

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 643 046**

51 Int. Cl.:

C12N 9/90 (2006.01)

C12P 7/06 (2006.01)

C12P 19/24 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.02.2013 PCT/EP2013/052407**

87 Fecha y número de publicación internacional: **15.08.2013 WO13117631**

96 Fecha de presentación y número de la solicitud europea: **07.02.2013 E 13706186 (7)**

97 Fecha y número de publicación de la concesión europea: **06.09.2017 EP 2812430**

54 Título: **Microorganismos fermentadores de pentosa**

30 Prioridad:

07.02.2012 EP 12000783

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.11.2017

73 Titular/es:

**CLARIANT PRODUKTE (DEUTSCHLAND) GMBH
(100.0%)
Brüningstrasse 50
65929 Frankfurt am Main, DE**

72 Inventor/es:

**DRAGOVIC, ZDRAVKO;
GAMAUF, CHRISTIAN;
REISINGER, CHRISTOPH y
KETTLING, ULRICH**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 643 046 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

Microorganismos fermentadores de pentosa

Campo de la invención

5 La xilosa es un componente fundamental de la biomasa vegetal, y se encuentra enlazada en una serie de materias primas en el enfoque de los conceptos actuales de biorefinería. Ejemplos de tales materiales ricos en xilosa incluyen paja de trigo, hojas de maíz o astillas de madera u otros subproductos de madera (Blake A Simmons y colaboradores. *Genome Biol.* 2008; 9 (12): 242).

10 Como consecuencia una hidrólisis llevada a cabo del material de partida mediante procedimientos enzimáticos, químicos o químico/enzimáticos conduce a productos intermedios ricos en xilosa, además de otros azúcares valiosos (Deepak Kumar y colaboradores. *Biotechnol Biocombustibles.* 2011; 4: 27). La utilización eficiente de soluciones de azúcares ricos en C5 en líneas de fermentación acopladas es a la vez crucial y exigente para las cepas de fermentación aplicadas (Sara Fernandes y Patrick Murray, *Bioeng Bugs.* 2010; 1(6): 424). Especialmente las levaduras C6, tales como *Saccharomyces cerevisiae*, que son caballos de trabajos deseados debido a la larga historia de cría que terminó en rasgos con tolerancia extrema al etanol y altos rendimientos para la conversión de glucosa, dejan la xilosa completamente intacta, disminuyendo así el rendimiento potencial. Se conocen varias estrategias para evitar esta limitación. Una etapa clave aquí parece la alimentación exitosa de la xilosa por isomerización en xilulosa y la modificación posterior en cascada de la parte no reductora del derivado C5 en la ruta de glicolisis regular de *Saccharomyces cerevisiae*. Mientras que la fuerza de absorción a través de la membrana por transportadores específicos y la densidad de flujo alcanzable a través del derivado C5 están sujetas a posibles mejoras (David Runquist y colaboradores. *Microb Cell Fact.* 2009; 8: 49), la etapa de isomerización clave de xilosa a xilulosa plantea un problema importante en el proceso global. Se conocen dos rutas principales para llevar a cabo esta etapa. La primera, que emplea etapas posteriores de reducción a xilitol (mediante xilosa reductasa) y oxidación (mediante xilitol deshidrogenasa) a xilulosa, provoca un desequilibrio importante entre los cofactores NADH y NADPH y conduce a un aumento de la formación de xilitol bajo condiciones de fermentación (Maurizio Bettiga y colaboradores. *Biotechnol Biocombustibles.* 2008; 1: 16). La isomerización directa alternativa mediante la aplicación de xilosa isomerasa sufre de la falta de disponibilidad de genes de xilosa isomerasa que combinan una expresión activa en microorganismos eucarióticos (en particular levaduras como *Saccharomyces cerevisiae*), una eficacia catalítica elevada, una temperatura y un pH óptimos adaptados a la temperatura de fermentación y una baja inhibición por productos secundarios, especialmente xilitol. Un aspecto de la presente invención es la divulgación de una célula de levadura que expresa una proteína que presenta actividad de xilosa isomerasa en una célula de levadura, para cumplir este requisito.

35 La ruta de la xilosa isomerasa es originaria de especies bacterianas y de levaduras raras. En contraste con la ruta de la oxidoreductasa la ruta de la isomerasa no requiere cofactores. La ruta de la isomerasa consiste mínimamente en enzimas individuales, xilosa isomerasas heterólogas (XI), que convierten directamente xilosa a xilulosa. Como con la ruta de la oxidoreductasa, la mejora adicional del rendimiento se puede obtener por coexpresión de la xilulosa quinasa heteróloga (XK). El primer XI funcionalmente expresado fue un gen *xylA* del hongo anaeróbico *Piromyces* sp E2 (Kuyper M. y colaboradores. *FEMS Yeast Res.* 2003; 4(1): 69). Se construyó la cepa de levadura haploide con capacidad para fermentar xilosa como única fuente de carbono bajo condiciones anaeróbicas. La mayoría de las xilosa isomerasas son proteínas bacterianas y el principal obstáculo fue su expresión en levaduras. Sin embargo, trabajos recientes han demostrado la expresión funcional en la levadura (Tabla 1). Debido a la importancia clave de la actividad de la xilosa isomerasa dentro del concepto de los organismos fermentadores de C5, es deseable utilizar xilosa isomerasas óptimas. De los informes anteriores se deduce que la xilosa isomerasa de *Clostridium phytofermentans* proporciona un estándar técnico bajo pero más alto disponible con este respecto. Las propiedades beneficiosas mejoradas de xilosa isomerasas expresadas por una célula de levadura en el alcance de esta invención son por tanto altamente deseadas.

Muzny y colaboradores describen una cepa de xilosa isomerasa dentro de la entrada de datos del banco de genes XP-002681255 que es casi idéntica al *Eubacterium saburreum* de tipo nativo DSM3986.

50 Brat y colaboradores, Functional Expression of a Bacterial Xylose Isomerase in *Saccharomyces cerevisiae*, *Applied and Environmental Microbiology*, Apr. 2009, p. 2304-2311, describen maneras para clonar y expresar nuevas especies de XI de *Clostridium phytofermentans* en *S. cerevisiae*. Brat y colaboradores mostraron que las células de levadura tenían la capacidad para metabolizar D-xilosa y utilizarla como la única fuente de carbono y energía.

55 Dentro del documento de Patente WO2012/009272 se describe la fermentación de pentosa por un microorganismo recombinante. El documento de Patente WO2012/009272 describe además construcciones de ácido nucleico recombinante que comprenden un polinucleótido de xilosa isomerasa, una célula huésped de hongo recombinante que comprende un plinucleótido de xilosa isomerasa recombinante y métodos relacionados.

El documento de Patente WO2011/078262 describe células eucarióticas capaces de asimilar xilosa y métodos para la transformación de células eucarióticas tales como levaduras utilizando DNA que codifica xilosa isomerasa procedentes de termitas protistas y estas nuevas células eucarióticas capaces de asimilar xilosa.

Tabla 1: Ejemplos para xilosa isomerasas para la aplicación en levaduras

Organismo de origen	Referencia
<i>Clostridium phytofermentas</i>	WO 2010/000464
<i>Piromyces sp. E2</i>	EP 1 468093 B1
<i>Bacteroides thetaiotaomicron</i>	US 2012/0225451A1
Desconocido	WO2011078262
<i>Abiotrophia defectiva</i>	WO/2012/009272

5 El transporte de azúcar a través de la membrana no se limita la fermentación de los azúcares de hexosa, aunque puede limitar el metabolismo de la pentosa especialmente en el caso de cofermentaciones de hexosa y pentosa. Se han realizado varios estudios de expresión del transportador de pentosa.

Breve descripción de la invención

10 Un objetivo de la invención es proporcionar una célula de levadura capaz de utilizar azúcares C5, en particular xilosa. Se encontró sorprendentemente que la proteína descrita por SEQ ID NO. 2 (secuencia publicada previamente en el número de acceso al banco de genes del NCBI ZP_07904696.1) o una versión truncada en el extremo N desprovista de los primeros 18 aminoácidos (mktknniicialkgdif) (SEQ ID NO. 8) se expresa funcionalmente en células microbianas eucarióticas, en particular levaduras como *Saccharomyces cerevisiae*, cuando estas células se transforman con un vector que lleva un casete de expresión que comprende una secuencia de DNA que codifica dicha proteína SEQ ID NO. 2, por ejemplo la molécula la molécula de DNA descrita en SEQ ID NO. 1 (publicada previamente como parte del número de acceso al banco de genes NZ_AEPW01000073.1 GI:315651683).

15 La presente invención proporciona también una célula de levadura que expresa una proteína que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad, lo más preferiblemente un 90% de identidad, más altamente preferible un 95% de identidad con SEQ ID NO. 2 o SEQ ID NO.8 y que tiene actividad de xilosa isomerasa en una célula eucariótica.

20 Se encontró además que las células transformadas muestran una mayor velocidad de consumición de xilosa cuando se compara con las células no transformadas.

Como un aspecto adicional, se mejoró la fermentación de la biomasa a partir de medios que contenían fuente de carbono de xilosa y se incrementó la cantidad de metabolitos formados por dichas cepas transformadas en estas condiciones en comparación con los controles transformados.

25 Otro aspecto se refiere a las propiedades biocatalíticas de la proteína expresada y su aplicación como biocatalizador *in situ* o en forma purificada para la producción de productos de azúcar o intermedios isomerizados.

Figuras

Figura 1: Ruta de utilización de xilosa.

30 Figura 2: Comparación del crecimiento de las colonias entre *Saccharomyces cerevisiae* transformado con *Eubacterium saburreum* (Es XI), *Piromyces sp.* (Pi XI) y *Clostridium phytofermentas* (Cp XI). Como control negativo se utilizó la cepa transformada con el vector de expresión simple pSCMB454 (Vector). Un 1 en la escala corresponde al crecimiento de la semana después de 6 días mientras un 4 corresponde a un crecimiento muy fuerte después de 6 días.

Figura 3: Mapa del plásmido de expresión de levadura

Figura 4: Plásmido de expresión para EsXI.

35 Figura 5: Comparación del crecimiento del cultivo entre *Saccharomyces cerevisiae* transformada con *Eubacterium saburreum* (Es-sh XI), y *Clostridium phytofermentas* (Cp XI). Como control negativo se utilizó la cepa transformada con el vector de expresión simple pSCMB454 (Vector).

40 Figura 6: Actividad de la xilosa isomerasa en extractos celulares de *Saccharomyces cerevisiae* que expresan *Eubacterium saburreum* (Es-sh XI, diamantes), y *Clostridium phytofermentas* (Cp XI, círculos). Como control negativo se utilizó la cepa transformada con el vector de expresión simple pSCMB454 (Vector). Las fórmulas para ajustes de curvas lineales ($Abs_{340nm} = V * \text{Tiempo} + Abs_{3min}$) se muestran debajo de la leyenda. Por razones de claridad solo se trazó el ajuste de la curva para el control negativo (Vector).

Figura 7: Actividad específica de xilasas isomerasas purificadas: *Eubacterium saburreum* Es-sh XI y *Clostridium phytofermentas* Cp XI. Como control negativo se utilizó la mezcla de reacción sin enzima (tampón solamente). La actividad específica se expresa como % de relación de xilosa convertida a enzima a sustrato (E/S) de 0,05% peso/peso.

- 5 Figura 8: Determinación del pH óptimo para las xilasas isomerasas purificadas: *Eubacterium saburreum* (Es-sh XI, diamantes), y *Clostridium phytofermentas* (Cp XI, círculos). Como control negativo se utilizó la mezcla de reacción sin enzima (Buff, triángulos). La actividad se expresa como % de la actividad máxima.

Figura 9: Determinación de la temperatura óptima para xilasas isomerasas purificadas: *Eubacterium saburreum* (Es-sh XI, diamantes), y *Clostridium phytofermentas* (Cp XI, círculos). Como control negativo se utilizó la mezcla de reacción sin enzima (Buff, triángulos). La actividad se expresa como % de la actividad máxima.

Figura 10: Determinación de Km para xilasas isomerasas purificadas: *Eubacterium saburreum* (Es-sh XI), y *Clostridium phytofermentas* (Cp XI).

Descripción detallada

Definiciones

15 La actividad de la xilosa isomerasa se define aquí como la actividad enzimática de un enzima que pertenece a la clase de xilosa isomerasas (EC 5.3.1.5), catalizando así la isomerización de diversos azúcares de aldosa y cetosa y otras reacciones secundarias enzimáticas inherentes a esta clase de enzimas. La asignación de una proteína a la clase de xilosa isomerasa se lleva a cabo según el patrón de actividad o consideraciones de homología, lo que sea más relevante en cada caso. La actividad de la xilosa isomerasa se puede determinar mediante el uso de un ensayo fotométrico enzimático acoplado que utiliza sorbitol deshidrogenasa.

