

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 438 001**

51 Int. Cl.:

**C12N 9/10** (2006.01)  
**C12N 15/54** (2006.01)  
**C12N 15/82** (2006.01)  
**C12N 1/21** (2006.01)  
**C12N 5/04** (2006.01)  
**A01H 5/00** (2006.01)  
**A01H 5/10** (2006.01)  
**C08B 37/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.02.2008 E 08735411 (4)**

97 Fecha y número de publicación de la concesión europea: **23.10.2013 EP 2121908**

54 Título: **Moléculas de ácido nucleico que codifican alternano sacarosas truncadas**

30 Prioridad:

**14.02.2007 EP 07090022**  
**15.02.2007 US 901532 P**

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

**15.01.2014**

73 Titular/es:

**BAYER CROPSCIENCE AG (100.0%)**  
**Alfred-Nobel-Strasse 50**  
**40789 Monheim, DE**

72 Inventor/es:

**FROHBERG, CLAUS y**  
**VAN LIPZIG, ROSALINDE**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 438 001 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

Moléculas de ácido nucleico que codifican alternano sacarosas truncadas

5 La presente invención se refiere a moléculas de ácido nucleico que codifican una alternano sacarosa truncada. La invención se refiere además a alternano sacarosas truncadas, a vectores, a células huésped y a células vegetales transformadas con las moléculas de ácido nucleico descritas y a plantas que comprenden estas células. La presente invención se refiere además a polímeros de alternano novedosos con propiedades ventajosas y a procedimientos para su producción.

10 El alternano es un polisacárido compuesto por unidades de glucosa. Las unidades de glucosa están unidas entre sí por enlaces  $\alpha$ -1,3 y  $\alpha$ -1,6-glucosídicos, con estos dos tipos de enlace alternándose de forma predominante (Miasaki et al., Carbo. Res., 84:273- 285 (1980)). Además, el alternano puede comprender ramificaciones (Seymour et al., Carbohydrate Research 74, (1979); 41-62). El alternano nativo tiene un peso molecular medio  $P_m$  de  $10^6$ - $10^7$  (documento WO 03/010177).

15 Debido a sus propiedades físicoquímicas, se han analizado posibles usos del alternano, no sólo en la industria farmacéutica, por ejemplo como excipiente para principios activos en medicamentos, sino también como aditivo en la industria textil, cosmética y alimentaria (Lopez-Munguia et al., Enzyme Microb. Technol. 15, (1993), 77-85; Leathers et al., Journal of Industrial Microbiology & Biotechnology 18, (1997), 278-283). Otro posible uso es como sustituto de la goma arábiga (Cote, Carbohydrate Polymers 19, (1992), 249-252).

20 El alternano se puede producir enzimáticamente usando enzimas con la actividad biológica de las alternano sacarosas. Las alternano sacarosas pertenecen al grupo de las glucosil transferasas que, partiendo de la sacarosa, pueden catalizar la formación de glucano y fructosa. Por lo general, las cepas naturales también producirán dextrano además de alternano. La producción de alternano sacarosa a partir de las cepas naturales supone una desventaja porque no se ha podido obtener la enzima en su forma pura, a pesar de los intentos de realizar complicados procedimientos de purificación; véase, por ejemplo, la descripción general de la introducción del documento WO 00/4772.

25 En el pasado se han realizado gran cantidad de intentos para proporcionar alternano sacarosa ultrapura por medio de rutas alternativas.

30 El documento WO 00/47727 describe moléculas de ácido nucleico que codifican una alternano sacarosa. Además, el documento WO 00/47727 se refiere a vectores, a células huésped y a células vegetales transformadas con las moléculas de ácido nucleico descritas y a plantas que comprenden estas células. Con ayuda de las moléculas de ácido nucleico del documento WO 00/47727, se pueden generar células huésped que producen proteína alternano sacarosa recombinante de gran pureza y/o en cantidades suficientes y generar plantas modificadas genéticamente con una actividad de estas enzimas, de modo que se forma alternano en la planta.

35 La solicitud de patente US20060127328A1 describe secuencias de ácidos nucleicos de alternano sacarosas truncadas o mutadas y vectores y células huésped transformadas. Se obtienen alternano sacarosas truncadas ultrapuras que mantienen su actividad catalítica. Además, se reivindica que las alternano sacarosas truncadas se expresan mejor y se degradan con menos rapidez.

A la vista de la técnica anterior, es un objeto de la presente invención proporcionar alternano sacarosas con propiedades aún mejores.

40 En particular, era un objeto de la invención proporcionar moléculas de ácido nucleico que se expresen particularmente bien en células vegetales con el fin de lograr una producción eficaz de alternano sacarosa en plantas.

Otro objeto era proporcionar una alternano sacarosa que provoque una producción de alternano en una planta transgénica con rendimiento alto.

45 Además, era un objeto de la invención proporcionar una alternano sacarosa con ayuda de la cual se pueda preparar un alternano con propiedades particularmente ventajosas.

Estos objetivos se logran proporcionando las formas de uso que se caracterizan en las reivindicaciones de la patente.

La presente memoria descriptiva se refiere a una molécula de ácido nucleico que codifica una proteína con la actividad enzimática de una alternano sacarosa, donde la molécula de ácido nucleico se selecciona de entre

50 a) una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 118 hasta, e incluido, el nucleótido de la posición 4140, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o la SEQ ID NO 3,

b) una molécula de ácido nucleico que codifica una proteína cuya secuencia de aminoácidos tiene al menos el

70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1380, inclusive, de la SEQ ID NO 2,

5 c) una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 118 hasta el nucleótido de la posición 3945, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o la SEQ ID NO 3,

10 d) una molécula de ácido nucleico que codifica una proteína cuya secuencia de aminoácidos tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2,

e) una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 118 hasta el nucleótido de la posición 4518, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o la SEQ ID NO 3,

15 f) una molécula de ácido nucleico que codifica una proteína cuya secuencia de aminoácidos tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1506, inclusive, de la SEQ ID NO 2,

20 g) una molécula de ácido nucleico cuya secuencia tenga al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia de a), c) o e), donde dicha secuencia tiene el mismo número de nucleótidos que la secuencia de a), c) o e),

h) una moléculas de ácido nucleico cuya secuencia de ácidos nucleicos, como consecuencia de la degeneración del código genético, se desvía de la secuencia de las moléculas de ácido nucleico de a), b), c), d), e), f) o g),

25 i) una molécula de ácido nucleico cuya secuencia de ácidos nucleicos es complementaria a la secuencia de longitud completa de una de las moléculas de ácido nucleico de a), b), c), d), e), f), g) o h), donde la secuencia complementaria tiene el mismo número de nucleótidos que las moléculas de ácido nucleico de a), b), c), d), e), f), g) o h).

La presente invención tiene, en particular, las siguientes ventajas:

30 Con ayuda de las moléculas de ácido nucleico de acuerdo con la invención, se pueden preparar proteínas con la actividad enzimática de una alternano sacarasa que, en comparación con la alternano sacarasa de longitud completa (por ejemplo, de origen natural), tienen una secuencia de aminoácidos truncada.

35 En la presente invención, "alternano sacarasa de longitud completa" se refiere a la proteína que tiene una secuencia de aminoácidos desde el aminoácido 40 al 2057 de la SEQ ID NO 2. La alternano sacarasa de longitud completa de acuerdo con la definición anterior difiere de la alternano sacarasa descrita en la literatura (por ejemplo, en el documento US20060127328A1) por el hecho de que carece del péptido señal de los aminoácidos 1-39 de la SEQ ID NO 2.

40 Las proteínas de acuerdo con la invención con la actividad enzimática de una alternano sacarasa también se denominan en adelante en el presente documento "alternano sacarasa truncada" o "enzima truncada". El término "truncada" significa un truncamiento en el extremo N- y/o C-terminal de la secuencia de aminoácidos de una alternano sacarasa de longitud completa, o en el extremo 5' o 3' de la secuencia de ácidos nucleicos en cuestión. El truncamiento se puede llevar a cabo por medio de enzimas de restricción, enzimas proteolíticas en el caso de proteínas, por síntesis completa (por ejemplo, de Merrifield) o, preferentemente, como se describe en los ejemplos de la presente invención. Por ejemplo, las moléculas de ácido nucleico codificadas por la reivindicación 1 y sus modos de realización preferentes comprenden alternano sacarasas truncadas de acuerdo con la presente invención.

45 Preferentemente, las moléculas de ácido nucleico de acuerdo con la invención adoptan la forma de moléculas de ácido nucleico aisladas.

Las alternano sacarasas truncadas de la presente invención tienen una actividad mayor en células vegetales que las enzimas de longitud completa.

Además, las alternano sacarasas truncadas de la presente invención se expresan muy bien.

50 Las alternano sacarasas truncadas se pueden introducir más fácilmente en una vacuola vegetal o en otros compartimentos vegetales, tales como plastos o apoplastos, que la enzima de longitud completa. Esto es ventajoso para la producción de alternano en la planta, ya que, en función de la especie vegetal, la sacarosa material de partida está concentrada en la vacuola. Frecuentemente, es ventajoso expresar enzimas que sintetizan polímeros en compartimentos distintos del citosol.

La presente invención demostró además que se puede obtener un alternano con propiedades ventajosas con ayuda de alternano sacarosas truncadas codificadas por las moléculas de ácido nucleico anteriores.

También se puede demostrar que la enzima alternano sacarasa truncada produce menos polímero de alternano como subproducto en la producción de oligosacárido de alternano que la alternano sacarasa de longitud completa. Los oligosacáridos de alternano se producen por una reacción catalizada por una enzima alternano sacarasa de sacarosa con diversos azúcares aceptores y se han descrito como prebióticos para controlar patógenos bacterianos entéricos. Al usar enzima alternano sacarasa truncada en la producción de oligosacáridos de alternano, se obtiene un rendimiento mayor del producto de oligosacárido deseado. Además, se facilita la purificación de membrana del oligosacárido, ya que el alternano polimérico puede dar lugar a incrustaciones en la membrana.

- 5
- 10
- 15
- Por último, la presente invención proporciona también alternano sacarosas truncadas de gran pureza que mantienen su actividad catalítica. Con ayuda de las moléculas de ácido nucleico de acuerdo con la invención, se pueden generar esas células que producen una alternano sacarasa truncada recombinante de gran pureza y/o en cantidades relativamente grandes, y también generar plantas modificadas genéticamente con una actividad de estas enzimas, de modo que se produce la formación de alternano en la planta: en el contexto de la presente invención, el término "gran pureza" significa que la proteína de acuerdo con la invención tiene un grado de pureza de al menos el 80 %, preferentemente de al menos el 90 % y aún más preferentemente de al menos el 95 %.

Preferentemente, se excluyen de la presente invención las moléculas de ácido nucleico y moléculas de ácido nucleico truncadas descritas en la solicitud de patente publicada US20060127328A1.

- 20
- Preferentemente, se excluyen de la presente invención las moléculas de ácido nucleico descritas en los párrafos [0062], [0065] y [0094] del documento US20060127328A 1. De acuerdo con la numeración de la SEQ ID NO 1, éstas son las siguientes moléculas de ácido nucleico:

1. desde el nucleótido de la posición 1 hasta el nucleótido de la posición 4275, inclusive, de la SEQ ID NO 1 (= ácido nucleico 195-4469 de acuerdo con la numeración del documento US20060127328A 1)

- 25
2. desde el nucleótido de la posición 1 hasta el nucleótido de la posición 4047, inclusive, de la SEQ ID NO 1 (= ácido nucleico 195-4241 de acuerdo con la numeración del documento US20060127328A 1)

3. desde el nucleótido de la posición 1024 hasta el nucleótido de la posición 4275, inclusive, de la SEQ ID NO 1 (= ácido nucleico 1218-4469 de acuerdo con la numeración del documento US20060127328A 1)

4. desde el nucleótido de la posición 1 hasta el nucleótido de la posición 3870, inclusive, de la SEQ ID NO 1 (= ácido nucleico 195-4064 de acuerdo con la numeración del documento US20060127328A 1)

- 30
5. desde el nucleótido de la posición 1024 hasta el nucleótido de la posición 3870, inclusive, de la SEQ ID NO 1 (= ácido nucleico 1218-4064 de acuerdo con la numeración del documento US20060127328A 1)

Preferentemente, se excluyen asimismo de la presente invención las moléculas de ácido nucleico que se mencionan en la memoria descriptiva del documento WO 00/47727 y que codifican una alternano sacarasa.

- 35
- En contraste, otras moléculas de ácido nucleico ventajosas que codifican una proteína con la actividad enzimática de alternano sacarasa son:

1. una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 502 hasta el nucleótido de la posición 3945, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o la SEQ ID NO 3,

- 40
2. una molécula de ácido nucleico que codifica una proteína cuya secuencia de aminoácidos tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia desde el aminoácido de la posición 168 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2,

45

3. una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 1024 hasta el nucleótido de la posición 4047, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o la SEQ ID NO 3,

4. una molécula de ácido nucleico que codifica una proteína cuya secuencia de aminoácidos tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia desde el aminoácido de la posición 342 hasta el aminoácido de la posición 1349, inclusive, de la SEQ ID NO 2,

- 50
5. una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 1024 hasta el nucleótido de la posición 3945, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o la SEQ ID NO 3, y

6. una molécula de ácido nucleico que codifica una proteína cuya secuencia de aminoácidos tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia desde el aminoácido de la posición 342 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2.

5 También se proporcionan en cada caso moléculas de ácido nucleico cuya secuencia de ácidos nucleicos deriva de la secuencia de una de las moléculas de ácido nucleico especificadas anteriormente, debido a la degeneración del código genético. Asimismo, se proporcionan moléculas de ácido nucleico cuya secuencia tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una de las secuencias especificadas  
10 anteriormente. El término "identidad" se define en otro punto de la presente divulgación y se detalla un procedimiento de identificación de la identidad. Además, la presente invención proporciona moléculas de ácido nucleico cuya secuencia de ácidos nucleicos es complementaria a la secuencia de longitud completa de una de las moléculas de ácido nucleico mencionadas anteriormente, donde la secuencia complementaria, preferentemente, tiene el mismo número de nucleótidos que las moléculas de ácido nucleico mencionadas anteriormente.

15 El trabajador experto puede encontrar en la literatura sugerencias relativas a la posición de la secuencia de nucleótidos o la secuencia de aminoácidos donde se pueden llevar a cabo modificaciones al mismo tiempo que la proteína codificada, o la proteína con la secuencia de aminoácidos, mantiene aún la actividad enzimática de una alternano sacarasa. G. Joucla et al., FEBS Letters 580 (2006) 763-768 y el documento US 2006/0127328A1 describen diversos dominios de una enzima alternano sacarasa, véanse, por ejemplo, la fig. 2 de Joucla et al. y las  
20 figs. 3A y 3B del documento US 2006/0127328A1. En primer lugar, esta referencia divulga una región variable N-terminal en el intervalo desde el aminoácido 40 hasta el 341 de la proteína, que corresponde a los nucleótidos 118 a 1023 de una secuencia codificante adecuada, tal como, por ejemplo, los nucleótidos 118 a 1023 de la SEQ ID NO 1 o 3 de la presente invención. La región variable N-terminal va seguida de un dominio catalítico muy conservado desde el aminoácido 342 hasta el 1290 de la proteína, que corresponde a los nucleótidos 1024 a 3870 de una  
25 secuencia codificante correspondiente, por ejemplo, con la SEQ ID NO 1 o 3 (G. Joucla et al., p. 763, columna de la izquierda). Partiendo de la presente divulgación, un trabajador experto esperaría que las modificaciones de nucleótidos/aminoácidos en la región variable mantuvieran la actividad enzimática de una alternano sacarasa, en lugar de las modificaciones en el dominio catalítico muy conservado. Por tanto, el trabajador experto, realizaría en primer lugar modificaciones en la sección de la molécula de ácido nucleico que codifica la región variable de  
30 proteínas alternano sacarasa (nucleótidos 118-1023 de la SEQ ID 1 o 3) con el fin de generar una molécula de ácido nucleico cuya secuencia tenga al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia de la reivindicación 1 a), c) o e) y que todavía codifique una proteína con actividad alternano sacarasa. Esto mismo se cumple para una modificación de la secuencia de aminoácidos con el fin de obtener una molécula de  
35 ácido nucleico de acuerdo con la reivindicación 1 b), d) o f), que el trabajador experto realizaría en primer lugar en la región de los aminoácidos 40-341 de la SEQ ID NO 2.

Además, en G. Joucla et al., FEBS Letters 580 (2006) 763-768 y el documento US 2006/0127328A1 se detalla un dominio C-terminal (dominio de unión a glucano, GBD) que consiste en 709 aminoácidos y, por lo tanto, comprende los aminoácidos 1348-2057 de una enzima alternano sacarasa. El dominio C-terminal contiene lo que se conoce  
40 como repeticiones CW y repeticiones APY. G. Joucla et al. explican que un alternano sintetizado por una alternano sacarasa que carece de las repeticiones APY (aminoácidos 1507-2057) es idéntico a un alternano obtenido con alternano sacarasa nativa. Por tanto, los motivos APY no participan en la actividad polimerasa de la enzima (Joucla et al., p. 767, columna de la izquierda). Además, la alternano sacarasa truncada estudiada por Joucla et al. (denominada "ASR C- APY- del") contiene todavía sólo 4 de las 7 repeticiones CW originales (Joucla et al. p. 765, columna de la izquierda), sin que esto afecte a la actividad polimerasa de la enzima. Por tanto, la retirada de los  
45 aminoácidos de la posición 1426-1481, que corresponde aproximadamente a las unidades CW de 5 a 7 (véase, Joucla et al., fig. 1) tampoco destruye la actividad de la alternano sacarasa, como confirman los resultados de Joucla et al. Partiendo de una molécula de ácido nucleico de acuerdo con la reivindicación 1 e), por lo tanto, un trabajador experto, preferentemente llevaría a cabo modificaciones en la sección de 4276-4518 de la SEQ ID NO 1 o 3, que  
50 codifica los aminoácidos 1426-1506 de una alternano sacarasa de acuerdo con la SEQ ID NO 2, con el fin de generar una molécula de ácido nucleico cuya secuencia tenga al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia de la reivindicación 1 e). Esto mismo se cumple para una modificación de la secuencia de aminoácidos con el fin de obtener una molécula de ácido nucleico de acuerdo con la  
55 reivindicación 1 f), que el trabajador experto realizaría en primer lugar en la región de los aminoácidos 1426-1506 de la SEQ ID NO 2.

Sin embargo, lo dicho anteriormente no excluye en modo alguno las modificaciones en otras regiones de las secuencias de ácidos nucleicos/aminoácidos de acuerdo con la invención, que todavía mantienen la actividad enzimática de una alternano sacarasa. Por lo tanto, el ámbito de la presente invención comprende expresamente  
60 modificaciones en otras secciones distintas de las indicadas anteriormente.

En la técnica anterior de G. Joucla et al., se mantiene que la retirada de todo el dominio C-terminal es desfavorable para la expresión de la proteína (p. 765, columna de la derecha), donde, de acuerdo con Joucla, el dominio C-

terminal consiste en 709 aminoácidos y, por tanto, comprende los aminoácidos 1348-2057 de una alternano sacarasa (Joucla, p. 763, columna de la derecha). Sin embargo, en la presente invención se demostró sorprendentemente que una alternano sacarasa truncada que comprende los aminoácidos 40-1315 de la SEQ ID NO 2 se expresa y es catalíticamente activa, cosa que no cabía esperar a la vista de las enseñanzas de Joucla et al.

- 5 En otro aspecto, la presente memoria descriptiva describe una molécula de ácido nucleico que codifica una proteína con actividad alternano sacarasa seleccionada del grupo que consiste en
- a) una molécula de ácido nucleico que comprende la secuencia de nucleótidos mostrada en la SEQ ID NO 3,
  - b) una molécula de ácido nucleico cuya secuencia tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia mostrada en la SEQ ID NO 3,
  - c) una molécula de ácido nucleico cuya secuencia de ácidos nucleicos se desvía de la secuencia de las moléculas de ácido nucleico de a) o b) debido a la degeneración del código genético,
  - d) una molécula de ácido nucleico cuya secuencia de ácidos nucleicos es complementaria a la secuencia de una de las moléculas de ácido nucleico de a), b), c) o parte de las mismas,
  - 15 e) moléculas de ácido nucleico que son fragmentos, variantes alélicas y/o derivados de las moléculas de ácido nucleico mencionadas en a), b), c) o d).

Como consecuencia de su secuencia con codones optimizados, las moléculas de ácido nucleico basadas en la SEQ ID NO 3 e indicadas en el párrafo anterior muestran una expresión mejorada en células vegetales en comparación con las moléculas de ácido nucleico de la SEQ ID NO 1. Los codones se modifican en el sentido de que se adaptan a la frecuencia de uso de los codones de la célula vegetal o de la planta en cuyo genoma están integrados o se van a integrar. El uso de codones de la SEQ ID NO 3 se optimizó específicamente para el maíz y la caña de azúcar.

Debido a la degeneración del código genético, los aminoácidos pueden estar codificados por uno o más codones. Los diferentes organismos usan los codones, cada uno de los cuales codifican un aminoácido, con diferente frecuencia. La adaptación de los codones de una secuencia de ácidos nucleicos codificante a su frecuencia de uso en la célula vegetal o en la planta en cuyo genoma se va a integrar la secuencia a expresar puede contribuir a aumentar la cantidad de proteína traducida y/o a la estabilidad del ARNm correspondiente en las células vegetales o plantas en cuestión. El trabajador experto puede determinar la frecuencia de uso de los codones en las células vegetales o plantas en cuestión mediante el estudio del mayor número posible de secuencias de ácidos nucleicos codificantes del organismo en cuestión para la frecuencia a la que se usan determinados codones para codificar un aminoácido determinado. El trabajador experto conoce la frecuencia de uso de los codones de determinados organismos y se puede llevar a cabo de forma sencilla y rápida con ayuda de programas informáticos. Existen programas informáticos adecuados disponibles públicamente y están disponibles de forma gratuita, entre otros, en Internet (por ejemplo, <http://gcua.schoedl.de/>; <http://www.kazusa.o.jp/codon/>; <http://www.entelechon.com/eng/cutanalysis.html>).

La adaptación de los codones de una secuencia de ácidos nucleicos codificante a su frecuencia de uso en la célula vegetal o en la planta en cuyo genoma se va a integrar la secuencia a expresar se puede efectuar por mutagénesis *in vitro* o, preferentemente, por síntesis *de novo* de la secuencia génica. El trabajador experto conoce procedimientos para la síntesis *de novo* de secuencias de ácidos nucleicos. Una síntesis *de novo* se puede efectuar, por ejemplo, sintetizando en primer lugar oligonucleótidos de ácidos nucleicos individuales, hibridándolos con oligonucleótidos complementarios de modo que formen una doble hebra de ADN y ligando después los oligonucleótidos bicatenarios individuales entre sí de tal forma que se obtenga la secuencia de ácidos nucleicos deseada. La síntesis *de novo* de secuencias de ácidos nucleicos que incluye la adaptación de la frecuencia de uso de los codones a un organismo objetivo específico también se puede encargar a empresas que ofrecen este tipo de servicio (por ejemplo, Entelechon GmbH, Regensburg, Alemania).

Preferentemente, también se excluyen de la presente invención, entre las moléculas de ácido nucleico descritas anteriormente con secuencia con codones optimizados, las moléculas de ácido nucleico que se describen en la memoria descriptiva del documento WO 00/47727.

El término "derivado" significa que las secuencias de estas moléculas difieren de las secuencias de las moléculas de ácido nucleico descritas anteriormente en una o más posiciones y que tienen un alto grado de identidad con estas secuencias. Las desviaciones de las moléculas de ácido nucleico descritas anteriormente se pueden haber originado, por ejemplo, como consecuencia de una eliminación, adición, sustitución, inserción o recombinación. Las variantes alélicas pueden adoptar la forma de variantes naturales o, si no, de variantes generadas sintéticamente o variantes que se han generado por tecnologías de ADN recombinante.

Preferentemente, el triplete de ácido nucleico ATG está unido cadena arriba del extremo 5' de las moléculas de ácido nucleico de acuerdo con la invención descritas anteriormente y a continuación en el presente documento, a menos que ya haya un triplete de este tipo presente en el extremo 5' de la secuencia. En este modo de realización

preferente, ATG forma por tanto el nuevo extremo 5' de la secuencia. ATG actúa como el codón de iniciación de la traducción del ácido nucleico. El triplete ATG puede estar unido al extremo 5' de una molécula de ácido nucleico de acuerdo con la invención por procedimientos conocidos por el trabajador experto. Un procedimiento de uso extendido es la PCR con el uso de cebadores adaptados específicamente, como se describe en los ejemplos adjuntos.

Preferentemente, un triplete de ácido nucleico seleccionado de entre TAA, TAG o TGA está unido cadena abajo del extremo 3' de las moléculas de ácido nucleico de acuerdo con la invención descritas anteriormente y a continuación en el presente documento, a menos que ya haya un triplete de este tipo presente en el extremo 3' de la secuencia. En este modo de realización preferente, TAA, TAG o TGA forman por tanto el nuevo extremo 3' de la secuencia. Los tripletes mencionado actúan como codones de detención en la transcripción/traducción de las moléculas de ácido nucleico. Un triplete TAA, TAG o TGA puede estar unido al extremo 3' de una molécula de ácido nucleico de acuerdo con la invención por procedimientos conocidos por el trabajador experto. Un procedimiento de uso extendido es la PCR con el uso de cebadores adaptados específicamente, como se describe en los ejemplos adjuntos.

Las proteínas codificadas por las moléculas de ácido nucleico de acuerdo con la invención se expresan muy bien en células vegetales y en sistemas de expresión bacterianos.

Las moléculas de ácido nucleico de acuerdo con la invención adoptan la forma de moléculas de ADN, en particular moléculas genómicas. Además, las moléculas de ácido nucleico de acuerdo con la invención pueden ser moléculas de ARN.

Las moléculas de ácido nucleico de acuerdo con la invención que codifican una alternano sacarasa troncada se pueden haber preparado sintéticamente o por técnicas recombinantes. Con este fin, por ejemplo, se puede preparar sintéticamente una molécula de ácido nucleico con una secuencia con codones optimizados de acuerdo con la SEQ ID NO 3 que codifica una alternano sacarasa de longitud completa, como ya se ha descrito anteriormente. Después, se usa la molécula de ácido nucleico con la secuencia de SEQ ID NO 3 para preparar las moléculas de ácido nucleico de acuerdo con la invención, preferentemente por medio de técnicas de PCR usando cebadores adaptados adecuadamente, moléculas de ácido nucleico que codifican una alternano sacarasa troncada. En los ejemplos y los listados de secuencias adjuntos se detallan el procedimiento de PCR y cebadores adecuados.

Las moléculas de ácido nucleico que codifican una alternano sacarasa de longitud completa también se pueden originar a partir de una fuente natural, preferentemente a partir de microorganismos, más preferentemente a partir de bacterias, aún más preferentemente a partir de bacterias grampositivas y lo más preferentemente a partir de bacterias de especies de *Leuconostoc*, en particular de *Leuconostoc mesenteroides*. En el documento WO 00/47727 se describen moléculas de ácido nucleico que codifican una alternano sacarasa de longitud completa a partir de *Leuconostoc mesenteroides* y que tienen una secuencia como se muestra en la SEQ ID NO 1. Por lo tanto, el trabajador experto puede consultar la divulgación del documento WO 00/47727 para obtener una molécula de ácido nucleico como se muestra en la SEQ ID NO 1. Una molécula de ácido nucleico con la SEQ ID NO 1, a su vez, se puede usar para preparar las moléculas de ácido nucleico de acuerdo con la invención que codifican una alternano sacarasa troncada. Preferentemente, la preparación de las moléculas de ácido nucleico de acuerdo con la invención a partir de moléculas de ácido nucleico con una secuencia como se muestra en la SEQ ID NO 1 se lleva a cabo por medio de técnicas de PCR. También se pueden preparar moléculas de ácido nucleico de acuerdo con la invención a partir de una molécula de ácido nucleico con la SEQ ID NO 3 por medio de PCR, de forma análoga a la preparación de las moléculas de ácido nucleico de acuerdo con la invención a partir de una molécula de ácido nucleico con la SEQ ID NO 1 como se describe en los ejemplos. La preparación de los cebadores adaptados adecuadamente forma parte de las competencias de un experto en la técnica.

