyr Gly Lys Val Asn Pro Ala Ala Arg Ser Pro Phe Thr Trp Gly Lys  
 580 585 590  
 Thr Arg Glu Ser Tyr Gly Ala Asp Val Leu Tyr Lys Pro Asn Asn Gly  
 595 600 605  
 Asn Gly Ala Pro Gln Gln Asp Phe Thr Glu Gly Val Phe Ile Asp Tyr  
 610 615 620  
 Arg Tyr Phe Asp Lys Val Asp Asp Asp Ser Val Ile Tyr Glu Phe Gly  
 625 630 635 640  
 His Gly Leu Ser Tyr Thr Thr Phe Glu Tyr Ser Asn Ile Arg Val Val  
 645 650 655  
 Lys Ser Asn Val Ser Glu Tyr Arg Pro Thr Thr Gly Thr Thr Ala Gln  
 660 665 670

ES 2 521 940 A1

Ala Pro Thr Phe Gly Asn Phe Ser Thr Asp Leu Glu Asp Tyr Leu Phe  
675 680 685

Pro Lys Asp Glu Phe Pro Tyr Ile Tyr Gln Tyr Ile Tyr Pro Tyr Leu  
690 695 700

Asn Thr Thr Asp Pro Arg Arg Ala Ser Ala Asp Pro His Tyr Gly Gln  
705 710 715 720

Thr Ala Glu Glu Phe Leu Pro Pro His Ala Thr Asp Asp Asp Pro Gln  
725 730 735

Pro Leu Leu Arg Ser Ser Gly Gly Asn Ser Pro Gly Gly Asn Arg Gln  
740 745 750

Leu Tyr Asp Ile Val Tyr Thr Ile Thr Ala Asp Ile Thr Asn Thr Gly  
755 760 765

Ser Val Val Gly Glu Glu Val Pro Gln Leu Tyr Val Ser Leu Gly Gly  
770 775 780

Pro Glu Asp Pro Lys Val Gln Leu Arg Asp Phe Asp Arg Met Arg Ile  
785 790 795 800

Glu Pro Gly Glu Thr Arg Gln Phe Thr Gly Arg Leu Thr Arg Arg Asp  
805 810 815

Leu Ser Asn Trp Asp Val Thr Val Gln Asp Trp Val Ile Ser Arg Tyr  
820 825 830

Pro Lys Thr Ala Tyr Val Gly Arg Ser Ser Arg Lys Leu Asp Leu Lys  
835 840 845

Ile Glu Leu Pro  
850

<210> 5  
<211> 871  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Preproteína Bgl1Q211H

<220>  
<221> SIGNAL  
<222> (1)..(19)

<400> 5

Met Gln Leu Pro Ala Ala Ala Gln Trp Leu Leu Thr Leu Pro Ala Lys  
1 5 10 15

ES 2 521 940 A1

Ala Ser Leu Ala Asp Asn His Arg Gln Val His Gln Lys Pro Leu Ala  
20 25 30

Arg Ser Glu Pro Phe Tyr Pro Ser Pro Trp Met Asn Pro Asn Ala Asp  
35 40 45

Gly Trp Ala Glu Ala Tyr Ala Gln Ala Lys Ser Phe Val Ser Gln Met  
50 55 60

Thr Leu Leu Glu Lys Val Asn Leu Thr Thr Gly Val Gly Trp Gly Ala  
65 70 75 80

Glu Gln Cys Val Gly Gln Val Gly Ala Ile Pro Arg Leu Gly Leu Arg  
85 90 95

Ser Leu Cys Met His Asp Ser Pro Leu Gly Ile Arg Gly Ala Asp Tyr  
100 105 110

Asn Ser Ala Phe Pro Ser Gly Gln Thr Val Ala Ala Thr Trp Asp Arg  
115 120 125

Gly Leu Met Tyr Arg Arg Gly Tyr Ala Met Gly Gln Glu Ala Lys Gly  
130 135 140

Lys Gly Ile Asn Val Leu Leu Gly Pro Val Ala Gly Pro Leu Gly Arg  
145 150 155 160

Met Pro Glu Gly Gly Arg Asn Trp Glu Gly Phe Ala Pro Asp Pro Val  
165 170 175

Leu Thr Gly Ile Gly Met Ser Glu Thr Ile Lys Gly Ile Gln Asp Ala  
180 185 190

Gly Val Ile Ala Cys Ala Lys His Phe Ile Gly Asn Glu Gln Glu His  
195 200 205

Phe Arg His Val Pro Glu Ala Gln Gly Tyr Gly Tyr Asn Ile Ser Glu  
210 215 220

Thr Leu Ser Ser Asn Ile Asp Asp Lys Thr Met His Glu Leu Tyr Leu  
225 230 235 240