Una construcción de expresión se define aquí como una secuencia de DNA que comprende todos los elementos de secuencia necesarios para establecer la expresión de un marco de lectura abierto (ORF) comprendido en la célula huésped que incluye secuencias para el inicio de la transcripción (promotores), terminación y regulación, sitios para el inicio de la traducción, regiones para la replicación estable o integración en el genoma huésped y un marcador genético seleccionable. De este modo la configuración funcional puede establecerse o alcanzarse mediante la organización del suceso (integración, etc.) en la célula huésped. En una realización preferida la construcción de expresión contine un promotor funcionalmente unido al marco de lectura abierto seguido por una secuencia de terminación opcional. Las secuencias reguladoras para la expresión en células eucarióticas comprenden secuencias promotoras, sitios de unión al factor de regulación de la transcripción, secuencias para el inicio de la traducción y secuencias terminadoras. Las secuencias reguladoras para la expresión en células eucarióticas se entienden como regiones codificadas por DNA o RNA que permanecen en conexión funcional con el proceso de transcripción y/o traducción de las cadenas de DNA codificantes en células eucarióticas, cuando se encuentran conectadas a cadenas de DNA codificantes solas o en combinación con otras secuencias reguladoras. En el enfoque están las secuencias promotoras acopladas a los genes de la xilosa isomerasa de la invención permitiendo así su expresión en una levadura eucariótica o una célula fúngica seleccionada. La combinación del promotor eucariótico y las secuencias de DNA que codifican la xilosa isomerasa conduce a la expresión de la xilosa isomerasa en la célula de levadura transformada. Los promotores preferidos son promotores de *Sacharomyces cerevisiae*, de intensidad media a elevada, activos bajo condiciones fermentativas. Ejemplos para dichos promotores preferidos son promotores de la ruta glicolítica o del transporte de azúcar, particularmente los promotores de los genes conocidos como PFK1, FBA1, PGK1, ADH1, ADH2, TDH3 así como las variantes truncadas o mutadas de los mismos. Los elementos para el establecimiento de la estabilidad mitótica son conocidos en la técnica y comprenden el origen de replicación del plásmido *S. cerevisiae* 2 μ , las secuencias centroméricas (CEN), la secuencia de replicación autónoma (ARS) o secuencias homólogas de cualquier longitud para la promoción de la integración cromosómica a través de la ruta de unión del cromosoma homólogo. Los marcadores seleccionables incluyen elementos genéticos que atribuyen resistencia antibiótica a la célula huésped. Ejemplos son los genes marcadores *kan* y *ble*. Se pueden utilizar marcador de auxotrofia que complementan las auxotrofias definidas de la cepa huésped. Ejemplos de tales marcadores a mencionar son genes y mutaciones que reflejan la ruta de la leucina (LEU2) o uracilo (URA3), pero también xilosa isomerasa.

El consumo de xilosa mejorado se define aquí como cualquier velocidad de consumo de xilosa que da como resultado el crecimiento y proliferación celular, la formación de metabolitos y/o la generación de energía calórica que aumenta comparado con la velocidad de consumo de xilosa de células (cultivos) no modificadas con respecto al rasgo considerado. La velocidad de consumo se puede determinar por ejemplo fenomenológicamente considerando la densidad de células formadas o el tamaño de las colonias, por determinación de la velocidad de consumo de oxígeno, velocidad de formación de etanol o por medida directa de la concentración de xilosa en el medio de crecimiento a lo largo del tiempo. El consumo en este contexto equivale a los términos utilización, fermentación o degradación.

Los genes implicados en el metabolismo de la xilosa fueron descritos por varios autores y codificaron transportadores de hexosa y pentosa, xiluloquinasa, ribulosa-5-fosfato-3-epimerasa, ribulosa-5-fosfato isomerasa,

transcetolasa, transaldolasa y genes homólogos.

La célula que expresa xilosa isomerasa se refiere aquí como una célula eucariótica microbiana que se modificó genéticamente al llevar una construcción de expresión para la expresión de la xilosa isomerasa descrita. En una realización preferida la xilosa isomerasa que expresa la célula de levadura se selecciona del grupo de *Pichia*, *Pachysolen*, *Yarrowia*, *Saccharomyces*, *Candida*, *Arxula*, *Ashbya*, *Debaryomyces*, *Hansenula*, *Hartaea*, *Kluyveromyces*, *Schwanniomyces*, *Trichosporon*, *Xanthophylomyces*, *Schizosaccharomyces*, *Zygosaccharomyces*, lo más preferible que sea *Saccharomyces cerevisiae*.

Descripción detallada de la invención

La presente solicitud proporciona soluciones para la construcción genética de células eucarióticas con un metabolismo de xilosa mejorado, una formación de biomasa mejorada en presencia de xilosa y/o una formación de metabolitos mejorada. Estas son propiedades deseables y presentan cuellos de botellas para muchas producciones industriales de cepas, especialmente producción de cepas del género *Saccharomyces*, por nombrar *Saccharomyces cerevisiae* como ejemplo no limitante. La solicitud resuelve este problema proporcionando secuencias de proteínas y DNA de genes de xilosa isomerasa que se expresan funcionalmente en células eucarióticas inferiores, especialmente levaduras siendo un ejemplo destacado las levaduras del género *Saccharomyces*, de nuevo nombrada *Saccharomyces cerevisiae* como ejemplo no limitante. La cepa creada es una célula que expresa xilosa isomerasa que muestra un consumo potencialmente mejorado de xilosa. La propiedad deseada de la actividad de xilosa isomerasa producida por la célula que expresa xilosa isomerasa es difícil de realizar de una manera satisfactoria con los medios conocidos en la técnica.

La presente invención proporciona por tanto una célula de levadura que expresa una proteína que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad, preferiblemente un 85% de identidad, lo más preferible un 90% de identidad, lo más altamente preferible un 95% de identidad con SEQ ID NO. 2 o SEQ ID NO. 8 y que tiene actividad de xilosa isomerasa en una célula de levadura. En una realización preferida, la proteína consiste en dicha secuencia de aminoácidos. En otra realización preferida, la proteína consiste en dicha secuencia de aminoácidos unida a otra parte de otras proteínas, preferiblemente partes de tales proteínas que muestran altos niveles de identidad con xilosa isomerasas conocidas o actividad de xilosa isomerasa demostrada por sí mismas.

En una realización preferida, la proteína consiste en la secuencia de SEQ ID NO. 2, o de una secuencia de aminoácidos que tiene al menos un 80% de identidad, lo más preferible un 90% de identidad, lo más altamente preferible un 95% de identidad con SEQ ID NO. 2 y que tiene actividad de xilosa isomerasa en una célula eucariótica.

Las proteínas homólogas comprenderán también secuencias de proteínas truncadas con actividad de xilosa isomerasa conservada. Un ejemplo específico de dichas secuencias de proteínas truncadas se da como SEQ ID NO. 8 o variantes de las mismas que muestran al menos un 75%, 80%, 85%, 90% o 95% de identidad con SEQ ID NO. 8.

La proteína o una composición que contiene dicha proteína, es preferiblemente diferente de una proteína o composición que se obtiene por expresión a partir de una célula procariótica. Por lo tanto, la proteína es generalmente una que se puede obtener mediante la expresión de una célula eucariótica.

La proteína muestra preferiblemente una actividad de xilosa isomerasa óptima dentro de un intervalo de pH de 7,5 a 8,5, según se determina por el método descrito en los Ejemplos.

Los niveles de identidad se pueden determinar por el programa informático AlignX, vendido en el paquete Vector-NTI por Life[™] Technology. Se aplican los valores predeterminados del componente del paquete en la versión 10.3.0.

Está claro para una persona experta que un número elevado de moléculas de DNA variables traducen a la misma secuencia de proteína y serán cubiertas como tales. También proporciona por tanto una molécula de DNA que comprende (preferiblemente consiste en) una secuencia de DNA que codifica la proteína, es decir, una proteína que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad, preferiblemente un 85% de identidad, lo más preferible un 90% de identidad, lo más altamente preferible un 95% de identidad con SEQ ID NO. 2 y que tiene actividad de xilosa isomerasa en una célula de levadura, o una realización preferida como se ha ilustrado, en donde la secuencia de DNA está operativamente unida a una secuencia reguladora eucariótica, es decir, una secuencia reguladora que permite la expresión de la célula de levadura. Ejemplos no limitantes de secuencias de DNA se dan en SEQ ID NO. 1 o SEQ ID NO. 7. Se conocen métodos para la potenciación computacional de una secuencia de DNA con respecto a los niveles de producción de la proteína. Incluyen no exclusivamente métodos que emplean la evaluación estadística de los codones preferidos (tablas de uso de codones), algoritmos que predicen la estructura secundaria del m-RNA y conocimientos basados en modelos basados en HMM o NN. De este modo se prefieren las secuencias de DNA optimizadas calculadas a partir de la secuencia de proteína objetivo. También se incluyen secuencias de DNA obtenidas por etapas recursivas y no recursivas de mutagénesis y selección o cribado de variantes mejoradas. Esta es una técnica regular para la mejora de las secuencias de DNA y proteína y las secuencias obtenidas por dichos métodos no pueden excluirse del concepto de la invención. Esto se verá independientemente de la cuestión de si la secuencia de DNA venidera de

dicho experimento deja intacta la secuencia de proteína traducida o traduce las mutaciones en ellos, siempre y cuando los niveles de identidad no bajen por debajo preferentemente del 75%, 80%, 85%, 90% o 95% con SEQ ID NO. 2 o SEQ ID NO. 8, respectivamente. Una realización preferida de hecho aplica tales procesos de mejoras para el ajuste de las moléculas de ácido nucleico descritas y las secuencias de proteínas para el problema particular.

5 Otro aspecto de la invención se refiere a secuencias quiméricas generadas por fusiones de partes de la secuencia de xilosa isomerasa con partes de otras proteínas, preferiblemente partes de dichas proteínas que muestran altos niveles de identidad con las xilosa isomerasas conocidas o una actividad demostrada de xilosa isomerasa por sí mismas así como moléculas de ácidos nucleicos que codifican tales proteínas quiméricas. Especialmente se
10 destacarán como preferidas las fusiones de la parte N-terminal de SEQ ID NO. 2 o SEQ ID NO.8 de la proteína o la parte 5' de SEQ ID NO. 1 o SEQ ID NO.7 de la molécula de ácido nucleico. Ha sido en el campo de visión de los inventores que la etapa de isomerización de xilosa es un cambio central requerido y para la configuración de un flujo de carbono eficiente con xilosa como bloque de salida cambia adicionalmente la xilosa isomerasa que expresa la célula debe ser necesario. Las cuestiones conocidas por los autores incluyen el transporte de membrana de xilosa, especialmente la captación del medio de crecimiento, la fosforilación y las etapas metabólicas del C5 de paso (parte
15 no oxidativa del orificio de paso de pentosa fosfato). Por lo tanto, los cambios adicionales introducidos en la célula de levadura, especialmente aquellos que reflejan la emulación de las cuestiones conocidas y alteran los niveles de expresión de los genes implicados en el metabolismo de xilosa, presentan una realización preferida de la invención. El orden de las introducciones de dichos cambios en las células de levadura, que pueden hacerse subsiguientes o paralelas en formas aleatorias u ordenadas, no se distinguirá en este punto y todas las posibles estrategias se
20 considerarán una parte integral y como realizaciones especiales de la invención.

La presente invención proporciona una célula de levadura que expresa la proteína que consiste preferentemente en la secuencia SEQ ID NO. 2 o SEQ ID NO.8. La célula de levadura se selecciona preferiblemente del grupo de
25 *Pichia*, *Pachysolen*, *Yarrowia*, *Saccharomyces*, *Candida*, *Arxula*, *Ashbya*, *Debaryomyces*, *Hansenula*, *Hartaea*, *Kluyveromyces*, *Schwanniomyces*, *Trichosporon*, *Xanthophylomyces*, *Schizosaccharomyces*, *Zygosaccharomyces*, siendo lo más preferible *Saccharomyces cerevisiae*. Por lo tanto, la invención proporciona también una célula de levadura genéticamente modificada que comprende un gen exógeno de xilosa isomerasa funcional en dicha célula de levadura, preferiblemente en donde el gen exógeno de xilosa isomerasa está operativamente unido a secuencias promotoras y terminadoras que son funcionales en dicha célula de levadura. En una realización preferida, el gen exógeno de xilosa isomerasa es una molécula de DNA. La célula de levadura modificada genéticamente se
30 selecciona preferiblemente del grupo de *Pichia*, *Pachysolen*, *Yarrowia*, *Saccharomyces*, *Candida*, *Arxula*, *Ashbya*, *Debaryomyces*, *Hansenula*, *Hartaea*, *Kluyveromyces*, *Schwanniomyces*, *Trichosporon*, *Xanthophylomyces*, *Schizosaccharomyces*, *Zygosaccharomyces*, siendo preferible *Saccharomyces cerevisiae*.