Se entiende que una proteína o enzima con la actividad enzimática o biológica de una alternano sacarasa (E.C. 2.4.1.140) quiere decir una enzima que cataliza la conversión de sacarosa en alternano y fructosa. Esta conversión se puede producir en presencia o ausencia de aceptores externos (por ejemplo, maltosa, isomaltosa, isomaltotriosa y similares). En ausencia de aceptores externos, las alternano sacarosas, partiendo de la sacarosa; catalizarán la liberación de fructosa y alternano de alto peso molecular, un polisacárido compuesto por unidades de glucosa, cuyo esqueleto consiste en unidades de glucosa que están unidas entre sí, preferentemente de forma alterna, por enlaces  $\alpha$ -1,3- y  $\alpha$ -1,6-glucosídicos, donde la estructura puede contener también ramificaciones. En presencia de aceptores externos tales como, por ejemplo, maltosa, isomaltosa, isomaltotriosa y metil-alfa-D-glucano, la alternano sacarasa puede catalizar, en esos aceptores de polisacáridos, la síntesis de cadenas de alfa-D-glucano en las que los residuos de glucosa están unidos, predominantemente de forma alterna, con enlaces alfa-1,6- y alfa-1,3-glucosídicos, y de fructosa. En función del aceptor empleado, los productos obtenidos tienen diferentes estructuras. La actividad enzimática de una alternano sacarasa se puede detectar, por ejemplo, como se describe por Lopez-Munguia (Annals New York Academy of Sciences 613 (1990), 717-722) o en los ejemplos de la presente solicitud. Se puede definir una unidad de actividad (1 u) como la cantidad de enzima que, en un minuto, da lugar a la liberación de 1  $\mu$ mol de fructosa.

En el contexto de la presente invención, el término "identidad" significa una identidad de secuencia a lo largo de toda la extensión de una secuencia de nucleótidos o una identidad de secuencia a lo largo de toda la extensión de una secuencia de aminoácidos de al menos el 60 %, preferentemente al menos el 70 %, más preferentemente al menos

el 80 %, en particular al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 %. Por tanto, la presente invención comprende también todas las modificaciones de todas las secuencias de ácidos nucleicos o de aminoácidos de acuerdo con la invención que se describen en el presente documento que tienen el porcentaje de identidad mencionado anteriormente.

- 5 En relación con la presente invención, se entiende que el término "identidad" significa el número de aminoácidos/nucleótidos coincidentes (identidad) con otras proteínas/aminoácidos, expresado como porcentaje. Preferentemente, la identidad entre dos secuencias de aminoácidos pertinentes o entre dos secuencias de ácidos nucleicos pertinentes se determina con la ayuda de programas informáticos. Si las secuencias que se comparan entre sí son de distinta longitud, la identidad se debe determinar de tal forma que el número de aminoácidos que la
- 10 secuencia más corta comparte con la secuencia más larga determine el porcentaje de identidad. Preferentemente, la identidad se determina por medio del programa informático conocido y disponible públicamente ClustalW (Thompson et al., Nucleic Acids Research 22 (1994), 4673-4680). La definición de identidad en el programa ClustalW y los procedimientos para determinarla, que están disponible públicamente, se incorporan expresamente por referencia. ClustalW ha sido puesto a disposición del público por Julie Thompson (Thompson@EMBL-Heidelberg.DE) y Toby
- 15 Gibson (Gibson@EMBL-Heidelberg.DE), del European Molecular Biology Laboratory, Meyerhofstrasse 1, D 69117 Heidelberg, Alemania. Clustal W también se puede descargar de diversas páginas de Internet, entre otras, la del IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire, B.P. 163, 67404 Illkirch Cedex, Francia; ftp://ftp-igbmc.u-strasbg.fr/pub/) y el EBI (ftp://ftp.ebi.ac.uk/pub/software/) y todas las páginas de Internet del EBI reflejadas (European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, RU).
- 20 Preferentemente, para determinar la identidad entre proteínas descritas dentro del ámbito de la presente invención y otras proteínas se usa la versión 1.8 del programa informático ClustalW. En este contexto, se configurarán los siguientes parámetros: KTUPLE=1, TOPDIAG=5, WINDOW=5, PAIRGAP=3, GAOPEN=10, GAPEXTEND=0,05, GAPDIST=8, MAXDIV=40, MATRIX=GONNET, ENDGAPS (OFF), NOPGAP, NOHGAP. Preferentemente, para determinar la identidad entre, por ejemplo, la secuencia de nucleótidos de las moléculas de ácido nucleico descritas dentro del ámbito de la presente invención y la secuencia de nucleótidos de otras moléculas de ácido nucleico, se
- 25 usa la versión 1.8 del programa informático ClustalW. En este contexto, se configurarán los siguientes parámetros: KTUPLE=2, TOPDIAGS=4, PAIRGAP=5, DNAMATRIX: IUB, GAOPEN=10, GAPEXT=5, MAXDIV=40, TRANSITIONS: no ponderadas.

30 En la presente invención, preferentemente, la secuencia de comparación ya la secuencia modificada (es decir, las secuencias comparadas) tienen el mismo número de nucleótidos o aminoácidos. Aún más preferentemente, la secuencia de ácidos nucleicos o de aminoácidos modificada se obtiene por intercambio de ácidos nucleicos o aminoácidos en la secuencia de comparación. Esto es, la secuencia de ácidos nucleicos o aminoácidos modificada se basa en la misma sección de la SEQ 1, 2 o 3 que la secuencia de comparación. Por ejemplo, una molécula de ácido nucleico modificada cuya secuencia tiene al menos el 70 %, preferentemente al menos el 80 %, más

35 preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia se basa en una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 118 hasta el nucleótido de la posición 4140, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o SEQ ID NO 3 (secuencia de comparación), pero dicha molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 118 hasta el nucleótido de la posición 4140, inclusive, se

40 modificó por intercambio de nucleótidos con el fin de obtener la secuencia modificada en la que se mantiene el número de nucleótidos. Preferentemente, se puede aplicar el mismo principio a todas las demás secuencias de ácidos nucleicos o aminoácidos modificadas descritas en el presente documento, que se definen por un determinado porcentaje de identidad en relación con una secuencia de comparación.

45 También se describe una molécula de ácido nucleico cuya secuencia de ácidos nucleicos difiere de la secuencia de una de las moléculas de ácido nucleico descritas anteriormente, en particular de una molécula de ácido nucleico como se especifica en la reivindicación 1, por una o más mutaciones puntuales dentro de la secuencia de ácidos nucleicos, manteniéndose la actividad enzimática de una alternano sacarasa en la proteína codificada.

50 Es posible introducir las mutaciones mencionadas anteriormente en las moléculas de ácido nucleico por medio de técnicas convencionales de biología molecular (véase, por ejemplo, Sambrook et al., 2001, Molecular Cloning, A Laboratory Manual, 3ª edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Un ejemplo que es factible es la introducción de mutaciones puntuales donde una modificación de la secuencia de aminoácidos tiene un efecto, por ejemplo, en la actividad de la enzima o en la regulación de la enzima. De este modo, es posible generar, por ejemplo, mutantes con una estereoselectividad o una regioselectividad modificadas o un valor de Km modificado, o que ya no están sujetos a los mecanismos reguladores que normalmente están

55 presentes en la célula y que actúan a través de regulación alostérica o modificación covalente. Mutaciones de este tipo, que también se pueden introducir en una molécula de ácido nucleico de acuerdo con la invención, se describen, por ejemplo, en la solicitud de patente publicada US 20060127328A1. Preferentemente, se excluyen de la presente invención las moléculas de ácido nucleico descritas en el documento US20060127328A1, en particular en los párrafos [0063], [0064] y [0066]. Preferentemente, también se excluyen de la presente invención las alternano

60 sacarosas de longitud truncada que están codificadas por estas moléculas de ácido nucleico y que se describen en particular en el párrafo [0091] del documento US20060127328A1.

Con el fin de someter células procariotas a manipulación recombinante, se pueden introducir las moléculas de ácido nucleico de acuerdo con la invención o partes de estas moléculas en plásmidos que permitan una mutación o modificación de secuencia por medio de recombinación de secuencias de ADN. Se pueden llevar a cabo intercambios de bases o se pueden añadir secuencias naturales o sintéticas, con ayuda de procedimientos ordinarios (véase Sambrook et al., 2001, Molecular Cloning: A laboratory manual, 3ª edición, Cold Spring Harbor Laboratory Press, NY, EE. UU.). Para conectar los fragmentos de ADN entre sí, se pueden añadir adaptadores o enlazadores a los fragmentos. Además, es posible emplear manipulaciones que proporcionan sitios de escisión de restricción o que retiran el exceso de ADN o sitios de escisión de restricción. Cuando sean adecuadas inserciones, eliminaciones o sustituciones, se puede usar mutagénesis *in vitro*, "reparación de cebadores", restricción; PCR o ligación (véase Sambrook et al., 2001, Molecular Cloning: A laboratory manual, 3ª edición, Cold Spring Harbor Laboratory Press, NY, EE. UU.). Los procedimientos analíticos que se llevan a cabo en general son el análisis de secuencias, los análisis de restricción y otros procedimientos bioquímicos/de biología molecular.

La invención se refiere también a moléculas de ácido nucleico que hibridan en condiciones rigurosas con las moléculas de ácido nucleico de acuerdo con la invención descritas anteriormente, con la condición de que la proteína codificada tenga la actividad enzimática de una alternano sacarasa.

Para los propósitos de la presente invención, el término "hibridación" significa hibridación en condiciones de hibridación convencionales, preferentemente en condiciones rigurosas, como se describen, por ejemplo, en Sambrook et al., Molecular Cloning, A Laboratory Manual, 3ª edición, (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. De forma particularmente preferente "hibridación" significa que se produce una hibridación en las siguientes condiciones: tampón de hibridación: 2 x SSC; 10 x solución de Denhardt (Ficoll 400 + PEG + BSA; proporción 1:1:1); SDS al 0,1 %; EDTA 5 mM; Na<sub>2</sub>HPO<sub>4</sub> 50 mM; 250 µg/ml de ADN de esperma de Herring; 50 µg/ml de ARNt; o tampón fosfato de sodio 0,25 M a pH 7,2; EDTA 1 mM, SDS al 7 %; T=65 a 68 °C, tampón de lavado: 0,1 x SSC; SDS al 0,1 %, temperatura de lavado: T=65 a 68 °C.

Las moléculas de ácido nucleico que hibridan con moléculas de ácido nucleico de acuerdo con la invención pueden derivar de cualquier organismo deseado; pueden derivar, por lo tanto, bacterias, hongos, animales, plantas o de virus.

Las moléculas de ácido nucleico que hibridan con las moléculas mencionadas anteriormente se pueden aislar, por ejemplo, a partir de colecciones genómicas de colecciones de ADNc. La identificación y el aislamiento de estas moléculas de ácido nucleico se pueden efectuar usando las moléculas de ácido nucleico mencionadas anteriormente o partes de estas moléculas, o los complementos inversos de estas moléculas, por ejemplo, por medio de hibridación por procedimientos ordinarios (véase, por ejemplo, Sambrook et al., 2001, Molecular Cloning, A Laboratory Manual, 3ª edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), o por amplificación por medio de PCR.

Los fragmentos usados como sonda de hibridación también pueden adoptar la forma de fragmentos sintéticos u oligonucleótidos que se han preparado con ayuda de técnicas de síntesis habituales y cuya secuencia coincide esencialmente con la de una molécula de ácido nucleico descrita dentro del ámbito de la presente invención.

Las moléculas que hibridan con las moléculas de ácido nucleico descritas comprenden en particular fragmentos, derivados y variantes alélicas de las moléculas de ácido nucleico mencionadas anteriormente. El término "derivado" significa que las secuencias de estas moléculas difieren de las secuencias de las moléculas de ácido nucleico descritas anteriormente en una o más posiciones y que tienen un alto grado de identidad con estas secuencias. Las desviaciones de las moléculas de ácido nucleico descritas anteriormente se pueden haber originado, por ejemplo, por eliminación, adición, sustitución, inserción o recombinación.

También se describen oligonucleótidos que hibridan específicamente con una molécula de ácido nucleico de acuerdo con la invención. Preferentemente, estos oligonucleótidos tienen una longitud de al menos 10, en particular de al menos 15 y de forma particularmente preferente de al menos 50 nucleótidos. Se caracterizan por que hibridan específicamente con moléculas de ácido nucleico de acuerdo con la invención, es decir, no lo hacen, o lo hacen en un grado muy bajo, con secuencias de ácidos nucleicos que codifican otras proteínas, en particular otras glucosil transferasas. Los oligonucleótidos se pueden usar, por ejemplo, como cebadores para técnicas de amplificación tales como la reacción de PCR, o como sonda de hibridación para el aislamiento de genes relacionados.

En otro modo de realización preferente, las moléculas de ácido nucleico de acuerdo con la invención están enlazadas a elementos reguladores que garantizan la transcripción y la síntesis de un ARN traducible en células procariotas o eucariotas. Estas secuencias reguladoras se detallan en otra parte de la presente divulgación, tanto para microorganismos como para plantas.

La invención se refiere además a vectores, en particular plásmidos, cósmidos, virus, bacteriófagos y otros vectores usados convencionalmente en ingeniería genética, que comprenden las moléculas de ácido nucleico de acuerdo con la invención descritas anteriormente. En un modo de realización preferente de la invención, los vectores de acuerdo con la invención son adecuados para la transformación de células de microorganismos. Preferentemente, estos vectores son adecuados para la transformación de células vegetales. En otro modo de realización preferente, estos

vectores permiten la integración en el genoma de la célula vegetal de las moléculas de ácido nucleico de acuerdo con la invención, en caso apropiado junto con regiones reguladoras flanqueantes. Son ejemplos los vectores binarios que se pueden emplear en la transferencia de genes mediada por agrobacterias, algunos de los cuales ya están disponibles comercialmente. Otros vectores que se pueden usar de forma ventajosa son los que comprenden

5 secuencias que se integran en el genoma de la planta por medio de administración directa de genes. Un ejemplo de una referencia sobre la administración directa de genes en la caña de azúcar es: Bower R, Elliot AR, Potier BAM y Birch RG (1996) High efficiency, microprojectile-mediated cotransformation of sugarcane, using visible o selectable markers. *Molecular Breeding* 2: 239-249.

En un aspecto de la presente invención, los vectores son plásmidos como se describe en el documento

10 US20060127328A1, párrafo [0076].

La expresión de los ácidos nucleicos de la presente invención se puede producir en células procariotas o eucariotas. En el documento US20060127328A1, párrafo [0077], se detallan ejemplos de células adecuadas.

Otros vectores preferentes son los vectores usados en los ejemplos de uso adjuntos.

Una expresión particularmente ventajosa es la expresión de las moléculas de ácido nucleico de acuerdo con la

15 invención en células procariotas o eucariotas que no tienen enzimas que interfieran, tales como, por ejemplo, dextrano sacarasa u otras enzimas formadoras o degradadoras de polisacáridos.

En otro modo de realización, la invención se refiere a células huésped, en particular células procariotas o eucariotas, que están transformadas con una molécula de ácido nucleico de acuerdo con la invención descrita anteriormente o un vector de acuerdo con la invención, y a células que derivan de estas células transformadas y que comprenden

20 una molécula de ácido nucleico de acuerdo con la invención o un vector de acuerdo con la invención. Preferentemente, la célula de acuerdo con la invención es una célula huésped aislada.

En un modo de realización preferente, las células huésped son células de microorganismos. En el contexto de la presente invención, el término "microorganismo" comprende bacterias y todos los protistas (por ejemplo, hongos, en particular levaduras, algas) como se define en Schlegel "Allgemeine Mikrobiologie" (Georg Thieme Verlag, 1985, 1-2). En un modo de realización preferente, la invención se refiere a células de algas y células huésped de los géneros

25 *Aspergillus*, *Bacillus*, en particular *Bacillus subtilis*, *Saccharomyces*, tal como *Saccharomyces cerevisiae*, o *Pichia* (Rodríguez, *Journal of Biotechnology* 33 (1994), 135-146, Romanos, *Vaccine*, Vol. 9 (1991), 901 y sig.). Otras células adecuadas que se pueden usar en la presente invención son *Salmonella typhimurium*, cepas derivadas de *Pseudomonas*, *Streptomyces* y *Staphylococcus*, parásitos tales como parásitos apicomplejos (*Plasmodia*, *Toxoplasma*, *Cryptosporidia*), *Leishmania* o *Trypanosoma*. En un modo de realización particularmente preferente, la invención se refiere a células de *E. coli*.

Es particularmente preferente que la alternano sacarasa truncada se secrete desde la célula huésped. La generación de células huésped de este tipo para la producción de una alternano sacarasa truncada se puede efectuar por procedimientos conocidos por el trabajador experto. Normalmente, la secreción de proteínas por microorganismos

35 está mediada por péptidos señal N-terminales (secuencia señal, péptido líder, péptido de tránsito). Las proteínas con esta secuencia señal pueden atravesar la membrana celular del microorganismo. Se puede lograr una secreción de proteínas uniendo la secuencia de ADN que codifica el péptido señal a la región codificante de alternano sacarasa correspondiente. La secreción la garantiza una secuencia señal que comprende los primeros 39 residuos de aminoácido N-terminales de la SEQ ID NO 2.

En un modo de realización preferente de la invención, las células huésped de acuerdo con la invención no tienen actividades enzimáticas que interfieren, tales como, por ejemplo, de enzimas formadoras o degradadoras de

40 polisacáridos.

Se encuentra una descripción general de los diversos sistemas de expresión, por ejemplo, en *Methods in Enzymology* 153 (1987), 385-516, en Bitter et al. (*Methods in Enzymology* 153 (1987), 516-544) y en Sawers et al. (*Applied Microbiology and Biotechnology* 46 (1996), 1-9), Billman-Jacobe (*Current Opinion in Biotechnology* 7 (1996), 500-4), Hockney (*Trends in Biotechnology* 12, (1994), 456-463), Griffiths et al., *Methods in Molecular Biology* 75 (1997), 427-440). Se puede encontrar una descripción general de sistemas de expresión de levaduras, por ejemplo, en Hensing et al. (*Antonie von Leeuwenhoek* 67(1995), 261-279), Bussineau et al. (*Developments in Biological Standardization* 83 (1994), 13-19), Gellissen et al. (*Antonie von Leeuwenhoek* 62 (1992), 79-93), Fleer (*Current Opinion in Biotechnology* 3 (1992), 486-496), Vedvick (*Current Opinion in Biotechnology* 2 (1991), 742-745) y Buckholz (*Bio/Technology* 9 (1991), 1067-1072).

45

50

En la literatura se describen ampliamente vectores de expresión. Además de un gen marcador de selección y un origen de replicación que garantiza la replicación en el huésped escogido, comprenden, por norma, un promotor bacteriano o vírico y, en la mayoría de los casos, una señal de transcripción. Entre el promotor y la señal de

55 terminación se localiza al menos un sitio de escisión de restricción o un polienlazador, elementos que hacen posible la inserción de una secuencia de ADN codificante. Un elemento que se puede usar como secuencia promotora es la secuencia de ADN que controla de forma natural la transcripción del gen en cuestión, siempre que esté activa en el organismo huésped escogido. Sin embargo, también se pueden sustituir las demás secuencias promotoras por esta

secuencia. No sólo es posible usar promotores que provocan la expresión constitutiva del gen, sino también promotores inducibles, que permiten una regulación dirigida de la expresión del gen cadena abajo.

En la literatura se describen ampliamente secuencias promotoras bacterianas y víricas con estas propiedades. En la literatura se describen suficientemente secuencias reguladoras para su expresión en microorganismos (por ejemplo, *E. coli*, *S. cerevisiae*). Son promotores que permiten una expresión particularmente intensa del gen cadena abajo, por ejemplo, el promotor T7 (Studier et al., *Methods in Enzymology* 185 (1990), 60-89), *lacuv5*, *trp*, *trp-lacUV5* (DeBoer et al., en Rodríguez y Chamberlin (Eds), *Promoters, Structure and Function*; Praeger, Nueva York, (1982), 462-481; DeBoer et al. *Proc. Natl. Acad. Sci. USA* (1983), 21-25), *lpl*, *rac* (Boros et al., *Gene* 42 (1986), 97- 100). Por norma, las cantidades de proteína alcanzan su máximo desde la mitad hacia el final de la fase logarítmica del ciclo de crecimiento de los microorganismos. Por lo tanto, la síntesis de proteínas se lleva a cabo preferentemente usando promotores inducibles. Frecuentemente, dan lugar a mayores rendimientos de proteína que los promotores constitutivos. Como consecuencia de la transcripción y traducción permanentes de un gen clonado, el uso de promotores constitutivos fuertes conduce frecuentemente al hecho de la pérdida de energía para otras funciones celulares, de modo que se ralentiza el crecimiento celular (Bernard R. Glick/Jack J. Pasternak, *Molekulare Biotechnologie* (1995), Spektrum Akademischer Verlag GmbH, Heidelberg Berlín Oxford, p. 342.). Para lograr una cantidad óptima de proteína, por lo tanto, se recurre frecuentemente a un procedimiento en dos etapas. En primer lugar, se cultivan las células huésped en condiciones óptimas hasta alcanzar una densidad celular relativamente alta. En la segunda etapa, se induce la transcripción, en función del tipo de promotor empleado. Un promotor que es particularmente adecuado en este contexto es un promotor *tac* inducible por lactosa o IPTG (= isopropil-beta-D-tiogalactopiranosido) (deBoer et al., *Proc. Natl. Acad. Sci. USA* 80 (1983), 21-25). En la literatura se describen del mismo modo señales de terminación de la transcripción (Young RA (1979) *Transcription termination in the Escherichia coli ribosomal RNA operon rrnC*, *J Biol Chem.* 254: 12725 31; Lynn SP, Bauer CE, Chapman K, Gardner JF. (1985) *Identification and characterization of mutants affecting transcription termination at the threonine operon attenuator.* *J Mol Biol.* 183: 529- 41).

Por norma, la transformación de la célula huésped microbiana con el ADN que codifica una alternano sacarasa truncada se puede llevar a cabo por procedimientos ordinarios, tales como, por ejemplo, los descritos en Sambrook et al. (*Molecular Cloning: A Laboratory Course Manual*, 3ª edición, (2001), Cold Spring Harbor Press, Nueva York; *Methods in Yeast Genetics, A Laboratory Course Manual*, Cold Spring Harbor Laboratory Press, 1990). La célula huésped se cultiva en medios nutritivos que cumplen los requisitos de la célula huésped en particular usada en cada caso, teniendo en cuenta en particular el pH, la temperatura, la concentración de sales, la oxigenación, los antibióticos, las vitaminas, los oligoelementos y similares.

La invención se refiere además a proteínas con la actividad enzimática de una alternano sacarasa que están truncadas en comparación con una alternano sacarasa de longitud completa definida anteriormente en el extremo N-y/o C-terminal de la secuencia de aminoácidos (en adelante en el presente documento, denominadas también "alternano sacarasa truncada") y que están codificadas por moléculas de ácido nucleico de acuerdo con la invención con la secuencia correspondiente.

En particular, se describen las siguientes proteínas con la actividad enzimática de una alternano sacarasa:

Una proteína con la actividad enzimática de una alternano sacarasa, proteína que se selecciona de entre

a) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y lo más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1295 hasta el aminoácido de la posición 1345 de la SEQ ID NO 2,

b) una proteína con una secuencia de aminoácidos que tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de a), donde, preferentemente, la secuencia de aminoácidos comparada tiene el mismo número de aminoácidos,

c) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y lo más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1355 hasta el aminoácido de la posición 1420 de la SEQ ID NO 2,

d) una proteína con una secuencia de aminoácidos que tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de c), donde, preferentemente, la secuencia de aminoácidos comparada tiene el mismo número de aminoácidos,

e) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y lo más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1430 hasta el aminoácido de la posición 1520 de la SEQ ID NO 2, o

f) una proteína con una secuencia de aminoácidos que tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de e), donde, preferentemente, la secuencia de aminoácidos comparada tiene el mismo número de aminoácidos.

Se divulga además una proteína con la actividad enzimática de una alternano sacarasa, proteína que se selecciona de entre

a) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y lo más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1305 hasta el aminoácido de la posición 1325 de la SEQ ID NO 2,

b) una proteína con una secuencia de aminoácidos que tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de a), donde, preferentemente, la secuencia de aminoácidos comparada tiene el mismo número de aminoácidos,

c) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y lo más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1370 hasta el aminoácido de la posición 1390 de la SEQ ID NO 2,

d) una proteína con una secuencia de aminoácidos que tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de c), donde, preferentemente, la secuencia de aminoácidos comparada tiene el mismo número de aminoácidos,

e) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y lo más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1470 hasta el aminoácido de la posición 1520 de la SEQ ID NO 2,

f) una proteína con una secuencia de aminoácidos que tiene al menos el 70 %, preferentemente al menos el 80 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de a), donde, preferentemente, la secuencia de aminoácidos comparada tiene el mismo número de aminoácidos.

Adicionalmente, se describen proteínas con la actividad enzimática de una alternano sacarasa:

a) una proteína con una secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1380, inclusive, de la SEQ ID NO 2,

b) una proteína con una secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2,

c) una proteína con una secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1506, inclusive, de la SEQ ID NO 2.

Se divulgan además las siguientes proteínas con la actividad enzimática de una alternano sacarasa:

d) una proteína con una secuencia desde el aminoácido de la posición 168 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2,

e) una proteína con una secuencia desde el aminoácido de la posición 342 hasta el aminoácido de la posición 1349, inclusive, de la SEQ ID NO 2,

f) una proteína con una secuencia desde el aminoácido de la posición 342 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2.

Se describen además proteínas con la actividad enzimática de una alternano sacarasa cuyas secuencias de aminoácidos tienen al menos el 70 %, más preferentemente al menos el 80 %, en particular al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con las secuencias de aminoácidos especificadas anteriormente, donde, preferentemente, las secuencias de aminoácidos comparadas tienen el mismo número de aminoácidos. El término "identidad" se define como se describe anteriormente, y anteriormente se ha descrito un procedimiento de identificación de la identidad.

En una divulgación, hay una metionina situada cadena arriba del extremo N de las proteínas (o secuencias proteínicas) descritas anteriormente y a continuación en el presente documento que tienen la actividad enzimática de una alternano sacarasa. Por tanto, la metionina forma el extremo N de la secuencia de aminoácidos de acuerdo con la invención. Esta metionina deriva habitualmente de la traducción del ARN que codifica la proteína, donde está codificada por el triplete de ácido nucleico ATG. A menos que ya esté presente en la secuencia de ácidos nucleicos, el triplete ATG se puede unir al extremo 5' de una molécula de ácido nucleico de acuerdo con la invención por el procedimiento conocido por el trabajador experto. Un procedimiento establecido es la PCR usando cebadores adaptados específicamente, como se describe en los ejemplos adjuntos.

Preferentemente, se excluyen de la presente invención las proteínas con la actividad enzimática de una alternano sacarasa descritas en el documento US20060127328A1, en particular las proteínas descritas en el párrafo [0089] y en las figuras 3A y 3B del documento US20060127328A1. Por tanto, preferentemente, se excluyen de la presente invención las proteínas desde el aminoácido de la posición 1 a la 1425, de la posición 1 a la 1349, de la posición 1 a la 1290, de la posición 342 a la 1425 y de la posición 342 a la 1290, estando definidas las posiciones con referencia a la SEQ ID NO 2 adjunta.