Trp Pro Phe Ala Asp Ala Val Arg Ala Gly Val Gly Ser Val Met Cys  
245 250 255

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

Leu Asn Asp Leu Leu Lys Asn Glu Leu Gly Phe Gln Gly Phe Val Met  
275 280 285

ES 2 521 940 A1

Ser Asp Trp Gln Ala Gln His Thr Gly Ala Ala Ser Ala Val Ala Gly  
 290 295 300

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

Phe Trp Gly Ala Asn Leu Thr Leu Ala Val Leu Asn Gly Thr Val Pro  
 325 330 335

Ala Tyr Arg Leu Asp Asp Met Ala Met Arg Ile Met Ala Ala Leu Phe  
 340 345 350

Lys Val Thr Lys Thr Thr His Leu Glu Pro Ile Asn Phe Ser Phe Trp  
 355 360 365

Thr Asp Asp Thr Tyr Gly Pro Ile His Trp Ala Ala Lys His Gly Tyr  
 370 375 380

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

Ile Arg Glu Ile Ala Ala Lys Gly Thr Val Leu Leu Lys Asn Thr Gly  
 405 410 415

Ser Leu Pro Leu Asn Lys Pro Lys Phe Val Ala Val Ile Gly Glu Asp  
 420 425 430

Ala Gly Ser Ser Pro Asn Gly Pro Asn Gly Cys Ser Asp Arg Gly Cys  
 435 440 445

Asn Glu Gly Thr Leu Ala Met Gly Trp Gly Ser Gly Thr Ala Asn Tyr  
 450 455 460

Pro Tyr Leu Val Ser Pro Asp Ala Ala Leu Gln Ala Arg Ala Ile Gln  
 465 470 475 480

Asp Gly Thr Arg Tyr Glu Ser Val Leu Ser Asn Tyr Ala Glu Glu Lys  
 485 490 495

Thr Lys Ala Leu Val Ser Gln Ala Asn Ala Thr Ala Ile Val Phe Val  
 500 505 510

Asn Ala Asp Ser Gly Glu Gly Tyr Ile Asn Val Asp Gly Asn Glu Gly  
 515 520 525

Asp Arg Lys Asn Leu Thr Leu Trp Asn Asn Gly Asp Thr Leu Val Lys  
 530 535 540

Asn Val Ser Ser Trp Cys Ser Asn Thr Ile Val Val Ile His Ser Val  
 545 550 555 560

ES 2 521 940 A1

Gly Pro Val Leu Leu Thr Asp Trp Tyr Asp Asn Pro Asn Ile Thr Ala  
565 570 575

Ile Leu Trp Ala Gly Leu Pro Gly Gln Glu Ser Gly Asn Ser Ile Thr  
580 585 590

Asp Val Leu Tyr Gly Lys Val Asn Pro Ala Ala Arg Ser Pro Phe Thr  
595 600 605

Trp Gly Lys Thr Arg Glu Ser Tyr Gly Ala Asp Val Leu Tyr Lys Pro  
610 615 620

Asn Asn Gly Asn Gly Ala Pro Gln Gln Asp Phe Thr Glu Gly Val Phe  
625 630 635 640

Ile Asp Tyr Arg Tyr Phe Asp Lys Val Asp Asp Asp Ser Val Ile Tyr  
645 650 655

Glu Phe Gly His Gly Leu Ser Tyr Thr Thr Phe Glu Tyr Ser Asn Ile  
660 665 670

Arg Val Val Lys Ser Asn Val Ser Glu Tyr Arg Pro Thr Thr Gly Thr  
675 680 685

Thr Ala Gln Ala Pro Thr Phe Gly Asn Phe Ser Thr Asp Leu Glu Asp  
690 695 700

Tyr Leu Phe Pro Lys Asp Glu Phe Pro Tyr Ile Tyr Gln Tyr Ile Tyr  
705 710 715 720