La célula de levadura que tiene aumentado los niveles de actividad de la xilosa isomerasa se obtiene preferiblemente por transformación de una cepa de levadura de tipo nativo con un DNA. Un aspecto adicional se
35 refiere a la aplicación de la xilosa isomerasa o la xilosa isomerasa que expresa una célula de levadura de la invención mediante la producción de productos bioquímicos basados en materias primas que contienen xilosa tal como por fermentación de la biomasa. Los productos bioquímicos incluyen biocombustibles como el etanol o butanol así como así como materias primas de base biológica para productos químicos a granel como al ácido láctico, ácido itacónico para nombrar algún ejemplo. El departamento de energía de Estados Unidos publicó una lista
40 de posibles productos bioquímicos. La proteína o la célula de levadura de la invención se puede utilizar también como un biocatalizador *in situ* o en forma purificada para la producción de productos o intermedios de azúcar isomerizados, preferiblemente para productos de azúcar isomerizados.

Un aspecto adicional de la invención se refiere al uso de la enzima de xilosa isomerasa aislada de un huésped de
45 expresión de levadura, donde dicha xilosa isomerasa está libre de contaminantes bacterianos o fragmentos de materia bacteriana. Posibles aplicaciones de tales xilosa isomerasas comprenden aplicaciones de alimentos y piensos, donde la presencia de los contaminantes mencionados incluso a niveles muy bajos representa un riesgo para la seguridad del producto. En este punto se debe plantear las cuestiones frente a la aplicación directa de *Eubacterium sabbureum* como huésped de la producción. La aplicación de la xilosa isomerasa en una levadura, es claramente ventajosa.

50 Ejemplos

1. Identificación de las secuencias genéticas candidatas con función de xilosa isomerasa

Para el descubrimiento de secuencias de xilosa isomerasa dentro del banco de genes se eligió el programa BlasP (Stephen F. Altschul y colaboradores, Nucleic Acids Res. 1997; 25; 2239-3402) en el sitio BLAST genómico del
55 NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). Como una secuencia de ensayo se tomó la secuencia de proteínas del gen de xilosa isomerasa de la *Escherichia coli* K12 (SEQ ID NO. 3). Los parámetros estándar del programa no se modificaron y las cuestiones se filtraron frente a las bases de datos de proteínas bacterianas incluyendo la base de datos de *Eubacterium saburreum* DSM 3986 construido sobre los resultados de la secuencia de escopeta del organismo (número de registro NZ_AEPW01000000). Se tuvieron en cuenta las secuencias con nivel de homología significativa sobre la longitud total de la secuencia. La búsqueda reveló un número de genes
60 candidato potenciales que fueron posteriormente clonados en *S. cerevisiae* y probados para la expresión funcional.

Todos dichos genes candidato se trataron como en la entrada ZP_07904696.1 (SEQ ID NO. 2) que es xilosa isomerasa (EsXI) de *Eubacterium saburreum* DSM 3986 como se describe en los párrafos siguientes. La entrada de secuencia de codificación unida (NZ_AEPW01000073 REGION: 2583..3956: SEQ ID NO.1) se tomó como base para la construcción de cebadores de clonación.

5 2. Amplificación del gen de xilosa isomerasa (EsXI) de *Eubacterium saburreum* DSM 3986

Los métodos para la manipulación de moléculas de ácido nucleico son generalmente conocidos por la persona experta en la materia y se introducen aquí por referencia (1. *Molecular cloning: a laboratory manual*, Joseph Sambrook, David William Russell; 2. *Current Protocols in Molecular Biology*. Last Update: January 11, 2012. Page Count: approx, 5300, Print ISSN: 1934-3639:). Las plantillas del DNA genómico del *Eubacterium saburreum* DSM 3986 se obtuvo de DSMZ (Deutsche Stammsammlung für Mikroorganismen und Zellkulturen). Los pares de cebadores flaqueantes se diseñaron para que coincidieran con la terminación N- y C-terminal de SEQ ID NO. 1. Para la amplificación de una versión truncada en el extremo N de SEQ ID NO.1 la región de unión del cebador sensorial se desplazó 54 bp posteriores (empezando con A55). La reacción de PCR se establece utilizando Finnzymes Phusion™ High Fidelity Polymerase (sistema tampón HF) siguiendo las recomendaciones del suministrador para las concentraciones de dNTP, cebador y tampón. La amplificación de los productos de PCR se realiza en un termociclador Eppendorf utilizando el programa estándar para Phusion Polymerase (98°C 30" de desnaturalización inicial seguido de 35 ciclos de etapas de 98°C (20") – 60°C (20") – 72°C (1'20") y una fase de elongación final a 72°C durante 10 minutos. Los productos de PCR de tamaño esperado se purifican mediante electroforesis en gel TAE-agarosa teñido con bromuro de etidio preparativo y se recuperan del gel utilizando el kit SV-PCR y de purificación en gel de Promga Wizard. Para la unión de un 6xHis-Tag C-terminal se utilizan los productos de PCR primarios como plantilla para la reamplificación del fragmento de DNA entero utilizando la imprimación inversa extendida con la correspondiente extensión 5', bajo idénticas condiciones (fusión PCR de 6xHIS-Tag). Los productos de PCR obtenidos se purifican otra vez por electroforesis en gel de Agarosa y se recuperan utilizando el kit SV-PCR y de purificación en gel de Promga Wizard. Contiene la Version 6xHis-TAG C-terminal del gen EsXI o 6xHis-TAG C-terminal del gen-gen (EsXI truncado) *Es-sh XI*, respectivamente.

La amplificación de genes de xilosa isomerasa optimizados con codones se hizo a partir de plantillas genéticas optimizadas solicitadas por Geneart Regensburg, Alemania. Los algoritmos de optimización para la optimización de la secuencia se utilizaron según lo proporcionado por la empresa.

3. Clonación del *EsXI* y *Es-shXI ORF* en el plásmido de expresión de *Saccharomyces cerevisiae*

Una preparación de plásmido del plásmido pSCMB454 aislado del cultivo de *Escherichia coli* se linealizó mediante restricción con endonucleasa *XmnI* y los fragmentos digeridos separados de las especies no procesadas por electroforesis en gel de Agarosa. El esqueleto del vector linealizado se recuperó del gel siguiendo las instrucciones del kit SV-PCR y de purificación en gel de Promga Wizard. El producto de PCR amplificado se clonó en el esqueleto del vector digerido con *XmnI* utilizando métodos de clonación estándar. La transformación se llevó a cabo en células de *Escherichia coli* W Mach1 químicamente competentes de acuerdo con el protocolo del proveedor. Los transformantes se hicieron crecer durante una noche en placas de LB-Ampicilina y se probaron para la corrección mediante la preparación del plásmido MINI y el control de la digestión así como la secuenciación del DNA. Se preparó una mayor cantidad de ADN plasmídico a partir de un clon confirmado utilizando el sistema PureYield™ Plasmid Midiprep System de Promega. Un ejemplo de la secuencia del casete de expresión resultante que incluye la secuencia promotora GPD y terminadora *cyc1* se da en SEQ ID NO. 6.

4. Transformación en *Saccharomyces cerevisiae*

La cepa ATCC 204667 de *Saccharomyces cerevisiae* (MATa, *ura3-52*, mal GAL+, CUP(r)) se utilizó como huésped para todos los experimentos de transformación.

La transformación se llevó a cabo utilizando métodos estándar conocidos por los expertos en la técnica (por ejemplo, véase Gietz, R.D. y R.A. Woods. (2002) Transformation of yeast by the LiAc/ss Carrier DNA/PEG method. *Methods in Enzymology* 350: 87-96). Una versión intacta del gen *ura3* de *S. cerevisiae* contenido en el vector de expresión se utilizó como marcador de selección y los transformantes se seleccionaron para el crecimiento en medio mínimo sin uracilo. El medio mínimo consistió en 20 g·l⁻¹ de glucosa, 6,7 g·l⁻¹ de base nitrogenada de levadura sin aminoácidos, 40 mg·l⁻¹ de L-tirosina, 70 mg·l⁻¹ de L-fenilalanina, 70 mg·l⁻¹ de L-triptófano, 200 mg·l⁻¹ de L-valina y 50 mg·l⁻¹ de hemisulfato de adenina, clorhidrato de L-arginina, clorhidrato de L-histidina monohidrato, L-isoleucina, L-leucina, clorhidrato de L-lisina, L-metionina, L-serina y L-treonina. El pH se ajustó a 5,6 y se añadieron 15 g·l⁻¹ de agar para medio sólido.

5. Crecimiento de xilosa isomerasas que expresan cepas de *Saccharomyces* en medio de xilosa

A) Colonias individuales de cepas de *Saccharomyces* transformadas con el vector de expresión para xilosa isomerasa de *Eubacterium saburreum* (Es XI), y *Clostridium phytofermentas* (Cp XI) así como el vector de expresión simple pSCMB454 se transfirieron en placas de medio mínimo con glucosa como única fuente de carbono. Las colonias individuales se transfirieron después en placas de medio mínimo con xilosa como única fuente de carbono (20 g·l⁻¹) y se incubaron a 30°C. Después de 7 días solo los transformantes con los vectores de expresión de xilosa

isomerasa estaban creciendo visiblemente (Fig. 2).

El examen del tamaño medio de las colonias de las cepas de *Saccharomyces* que expresan diferentes xilosa isomerasas indica que el efecto más fuerte se observó con Es XI. Cp XI y Pi XI tuvieron un efecto fuerte similar en este ensayo fisiológico pero el efecto fue notablemente más débil que para Es XI. El control negativo, la cepa de *Saccharomyces* transformada pSCMB454 solamente (vector simple), mostró sólo un crecimiento base semanal indistinguible del crecimiento base de *Saccharomyces* no transformado.

B) El crecimiento de las cepas se evaluó también en medio líquido. El medio mínimo con 20 g·l⁻¹ de xilosa como única fuente de carbono, ajustado a pH 5,6 se inoculó con una sola colonia. Después de 7 días los cultivos se dividieron en alícuotas, se almacenaron a -80°C y se utilizaron como cultivos iniciadores para el experimento de crecimiento. El experimento de crecimiento se realizó en el mismo medio mínimo y se inoculó con los cultivos iniciadores. La incubación se hizo durante 10 días en matraces agitados a 250 rpm, 30°C. El crecimiento se evaluó midiendo OD_{600nm} (Fig. 5).

Como puede deducirse de la Figura 5, el crecimiento de la cepa de *Saccharomyces* transformada con Es XI es ligeramente más fuerte que con la cepa con Cp XI. Las barras de error presentan una desviación estándar de 3 matraces agitados medidos por cepa e indican la significación estadística de la medición.

6. Preparación de extractos libres de células de levadura

Colonias individuales de cepas de *Saccharomyces* que expresan Es XI, Cp XI, se transfirieron a medio mínimo que contiene 20 g·l⁻¹ de xilosa, 6,7 g·l⁻¹ de base nitrogenada de levadura sin aminoácidos. El pH se ajustó a 5,6. Los cultivos se incubaron aeróbicamente a 30°C, 250 rpm durante 7 a 10 días. Las células se recogieron por centrifugación y se lavaron una vez con agua estéril a temperatura ambiente, se resuspendieron en agua estéril dd con OD_{600nm}>200 y se congelaron a -80°C.

La suspensión de células congelada se descongeló en hielo y se ajustó a OD_{600nm}=200. Se añadieron 100 µl de tampón de reserva NMDT (NaCl 250 mM, MnCl₂ 10 mM, DTT 1 mM, Tris/HCl 250 mM pH 7,5) y 11 µl de reserva PMSF (100 mM en isopropanol) en 1000 µl de suspensión celular. 500 µl de la suspensión celular tamponada se transfirió a un kit Precellys-Glass 0,5 mM (orden#91 PCSVK05) y fue lisado mecánicamente en un homogenizador Precellys 24 (Peqlab). La lisis se realizó 2 x 15 segundos a 5.500 rpm. Los restos celulares se eliminaron por centrifugación a 13.200 g/4°C. El lisado obtenido se dividió en alícuotas, se congeló en nitrógeno líquido y se almacenó a -80°C.