La invención se refiere además a un procedimiento de producción de una proteína descrita anteriormente de acuerdo con la invención con la actividad enzimática de una alternano sacarasa, en el que se transforma una célula huésped con

a) una molécula de ácido nucleico cuya secuencia codifica la secuencia de aminoácidos de una proteína descrita anteriormente con la actividad enzimática de una alternano sacarasa, donde la molécula de ácido nucleico se puede obtener a partir de una molécula de ácido nucleico con una secuencia mostrada en la SEQ ID NO 1 o 3, o

b) un vector o plásmido que comprende una molécula de ácido nucleico como se describe en a),

2) se cultiva la célula huésped en medio de cultivo y

3) se aísla la proteína de las células cultivadas y/o del medio de cultivo.

Las secuencias de aminoácidos de las proteínas de acuerdo con la invención con la actividad enzimática de una alternano sacarasa se detallan más arriba y se definen con referencia a la SEQ ID NO 2. Basándose en las secuencias de aminoácidos, el trabajador experto puede determinar secuencias de ácidos nucleicos codificantes adecuadas, que pueden diferir entre sí debido a la degeneración del código genético. Por ejemplo, a partir de la SEQ ID NO 1, que comprende la secuencia de aminoácidos de una alternano sacarasa de longitud completa y una secuencia de nucleótidos codificante, es posible determinar la correspondiente secuencia de ácidos nucleicos que codifica la proteína de acuerdo con la invención (alternano sacarasa truncada). Sobre la base de las mismas posiciones de nucleótidos que en la SEQ ID NO 1, que determinan el comienzo y el final de la secuencia de nucleótidos para la proteína de acuerdo con la invención que se desea en cada caso, también es posible determinar, sobre la base de la SEQ ID NO 3, una secuencia de ácidos nucleicos que codifica la proteína. Del mismo modo, es posible traducir la secuencia de nucleótidos de acuerdo con la SEQ ID NO 3 en una secuencia de aminoácidos y, con referencia al comienzo y al final de la secuencia de aminoácidos deseada, como se detalla para las proteínas de acuerdo con la invención, determinar la secuencia de nucleótidos codificante correspondiente.

Una vez determinada la secuencia, se puede generar una molécula de ácido nucleico con esta secuencia a partir de una molécula de ácido nucleico con una secuencia de acuerdo con la SEQ ID NO 1 o la SEQ ID NO 3. Preferentemente, la generación de una molécula de ácido nucleico cuya secuencia codifica la secuencia de aminoácidos de una proteína de acuerdo con la invención con la actividad enzimática de una alternano sacarasa se lleva a cabo por medio de una técnica de PCR, partiendo de moléculas de ácido nucleico con una secuencia de acuerdo con la SEQ ID NO 1 o 3. La generación de cebadores adaptados adecuadamente forma parte de los conocimientos de un experto en la técnica.

En el procedimiento descrito anteriormente, también es posible emplear todas moléculas de ácido nucleico de acuerdo con la invención descritas anteriormente que codifican una proteína con la actividad enzimática de una alternano sacarasa y que está truncada en el extremo N- y/o C-terminal de la secuencia de aminoácidos en comparación con una alternano sacarasa de longitud completa. En particular, es posible emplear las moléculas de ácido nucleico definidas en la reivindicación 1.

En cuanto a los vectores y plásmidos que se pueden emplear en el procedimiento anterior, en otro lugar se hace referencia a la divulgación en cuestión. Además, en otro lugar de la presente descripción se divulgan con detalle células huésped y técnicas para la transformación de células huésped microbianas y vegetales que se pueden emplear en el procedimiento anterior. Son preferentes células huésped vegetales o células huésped microbianas, por ejemplo, células de *Escherichia coli*.

Las células huésped se hacen crecer, es decir, se cultivan y propagan, en un medio nutriente adecuado y en condiciones de cultivo adecuadas. Los medios nutriente y las condiciones de cultivo para células microbianas y vegetales habituales son conocidas por el trabajador experto y se describen en la literatura especializada (por ejemplo, Sambrook y Russell Molecular cloning: a laboratory manual (tercera edición) Cold Spring Harbor Laboratory Press, NY, EE. UU. (2001)).

Por último, se aísla la proteína a partir de la célula huésped y/o a partir del medio de cultivo. Con este fin, por norma, el material celular producido se homogeneiza con ayuda de procedimientos físicos, químicos y/o enzimáticos adecuados. Las proteínas con la actividad enzimática de una alternano sacarasa se pueden aislar a partir del material homogeneizado.

Se puede llevar a cabo la purificación adicional de las proteínas con ayuda de procedimientos habituales conocidos en la técnica. Para obtener un extracto en bruto, se trata la solución obtenida en el procedimiento de disrupción, por ejemplo, con ayuda de procedimientos de extracción, procedimientos de centrifugación y procedimientos de filtración. La precipitación de las proteínas a partir del extracto en bruto y los procedimientos de ultrafiltración son procedimientos eficaces para concentrar proteínas a partir de una cantidad sustancial de fluido. Son agentes de precipitación adecuados, entre otros, las sales inorgánicas, por ejemplo, sulfato de sodio y sulfato de amonio, los disolventes orgánicos, por ejemplo, alcoholes, y los polímeros, por ejemplo, polietilenglicol. Para retirar el agente de precipitación usado, se puede llevar a cabo posteriormente un procedimiento de diálisis. Se puede realizar una segunda purificación adicional con ayuda de procedimientos cromatográficos y procedimientos de distribución, por ejemplo, usando sistemas de fase acuosa. Estos procedimientos incluyen, entre otros, la cromatografía de adsorción, la cromatografía de intercambio iónico, la cromatografía en gel y la cromatografía de afinidad.

En otro aspecto, la proteína de acuerdo con la invención con la actividad enzimática de una alternano sacarasa se puede sintetizar por el procedimiento de síntesis de proteínas de Merrifield. Por tanto, las alternano sacarosas truncadas producidas sintéticamente son otro aspecto de la presente invención.

En otro aspecto, la proteína de acuerdo con la invención con la actividad enzimática de una alternano sacarasa también se puede preparar con sistemas de expresión/traducción *in vitro* sin células. Es posible una traducción *in vitro*, por ejemplo, en extracto de germen de trigo. Sin embargo, los procedimientos *in vitro* no están restringidos en modo alguno y en principio se puede usar cualquier procedimiento conocido en la técnica. De acuerdo con la invención, es posible usar células huésped y vectores que permitan la producción de la alternano sacarasa truncada en ausencia de sacarosa, de modo que ya no haga falta una separación adicional de la alternano sacarasa truncada de los polisacáridos.

La purificación de la alternano sacarasa truncada producida por las células huésped se puede efectuar con ayuda de procedimientos tradicionales de purificación tales como la precipitación, la cromatografía de intercambio iónico, la cromatografía de afinidad, la filtración en gel, la HPLC, la cromatografía en fase inversa y similares.

Como consecuencia de la modificación de las moléculas de ácido nucleico de acuerdo con la invención que se expresan en las células huésped y que codifican una alternano sacarasa truncada, se puede producir un polipéptido en la célula huésped que, como consecuencia de determinadas características, se puede aislar más fácilmente del medio de cultivo. Por tanto, es posible expresar la proteína a expresar como una proteína de fusión con otra secuencia polipeptídica (marca) cuyas propiedades de unión específicas hagan posible el aislamiento de la proteína de fusión por cromatografía de afinidad (por ejemplo, Hopp et al., *Bio/Technology* 6 (1988), 1204-1210; Sassenfeld, *Trends Biotechnol.* 8 (1990), 88-93). Son marcas usadas convencionalmente, por ejemplo, la marca Estrep, la marca His, la marca FLAG, la marca T7, la marca S.

Además, las secuencias de ácidos nucleicos de la presente invención se pueden modificar en el extremo 5' o 3' por secuencias que codifican marcas en el extremo N o C de la proteína, por ejemplo, también por las que se describen en el párrafo [0071] del documento US20060127328A1.

Proporcionando las moléculas de ácido nucleico de acuerdo con la invención, con ayuda de procedimientos recombinantes es posible generar células vegetales que expresan alternano sacarasa truncada. Por tanto, la invención se refiere también a células vegetales transgénicas que se han transformado con una molécula de ácido nucleico de acuerdo con la invención o un vector de acuerdo con la invención o que derivan de estas células, donde la molécula de ácido nucleico que codifica la proteína con la actividad enzimática de una alternano sacarasa está bajo el control de elementos reguladores que inician la transcripción de un ARNm traducible en células vegetales (promotores). Pueden adoptar la forma de promotores homólogos o heterólogos. Los promotores pueden adoptar la forma de promotores constitutivos, de promotores específicos de tejido, de promotores específicos del desarrollo o, si no, de promotores que están regulados por factores externos (por ejemplo, después de la aplicación de sustancias

químicas, como consecuencia del efecto de factores abióticos tales como el calor y/o el frío, la sequía, la incidencia de enfermedades y similares).

En general, un promotor que es adecuado para la expresión de una molécula de ácido nucleico exógena es un promotor que es activo en células vegetales. Son ejemplos de promotores adecuados el promotor de ARN 35S del virus del mosaico de la coliflor o el promotor de la ubiquitina del maíz o el promotor YLCV de *Cestrum* (virus del rizado amarillo de la hoja; en el documento WO 01 73087; Stavolone et al., 2003, *Plant Mol. Biol.* 53, 703-713) para la expresión constitutiva, el promotor del gen de la patatina B33 (Rocha-Sosa et al., *EMBO J.* 8 (1989), 23- 29) para una expresión específica de tubérculo en patatas, o un promotor específico de fruto para tomates tal como, por ejemplo, el promotor de la poligalacturonasa del tomate (Montgomery et al., 1993, *Plant Cell* 5, 1049-1062) o el promotor E8 del tomate (Metha et al., 2002, *Nature Biotechnol.* 20 (6), 613-618) o el promotor de la ACC oxidasa del melocotón (Moon y Callahan, 2004, *J. Experimental Botany* 55 (402), 1519-1528), o un promotor que garantice la expresión únicamente en tejidos fotosintéticamente activos, por ejemplo, el promotor ST-LS I (Stockhaus et al., *Proc. Natl. Acad. Sci. USA* 84 (1987), 7943-7947; Stockhaus et al., *EMBO J.* 8 (1989), 2445-2451), o, para la expresión específica de endospermo, el promotor HMWG del trigo, el promotor USP, el promotor de la faseolina, los promotores de los genes Zein del maíz (Pedersen et al., *Cell* 29 (1982), 1015-1026; Quatroccio et al., *Plant Mol. Biol.* 15 (1990), 81-93), un promotor de la glutelina (Leisy et al., *Plant Mol. Biol.* 14 (1990), 41-50; Zheng et al., *Plant J.* 4 (1993), 357- 366; Yoshihara et al., *FEBS Lett.* 383 (1996), 213-218), un promotor de la globulina (Nakase et al., 1996, *Gene* 170 (2), 223-226), un promotor de la prolamina (Qu y Takaiwa, 2004, *Plant Biotechnology Journal* 2 (2), 113-125) o un promotor Shrunken-1 (Werr et al., *EMBO J.* 4 (1985), 1373-1380). Sin embargo, también es posible usar promotores que se activan únicamente en un punto temporal que viene determinado por factores externos (véase, por ejemplo, el documento WO 9307279). En este contexto pueden ser de especial interés los promotores de proteínas de choque térmico, que permiten una inducción sencilla. Además, es posible, usar promotores específicos de semilla tales como, por ejemplo, el promotor USP de *Vicia faba*, que garantiza una expresión específica de semilla en *Vicia faba* y otras plantas (Fiedler et al., *Plant Mol. Biol.* 22 (1993), 669-679; Baumlein et al., *Mol. Gen. Genet.* 225 (1991), 459-467).

El uso de promotores que se encuentran en el genoma de virus que infectan algas también es adecuado para la expresión de secuencias de ácidos nucleicos en plantas (Mitra et al., 1994, *Biochem. Biophys Res Commun* 204 (1), 187-194; Mitra y Higgins, 1994, *Plant Mol Biol* 26 (1), 85-93, Van Etten et al., 2002, *Arch Virol* 147, 1479-1516).

En el contexto de la presente invención, se entiende que el término "específico de tejido" significa la limitación predominante de una infección (por ejemplo, iniciación de la transcripción) a un tejido específico.

En el contexto de la presente invención, se ha de entender que los términos "células de tubérculo, célula de fruto o célula de endospermo" significan todas las células que están presentes en un tubérculo, un fruto o en una semilla.

En el contexto de la presente invención, se entiende que el término "promotor homólogo" significa un promotor presente de forma natural en células vegetales o plantas que se han usado para la generación de células vegetales modificadas genéticamente de acuerdo con la invención o plantas modificadas genéticamente de acuerdo con la invención (homólogo respecto a la célula vegetal o planta), o un promotor que regula la regulación de la expresión de un gen del organismo a partir del cual se ha aislado la correspondiente molécula de ácido nucleico exógena que codifica una proteína (homólogo respecto a la molécula de ácido nucleico a expresar).

En el contexto de la presente invención, se entiende que el término "promotor heterólogo" significa un promotor que no está presente de forma natural en células vegetales o plantas que se han usado para la generación de células vegetales modificadas genéticamente de acuerdo con la invención o plantas modificadas genéticamente de acuerdo con la invención (heterólogo respecto a la célula vegetal o planta), o un promotor que no regula de forma natural la expresión de dicha molécula de ácido nucleico exógena en el organismo a partir del cual se ha aislado la correspondiente molécula de ácido nucleico exógena que codifica una proteína (heterólogo respecto a la molécula de ácido nucleico a expresar).

También es posible que esté presente una secuencia de terminación (señal de poliadenilación), que sirve para añadir una cola poli-A al transcrito. Se supone que la cola poli-A ejerce una función en la estabilización del transcrito. Estos elementos se describen en la literatura (véase Gielen et al., *EMBO J.* 8 (1989), 23-29) y se pueden sustituir como se desee.

La presente divulgación también se refiere de forma explícita a los promotores y secuencias de terminación que se describen en el documento WO 00/47727, páginas 21/22.

También es posible que estén presente secuencias intrónicas entre el promotor y la región codificante de la molécula de ácido nucleico exógena. Estas secuencias intrónicas pueden dar lugar a la estabilidad de la expresión y a un aumento de la expresión en plantas (Callis et al., 1987, *Genes Devel.* 1, 1183-1200; Luehrsen, y Walbot, 1991, *Mol. Gen. Genet.* 225, 81-93; Rethmeier et al., 1997, *Plant Journal* 12(4), 895-899; Rose y Beliakoff, 2000, *Plant Physiol.* 122 (2), 535-542; Vasil et al., 1989, *Plant Physiol.* 91, 1575-1579; XU et al., 2003, *Science in China Series C* vol.46 N.º 6, 561-569). Son secuencias intrónicas adecuadas, por ejemplo, el primer intrón del gen *sh1* del maíz, el primer intrón del gen de la poliubiquitina 1 del maíz, el primer intrón del gen *EPSPS* del arroz, o uno de los dos primeros

intrones del gen PAT1 de Arabidopsis.

Introduciendo la actividad de las proteínas de acuerdo con la invención, por ejemplo, mediante la expresión de moléculas de ácido nucleico adecuadas, es posible producir alternano en células vegetales modificadas genéticamente de forma adecuada. Esto permite la expresión de las moléculas de ácido nucleico de acuerdo con la invención en células vegetales, de modo que se puede introducir una actividad adicional de la alternano sacarasa en cuestión, que no está presente en la natural. Además, es posible modificar las moléculas de ácido nucleico de acuerdo con la invención por procedimientos conocidos por el trabajador experto con el fin de obtener alternano sacarosas truncadas de acuerdo con la invención, por ejemplo, con dependencias de la temperatura modificadas, o especificidades de sustrato o de producto. Estos procedimientos se describen con más detalle en la solicitud de patente WO00/47727.

Existen múltiples técnicas disponibles para la generación de células vegetales de acuerdo con la invención, o de plantas que comprenden células vegetales de acuerdo con la invención, para la integración (estable) de moléculas de ácido nucleico en una célula huésped vegetal. Estas técnicas comprenden la transformación de células vegetales con ADN-T usando *Agrobacterium tumefaciens* o *Agrobacterium rhizogenes* como agentes de transformación, la fusión de protoplastos, la inyección, la electroporación de ADN, la introducción del ADN por medio de un enfoque biolístico, y otras posibilidades (revisión en "Transgenic Plants", Leandro ed., Humana Press 2004, ISBN 1-59259-827-7).

El uso de la transformación de células vegetales mediada por agrobacterias se ha estudiado extensamente y se ha descrito suficientemente en el documento EP 120516 y Hoekema, en: *The Binary Plant Vector System* Offsetdrukkerij Kanters B.V. Alblasterdam (1985), Chapter V; Fraley et al., *Crit. Rev. Plant Sci.* 4, 1-46 y en An et al. *EMBO J.* 4, (1985), 277-287). Para la transformación de la patata, véase, por ejemplo, Rocha-Sosa et al., *EMBO J.* 8, (1989), 29-33), para la transformación de tomates, véase, por ejemplo, el documento US 5.565.347.

Para la transformación de plantas monocotiledóneas por medio de vectores, también se han descrito transformaciones basadas en *Agrobacterium* (Chan et al., *Plant Mol. Biol.* 22, (1993), 491-506; Hiei et al., *Plant J.* 6, (1994) 271-282; Deng et al, *Science en China* 33, (1990), 28-34; Wilpink et al., *Plant Cell Reports* 11, (1992), 76-80; May et al., *Bio/Technology* 13, (1995), 486-492; Conner y Domisse, *Int. J. Plant Sci.* 153 (1992), 550-555; Ritchie et al, *Transgenic Res.* 2, (1993), 252-265). Un sistema alternativo para la transformación de plantas monocotiledóneas es la transformación por medio de un enfoque biolístico (Wan y Lemaux, *Plant Physiol.* 104, (1994), 37-48; Vasil et al., *Bio/Technology* 11 (1993), 1553-1558; Ritala et al., *Plant Mol. Biol.* 24, (1994), 317-325; Spencer et al., *Theor. Appl. Genet.* 79, (1990), 625-631), las transformaciones de protoplastos, la electroporación de células parcialmente permeabilizadas, la introducción de ADN por medio de fibras de vidrio. La transformación del maíz, en particular, se describe repetidamente en la literatura (véanse, por ejemplo, los documentos WO95/06128, EP0513849, EP0465875, EP0292435; Fromm et al., *Biotechnology* 8, (1990), 833-844; Gordon-Kamm et al., *Plan Cell* 2, (1990), 603-618; Koziel et al., *Biotechnology* 11 (1993), 194-200; Moroc et al., *Theor. Appl. Genet.* 80, (1990), 721-726). Se ha descrito igualmente la transformación de otras hierbas, tales como, por ejemplo, el pasto varilla (*Panicum virgatum*) (Richards et al., 2001, *Plant Cell Reporters* 20, 48-54).

También se ha descrito la transformación con éxito de otras especies de cereales, por ejemplo, para la cebada (Wan y Lemaux, véase anteriormente; Ritala et al., véase anteriormente; Krens et al., *Nature* 296, (1982), 72-74) y para el trigo (Nehra et al., *Plant J.* 5, (1994), 285-297; Becker et al., 1994, *Plant Journal* 5, 299-307). Todos los procedimientos anteriores son adecuados para los fines de la presente invención.

Las células vegetales modificadas genéticamente y las plantas modificadas genéticamente que contienen una molécula de ácido nucleico exógena se pueden distinguir de las células vegetales naturales, o las plantas naturales que no comprenden dicha molécula de ácido nucleico exógena, entre otros, por el hecho de que comprenden una molécula de ácido nucleico exógena que no está presente de forma natural en células vegetales naturales o plantas naturales. La integración de una molécula de ácido nucleico exógena en una célula vegetal o planta se puede detectar por procedimientos que son conocidos por el trabajador experto, tales como, por ejemplo, análisis de transferencia Southern o por medio de PCR.

Preferentemente, la molécula de ácido nucleico exógena está integrada en el genoma; de forma especialmente preferente, la molécula de ácido nucleico exógena está integrada de forma estable en el genoma de las células vegetales de acuerdo con la invención o plantas de acuerdo con la invención. En el contexto de la presente invención, se entiende que el término "molécula de ácido nucleico integrada de forma estable" significa la integración de una molécula de ácido nucleico en el genoma de la planta. Una molécula de ácido nucleico integrada de forma estable se distingue por el hecho de que se replica al replicarse el sitio de integración en cuestión junto con las secuencias de ácidos nucleicos homólogas que flanquean el sitio de integración, de modo que el sitio de integración de la hebra de ADN hija replicada está rodeado por las mismas secuencias de ácidos nucleicos que en la hebra original leída, que actúa como matriz para la replicación.

La integración de una molécula de ácido nucleico en el genoma de una célula vegetal o una planta se puede detectar por procedimientos genéticos y/o de biología molecular. Una integración estable de una molécula de ácido nucleico en el genoma de una célula vegetal o en el genoma de una planta se distingue por el hecho de que la

molécula de ácido nucleico integrada de forma estable está presente en el mismo entorno genómico en la progenie a la que se ha pasado dicha molécula de ácido nucleico que en la generación original. Para que se pueda detectar la presencia de una integración estable de una secuencia de ácidos nucleicos en el genoma de una célula vegetal o en el genoma de una planta por procedimientos conocidos por el trabajador experto, entre otros, con ayuda del análisis de bandas Southern, el análisis de RFLP (polimorfismos en la longitud de los fragmentos de restricción) (Nam et al., 1989, *The Plant Cell* 1, 699-705; Leister y Dean, 1993, *The Plant Journal* 4 (4), 745-750), procedimientos basados en PCR tales como, por ejemplo, el análisis de polimorfismos en la longitud de fragmentos amplificados (AFLP) (Castiglioni et al., 1998, *Genetics* 149, 2039-2056; Meksem et al., 2001, *Molecular Genetics and Genomics* 265, 207-214; Meyer et al., 1998, *Molecular and General Genetics* 259, 150-160), o el uso de fragmentos amplificados que se han escindido con endonucleasas de restricción (secuencias polimórficas amplificadas y escindidas, CAPS) (Koniczny y Ausubel, 1993, *The Plant Journal* 4, 403-410; Jarvis et al., 1994, *Plant Molecular Biology* 24, 685-687; Bachem et al., 1996, *The Plant Journal* 9 (5), 745-753).

En el contexto de la presente invención, se entiende que el término "genoma" significa todo el material hereditario que está presente en una célula vegetal. El trabajador experto sabe que no sólo el núcleo comprende material hereditario, sino también otros compartimentos (por ejemplo, los plastos, las mitocondrias). Una molécula de ácido nucleico de acuerdo con la invención se puede integrar en el genoma nuclear, pero también en el genoma de otros compartimentos tales como, por ejemplo, los plastos.

Además, preferentemente, las células vegetales de acuerdo con la invención se pueden distinguir de células vegetales naturales por al menos una de las siguientes características: si las células vegetales transgénicas comprenden transcritos de las moléculas de ácido nucleico de acuerdo con la invención que se han introducido, esto se puede comprobar, por ejemplo, por análisis de bandas Northern. Preferentemente, las células vegetales de acuerdo con la invención comprenden una alternano sacarasa truncada que está codificada por una molécula de ácido nucleico de acuerdo con la invención que se ha introducido. Esto se puede detectar, por ejemplo, por procedimientos inmunológicos, en particular por un análisis de bandas Western. Si una célula vegetal o planta de acuerdo con la invención comprende una alternano sacarasa truncada, de forma ventajosa, también se puede distinguir una célula vegetal o planta de acuerdo con la invención por el hecho de que comprende una actividad alternano sacarasa en comparación con las células naturales/plantas naturales. La actividad enzimática de una alternano sacarasa se puede medir como se describe, por ejemplo, en Lopez-Munguia et al. (*Annals New York Academy of Sciences* 613, (1990), 717-722) o como se especifica en los ejemplos de la presente invención.

En otro modo de realización preferente, las células vegetales son células de patata, caña de azúcar, maíz, tabaco, remolacha azucarera, sorgo azucarero, arroz o trigo.

Las células vegetales transgénicas se pueden regenerar por técnicas conocidas por el trabajador experto para dar plantas intactas.

Las plantas que se pueden obtener por regeneración de las células vegetales transgénicas de acuerdo con la invención con otro asunto objeto de la presente invención.

En la mayoría de las plantas, los fotoasimilados que se han formado en una planta durante la fotosíntesis en forma de azúcares se transportan a los órganos objetivo en cuestión, principalmente en forma de sacarosa. Dado que la sacarosa es el sustrato para la reacción de polimerización de la alternano sacarasa truncada de acuerdo con la invención, en principio, es posible modificar todas las plantas, tanto plantas monocotiledóneas como dicotiledóneas, con ayuda de la molécula de ácido nucleico de acuerdo con la invención con respecto a la expresión de la alternano sacarasa truncada.

La expresión de las moléculas de ácido nucleico de acuerdo con la invención es particularmente ventajosa en los órganos vegetales que tienen mayor contenido en sacarosa o que almacenan sacarosa. Estos órganos son, por ejemplo, la raíz de la remolacha azucarera, el tallo de la caña de azúcar, el tallo de las plantas de maíz o el tallo del sorgo azucarero.

El sitio de biosíntesis de sacarosa en las células vegetales es el citosol. En contraste, el sitio de almacenamiento es la vacuola. En el transporte al tejido de almacenamiento de la remolacha azucarera o de la patata, o en el transporte al endospermo de las semillas, la sacarosa debe migrar a través del apoplasto. Por tanto, los tres compartimentos son adecuados para la expresión de las moléculas de ácido nucleico de acuerdo con la invención para la síntesis de alternano, es decir, el citosol, la vacuola y el apoplasto. Además, los plastos también son adecuados, como se ha demostrado, por ejemplo, por medio de la expresión amiloplastídica de fructosil transferasas bacterianas. Estas fructosil transferasas, que requieren asimismo sacarosa como sustrato, han podido mediar la formación de "amilofructano" en amiloplastos (Smeekens, *Trends in Plant Science* vol. 2, N.º 8, (1997), 286-288). Las publicaciones WO2006/072603 y WO 2006/063862 se refieren a células vegetales y plantas modificadas genéticamente en las que la modificación genética da lugar a la expresión de mutano sacarasa o dextrano sacarasa, respectivamente, en plastos.

En el caso de plantas productoras de almidón tales como, por ejemplo, la patata y el maíz, en las que la biosíntesis de almidón y el almacenamiento de almidón se produce normalmente en los amiloplastos, una expresión de la

alternano sacarasa truncada en el apoplasto, en el citosol o en la vacuola daría lugar a una síntesis adicional de oligo- y/o polisacáridos en estos compartimentos, que, en total, puede suponer un aumento del rendimiento. Este también puede ser el caso de los plastos, ya que aquí no se utiliza ADP-glucosa, sino la sacarosa plastídica.

5 Dado que es posible que las plantas, en particular la patata, separen el almidón sintetizado en los amiloplastos del alternano sintetizado en el apoplasto, en el citosol o en la vacuola, se puede usar la misma planta para obtener almidón y alternano.

10 Además, se conocen plantas transgénicas de patata y maíz en las que, como consecuencia de la inhibición de la enzima ADP-glucosa pirofosforilasa por una construcción antisentido, se inhibe completamente la síntesis de almidón en los tubérculos o en los granos. En su lugar, se acumulan azúcares solubles, en particular sacarosa y glucosa, por ejemplo en los tubérculos, en el caso de la patata (Müller-Röber et al., EMBO J. 11, (1992), 1229-1238). Como consecuencia de la expresión de una alternano sacarasa truncada que utiliza sacarosa como sustrato, se puede producir alternano en el citosol, en la vacuola o en el apoplasto de estas plantas.

15 Es aquí donde, en otro modo de realización de la invención, las células vegetales o plantas de acuerdo con la invención tienen una actividad ADP-glucosa pirofosforilasa reducida (AGPasa), en comparación con células correspondientes de plantas naturales. Las moléculas de ADN que codifican la AGPasa son bien conocidas por el experto en la técnica y se describen, por ejemplo, en Müller-Röber et al. (Mol. Gen. Genet. 224 (1) (1990), 136-146). Como consecuencia de las moléculas de ADN que codifican AGPasa, es posible generar, usando técnicas de ADN recombinante (por ejemplo, un enfoque antisentido, de ribozima o de cosupresión), plantas que tienen una actividad AGPasa reducida. Además, el experto en la técnica conoce mutantes de AGPasa, por ejemplo, del maíz (brittle-2 y shrunken-2) con una actividad AGPasa reducida. Preferentemente, el término "reducida o disminuida" significa una reducción de la actividad de la AGPasa de al menos el 10 %, más preferentemente de al menos el 50 % y aún más preferentemente de al menos el 80 % en comparación con las correspondientes células naturales.

