

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 906 870**

51 Int. Cl.:

**C07K 14/435** (2006.01)

**C12P 19/24** (2006.01)

**C12N 9/90** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.09.2018 PCT/EP2018/073747**

87 Fecha y número de publicación internacional: **07.03.2019 WO19043252**

96 Fecha de presentación y número de la solicitud europea: **04.09.2018 E 18762311 (1)**

97 Fecha y número de publicación de la concesión europea: **17.11.2021 EP 3679059**

54 Título: **Nuevas glucosa isomerasas mejoradas**

30 Prioridad:

**04.09.2017 EP 17189270**

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

**20.04.2022**

73 Titular/es:

**METGEN OY (100.0%)  
Rakentajantie 26  
20780 Kaarina, FI**

72 Inventor/es:

**BIRIKH, KLARA y  
SUONPÄÄ, ANU MINNA MAARET**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

ES 2 906 870 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Nuevas glucosa isomerasas mejoradas

## 5 Campo de la invención

La invención se encuentra en el campo de la enzimología. Más en particular, proporciona un método para la isomerización de glucosa en fructosa en el que la glucosa se deriva de material lignocelulósico. También proporciona una enzima con una mayor actividad de glucosa isomerasa en presencia de xilosa. Más en particular, la invención proporciona polipéptidos que codifican enzimas de glucosa isomerasa mutantes con actividad de glucosa isomerasa mejorada en comparación con la correspondiente enzima de tipo silvestre. Los polipéptidos divulgados son particularmente adecuados para convertir glucosa en fructosa en presencia de xilosa.

## 15 Antecedentes de la invención

El jarabe de maíz de alta fructosa (HFCS) (también llamado glucosa-fructosa, isoglucosa y jarabe de glucosa-fructosa) es un edulcorante hecho de almidón de maíz que ha sido procesado por una enzima llamada glucosa isomerasa (EC 5.3.1.5) para convertir algunas de sus glucosas en fructosa. El HFCS fue comercializado por primera vez a principios de la década de 1970 por Clinton Corn Processing Company, junto con el instituto de investigación japonés donde se descubrió la enzima.

Como edulcorante, el HFCS a menudo se compara con el azúcar granulada. Las ventajas del HFCS sobre el azúcar granulada incluyen que es más fácil de manejar y menos costoso en algunos países. En los EE. UU. el HFCS se encuentra entre los edulcorantes que reemplazaron principalmente a la sacarosa (azúcar de mesa) en la industria alimentaria.

En un proceso contemporáneo, el maíz (maíz) se muele para producir almidón de maíz y se utiliza un proceso de "ácido-enzima" en el que la solución de almidón de maíz se acidifica para comenzar a descomponer los carbohidratos existentes, y luego se agregan enzimas para metabolizar aún más el almidón y convertir los azúcares resultantes en fructosa.

La glucosa isomerasa es una enzima que convierte la glucosa en fructosa en una reacción reversible con un equilibrio de alrededor de 1:1 de glucosa a fructosa. La enzima se puede obtener de muchas especies diferentes de bacterias como *Streptomyces*, *Actinoplanes*, *Microbacterium* y *Bacillus*, y la enzima es o ha sido comercializada por compañías como Enzyme Bio-systems, Genencor, Gist-Brocades, Solvay Enzyme Inc y Novo Nordisk.

La mayoría de las glucosa isomerasas comerciales exitosas se inmovilizan en un soporte sólido y, como consecuencia, son muy estables con una vida media extremadamente larga. En un proceso típico, la isomerasa inmovilizada se carga en una columna y el sustrato (materia prima) pasa a través de ella a una velocidad que produce un efluente que contiene un 42 % de fructosa. Sin embargo, el requisito previo es que la materia prima sea un hidrolizado refinado que contenga 93-96 % de glucosa. Se requiere un refinado eficiente para eliminar las impurezas que podrían causar la inactivación de la glucosa isomerasa.

La glucosa también se puede obtener a partir de material lignocelulósico. El término "lignocelulosa" se refiere a la materia seca vegetal, denominada biomasa lignocelulósica. Es la fuente de carbono disponible más abundante en la tierra para la producción de biocombustibles, principalmente bioetanol y materiales potencialmente de base biológica, como polímeros y plásticos. Está compuesto por polímeros de carbohidratos (celulosa, hemicelulosa) y un polímero aromático (lignina). Estos polímeros de carbohidratos contienen diferentes monómeros de azúcar (azúcares de seis y cinco carbonos) y están fuertemente unidos a la lignina.

El uso de las glucosa isomerasas actualmente disponibles en la conversión de glucosa derivada de lignocelulosa en fructosa se ve obstaculizado por las impurezas que están presentes en la glucosa derivada de lignocelulosa. Estas impurezas conducen a una disminución significativa en la estabilidad de la enzima y causan costes significativos para la purificación de la materia prima.

Por lo tanto, para evitar etapas de purificación engorrosas y de costes elevados en la producción de fructosa a partir de material lignocelulósico, es deseable tener una isomerasa de glucosa que sea resistente a algunas o la mayoría, si no todas, las impurezas de la glucosa derivada de la lignocelulosa, más específicamente la lignina y xilosa.

En una solicitud en trámite conjunto (Solicitud de Patente Europea EP 16175234.0) identificamos una familia de glucosa isomerasas derivadas del género *Diktyoglomus* que demostraron ser resistentes frente a la disminución de la estabilidad provocada por la presencia de lignina.

La xilosa es un sustrato ampliamente preferido para la glucosa isomerasa, en comparación con la glucosa. Por lo tanto, la xilosa compite con la glucosa por la enzima y, por lo tanto, la tasa de conversión de glucosa se reduce en

gran medida. Por lo tanto, es deseable tener una enzima con una actividad de glucosa isomerasa que sea menos o nada inhibida por la xilosa.

Resumen de la invención

5 La presente invención divulga glucosa isomerasas que son resistentes a la inhibición por la presencia de xilosa en la mezcla de reacción y variantes de las mismas. Por lo tanto, la invención proporciona polipéptidos que codifican enzimas de glucosa isomerasa con una actividad de glucosa isomerasa aumentada o mejorada.

10 El término “actividad de glucosa isomerasa mejorada” o “actividad de glucosa isomerasa aumentada”, como se usa en el presente documento, se refiere a una enzima con una actividad de glucosa isomerasa superior en comparación con una enzima de control. En otras palabras, esto significa que la misma cantidad de enzima (expresada como masa de proteína) es capaz de convertir más glucosa en fructosa por minuto en comparación con una enzima de control.

15 Más en particular, la invención proporciona un polipéptido con actividad glucosa isomerasa que comprende una secuencia de aminoácidos que es al menos un 90 % idéntica a la secuencia de aminoácidos de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2, en el que el polipéptido comprende un diminuto residuo de aminoácido en una posición de aminoácido correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2.

20 Esta enzima presenta una actividad glucosa isomerasa mejorada en comparación con la enzima de control; es decir, en el que la enzima de control es la glucosa isomerasa de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2 en la que el aminoácido correspondiente a la posición 104 no es un aminoácido diminuto.