7. Ensayos para la medición de la actividad de xilosa isomerasa

A) Para algunas mediciones de la actividad de xilosa isomerasa aplicamos un ensayo espectrofotométrico basado en sorbitol deshidrogenasa (SD). Como producto de xilosa isomerasa, se forma el azúcar xilulosa isomérico. En el ensayo isomérico, se midió la cantidad de xilulosa producida. Para la medición de la actividad isomerasa en los lisados de células totales, el ensayo enzimático se realizó en forma de ensayo XI-SD acoplado (Fig. 6). Para todos los otros experimentos el ensayo enzimático se realizó en dos etapas, isomerización de la xilosa como primera etapa seguido por la determinación de la concentración de la xilulosa como segunda etapa. En el protocolo de inactivación de xilosa isomerasa en dos pasos se llevó a cabo (95°C, 10 min) después de la primera etapa. La composición de mezcla del ensayo enzimático se da a continuación:

Componentes	Concentración final
Tris-Cl pH 7,5	100 mM
MgCl ₂	10,5 mM
MnCl ₂	10 mM
DTT	1 mM
NADH	0,25 mM
Sorbitol Deshidrogenasa (Enzymstock 100 U/ml Sigma#S3764)	2 U/ml
Xilosa	1% (w/v)
ddH ₂ O	/

El ensayo se realizó en placas de microtitulación de 96 pocillos y la cinética se siguió a 340 nm.

La evaluación de la actividad enzimática en extractos celulares de *Saccharomyces cerevisiae* que expresan diferentes Xis (Fig. 6) mostró que la actividad más alta se obtuvo con *Eubacterium saburreum* XI. La actividad del *Clostridium phytofermentas* XI era medible y superior al nivel base del extracto celular libre de XI pero significativamente más bajo si se compara con los otros dos extractos.

- 5 B) Para algunas mediciones de la actividad de xilosa isomerasa se aplicó un método basado en HPLC. La cantidad de xilulosa producida se midió indirectamente a través de la disminución de la concentración de xilosa (Fig. 7). La medición se realizó con una columna H y un instrumento Dionex Ultimate 3000.

8. Expresión de xilosa isomerasa en *E. coli*, purificación de xilosa isomerasa y mediciones de la actividad

10 Todas las xilosa isomerasas se expresaron en células K12 Top10 de *E. coli* bajo promotores inducibles por arabinosa usando técnicas de biología molecular estándar. La expresión de XI se indujo con 0,02% de arabinosa a 25°C y 200 rpm durante 14 horas. Los cultivos se recogieron por centrifugación, los sobrenadantes se desecharon y las células se resuspendieron en tampón fosfato 100 mM pH 7,0 a OD_{600nm} entre 200 y 300. Las células se lisaron mediante ultrasonidos de acuerdo con métodos de purificación estándar. Las células lisadas se centrifugaron durante 30 minutos a 20.000 g a 4°C. Los sobrenadantes claros se dividieron en alícuotas y se congelaron a -80°C.

15 Antes de la purificación de los lisados se descongelaron en hielo y se añadió imidazol a 10 mM final. La purificación se hizo en columnas giratorias con 500 µl Ni-NTA (Biorad). Las columnas se equilibraron con tampón fosfato 100 mM pH 7,0, y las células lisadas se cargaron en las columnas. Las columnas se lavaron una vez con fosfato 100 mM pH 7,0 con imidazol 20 mM y se eluyó con el mismo tampón que contenía imidazol 250 mM. Se separó el imidazol y se intercambió el tampón (de fosfato a Tris-Cl se hizo con columnas giratorias Micro Bio (Biorad) de acuerdo con el manual de instrucciones. El análisis SDS-PAGE se hizo en geles al 10% (Biorad Criterion XT) de acuerdo con el manual de instrucciones. Todas las proteínas se purificaron hasta homogeneidad (>99%). La concentración de proteínas se determinó mediante el reactivo Bradford de Biorad de acuerdo con el manual de instrucciones. La albúmina de suero bovino se utilizó como un estándar. Todas las proteínas purificadas se obtuvieron a una concentración final de aproximadamente 2 g/l.

25 Las mediciones de actividad inicial de las proteínas purificadas se realizaron con el método basado en el punto final por HPLC (por favor véase más arriba). Las mediciones se llevaron a cabo a una relación enzima a sustrato (relación E/S) de 0,05%, 60°C y 2 horas. Después de la isomerización las reacciones se inactivaron como se describió anteriormente.

30 El ensayo se utilizó para obtener información sobre la actividad específica de las enzimas purificadas. Como se muestra en la Figura 7 la actividad específica más alta se observó con *Eubacterium saburreum* XI (11,9%). La actividad específica más baja se obtuvo con *Clostridium phytofermentas* XI (2,1%). Los datos obtenidos son consistentes con las mediciones de actividad de XI en los extractos celulares crudos. En ambos casos dos nuevas isomerasas descritas en esta solicitud tenían mayor actividad que el referente Cp XI.

9. Determinación del pH óptimo para xilosa isomerasas purificadas

35 El pH óptimo para XIs purificadas se determinó con el ensayo enzimático basado en el punto final de sorbitol deshidrogenasa descrito anteriormente. Se utilizó el protocolo de dos etapas descrito. Como medida de la actividad de isomerasa se utilizó la cantidad de NADH oxidado (NAD⁺; seguido como disminución a 340 nm) en el punto final de la reacción. La cantidad de NADH oxidado es equimolar con la cantidad de moléculas de xilulosa formadas durante la etapa de isomerización. Se tuvo cuidado para que el NADH no se agotase en ninguna de las reacciones utilizadas para la determinación del pH óptimo. Se utilizaron dos sistemas de tampón para la determinación del pH óptimo: BisTris para pH 5,5-7,5 y Tris de 7,5-9,5. La comparación de las actividades enzimáticas en los dos sistemas tampón se hizo a pH 7,5. No se observaron diferencias significantes.

45 El pH óptimo determinado, como se muestra en la Figura 8, muestra varias diferencias entre el referente Cp XI y los dos nuevos Xis descritos en esta invención. Primero: el pH óptimo para Cp XI es neutro (pH=7,0) y el pH óptimo para Es XI y Cp XI está en la región alcalina (pH=8,0). Segundo: la actividad residual de Es XI (pH=5,5) se encuentra al 50% (flecha inferior) de la actividad máxima. La actividad residual de Cp XI a pH=5,5 equivale prácticamente a cero. Tercero: los dos nuevos XIs forman picos relativamente anchos entre pH 7,0 (>90%, flecha superior) y pH 8,0 (=100%). En comparación Cp XI conserva <80% de la actividad a pH=8,0.

10. Determinación de la temperatura óptima para xilosa isomerasas purificadas

50 La temperatura óptima para XIs purificadas se determinó con el ensayo enzimático de punto final descrito anteriormente. También en este experimento se tuvo cuidado para que el NADH no se agotase en ninguna de las reacciones utilizadas para la determinación de la temperatura óptima. Los gradientes de temperatura se generaron con cicladores comunes de PCR de laboratorio (Eppendorf).

55 La determinación de la temperatura óptima (Figura 9) reveló varias diferencias entre el referente Cp XI y Es XI descrito en esta invención. Primero: la temperatura óptima para Cp XI se define con un pico relativamente afilado a 56,2°C. Es XI muestra picos significativamente más anchos que van desde 53,8°C hasta 61,6°C. Segundo: la actividad de Es XI a 67°C es de alrededor del 50% y para el Cp XI es prácticamente igual a cero. En su conjunto los

datos de la temperatura óptima muestran que los XIs descritos en esta solicitud poseen una temperatura de estabilidad significativamente más alta que la referencia Cp XI.

11. Determinación de Km para xilosa isomerasas purificadas

5 Los valores de Km se determinaron con el ensayo enzimático descrito en los ejemplos previos. Para el experimento se utilizó xilosa isomerasas purificadas de *E. coli*.

La determinación de Km para xilosa isomerasas purificadas revelan un Km para Es-sh XI de 18,4 mM. El Km para Cp XI (Km=36,6 mM), (Fig. 10).

Listado de secuencias

10 SEQ ID NO. 1: Secuencia de *Eubacterium sabbureum* DSM 3986 Secuencia de DNA que codifica xilosa isomerasa (NZ_AEPW01000073.1 GI:315651683)

gtgaaaacaaaaacaacattatatgtactattgcattgaaaggagacatatttatgaaagaatttttcccggcatatcacctgtaaagttt
 gagggcagagatagtaaaaatccacttagtttcaaatattatgatgccaaaagggtgataatgggcaaaacaatggaggaacatttacc
 atttgctatggcatggggcataatctttgtgcctgtggtggtgatattgctggacagggtactgtcgataaaagtttggtaaaagctccg
 gtactatggagcatgcaagggtcaaaagtgatgcaggcattgaatttatgaaaaagcttggataaaagtattattgcttccatgatacggga
 tattgtacctgaggatcaggaagatataaatgttaccatgcacgtttggatgagattacagactatatcttagaaaaacaaaggataacc
 gatataaatgtctttggacaacctgcaaatgttcaagtaaatccaagatttatgaacgggtgcaggaagctcaaacagtgcagatgtattttg
 cttgcagcggcacaggcaaaagaaggtcttgaaaatgccgtaaaacttggagcaaaagggttattctggggaggcagagaagg
 ttatgagacacttctaaatacagatatgaagcctgaagaggaaaatagcaaacactcttacaatgtgcagagattatggacgcagtata
 ggctttatgggagattttatattgagcctaagccgaaggagcctatgaagcatcagatgattttgatcgggcaactgcaatcggttttta
 agaaaatattggacttgataaagatttcaactaaatattgaggcaaatcacgctacacttgcaggtcatactttcagcatgagttaaagat
 atgtgcagtcaacggatgatggggctcggtagatccaatcaaggagatacattcttggatgggacactgatcaattccctacaaatgt
 ctatgatactacattggctatgtatgaaatattaaggcagcggactccgtggaggtctgaactttgattcaagaatcgcagaccaag
 taatacagccgatgatatttctatggctttatagcaggtatggacacatttgcacttggacttattaagcggcggaattatagaagac
 ggaagaatagatgattttgtaaaagaagatatgcaagttataattcaggaataggtaaagaagataagaacagaaaagtgactgat
 agagtggtccgagatgccgcaagcttaaaaagcctgaactgccggaatcaggaagacaggaatatcttgagagcgtagtgaataa
 tatattgttcggataa

SEQ ID NO. 2: Secuencia de proteína de *Eubacterium sabbureum* DSM 3986 traducido Secuencia de DNA que codifica xilosa isomerasa (ZP_07904696.1) (EsXI)

mktknniictialkgdifmkeffpgispvkfegrdsknplsfkyydakrvimgktmehlsfamawwhnlcacgvdmfmgqgt
 vdksgessgtmeharakvdagiefmkklgikyycfhdtdivpedqedinvtnarldeitdyilektdtdikclwttnmfsnpr
 fmngagssnsadvfcfaaaqakkglenavklgakgfvfwggregyettlntdmkleeeniatlftmcrdygrsigfmgdfyiep
 kpkepmkhqydfdaataigflrkygldkdfklnieanhatlaghtfqhelrvcavngmmgsvdanqgdllgwtdtdqfptnvy
 dttlamyeilkgglrgglndfsknrrpsntaddmfygfiaqmdtfalglkaaeiiedgriddfvkeryasynsigkkirnrvktli
 15 ecaeyaaklkkpelpesgrqeylesvvnnilfg*

SEQ ID NO. 3: Xilosa isomerasa de *Escherichia coli* (Proteína)

mqa yfdqldr vryegskssnplaf rhynpd elvl gkrmeehl rfaacywh tfcwn gadm fgv gafnrpwqqpgealal akrk
advafeffhklhvpfycfhdvdspegaslkeyinnfaqmvdvlagkqeesgvkllwgtancftnprygagaatnpdpevfsw
aatqvvtameathklggenyvfwggregyettlntdlrqereqlgrfmqmvvehkhkigfqgtlliepkpseptkhqydydaat
vygflkqfglekeiklnieanhatlaghsfhheiataialglfgsvdanrgdaqlgwtdqfpnsveenalvmyeilkaggfttgln
fdakvrrqstdkydlfyghigamdtmalalkiaarmedgeldkriaqrys gwnselgqqilkqgmsladlakyae hhlspvh
qsgrqeqlenlvnhylfdk*

SEQ ID NO. 4: Xilosa isomerasa de *Clostridium phytofermentas* (Proteína) (Cp XI)

mknyfnpvpevkeyegpnstnpfakyydankvvagktmkehcrfalswwhtlcaggadpfgvttmdrtygnitdpmelaka
kvdagfelmtklgieffcfhdadiapegdtfeeskknlfeivdyikekmdqgtikllwgtannfshprfmhgastscnadvfayaa
akiknaldatiklggkgyvfwggregyettlntdlgleldnmarlkmaveygrangfdgdfyiepkpseptkhqydfdatvla
flrkyglekdfkmnieanhatlaghtfeh elamarvngafgsvdanqgd pnlgwtdqfpdvhsatlamlevlkaggftngln
fdakvrrgsfefddiaygyiagmdtfalglikaaeiiddgriakfvddryasyktgigkaivdgtsleeleqyvlt hsepvmqsgrq
evletivnnilfr*