20 Cuando se expresan las moléculas de ácido nucleico de acuerdo con la invención en plantas, en principio, es posible que la proteína sintetizada que tiene la actividad enzimática de una alternano sacarasa esté localizada en cualquier compartimento de la célula vegetal (por ejemplo, el citosol, los plastos, la vacuola, las mitocondrias) o la planta (por ejemplo, el apoplasto). Para lograr la localización en un compartimento específico, se debe enlazar la región codificante, en caso apropiado, con secuencias de ADN que garanticen la localización en el compartimento en cuestión. Se conocen secuencias de este tipo (véanse, por ejemplo, Braun, EMBO J. 11 (1992), 3219-3227; Sonnewald, Plant J. 1 (1991), 95-106; Rocha-Sosa, EMBO J. 8 (1989), 23-29, Neuhaus y Rodgers, 1998, Plant Molecular Biology 38: 127-144). La unión de las secuencias señal con la región codificante de la proteína se lleva a cabo de tal forma que las secuencias de ácidos nucleicos fusionadas al péptido señal y las codificantes de la proteína forman conjuntamente un marco de lectura abierto durante la transcripción (fusión traduccional). Esto es por lo que, después de la traducción del ARN, que es la consecuencia de la transcripción de las secuencias de ácidos nucleicos fusionadas, se forma una proteína que comprende un péptido señal además de la región codificante de la proteína.

25 En principio, una secuencia señal plastídica que se puede usar es cualquier secuencia señal que garantice la importación de proteínas codificadas en el núcleo al interior de los plastos. El trabajador experto sabe cómo obtener secuencias señal plastídicas. Por tanto, es posible, por ejemplo, examinar secuencias codificantes de proteína con ayuda de programas informáticos (1999, Emanuelsson et al., Protein Science 8, 978-984) para determinar la presencia de una secuencia señal plastídica. Existen programas informáticos que realizan este tipo de análisis tanto disponibles comercialmente como disponibles públicamente en Internet (por ejemplo, en <http://www.cbs.dtu.dk/services/ChloroP/>).

30 Por ejemplo, es posible usar la secuencia señal plastídica de la ferredoxina:NADP+ oxidoreductasa (FNR) de la espinaca. Esta secuencia comprende la región no traducida de 5' y la secuencia de péptido de tránsito flanqueante del ADNc de la proteína plastídica ferredoxina:NADP+ oxidoreductasa de la espinaca (nucleótido -171 a +165; Jansen et al., Current Genetics 13 (1988), 517-522).

35 Otra secuencia señal plastídica que se puede usar es, por ejemplo, el péptido de tránsito de la proteína cerosa del maíz más los primeros 34 aminoácidos de la proteína cerosa madura (Klösgen et al., Mol Gen Genet. 217 (1989), 155-161). Además, también es posible usar el péptido de tránsito de la proteína cerosa del maíz (véase anteriormente) sin los primeros 34 aminoácidos de la proteína cerosa madura.

40 También es factible el uso de las siguientes secuencias señal plastídicas: la secuencia señal de la subunidad pequeña de la ribulosa bifsosfato carboxilasa (Wolter et al., Proc. Natl. Acad. Sci. USA 85 (1988), 846-850; Nawrath et al., Proc. Natl. Acad. Sci. USA 91 (1994), 12760-12764); la secuencia señal de la NADP-malato deshidrogenasa (Gallardo et al., Planta 197 (1995), 324-332); la secuencia señal de la glutatión reductasa (Creissen et al., Plant J. 8 (1995), 167-175), la secuencia señal de la EPSPS (en el documento US 5.188.642). Se prefiere el uso de la secuencia señal del gen dul del arroz.

45 La molécula de ácido nucleico recombinante puede comprender una o más secuencias señal, preferentemente dos. En este contexto, las secuencias señal pueden estar presentes en la dirección de la transcripción directamente una

detrás de otra o, si no, pueden estar separadas entre sí en cada caso por lo que se conoce como un espaciador. Se describen secuencias señal plastídicas que comprenden dos secuencias señal, por ejemplo, en los documentos EP 0 508 909 y EP 0 924 299, US 5.510.471.

5 En un modo de realización particularmente preferente, las moléculas de ácido nucleico de acuerdo con la invención están unidas, por tanto, a una secuencia que dirige a la vacuola con el fin de provocar el transporte de la proteína formada al interior de la vacuola vegetal, donde se puede producir la síntesis de alternano con un rendimiento particularmente alto debido a la concentración de sacarosa, que es alta localmente. Para garantizar la localización en la vacuola, es factible el uso de una de las siguientes proteínas de tránsito: la secuencia N-terminal (146 aminoácidos) de la proteína patatina (Sonnewald et al., *Plant J.* 1 (1991), 95-106) o las secuencias señal que se describen en Matsuoka y Neuhaus, *Journal of Experimental Botany* 50 (1999), 165-174; Chrispeels y Raikhel, *Cell* 68 (1992), 613-616; Matsuoka y Nakamura, *Proc. Natl. Acad. Sci. USA* 88 (1991), 834-838; Bednarek y Raikhel, *Plant Cell* 3 (1991), 1195-1206; Nakamura y Matsuoka, *Plant Phys.* 101 (1993), 1-5.

15 Para garantizar la localización en los plastos, en la mitocondria o en los apoplastos, también es posible el uso de uno de los péptidos de tránsito que se describen en el documento WO 00/47727 en la página 25/26 y en la solicitud de patente de EE. UU. 5.510.471 (péptido señal OTP para dirigir a plastos).

20 En principio, las plantas transgénicas pueden ser plantas de cualquier especie vegetal, es decir, plantas tanto monocotiledóneas como dicotiledóneas. Preferentemente, son plantas útiles, es decir, plantas que cultivadas por el ser humano como alimento o con fines técnicos, en particular industriales. Preferentemente, son plantas que almacenan almidón, por ejemplo, especies de cereales tales como centeno, cebada, avena, trigo, mijo/sorgo, sagú y similares, arroz, guisante, guisante de grano arrugado, mandioca, patata, tomate, colza, soja, cáñamo, lino, tabaco, girasol, judía de vaca o arrurruz, plantas formadoras de fibras tales como, por ejemplo, lino, cáñamo, algodón, plantas que almacenan aceite tales como, por ejemplo, colza, girasol, soja, y plantas que almacenan proteínas tales como, por ejemplo, plantas leguminosas, cereales, soja. La invención también se refiere a árboles frutales y palmeras. La invención se refiere además a plantas para piensos tales como, por ejemplo, hierbas para piensos y hierbas para forraje (tales como, por ejemplo, alfalfa, trébol, ballico), y plantas hortalizas, tales como tomate, lechuga, achicoria, y ornamentales tales como, por ejemplo, tulipanes y jacintos.

25 Del modo más preferente, las plantas de acuerdo con la invención son caña de azúcar, remolacha azucarera, sorgo azucarero, plantas de patata, maíz, arroz, trigo o tabaco.

30 La invención también se refiere a un procedimiento de generación de células vegetales modificadas genéticamente y plantas modificadas genéticamente que, en comparación con las células naturales no modificadas genéticamente/plantas naturales no modificadas genéticamente, sintetizan alternano.

El procedimiento de acuerdo con la invención comprende las siguientes etapas:

- 35 a) introducir en una célula vegetal una molécula de ácido nucleico que codifica una alternano sacarasa troncada o un vector que comprende una molécula de ácido nucleico de este tipo,
- b) regenerar una planta a partir de células vegetales obtenidas en la etapa a), y
- c) en caso apropiado, generar más plantas con ayuda de las plantas obtenidas en la etapa b).

Las moléculas de ácido nucleico que codifican una alternano sacarasa troncada y los vectores que comprenden una molécula de ácido nucleico de este tipo se han descrito anteriormente.

40 En cuanto a la introducción en una célula vegetal de moléculas de ácido nucleico exógenas de acuerdo con la etapa a) del procedimiento de generación de una planta modificada genéticamente, en principio, puede adoptar la forma de cualquier tipo de introducción de moléculas de ácido nucleico que sea adecuada para integrar una molécula de ácido nucleico exógena en una célula vegetal o planta. Ya se han descrito procedimientos de este tipo anteriormente y se pueden aplicar aquí de forma análoga.

45 La regeneración de plantas, en función de los procedimientos de la etapa b) y/o c) de los procedimientos de acuerdo con la invención, se puede llevar a cabo por procedimientos que son conocidos para el trabajador experto (por ejemplo, los descritos en "Plant Cell Culture Protocols", 1999, ed. por R.D. Hall, Humana Press, ISBN 0-89603-549-2).

50 La generación de más plantas en función del procedimiento de acuerdo con la etapa c) o d) de los procedimientos de acuerdo con la invención se puede llevar a cabo, por ejemplo, por propagación vegetativa (por ejemplo, por esquejes, tubérculos o cultivo de callos y regeneración de plantas intactas) o por propagación generativa. En este contexto, la propagación generativa se produce, preferentemente, en condiciones controladas, es decir, se cruzan plantas seleccionadas con propiedades específicas entre sí y se propagan. Preferentemente, la selección se lleva a cabo de tal forma que las plantas, en función del procedimiento de acuerdo con la etapa b) o d), comprenden las modificaciones introducidas en la etapa a).

La invención también se refiere a material de propagación de las plantas de acuerdo con la invención que comprende células vegetales de acuerdo con la invención. En este contexto, el término "material de propagación" comprende los componentes de la planta que son adecuados para generar una progenie por medio de una vía vegetativa o generativa. El material adecuado para la propagación vegetativa es, por ejemplo, esquejes, cultivos de callos, rizomas o tubérculos. Otro material de propagación comprende, por ejemplo, frutos, semillas, plántulas, cultivos celulares y similares. Preferentemente, el material de propagación es material de propagación de caña de azúcar, remolacha azucarera, sorgo azucarero, plantas de patata, maíz, arroz, trigo o tabaco.

En otro aspecto, la presente invención se refiere a un procedimiento de producción de una proteína de acuerdo con la invención con la actividad enzimática de una alternano sacarasa, en el que se expresa la proteína en una planta de acuerdo con la invención y, en una etapa posterior, se extrae y se aísla. Con este fin, se puede hacer crecer la planta de acuerdo con la invención y cultivarla en condiciones de crecimiento adecuadas hasta que, como consecuencia de la biosíntesis de proteínas, se forme la proteína con la actividad enzimática de una alternano sacarasa en sus células, que comprenden una molécula de ácido nucleico de acuerdo con la invención. Después de esto, se puede obtener la proteína de acuerdo con la invención por medio de procedimientos habituales, a partir de la planta o a partir de partes de la planta o a partir de material de propagación de la planta en el que esté presente la proteína con la actividad enzimática de una alternano sacarasa. Las etapas de procedimiento convencionales comprenden, por ejemplo, triturar el material vegetal y extraer y purificar la proteína.

Preferentemente, el procedimiento de producción de una proteína con la actividad enzimática de una alternano sacarasa, comprende las etapas de

1) se transforma una célula vegetal con

a) una molécula de ácido nucleico cuya secuencia codifica la secuencia de aminoácidos de una proteína descrita anteriormente con la actividad enzimática de una alternano sacarasa, pudiendo obtenerse la molécula de ácido nucleico a partir de una molécula de ácido nucleico con una secuencia mostrada en la SEQ ID NO 1 o 3, o

b) un vector o plásmido que comprende una molécula de ácido nucleico de acuerdo con a),

2) se regenera una planta a partir de las células vegetales obtenidas en la etapa 1),

3) se expresa la proteína en la planta y posteriormente se extrae y se aísla.

Las secuencias de aminoácidos de las proteínas de acuerdo con la invención con la actividad enzimática de una alternano sacarasa se indican más arriba y se definen con referencia a la SEQ ID NO 2. Partiendo de las secuencias de aminoácidos, un experto en la técnica puede determinar secuencias codificantes de ácidos nucleicos adecuadas que, debido a la degeneración del código genético, pueden diferir. La determinación de una secuencia de ácidos nucleicos codificante y la generación de dicha secuencia de ácidos nucleicos por medio de PCR ya se han descrito anteriormente en el procedimiento de producción de la proteína de acuerdo con la invención a partir de células huésped transformadas que se hacen crecer en medio de cultivo. En cuanto a los vectores y plásmidos que se pueden emplear en el procedimiento anterior, en otro punto se hace referencia a la divulgación en cuestión. Además, en otro punto de la presente descripción se divulgan ampliamente células huésped y técnicas para la transformación de células huésped vegetales que se pueden emplear en el procedimiento anterior.

La invención también se refiere a partes de plantas de las plantas de acuerdo con la invención descritas anteriormente que se pueden cosechar, tales como, por ejemplo, de frutos, semillas, tubérculos, hojas o rizomas de tallos, donde las partes de la planta cosechables comprenden células vegetales de acuerdo con la invención.

En el presente documento se divulga un alternano con un promedio en peso del peso molecular Pm en el intervalo de desde 12 000 000 hasta 30 000 000 g/mol.

El término "alternano" se refiere a un polisacárido compuesto por unidades de glucosa. Las unidades de glucosa están unidas entre sí por enlaces  $\alpha$ -1,3 y  $\alpha$ -1,6-glucosídicos, donde estos dos tipos de enlace se alternan de forma predominante. Además, el alternano puede comprender ramificaciones.

Preferentemente, el alternano descrito en el presente documento tiene un promedio en peso del peso molecular Pm en el intervalo de desde 14 000 000 hasta 28 000 000 g/mol, aún más preferentemente de 16 000 000 a 26 000 000 g/mol, más preferentemente de 18 000 000 a 23 000 000 g/mol y lo más preferentemente de 19 000 000 a 23 000 000 g/mol. El peso molecular se determina por medio del procedimiento GPC-MALLS, que se describe con mayor detalle en los ejemplos de uso.

Por tanto, el alternano descrito en el presente documento tiene un peso molecular promedio notablemente mayor que el descrito para el alternano nativo (Pm  $10^6$ - $10^7$ , p. ej., en el documento WO 03/010177) y también un peso molecular notablemente mayor que el descrito en los documentos WO 2004/023891 y US 2006/0127328.

La presente invención se refiere además a diversos procedimientos para la producción del alternano de acuerdo con la invención.

Se divulga un procedimiento para la producción del alternano de acuerdo con la invención en el que

1) se pone en contacto una solución que comprende sacarosa con una proteína con la actividad enzimática de una alternano sacarasa, y

2) se aísla alternano a partir de la solución,

5 donde, en el procedimiento, la proteína

a) tiene una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y lo más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y termina en la región desde el aminoácido de la posición 1295 al aminoácido de la posición 1520 de la SEQ ID NO 2, o

b) tiene una secuencia de aminoácidos (secuencia modificada) con al menos el 70 %, preferentemente al menos el 80 %, en particular al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de a), donde, preferentemente, la secuencia de aminoácidos modificada tiene el mismo número de aminoácidos que la secuencia de a).

Por tanto, la invención se refiere a un procedimiento para la producción de alternano *in vitro* con ayuda de una preparación de una enzima sin células. En comparación con la síntesis de alternano en un sistema que sí tiene células, este procedimiento tiene la ventaja de que se pueden controlar mejor las condiciones de reacción y de que los productos de la reacción son considerablemente más puros y se pueden purificar con más facilidad.

La proteína con la actividad enzimática de una alternano sacarasa se puede obtener, por ejemplo, transformando una célula huésped con un vector de acuerdo con la invención como se describe anteriormente y propagando las células huésped en condiciones adecuadas. Después, se pueden obtener la proteína expresada a partir de las células huésped. Se puede formar un extracto de proteína destruyendo las células huésped, por ejemplo, usando una prensa francesa, y se emplea este extracto de proteína en el procedimiento de acuerdo con la invención. En principio, es posible usar todas las técnicas conocidas para aislar proteínas a partir de células huésped. De forma muy particularmente preferente, la célula huésped es *E. coli*.

En otra variante para obtener la proteína con la actividad enzimática de una alternano sacarasa, se cultivan microorganismos que secretan una proteína de este tipo en un medio sin sacarosa, lo que permite la formación de la proteína, hasta alcanzar la fase estacionaria. Una vez se han separado las células del medio de cultivo por centrifugación, se puede obtener la enzima secretada a partir del sobrenadante. Posteriormente se pueden añadir la enzima a soluciones que comprenden sacarosa con el fin de sintetizar alternano.

Una proteína que se usa preferentemente en el procedimiento descrito anteriormente es una proteína con la actividad enzimática de una alternano sacarasa que

a) tiene una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1300 hasta el aminoácido de la posición 1420 de la SEQ ID NO 2, o que

b) tiene una secuencia de aminoácidos (secuencia modificada) que tiene al menos el 70 %, más preferentemente al menos el 80 %, en particular al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de a), donde, preferentemente, la secuencia modificada tiene el mismo número de aminoácidos que la secuencia de a).

Una proteína que se usa aún más preferentemente en el procedimiento descrito anteriormente es una proteína con la actividad enzimática de una alternano sacarasa que

a) tiene una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1300 hasta el aminoácido de la posición 1340 de la SEQ ID NO 2, o que

b) tiene una secuencia de aminoácidos (secuencia modificada) que tiene al menos el 70 %, más preferentemente al menos el 80 %, en particular al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de a), donde,

preferentemente, la secuencia modificada tiene el mismo número de aminoácidos que la secuencia de a).

En otro modo de realización ventajoso, el procedimiento descrito anteriormente emplea una proteína con la actividad enzimática de una alternano sacarasa que

5 a) tiene una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 350, preferentemente en la región desde el aminoácido 1 hasta el 300, más preferentemente en la región desde el aminoácido 1 hasta el 200, aún más preferentemente en la región desde el aminoácido 1 hasta el 100 y más preferentemente en la región desde el aminoácido 1 hasta el 40 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1310 hasta el aminoácido de la posición 1320 de la SEQ ID NO 2, o que

10 b) tiene una secuencia de aminoácidos (secuencia modificada) que tiene al menos el 70 %, más preferentemente al menos el 80 %, en particular al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con una secuencia de aminoácidos de a), donde, preferentemente, la secuencia modificada tiene el mismo número de aminoácidos que la secuencia de a).

15 Lo más preferentemente, la proteína con la actividad enzimática de una alternano sacarasa, proteína que se emplea en el procedimiento de acuerdo con la invención, tiene una de las siguientes secuencias:

a) una secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1380, inclusive, de la SEQ ID NO 2

b) una secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2

20 c) una secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1506, inclusive, de la SEQ ID NO 2

25 d) una secuencia (secuencia modificada) que tiene al menos el 70 %, más preferentemente al menos el 80 %, en particular al menos el 90 %, aún más preferentemente al menos el 95 % y lo más preferentemente al menos el 98 % de identidad con la secuencia de a), b) o c), donde, preferentemente, la secuencia modificada tiene el mismo número de aminoácidos que la secuencia de a).

30 Además, el procedimiento descrito anteriormente también puede emplear todas las proteínas descritas anteriormente con la actividad enzimática de una alternano sacarasa que están truncadas en comparación con una alternano sacarasa de longitud completa definida anteriormente del extremo N-y/o C-terminal de la secuencia de aminoácidos ("alternano sacarasa truncada") y que están codificadas por moléculas de ácido nucleico de acuerdo con la invención con la secuencia correspondiente.

35 En una divulgación, hay una metionina situada cadena arriba del extremo N de las proteínas descritas anteriormente y a continuación en el presente documento, con la actividad enzimática de una alternano sacarasa, proteínas que se emplean en el procedimiento de acuerdo con la invención. En este modo de realización preferente, por tanto, la metionina forma el nuevo extremo N de la proteína empleada en el procedimiento de acuerdo con la invención. Esta metionina se origina habitualmente en la traducción del ARN que codifica la proteína y está codificada en la traducción por el triplete de ácido nucleico ATG.

40 Preferentemente, se excluyen del procedimiento de acuerdo con la invención las proteínas truncadas con la actividad enzimática de una alternano sacarasa que se describen en el documento US20060127328A1, en particular las proteínas descritas en el párrafo [0089], y en la figura 3A y 3B del documento US20060127328A 1. Por tanto, preferentemente, se excluyen de la presente invención las proteínas desde el aminoácido de la posición 1 a la 1425, de la posición 1 a la 1349, de la posición 1 a la 290, de la posición 342 a la 1425 y de la posición 342 a la 1290, estando definidas las posiciones con referencia a la SEQ ID NO 2 adjunta.

45 Un modo de realización preferente del procedimiento de acuerdo con la invención emplea una proteína purificada con la actividad enzimática de una alternano sacarasa. Se entiende que una proteína purificada significa una proteína que, en gran medida, no tiene componentes celulares de las células en las que se sintetiza la proteína y/o que carece de contaminación con proteínas que tienen actividades sintetizadoras de polisacáridos (por ejemplo, dextrano sacarosas) o actividades degradadoras de polisacáridos y/o que comprenden aceptores (de polisacáridos). Preferentemente, el término "proteína purificada con la actividad enzimática de una alternano sacarasa" indica una proteína con un grado de pureza de al menos el 70 %, preferentemente de al menos el 85 % y de forma particularmente preferente de al menos el 95 %.

50 El uso de una proteína purificada para la producción de alternano tiene una variedad de ventajas. En comparación con procedimientos que emplean extractos de proteínas parcialmente purificados, el medio de reacción del procedimiento de acuerdo con la invención no comprende restos de la cepa de producción (microorganismo) que se usa para producir genéticamente la proteína.

- Además, el uso de la proteína purificada conlleva ventajas para el uso en las industrias alimentaria y farmacéutica. Como consecuencia de la composición definida del medio de reacción, composición que se ha liberado de todos los componentes innecesarios, el producto también se define de forma más precisa en cuanto a sus componentes. Esto da lugar a un proceso de obtención de licencias considerablemente menos complicado para estos productos biotecnológicos en la industrias alimentaria y farmacéutica, en particular porque estos productos no deberían comprender trazas de un microorganismo transgénico.
- Además, con ayuda de este procedimiento de acuerdo con la invención, es posible producir alternano con rendimientos altos, sin pérdidas como consecuencia, por ejemplo, de la interferencia de la actividad secundaria de las enzimas de otra célula huésped que puede formar productos secundarios que interfieren.
- En otro modo de realización preferente del procedimiento de acuerdo con la invención, se usa una proteína producida de forma recombinante con la actividad enzimática de una alternano sacarasa.
- En otro modo de realización preferente más, la proteína con la actividad enzimática de una alternano sacarasa se inmoviliza sobre un soporte.
- La inmovilización de la proteína de acuerdo con la invención con la actividad enzimática de una alternano sacarasa tiene la ventaja de que se puede recuperar la enzima, que es un catalizador de la reacción de síntesis, de forma sencilla a partir de la mezcla de reacción y volverse a emplear. Dado que la purificación de las enzimas, por norma general, es cara y consume tiempo, una inmovilización y recuperación de la enzima permite realizar ahorros considerables. Otra ventaja es el grado de pureza de los productos de la reacción, que no comprenden proteína residual.
- Existen múltiples soportes disponibles para la inmovilización de proteínas, siendo posible su acoplamiento al soporte por medio de enlaces covalentes o no covalentes (para una descripción general, véase: Methods in Enzymology 135, 136, 137). Son soportes de uso extendido, por ejemplo, la agarosa, el alginato, la celulosa, la poli(acrilamida), la sílice o el nailon.
- El soporte puede ser, por ejemplo, un soporte inerte, un soporte cargado eléctricamente, un soporte inorgánico o un soporte orgánico. Son soportes adecuados que se pueden usar los materiales inorgánicos tales como vidrios porosos, gel de sílice, alúmina, hidroxiapatita o diversos óxidos metálicos, los polímeros naturales, por ejemplo, celulosa, almidón, alginatos, agarosa o colágeno, o los polímeros sintéticos tales como poli(acrilamida), poli(alcohol vinílico), acrilato de metilo, nailon u oxiranos. La inmovilización se lleva a cabo por fuerzas de unión físicas tales como, por ejemplo, las fuerzas de Van der Waals, las interacciones hidrófobas y los enlaces iónicos. Como se menciona anteriormente, la proteína con la actividad enzimática de una alternano sacarasa también se puede inmovilizar sobre soportes por medio de un enlace covalente. Con este fin, los soportes deben presentar grupos reactivos, que pueden formar enlaces homopolares con cadenas laterales de aminoácidos. Son ejemplos de grupos adecuados los grupos carboxilo, los grupos hidroxilo y los grupos sulfuro. Por ejemplo, se puede activar la superficie de los vidrios porosos por tratamiento con silanos y, posteriormente, hacerla reaccionar con proteínas. Se pueden activar los grupos hidroxilo de los polímeros naturales con bromuro de cianógeno y los grupos carboxilo con cloruro de tionilo y, posteriormente, acoplarlos con enzimas. Otra posibilidad de inmovilización de la proteína con la actividad enzimática de una alternano sacarasa es su inclusión en redes tridimensionales. Una ventaja es que las enzimas están presentes en forma libre, sin unir, en la red. Los poros de la matriz circundante deben ser lo suficientemente pequeños para retener la enzima.
- En un modo de realización preferente del procedimiento de acuerdo con la invención descrito anteriormente, la solución que comprende sacarosa es una solución con una concentración de sacarosa de al menos el 12 % en peso, aún más preferentemente de al menos el 15 % en peso, en base al peso de agua. De forma particularmente preferente, la concentración de sacarosa es del 15 % en peso, del 40 % en peso en base al peso de agua, más preferentemente, del 15 % en peso al 30 % en peso y, lo más preferentemente, del 18 % en peso al 25 % en peso.
- La etapa 1) del procedimiento, en la que se pone en contacto una solución que comprende sacarosa con una proteína con la actividad enzimática de una alternano sacarasa, se lleva a cabo preferentemente a una temperatura de 35-45 °C, aún más preferentemente a 35-40 °C.
- Preferentemente, el pH en el procedimiento de acuerdo con la invención es de 5-6, aún más preferentemente 5,0-5,3.
- Preferentemente, el tiempo de reacción en la etapa 1 del procedimiento supone al menos 12 horas, más preferentemente al menos 24 horas, aún más preferentemente al menos 36 horas, de forma particularmente preferente 40-60 horas.
- En otra variante ventajosa, el procedimiento se lleva a cabo sin agitación.
- En una variante del procedimiento descrito anteriormente, se añade una molécula aceptora en la etapa 1) del procedimiento. En el contexto de la presente invención, se entiendo que una molécula aceptora significa una molécula en la que una alternano sacarasa troncada puede catalizar una reacción de elongación de la cadena,

formándose una cadena alternano en la molécula aceptora. Preferentemente, el aceptor que se puede añadir a la mezcla de reacción al comienzo de la reacción es un hidrato de carbono o un derivado de hidrato de carbono. Preferentemente, el aceptor de hidrato de carbono es un oligo o polisacárido, en particular un polisacárido ramificado tal como, por ejemplo, dextrina, glucógeno o amilopectina, preferentemente un polisacárido lineal y de forma particularmente preferente un sacárido seleccionado del grupo que consiste en maltosa, isomaltosa, isomaltotriosa y metil-alfa-D-glucano o gentobiosa. Si la elongación de una cadena de alternano se produce en estos aceptores, esto da lugar a productos con un peso molecular mayor en comparación con el material de partida. Cuando se usan maltosa, isomaltosa, isomaltotriosa y metil-alfa-D-glucano, se obtienen productos que tienen un peso molecular reducido en comparación con el alternano que se puede preparar en ausencia de aceptores de hidrato de carbono externos, productos que se denominan oligoaltermanos. El tamaño del peso molecular de los oligoaltermanos preparados depende de la proporción de sacarosa:aceptor empleada. Por tanto, el grado de polimerización de los productos aumenta con un aumento de la proporción de sacarosa:aceptor. Además, la proporción de sacarosa:aceptor afecta al rendimiento de oligoalternano. Por tanto, en el caso del aceptor de isomaltosa, el rendimiento de oligoalternano aumenta al disminuir la proporción de sacarosa:isomaltosa.