25 En una realización preferida, el término “aminoácido diminuto”, como se usa en el presente documento, se refiere a aminoácidos seleccionados del grupo que consiste en Glicina, Alanina, Serina y Cisteína.

La invención también se refiere a una composición que comprende un polipéptido como se describió anteriormente, un ácido nucleico que codifica un polipéptido como se describió anteriormente, un vector que comprende dicho ácido nucleico y una composición que comprende dicho ácido nucleico o un vector.

30 La invención también proporciona una célula huésped recombinante que comprende un ácido nucleico, un vector o una composición como se describió anteriormente.

35 Además, la invención se relaciona con un método para producir un polipéptido como se describió anteriormente, que comprende las etapas de: cultivar una célula huésped recombinante como se describió anteriormente, en condiciones adecuadas para la producción del polipéptido, y recuperar el polipéptido obtenido, y opcionalmente purificar el polipéptido.

40 Además, la invención se refiere a un método de uso de un polipéptido como se describió anteriormente para convertir glucosa en fructosa.

45 La invención también se refiere a un método para mejorar la conversión de glucosa en fructosa en presencia de xilosa mediante un polipéptido que comprende una secuencia de aminoácidos que es al menos un 90 % idéntica al aminoácido de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2, el método comprende la etapa de alterar el aminoácido en una posición correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2 a un diminuto residuo de aminoácido.

50 En el presente documento se ejemplifican glucosa isomerasas mejoradas que comprenden una secuencia de aminoácidos de acuerdo con SEQ ID NO: 1, en el que se han realizado sustituciones de un solo aminoácido para llegar a las glucosa isomerasas P104G, P104S, P104A y P104C. Esta anotación se usa aquí para indicar un reemplazo del residuo de aminoácido Prolina (P), correspondiente a la posición 104 de SEQ ID NO: 1, con cualquiera de los residuos G (Glicina), S (Serina), A (Alanina) o C (Cisteína), obteniendo así los polipéptidos de acuerdo con SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6, respectivamente.

55 También se ejemplifican sustituciones similares en una glucosa isomerasa que comprende la secuencia de aminoácidos de acuerdo con SEQ ID NO: 2, en el que se obtienen las glucosa isomerasas P104G, P104S, P104A y P104C, cuyas secuencias de aminoácidos están representadas por SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9 y SEQ ID NO: 10 respectivamente.

60 Secuencias de ADN que codifican secuencias de aminoácidos SEQ ID NO: 1 a SEQ ID NO: 10 están representados por SEQ ID NO: 11 a SEQ ID NO: 20

Leyenda de la figura

65 Figura 1: Diagrama que muestra la actividad glucosa isomerasa relativa de los polipéptidos que comprenden la secuencia de aminoácidos de acuerdo con SEQ ID NO: 1 (WT, anotado como GI1), SEQ ID NO: 2 (WT, anotado como

GI2) y SEQ ID NO: 3 (GI1/P104G), SEQ ID NO: 4 (GI1/P104S), SEQ ID NO: 5 (GI1/P104A), SEQ ID NO: 6 (GI1/P104C), SEQ ID NO: 7 (GI2/P104G), SEQ ID NO: 8 (GI2/P104S), SEQ ID NO: 9 (GI2/P104A), SEQ ID NO: 10 (GI2/P104C) en varios sustratos, véase el ejemplo 4.

5 La actividad de glucosa isomerasa se determinó usando un sustrato que contenía 200 mM de glucosa (etiquetado como "Glucosa") y se comparó con un sustrato que contenía 200 mM de glucosa y 20 mM de xilosa (etiquetado como "Glucosa xilosa", véase el Ejemplo 4), así como un hidrolizado de madera dura que comprende 42 xilosa mM (véase el ejemplo 6).

10 Descripción detallada de la invención

En enzimología, una glucosa isomerasa (EC 5.3.1.5) es una enzima que cataliza la interconversión de D-glucosa y D-fructosa. Esta enzima pertenece a la familia de las isomerasas, concretamente a aquellas oxidorreductasas intramoleculares que interconvierten aldosas y cetosas. La glucosa isomerasa ahora se ha observado en casi cien especies de bacterias. El nombre sistemático de esta clase de enzimas es D-xilosa aldosa-cetosa-isomerasa. Otros nombres de uso común incluyen D-xilosa isomerasa, D-xilosa cetoisomerasa y D-xilosa cetol-isomerasa. En la industria, estas enzimas se denominan principalmente glucosa-isomerasas debido a su uso industrial para producir fructosa a partir de glucosa. Sin embargo, la xilosa es un sustrato preferido para estas enzimas y la presencia de xilosa inhibe en gran medida la isomerización de la glucosa.

20 Las enzimas glucosa isomerasa disponibles en el mercado se han utilizado con éxito en la producción de jarabe de maíz con alto contenido de fructosa (HFCS) a partir de almidón, pero no son adecuadas para la isomerización de glucosa obtenida a partir de material lignocelulósico. Dicha glucosa derivada de lignocelulosa se caracteriza por la presencia de lignina y otros azúcares derivados de hemicelulosas, incluida la xilosa.

25 La biomasa lignocelulósica es la fuente de carbono disponible más abundante en la tierra para la producción de biocombustibles, principalmente bioetanol y potencialmente también para la producción de biomateriales como polímeros y plásticos. Está compuesto por polímeros de carbohidratos (celulosa, hemicelulosa) y un polímero aromático (lignina). La celulosa consiste en polímeros de glucosa lineales, mientras que la hemicelulosa es un polímero heterogéneo ramificado que consta de varios azúcares de 6 y 5 carbonos según la especie biológica.

30 La hemicelulosa derivada de la madera blanda contiene principalmente glucomananos, mientras que la hemicelulosa derivada de la madera dura contiene principalmente glucuronoxilanos. Por lo tanto, el hidrolizado de madera dura generalmente contiene una cantidad considerable de xilosa (Sjostrom, E., Wood Chemistry. Fundamentals and Applications. Second edition ed. 1993, San Diego: Academic press. 292.).

35 En nuestra Solicitud de Patente Europea en trámite conjunto EP 16175234.0 se demostró que las glucosa isomerasas GI1 y GI2 de tipo silvestre funcionan eficazmente en hidrolizados de madera blanda, donde son resistentes a la lignina y otras impurezas. Sin embargo, la madera dura presenta un obstáculo adicional, a saber, la xilosa. La xilosa es un sustrato muy preferido para este tipo de enzimas, y al ser la reacción reversible, la xilosa no se gasta durante el curso de la reacción y presenta una inhibición continua.

40 De hecho, aquí confirmamos que las glucosa isomerasas derivadas de *Dictyoglomus thermophilum* y *Dictyoglomus turgidum* de acuerdo con SEQ ID NO: 1 y SEQ ID NO: 2 eran adecuados para la conversión de glucosa en fructosa en una solución que contenía glucosa; sin embargo, estaban fuertemente inhibidos por la presencia de xilosa, como la presente en los hidrolizados de madera dura (figura 1).