SEQ ID NO. 5: Xilosa isomerasa de *Piromyces* sp (Proteína – PI_XI)

makeyfpqi qkikfegkdsknplafhydae kevmgkkmkdwlrfamawwh tcaegadqfgggtksfpwnegtdaieia
kqkvdagfeimqklgipyycfhdvdlvsegn sieeyesnlkavvaylkekqketgikllwstanvfghkrymngastnpdfdv
varai vqiknaidagielgaenyvfwggregyemslntdqkrekehmatmltmardyarskgfkgf liepkpmeptkhqydv
dtetaigflkahnldkdfkvnievnhatlaghtfeh elacavdagmlgsidanrgdyqngwtdqfpidqyelvqawmeiirgg
gfvttgtnfdaktrrntdlediiiahvsgmdamaralenaakllqespytkmkkeryasfdsgigkdfedgkltleqvyyegkkn
5 gepkqts gkqelyeai vamyq*

SEQ ID NO. 6: Casete de expresión Ex XI (Negrita y mayúsculas: secuencia de codificación del gen EsXI 6x-His-tag C-terminal y un enlazador de fusión; Mayúsculas pequeñas: Promotor GPD; subrayado: restos del sitio Xnml; cursiva: terminador CYCI)

CTCGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATT
 AGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGGGTTACACAGAATATATAACAT
 CGTAGGTGTCTGGGTGAACGTTTTATTCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGG
 CATCCAGAAAAAAGAATCCCAGCACAAAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCC
 ATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCAAAAACGGGCACAACCTCAATGGAG
 TGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTCTCT
 TACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAAGCAGTTCCC
 TCAAATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCGTAAATCT
 ATTTCTTAAACTTCTTAAATCTACTTTTATAGTTAGTCTTTTTTTTAGTTTTTAAACACCAAGAACTTAG
 TTTCGAATaaacacacagaacaagaaa**ATGAAAACAAAAACAACATTATATGTACTATT**
GCATTGAAAGGAGACATATTTATGAAAGAATTTTTTCCCGGCATATCACCTG
TAAAGTTTGAGGGCAGAGATAGTAAAAATCCACTTAGTTTCAAATATTATGA
TGCCAAAAGGGTGATAATGGGCAAAACAATGGAGGAACATTTATCATTGCT
ATGGCATGGTGGCATAATCTTTGTGCCTGTGGTGTGGATATGTTTCGGACAG
GGTACTGTGATAAAAAGTTTTGGTGAAAGCTCCGGTACTATGGAGCATGCAA
GGGCTAAAGTGGATGCAGGCATTGAATTTATGAAAAAGCTTGGTATAAAGTA
TTATTGCTTCCATGATACGGATATTGTACCTGAGGATCAGGAAGATATAAAT
GTTACCAATGCACGTTTGGATGAGATTACAGACTATATCTTAGAAAAACAA
AGGATACCGATATTAATGTCTTTGGACAACCTGCAATATGTTTCAGTAATCC
AAGATTTATGAACGGTGCAGGAAGCTCAAACAGTGCAGATGTATTTTGCTTT
GCAGCGGCACAGGCAAAGAAAGGTCTTGAAAATGCCGTAAAACCTGGAGCA
AAGGGATTTGTATTCTGGGGAGGCAGAGAAGGTTATGAGACACTTCTAAAT
ACAGATATGAAGCTTGAAGAGGAAAATATAGCAACACTCTTTACAATGTGCA
GAGATTATGGACGCAGTATAGGCTTTATGGGAGATTTTTATATTGAGCCTAA
GCCGAAGGAGCCTATGAAGCATCAGTATGATTTTGATGCGGCAACTGCAAT
CGGTTTTTTAAGAAAATATGGACTTGATAAAGATTTCAAACATAAATATTGAG
GCAAATCACGCTACACTTGCAGGTCATACTTTTCAGCATGAGTTAAGAGTAT
GTGCAGTCAACGGTATGATGGGGTCGGTAGATGCCAATCAAGGAGATACAT
TACTTGGATGGGACACTGATCAATTCCCTACAAATGTCTATGATACTACATT
GGCTATGTATGAAATATTAAGGCAGGCGGACTCCGTGGAGGTCTGAACTT
TGATTCAAAGAATCGCAGACCAAGTAATACAGCCGATGATATGTTCTATGGC
TTTATAGCAGGTATGGACACATTTGCACTTGGACTTATTAAGGCGGCGGAAA
TTATAGAAGACGGAAGAATAGATGATTTTGTAAAGAAAGATATGCAAGTTA
TAATTCAGGAATAGGTAAGAAGATAAGAAACAGAAAAGTGACACTGATAGA
GTGTGCCGAGTATGCCGCAAAGCTTAAAAAGCCTGAACTGCCGGAATCAGG
AAGACAGGAATATCTTGAGAGCGTAGTGAATAATATATTGTTTCGGAGGATCT
GGCCATCACCACCATCATCTAA*gttcgctcctgtagttatgtcacgcttacattcacgccctccccc*
catccgctctaaccgaaaaggaaggagtagacaacctgaagtctaggtccctattttttatagttatgtagtattaagaacgt
tatttatattcaaatctttttttctgtacagacgcgtgtacgatgaacattatactgaaaacctgcttgagaagggtttgggacg
ctcgaaggcttaatttg

SEQ ID NO. 7: DNA optimizado de *S. cerevisiae* que codifica la versión truncada de Es-sh_XI con unión C-terminal de un 6x His TAG

atgaaggaattctcccaggtatctcccagttaagttgaaggtagagattctaagaaccattgtccttcaagtactacgatgccaaaga
gagttattatgggtaagaccatggaagaacattgtcttttctatggcttgggtggcataattgtgtgcttgggtgtgatggtcgggca
ggactgttgataagctttcgggtgaacttctggfactatggaacatgctagagctaaagtgtatgccggattgaattcatgaagaagttg
ggattaagtactactgctccacgatactgatacgttccagaagatcaagaa gatatcaacgttaccatgccagattgacgaaatta
ccgattacatcttgaaaagactaaggacaccgatatcaagtggttggactactgtaacatgttctccaaccaagattcatgaacgg
tgctggttcttaattctgctgatgtttttgttctgctgctcaagctaaaaagggttggaaaatgctgtaagtgggtgctaagggtt
tgtttttgggggtgtagagaaggttacgaaacttggtaacactgacatgaagttggaagaagaaaacattgctacctgttcaccatgt
gtagagattacggtagatccattggttcatgggtgatttctacattgaacctaagccaaaagaacctatgaagcaccaatac gattttgat
gctgctactgctattggttcttgagaaagtatggttggacaaggacttcaagttgaacattgaagctaacctgctactttggctggtcat
actttcaacacgaattgagagtttggctgtcaatggatgatgggttctgttgatgctaatacaggtgatacttgttgggttgggatactg
atcaattccaactaacgtttacgataccacctggccatgtacgaaatttgaagctggtggttgagaggtggtttaaactttgactcta
agaacagaagaccatccaactgctgatgatatgtttacggttcattgctggtatggatacttctgcttgggttattaaggccgccg
aaattattgaagatggtagaattgatgacttgcataagaaagatacgcctcttacaattccgggtatcggtaagaagattagaacagaa
aggttacctgatcgaatgcgctgaatgctgctaaattgaagaaccagaattgccagaatccggtagacaagaatatttgaatctgt
cgtcaacaacatctgtttggtggttctggtcatcatcaccatcattaa

5 SEQ ID NO. 8: Es-sh_XI truncado en el extremo N de *Eubacterium sabbureum* DSM 3986

mkeffpgispvkfegrsknplsfkyydakrvimgktmeehlsfamawwhnlcacgvdmfqqgtvdkfsfessgtmehara
kvdagiefmkklgikyycfhdtdivpedqedinvtnarldeitdyilektkdtdikclwttcnmfsnprfmngagssnsadvfca
aaqakkglenavklgakgfvfwggregyetllntdmkleeniatlftmcrdygrsigfmgdfyiepkpkpepmkhqydfdaata
igflrkygldkdfklnieanhatlaghtfqhelrvcaavngmmgsvdanqgdttl gwtdtdqfptnvdydtllamyeilkagglrgln
fdsknrpsntaddmfygfia gmdtfalglikaeeiiedgriddfvkeryasynsiggkknrvktliecaeyaaklkkpelpesgr
qeylesvvnnilfg*

Listado de secuencias

<110> Clariant Produkte GmbH

<120> Microorganismos que fermentan pentosa

10 <130> 162 373

<150> EP12000783.6

<151> 2012-02-07

<160> 8

<170> PatentIn version 3.5

15 <210> 1

<211> 1374

<212> DNA

<213> *Eubacterium sabbureum* DSM 3986

ES 2 643 046 T3

<400> 1
 gtgaaaaaaa aaaacaacat tatatgtact attgcattga aaggagacat atttatgaaa 60
 gaatTTTTTC cCGGCATATC acctgtaaag tttgagggca gagatagtaa aaatccactt 120
 agtttcaaat attatgatgc caaaaggggtg ataatgggca aaacaatgga ggaacattta 180
 tcatttgcta tggcatgggtg gcataatctt tgtgcctgtg gtgtggatat gttcggacag 240
 ggtactgtcg ataaaagttt tggtgaaaagc tccggtacta tggagcatgc aagggtctaaa 300
 gtggatgcag gcattgaatt tatgaaaaag cttgggtataa agtattattg cttccatgat 360
 acggatattg tacctgagga tcaggaagat ataaatgtta ccaatgcacg tttggatgag 420
 attacagact atactttaga aaaaacaaag gataccgata ttaaattgtct ttggacaacc 480
 tgcaatatgt tcagtaatcc aagatttatg aacgggtgcag gaagctcaaa cagtgcagat 540
 gtatTTTtGCT ttgcagcggc acaggcaaag aaaggctcttG aaaatgccgt aaaacttggA 600
 gcaaagggat ttgtattctg gggaggcaga gaaggttatg agacacttct aaatacagat 660
 atgaagcttg aagaggaaaa tatagcaaca ctctttacaa tgtgcagaga ttatggacgc 720
 agtataggct ttatgggaga tttttatatt gagcctaagc cgaaggagcc tatgaagcat 780
 cagtatgatt ttgatgcggc aactgcaatc ggtTTTTttaa gaaaatatgg acttgataaa 840
 gatttcaaac taaatattga ggcaaatcac gctacacttg caggtcatac ttttcagcat 900
 gagttaagag tatgtgcagt caacgggatg atggggtcgg tagatgccaa tcaaggagat 960
 acattacttg gatgggacac tgatcaattc cctacaaatg tctatgatac tacattggct 1020
 atgtatgaaa tattaagggc aggcggactc cgtggaggtc tgaactttga ttcaaagaat 1080
 cgcagaccaa gtaatacagc cgatgatatg ttctatggct ttatagcagg tatggacaca 1140
 tttgcaacttg gacttattaa ggcggcggaa attatagaag acggaagaat agatgatTTT 1200
 gttaaagaaa gatatgcaag ttataattca ggaataggta agaagataag aaacagaaaa 1260
 gtgacactga tagagtgtgc cgagtatgcc gcaaagctta aaaagcctga actgccggaa 1320
 tcaggaagac aggaatatct tgagagcgta gtgaataata tattgttcGg ataa 1374

- 5 <210> 2
 <211> 457
 <212> PRT
 <213> Eubacterium sabbureum DSM 3986
 <400> 2

ES 2 643 046 T3

Met Lys Thr Lys Asn Asn Ile Ile Cys Thr Ile Ala Leu Lys Gly Asp
 1 5 10 15

Ile Phe Met Lys Glu Phe Phe Pro Gly Ile Ser Pro Val Lys Phe Glu
 20 25 30

Gly Arg Asp Ser Lys Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Ala Lys
 35 40 45