En otro procedimiento, se obtiene alternano extrayendo alternano de una célula huésped de acuerdo con la invención o una planta de acuerdo con la invención o su material de propagación o de las partes de plantas cosechadas descritas anteriormente, y, si es apropiado, aislándolo.

Un modo de realización ventajoso proporciona que el alternano se aísla de las plantas de acuerdo con la invención, en particular de sus vacuolas, y se purifica. Es particularmente preferente aislar, y obtener, alternanos a partir de los órganos de almacenamiento de las plantas de patata, maíz, remolacha azucarera, sorgo azucarero o caña de azúcar. En un modo de realización, el procedimiento comprende la trituración mecánica de la biomasa de células de las plantas más grandes con ayuda de un mezclador de turbina, por ejemplo, una mezcladora Waring, donde, preferentemente, el desmenuzamiento se lleva a cabo en un volumen grande de medio acuoso a bajas temperaturas, por ejemplo, 4 °C. Se puede llevar a cabo una disrupción adicional con ayuda de procedimientos físicos de disrupción, por ejemplo, por medio de ultrasonidos, un dispositivo de prensa francesa, con molinos o prensas, con ayuda de procedimientos químicos de disrupción, por ejemplo, liofilización o usando detergentes o usando cambios en la presión osmótica, o con ayuda de procedimientos biológicos de disrupción, por ejemplo, usando enzimas que atacan la pared celular, o por medio de un tratamiento con ácido/álcali. En un modo de realización particularmente preferente, se obtiene el alternano mediante la producción de un extracto acuoso, en particular de los órganos de almacenamiento, y la posterior precipitación con alcohol, o a bajas temperaturas, por ejemplo, 4 C.

El trabajador experto puede encontrar otras variantes del procedimiento de obtención de alternano por extracción/aislamiento a partir de una célula huésped de acuerdo con la invención, una planta de acuerdo con la invención o su material de propagación o partes de plantas cosechables en la literatura especializada conocida, tal como, por ejemplo, Loncin, Marcel Grundlagen der Verfahrenstechnik in der Lebensmittelindustrie [Fundamentos de ingeniería de procesos en la industria alimentaria]. Fráncfort/Main: Verlag Sauerländer 1969; Tscheuschner, H.D. Grundzüge der Lebensmitteltechnik [Fundamentos de tecnología de los alimentos]. Hamburgo: Behrs Verlag 2004; Kessler, H.G. Lebensmittel und Bioverfahrenstechnik [Tecnología de los alimentos y bioprocesos]. Munich: Verlag A. Kessler 1996; Martin, A.M. Bioconversion of waste materials to industrial products. Londres, Nueva York: Blackie Academic & Professional 1998.

Si la célula vegetal/planta de acuerdo con la invención adopta la forma de una planta de almacenamiento de azúcares tal como, por ejemplo, la remolacha azucarera o la caña de azúcar, o de material de propagación o partes de plantas cosechables de las mismas, el trabajador experto dispone de protocolos de extracción establecidos para azúcares, tales como los descritos, por ejemplo, en Belitz y Grosch. Lehrbuch der Lebensmittelchemie [Manual de química de los alimentos], Berlín Heidelberg, Springer-Verlag 1992, p. 786.

La extracción y el aislamiento del alternano a partir de una planta de acuerdo con la invención o su material de propagación se puede realizar por medio de procedimientos ordinarios tales como, por ejemplo, precipitación, extracción y procedimientos cromatográficos.

En otro procedimiento más de acuerdo con la presente invención, se secreta una proteína de acuerdo con la invención con la actividad enzimática de una alternano sacarasa, como se describe anteriormente (en adelante en el presente documento, alternano sacarasa), producida por una célula huésped de acuerdo con la invención, preferentemente E. coli, en un medio de cultivo que comprende sacarosa, y se aísla alternano a partir del medio de cultivo.

Este procedimiento se describe en el documento WO 00/47727. Las partes pertinentes de la divulgación incluyen los modos de realización preferentes, en particular de la p. 28, 3<sup>er</sup> párrafo hasta la p. 31, 1<sup>er</sup> párrafo, inclusive, del documento WO 00/47727.

El alternano es un polvo seco blanco que es de transparente a opaco en solución. Presenta una solubilidad alta y la solución tiene una viscosidad baja, como se muestra en los ejemplos adjuntos. El alternano de la presente invención no tiene aroma y tiene un sabor limpio. Además, el alternano de la presente invención es inerte frente a los

microbios, no digerible, tiene pocas o ninguna caloría y tiene una respuesta glucémica y un índice glucémico muy bajos.

5 También se describe el uso de alternano con un peso molecular promedio en peso Pm en el intervalo de desde 12.000.000 g/mol hasta 30.000.000 g/mol como adición a alimentos o complementos alimenticios. De acuerdo con la presente invención, el término alimentos incluye tanto alimentos para seres humanos como alimentos para animales o pienso para animales. Además, el término alimentos incluye bebidas. Los complementos alimenticios incluyen complementos alimenticios para seres humanos y para animales.

10 En los alimentos y los complementos alimenticios, se puede usar alternano con un peso molecular promedio en peso Pm en el intervalo de desde 12.000.000 g/mol hasta 30.000.000 g/mol en particular como emulsionante, sustituto de la goma arábica, carga, agente voluminizador, agente descompactante, diluyente, agente prebiótico y/o fibra alimenticia. Son alimentos, sin limitación, bebidas tales como bebidas carbonatadas, zumos, especialmente zumos de fruta, batidos de fruta, gaseosas, té helado, bebidas dietéticas, bebidas saciantes, bebidas terminadas, bebidas para deportistas, bebidas de resistencia, mezclas en polvo de bebidas para complementos alimenticios, productos de panadería, tales como pan, pasteles, galletas, pastas, galletas saladas, cruasanes, fideos y pasta, 15 barras nutritivas, barras energéticas, barras de desayuno, productos de aperitivo y cereales, tales como cereales extrudidos, cereales de desayuno, virutas de cereal, alimentos infantiles y para bebés, sopas, congelados, comida precocinada y sustitutivos de las comidas.

20 Como fibra alimenticia se usa alternano con un peso molecular promedio en peso Pm en el intervalo de desde 12.000.000 g/mol hasta 30.000.000 g/mol. En relación con esto, es particularmente útil para el enriquecimiento en fibra en los alimentos mencionados anteriormente.

Se divulgan además alimentos y complementos alimenticios que comprenden alternano con un peso molecular promedio en peso Pm en el intervalo de desde 12.000.000 g/mol hasta 30.000.000 g/mol como emulsionante, sustituto de la goma arábica, carga, agente voluminizador, agente descompactante, diluyente, agente prebiótico y/o fibra alimenticia.

25 También se divulga el uso de alternano con un peso molecular promedio en peso Pm en el intervalo de desde 12.000.000 g/mol hasta 30.000.000 g/mol como adición a productos cosméticos o farmacéuticos. En estos productos se puede usar alternano con un peso molecular promedio en peso Pm en el intervalo de desde 12.000.000 g/mol hasta 30.000.000 g/mol como emulsionante, carga, agente descompactante, diluyente y/o como excipiente para principios activos.

### 30 Descripción de las secuencias

SEQ ID NO 1: secuencia de ácidos nucleicos de *Leuconostoc mesenteroides*, que codifica una proteína con la actividad enzimática de una alternano sacarasa

35 SEQ ID NO 2: secuencia de aminoácidos de una proteína con la actividad enzimática de una alternano sacarasa de *Leuconostoc mesenteroides*. La secuencia de aminoácidos mostrada puede derivar de la SEQ ID NO 1.

40 SEQ ID NO 3: Secuencia de ácidos nucleicos sintética que codifica una proteína con la actividad enzimática de una alternano sacarasa. La síntesis de los codones de la secuencia mostrada se llevó a cabo de tal forma que está adaptada al uso de codones en células vegetales. La secuencia de ácidos nucleicos mostrada codifica una proteína con la secuencia de aminoácidos mostrada en la SEQ ID NO 2.

SEQ ID NO 4: Cebador sintético GUS-ForwI

SEQ ID NO 5: Cebador sintético GUS-Rev

SEQ ID NO 6: Cebador sintético GUS-ForwII

SEQ ID NO 7: Secuencia de enlazador sintético Linker1

45 SEQ ID NO 8: Secuencia de enlazador sintético Linker2

SEQ ID NO 9: Cebador sintético Nos-Forw

SEQ ID NO 10: Cebador sintético Nos-Rev

SEQ ID NO 11: Oligonucleótido sintético, extensión 5' antes del nucleótido 118 de la SEQ ID NO 3

SEQ ID NO 12: Oligonucleótido sintético, extensión 3' de la SEQ ID NO 3

50 SEQ ID NO 13: Cebador sintético RLP3

- SEQ ID NO 14: Cebador sintético RLP4  
 SEQ ID NO 15: Cebador sintético AISul-Forw  
 SEQ ID NO 16: Cebador sintético AISul-Rev  
 SEQ ID NO 17: Cebador sintético AISull-Forw  
 5 SEQ ID NO 18: Cebador sintético AISull-Rev  
 SEQ ID NO 19: Cebador sintético SIP6  
 SEQ ID NO 20: Cebador sintético SIP7  
 SEQ ID NO 21: Cebador sintético SIP1a  
 SEQ ID NO 22: Cebador sintético SIP2  
 10 SEQ ID NO 23: Cebador sintético SIP3  
 SEQ ID NO 24: Cebador sintético SIP4  
 SEQ ID NO 25: Cebador sintético SIP5

## 1. Procedimientos

### 1.1 Construcción de plásmidos

- 15 Todos los procedimientos de biología molecular se llevaron a cabo por medio de procedimientos ordinarios (Sambrook y Russell Molecular cloning: a laboratory manual (tercera edición) Cold Spring Harbor Laboratory Press, NY, EE. UU. (2001)). Las enzimas de restricción se obtuvieron de New England Biolabs, Inc., Ipswich, MA, EE. UU.

#### pUbiGusNos

- 20 Se escindió el promotor de la ubiquitina de Zea mays (n.º de acceso de GenBank 94464, Christensen et al., Plant Mol. Biol. 18, 1992, 675)) a partir de un pML8 (en el documento WO 2006 066969) con PstI y se ligó en un pUC18 escindido con PstI (n.º de acceso de GenBank L09136) en una orientación tal que el extremo 5' del promotor estaba en el lado del sitio HindIII, lo que dio lugar a un pUC18-Ubi. Se amplificó la secuencia codificante de la beta-glucuronidasa (GUS) de E. coli (número de acceso de GenBank S69414) con los cebadores GUS-ForwI (SEQ ID NO 4; GAAAGCTTGGCTGCAGGTCAGTCCCTTATGTTACGTCCTGTAG) y GUS-Rev (SEQ ID NO 5; CTGGTCCC GGGATTTCATTGTTTGCCTCCCTGCT), condiciones de PCR: 5 ciclos en cada caso de 30 s a 94 °C, 30 s a 60 °C, 2 min a 72 °C, y 25 ciclos en cada caso con 30 s a 94 °C, 30 s a 73 °C, 2 min a 72 °C, usando la polimerasa de ADN Taq Platinum de alta fidelidad (Invitrogen Corp, Carlsbad, CA, EE. UU.) y realizando el procedimiento en las condiciones descritas por el fabricante. El fragmento resultante se amplificó de nuevo con los cebadores GUS-ForwII (SEQ ID NO 6; TCGAGCTCGCGAAAGCTTGGCTGCAGGT) y GUS-Rev en las mismas condiciones y se escindió con XmaI. Se escindió el pUC18-Ubi con BamHI, se rellenaron los salientes con la polimerasa de ADN T4 (New England Biolabs, Inc., Ipswich, MA, EE. UU.) y se escindió el fragmento con XmaI. El producto de PCR escindido con XmaI GUS se ligó en este vector y el plásmido resultante se llamó pUC18-Ubi-GUS. Se escindió el pUC18-Ubi-GUS con XmaI y se retiraron los salientes con nucleasa de judía mungo (New England Biolabs, Inc., Ipswich, MA, EE. UU.) y se escindió el fragmento con EcoRI. Se alinearon los oligonucleótidos fosforilados Linker1 (SEQ ID NO 7; AACAACTCTCCTGGCGCACCATCGTCG-  
 25 GCTACAGCCTCGGGAATTGCTGCAGGTC GACGGATCCG) y Linker2 (SEQ ID NO 8; AATTCGGATCCGTCGACCTGCAGCAATTCCTCCGAGGCTGTAGCCGACGATGGTGCG CCAGGAGAGTTGTT) y se ligó el enlazador resultante en este fragmento de vector escindido con XmaI/EcoRI, dando lugar al pUC18-Ubi-GUS-PL. Se amplificó el terminador nos de Agrobacterium tumefaciens (Depicker et al., 1982, Journal of Molecular and Applied Genetics 1, 561-573) a partir de un pML8 usando los cebadores Nos-Forw (SEQ ID NO 9; CGGATCGGATCCGTCGACCTGCAGATCGTTCAAACATTTGGCAAT) y Nos-Rev (SEQ ID NO 10; CCACGGAATTCGATCTAGTAACATAGATGACACCGCG). Condiciones de PCR: 5 ciclos en cada caso con 30 s a 94 °C, 30 s a 58 °C, 3 min a 72 °C, y 25 ciclos en cada caso con 30 s a 94 °C, 30 s a 70 °C, 3 min a 72 °C, usando la polimerasa de ADN Taq Platinum de alta fidelidad (Invitrogen Corp, Carlsbad, CA, EE. UU.) y realizando el  
 40 procedimiento en las condiciones descritas por el fabricante. Se escindió el producto de PCR con BamHI y EcoRI y se ligó en el pUC18-Ubi-GUS-PL escindido con BamHI/EcoRI, lo que dio lugar al vector pUbiGusNos.

#### pSM13

- 50 La secuencia codificante de alternano sacarasa con codones optimizados se sintetizó en Entelechon (Regensburg, Alemania). Se retiraron los sitios de ajuste predichos y las señales de poli(A). Además, en la secuencia con codones optimizados no existen tramos de GC/AT de más de 5 o 6 nucleótidos de longitud, respectivamente, ni secuencias CAAT. Se retiraron o añadieron determinados sitios de escisión de restricción. La secuencia sintética comprende los

nucleótidos 118-6174 de la SEQ ID NO 3, con la extensión en 5' GCGGCCGCGGCAGCCATG [SEQ ID NO 11], de modo que se introduce un codón de iniciación de las traducciones ATG [subrayado] cadena arriba del nucleótido 118 de la SEQ ID NO 3, y con la extensión en 3' ACCGGATATCGCGGCCGC [SEQ ID NO 12]). Se clonó la secuencia sintética en un pUC19 (New England Bio labs; número de acceso de GenBank M77789) usando EcoRI/HindIII y usando el sitio de escisión de restricción de HindIII en la secuencia codificante de alternano sacarasa sintética, lo que dio lugar a un pRVL110. Se introdujo de nuevo el codón de detención (subrayado en RLP3) alineando los cebadores fosforilados RLP3 (SEQ ID NO 13; AGCTTGAACCGGATATCGCGGCCGC) y RLP4 (SEQ ID NO 14; AGCTGCGGCCGCGATATCCGGTTCA) y ligando el enlazador resultante en un pRVL110 escindido con HindIII, con la orientación del enlazador de tal modo que el RLP3 estaba en orientación sentido cadena abajo de la secuencia codificante de la alternano sacarasa. El plásmido resultante se llamó pSM12.

Se amplificó el segmento 5' de la alternano sacarasa a partir de pSM12 con los cebadores AISul-Forw (SEQ ID NO 15; CAACTGAGCTCTGCAGAATTCGGCTTGTTCGGCAGCCATGGACACCAACTCC) y AISul-Rev (SEQ ID NO 16; GCCGGAGTTGTCGGTGCCGAAGCCCTTGCG). Condiciones de PCR: 5 ciclos en cada caso con 30 s a 94 °C, 30 s a 65°C, 3 min a 72 °C, y 25 ciclos en cada caso con 30 s a 94 °C, 30 s a 72°C, 5 min a 72 °C, usando la polimerasa de ADN Taq Platinum de alta fidelidad (Invitrogen Corp, Carlsbad, CA, EE. UU.) en las condiciones especificadas por el fabricante. Se escindió el producto con SacI. Se clonó el fragmento resultante en un pUbiGusNos abierto con SacI en una orientación tal que el extremo 5' de la alternano sacarasa estaba cadena abajo del promotor de la ubiquitina, lo que dio el pUbiAISul. Se amplificó el extremo 3' de la alternano sacarasa en un pSM12 usando los cebadores AISull-Forw (SEQ ID NO 17; TCCTACGACAAGTCCTCCTTCGAGAACGTG) y AISull-Rev (SEQ ID NO 18; CTTACGGATCCACTAGTAACGGCCGCGATATCCGGTTC). Condiciones de PCR: 5 ciclos en cada caso con 30 s a 94 °C, 30 s a 65°C, 3 min a 72 °C, y 25 ciclos en cada caso con 30 s a 94 °C, 30 s a 72°C, 5 min a 72 °C, usando la polimerasa de ADN Taq Platinum de alta fidelidad (Invitrogen Corp, Carlsbad, CA, EE. UU.) en las condiciones especificadas por el fabricante. Se escindió el fragmento resultante con KpnI y BamHI y se ligó en el pUbiAISul digerido con KpnI/BamHI, lo que dio un pSM13.

#### pS13

Se amplificó la alternano sacarasa del pSM13 usando los cebadores SIP6 (SEQ ID NO 19; GCCATCGAGCAGCAGATCTCCCTCAAG) y SIP7 (SEQ ID NO 20; CATAGGATCCCTACACGCCGGAGGCGTC). Al hacerlo, se introdujo un codón de detención TAG (secuencia complementaria subrayada en el SIP7) después del nucleótido 3945 de la SEQ ID NO 3. Condiciones de PCR: 30 ciclos en cada caso con 10 s a 95 °C, 2 min a 71 °C, 3 min + 15 s de autoextensión por ciclo a 68 °C, seguido de 5 min a 72 °C, usando la polimerasa de ADN Taq Platinum de alta fidelidad (Invitrogen Corp, Carlsbad, CA, EE. UU.) y realizando el procedimiento como especifica el fabricante. Se escindió el producto de PCR con DraIII y BamHI y se clonó en el fragmento DraIII/BamHI del vector pSM13. El plásmido resultante se llamó pS13.

#### pSR16

El pSR16 se generó por escisión del pSM13 con XbaI y Sall y ligando el fragmento que contiene la alternano sacarasa en el fragmento XbaI-Sall del vector pAIsu-pET24a (en el documento WO 00/47727). La proteína codificada por el pSR16 comprende los aminoácidos 40-2057 de la SEQ ID NO 2 más una metionina den el extremo N. Esta proteína se llama alternano sacarasa de longitud completa en los ejemplos que figuran a continuación.

#### pSR17

El pSR17 se clonó por escisión del pS13 con XbaI y Sall y ligando el fragmento que comprende la alternano sacarasa en el fragmento XbaI-Sall del vector pAIsu-pET24a (en el documento WO 00/47727). La proteína codificada por el pSR17 comprende los aminoácidos 40-1315 de la SEQ ID NO 2 más una metionina den el extremo N. Esta proteína se llama alternano sacarasa truncada en los ejemplos que figuran a continuación.

### **1.2 Expresión de alternano sacarasa en extracto de germen de trigo**

#### Transcripción *in vitro*

Se generaron moldes para la transcripción *in vitro* por medio de PCR. Las condiciones de PCR fueron las siguientes: 35 ciclos en cada caso con 10 s a 94 °C, 2 min a 71 °C, 3 min a 68 °C, seguido de 5 min a 72 °C. La polimerasa usada fue la polimerasa de ADN Taq Platinum de alta fidelidad (Invitrogen Corp, Carlsbad, CA, EE. UU.) realizando el procedimiento como especifica el fabricante. La PCR se llevó a cabo sobre el pSM13 usando los siguientes cebadores: SIP1a (SEQ ID NO 21; TAATACGAGTCACTATAGGCAGAAATTCGGCTTGTTCGGGC) y SIP2 (SEQ ID NO 22; AGTAACGGCCGCGATATCCGGTTCAAGC) (el producto codifica los aminoácidos 40-2057 de la SEQ ID NO 2 más una metionina en el extremo N); SIP1a-SIP3 (SEQ ID NO 23; CGAAGAGGCCGTCTAGCGGAGGGAG), donde se incorporó un codón de detención TAG después del nucleótido 4518 de la SEQ ID NO 3 (secuencia complementaria subrayada en el SIP3) (el producto codifica los aminoácidos 40 a 1506 de la SEQ ID NO 2 más una metionina en el extremo N); SIP1a-SIP4 (SEQ ID NO 24; GGTGCCGAAGCCCTAGCGGAGGCTC), donde se incorporó un codón de detención TAG después del nucleótido 4140 de la SEQ ID NO 3 (secuencia complementaria subrayada en el SIP4) (el producto codifica los aminoácidos 40 a 1380 de la SEQ ID NO 2 más una metionina en el extremo N); SIP1a-SIP5 (SEQ ID NO 25; CCGGAGATGGAGTAGTACTACACGCCGGA), donde se incorporó un

codón de detención TAG después del nucleótido 3945 de la SEQ ID NO 3 (secuencia complementaria subrayada en el SIP5) (el producto codifica los aminoácidos 40 a 1315 de la SEQ ID NO 2 más una metionina en extremo N).

La transcripción *in vitro* se llevó a cabo con el kit mMessage mMachine (Ambion, Austin, TX, EE. UU.) usando la polimerasa de ARN T7 y siguiendo los protocolos proporcionados por el fabricante.

#### 5 Traducción *in vitro*

Se tradujeron los ARNm descritos anteriormente en un extracto de germen de trigo de Promega (Promega, Madison, WI, EE. UU.), siguiendo los protocolos proporcionados por el fabricante.

#### Determinación de la cantidad de alternano sacarasa

10 La traducción *in vitro* se llevó a cabo como se describe anteriormente en presencia de metionina marcada con <sup>35</sup>S. Se separaron las proteínas en gel de SDS/poliacrilamida al 8 % y se visualizaron por medio de autorradiografía.

#### **1.3 Expresión de alternano sacarasa en protoplastos de patata**

15 Se generaron protoplastos de patata mediante la incubación de material foliar cortado de plantas de patata *in vitro* durante la noche en celulasa al 0,5 % y macerozima al 0,2 % en manitol 0,6 M en oscuridad a temperatura ambiente. Al día siguiente, se recogieron los protoplastos colándolos con un colador de 125 µm y uno de 63 µm. Se lavaron los protoplastos con manitol 0,6 M, se contaron, se sedimentaron y se suspendieron en medio MaMg [MaMg<sub>2</sub>x6H<sub>2</sub>O, MES al 0,1 %, a pH 5,6] a una concentración de 2x10<sup>6</sup> protoplastos por ml. Se incubaron los protoplastos en hielo durante 30 minutos. Para la transformación, se incubó un volumen típico de 40 µl de ADN con 10<sup>6</sup> protoplastos y 1 ml [PEG 6000 al 40 %, manitol 0,4 M, Ca(NO<sub>3</sub>) x 4H<sub>2</sub>O 0,1 M, a pH 7-9] durante 20 minutos a temperatura ambiente. Se añadieron por etapas 10 ml de medio B5 SMX [vitaminas y sales B5 de Gamborg (Duchefa) complementado con sacarosa al 1 %, manitol 0,4 M, 1,25 g/l de xilosa, 0,1 mg/l de 2,4-D, 0,2 mg/l de BAP, 1,0 mg/l de NAA, a pH 5,7]. Se recogieron los protoplastos por centrifugación, se resuspendieron en 5 ml de medio B5 SMX por 10<sup>6</sup> protoplastos transformados y se incubaron durante la noche en oscuridad a 23 °C. Se recogieron los protoplastos por centrifugación. Para analizar la actividad alternano sacarasa, se resuspendió el sedimento en 100 µl [citrato de sodio 100 mM a pH 6,5, EDTA 1 mM, DTT 5 mM], se congeló rápidamente en nitrógeno líquido para romper las células y se centrifugó. El sobrenadante se usó para el análisis posterior.

25 Para analizar la actividad beta-glucuronidasa (GUS), se homogeneizó el sedimento en tampón de extracción de GUS [tampón NaP 50 mM a pH 7, EDTA 10 mM, SDS al 0,1 %, Triton X-100 al 0,1 %, beta-mercaptoetanol 10 mM] y se centrifugó. El sobrenadante se usó para el análisis posterior.

#### Determinación de la eficacia de transformación de los protoplastos

30 Se determinó la eficacia de la transformación cuantificando la cantidad de GUS sintetizada. La cantidad de GUS se cuantificó incubando 100 µg del extracto de proteína en 4-metilumbeliferil-glucurónido (4-MUG) a una concentración final de 1 mM. Se detuvo la reacción mediante la adición de 900 µl de Na<sub>2</sub>CO<sub>3</sub> 0,2 M a 100 µl de solución de reacción. Se tomaron muestras en diversos puntos en el tiempo y se determinó la cantidad de 4-metilumbeliferilo (MU) producida en un fluorímetro con una absorción a 356 nm y una emisión a 455 nm en diversos puntos en el tiempo como medida de la actividad GUS. La eficacia de transformación relativa se expresó como el aumento relativo de la cantidad de MU formada por 100 µg de extracto de proteína.

#### **1.4 Determinación de la actividad alternano sacarasa**

Se determinó la actividad alternano sacarasa como se describe en el documento WO 00/47727, excepto porque se tiñó el polímero con azul alcian al 0,5 %, solución de HAC al 3 %.

#### 40 **1.5 Producción de alternano sacarasa truncada y de longitud completa recombinante y alternanos en E. coli**

45 Se transformaron el plásmido pSR16 (la proteína codificada por el pSR16 comprende los aminoácidos 40-2057 de la SEQ ID NO 2 más una metionina en el extremo N-alternano sacarasa de longitud completa) y el pSR17 (la proteína codificada por el pSR17 comprende los aminoácidos 40-1315 de la SEQ ID NO 2-alternano sacarasa truncada) en BL21 de E. coli (DE3) (Stratagene), siguiendo procedimientos ordinarios como se describe en Sambrook y Russell, Molecular cloning: a laboratory manual (tercera edición) Cold Spring Harbor Laboratory Press, NY, EE. UU. (2001). Se inocularon 50 ml de medio YT + 50 µg/ml de kanamicina con 1 ml de precultivo de estas células y se hicieron crecer a 37 °C hasta que la DO<sub>600</sub> fue de 0,4. Se añadieron IPTG hasta una concentración final de 5 mM y glucosa hasta una concentración final del 1 %, y se hizo crecer el cultivo durante la noche a temperatura ambiente. Se recogieron las células (15 min a 4000 rpm, a 4 °C), se resuspendieron en 3 ml de tampón de extracción (citrate de sodio 100 mM a pH 5,2, EDTA 1 mM, DTT 10 mM) y se disruptó dos veces usando una prensa francesa. Se centrifugó el extracto en bruto durante 10 min a 13 000 rpm y se esterilizó el sobrenadante (Sterivex GV 0,2 µm, Millipore) y después se usó para el análisis adicional.

Típicamente, se incubaron 5 ml de extracto de proteína de la alternano sacarasa truncada o de longitud completa de matraces de agitación Erlenmeyer en un volumen total de 1 litro (acetato de sodio 100 mM a pH 5,3, sacarosa al

20 %) durante 48 horas a 37 °C sin agitación. Se obtuvo alternano mediante la adición de EtOH hasta una concentración final del 50 %, centrifugación durante 10 minutos a 4000 rpm, lavado de los sedimentos de alternano tres veces con EtOH al 50 % y secado de los sedimentos a 37 °C. El alternano seco se molió en un mortero con mano.