Pro Tyr Leu Asn Thr Thr Asp Pro Arg Arg Ala Ser Ala Asp Pro His  
725 730 735

Tyr Gly Gln Thr Ala Glu Glu Phe Leu Pro Pro His Ala Thr Asp Asp  
740 745 750

Asp Pro Gln Pro Leu Leu Arg Ser Ser Gly Gly Asn Ser Pro Gly Gly  
755 760 765

Asn Arg Gln Leu Tyr Asp Ile Val Tyr Thr Ile Thr Ala Asp Ile Thr  
770 775 780

Asn Thr Gly Ser Val Val Gly Glu Glu Val Pro Gln Leu Tyr Val Ser  
785 790 795 800

Leu Gly Gly Pro Glu Asp Pro Lys Val Gln Leu Arg Asp Phe Asp Arg  
805 810 815

Met Arg Ile Glu Pro Gly Glu Thr Arg Gln Phe Thr Gly Arg Leu Thr  
820 825 830



ES 2 521 940 A1

Arg Arg Asp Leu Ser Asn Trp Asp Val Thr Val Gln Asp Trp Val Ile  
 835 840 845

Ser Arg Tyr Pro Lys Thr Ala Tyr Val Gly Arg Ser Ser Arg Lys Leu  
 850 855 860

Asp Leu Lys Ile Glu Leu Pro  
 865 870

<210> 6  
 <211> 2616  
 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Polinucleótido codificante para SEQ ID NO: 5

<400> 6  
 atgcaacttc cagccgcagc ccaatggctg ctcacgcttc ccgcgaaagc ctcacttgct 60  
 gacaatcatc gtcaggttca ccagaagccc ctcgcgagat ctgaaccttt ttacccgtcg 120  
 ccatggatga atccaacgc cgacggctgg gcggaggcct atgcccaggc caagtccttt 180  
 gtctccaaa tgactctgct agagaaggtc aacttgacca cgggagtcgg gtggggggct 240  
 gagcagtgcg tcggccaagt gggcgcgatc cctcgccttg gacttcgcag tctgtgcatg 300  
 catgactccc ctctcggcat ccgaggagcc gactacaact cagcgttccc ctctggccag 360  
 accgttgctg ctacctggga tcgcggtctg atgtaccgct gcggctacgc aatggggccag 420  
 gaggcaaaag gcaagggcat caatgtcctt ctcggaccag tcgccggccc ccttggccgc 480  
 atgcccgagg gcggtcgtaa ctgggaaggc ttcgctcccg atcccgtcct taccggcatc 540  
 ggcatgtccg agacgatcaa gggcattcag gatgctggcg tcatcgcttg tgcgaagcac 600  
 ttatttgaa acgagcagga gcacttcaga catgtgccag aagcccaggg atacggttac 660  
 aacatcagcg aaaccctctc ctccaacatt gacgacaaga ccatgcacga gctgtacctt 720  
 tggccgtttg ccgatgccgt ccgggccggc gtcggctctg tcatgtgctc gtaccagcag 780  
 gtcaacaact cgtacgcctg ccagaactcg aagctgctga acgacctcct caagaacgag 840  
 cttgggtttc agggcttcgt catgagcgac tggcaggcac agcacactgg cgcagcaagc 900  
 gccgtggctg gtctcgatat gtccatgccg ggcgacacc agttcaacac tggcgtcagt 960  
 ttctggggcg ccaatctcac cctcgccgtc ctcaacggca cagtccctgc ctaccgtctc 1020  
 gacgacatgg ccatgcgcat catggccgcc ctcttcaagg tcaccaagac caccacctg 1080  
 gaacctatca acttctcctt ctggaccgac gacacttatg gcccgatcca ctgggccgcc 1140  
 aagcatggct accagaagat taattccac gttgacgtcc gcgccgacca cggcaacctc 1200  
 atccgggaga ttgccgcaa gggtaggtg ctgctgaaga ataccggctc tctaccctg 1260  
 aacaagcaa agttcgtggc cgtcatcggc gaggatgctg ggtcgagccc caacgggccc 1320  
 aacggctgta gcgaccgcg ctgtaacgaa ggcacgctcg ccatgggctg gggttccggc 1380