Arg Val Ile Met Gly Lys Thr Met Glu Glu His Leu Ser Phe Ala Met
 50 55 60

Ala Trp Trp His Asn Leu Cys Ala Cys Gly Val Asp Met Phe Gly Gln
 65 70 75 80

Gly Thr Val Asp Lys Ser Phe Gly Glu Ser Ser Gly Thr Met Glu His
 85 90 95

Ala Arg Ala Lys Val Asp Ala Gly Ile Glu Phe Met Lys Lys Leu Gly
 100 105 110

Ile Lys Tyr Tyr Cys Phe His Asp Thr Asp Ile Val Pro Glu Asp Gln
 115 120 125

Glu Asp Ile Asn Val Thr Asn Ala Arg Leu Asp Glu Ile Thr Asp Tyr
 130 135 140

Ile Leu Glu Lys Thr Lys Asp Thr Asp Ile Lys Cys Leu Trp Thr Thr
 145 150 155 160

Cys Asn Met Phe Ser Asn Pro Arg Phe Met Asn Gly Ala Gly Ser Ser
 165 170 175

Asn Ser Ala Asp Val Phe Cys Phe Ala Ala Ala Gln Ala Lys Lys Gly
 180 185 190

Leu Glu Asn Ala Val Lys Leu Gly Ala Lys Gly Phe Val Phe Trp Gly
 195 200 205

ES 2 643 046 T3

Gly Arg Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu
 210 215 220

Glu Glu Asn Ile Ala Thr Leu Phe Thr Met Cys Arg Asp Tyr Gly Arg
 225 230 235 240

Ser Ile Gly Phe Met Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu
 245 250 255

Pro Met Lys His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe
 260 265 270

Leu Arg Lys Tyr Gly Leu Asp Lys Asp Phe Lys Leu Asn Ile Glu Ala
 275 280 285

Asn His Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Val
 290 295 300

Cys Ala Val Asn Gly Met Met Gly Ser Val Asp Ala Asn Gln Gly Asp
 305 310 315 320

Thr Leu Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp
 325 330 335

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

Gly Leu Asn Phe Asp Ser Lys Asn Arg Arg Pro Ser Asn Thr Ala Asp
 355 360 365

Asp Met Phe Tyr Gly Phe Ile Ala Gly Met Asp Thr Phe Ala Leu Gly
 370 375 380

Leu Ile Lys Ala Ala Glu Ile Ile Glu Asp Gly Arg Ile Asp Asp Phe
 385 390 395 400

Val Lys Glu Arg Tyr Ala Ser Tyr Asn Ser Gly Ile Gly Lys Lys Ile
 405 410 415

Arg Asn Arg Lys Val Thr Leu Ile Glu Cys Ala Glu Tyr Ala Ala Lys
 420 425 430

Leu Lys Lys Pro Glu Leu Pro Glu Ser Gly Arg Gln Glu Tyr Leu Glu
 435 440 445

Ser Val Val Asn Asn Ile Leu Phe Gly
 450 455

<210> 3

<211> 440

<212> PRT

5 <213> Escherichia coli

<400> 3

ES 2 643 046 T3

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Tyr Glu Gly Ser
 1 5 10 15

Lys Ser Ser Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu Leu
 20 25 30

Val Leu Gly Lys Arg Met Glu Glu His Leu Arg Phe Ala Ala Cys Tyr
 35 40 45

Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe Gly Val Gly Ala
 50 55 60

Phe Asn Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu Ala Leu Ala Lys
 65 70 75 80

Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys Leu His Val Pro
 85 90 95

Phe Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu Gly Ala Ser Leu
 100 105 110

Lys Glu Tyr Ile Asn Asn Phe Ala Gln Met Val Asp Val Leu Ala Gly
 115 120 125

Lys Gln Glu Glu Ser Gly Val Lys Leu Leu Trp Gly Thr Ala Asn Cys
 130 135 140

Phe Thr Asn Pro Arg Tyr Gly Ala Gly Ala Ala Thr Asn Pro Asp Pro
 145 150 155 160

Glu Val Phe Ser Trp Ala Ala Thr Gln Val Val Thr Ala Met Glu Ala
 165 170 175

Thr His Lys Leu Gly Gly Glu Asn Tyr Val Leu Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Arg Gln Glu Arg Glu Gln
 195 200 205

Leu Gly Arg Phe Met Gln Met Val Val Glu His Lys His Lys Ile Gly
 210 215 220

ES 2 643 046 T3

Phe Gln Gly Thr Leu Leu Ile Glu Pro Lys Pro Gln Glu Pro Thr Lys
225 230 235 240

His Gln Tyr Asp Tyr Asp Ala Ala Thr Val Tyr Gly Phe Leu Lys Gln
245 250 255

Phe Gly Leu Glu Lys Glu Ile Lys Leu Asn Ile Glu Ala Asn His Ala
260 265 270

Thr Leu Ala Gly His Ser Phe His His Glu Ile Ala Thr Ala Ile Ala
275 280 285

Leu Gly Leu Phe Gly Ser Val Asp Ala Asn Arg Gly Asp Ala Gln Leu
290 295 300

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

Val Met Tyr Glu Ile Leu Lys Ala Gly Gly Phe Thr Thr Gly Gly Leu
325 330 335

Asn Phe Asp Ala Lys Val Arg Arg Gln Ser Thr Asp Lys Tyr Asp Leu
340 345 350

Phe Tyr Gly His Ile Gly Ala Met Asp Thr Met Ala Leu Ala Leu Lys
355 360 365

Ile Ala Ala Arg Met Ile Glu Asp Gly Glu Leu Asp Lys Arg Ile Ala
370 375 380

Gln Arg Tyr Ser Gly Trp Asn Ser Glu Leu Gly Gln Gln Ile Leu Lys
385 390 395 400

Gly Gln Met Ser Leu Ala Asp Leu Ala Lys Tyr Ala Gln Glu His His
405 410 415

Leu Ser Pro Val His Gln Ser Gly Arg Gln Glu Gln Leu Glu Asn Leu
420 425 430

Val Asn His Tyr Leu Phe Asp Lys
435 440

<210> 4

<211> 438

<212> PRT

5 <213> Clostridium phytofermentas

<400> 4

ES 2 643 046 T3

Met Lys Asn Tyr Phe Pro Asn Val Pro Glu Val Lys Tyr Glu Gly Pro
 1 5 10 15

Asn Ser Thr Asn Pro Phe Ala Phe Lys Tyr Tyr Asp Ala Asn Lys Val
 20 25 30

Val Ala Gly Lys Thr Met Lys Glu His Cys Arg Phe Ala Leu Ser Trp
 35 40 45

Trp His Thr Leu Cys Ala Gly Gly Ala Asp Pro Phe Gly Val Thr Thr
 50 55 60

Met Asp Arg Thr Tyr Gly Asn Ile Thr Asp Pro Met Glu Leu Ala Lys
 65 70 75 80

Ala Lys Val Asp Ala Gly Phe Glu Leu Met Thr Lys Leu Gly Ile Glu
 85 90 95

Phe Phe Cys Phe His Asp Ala Asp Ile Ala Pro Glu Gly Asp Thr Phe
 100 105 110

Glu Glu Ser Lys Lys Asn Leu Phe Glu Ile Val Asp Tyr Ile Lys Glu
 115 120 125

Lys Met Asp Gln Thr Gly Ile Lys Leu Leu Trp Gly Thr Ala Asn Asn
 130 135 140

Phe Ser His Pro Arg Phe Met His Gly Ala Ser Thr Ser Cys Asn Ala
 145 150 155 160

Asp Val Phe Ala Tyr Ala Ala Ala Lys Ile Lys Asn Ala Leu Asp Ala
 165 170 175

Thr Ile Lys Leu Gly Gly Lys Gly Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Leu Glu Leu Asp Asn
 195 200 205

Met Ala Arg Leu Met Lys Met Ala Val Glu Tyr Gly Arg Ala Asn Gly
 210 215 220

Phe Asp Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys
 225 230 235 240

His Gln Tyr Asp Phe Asp Thr Ala Thr Val Leu Ala Phe Leu Arg Lys

ES 2 643 046 T3

Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Cys Ala Val
 275 280 285

Asp Ala Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr Gln
 290 295 300

Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Gln Tyr Glu Leu Val
 305 310 315 320

Gln Ala Trp Met Glu Ile Ile Arg Gly Gly Gly Phe Val Thr Gly Gly
 325 330 335

Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser Thr Asp Leu Glu Asp
 340 345 350

Ile Ile Ile Ala His Val Ser Gly Met Asp Ala Met Ala Arg Ala Leu
 355 360 365

Glu Asn Ala Ala Lys Leu Leu Gln Glu Ser Pro Tyr Thr Lys Met Lys
 370 375 380

Lys Glu Arg Tyr Ala Ser Phe Asp Ser Gly Ile Gly Lys Asp Phe Glu
 385 390 395 400

Asp Gly Lys Leu Thr Leu Glu Gln Val Tyr Glu Tyr Gly Lys Lys Asn
 405 410 415

Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala Ile
 420 425 430

Val Ala Met Tyr Gln
 435

<210> 6

<211> 2309

<212> DNA

5 <213> Secuencia artificial

<220>

<223> Secuencia codificante del gen EsXI con fusión de enlazador y etiqueta 6x-His C terminal

<400> 6

ctcgccattt caaagaatac gtaaataatt aatagtagtg attttcctaa ctttatttag 60

tcaaaaaatt agccttttaa ttctgctgta acccgtagat gcccaaaata gggggcgggt 120

tacacagaat atataacatc gtaggtgtct gggtagaacag tttattcctg gcatccacta 180

aatataatgg agcccgcttt ttaagctggc atccagaaaa aaaaagaatc ccagcaccaa 240

aatattgttt tcttcaccaa ccatcagttc ataggtccat tctcttagcg caactacaga 300

ES 2 643 046 T3

gaacaggggc acaaacaggc aaaaaacggg cacaacctca atggagtgat gcaacctgcc 360
 tggagtaaat gatgacacaa ggcaattgac ccacgcatgt atctatctca ttttcttaca 420
 ccttctatta ccttctgctc tctctgattt ggaaaaagct gaaaaaaaag gttgaaagca 480
 gttccctcaa attattcccc tacttgacta ataagtatat aaagacggta ggtattgatt 540
 gtaattctgt aatctatctt cttaaacttc ttaaattcta cttttatagt tagtcttttt 600
 tttagtttta aaacaccaag aacttagttt cgaataaaca cacagaaaca aagaaaatga 660
 aaacaaaaaa caacattata tgtactattg cattgaaagg agacatattt atgaaagaat 720
 tttttccogg catatcacct gtaaagtttg agggcagaga tagtaaaaat ccacttagtt 780
 tcaaataatta tgatgccaaa aggtgataa tgggcaaaac aatggaggaa catttatcat 840
 ttgctatggc atggtggcat aatctttgtg cctgtggtgt ggatatgttc ggacagggta 900
 ctgctgataa aagttttggt gaaagctccg gtactatgga gcatgcaagg gctaaagtgg 960
 atgcaggcat tgaatctatg aaaaagcttg gtataaagta ttattgcttc catgatcagg 1020
 atattgtacc tgaggatcag gaagatataa atggtaccaa tgcacgtttg gatgagatta 1080
 cagactatat cttagaaaaa acaaaggata ccgatattaa atgtcttttg acaacctgca 1140
 atatgttcag taatccaaga tttatgaacg gtgcaggaag ctcaaacagt gcagatgtat 1200
 tttgctttgc agcggcacag gcaaagaaag gtcttgaaaa tgccgtaaaa cttggagcaa 1260
 agggatttgt attctgggga ggcagagaag gttatgagac acttctaaat acagatatga 1320
 agcttgaaga ggaaaatata gcaacactct ttacaatgtg cagagattat ggacgcagta 1380
 taggctttat gggagatttt tatattgagc ctaagccgaa ggagcctatg aagcatcagt 1440
 atgattttga tgcggcaact gcaatcgggt ttttaagaaa atatggactt gataaagatt 1500
 tcaaactaaa tattgaggca aatcacgcta cacttgaggc tcatactttt cagcatgagt 1560
 taagagtatg tgcagtcaac ggtatgatgg ggtcggtaga tgccaatcaa ggagatacat 1620
 tacttgatg ggacactgat caattcccta caaatgtcta tgatactaca ttggctatgt 1680
 atgaaatatt aaaggcaggc ggactccgtg gaggtctgaa ctttgattca aagaatcgca 1740
 gaccaagtaa tacagccgat gatatgttct atggctttat agcaggtagt gacacatttg 1800
 cacttgact tattaaggcg gcggaatta tagaagcgg aagaatagat gattttgtta 1860
 aagaagata tgcaagttat aattcaggaa taggtaagaa gataagaaac agaaaagtga 1920
 cactgataga gtgtgccgag tatgccgcaa agcttaaaaa gcctgaactg ccggaatcag 1980
 gaagacagga atatcttgag agcgtagtga ataatatatt gttcggagga tctggccatc 2040
 accaccatca tcaactaatgt tcgtcctcgt ttagttatgt cacgcttaca ttcacgccct 2100
 cccccacat ccgctctaac cgaaaaggaa ggagttagac aacctgaagt ctaggtccct 2160
 atttattttt ttatagttat gttagtatta agaacgttat ttatatttca aatttttctt 2220
 tttttctgt acagacgcgt gtacgcatgt aacattatac tgaaaacctt gcttgagaag 2280
 gttttgggac gctcgaaggc ttttaatttg 2309