45 Encontramos que las actividades GI1 y GI2 se redujeron en presencia de xilosa (es decir, en una solución de glucosa más xilosa), se observaron reducciones al 10-15 % de su actividad en glucosa pura. De acuerdo con estos hallazgos, la actividad de ambas glucosas isomerasas se redujo a menos del 12 % cuando se usó como sustrato un hidrolizado de madera dura (Tabla 1, figura 1).

50 La madera dura proviene de las angiospermas, o plantas con flores, como el abedul, el eucalipto, el roble, el arce o el nogal, que no son monocotiledóneas. Otros ejemplos de madera dura incluyen, entre otros, aliso, balsa, haya, nogal americano, caoba y teca.

55 Sorprendentemente, encontramos que una sola mutación en las glucosa isomerasas GI1 y GI2 de tipo silvestre en la posición 104 sustituyendo el aminoácido de tipo silvestre Prolina con un aminoácido diminuto como Glicina, Alanina, Serina o Cisteína, hace que la isomerización de la glucosa sea mucho menos inhibida por xilosa. Además, mejoró enormemente la actividad de la glucosa isomerasa a un valor entre 150 y 250 % (Tabla 1).

Tabla 1 Actividad glucosa isomerasa relativa de GI1 y GI2 en diferentes sustratos

	Glucosa [% de actividad]	Glucosa + Xilosa [% de actividad]	Madera dura [% de actividad]
GI1	100	15	12
GI1/P104G	250	240	150
GI1/P104S	245	230	140
GI1/P104A	230	215	120
GI1/P104C	150	120	90
GI2	95	12	10
GI2/P104G	240	230	130
GI2/P104S	230	220	110
GI2/P104A	220	200	100
GI2/P104C	145	100	80

En conclusión, la actividad de isomerasa de glucosa de tipo silvestre de las enzimas GI1 y GI2 en glucosa 200 mM se redujo en gran medida en presencia de xilosa 20 mM. Por el contrario, las enzimas glucosa isomerasa que portaban una de las mutaciones mencionadas anteriormente en la posición 104 eran más activas que las enzimas de tipo silvestre en presencia de xilosa (tabla 1, figura 1, ejemplo 4). Además, las enzimas con mutaciones en la posición 104 mostraron una actividad de isomerización de glucosa de 1.5 a 2.5 veces mayor en comparación con las enzimas de tipo silvestre con la misma dosis de enzima (figura 1).

A continuación, probamos las enzimas mutadas en la isomerización de glucosa en hidrolizados lignocelulósicos crudos de madera dura (ejemplo 6). El hidrolizado de madera dura, que contiene niveles más altos de xilosa que la madera blanda, inhibió en gran medida la actividad de la glucosa isomerasa de las enzimas de tipo silvestre GI1 y GI2, pero no la de las enzimas mutadas. Las enzimas mutadas demostraron una clara actividad de isomerasa de glucosa en los hidrolizados de madera dura, lo que los hace excepcionalmente adecuados para la conversión de glucosa en fructosa en material derivado de lignocelulosa, en particular cuando la lignocelulosa se deriva de madera como la madera con un alto contenido de xilosa, como la madera dura.

El término “enzima mutada”, como se usa en el presente documento, se refiere a una enzima glucosa isomerasa que comprende una secuencia de aminoácidos de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2, en el que el aminoácido en una posición correspondiente a la posición 104 de SEQ ID NO: 1, o SEQ ID NO: 2 (prolina en SEQ ID NO: 1 o SEQ ID NO: 2) se reemplaza por un diminuto aminoácido.

El término “aminoácido diminuto”, como se usa en el presente documento, indica un aminoácido con un grupo lateral diminuto. En otras palabras, estos aminoácidos son más pequeños que 110 Angstrom o 11 nanómetros. [http://www.imgt.org/IMGTEducation/Aide-memoire/\\_UK/aminoacids/IMGTclasses.html](http://www.imgt.org/IMGTEducation/Aide-memoire/_UK/aminoacids/IMGTclasses.html). Los ejemplos preferidos de aminoácidos diminutos son los aminoácidos G (Glicina), A (Alanina), S (Serina) y C (Cisteína).

Glucosa isomerasas (GI) de acuerdo con SEQ ID NO: 1 y SEQ ID NO: 2 son secuencias homólogas con una identidad de secuencia de aminoácidos del 98 %. Por lo tanto, cabe esperar que las indicaciones geográficas estrechamente relacionadas, como las indicaciones geográficas con una secuencia de aminoácidos de al menos el 90 %, como el 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o incluso 99 % idénticos con SEQ ID NO: 1 o SEQ ID NO: 2, funcionará de la misma manera que GI1 y GI2 aquí ejemplificados. Dichos homólogos cercanos pueden obtenerse de fuentes naturales o mediante mutagénesis dirigida. El experto en la materia conoce muy bien los materiales y métodos para obtener tales homólogos próximos.

Como se usa aquí, el grado de identidad entre dos o más secuencias de aminoácidos es equivalente a una función del número de posiciones idénticas compartidas por las secuencias (es decir, % de identidad = número de posiciones idénticas dividido por el número total de posiciones alineadas x 100), excluyendo espacios, que deben introducirse para una alineación óptima de las dos secuencias, y salientes. La comparación de secuencias y la determinación del porcentaje de identidad entre dos o más secuencias se puede lograr utilizando métodos estándar conocidos en la técnica. Por ejemplo, un software gratuito utilizado convencionalmente para este propósito es la herramienta [http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE\\_TYPE=BlastSearch&BLAST\\_SPEC=blast2seq&LINK\\_LOC=align2seq](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=align2seq).

La introducción de una mutación específica en un gen recombinante también se encuentra entre las habilidades de rutina de un biólogo molecular. Se puede obtener orientación específica de *Methods in Molecular Biology* Vol 182, “In vitro mutagenesis protocols”, Eds Jeff Braman, Humana Press 2002. Existen kits comercialmente disponibles para realizar mutagénesis dirigida al sitio (por ejemplo, kit de mutagénesis dirigida al sitio QuikChange II XL Agilent Technologies, No. de catálogo 200521).

Por tanto, la invención se refiere a un polipéptido con actividad glucosa isomerasa que comprende una secuencia de aminoácidos que es al menos un 90 % idéntica al aminoácido de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2, en el que el polipéptido comprende un diminuto residuo de aminoácido en una posición de aminoácido correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO:2.

La frase "una posición de aminoácido correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2" es en sí mismo suficientemente claro para el experto en la materia. Para evitar cualquier malentendido, se proporciona lo siguiente. La frase "una posición de aminoácido correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2" se usa aquí para indicar una determinada posición en la secuencia de aminoácidos del polipéptido con actividad de glucosa isomerasa. Esa cierta posición debe determinarse alineando la secuencia del polipéptido con actividad de isomerasa de glucosa con la secuencia de SEQ ID NO: 1 o SEQ ID NO: 2 como se describió anteriormente. La posición del aminoácido en el polipéptido con glucosa isomerasa que se alinea con el aminoácido en la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2 se denomina entonces la posición de aminoácido correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2.