ES 2 521 940 A1

acagccaact atccgtacct cgtttcccc gacgccgcgc tccaggcacg ggccatccag 1440  
gacggcacga ggtacgagag cgtcctgtcc aactacgccg aggaaaagac aaaggctctg 1500  
gtctcgagg ccaatgcaac cgccatcgtc ttcgtcaatg ccgactcagg cgagggctac 1560  
atcaacgtgg acggtaacga gggcgaccgt aagaacctga ctctctggaa caacgggtgat 1620  
actctggtca agaacgtgtc gagctggtgc agcaacacca tcgtcgtcat ccactcggtc 1680  
ggcccggctc tcctgaccga ttggtacgac aacccaaca tcacggccat tctctgggct 1740  
ggtcttccgg gccaggagtc gggcaactcc atcaccgacg tgctttacgg caaggtaaac 1800  
cccgccgcc gctcgccctt cacttggggc aagaccgcg aaagctatgg cgcggacgctc 1860  
ctgtacaagc cgaataatgg caatggtgcg cccaacagg acttcaccga gggcgtcttc 1920  
atcgactacc gctacttcga caaggttgac gatgactcgg tcatctacga gttcggccac 1980  
ggcctgagct acaccacctt cgagtacagc aacatccgcg tcgtcaagtc caacgtcagc 2040  
gagtaccggc ccacgacggg caccacggcc caggccccga cgtttgcaa cttctccacc 2100  
gacctggagg actatctctt cccaaggac gagttcccct acatctacca gtacatctac 2160  
ccgtacctca acacgaccga ccccggagg gcctcggccg atccccacta cggccagacc 2220  
gccgaggagt tcctcccgcc ccacgccacc gatgacgacc cccagccgct cctccggctc 2280  
tcgggcggaa actccccgg cggcaaccgc cagctgtacg acattgtcta cacaatcacg 2340  
gccgacatca cgaatacggg ctccgttgta ggcgaggagg ttccgcagct ctacgtctcg 2400  
ctgggcggtc ccgaggacc caaggtgacg ctgacgacgact ttgacaggat gcggatcgaa 2460  
cccggcgaga cgaggcagtt caccggccgc ctgacgcgca gagatctgag caactgggac 2520  
gtcacggtgc aggactgggt catcagcagg tatccaaga cggcatatgt tgggaggagc 2580  
agccggaagt tggatctcaa gattgagctt ccttga 2616

<210> 7  
<211> 40  
<212> DNA  
<213> Secuencia Artificial

<220>  
<223> oligonucleótido 1 utilizado para la amplificación mutagénica de un fragmento de 2553 pb correspondiente a la secuencia bgl1 sin péptido señal ni codon de terminación.

<400> 7  
gctgacaatc atcgtcaggt tcaccagaag ccctcgcga 40

<210> 8  
<211> 40  
<212> DNA  
<213> Secuencia Artificial

<220>  
<223> oligonucleótido 2 utilizado para la amplificación mutagénica de un fragmento de 2553 pb correspondiente a la secuencia bgl1 sin péptido señal ni codón de terminación.

ES 2 521 940 A1

<400> 8  
aggaagctca atcttgagat ccaacttccg gctgctcctc 40

<210> 9  
<211> 30  
<212> DNA  
<213> Secuencia Artificial

<220>  
<223> Oligonucleótido 3 utilizado para la amplificación del gen bgl1 expresado  
en  
el mutante 31L2D.

<400> 9  
cgaggagctc cttacaaaaa aaaggtatcc 30

<210> 10  
<211> 53  
<212> DNA  
<213> Secuencia Artificial

<220>  
<223> Oligonucleótido 4 utilizado para la amplificación del gen bgl1 expresado  
en  
el mutante 31L2D.

<400> 10  
ttcctgcagc ccgggggatc ctcaaggaag ctcaatcttg agatccaact tcc 53



- ②① N.º solicitud: 201330678  
②② Fecha de presentación de la solicitud: 10.05.2013  
③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
A	WO 2011066457 A2 (CODEXIS LABORATORIES HUNGARY KFT) 03.06.2011, reivindicaciones.	1-20
A	WO 2011041594 A1 (CODEXIS, INC.) 07.04.2011, reivindicaciones.	1-20
A	BHAT, K.M. et al. "Purification and characterization of a n extracellular $\beta$ -glucosidase from the thermophilic fungus <i>Sporotrichum thermophile</i> and its influence on cellulase activity". JOURNAL OF GENERAL MICROBIOLOGY. 01.11.1993. Vol. 139, N.º. 11, páginas 2825-2832; todo el documento.	1-20
A	KIM, T.J. et al. "Role of the glutamate 332 residue in the transglycosilation activity of <i>Thermus maltogenic</i> amylase". BIOCHEMISTRY. 13.06.2000. Vol. 39, N.º. 23, páginas 6773-6780; resumen.	1-20