- 5 <210> 7
- <211> 1347
- <212> DNA
- <213> Secuencia artificial

ES 2 643 046 T3

<220>

<223> ADN optimizado de *S. cerevisiae* que codifica la versión truncada de Es-sh_XI con fusión C-terminal de a 6x etiqueta His

<400> 7

```

atgaaggaat tcttcccag tctctcccca gttaagtttg aaggtagaga ttctaagaac      60
ccattgtcct tcaagtacta cgatgccaaag agagttatta tgggtaagac catggaagaa      120
catttgtcct ttgctatggc ttggtggcat aatttgtgtg cttgtggtgt tgatatgttc      180
ggcacaagga ctgttgataa gtctttcggg gaactctctg gtactatgga acatgctaga      240
gctaaagttg atgccgggat tgaattcatg aagaagttgg gtattaagta ctactgcttc      300
cacgatactg atatcgttcc agaagatcaa gaagatatca acgttaccaa tgccagattg      360
gacgaaatta ccgattacat cttggaaaag actaaggaca ccgatatcaa gtgtttgtgg      420
actacttgta acatgttctc caaccaaga ttcattgaac gtgctgggtc ttctaattct      480
gctgatgttt tttgtttcgc tgctgctcaa gctaaaagg gtttgaaaa tgctgttaag      540
ttgggtgcta agggttttgt tttttggggg ggtagagaag gttacgaaac tttgttgaac      600
actgacatga agttggaaga agaaaacatt gctacctgtg tcaccatgtg tagagattac      660
ggtagatcca ttggtttcat gggtgatttc tacattgaac ctaagccaaa agaacctatg      720
aagcaccaat acgattttga tgctgctact gctattgggt tcttgagaaa gtatggtttg      780
gacaaggact tcaagttgaa cattgaagct aaccatgcta ctttggctgg tcatactttt      840
caacacgaat tgagagtttg tgctgtcaat ggtatgatgg gttctgttga tgctaataca      900
ggtgatactt tgttgggttg ggatactgat caatttccaa ctaacgttta cgataccacc      960
ttggccatgt acgaaatfff gaaagctggg ggtttgagag gtggttttaa ctttgactct     1020
aagaacagaa gaccatccaa cactgctgat gatatgtttt acggtttcat tgctggtatg     1080
gatactttcg ctttgggttt gattaaggcc gccgaaatta ttgaagatgg tagaattgat     1140
gacttcgtca aagaaagata cgcctcttac aattccggta tcggtaagaa gattagaaac     1200
agaaaggtta cttgatcga atgctgctgaa tatgctgcta aattgaagaa accagaattg     1260
ccagaatccg gtagacaaga atatttggaa tctgtcgtca acaacatctt gtttgggtgg     1320
5 tctggtcatc atcatcacca tcattaa                                     1347

```

<210> 8

<211> 439

<212> PRT

<213> Secuencia artificial

10

<220>

<223> *Eubacterium sabbureum* DSM 3986 Es-sh_XI truncada N-terminalmente

<400> 8

ES 2 643 046 T3

Met Lys Glu Phe Phe Pro Gly Ile Ser Pro Val Lys Phe Glu Gly Arg
 1 5 10 15

Asp Ser Lys Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Ala Lys Arg Val
 20 25 30

Ile Met Gly Lys Thr Met Glu Glu His Leu Ser Phe Ala Met Ala Trp
 35 40 45

Trp His Asn Leu Cys Ala Cys Gly Val Asp Met Phe Gly Gln Gly Thr
 50 55 60

Val Asp Lys Ser Phe Gly Glu Ser Ser Gly Thr Met Glu His Ala Arg
 65 70 75 80

Ala Lys Val Asp Ala Gly Ile Glu Phe Met Lys Lys Leu Gly Ile Lys
 85 90 95

Tyr Tyr Cys Phe His Asp Thr Asp Ile Val Pro Glu Asp Gln Glu Asp
 100 105 110

Ile Asn Val Thr Asn Ala Arg Leu Asp Glu Ile Thr Asp Tyr Ile Leu
 115 120 125

Glu Lys Thr Lys Asp Thr Asp Ile Lys Cys Leu Trp Thr Thr Cys Asn
 130 135 140

Met Phe Ser Asn Pro Arg Phe Met Asn Gly Ala Gly Ser Ser Asn Ser
 145 150 155 160

Ala Asp Val Phe Cys Phe Ala Ala Ala Gln Ala Lys Lys Gly Leu Glu
 165 170 175

Asn Ala Val Lys Leu Gly Ala Lys Gly Phe Val Phe Trp Gly Gly Arg
 180 185 190

Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu Glu Glu
 195 200 205

ES 2 643 046 T3

Asn Ile Ala Thr Leu Phe Thr Met Cys Arg Asp Tyr Gly Arg Ser Ile
 210 215 220

Gly Phe Met Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Met
 225 230 235 240

Lys His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe Leu Arg
 245 250 255

Lys Tyr Gly Leu Asp Lys Asp Phe Lys Leu Asn Ile Glu Ala Asn His
 260 265 270

Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Val Cys Ala
 275 280 285

Val Asn Gly Met Met Gly Ser Val Asp Ala Asn Gln Gly Asp Thr Leu
 290 295 300

Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp Thr Thr
 305 310 315 320

Leu Ala Met Tyr Glu Ile Leu Lys Ala Gly Gly Leu Arg Gly Gly Leu
 325 330 335

Asn Phe Asp Ser Lys Asn Arg Arg Pro Ser Asn Thr Ala Asp Asp Met
 340 345 350

Phe Tyr Gly Phe Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Leu Ile
 355 360 365

Lys Ala Ala Glu Ile Ile Glu Asp Gly Arg Ile Asp Asp Phe Val Lys
 370 375 380

Glu Arg Tyr Ala Ser Tyr Asn Ser Gly Ile Gly Lys Lys Ile Arg Asn
 385 390 395 400

Arg Lys Val Thr Leu Ile Glu Cys Ala Glu Tyr Ala Ala Lys Leu Lys
 405 410 415

Lys Pro Glu Leu Pro Glu Ser Gly Arg Gln Glu Tyr Leu Glu Ser Val
 420 425 430

Val Asn Asn Ile Leu Phe Gly
 435

REIVINDICACIONES

1. Una célula de levadura que expresa una proteína que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad, lo más preferiblemente un 90% de identidad, lo más altamente preferible un 95% de identidad con SEQ ID NO. 2 y que tiene actividad de xilosa isomerasa en una célula de levadura.
- 5 2. La célula de levadura de la reivindicación 1, en donde la proteína consiste en la secuencia SEQ ID NO. 2 y que tiene una actividad de xilosa isomerasa en la célula de levadura.
3. La célula de levadura de la reivindicación 1 o 2, en donde la proteína consiste en la secuencia SEQ ID NO. 8, o en la secuencia de aminoácidos que tiene al menos un 75% de identidad, preferiblemente un 80% de identidad, lo más preferiblemente un 90% de identidad, lo más altamente preferible un 95% de identidad con SEQ ID NO. 8 y que tiene una actividad de xilosa isomerasa en la célula.
- 10 4. La célula de levadura de una cualquiera de las reivindicaciones 1 a 3, en donde la proteína tiene una actividad óptima de xilosa isomerasa dentro de un intervalo de pH de 7,5 a 8,5.
5. La célula de levadura de una cualquiera de las reivindicaciones 1 a 4 que comprende una molécula de DNA que comprende una secuencia de DNA que codifica una proteína como se define en una cualquiera de las reivindicaciones 1 a 4, en donde la secuencia de DNA está unida operativamente a una secuencia reguladora eucariótica.
- 15 6. La célula de levadura de la reivindicación 5, en donde la molécula de DNA consiste en la secuencia SEQ ID NO. 1 o SEQ ID NO. 7.
7. La célula de levadura de una cualquiera de las reivindicaciones 1 a 6, en donde la célula de levadura se selecciona del grupo de *Pichia*, *Pachysolen*, *Yarrowia*, *Saccharomyces*, *Candida*, *Arxula*, *Ashbya*, *Debaryomyces*, *Hansenula*, *Hartaea*, *Kluyveromyces*, *Schwanniomyces*, *Trichosporon*, *Xanthophylomyces*, *Schizosaccharomyces*, *Zygosaccharomyces*, siendo lo más preferible *Saccharomyces cerevisiae*.
8. La célula de levadura de una cualquiera de las reivindicaciones 1 a 7 que comprende un gen exógeno de la xilosa isomerasa funcional en dicha célula de levadura, en donde el gen exógeno de la xilosa isomerasa está unido operativamente a las secuencias promotora y terminadora que son funcionales en dicha célula de levadura, conduciendo a la expresión de una proteína como se definió en una cualquiera de las reivindicaciones 1 a 4.
- 25 9. La célula de levadura de la reivindicación 8, en donde el gen exógeno de la xilosa isomerasa es una molécula de DNA como se definió en la reivindicación 5 o 6.
10. El uso de una célula de levadura de acuerdo con una cualquiera de las reivindicaciones 1 a 9 para la fermentación de biomasa a partir de un medio que contiene xilosa como fuente de carbono.
- 30 11. El uso de una célula de levadura de acuerdo con una cualquiera de las reivindicaciones 1 a 9 como un biocatalizador *in situ* o en una forma purificada para la producción de productos de azúcar isomerizados, preferiblemente para productos de azúcar isomerizados.
12. Un proceso para producir un producto de fermentación seleccionado del grupo del ácido láctico, ácido acético, ácido succínico, aminoácidos, 1,3-propano-diol, etileno, glicerol, antibióticos β -lactámicos, cefalosporinas, biocombustibles, butanol, etanol, ácido láctico, ácido itacónico, preferiblemente butanol, más preferiblemente etanol por lo que el proceso comprende las etapas de:
 - a) fermentar un medio que contiene una fuente de xilosa con una célula de levadura como se definió en una cualquiera de las reivindicaciones 1 a 9, y opcionalmente,
 - 40 b) recuperar el producto de fermentación.