La invención también se refiere a un método para la interconversión de D-glucosa y D-fructosa en presencia de una glucosa isomerasa, en el que la D-glucosa se deriva de biomasa que contiene lignocelulosa, y en el que la glucosa isomerasa es una glucosa isomerasa de acuerdo con la invención.

La frase "glucosa derivada de material que contiene lignocelulosa" es equivalente al término "glucosa derivada de lignocelulosa". Ambos se utilizan aquí para indicar que la glucosa está contenida en una solución que comprende otros azúcares, en particular xilosa, derivados del material lignocelulósico, como la biomasa lignocelulósica. Como tal, el término se utiliza para distinguir la glucosa de la glucosa purificada, que no contiene xilosa.

Las variantes mutadas de GI1 y GI2 como se divulga en este documento y sus homólogos con al menos un 90 % de identidad de secuencia proporcionan resultados ventajosos en comparación con otras GI en condiciones en las que la solución de sustrato comprende xilosa. Las enzimas mutantes no solo son resistentes frente a la presencia de xilosa en una composición que comprende glucosa y xilosa, sino que también son más activas, incluso hasta 2.5 veces más activas en la conversión de glucosa en fructosa.

En otros términos, la invención se relaciona con un proceso para convertir glucosa en fructosa que comprende las etapas de:

a) proporcionar una solución o suspensión que comprende glucosa y xilosa

b) convertir enzimáticamente la glucosa en fructosa en presencia de una isomerasa de glucosa,

c) opcionalmente purificar la fructosa de la solución,

en el que la glucosa isomerasa comprende una secuencia de aminoácidos que es al menos un 90 % idéntica a la secuencia de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2 y en el que la isomerasa de glucosa comprende un diminuto residuo de aminoácido en una posición de aminoácido correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2.

La solución o suspensión que comprende glucosa y xilosa se puede obtener ventajosamente por hidrólisis de una biomasa de madera dura. Tal hidrólisis se realiza ventajosamente enzimáticamente, por ejemplo empleando una celulasa.

Ventajosamente, la etapa de pretratamiento comprende una etapa de explosión de vapor y/o una etapa de pretratamiento con ácido.

Todas estas etapas son bien conocidos en la técnica y el experto es muy consciente de las medidas y límites de los términos utilizados en este documento.

La invención puede tener ventajas particulares cuando la enzima se usa en una solución o suspensión que contiene xilosa en una concentración que inhibe la actividad de la enzima de tipo silvestre de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2 para 10 % o más, como 20 %, 30 %, 40 %, 50 % o más, como 60 %, 70 %, 80 %, 90 % o incluso más como 95 % o más.

Los polipéptidos que se describen en el presente documento se pueden usar en composiciones que contienen varios componentes adicionales, como estabilizadores, rellenos, desechos celulares, medio de cultivo, etcétera. Por lo tanto, la invención proporciona una composición que comprende un polipéptido como se describe en este documento.

Los polipéptidos que se describen en el presente documento pueden obtenerse expresando un ADN recombinante en un sistema de expresión heterólogo. El término "sistema de expresión heterólogo" o equivalente significa un sistema

para expresar una secuencia de ADN de un organismo huésped en un organismo receptor de una especie o género diferente al del organismo huésped. Los receptores más frecuentes, conocidos como sistemas de expresión heterólogos, se eligen generalmente porque son fáciles de transferir ADN o porque permiten una evaluación más sencilla de la función de la proteína. Los sistemas de expresión heterólogos también se utilizan preferentemente porque permiten aumentar la producción de una proteína codificada por la secuencia de ADN en un proceso industrial. Los organismos receptores preferidos para usar como sistemas de expresión heterólogos incluyen organismos bacterianos, fúngicos y de levadura, como por ejemplo *Escherichia coli*, *Bacillus*, *Corynebacterium*, *Pseudomonas*, *Pichia pastoris*, *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, hongos filamentosos y muchos más sistemas bien conocidos en la técnica.

Los polipéptidos o proteínas divulgados en la presente pueden fusionarse con secuencias adicionales mediante la unión o la inserción, incluidos, entre otros, etiquetas de afinidad, que facilitan la purificación de proteínas (etiqueta S, dominio de unión a maltosa, dominio de unión a quitina), dominios o secuencias que ayudan al plegamiento (como el dominio de tiorredoxina, proteína SUMO), secuencias que afectan la localización de proteínas (señales de localización periplásmica, etc.), proteínas que tienen una función adicional, como la proteína fluorescente verde (GFP), o secuencias que representan otra actividad enzimática. Otros compañeros de fusión adecuados para los polipéptidos actualmente descritos son conocidos por los expertos en la técnica.

La presente invención también se refiere a polinucleótidos que codifican cualquiera de las variantes de glucosa isomerasa divulgadas en este documento. Los medios y métodos para clonar y aislar dichos polinucleótidos son bien conocidos en la técnica.

Además, la presente invención se refiere a un vector que comprende un polinucleótido de acuerdo con la invención, opcionalmente unido operativamente a una o más secuencias de control. Las secuencias de control adecuadas están fácilmente disponibles en la técnica e incluyen, pero no se limitan a, secuencias de promotor, líder, poliadenilación y señal.

Las variantes de glucosa isomerasa de acuerdo con diversas realizaciones de la presente invención pueden obtenerse mediante métodos recombinantes estándar conocidos en la técnica. Brevemente, dicho método puede comprender las etapas de: cultivar una célula huésped recombinante como se describió anteriormente en condiciones adecuadas para la producción del polipéptido y recuperar el polipéptido obtenido. A continuación, el polipéptido se puede purificar adicionalmente opcionalmente.

Se puede utilizar un gran número de sistemas vector-huésped conocidos en la técnica para la producción recombinante de las glucosa isomerasas como se describe en el presente documento. Los posibles vectores incluyen, pero no se limitan a, plásmidos o virus modificados que se mantienen en la célula huésped como molécula de ADN autónoma o integrada en el ADN genómico. El sistema de vector debe ser compatible con la célula huésped utilizada como es bien conocido en la técnica. Los ejemplos no limitantes de células huésped adecuadas incluyen bacterias (por ejemplo, *E. coli*, bacilos), levaduras (por ejemplo, *Pichia pastoris*, *Saccharomyces Cerevisiae*), hongos (por ejemplo, hongos filamentosos), células de insectos (por ejemplo, Sf9).

En otros términos, la invención se refiere a un método para mejorar la isomerización de glucosa, especialmente en presencia de xilosa, de un polipéptido con actividad de glucosa isomerasa que comprende una secuencia de aminoácidos que es al menos 90 % idéntica al aminoácido de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2, el método comprende la etapa de alterar el aminoácido en una posición correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2 a un diminuto residuo de aminoácido.

#### Ejemplos

Ejemplo 1: Preparación de polipéptidos de acuerdo con SEQ ID NO: 1 - 10 y nucleótidos de acuerdo con SEQ ID NO: 11 - 20.