## 5 1.6 Determinación del peso molecular del alternano

Se disolvió alternano en DMSO a una concentración del 0,1 % (p/v). Se incuban las soluciones durante la noche a TA en un agitador (Eppendorf) a 300 rpm y después se calientan durante 2 h a 95 °C en un agitador térmico (Eppendorf) a 300 rpm. Se filtró la muestra a través de una jeringuilla de 13 mm con un filtro de PTFE de 5 µm Whatman Puradisc™. Se analizaron los polímeros usando un sistema Dionex (Dionex Corporation, Sunnyvale, EE. UU.) que consistía en los siguientes componentes: una bomba de HPLC P680, un automuestreador AS50, un compartimento de columna termostatazada TCC-100. Se lleva a cabo la detección usando un detector de dispersión de luz DAWN-EOS (Wyatt Technology, Santa Bárbara, EE. UU.) con una  $\lambda_0 = 690$  nm y 15 detectores en el intervalo de ángulos desde 25,9 hasta 163,3° y una celda de flujo K5 acoplada a un detector Shodex RI-101 RI (Shodex Denko K.K., Kanagawa, Japón). Se fraccionan los polímeros por medio de las siguientes columnas: columna de guarda WATERS Styragel, Styragel 7, Styragel 6E y Styragel 2 (Waters Corp, Milford, MA, EE. UU.). Se inyectan 50 µl de solución. El fraccionamiento se produce a una temperatura de 60 °C y a un caudal de 0,5 ml/min con NaNO<sub>3</sub> 90 mM en DMSO como eluyente. Se analiza a distribución de pesos moleculares de las muestras por medio del programa Astra V 5.1.7.3, usando el programa de cálculo de masas Berry 2nd order (Wyatt Technology, Santa Bárbara, EE. UU.).

## 20 1.7 Determinación de la viscosidad

Se determinó la viscosidad de soluciones de alternano a 23 °C como una función de la concentración, usando el reómetro avanzado Gemini de BOHLIN a una velocidad de cizalladura constante de 39,6 s<sup>-1</sup>. Se transfiere la muestra a la placa/cono del sistema de medida CP4°/40 mm que se había precalentado hasta 23 °C. Los parámetros de medida son los siguientes: modo de viscosimetría con configuración de velocidad de cizalladura: precizalladura: 10 1/s, 60 s, tiempo de configuración 10 s; velocidad de cizalladura constante 39,6 s<sup>-1</sup>; por 10 datos, tiempo de retardo y tiempo de espera: 5 s, tiempo de integración 10 s; tiempo por punto 20 s. Se mide dos veces la misma muestra; en el caso de curvas indebidamente inestables, tres veces con el fin de obtener un mayor número de datos. Los datos que se desvían claramente de la curva lineal no se tienen en cuenta para calcular la media.

## 2. Ejemplos

### 30 Ejemplo 1: Actividad de mutantes de alternano sacarasa por delección C terminal en extracto de germen de trigo

Se generaron mutantes de alternano sacarasa por delección para su expresión en extracto de germen de trigo como se describe en los procedimientos. Se determinaron la actividad y la cantidad del alternano producido en el extracto de germen de trigo.

35 La fig.1 muestra la actividad y la cantidad de la alternano sacarasa en el extracto de germen de trigo. La fig. 1A muestra la actividad alternano sacarasa y que la alternano sacarasa generada usando los cebadores SIP1a y SIP5 (es decir, la proteína comprende los aminoácidos 40 a 1315 de la SEQ ID NO 2 más una metionina en el extremo N) es activa en extracto de germen de trigo.

40 La fig. 1B muestra la cantidad de la alternano sacarasa y que la eficacia de producción en extracto de germen de trigo aumenta con la disminución del tamaño de la proteína. El control es la traducción *in vitro* en ausencia de ARN. Los cebadores usados para generar moldes para la transcripción *in vitro* se muestran en las líneas anteriores. La alternano sacarasa de E. coli es una alternano sacarasa de longitud completa expresada en E. coli. El peso molecular de los marcadores proteínicos se indica a la derecha en la fig. 1B.

### Ejemplo 2: Actividad de mutantes de alternano sacarasa por delección C terminal en protoplastos de patata

45 Para expresar alternano sacarasa de longitud completa en protoplastos de patata, se transformaron 30 µg de pSM13 junto con 10 µg de pUbiGusNos por 10<sup>6</sup> protoplastos. Para expresar la alternano sacarasa truncada en el extremo C-terminal, se transformaron 30 µg de pS13 junto con 10 µg de pUbiGusNos por 10<sup>6</sup> protoplastos. Como control, se transformaron 30 µg de pBluescriptII KS(+) (Stratagene, La Jolla, CA, EE. UU.; número de acceso de GenBank X52327) junto con 10 µg de pUbiGusNos o 40 µg de pBluescriptII KS(+) por 10<sup>6</sup> protoplastos. Se determinó la actividad de las proteínas como se describe en los procedimientos. La cantidad de extracto de proteína analizada se corrigió mediante la eficacia de transformación relativa. Las fig. 2A y 2B muestran los resultados de los dos experimentos diferentes. Se aplicó la alternano sacarasa codificada por el pS13 al gel a diferentes concentraciones y se comparó con la alternano sacarasa codificada por el pSM13. El pKS-GUS es el control negativo en el que no se habían transformado las secuencias codificantes de alternano sacarasa.

55 El experimento mostrado en las fig. 2A/B demuestra que la actividad de la proteína codificada por el pS13 (la proteína comprende los aminoácidos 40 a 1315 de la SEQ ID NO 2 más una metionina en el extremo N) en

protoplastos de patata es de 5 a 40 veces mayor que la actividad de la alternano sacarasa codificada por el pSM13 (la proteína comprende los aminoácidos 40 a 2057 de la SEQ ID NO 2 más una metionina en el extremo N).

### Ejemplo 3: Determinación del peso molecular de alternanos

5 Se produjeron alternanos con alternano sacarasa truncada y con alternano sacarasa de longitud completa (ejemplo comparativo) como se describe en 1.5. El alternano producido con alternano sacarasa truncada también se abrevia en adelante en el presente documento como "SR17" y el alternano producido con alternano sacarasa de longitud completa también se abrevia en adelante en el presente documento como "SR16".

10 Se determinaron los pesos moleculares de los alternanos por GPC MALLS (véanse los procedimientos) y se muestran en la tabla siguiente. El promedio en peso del peso molecular Pm del alternano producido con la alternano sacarasa truncada fue de 21 200 000 g/mol (SR17) y el Pm del alternano producido con la alternano sacarasa de longitud completa (SR16) fue de 35 300 000 g/mol (ambos con una concentración inicial del 1 %).

Muestra	Concentración inicial	Pm (g/mol) *10 <sup>7</sup>	Recuperación (%)
<b>Alternano producido a partir de alternano sacarasa de longitud completa</b>	1 %	3,53	79,0 %
SR16-15	15 %	3,56	59,5 %
SR16-13	13 %	3,64	45,0 %
SR16-12	12 %	3,69	71,1 %
SR16-11	11 %	3,69	66,1 %
SR16-10	10 %	3,74	83,2 %
SR16-9	9 %	3,74	80,9 %
SR16-8	8 %	3,77	86,8 %
SR16-7	7 %	3,77	87,1 %
<b>Alternano producido a partir de alternano sacarasa truncada (ej. comparativo)</b>	1 %	2,12	86,4 %
SR17-15	15 %	2,18	66,5 %
SR17-13	13 %	2,18	67,8 %
SR17-12	12 %	2,22	90,8 %
SR17-11	11 %	2,21	95,2 %
SR17-10	10 %	2,20	88,0 %
SR17-9	9 %	2,19	87,4 %
SR17-8	8 %	2,21	89,3 %
SR17-7	7 %	2,23	91,2 %

15 El alternano producido a partir de alternano sacarasa truncada se midió usando espectrometría de RMN de <sup>13</sup>C. Con este fin, se disolvió el material de muestra en DMSO deuterado a 80 °C con agitación intensa y se midió a una temperatura igualmente de 80 °C usando un espectrómetro Unity INOVA 500 (Varian Inc., EE. UU.) a una frecuencia de medida de 125,69 MHz en condiciones de medida cuantitativa (desacoplamiento de paso inverso, retardo de relajación de 5 s). El tiempo de medida fue de aproximadamente 18 horas. Los desplazamientos químicos eran referidos a TMS (0 ppm). Las intensidades de las diferentes señales del grupo de carbono C1, C3 y C6 de los enlaces 1,3 y 1,6 en el espectro de RMN se usaron para calcular una proporción media de enlaces alfa-1,3 y enlaces alfa-1,6 de estos alternanos de 0,41/0,59 (error de 0,02).

### Ejemplo 4: Solubilidad del alternano

25 Se produjeron alternanos con alternano sacarasa truncada y con alternano sacarasa de longitud completa (ejemplo comparativo) como se describe en 1.5. El alternano producido con alternano sacarasa truncada también se abrevia en adelante en el presente documento como "SR17" y el alternano producido con alternano sacarasa de longitud completa también se abrevia en adelante en el presente documento como "SR16".

Se disolvieron los alternanos durante la noche a 50 °C a concentraciones de desde el 7 % hasta el 15 % (p/v). Se filtraron las muestras a través de una jeringuilla de 13 mm con filtro de PTFE de 5 µm Whatman Puradisc™ y se analizó la fracción de alternano disuelto por medio de GPC-MALLS. En la tabla 1 se muestra la fracción de alternano disuelto a diferentes concentraciones.

30 Tabla 1. Fracción de alternano disuelto a diferentes concentraciones, determinado por GPC-MALLS

	<b>Alternano producido a partir de alternano sacarasa truncada (SR17)</b>	<b>Alternano producido a partir de alternano sacarasa de longitud completa (SR16)</b>
Concentración (p/v)	% disuelto	% disuelto
7 %	91,2 %	87,1 %
8 %	89,3 %	86,8 %
9 %	87,4 %	80,9 %
10 %	88,0 %	83,2 %
11 %	95,2 %	66,1 %
12 %	90,8 %	71,1 %
13 %	67,8 %	45,0 %
15 %	66,5 %	59,5 %

### Ejemplo 5: Viscosidad del alternano

Se produjeron alternanos con alternano sacarasa truncada y con alternano sacarasa de longitud completa (ejemplo comparativo) como se describe en 1.5. El alternano producido con alternano sacarasa truncada también se abrevia en adelante en el presente documento como "SR17" y el alternano producido con alternano sacarasa de longitud completa también se abrevia en adelante en el presente documento como "SR16".

Se disolvieron los alternanos durante la noche a 50 °C a concentraciones de desde el 7 % hasta el 15 % (p/v). Se filtraron las muestras a través de una jeringuilla de 13 mm con un filtro de PTFE de 5 µm Whatman Puradisc™. Se determinó la viscosidad como se describe en los procedimientos. La viscosidad se muestra en la tabla 2.

Tabla 2. Viscosidad del alternano a diferentes concentraciones.

	<b>Alternano producido a partir de alternano sacarasa truncada (SR17)</b>		<b>Alternano producido a partir de alternano sacarasa de longitud completa (SR16)</b>	
Concentración (p/v)	Viscosidad (mPa*S)	Dev. estándar	Viscosidad (mPa*S)	Dev. estándar
8 %	9,4	1,9	17,7	1,8
9 %	13,9	2,1	39,3	2,0
10 %	29,9	3,2	120,6	3,8
11 %	60,7	2,5	367,9	11,1
12 %	240,3	9,3	927,2	7,4
13 %	702,0	14,1	1722,8	25,5
15 %	3605,7	94,5	6204,1	60,3

### Ejemplo 6: Prueba de diferencias en las reacciones con aceptor que afectan al desarrollo del alternano con relación a los oligosacáridos de alternano

#### 6.1 Materiales y procedimientos:

##### Expresión de proteínas:

El vector pAI-B-AISu contiene la secuencia codificante de la alternano sacarasa de longitud completa derivada de la cepa NRRL B-1355 de *Leuconostoc mesenteroides* que carece de los 39 aminoácidos N-terminales del péptido señal (la proteína comprende los aminoácidos 40-2057 de la SEQ ID NO 2), fusionada con un octapéptido con una marca estrep en el extremo C-terminal. La marca estrep se une a la proteína por medio de un dipéptido enlazador. La expresión de la alternano sacarasa se encuentra bajo en control transcripcional del promotor/operador y represor tetA. El promotor tetA está estrechamente regulado por el represor tet, que está codificado en el mismo plásmido y se expresa de forma constitutiva a partir del promotor de la β-lactamasa. De este modo, la expresión de la alternano sacarasa está reprimida de forma rigurosa hasta su inducción química eficaz por tetraciclina o anhidrotetraciclina, AHT.

Para la fermentación, se transformó el vector pAI-B-AISu en K12 DH5α de *E. coli* y se hicieron crecer de forma selectiva células bacterianas que albergaban el vector durante 28 h a 37 °C hasta una DO600 de 50 en medio mineral (Horn et al., 1996) complementado con ampicilina (100 µg/ml). Se indujo la expresión de la alternano sacarasa mediante la adición de anhidrotetraciclina (0,2 mg/l) y cultivo adicional durante 22 h a 25 °C hasta una DO600 de 140. Para la purificación de la enzima, se recogieron las células bacterianas por centrifugación (20000 rpm; 20 min) y se solubilizaron en tampón de resuspensión (acetato de sodio 100 mM, a pH 5,3).

Se provocó la disrupción de las células usando un homogeneizador de alta presión (dos ciclos, 1200 bar [ $11,8 \times 10^7$  Pa]). Se degradó el ácido nucleico bacteriano por tratamiento con DNasa/RNasa (3 mg/l) y se centrifugó el extracto resultante (9400 g durante 15 min a 4 °C) para recoger la materia celular insoluble. Se filtró el sobrenadante a través de Sterivex-GV de 0,22 µm (Millipore) y se almacenó a -20 °C hasta su uso.

- 5 El vector pAN94 contiene la versión truncada en C-terminal de la alternano sacarasa, que carece de los 742 aminoácidos C-terminales (la proteína comprende los aminoácidos 40-1315 de la SEQ ID NO 2). El vector pAN94 deriva del vector pAI-B y la expresión también se produce bajo el control transcripcional del promotor/operador y represor tetA. Para su expresión, se transformó el vector pAN94 en K12 DH5α de E. coli y se hicieron crecer de forma selectiva células bacterianas que albergaban el vector durante 4 h a 37 °C hasta una DO600 de 0,73 en 400 ml de medio YT (Sambrook, Fritsch, Maniatis) complementado con ampicilina (100 µg/ml). Se indujo la expresión de la alternano sacarasa truncada mediante la adición de anhidrotetraciclina (0,2 mg/l). También se añadieron IPTG 5 mM y glucosa al 1 %. El cultivo adicional se produce durante 25 h a 22 °C hasta una DO600 de 1,9. Para la purificación de la enzima, se recogieron las células bacterianas por centrifugación (4000 rpm; 20 min) y se solubilizaron en tampón de resuspensión (acetato de sodio 50 mM, a pH 5,3, DTT 2,5 mM). Se disrupción las células usando una prensa francesa (3 ciclos). Se centrifugó el extracto (33000 g durante 15 min a 4 °C) para recoger la materia celular insoluble. Se filtró el sobrenadante a través de Sterivex-GV de 0,22 µm (Millipore) y se almacenó a -20 °C hasta su uso.

#### Ensayo de actividad:

- 20 Se sometió a ensayo la actividad usando un ensayo fotométrico acoplado de NADP que medía la liberación de fructosa. Para las variantes de la alternano sacarasa, se define una unidad como la cantidad de enzima que cataliza la formación de 1 µmol de fructosa en un minuto a 37 °C en acetato de sodio 75 mM, a pH 5,2 y con 200 g/l de sacarosa.

#### Reacciones con aceptor

- 25 Se realizaron reacciones con aceptor en un volumen total de 20 ml con 0,22 unidades/ml, a 20 °C en tampón acetato de sodio 75 mM, a pH 5,2 y con 200 g/l de sacarosa. Como aceptor, se añadieron 80, 100, 120 g/l de maltosa monohidratada o isomaltulosa, respectivamente. Se realizó un seguimiento de la conversión de sacarosa después de 2,5 h, 18,5 h, 26 h, 42,5 h, y 48 h mediante el uso de un ensayo fotométrico acoplado a NADP que medía la liberación de fructosa y la sacarosa restante, y se detuvieron las reacciones por 10 min de incubación a 95 °C.

#### Análisis de glucanos

- 30 Tras la eliminación total de la sacarosa (48 h) se analiza la reacción con aceptor por s-MALLS-RI en columnas Styragel de WATERS (Styragel 7, Styragel 6E, Styragel 2 y Styragel de guarda). Como eluyente se usó nitrato de sodio 90 mM en dimetilsulfóxido (DMSO) a un caudal de 0,5 ml/min. Se calentaron las columnas hasta 60 °C. Se prepararon las muestras como sigue: Se añadieron 10 µl de muestra a 990 µl de DMSO. La muestra diluida se incubó a temperatura ambiente con rotación al mismo tiempo durante la noche. Después, se incubó la muestra diluida durante 2 horas a 95 °C mezclando al mismo tiempo a 1400 rpm en un termomezclador. Después de enfriarla, se centrifugó la muestra durante 5 min a 13000 rpm y se filtró a través de una jeringuilla con filtro de PTFE de 5 µm (Whatman). Se inyectaron 50 µl de la muestra filtrada. En caso posible, se determinaron y compararon las áreas de pico del índice de refracción (IR) del alternano y los productos con aceptor. En el caso de la maltosa, la cantidad de alternano es demasiado baja y, por lo tanto, solo fue posible una comparación de las señales de dispersión de luz, muchos más sensibles.

### **6.2 Resultados**

La cantidad de alternano producido por la variante de la alternano sacarasa de longitud completa es mayor en comparación con la cantidad producida por la alternano sacarasa truncada.

- 45 En el caso de la maltosa, los elugramas de s-MALLS-RI se muestran en la figura 3. La figura 3 muestra elugramas (IR y señales de dispersión de luz, DL) de muestras provenientes de reacciones con aceptor con maltosa al 10 % (AISa\_truncada = alternano sacarasa truncada, AISa\_completa = alternano sacarasa de longitud completa, OSA = oligosacáridos de alternano).

- 50 Debido a la elevada fuerza aceptora de la maltosa, no hay señales de IR provocadas por el alternano. Pero la diferencia de las señales de dispersión de luz muestra las diferentes cantidades de alternano. Usando la alternano sacarasa truncada, la señal de dispersión de luz resultante es mucho más baja en comparación con el uso de la variante de la alternano sacarasa de longitud completa en la reacción con aceptor.

- 55 En el caso de la isomaltulosa se muestran tanto los elugramas como las áreas de pico de IR en la figura 4 y la tabla 3, respectivamente. La figura 4 muestra elugramas (IR y señales de dispersión de luz, DL) de muestras provenientes de reacciones con aceptor con isomaltulosa al 10 % (AISa\_truncada = alternano sacarasa truncada, AISa\_completa = alternano sacarasa de longitud completa, OSA = oligosacáridos de alternano).

Debido a la baja fuerza como aceptor en comparación con la maltosa, se pueden detectar cantidades medibles de alternano. Usando la alternano sacarasa truncada, el área de pico de IR y la señal de dispersión de luz resultante es mucho más baja en comparación con el uso de la variante de la alternano sacarasa de longitud completa en la reacción con aceptor.

5 Tabla 3: Áreas de pico de IR determinadas por análisis s-MALLS-RI. El área relativa de alternano se calculó dividiendo el área de alternano entre el área total.

Enzima alternano sacarasa	Concentración de isomaltulosa	Área $\mu$ RIU*min de alternano	Área $\mu$ RIU*min de oligosacárido	Área $\mu$ RIU*min total	Área rel. (%) de alternano
Truncada	8 %	0,10	31,93	32,03	<b>0,30</b>
Truncada	10 %	0,04	37,25	37,29	<b>0,11</b>
Truncada	12 %	0,03	43,59	43,61	<b>0,06</b>
De longitud completa	8 %	0,35	31,52	31,88	<b>1,10</b>
De longitud completa	10 %	0,22	33,37	33,60	<b>0,67</b>
De longitud completa	12 %	0,24	38,80	39,04	<b>0,60</b>

**Listado de secuencias**

- <110> Bayer CropScience AG
- 10 <120> Moléculas de ácido nucleico que codifican alternano sacarosas truncadas
- <130> documento BCS 07 5003 PCT
- 15 <150> documento EP07090022 <151> 14-02-2007
- <150> documento US 60/901.532 <151> 15-02-2007
- <160> 25
- 20 <170> PatentIn versión 3.3
- <210> 1
- <211> 6174
- 25 <212> ADN
- <213> Leuconostoc mesenteroides
- <220>
- <221 > CDS
- 30 <222> (1).. (6171)
- <400> 1

ES 2 438 001 T3

atg	aaa	caa	caa	gaa	aca	gtt	acc	cgt	aaa	aaa	ctt	tat	aaa	tcc	ggt	48
Met	Lys	Gln	Gln	Glu	Thr	Val	Thr	Arg	Lys	Lys	Leu	Tyr	Lys	Ser	Gly	
1				5					10					15		
aag	gtt	tgg	gtt	gca	gca	gct	act	gca	ttt	gcg	gta	ttg	ggg	gtt	tca	96
Lys	Val	Trp	Val	Ala	Ala	Ala	Thr	Ala	Phe	Ala	Val	Leu	Gly	Val	Ser	
			20					25					30			
act	gta	aca	aca	gtc	cat	gcg	gat	aca	aat	tcg	aat	gtc	gct	gtt	aag	144
Thr	Val	Thr	Thr	Val	His	Ala	Asp	Thr	Asn	Ser	Asn	Val	Ala	Val	Lys	
			35				40					45				
caa	ata	aat	aat	aca	gga	acc	aat	gat	tct	ggc	gaa	aaa	aag	gta	ccg	192
Gln	Ile	Asn	Asn	Thr	Gly	Thr	Asn	Asp	Ser	Gly	Glu	Lys	Lys	Val	Pro	
			50			55					60					
gtt	cca	tca	act	aat	aat	gat	agt	ttg	aag	caa	gga	aca	gat	ggt	ttt	240
Val	Pro	Ser	Thr	Asn	Asn	Asp	Ser	Leu	Lys	Gln	Gly	Thr	Asp	Gly	Phe	
65					70					75					80	
tgg	tat	gat	tca	gac	ggc	aat	cgt	gtc	gat	cag	aag	acc	aat	cag	att	288
Trp	Tyr	Asp	Ser	Asp	Gly	Asn	Arg	Val	Asp	Gln	Lys	Thr	Asn	Gln	Ile	
				85					90					95		
ctg	ctt	act	gcg	gaa	caa	ctt	aaa	aaa	aat	aac	gaa	aaa	aat	tta	tca	336
Leu	Leu	Thr	Ala	Glu	Gln	Leu	Lys	Lys	Asn	Asn	Glu	Lys	Asn	Leu	Ser	
			100					105					110			
gta	atc	agt	gat	gat	aca	tca	aaa	aaa	gat	gat	gaa	aat	att	tct	aag	384
Val	Ile	Ser	Asp	Asp	Thr	Ser	Lys	Lys	Asp	Asp	Glu	Asn	Ile	Ser	Lys	
			115				120					125				

ES 2 438 001 T3

cag acc aaa att gct aat caa caa aca gta gat act gct aaa ggc ctg	432
Gln Thr Lys Ile Ala Asn Gln Gln Thr Val Asp Thr Ala Lys Gly Leu	
130 135 140	
act acc agt aat tta tct gat ccc atc act ggg ggt cac tat gaa aat	480
Thr Thr Ser Asn Leu Ser Asp Pro Ile Thr Gly Gly His Tyr Glu Asn	
145 150 155 160	
cac aat ggc tac ttt gtt tat ata gat gct tca gga aaa caa gta aca	528
His Asn Gly Tyr Phe Val Tyr Ile Asp Ala Ser Gly Lys Gln Val Thr	
165 170 175	
ggg ttg caa aat att gat ggt aat tta caa tat ttt gat gac aat gga	576
Gly Leu Gln Asn Ile Asp Gly Asn Leu Gln Tyr Phe Asp Asp Asn Gly	
180 185 190	
tat caa gtc aag gga tcc ttc cga gat gtc aac ggc aag cat atc tat	624
Tyr Gln Val Lys Gly Ser Phe Arg Asp Val Asn Gly Lys His Ile Tyr	
195 200 205	
ttt gat tca gta aca ggg aaa gct agt tca aat gtt gat att gtt aac	672
Phe Asp Ser Val Thr Gly Lys Ala Ser Ser Asn Val Asp Ile Val Asn	
210 215 220	
ggg aaa gct caa gga tat gat gcg caa ggc aac caa tta aag aaa agt	720
Gly Lys Ala Gln Gly Tyr Asp Ala Gln Gly Asn Gln Leu Lys Lys Ser	
225 230 235 240	
tat gtc gcc gat agt tct ggg caa act tac tat ttt gat ggt aat ggc	768
Tyr Val Ala Asp Ser Ser Gly Gln Thr Tyr Tyr Phe Asp Gly Asn Gly	
245 250 255	
caa ccg tta atc ggc ttg caa aca att gat ggg aac cta caa tat ttt	816
Gln Pro Leu Ile Gly Leu Gln Thr Ile Asp Gly Asn Leu Gln Tyr Phe	
260 265 270	
aac caa caa ggg gtt caa ata aag ggt ggt ttc caa gat gtt aac aat	864
Asn Gln Gln Gly Val Gln Ile Lys Gly Gly Phe Gln Asp Val Asn Asn	
275 280 285	
aaa cgt att tat ttt gca cca aac aca ggt aat gcc gtt gcc aat act	912
Lys Arg Ile Tyr Phe Ala Pro Asn Thr Gly Asn Ala Val Ala Asn Thr	
290 295 300	
gaa ata att aac ggt aaa tta cag ggg cgt gac gca aat ggt aac cag	960
Glu Ile Ile Asn Gly Lys Leu Gln Gly Arg Asp Ala Asn Gly Asn Gln	
305 310 315 320	
gta aag aat gca ttt agt aaa gat gtt gca gga aat aca ttt tat ttt	1008
Val Lys Asn Ala Phe Ser Lys Asp Val Ala Gly Asn Thr Phe Tyr Phe	
325 330 335	
gac gca aac ggt gtg atg tta aca ggg ttg caa act att tca gga aag	1056
Asp Ala Asn Gly Val Met Leu Thr Gly Leu Gln Thr Ile Ser Gly Lys	
340 345 350	
aca tat tat ctt gat gaa caa gga cac ctg aga aaa aat tac gcg gga	1104
Thr Tyr Tyr Leu Asp Glu Gln Gly His Leu Arg Lys Asn Tyr Ala Gly	
355 360 365	
aca ttc aat aat cag ttt atg tac ttc gat gct gat aca ggt gcg ggt	1152