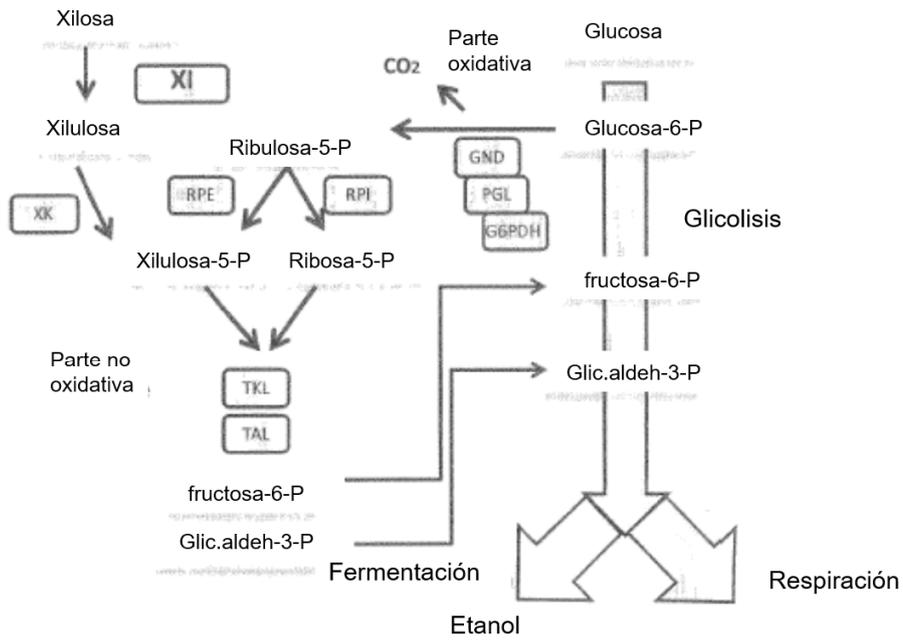


Figura 1

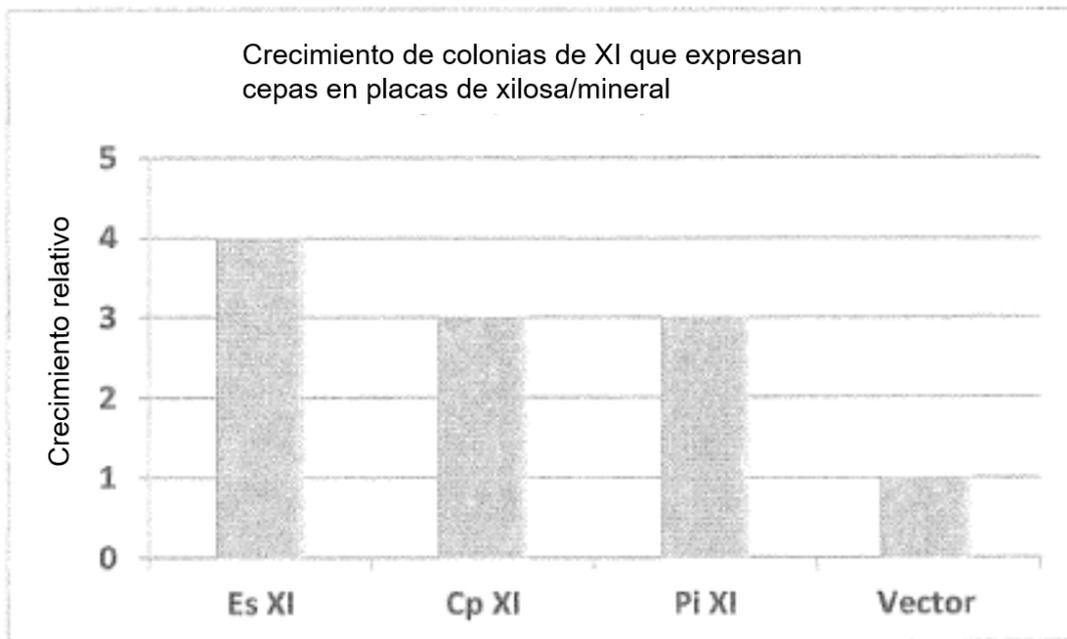


Figura 2

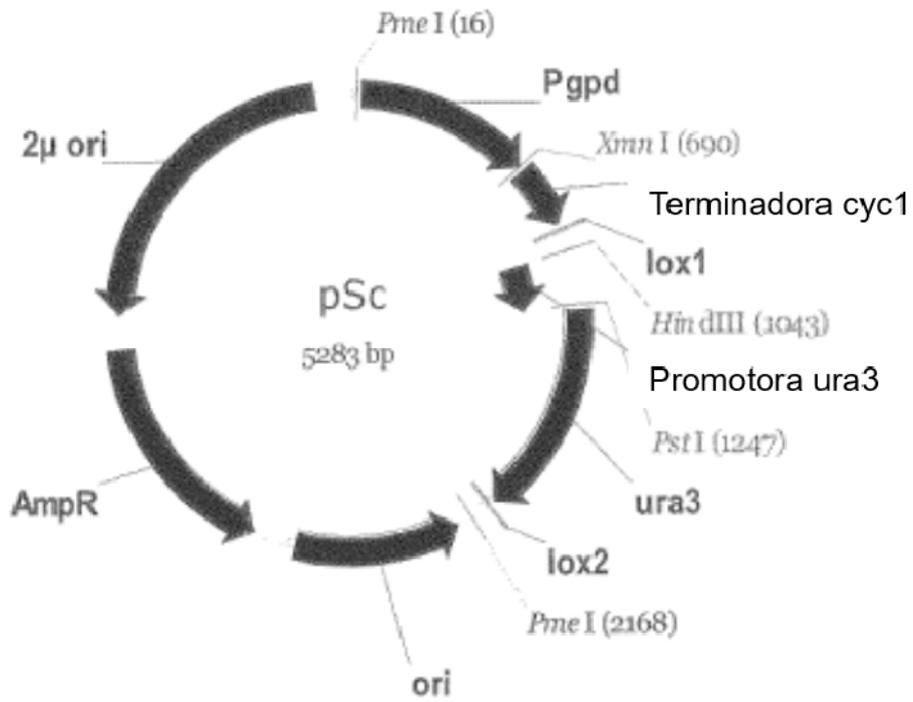


Figura 3



Figura 4

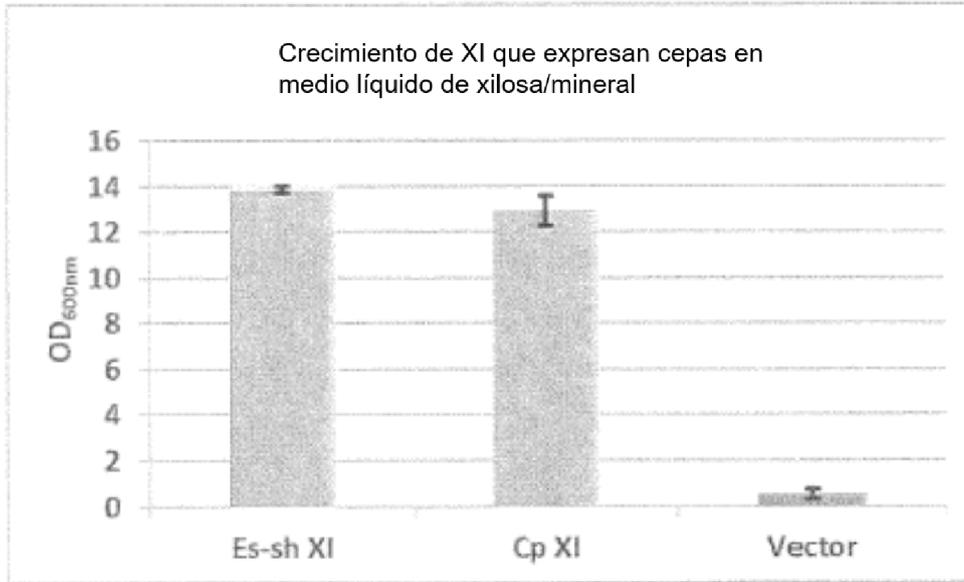


Figura 5

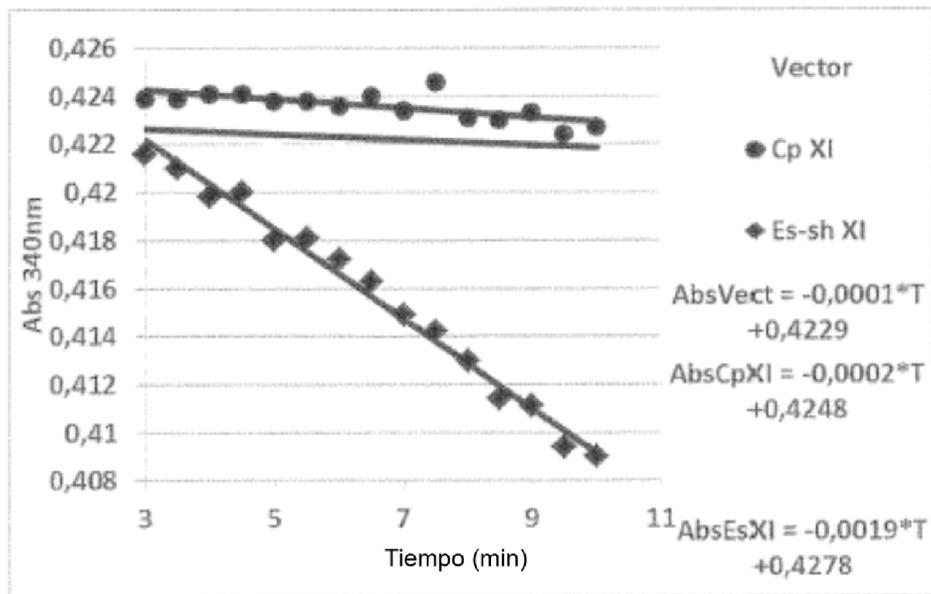


Figura 6

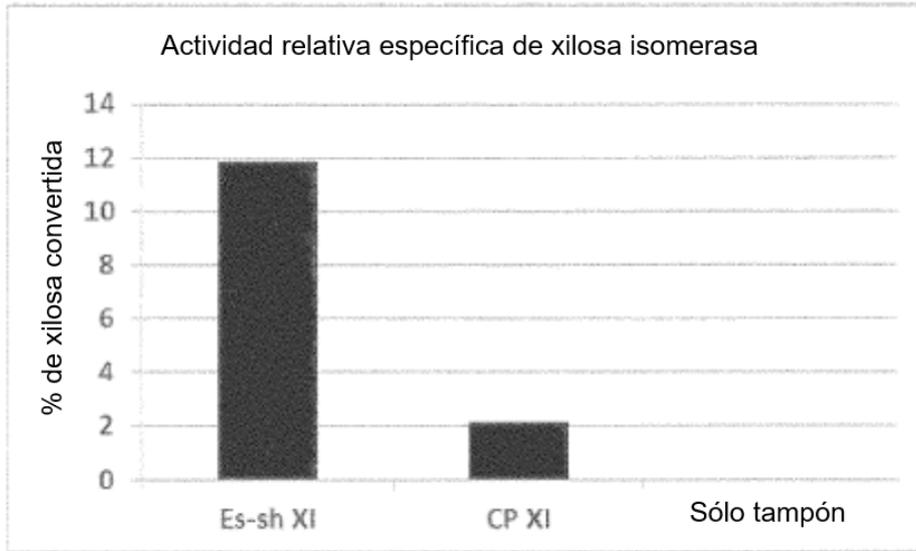


Figura 7

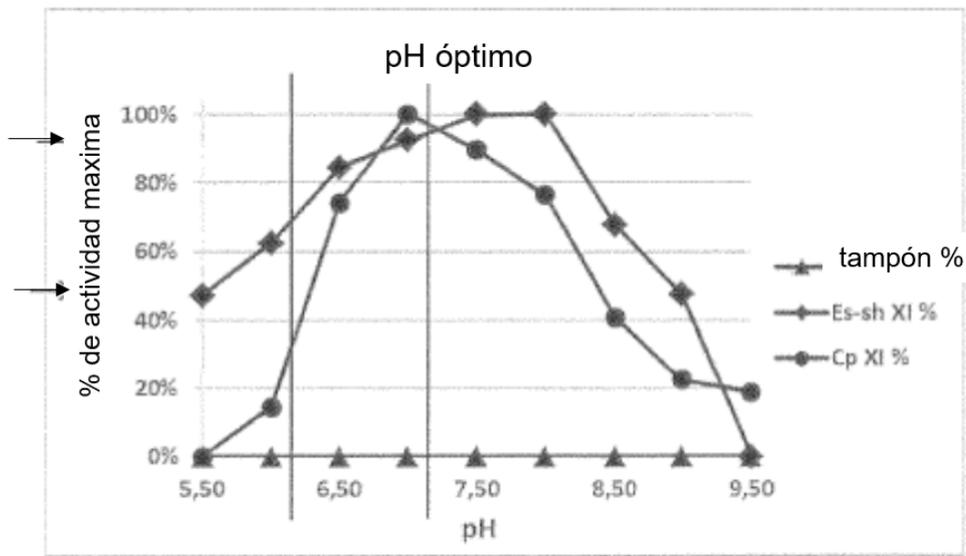


Figura 8

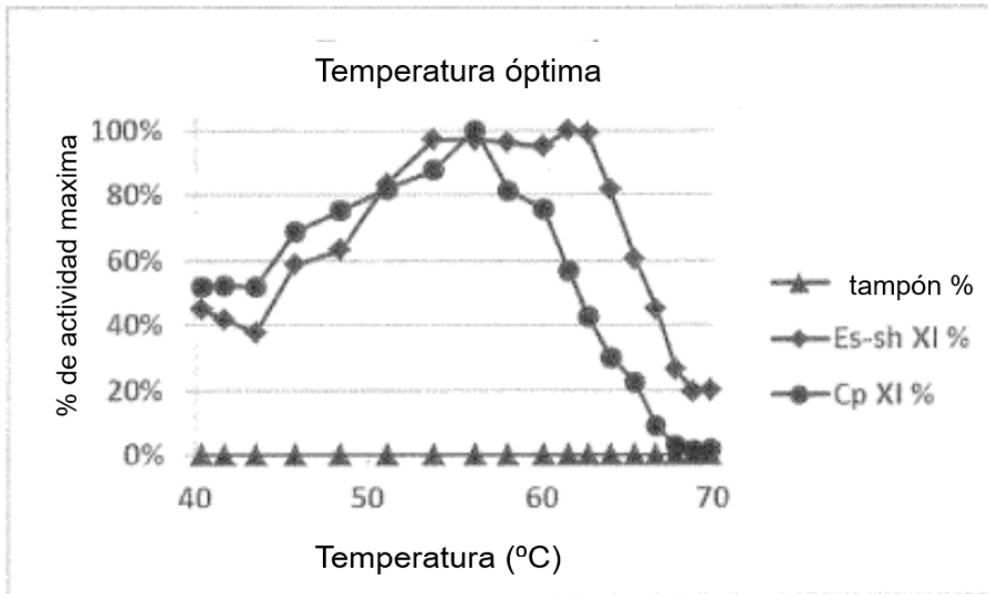


Figura 9

Xilosa isomerasa	Km [mM]
Es-sh XI	18.4
Cp XI	36.6

Figura 10