Las construcciones de ADN de acuerdo con SEQ ID NO: 11 y SEQ ID NO: 12 que codifica los polipéptidos de acuerdo con SEQ ID NO: 1, SEQ ID NO: 2 se diseñaron utilizando frecuencias de codones optimizadas para la expresión en *E. coli* y se sintetizaron comercialmente y se clonaron en un vector de plásmido basado en un plásmido pET28a+ estándar. El vector de plásmido contenía una secuencia de nucleótidos N-terminal que codificaba la peptidil-prolil isomerasa de Enterobacteriaceae (número de acceso al banco de datos de proteínas WP\_000255997.1). El gen recombinante se expresó en *Escherichia coli* BL21(DE3) bajo el control del promotor T7-RNA-polimerasa (ver Ejemplo 2). Esto dio como resultado la expresión de las proteínas recombinantes con una etiqueta de terminal N. Secuencias de nucleótidos de acuerdo con SEQ ID NO: 13 - 20 que codifica las glucosa isomerasas de acuerdo con SEQ ID NO: 3 - 10 se ordenaron comercialmente y se expresaron de la misma manera que se describe en este documento.

Ejemplo 2: Expresión heteróloga de polipéptidos con actividad glucosa isomerasa.

La producción de proteínas se llevó a cabo en la cepa *E. coli* BL21 (DE3) de acuerdo con el protocolo del fabricante del plásmido disponible en <http://richsingiser.com/4402/Novagen%20pET%20system%20manual.pdf>. La temperatura

de incubación para la producción de proteínas fue de 30 grados centígrados, que se consideró óptima para obtener el máximo rendimiento de la proteína activa. Las células se lisaron utilizando tampón de lisis (Tris-HCl 50 mM, pH 7.4, Triton X100 al 1 %, CoCl<sub>2</sub> 1 mM) y se calentaron a 70 grados Celsius durante 30 min. La actividad de glucosa isomerasa se detectó solo en la fracción insoluble y pudo recuperarse completamente por centrifugación. Por lo tanto, la isomerasa de glucosa recombinante termoestable se expresó en forma insoluble activa, lo que permitió la reutilización de la enzima en varios lotes de reacción. Las mutaciones en la posición correspondiente a la posición 104 de GI1 o GI-2 no afectaron de manera detectable el nivel de expresión de las proteínas recombinantes, fueron esencialmente los mismos que los niveles de expresión de las enzimas de tipo silvestre que comprenden secuencias de aminoácidos de acuerdo con SEQ ID NO: 1 o SEQ ID NO:2.

Ejemplo 3: Ensayo de actividad de glucosa isomerasa.

La actividad de glucosa isomerasa (tasa de reacción de isomerización) se determinó midiendo el nivel de fructosa en la mezcla de reacción de acuerdo con el protocolo descrito en Schenk and Bisswanger, A microplate assay for D-xylose/D-glucose isomerase. Enzyme and Microbial Technology (Elsevier Science Inc, NY, 1998), V22, pp.721-723.

La medición se realizó en la etapa lineal del curso de la reacción (la acumulación de producto es lineal con el tiempo). Se tomaron alícuotas de diez microlitros de la mezcla de reacción y se pipetearon en una placa de 96 pocillos, se añadieron 40 ul de agua dando como resultado 50 ul de muestra. En algunos casos, se usó una mayor dilución de la mezcla de reacción con agua para preparar 50 ul de la muestra diluida para igualar el rango dinámico del método. Se añadieron 150 ml de una mezcla 1:1 (v/v) recién preparada de solución A (resorcinol al 0.05 % en etanol) y solución B (0.216 g FeNH<sub>4</sub>(SO<sub>4</sub>)<sub>2</sub> \* 12 H<sub>2</sub>O en 1 l de HCl concentrado). Para el desarrollo del color, la placa se incubó a 80 °C durante 40 min. La absorbancia se midió con un lector de microplacas (Thermo) a 490 nm.

Cabe señalar que la presencia de xilosa o su producto de isomerización xilulosa no afecta la medición de fructosa por este método.

Ejemplo 4 Actividad de isomerización de glucosa e inhibición de xilosa de polipéptidos que comprenden SEQ ID NO: 1 - 10 en solución de glucosa pura con y sin xilosa.

En este experimento, comparamos glucosa isomerasas de tipo silvestre: GI1 (SEQ ID NO: 1) y GI2 (SEQ ID NO: 2) a los mutantes de glucosa isomerasas P104G, P104S, P104A y P104C de ambas secuencias.

La actividad enzimática de la enzima de tipo silvestre (GI1 y GI2) y mutantes de la misma se determinó primero en una solución de glucosa (glucosa 200 mM, MOPS 10 mM pH 8.0, MgCl<sub>2</sub> 1 mM).

Un conjunto paralelo de reacciones tenía la misma composición aparte de comprender adicionalmente xilosa 20 mM. La dosificación de enzima se seleccionó de modo que durante el tiempo de reacción de 1h (a 70 grados C), la formación del producto permaneciera lineal en todas las reacciones. Todas las enzimas se usaron a la misma dosis (en microgramos de proteína recombinante por ml de reacción). La actividad de isomerización de glucosa se midió como se describe en el Ejemplo 3.

Los resultados se muestran en la tabla 1 y se representan en la Figura 1. La actividad de glucosa isomerasa de la enzima de tipo silvestre sin xilosa en la mezcla de reacción se tomó como 100 %, y las actividades en presencia de xilosa así como las actividades de enzimas mutadas en presencia o ausencia de xilosa se calcularon como porcentaje de este valor.

Las glucosa isomerasas mutadas de acuerdo con la invención mostraron una actividad de isomerización de glucosa mucho mayor (150-250 % del tipo silvestre). Además, se encontró que las glucosa isomerasas de acuerdo con la invención mostraban poca o ninguna sensibilidad a la presencia de xilosa.

Ejemplo 5 Preparación de hidrolizado de lignocelulosa

Se sumergieron astillas de madera, obtenidas de abedul (madera dura), en ácido sulfúrico al 2 % con un contenido de materia seca del 20 % y se sometieron a un pretratamiento de explosión de vapor esencialmente como se describe en la solicitud de patente europea. EP 2623607A1. El material pretratado en su totalidad (sin eliminar las fracciones solubilizadas de hemicelulosa y lignina) se sometió a hidrólisis enzimática usando el producto de celulasa Cellic® CTec3 de Novozymes. La hidrólisis se llevó a cabo en las condiciones recomendadas por el fabricante (incubación durante 72 h a 55 grados centígrados, pH 5.5 al 10 % de contenido de sólidos), los sólidos restantes se eliminaron de la mezcla de hidrólisis mediante centrifugación, luego la fracción líquida se evaporó hasta aproximadamente 100 g/L de concentración de azúcar, y el pH se ajustó a 8 con hidróxido de sodio. La solución resultante se denomina en adelante "hidrolizado" o "hidrolizado de lignocelulosa" y se usa para la reacción de isomerización. El hidrolizado de madera dura resultante contenía la siguiente composición de azúcar: 68 g/L de glucosa, 12 g/L de xilosa y menos de 1 g/L de otros azúcares

Ejemplo 6 Actividad de isomerización de glucosa e inhibición de xilosa de polipéptidos que comprenden SEQ ID NO: 1 - 10 en hidrolizados de lignocelulosa de madera dura.