Thr	Phe	Asn	Asn	Gln	Phe	Met	Tyr	Phe	Asp	Ala	Asp	Thr	Gly	Ala	Gly		
370						375					380						
aaa	aca	gcg	att	gaa	tat	caa	ttt	gat	caa	gga	ttg	gta	tca	caa	agt	1200	
Lys	Thr	Ala	Ile	Glu	Tyr	Gln	Phe	Asp	Gln	Gly	Leu	Val	Ser	Gln	Ser	400	
385				390					395								
aat	gaa	aat	act	cct	cac	aat	gcc	gca	aag	tct	tat	gat	aaa	agt	agt	1248	
Asn	Glu	Asn	Thr	Pro	His	Asn	Ala	Ala	Lys	Ser	Tyr	Asp	Lys	Ser	Ser	415	
				405					410								
ttt	gaa	aat	ggt	gat	ggg	tac	tta	aca	gca	gat	aca	tgg	tat	cgt	cca	1296	
Phe	Glu	Asn	Val	Asp	Gly	Tyr	Leu	Thr	Ala	Asp	Thr	Trp	Tyr	Arg	Pro		
			420					425					430				
acc	gat	att	tta	aaa	aat	gga	gat	act	tgg	acg	gca	tct	acc	gaa	act	1344	
Thr	Asp	Ile	Leu	Lys	Asn	Gly	Asp	Thr	Trp	Thr	Ala	Ser	Thr	Glu	Thr		
		435					440					445					
gat	atg	cgt	ccg	ctt	tta	atg	aca	tgg	tgg	cct	gac	aaa	caa	aca	caa	1392	
Asp	Met	Arg	Pro	Leu	Leu	Met	Thr	Trp	Trp	Pro	Asp	Lys	Gln	Thr	Gln		
	450					455					460						
gca	aat	tac	ttg	aat	ttt	atg	tct	agt	aaa	gga	ctt	ggt	ata	acg	acc	1440	
Ala	Asn	Tyr	Leu	Asn	Phe	Met	Ser	Ser	Lys	Gly	Leu	Gly	Ile	Thr	Thr	480	
465					470					475							
act	tat	aca	gca	gct	acg	tca	caa	aaa	aca	cta	aat	gac	gca	gcc	ttt	1488	
Thr	Tyr	Thr	Ala	Ala	Thr	Ser	Gln	Lys	Thr	Leu	Asn	Asp	Ala	Ala	Phe		
				485				490						495			
gtt	att	caa	aca	gca	att	gaa	caa	caa	ata	tct	ttg	aaa	aaa	agt	act	1536	
Val	Ile	Gln	Thr	Ala	Ile	Glu	Gln	Gln	Ile	Ser	Leu	Lys	Lys	Ser	Thr		
			500					505					510				
gag	tgg	tta	cgt	gat	gca	att	gat	agt	ttt	gtg	aag	acg	caa	gct	aat	1584	
Glu	Trp	Leu	Arg	Asp	Ala	Ile	Asp	Ser	Phe	Val	Lys	Thr	Gln	Ala	Asn		
		515					520					525					
tgg	aat	aag	caa	aca	gaa	gat	gaa	gct	ttc	gat	ggt	ttg	cag	tgg	ctt	1632	
Trp	Asn	Lys	Gln	Thr	Glu	Asp	Glu	Ala	Phe	Asp	Gly	Leu	Gln	Trp	Leu		
	530					535					540						
caa	ggg	gga	ttc	cta	gct	tat	caa	gat	gat	tca	cat	cgg	acg	ccg	aat	1680	
Gln	Gly	Gly	Phe	Leu	Ala	Tyr	Gln	Asp	Asp	Ser	His	Arg	Thr	Pro	Asn	560	
545					550					555							
act	gat	tca	gga	aat	aac	aga	aaa	cta	gga	cgt	caa	cca	att	aat	atc	1728	
Thr	Asp	Ser	Gly	Asn	Asn	Arg	Lys	Leu	Gly	Arg	Gln	Pro	Ile	Asn	Ile		
				565					570					575			
gat	ggt	tcg	aaa	gat	aca	act	gat	ggt	aaa	ggc	tct	gaa	ttc	tta	tta	1776	
Asp	Gly	Ser	Lys	Asp	Thr	Thr	Asp	Gly	Lys	Gly	Ser	Glu	Phe	Leu	Leu		
			580					585					590				
gct	aac	gat	att	gac	aac	tca	aat	ccg	att	ggt	caa	gct	gag	caa	tta	1824	
Ala	Asn	Asp	Ile	Asp	Asn	Ser	Asn	Pro	Ile	Val	Gln	Ala	Glu	Gln	Leu		
		595					600					605					
aac	tgg	cta	cac	tat	tta	atg	aat	ttt	ggt	agt	att	aca	ggt	aat	aat	1872	
Asn	Trp	Leu	His	Tyr	Leu	Met	Asn	Phe	Gly	Ser	Ile	Thr	Gly	Asn	Asn		

ES 2 438 001 T3

610	615	620	
gac aat gcg aat ttt gat	ggc att cgt gta gat	gct gtt gat aat gtt	1920
Asp Asn Ala Asn Phe Asp	Gly Ile Arg Val Asp	Ala Val Asp Asn Val	
625	630	635 640	
gat gct gat tta cta aaa	ata gct ggc gat tat	ttt aaa gct cta tat	1968
Asp Ala Asp Leu Leu Lys	Ile Ala Gly Asp Tyr	Phe Lys Ala Leu Tyr	
	645	650 655	
ggt aca gat aaa agc gac	gcc aat gcc aat aag	cat ttg tct att tta	2016
Gly Thr Asp Lys Ser Asp	Ala Asn Ala Asn Lys	His Leu Ser Ile Leu	
	660	665 670	
gaa gac tgg aac ggt aaa	gat cct cag tat gtt	aat caa cag ggc aat	2064
Glu Asp Trp Asn Gly Lys	Asp Pro Gln Tyr Val	Asn Gln Gln Gly Asn	
	675	680 685	
gcg caa tta aca atg gat	tac aca gtt act tca	cag ttt ggc aat tct	2112
Ala Gln Leu Thr Met Asp	Tyr Thr Val Thr Ser	Gln Phe Gly Asn Ser	
	690 695	700	
cta aca cat ggc gcc aac	aac agg agt aac atg	tgg tat ttc tta gat	2160
Leu Thr His Gly Ala Asn	Asn Arg Ser Asn Met	Trp Tyr Phe Leu Asp	
	710	715 720	
act ggc tat tat ctt aat	gga gat ctt aat aag	aag ata gta gat aag	2208
Thr Gly Tyr Tyr Leu Asn	Gly Asp Leu Asn Lys	Lys Ile Val Asp Lys	
	725	730 735	
aac cgt cca aat tct ggc	act ttg gtt aac aga	att gct aat tca ggt	2256
Asn Arg Pro Asn Ser Gly	Thr Leu Val Asn Arg	Ile Ala Asn Ser Gly	
	740	745 750	
gat aca aaa gtt att cca	aat tat agt ttt gtt	aga gca cat gat tac	2304
Asp Thr Lys Val Ile Pro	Asn Tyr Ser Phe Val	Arg Ala His Asp Tyr	
	755	760 765	
gat gct caa gat cca att	aga aaa gcc atg att	gat cat ggt att att	2352
Asp Ala Gln Asp Pro Ile	Arg Lys Ala Met Ile	Asp His Gly Ile Ile	
	770 775	780	
aaa aac atg cag gat act	ttc act ttt gac caa	ctg gct cag gga atg	2400
Lys Asn Met Gln Asp Thr	Phe Thr Phe Asp Gln	Leu Ala Gln Gly Met	
	785 790	795 800	
gaa ttc tac tat aaa gat	caa gag aat ccg tct	ggt ttc aaa aag tat	2448
Glu Phe Tyr Tyr Lys Asp	Gln Glu Asn Pro Ser	Gly Phe Lys Lys Tyr	
	805	810 815	
aac gat tat aac tta cct	agt gct tat gca atg	ttg ttg act aat aag	2496
Asn Asp Tyr Asn Leu Pro	Ser Ala Tyr Ala Met	Leu Leu Thr Asn Lys	
	820	825 830	
gat act gta cct cgt gtc	tat tat gga gat atg	tac ctc gaa ggc ggg	2544
Asp Thr Val Pro Arg Val	Tyr Tyr Gly Asp Met	Tyr Leu Glu Gly Gly	
	835 840	845	
caa tat atg gaa aaa ggg	acg att tac aat cct	gtc att tca gcg ttg	2592
Gln Tyr Met Glu Lys Gly	Thr Ile Tyr Asn Pro	Val Ile Ser Ala Leu	
	850 855	860	



ES 2 438 001 T3

gca cct caa tat agg tct agt ggt gac act aat tac ggt ggc atg Ala Pro Gln Tyr Arg Ser Ser Gly Asp Thr Asn Tyr Gly Gly Met 1100 1105 1110	3339
tca ttc tta gat tct ttc tta aat aat ggt tat gca ttt acc gat Ser Phe Leu Asp Ser Phe Leu Asn Asn Gly Tyr Ala Phe Thr Asp 1115 1120 1125	3384
aga tat gat tta ggc ttt aac aaa gca gac ggg aat cct aac cca Arg Tyr Asp Leu Gly Phe Asn Lys Ala Asp Gly Asn Pro Asn Pro 1130 1135 1140	3429
aca aag tat gga aca gat caa gat tta cgt aat gca ata gag gca Thr Lys Tyr Gly Thr Asp Gln Asp Leu Arg Asn Ala Ile Glu Ala 1145 1150 1155	3474
tta cac aaa aac ggc atg cag gct ata gct gat tgg gtt cct gac Leu His Lys Asn Gly Met Gln Ala Ile Ala Asp Trp Val Pro Asp 1160 1165 1170	3519
caa ata tat gct tta cca gga aag gaa gtt gtt acc gct act aga Gln Ile Tyr Ala Leu Pro Gly Lys Glu Val Val Thr Ala Thr Arg 1175 1180 1185	3564
gta gac gaa cgg gga aat caa cta aaa gac aca gat ttt gtc aac Val Asp Glu Arg Gly Asn Gln Leu Lys Asp Thr Asp Phe Val Asn 1190 1195 1200	3609
tta ctc tat gtt gct aat act aaa agt agt ggt gtg gat tat cag Leu Leu Tyr Val Ala Asn Thr Lys Ser Ser Gly Val Asp Tyr Gln 1205 1210 1215	3654
gca aag tat ggc ggc gaa ttt tta gat aaa tta aga gaa gag tac Ala Lys Tyr Gly Gly Glu Phe Leu Asp Lys Leu Arg Glu Glu Tyr 1220 1225 1230	3699
cca tcg tta ttc aaa cag aac caa gta tcg aca ggt cag cca att Pro Ser Leu Phe Lys Gln Asn Gln Val Ser Thr Gly Gln Pro Ile 1235 1240 1245	3744
gat gct tct aca aaa att aag caa tgg tca gct aaa tat atg aat Asp Ala Ser Thr Lys Ile Lys Gln Trp Ser Ala Lys Tyr Met Asn 1250 1255 1260	3789
ggg acc aat att tta cat cga ggt gct tat tat gtt ttg aaa gac Gly Thr Asn Ile Leu His Arg Gly Ala Tyr Tyr Val Leu Lys Asp 1265 1270 1275	3834
tgg gct act aac cag tat ttt aac att gca aaa acg aat gaa gta Trp Ala Thr Asn Gln Tyr Phe Asn Ile Ala Lys Thr Asn Glu Val 1280 1285 1290	3879
ttt ttg cca cta cag ttg cag aat aaa gat gcg caa act ggt ttc Phe Leu Pro Leu Gln Leu Gln Asn Lys Asp Ala Gln Thr Gly Phe 1295 1300 1305	3924
att agt gat gcc tcc ggt gta aaa tat tac tca att agt ggt tat Ile Ser Asp Ala Ser Gly Val Lys Tyr Tyr Ser Ile Ser Gly Tyr 1310 1315 1320	3969
caa gca aaa gat act ttt att gaa gat ggt aat ggg aat tgg tat	4014

Gln Ala	Lys Asp Thr Phe Ile	Glu Asp Gly Asn Gly	Asn Trp Tyr	
1325	1330	1335		
tac ttt gat aaa gat ggt tac atg gtg cgt tcg cag caa gga gaa				4059
Tyr Phe Asp Lys Asp Gly Tyr Met Val Arg Ser Gln Gln Gly Glu				
1340	1345	1350		
aat cct ata aga aca gtc gaa act agt gtc aac aca cga aac ggt				4104
Asn Pro Ile Arg Thr Val Glu Thr Ser Val Asn Thr Arg Asn Gly				
1355	1360	1365		
aat tat tac ttt atg cca aat ggt gtc gag ttg cgc aaa ggc ttt				4149
Asn Tyr Tyr Phe Met Pro Asn Gly Val Glu Leu Arg Lys Gly Phe				
1370	1375	1380		
gga acg gat aat agt ggt aat gtc tat tat ttt gat gat caa ggt				4194
Gly Thr Asp Asn Ser Gly Asn Val Tyr Tyr Phe Asp Asp Gln Gly				
1385	1390	1395		
aag atg gtg aga gat aaa tac att aac gat gat gct aat aat ttt				4239
Lys Met Val Arg Asp Lys Tyr Ile Asn Asp Asp Ala Asn Asn Phe				
1400	1405	1410		
tat cac tta aat gtt gat ggg act atg tct cga gga cta ttt aaa				4284
Tyr His Leu Asn Val Asp Gly Thr Met Ser Arg Gly Leu Phe Lys				
1415	1420	1425		
ttt gat tct gat act cta cag tat ttt gct agt aat ggt gtc caa				4329
Phe Asp Ser Asp Thr Leu Gln Tyr Phe Ala Ser Asn Gly Val Gln				
1430	1435	1440		
ata aaa gat agt tat gcg aag gat agt aaa ggc aat aaa tat tat				4374
Ile Lys Asp Ser Tyr Ala Lys Asp Ser Lys Gly Asn Lys Tyr Tyr				
1445	1450	1455		
ttt gac tca gct aca gga aat aac gat act ggg aaa gcc caa act				4419
Phe Asp Ser Ala Thr Gly Asn Asn Asp Thr Gly Lys Ala Gln Thr				
1460	1465	1470		
tgg gat ggt aat ggc tac tat att act att gat tct gat gcg aac				4464
Trp Asp Gly Asn Gly Tyr Tyr Ile Thr Ile Asp Ser Asp Ala Asn				
1475	1480	1485		
aat aca att ggg gtt aac aca gac tac act gcc tac atc act agc				4509
Asn Thr Ile Gly Val Asn Thr Asp Tyr Thr Ala Tyr Ile Thr Ser				
1490	1495	1500		
tcg ctg cgc gaa gat ggc tta ttt gct aac gca cct tac ggt gtt				4554
Ser Leu Arg Glu Asp Gly Leu Phe Ala Asn Ala Pro Tyr Gly Val				
1505	1510	1515		
gta aca aaa gac caa aat ggt aac gat ctt aag tgg cag tat att				4599
Val Thr Lys Asp Gln Asn Gly Asn Asp Leu Lys Trp Gln Tyr Ile				
1520	1525	1530		
aac cat acg aaa cag tac gaa ggg caa caa gtg caa gtc acg cgt				4644
Asn His Thr Lys Gln Tyr Glu Gly Gln Gln Val Gln Val Thr Arg				
1535	1540	1545		
caa tac aca gac agt aag gga gtc agc tgg aac tta att acc ttt				4689
Gln Tyr Thr Asp Ser Lys Gly Val Ser Trp Asn Leu Ile Thr Phe				

ES 2 438 001 T3

1550		1555		1560		
gct ggt ggt gat tta caa gga	caa agg ctt tgg gtg	gat agt cgt	4734			
Ala Gly Gly Asp Leu Gln Gly	Gln Arg Leu Trp Val	Asp Ser Arg				
1565	1570	1575				
gcg tta act atg aca cca ttt	aaa acg atg aac caa	ata agc ttc	4779			
Ala Leu Thr Met Thr Pro Phe	Lys Thr Met Asn Gln	Ile Ser Phe				
1580	1585	1590				
att agt tat gct aac cgc aat	gat ggg ttg ttt ttg	aat gcg cca	4824			
Ile Ser Tyr Ala Asn Arg Asn	Asp Gly Leu Phe Leu	Asn Ala Pro				
1595	1600	1605				
tac caa gtc aag ggg tat caa	tta gct ggg atg tcc	aac caa tac	4869			
Tyr Gln Val Lys Gly Tyr Gln	Leu Ala Gly Met Ser	Asn Gln Tyr				
1610	1615	1620				
aag ggc caa caa gtg acc att	gct ggg gtg gcg aac	gtt tct gga	4914			
Lys Gly Gln Gln Val Thr Ile	Ala Gly Val Ala Asn	Val Ser Gly				
1625	1630	1635				
aaa gac tgg agt ctg att agt	ttt aat ggg aca cag	tac tgg att	4959			
Lys Asp Trp Ser Leu Ile Ser	Phe Asn Gly Thr Gln	Tyr Trp Ile				
1640	1645	1650				
gat agt cag gca ttg aat acc	aat ttc aca cat gac	atg aac caa	5004			
Asp Ser Gln Ala Leu Asn Thr	Asn Phe Thr His Asp	Met Asn Gln				
1655	1660	1665				
aag gtc ttt gtc aat aca act	agt aat ctt gat ggg	tta ttc tta	5049			
Lys Val Phe Val Asn Thr Thr	Ser Asn Leu Asp Gly	Leu Phe Leu				
1670	1675	1680				
aat gcg cca tac cgt caa ccg	ggt tat aag tta gcc	ggt ttg gct	5094			
Asn Ala Pro Tyr Arg Gln Pro	Gly Tyr Lys Leu Ala	Gly Leu Ala				
1685	1690	1695				
aaa aat tac aac aac caa acg	gtt act gtt agt caa	cag tac ttt	5139			
Lys Asn Tyr Asn Asn Gln Thr	Val Thr Val Ser Gln	Gln Tyr Phe				
1700	1705	1710				
gat gat caa ggc acg gtc tgg	agt cag gtt gtc ctt	ggg ggt cag	5184			
Asp Asp Gln Gly Thr Val Trp	Ser Gln Val Val Leu	Gly Gly Gln				
1715	1720	1725				
acg gtc tgg gtt gat aac cat	gca ttg gca cag atg	caa gtt agt	5229			
Thr Val Trp Val Asp Asn His	Ala Leu Ala Gln Met	Gln Val Ser				
1730	1735	1740				
gat aca gac caa cag ctg tat	gtg aat agc aat ggt	cgg aat gat	5274			
Asp Thr Asp Gln Gln Leu Tyr	Val Asn Ser Asn Gly	Arg Asn Asp				
1745	1750	1755				
ggg tta ttc ttg aat gcg cca	tat cgt ggt caa ggg	tca caa ctg	5319			
Gly Leu Phe Leu Asn Ala Pro	Tyr Arg Gly Gln Gly	Ser Gln Leu				
1760	1765	1770				
ata ggc atg acg gca gat tat	aat ggg caa cat gta	caa gtg acc	5364			
Ile Gly Met Thr Ala Asp Tyr	Asn Gly Gln His Val	Gln Val Thr				
1775	1780	1785				

aag caa ggg caa gat gcc tat ggt gca caa tgg cgt ctt att acg	5409
Lys Gln Gly Gln Asp Ala Tyr Gly Ala Gln Trp Arg Leu Ile Thr	
1790 1795 1800	
cta aat aat caa cag gtc tgg gtt gat agt cgc gct ttg agc aca	5454
Leu Asn Asn Gln Gln Val Trp Val Asp Ser Arg Ala Leu Ser Thr	
1805 1810 1815	
aca atc atg caa gcc atg aat gat aat atg tat gta aat agc agc	5499
Thr Ile Met Gln Ala Met Asn Asp Asn Met Tyr Val Asn Ser Ser	
1820 1825 1830	
caa cgg aca gat ggc ttg tgg tta aac gca cct tat acg atg agt	5544
Gln Arg Thr Asp Gly Leu Trp Leu Asn Ala Pro Tyr Thr Met Ser	
1835 1840 1845	
ggg gct aaa tgg gct ggt gat aca cgt tca gct aat ggg cgc tat	5589
Gly Ala Lys Trp Ala Gly Asp Thr Arg Ser Ala Asn Gly Arg Tyr	
1850 1855 1860	
gtc cat att tca aaa gct tat tca aac gaa gtc ggc aat aca tat	5634
Val His Ile Ser Lys Ala Tyr Ser Asn Glu Val Gly Asn Thr Tyr	
1865 1870 1875	
tac ttg acg aat ttg aat ggt caa agc aca tgg att gac aag cgg	5679
Tyr Leu Thr Asn Leu Asn Gly Gln Ser Thr Trp Ile Asp Lys Arg	
1880 1885 1890	
gcg ttt act gtg acc ttc gat cag gtg gtg gca tta aat gca acg	5724
Ala Phe Thr Val Thr Phe Asp Gln Val Val Ala Leu Asn Ala Thr	
1895 1900 1905	
att gtg gca cgc caa cga cca gat ggg atg ttt aag aca gca cca	5769
Ile Val Ala Arg Gln Arg Pro Asp Gly Met Phe Lys Thr Ala Pro	
1910 1915 1920	
tat ggt gaa gcg ggg gcg cag ttt gtc gat tat gtg aca aac tat	5814
Tyr Gly Glu Ala Gly Ala Gln Phe Val Asp Tyr Val Thr Asn Tyr	
1925 1930 1935	
aac cag caa acc gtg cca gta aca aag caa cat tca gat gct cag	5859
Asn Gln Gln Thr Val Pro Val Thr Lys Gln His Ser Asp Ala Gln	
1940 1945 1950	
ggg aat caa tgg tac tta gcg aca gtg aat ggg aca caa tac tgg	5904
Gly Asn Gln Trp Tyr Leu Ala Thr Val Asn Gly Thr Gln Tyr Trp	
1955 1960 1965	
att gat caa cgg tca ttt tca cca gta gta acg aag gtg gtt gat	5949
Ile Asp Gln Arg Ser Phe Ser Pro Val Val Thr Lys Val Val Asp	
1970 1975 1980	
tat caa gct aag att gtg cca cgg aca aca cgt gat ggt gtg ttt	5994
Tyr Gln Ala Lys Ile Val Pro Arg Thr Thr Arg Asp Gly Val Phe	
1985 1990 1995	
agt ggc gca ccc tat ggg gaa gtg aat gct aag cta gtt aac atg	6039
Ser Gly Ala Pro Tyr Gly Glu Val Asn Ala Lys Leu Val Asn Met	
2000 2005 2010	

ES 2 438 001 T3

```

gca act gcg tat caa aat caa gtt gtc cat gcg aca ggg gaa tat      6084
Ala Thr Ala Tyr Gln Asn Gln Val Val His Ala Thr Gly Glu Tyr
    2015                                2020                                2025

acg aat gct tca ggg atc aca tgg agt cag ttc gcg tta agc ggg      6129
Thr Asn Ala Ser Gly Ile Thr Trp Ser Gln Phe Ala Leu Ser Gly
    2030                                2035                                2040

caa gaa gac aag cta tgg att gat aag cgt gct ttg caa gct taa      6174
Gln Glu Asp Lys Leu Trp Ile Asp Lys Arg Ala Leu Gln Ala
    2045                                2050                                2055

```

5 <210>2  
 <211> 2057  
 <212> PRT  
 <213> Leuconostoc mesenteroides  
 <400> 2

```

Met Lys Gln Gln Glu Thr Val Thr Arg Lys Lys Leu Tyr Lys Ser Gly
 1                                5                                10                                15

Lys Val Trp Val Ala Ala Ala Thr Ala Phe Ala Val Leu Gly Val Ser
                20                                25                                30

Thr Val Thr Thr Val His Ala Asp Thr Asn Ser Asn Val Ala Val Lys
        35                                40                                45

Gln Ile Asn Asn Thr Gly Thr Asn Asp Ser Gly Glu Lys Lys Val Pro
 50                                55                                60

Val Pro Ser Thr Asn Asn Asp Ser Leu Lys Gln Gly Thr Asp Gly Phe
 65                                70                                75                                80

Trp Tyr Asp Ser Asp Gly Asn Arg Val Asp Gln Lys Thr Asn Gln Ile
                85                                90                                95

Leu Leu Thr Ala Glu Gln Leu Lys Lys Asn Asn Glu Lys Asn Leu Ser
                100                                105                                110

Val Ile Ser Asp Asp Thr Ser Lys Lys Asp Asp Glu Asn Ile Ser Lys
        115                                120                                125

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

Thr Thr Ser Asn Leu Ser Asp Pro Ile Thr Gly Gly His Tyr Glu Asn
 145                                150                                155                                160

His Asn Gly Tyr Phe Val Tyr Ile Asp Ala Ser Gly Lys Gln Val Thr

```

ES 2 438 001 T3

				165							170								175
Gly	Leu	Gln	Asn	Ile	Asp	Gly	Asn	Leu	Gln	Tyr	Phe	Asp	Asp	Asn	Gly				
			180					185					190						
Tyr	Gln	Val	Lys	Gly	Ser	Phe	Arg	Asp	Val	Asn	Gly	Lys	His	Ile	Tyr				
		195					200					205							
Phe	Asp	Ser	Val	Thr	Gly	Lys	Ala	Ser	Ser	Asn	Val	Asp	Ile	Val	Asn				
	210					215					220								
Gly	Lys	Ala	Gln	Gly	Tyr	Asp	Ala	Gln	Gly	Asn	Gln	Leu	Lys	Lys	Ser				
225					230					235					240				
Tyr	Val	Ala	Asp	Ser	Ser	Gly	Gln	Thr	Tyr	Tyr	Phe	Asp	Gly	Asn	Gly				
				245					250					255					
Gln	Pro	Leu	Ile	Gly	Leu	Gln	Thr	Ile	Asp	Gly	Asn	Leu	Gln	Tyr	Phe				
			260					265					270						
Asn	Gln	Gln	Gly	Val	Gln	Ile	Lys	Gly	Gly	Phe	Gln	Asp	Val	Asn	Asn				
		275					280					285							
Lys	Arg	Ile	Tyr	Phe	Ala	Pro	Asn	Thr	Gly	Asn	Ala	Val	Ala	Asn	Thr				
	290					295					300								
Glu	Ile	Ile	Asn	Gly	Lys	Leu	Gln	Gly	Arg	Asp	Ala	Asn	Gly	Asn	Gln				
305					310					315					320				
Val	Lys	Asn	Ala	Phe	Ser	Lys	Asp	Val	Ala	Gly	Asn	Thr	Phe	Tyr	Phe				
				325					330					335					
Asp	Ala	Asn	Gly	Val	Met	Leu	Thr	Gly	Leu	Gln	Thr	Ile	Ser	Gly	Lys				
			340					345					350						
Thr	Tyr	Tyr	Leu	Asp	Glu	Gln	Gly	His	Leu	Arg	Lys	Asn	Tyr	Ala	Gly				
		355					360					365							
Thr	Phe	Asn	Asn	Gln	Phe	Met	Tyr	Phe	Asp	Ala	Asp	Thr	Gly	Ala	Gly				
	370					375					380								
Lys	Thr	Ala	Ile	Glu	Tyr	Gln	Phe	Asp	Gln	Gly	Leu	Val	Ser	Gln	Ser				
385					390					395					400				
Asn	Glu	Asn	Thr	Pro	His	Asn	Ala	Ala	Lys	Ser	Tyr	Asp	Lys	Ser	Ser				
				405					410					415					

ES 2 438 001 T3

Phe Glu Asn Val Asp Gly Tyr Leu Thr Ala Asp Thr Trp Tyr Arg Pro  
 420 425 430

Thr Asp Ile Leu Lys Asn Gly Asp Thr Trp Thr Ala Ser Thr Glu Thr  
 435 440 445

Asp Met Arg Pro Leu Leu Met Thr Trp Trp Pro Asp Lys Gln Thr Gln  
 450 455 460

Ala Asn Tyr Leu Asn Phe Met Ser Ser Lys Gly Leu Gly Ile Thr Thr  
 465 470 475 480

Thr Tyr Thr Ala Ala Thr Ser Gln Lys Thr Leu Asn Asp Ala Ala Phe  
 485 490 495

Val Ile Gln Thr Ala Ile Glu Gln Gln Ile Ser Leu Lys Lys Ser Thr  
 500 505 510

Glu Trp Leu Arg Asp Ala Ile Asp Ser Phe Val Lys Thr Gln Ala Asn  
 515 520 525

Trp Asn Lys Gln Thr Glu Asp Glu Ala Phe Asp Gly Leu Gln Trp Leu  
 530 535 540

Gln Gly Gly Phe Leu Ala Tyr Gln Asp Asp Ser His Arg Thr Pro Asn  
 545 550 555 560

Thr Asp Ser Gly Asn Asn Arg Lys Leu Gly Arg Gln Pro Ile Asn Ile  
 565 570 575

Asp Gly Ser Lys Asp Thr Thr Asp Gly Lys Gly Ser Glu Phe Leu Leu  
 580 585 590

Ala Asn Asp Ile Asp Asn Ser Asn Pro Ile Val Gln Ala Glu Gln Leu  
 595 600 605

Asn Trp Leu His Tyr Leu Met Asn Phe Gly Ser Ile Thr Gly Asn Asn  
 610 615 620

Asp Asn Ala Asn Phe Asp Gly Ile Arg Val Asp Ala Val Asp Asn Val  
 625 630 635 640