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

**El presente informe ha sido realizado**

para todas las reivindicaciones

para las reivindicaciones n.º:

Fecha de realización del informe  
21.03.2014

Examinador  
M. Novoa Sanjurjo

Página  
1/4

CLASIFICACIÓN OBJETO DE LA SOLICITUD

**C12N9/42** (2006.01)

**C12P19/02** (2006.01)

**C12P7/10** (2006.01)

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12N, C12P

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, BIOSIS, REGISTRY, HCAPLUS, GOOGLE

Fecha de Realización de la Opinión Escrita: 21.03.2014

#### Declaración

<b>Novedad (Art. 6.1 LP 11/1986)</b>	Reivindicaciones 1-20	<b>SI</b>
	Reivindicaciones	<b>NO</b>
<b>Actividad inventiva (Art. 8.1 LP11/1986)</b>	Reivindicaciones 1-20	<b>SI</b>
	Reivindicaciones	<b>NO</b>

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

#### Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

#### Consideraciones:

La invención consiste en una variante de la beta-glucosidasa de *Myceliophthora thermophila* C1 que presenta respecto a la enzima de SEQ ID nº 2, una mutación Q211H. La variante de la invención presenta actividad de transglicosilación reducida respecto a la enzima salvaje. La nueva variante se utiliza en procedimientos de degradación de materiales celulósicos con la finalidad de obtener azúcares fermentables y bioetanol.

**1. Documentos considerados.-**

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	WO 2011066457 A2 (CODEXIS LABORATORIES HUNGARY KFT)	03.06.2011
D02	WO 2011041594 A1 (CODEXIS, INC.)	07.04.2011
D03	BHAT, K.M. et al. "Purification and characterization of a n extracellular beta-glucosidase from the thermophilic fungus <i>Sporotrichum thermophile</i> and its influence on cellulase activity". JOURNAL OF GENERAL MICROBIOLOGY. 01.11.1993. Vol. 139, N°. 11, páginas 2825-2832.	
D04	KIM, T.J. et al. "Role of the glutamate 332 residue in the transglycosilation activity of <i>Thermus maltogenic amylase</i> ". BIOCHEMISTRY. 13.06.2000. Vol. 39, N°. 23, páginas 6773-6780.	

El documento D01, describe variantes de la  $\beta$ -glucosidasa de *Myceliophthora thermofila* C1 obtenidas por sustituciones de uno o más aminoácidos en posiciones diferentes de la posición Q211 de la variante de la presente solicitud. Las variantes obtenidas, presentan mayor termoestabilidad respecto a la enzima salvaje.

El documento D02, describe la secuencia corregida de la  $\beta$ -glucosidasa de *Myceliophthora thermofila* C1, respecto a la descrita previamente en el estado de la técnica. Las secuencias correctas se presentan en las SEQ ID n°s 3 y 4. Las correcciones en los péptidos señal, se encuentran resumidas en la página 9, Tabla 1.

El documento D03 describe la caracterización de una  $\beta$ -glucosidasa de *Sporotrichum thermophile* (nombre que se le daba anteriormente a *Myceliophthora thermofila*), que también cataliza una reacción de transglicosilación. Esta enzima es diferente de la  $\beta$ -glicosidasa de la invención (ver página 2830, discusión).

El documento D04, describe la importancia del E332 en las propiedades de transglicosilación de una amilasa de una bacteria del género *Thermus*. Los cambios E332Q y E332H, reducen de forma significativa la actividad de transglicosilación.

**2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración**

**NOVEDAD Y ACTIVIDAD INVENTIVA**

**Reivindicaciones 1-20**

No se ha encontrado en el estado de la técnica, ningún documento que describa una variante de la  $\beta$ -glucosidasa de *Myceliophthora thermofila* C1, que presenta capacidad de transglicosilación reducida. Los documentos tenidos en consideración más próximos al estado de la técnica, describen variantes de la  $\beta$ -glucosidasa localizados en posiciones diferentes al mutante de la invención y que afectan a la termoestabilidad (ver documentos D01 y D02). Los documentos D01-D04, solo muestran el estado general de la técnica y no se consideran de particular relevancia, ya que para una persona experta en la materia, no sería obvio aplicar las características de los documentos citados y llegar a la invención tal y como se menciona en las reivindicaciones 1-20. Por lo tanto, el objeto de la presente solicitud cumple los requisitos de novedad y actividad inventiva de acuerdo con los Artículos 6 y 8 de la Ley de Patentes 11/1986.