5 En este experimento, comparamos glucosa isomerasas de tipo silvestre: GI1 (SEQ ID NO: 1) y GI2 (SEQ ID NO: 2) con variantes mutantes P104G, P104S, P104A y P104C de ambas glucosas isomerasas para la isomerización de glucosa a fructosa en hidrolizado crudo de madera dura.

10 Para este experimento, los hidrolizados de lignocelulosa (véase el ejemplo 5) se diluyeron con agua hasta una concentración final de glucosa 200 mM (dando como resultado una concentración de xilosa 42 mM) y se llevaron a MOPS 10 mM pH 8.0 y MgCl<sub>2</sub> 1 mM. Se agregaron enzimas a las mezclas de reacción en las mismas dosis que en el Ejemplo 4 y las reacciones se llevaron a cabo durante 1 hora a 70 grados C. Los resultados se muestran en la Figura 1. Para cada enzima, la actividad en solución de glucosa 200 mM sin xilosa como se describió anteriormente se tomó como 100 % y la actividad en hidrolizado de madera dura se representó gráficamente como un porcentaje de eso.

15 Ambas enzimas de tipo silvestre fueron fuertemente inhibidas en el hidrolizado de madera dura, como lo fueron en la solución de glucosa más xilosa descrita anteriormente. Sin embargo, las enzimas mutantes mostraban actividades más altas tanto en la solución modelo de glucosa más xilosa como en el hidrolizado de madera dura (tabla 1, figura 1).

20 LISTA DE SECUENCIAS

<110> MetGen OY

25 <120> NUEVAS GLUCOSA ISOMERASAS MEJORADAS

<130> 375 OT

<160> 20

30 <170> PatentIn versión 3.5

<210> 1

35 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

40 <400> 1

ES 2 906 870 T3

Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu  
 1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
 20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
 35 40 45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
 50 55 60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
 65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala  
 85 90 95

Gly Ile Arg Leu Val Ala Val Pro Leu Ser Leu Phe Tyr Asp Lys Ile  
 100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val  
 115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn  
 130 135 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser  
 145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val  
 165 170 175

ES 2 906 870 T3

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro  
180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp  
195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu  
210 215 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro  
225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr  
245 250 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser  
260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp  
275 280 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
290 295 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile  
305 310 315 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys  
355 360 365

<210> 2

5 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

10

<400> 2

Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu  
1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
20 25 30

ES 2 906 870 T3

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
 35 40 45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
 50 55 60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
 65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr  
 85 90 95

Gly Ile Arg Leu Val Ala Val Pro Leu Ser Leu Phe Tyr Asp Lys Ile  
 100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile  
 115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn  
 130 135 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser  
 145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val  
 165 170 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro  
 180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp  
 195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu  
 210 215 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro  
 225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr  
 245 250 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser  
 260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp  
 275 280 285

ES 2 906 870 T3

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
 290 295 300

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

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
 325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
 340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg  
 355 360 365

<210> 3

5 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

10

<400> 3

Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu  
 1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
 20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
 35 40 45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
 50 55 60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
 65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala  
 85 90 95

Gly Ile Arg Leu Val Ala Val Gly Leu Ser Leu Phe Tyr Asp Lys Ile  
 100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val  
 115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn  
 130 135 140

ES 2 906 870 T3

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser  
145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val  
165 170 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro  
180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp  
195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu  
210 215 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro  
225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr  
245 250 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser  
260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp  
275 280 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
290 295 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile  
305 310 315 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys  
355 360 365

<210> 4

5 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

10 <400> 4

ES 2 906 870 T3

Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu  
 1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
 20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
 35 40 45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
 50 55 60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
 65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala  
 85 90 95

Gly Ile Arg Leu Val Ala Val Ser Leu Ser Leu Phe Tyr Asp Lys Ile  
 100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val  
 115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn  
 130 135 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser  
 145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val  
 165 170 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro  
 180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp  
 195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu  
 210 215 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro  
 225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr  
 245 250 255

ES 2 906 870 T3

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser  
260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp  
275 280 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
290 295 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile  
305 310 315 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys  
355 360 365

<210> 5

5 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

10

<400> 5

Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu  
1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
35 40 45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
50 55 60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala  
85 90 95

Gly Ile Arg Leu Val Ala Val Ala Leu Ser Leu Phe Tyr Asp Lys Ile  
100 105 110

ES 2 906 870 T3

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

5 <210> 6

<211> 368

<212> PRT

10

<213> Dictyoglomus thermophilum

ES 2 906 870 T3

<400> 6

Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu  
1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
35 40 45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
50 55 60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala  
85 90 95

Gly Ile Arg Leu Val Ala Val Cys Leu Ser Leu Phe Tyr Asp Lys Ile  
100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val  
115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn  
130 135 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser  
145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val  
165 170 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro  
180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp  
195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu

ES 2 906 870 T3

210

215

220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro  
225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr  
245 250 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser  
260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp  
275 280 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
290 295 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile  
305 310 315 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys  
355 360 365

<210> 7

5 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

10

<400> 7

Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu  
1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
35 40 45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
50 55 60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu



ES 2 906 870 T3

<210> 8

<211> 368

5 <212> PRT

<213> Dictyoglomus thermophilum

<400> 8

10

Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu  
1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
35 40 45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
50 55 60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr  
85 90 95

Gly Ile Arg Leu Val Ala Val Ser Leu Ser Leu Phe Tyr Asp Lys Ile  
100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile  
115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn  
130 135 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser  
145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val  
165 170 175

ES 2 906 870 T3

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro  
180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp  
195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu  
210 215 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro  
225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr  
245 250 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser  
260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp  
275 280 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
290 295 300

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

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg  
355 360 365

<210> 9

5 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

10

<400> 9

Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu  
1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
20 25 30

ES 2 906 870 T3

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

ES 2 906 870 T3

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
 290 295 300

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

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
 325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
 340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg  
 355 360 365

<210> 10

5 <211> 368

<212> PRT

<213> Dictyoglomus thermophilum

10

<400> 10

Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu  
 1 5 10 15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser  
 20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro  
 35 40 45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu  
 50 55 60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu  
 65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr  
 85 90 95

Gly Ile Arg Leu Val Ala Val Cys Leu Ser Leu Phe Tyr Asp Lys Ile  
 100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile  
 115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn  
 130 135 140

ES 2 906 870 T3

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser  
145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val  
165 170 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro  
180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp  
195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu  
210 215 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro  
225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr  
245 250 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser  
260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp  
275 280 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr  
290 295 300

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

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg  
325 330 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro  
340 345 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg  
355 360 365

<210> 11

5 <211> 1104

<212> ADN

<213> Dictyoglomus thermophilum

10

<400> 11

ES 2 906 870 T3

atgccgtttg ttgatcatcg tgcacagaaa attcgtcgca gcaaagaaga actgctgaaa 60  
 catatgcaga ccttcaaact ggatctgaaa tttagcgtgg gcatctggta ttttacaccg 120  
 ggtggtggtc gttttcatga accgtatggt gaacagaaaa gcattccgga acgtattgaa 180  
 atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa 240  
 gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg 300  
 gttgcagttc cgctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
 ccgatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
 aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
 tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540  
 gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
 aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660  
 aaaaatccgg aaaacctgaa actgctgcaa gaagtcacg cactgggttg tctgaatccg 720  
 gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780  
 cgtgaaggtc gtctgtttca taccatttgg aatagccagc cgctgggtaa ttatgatcag 840  
 gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
 atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960  
 aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaactgctg 1020  
 ccgatgatc gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080  
 ctgattctgg ccgaaaatca taaa 1104

<210> 12

5 <211> 1104

<212> ADN

<213> Dictyoglomus thermophilum

10

<400> 12

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60  
 cacatgcaga cctttaaact ggatctgaaa tttagcgttg gcatctggta ttttaccctt 120  
 ggtggtggtc gttttcatga accgtatggt gaacagaaag gtattccgga acgtattgaa 180  
 atggcagcag aaatggcaaa atatggcgtt aaaggtatcg aagcacatta tccggctgaa 240  
 gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300  
 gttgcagttc cgctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
 ccgatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
 aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
 tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

ES 2 906 870 T3

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
aacatttatac gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660  
aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720  
gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780  
cgtgaaggtc gtctgtttca taccattgg aatagccagc cgctgggtaa ttatgatcag 840  
gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggtggt 960  
aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020  
ccgcatgatac gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080  
ctgattctgg cggaaaacca tcgt 1104

<210> 13

5 <211> 1104

<212> ADN

<213> Dictyoglomus thermophilum

10

<400> 13

atgccgtttg ttgatcatcg tgcacagaaa attcgtogca gcaaagaaga actgctgaaa 60  
catatgcaga ccttcaaact ggatctgaaa ttttagcgtgg gcatctggta ttttacaccg 120  
ggtggtggtc gttttcatga accgtatggt gaacagaaaa gcattccgga acgtattgaa 180  
atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa 240  
gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg 300  
gttgacgttg gtctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
ccgatgaaa aatatacgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
aaagaagcaa acgccgatatac ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540  
gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
aacatttatac gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660  
aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720  
gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780  
cgtgaaggtc gtctgtttca taccattgg aatagccagc cgctgggtaa ttatgatcag 840  
gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
atggttggtt atcagggcta ttttggcatc gatataatc cggaacgcat gccggttatt 960  
aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020  
ccgcatgatac gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080  
ctgattctgg cggaaaatca taaa 1104

15

<210> 14

ES 2 906 870 T3

<211> 1104

<212> ADN

5 <213> Dictyoglomus thermophilum

<400> 14

atgccgtttg	ttgatcatcg	tgacacagaaa	attcgtcgca	gcaaagaaga	actgctgaaa	60
catatgcaga	ccttcaaact	ggatctgaaa	tttagcgtgg	gcatctggta	ttttacaccg	120
ggtggtggtc	gttttcatga	accgtatggt	gaacagaaaa	gcattccgga	acgtattgaa	180
atggcagcag	aaatggcaaa	at ttggcgtg	aaaggtattg	aagcacatta	tccggctgaa	240
gtgaatgaag	aaaatctgca	cctgtataaa	cagctggaaa	aagaagcagg	tattcgtctg	300
gttgacagttt	ctctgagcct	gttttatgat	aaaatctttg	aatttggcag	cctgagcaac	360
ccgtatgaaa	aatatcgtaa	agttgcctat	gaaccgctgg	tgaatggtct	gaaactggtt	420
aaagaagcaa	acgccgatat	ttgcattatt	tgccctggta	ttgatggcta	tacctatagc	480
tatggtcacc	tgtattatca	catgtgggat	accttgaag	aactggttgc	acaggcaatg	540
gatgaagttc	cgggtgttca	ggttgcaatt	gaaccgaaac	cgtatgaacc	ggcaccgaat	600
aacatttatc	gtaccaccgc	agatggtatt	ctggcagcac	gtgatattga	agcgcgtctg	660
aaaaatccgg	aaaacctgaa	actgctgcaa	gaaggtcacg	cactggttgg	tctgaatccg	720
gaagttggtc	atgttcgtat	gggttttgaa	gatctgccgt	atgcatatgc	ccgtgttgca	780
cgtgaaggtc	gtctgtttca	taccattggg	aatagccagc	cgctgggtaa	ttatgatcag	840
gatctgaata	ttggtgtggt	ggattgggat	agcaccgaag	cactgctgta	taccctgaaa	900
atggttggtt	atcagggcta	ttttggcatc	gatatcaatc	cggaacgcat	gccggttatt	960
aaagccattg	aaattaacac	caaagtctg	cagattatga	acgaacgcat	tgaacgtctg	1020
ccgcatgatc	gtattattga	gtgttatttt	gaccctgaga	atcatcgtgg	tgaactggaa	1080
ctgattctgg	ccgaaaatca	taaa				1104

10

<210> 15

<211> 1104

15 <212> ADN

<213> Dictyoglomus thermophilum

<400> 15

20

atgccgtttg	ttgatcatcg	tgacacagaaa	attcgtcgca	gcaaagaaga	actgctgaaa	60
catatgcaga	ccttcaaact	ggatctgaaa	tttagcgtgg	gcatctggta	ttttacaccg	120
ggtggtggtc	gttttcatga	accgtatggt	gaacagaaaa	gcattccgga	acgtattgaa	180
atggcagcag	aaatggcaaa	at ttggcgtg	aaaggtattg	aagcacatta	tccggctgaa	240
gtgaatgaag	aaaatctgca	cctgtataaa	cagctggaaa	aagaagcagg	tattcgtctg	300

ES 2 906 870 T3

gttgcagttg ctctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
 ccgatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
 aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
 tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540  
 gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
 aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660  
 aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720  
 gaagttggtc atgttcgfat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780  
 cgtgaaggtc gtctgtttca taccattgg aatagccagc cgctgggtaa ttatgatcag 840  
 gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
 atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960  
 aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020  
 ccgcatgacg gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080  
 ctgattctgg ccgaaaatca taaa 1104

<210> 16

5 <211> 1104

<212> ADN

<213> Dictyoglomus thermophilum

10

<400> 16

atgccgtttg ttgatcatcg tgcacagaaa attcgtcgca gcaaagaaga actgctgaaa 60  
 catatgcaga ccttcaaact ggatctgaaa tttagcgtgg gcatctggta ttttacaccg 120  
 ggtggtggtc gttttcatga accgtatggt gaacagaaaa gcattccgga acgtattgaa 180  
 atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa 240  
 gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg 300  
 gttgcagttt gcctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
 ccgatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
 aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
 tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540  
 gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
 aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660  
 aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720  
 gaagttggtc atgttcgfat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780  
 cgtgaaggtc gtctgtttca taccattgg aatagccagc cgctgggtaa ttatgatcag 840

ES 2 906 870 T3

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960  
aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020  
ccgcatgacg gtattattga gtggtatttt gaccctgaga atcatcgtgg tgaactggaa 1080  
ctgattctgg ccgaaaatca taaa 1104

<210> 17

5 <211> 1104

<212> ADN

<213> Dictyoglomus thermophilum

10

<400> 17

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60  
cacatgcaga cctttaaact ggatctgaaa tttagcgttg gcatctggta ttttaccctt 120  
ggtggtggtc gttttcatga accgtatggt gaacagaaag gtattccgga acgtattgaa 180  
atggcagcag aaatggcaaa atatggcgtt aaaggtatcg aagcacatta tccggctgaa 240  
gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300  
gttgacgttg gtctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
ccgatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540  
gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660  
aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactgggttg tctgaatccg 720  
gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780  
cgtgaaggtc gtctgtttca taccattggg aatagccagc cgctgggtaa ttatgatcag 840  
gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960  
aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020  
ccgcatgacg gtattattga gtggtatttt gaccccgaaa atcatcgtgg tgaactggaa 1080  
ctgattctgg ccgaaaacca tcgt 1104

15 <210> 18

<211> 1104

<212> ADN

20

<213> Dictyoglomus thermophilum

<400> 18

ES 2 906 870 T3

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60  
cacatgcaga cctttaaact ggatctgaaa tttagcggtg gcatctggta ttttaccct 120  
ggtggtggtc gttttcatga accgtatggt gaacagaaag gtattccgga acgtattgaa 180  
atggcagcag aaatggcaaa atatggcggt aaaggtatcg aagcacatta tccggtgaa 240  
gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300  
gttgcagttt ctctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
ccgtatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540  
gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660  
aaaaatccgg aaaacctgaa actgctgcaa gaagtcacg cactggttgg tctgaatccg 720  
gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780  
cgtgaaggtc gtctgtttca taccatttgg aatagccagc cgctgggtaa ttatgatcag 840  
gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960  
aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaactgctg 1020  
ccgcatgacg gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080  
ctgattctgg cggaaaacca tcgt 1104

<210> 19

5 <211> 1104

<212> ADN

<213> Dictyoglomus thermophilum

10

<400> 19

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60  
cacatgcaga cctttaaact ggatctgaaa tttagcggtg gcatctggta ttttaccct 120  
ggtggtggtc gttttcatga accgtatggt gaacagaaag gtattccgga acgtattgaa 180  
atggcagcag aaatggcaaa atatggcggt aaaggtatcg aagcacatta tccggtgaa 240  
gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300  
gttgcagttg ctctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
ccgtatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

ES 2 906 870 T3

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
aacatttatac gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660  
aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720  
gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780  
cgtgaaggtc gtctgtttca taccattgg aatagccagc cgctgggtaa ttatgatcag 840  
gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960  
aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020  
ccgcatgatac gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080  
ctgattctgg cgaaaaacca tcgt 1104

<210> 20

5 <211> 1104

<212> ADN

<213> Dictyoglomus thermophilum

10

<400> 20

atgccgtttg tggatcatcg taatcagaaa attcgtogca gcaaagaaga actgctgaaa 60  
cacatgcaga cctttaaact ggatctgaaa ttttagcgttg gcatctggta ttttaccocct 120  
ggtggtggtc gttttcatga accgtatggt gaacagaaag gtattccgga acgtattgaa 180  
atggcagcag aaatggcaaa atatggcggt aaaggtatcg aagcacatta tccggctgaa 240  
gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300  
gttgacggtt gcctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360  
ccgatgaaa aatatacgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420  
aaagaagcaa acgccgatatac ttgcattatt tggcctggta ttgatggcta tacctatagc 480  
tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540  
gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600  
aacatttatac gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660  
aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720  
gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780  
cgtgaaggtc gtctgtttca taccattgg aatagccagc cgctgggtaa ttatgatcag 840  
gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900  
atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960  
aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020  
ccgcatgatac gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080  
ctgattctgg cgaaaaacca tcgt 1104

15

**REIVINDICACIONES**

- 5 1. Un polipéptido con actividad de glucosa isomerasa que comprende una secuencia de aminoácidos que es al menos un 90 % idéntica a la secuencia de aminoácidos de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2, en el que el polipéptido comprende un residuo de Glicina, Serina, Alanina o Cisteína en una posición correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2.
- 10 2. Polipéptido de acuerdo con la reivindicación 1 que comprende un residuo de Glicina en una posición correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2.
3. Polipéptido de acuerdo con la reivindicación 1 o 2 que comprende una secuencia de aminoácidos que es al menos un 90 % idéntica al aminoácido de acuerdo con SEQ ID NO: 1, tal como al menos 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, o 99 %.
- 15 4. Polipéptido de acuerdo con una cualquiera de las reivindicaciones 1 - 3, en el que el polipéptido es un polipéptido aislado.
5. Composición que comprende un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1 - 4.
- 20 6. Ácido nucleico que codifica un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1 - 4.
7. Vector que comprende un ácido nucleico de acuerdo con la reivindicación 6.
- 25 8. Composición que comprende un ácido nucleico de acuerdo con la reivindicación 6 o un vector de acuerdo con la reivindicación 7.
9. Célula huésped recombinante que comprende un ácido nucleico de acuerdo con la reivindicación 6, un vector de acuerdo con la reivindicación 7 o una composición de acuerdo con la reivindicación 8.
- 30 10. Célula huésped recombinante de acuerdo con la reivindicación 9 seleccionada del grupo que consiste en *Escherichia coli*, *Bacillus*, *Corynebacterium*, *Pseudomonas*, *Pichia pastoris*, *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, hongos filamentosos, levaduras y células de insectos.
- 35 11. Método para producir un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1 - 4, que comprende las etapas de:
- a. cultivar una célula huésped recombinante de acuerdo con la reivindicación 9 o 10 en condiciones adecuadas para la producción del polipéptido, y
- 40 b. recuperar el polipéptido obtenido, y
- c. purificar opcionalmente dicho polipéptido.
- 45 12. Uso de un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1 - 4 o una composición de acuerdo con la reivindicación 5 para la conversión de glucosa en fructosa, en el que la glucosa se deriva de biomasa lignocelulósica.
13. Uso de acuerdo con la reivindicación 12 en el que la biomasa contiene xilano o glucuronoxilano.
- 50 14. Uso de acuerdo con la reivindicación 12 o 13 en el que la biomasa de lignocelulosa es madera, tal como madera dura.
- 55 15. Método para aumentar la actividad de glucosa isomerasa de un polipéptido que comprende una secuencia de aminoácidos que es al menos un 90 % idéntica al aminoácido de acuerdo con SEQ ID NO: 1 o SEQ ID NO: 2, el método comprende la etapa de reemplazar el aminoácido en una posición correspondiente a la posición 104 en SEQ ID NO: 1 o SEQ ID NO: 2 con un residuo de Glicina, Serina, Alanina o Cisteína.