Asp Ala Asp Leu Leu Lys Ile Ala Gly Asp Tyr Phe Lys Ala Leu Tyr  
 645 650 655

ES 2 438 001 T3

Gly Thr Asp Lys Ser Asp Ala Asn Ala Asn Lys His Leu Ser Ile Leu  
 660 665 670  
 Glu Asp Trp Asn Gly Lys Asp Pro Gln Tyr Val Asn Gln Gln Gly Asn  
 675 680 685  
 Ala Gln Leu Thr Met Asp Tyr Thr Val Thr Ser Gln Phe Gly Asn Ser  
 690 695 700  
 Leu Thr His Gly Ala Asn Asn Arg Ser Asn Met Trp Tyr Phe Leu Asp  
 705 710 715 720  
 Thr Gly Tyr Tyr Leu Asn Gly Asp Leu Asn Lys Lys Ile Val Asp Lys  
 725 730 735  
 Asn Arg Pro Asn Ser Gly Thr Leu Val Asn Arg Ile Ala Asn Ser Gly  
 740 745 750  
 Asp Thr Lys Val Ile Pro Asn Tyr Ser Phe Val Arg Ala His Asp Tyr  
 755 760 765  
 Asp Ala Gln Asp Pro Ile Arg Lys Ala Met Ile Asp His Gly Ile Ile  
 770 775 780  
 Lys Asn Met Gln Asp Thr Phe Thr Phe Asp Gln Leu Ala Gln Gly Met  
 785 790 795 800  
 Glu Phe Tyr Tyr Lys Asp Gln Glu Asn Pro Ser Gly Phe Lys Lys Tyr  
 805 810 815  
 Asn Asp Tyr Asn Leu Pro Ser Ala Tyr Ala Met Leu Leu Thr Asn Lys  
 820 825 830  
 Asp Thr Val Pro Arg Val Tyr Tyr Gly Asp Met Tyr Leu Glu Gly Gly  
 835 840 845  
 Gln Tyr Met Glu Lys Gly Thr Ile Tyr Asn Pro Val Ile Ser Ala Leu  
 850 855 860  
 Leu Lys Ala Arg Ile Lys Tyr Val Ser Gly Gly Gln Thr Met Ala Thr  
 865 870 875 880  
 Asp Ser Ser Gly Lys Asp Leu Lys Asp Gly Glu Thr Asp Leu Leu Thr  
 885 890 895

ES 2 438 001 T3

Ser Val Arg Phe Gly Lys Gly Ile Met Thr Ser Asp Gln Thr Thr Thr  
 900 905 910

Gln Asp Asn Ser Gln Asp Tyr Lys Asn Gln Gly Ile Gly Val Ile Val  
 915 920 925

Gly Asn Asn Pro Asp Leu Lys Leu Asn Asn Asp Lys Thr Ile Thr Leu  
 930 935 940

His Met Gly Lys Ala His Lys Asn Gln Leu Tyr Arg Ala Leu Val Leu  
 945 950 955 960

Ser Asn Asp Ser Gly Ile Asp Val Tyr Asp Ser Asp Asp Lys Ala Pro  
 965 970 975

Thr Leu Arg Thr Asn Asp Asn Gly Asp Leu Ile Phe His Lys Thr Asn  
 980 985 990

Thr Phe Val Lys Gln Asp Gly Thr Ile Ile Asn Tyr Glu Met Lys Gly  
 995 1000 1005

Ser Leu Asn Ala Leu Ile Ser Gly Tyr Leu Gly Val Trp Val Pro  
 1010 1015 1020

Val Gly Ala Ser Asp Ser Gln Asp Ala Arg Thr Val Ala Thr Glu  
 1025 1030 1035

Ser Ser Ser Ser Asn Asp Gly Ser Val Phe His Ser Asn Ala Ala  
 1040 1045 1050

Leu Asp Ser Asn Val Ile Tyr Glu Gly Phe Ser Asn Phe Gln Ala  
 1055 1060 1065

Met Pro Thr Ser Pro Glu Gln Ser Thr Asn Val Val Ile Ala Thr  
 1070 1075 1080

Lys Ala Asn Leu Phe Lys Glu Leu Gly Ile Thr Ser Phe Glu Leu  
 1085 1090 1095

Ala Pro Gln Tyr Arg Ser Ser Gly Asp Thr Asn Tyr Gly Gly Met  
 1100 1105 1110

Ser Phe Leu Asp Ser Phe Leu Asn Asn Gly Tyr Ala Phe Thr Asp  
 1115 1120 1125

Arg Tyr Asp Leu Gly Phe Asn Lys Ala Asp Gly Asn Pro Asn Pro

ES 2 438 001 T3

1130		1135		1140
Thr Lys Tyr Gly Thr Asp Gln Asp Leu Arg Asn Ala Ile Glu Ala 1145 1150 1155				
Leu His Lys Asn Gly Met Gln Ala Ile Ala Asp Trp Val Pro Asp 1160 1165 1170				
Gln Ile Tyr Ala Leu Pro Gly Lys Glu Val Val Thr Ala Thr Arg 1175 1180 1185				
Val Asp Glu Arg Gly Asn Gln Leu Lys Asp Thr Asp Phe Val Asn 1190 1195 1200				
Leu Leu Tyr Val Ala Asn Thr Lys Ser Ser Gly Val Asp Tyr Gln 1205 1210 1215				
Ala Lys Tyr Gly Gly Glu Phe Leu Asp Lys Leu Arg Glu Glu Tyr 1220 1225 1230				
Pro Ser Leu Phe Lys Gln Asn Gln Val Ser Thr Gly Gln Pro Ile 1235 1240 1245				
Asp Ala Ser Thr Lys Ile Lys Gln Trp Ser Ala Lys Tyr Met Asn 1250 1255 1260				
Gly Thr Asn Ile Leu His Arg Gly Ala Tyr Tyr Val Leu Lys Asp 1265 1270 1275				
Trp Ala Thr Asn Gln Tyr Phe Asn Ile Ala Lys Thr Asn Glu Val 1280 1285 1290				
Phe Leu Pro Leu Gln Leu Gln Asn Lys Asp Ala Gln Thr Gly Phe 1295 1300 1305				
Ile Ser Asp Ala Ser Gly Val Lys Tyr Tyr Ser Ile Ser Gly Tyr 1310 1315 1320				
Gln Ala Lys Asp Thr Phe Ile Glu Asp Gly Asn Gly Asn Trp Tyr 1325 1330 1335				
Tyr Phe Asp Lys Asp Gly Tyr Met Val Arg Ser Gln Gln Gly Glu 1340 1345 1350				
Asn Pro Ile Arg Thr Val Glu Thr Ser Val Asn Thr Arg Asn Gly 1355 1360 1365				

ES 2 438 001 T3

Asn Tyr Tyr Phe Met Pro Asn Gly Val Glu Leu Arg Lys Gly Phe  
 1370 1375 1380

Gly Thr Asp Asn Ser Gly Asn Val Tyr Tyr Phe Asp Asp Gln Gly  
 1385 1390 1395

Lys Met Val Arg Asp Lys Tyr Ile Asn Asp Asp Ala Asn Asn Phe  
 1400 1405 1410

Tyr His Leu Asn Val Asp Gly Thr Met Ser Arg Gly Leu Phe Lys  
 1415 1420 1425

Phe Asp Ser Asp Thr Leu Gln Tyr Phe Ala Ser Asn Gly Val Gln  
 1430 1435 1440

Ile Lys Asp Ser Tyr Ala Lys Asp Ser Lys Gly Asn Lys Tyr Tyr  
 1445 1450 1455

Phe Asp Ser Ala Thr Gly Asn Asn Asp Thr Gly Lys Ala Gln Thr  
 1460 1465 1470

Trp Asp Gly Asn Gly Tyr Tyr Ile Thr Ile Asp Ser Asp Ala Asn  
 1475 1480 1485

Asn Thr Ile Gly Val Asn Thr Asp Tyr Thr Ala Tyr Ile Thr Ser  
 1490 1495 1500

Ser Leu Arg Glu Asp Gly Leu Phe Ala Asn Ala Pro Tyr Gly Val  
 1505 1510 1515

Val Thr Lys Asp Gln Asn Gly Asn Asp Leu Lys Trp Gln Tyr Ile  
 1520 1525 1530

Asn His Thr Lys Gln Tyr Glu Gly Gln Gln Val Gln Val Thr Arg  
 1535 1540 1545

Gln Tyr Thr Asp Ser Lys Gly Val Ser Trp Asn Leu Ile Thr Phe  
 1550 1555 1560

Ala Gly Gly Asp Leu Gln Gly Gln Arg Leu Trp Val Asp Ser Arg  
 1565 1570 1575

Ala Leu Thr Met Thr Pro Phe Lys Thr Met Asn Gln Ile Ser Phe  
 1580 1585 1590

ES 2 438 001 T3

Ile Ser Tyr Ala Asn Arg Asn Asp Gly Leu Phe Leu Asn Ala Pro  
 1595 1600 1605

Tyr Gln Val Lys Gly Tyr Gln Leu Ala Gly Met Ser Asn Gln Tyr  
 1610 1615 1620

Lys Gly Gln Gln Val Thr Ile Ala Gly Val Ala Asn Val Ser Gly  
 1625 1630 1635

Lys Asp Trp Ser Leu Ile Ser Phe Asn Gly Thr Gln Tyr Trp Ile  
 1640 1645 1650

Asp Ser Gln Ala Leu Asn Thr Asn Phe Thr His Asp Met Asn Gln  
 1655 1660 1665

Lys Val Phe Val Asn Thr Thr Ser Asn Leu Asp Gly Leu Phe Leu  
 1670 1675 1680

Asn Ala Pro Tyr Arg Gln Pro Gly Tyr Lys Leu Ala Gly Leu Ala  
 1685 1690 1695

Lys Asn Tyr Asn Asn Gln Thr Val Thr Val Ser Gln Gln Tyr Phe  
 1700 1705 1710

Asp Asp Gln Gly Thr Val Trp Ser Gln Val Val Leu Gly Gly Gln  
 1715 1720 1725

Thr Val Trp Val Asp Asn His Ala Leu Ala Gln Met Gln Val Ser  
 1730 1735 1740

Asp Thr Asp Gln Gln Leu Tyr Val Asn Ser Asn Gly Arg Asn Asp  
 1745 1750 1755

Gly Leu Phe Leu Asn Ala Pro Tyr Arg Gly Gln Gly Ser Gln Leu  
 1760 1765 1770

Ile Gly Met Thr Ala Asp Tyr Asn Gly Gln His Val Gln Val Thr  
 1775 1780 1785

Lys Gln Gly Gln Asp Ala Tyr Gly Ala Gln Trp Arg Leu Ile Thr  
 1790 1795 1800

Leu Asn Asn Gln Gln Val Trp Val Asp Ser Arg Ala Leu Ser Thr  
 1805 1810 1815

ES 2 438 001 T3

Thr Ile Met Gln Ala Met Asn Asp Asn Met Tyr Val Asn Ser Ser  
 1820 1825 1830

Gln Arg Thr Asp Gly Leu Trp Leu Asn Ala Pro Tyr Thr Met Ser  
 1835 1840 1845

Gly Ala Lys Trp Ala Gly Asp Thr Arg Ser Ala Asn Gly Arg Tyr  
 1850 1855 1860

Val His Ile Ser Lys Ala Tyr Ser Asn Glu Val Gly Asn Thr Tyr  
 1865 1870 1875

Tyr Leu Thr Asn Leu Asn Gly Gln Ser Thr Trp Ile Asp Lys Arg  
 1880 1885 1890

Ala Phe Thr Val Thr Phe Asp Gln Val Val Ala Leu Asn Ala Thr  
 1895 1900 1905

Ile Val Ala Arg Gln Arg Pro Asp Gly Met Phe Lys Thr Ala Pro  
 1910 1915 1920

Tyr Gly Glu Ala Gly Ala Gln Phe Val Asp Tyr Val Thr Asn Tyr  
 1925 1930 1935

Asn Gln Gln Thr Val Pro Val Thr Lys Gln His Ser Asp Ala Gln  
 1940 1945 1950

Gly Asn Gln Trp Tyr Leu Ala Thr Val Asn Gly Thr Gln Tyr Trp  
 1955 1960 1965

Ile Asp Gln Arg Ser Phe Ser Pro Val Val Thr Lys Val Val Asp  
 1970 1975 1980

Tyr Gln Ala Lys Ile Val Pro Arg Thr Thr Arg Asp Gly Val Phe  
 1985 1990 1995

Ser Gly Ala Pro Tyr Gly Glu Val Asn Ala Lys Leu Val Asn Met  
 2000 2005 2010

Ala Thr Ala Tyr Gln Asn Gln Val Val His Ala Thr Gly Glu Tyr  
 2015- 2020 - 2025

Thr Asn Ala Ser Gly Ile Thr Trp Ser Gln Phe Ala Leu Ser Gly  
 2030 2035 2040

Gln Glu Asp Lys Leu Trp Ile Asp Lys Arg Ala Leu Gln Ala

2045

2050

2055

ES 2 438 001 T3

<210>3  
 <211> 6174  
 <212> ADN  
 <213> Artificial

5

<220>  
 <223> secuencia sintética que codifica una proteína con la actividad de una alternano sacarasa

10 <400> 3

```

atgaaacaac aagaacagc taccgtaaa aaactttata aatccggtaa ggtttggggtt      60
gcagcagcta ctgcatttgc ggtattgggg gtttcaactg taacaacagc ccatgcggac      120
accaactcca acgtggccgt gaagcagatc aacaacaccg gcaccaacga ctccggcgag      180
aagaaggtgc cagtgccatc caccaacaac gactccctca agcagggcac cgaocgcttc      240
tggtacgact ccgacggcaa ccgcgtggac cagaagacca accagatcct cctcaccgcc      300
gagcagctca agaagaacaa cgagaagaac ctctccgtga tctccgacga cacctccaag      360
aaggacgacg agaacatctc caagcagacc aagatcgcca accagcagac cgtggacacc      420
gccaaagggc tcaccacctc caacctctcc gaccocatca ccggtggcca ctacgagaac      480
cacaacggct acttcgtgta catcgacgcc tccggcaagc aggtgaccgg cctccagaac      540
atcgacggca acctccagta ctccgacgac aacggctacc aggtgaaggg ctctctccgc      600
gacgtgaacg gcaagcacat ctacttcgac tccgtgaccg gcaaggcctc ctccaacgtg      660
gacatcgtga acggcaaggc ccagggctac gacgcccagg gcaaccagct caagaagtcc      720
tacgtggccg actcctccgg ccagacctac tacttcgacg gcaacggcca gccactcatc      780
ggcctccaga ccatcgacgg caacctccag tacttcaacc agcagggcgt gcagatcaag      840
ggtggcttcc aggacgtgaa caacaagegc atctacttcg ctccaacac cggcaacgcc      900
gtggccaaca ccgagatcat caacggcaag ctccagggca gggacgcca cggcaaccag      960
gtgaagaacg ccttctccaa ggacgtggtt ggcaaacact tctacttcga cgccaacggc     1020
gtgatgctca ccggcctcca gaccatctcc ggcaagacct actacctcga cgagcagggc     1080
cacctccgca agaactacgc cggcaccttc aacaaccagt tcatgtactt cgacgccgac     1140
actggcgctg gcaagaccgc catcgagtac cagttcgacc agggcctcgt gtcccagtcc     1200
aacgagaaca cccacacaaa cgctgccaag tcttacgaca agtcctcctt cgagaacgtg     1260
gacggttacc tcaccgccga cacctggtac cgccaaccg acatcctcaa gaacggcgac     1320
acctggaccg cctccaccga gaccgacatg aggccactcc tcatgacctg gtggccagac     1380
aaqcagaccc aggccaaacta cctcaacttc atgtcctcca agggcctcgg catcaccacc     1440
    
```

ES 2 438 001 T3

acctacaccg ctgccacctc ccagaagacc ctcaacgacg ctgccttcgt gatccagacc 1500  
 gccatcgagc agcagatctc cctcaagaag tccaccgagt ggctccgcga cgccatcgac 1560  
 tccttcgtga agaccacaggc caactggaac aagcagaccg aggacgaggc cttecgacggc 1620  
 ctccagtggc tccaggtggc ctctctcgcc taccaggacg actcccaccg caccctaac 1680  
 accgactccg gcaacaaccg caagctcggc aggcagccca tcaacatcga cggctccaag 1740  
 gacaccaccg acggcaaggc ctccgagttc ctctctgcca acgacatcga caactccaac 1800  
 ccgatcgtgc aggccgagca gctcaactgg ctccactacc tcatgaactt cggctccatc 1860  
 accggcaaca acgacaacgc caacttcgac ggcatccgcg tggacgccgt ggacaacgtg 1920  
 gacgccgacc tcctcaagat cgctggcgac tacttcaagg ccctctacgg caccgacaag 1980  
 tccgacgcca acgccaacaa gcacctctcc atcctcgagg actggaacgg caaggacca 2040  
 cagtacgtga accagcaggc caacgccagc ctactatgg actacaccgt gacctcccag 2100  
 ttccgcaact cctcaccga cggcgccaac aaccgctcca acatgtggta ctctctgac 2160  
 accggctact acctcaacgg cgacctcaac aagaagatcg tggacaagaa caggcctaac 2220  
 tccggcacc cctgtaaccg catcgccaac tccggcgaca ccaaggtgat cccaactac 2280  
 tccttcgtga gggcccacga ctacgacgcc caggaccgga tccgcaaggc catgatcgac 2340  
 cacggcatca tcaagaacat gcaggacacc ttcaccttcg accagctcgc ccagggcatg 2400  
 gagttctact acaaggacca ggagaaccgc tccggcttca agaagtaca cgactacaac 2460  
 ctcccattcg cctacgccat gctcctcacc aacaaggaca ccgtgccagc ggtgtactac 2520  
 ggcgacatgt acctcgaggc tggccagtac atggagaagg gcaccatcta caaccgggtg 2580  
 atctccgcc tcctcaaggc ccgcatcaag tacgtgtccg gtggccagac tatggccacc 2640  
 gactcctccg gcaaggacct caaggacggc gagaccgacc tcctcacctc cgtgcgcttc 2700  
 ggcaaggga tcatgacctc cgaccagacc accaccagc acaactcca ggactacaag 2760  
 aaccaggga tcggcgtgat cgtgggcaac aaccagacc tcaagctca caacgacaag 2820  
 accatcacc tcacatggg caaggcccac aagaaccagc tctacaggc cctcgtgctc 2880  
 tccaacgact ccggcatcga cgtgtacgac tccgacgaca aggccccgac cctccgcacc 2940  
 aacgacaacg ggcacctcat ctccacaag accaacacct tcgtgaagca ggacggcacc 3000  
 atcatcaact acgagatgaa gggctccctc aacgccctca tctccggcta cctcggcgtg 3060  
 tgggtgccag tgggcgcctc cgactcccag gacgcccgca ccgtggccac cgagtccctc 3120  
 tcctccaacg acggctccgt gttccactcc aacgctgccc tcgactcca cgtgatctac 3180  
 gagggcttct ccaacttcca ggccatgccc acctccccgg agcagtcac caacgtggtg 3240  
 atcgccacca aggccaacct ctcaaggag ctgggcatca cctccttcga gctggcccca 3300

ES 2 438 001 T3

cagtaccgct cctccggcga caccaactac ggtggcatgt ccttcctcga ctcttcctc 3360  
 aacaacggct acgccttcac cgaccgctac gacctcggct tcaacaaggc cgacggcaac 3420  
 ccgaacccaa ctaagtacgg caccgaccag gacctccgca acgccatcga ggccctccac 3480  
 aagaacggca tgcaggccat cgccgactgg gtgccggacc agatctacgc cctcccaggc 3540  
 aaggaggtgg tgaccgccac ccgcgtaggac gagcgcggca accagctcaa ggacaccgac 3600  
 ttcgtgaacc tcctctacgt ggccaacacc aagtccctcg gcgtggacta ccaggccaag 3660  
 tacggtggcg agttcctcga caagctccgc gaggagtacc catccctctt caagcagaac 3720  
 caggtgtcca ccggccagcc tatcgacgcc tccaccaaga tcaagcagtg gtccgccaag 3780  
 tacatgaacg gcaccaacat cctccacagg ggtgcctact acgtgctcaa ggactgggcc 3840  
 accaaccagt acttcaacat cgccaagacc aacgaggtgt tcctcccact ccagctccag 3900  
 aacaaggacg ccagaccgg cttcatctcc gacgcctccg gcgtgaagta ctactccatc 3960  
 tccggctacc aggccaagga caocttcato gaggacggca acggcaactg gtactacttc 4020  
 gacaaggacg gctacatggt gcgctcccag cagggcgaga acccgatccg caccgtggag 4080  
 acctccgtga acaccgcaa cggcaactac tacttcatgc ctaacggcgt ggagctccgc 4140  
 aagggtctcg gcaccgaaa ctccggcaac gtgtactact tcgacgacca gggcaagatg 4200  
 gtgcgcgaca agtacatcaa cgacgacgcc aacaactct accacctcaa cgtggacggc 4260  
 accatgtcca gggcctctt caagttcgac tccgacacc tccagtactt cgctccaac 4320  
 ggctgcaga tcaaggactc ctacgccaa gactccaagg gcaacaagta ctacttcgac 4380  
 tccgccaccg gcaacaacga caccggcaag gccagacct gggacggcaa cggctactac 4440  
 atcaccatcg actccgacgc caacaacacc atcggcgtga acaccgacta caccgcctac 4500  
 atcaacctct ccctccgca ggacggcctc ttgcocaaag cccctacgg cgtggtgacc 4560  
 aaggaccaga acggcaacga cctcaagtgg cagtacatca accacaccaa gcagtacgag 4620  
 ggccagcagg tgcagggtac ccgccagtac accgactcca agggcgtgtc ctggaacctc 4680  
 atcaccttcg ccggtggcga cctgcagggc cagcgcctct ggggtggactc cagggccctc 4740  
 accatgaccc cttcaagac catgaaccag atctccttea tctcctacgc caaccgcaac 4800  
 gacggcctct tcctcaacgc cccataccag gtgaagggt accagctcgc cggcatgtcc 4860  
 aaccagtaca agggccagca ggtgaccatc gctggcgtgg ccaacgtgtc cggcaaggac 4920  
 tggccctca tctcctcaa cggcaccag tactggatcg actcccaggc cctcaacacc 4980  
 aacttcaacc acgacatgaa ccagaaggtg ttcgtgaaca ccacctcaa cctcgacggc 5040  
 ctcttcctca acgctcccta ccgccagccg ggctacaagc tcgctggcct cgccaagaac 5100

ES 2 438 001 T3

tacaacaacc agaccgtgac cgtgtcccag cagtacttcg acgaccaggg caccgtgtgg 5160  
 tcccaggtgg tgctcggtag ccagaccgtg tgggtggaca accacgccct cgcccagatg 5220  
 caggtgtccg acaccgacca gcagctctac gtgaactcca acggtcgcaa cgacggcctc 5280  
 ttctcaacg ctccgtacag gggccagggc tcccagctca tcggcatgac cgccgactac 5340  
 aacggccagc acgtgcaggt gaccaagcag ggccaggacg cctacgggtg ccagtggcgc 5400  
 ctcatcacc tcaacaacca gcaggtgtgg gtggactcca gggctctctc caccaccatc 5460  
 atgcaggcca tgaacgacaa catgtacgtg aactcctccc agcgcaccga cggcctctgg 5520  
 ctcaacgctc cctacaccat gtctggcgcc aagtgggctg gcgacacccg ctccgccaac 5580  
 ggtaggtacg tgcacatctc caaggcctac tccaacgagg tgggcaaac ctactacctc 5640  
 accaacctca acggccagtc cacctggatc gacaagcgcg ctttcaccgt gaccttcgac 5700  
 caggtggtag ccctcaacgc caccatcgtg gccaggcaga ggccagacgg catgttcaag 5760  
 accgctccct acggcgaggc cggcgcccag ttctgtggact acgtgaccaa ctacaaccag 5820  
 cagaccgtgc cggtgaccaa gcagcactcc gacgcccagg gcaaccagtg gtacttggcc 5880  
 accgtgaacg gcacccagta ctggatcgac cagcgtcct tctccccggg ggtgaccaag 5940  
 gtgtgggact accaggccaa gatcgtgcca aggaccaccc gcgacggcgt gttctccggg 6000  
 gccccgtacg gcgaggtgaa cgccaagctc gtgaacatgg ccaccgccta ccagaaccag 6060  
 gtgtgtgcacg ccaccggcga gtacaccaac gcttcggca tcacctggtc ccagttcgcc 6120  
 ctctccggcc aggaggacaa gctctggatc gacaagcgcg ctctccaagc ttga 6174

5 <210>4  
 <211> 43  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Cebador  
 <400> 4

gaaagcttg ctgcaggtca gtccttatg ttacgtcctg tag 43

15 <210>5  
 <211> 33  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Cebador  
 <400> 5

25 ctggtcccgg gattcattgt ttgctcct gct 33

30 <210>6  
 <211> 28  
 <212> ADN  
 <213> Artificial

ES 2 438 001 T3

<220>  
<223> Cebador

<400> 6

5 tcgagctcgc gaaagcttgg ctgcaggt 28

<210>7  
<211> 65  
10 <212> ADN  
<213> Artificial

<220>  
<223> secuencia de enlazador sintético

15 <400> 7

aacaaactctc ctggcgcacc atcgctcggct acagcctcgg gaattgctgc aggtcgcacgg 60  
atccg 65

<210>8  
<211> 69  
20 <212> ADN  
<213> Artificial

<220>  
<223> secuencia de enlazador sintético

25 <400> 8

aattcggatc cgtcgcacctg cagcaattcc cgaggctgta gccgacgatg gtgcgccagg 60  
agagttggtt 69

<210>9  
<211> 45  
35 <212> ADN  
<213> Artificial

<220>  
<223> Cebador

40 <400> 9

cggatcggat ccgtcgacct gcagatcgtt caaacattg gcaat 45

<210> 10  
45 <211> 37  
<212> ADN  
<213> Artificial

<220>  
50 <223> Cebador

<400> 10

ccacggaatt cgatctagta acatagatga caccgcg 37

55 <210> 11  
<211> 18  
<212> ADN  
60 <213> Artificial

ES 2 438 001 T3

<220>  
<223> oligonucleótido sintético

<400> 11

5 gcggccgagg cagccatg 18

<210> 12  
<211> 18  
10 <212> ADN  
<213> Artificial

<220>  
<223> oligonucleótido sintético

15 <400> 12

accggatgc gggccgc 18

20 <210> 13  
<211> 25  
<212> ADN  
<213> Artificial

25 <220>  
<223> Cebador

<400> 13

30 agcttgaacc ggatcgcg gccgc 25

<210> 14  
<211> 25  
35 <212> ADN  
<213> Artificial

<220>  
<223> Cebador

40 <400> 14

agctggcc gcatcgc gttca 25

45 <210> 15  
<211> 53  
<212> ADN  
<213> Artificial

<220>  
50 <223> Cebador

<400> 15

caactgagct ctgcagaatt cggctgttc gggcagccat ggacaccaac tcc 53

55 <210> 16  
<211> 30  
<212> ADN  
<213> Artificial

60 <220>  
<223> Cebador

<400> 16

65 gccggagtg tcggtgccga agcccttgcg 30

<210> 17  
 <211> 30  
 <212> ADN  
 5 <213> Artificial  
  
 <220>  
 <223> Cebador  
 <400> 17  
 10 tcctacgaca agtctcctt cgagaacgtg 30  
  
 <210> 18  
 <211> 38  
 15 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> Cebador  
 20 <400> 18  
  
 ctacggatc cactagtaac ggccgcgata tccggttc 38  
 25 <210> 19  
 <211> 27  
 <212> ADN  
 <213> Artificial  
 30 <220>  
 <223> Cebador  
  
 <400> 19  
 35 gccatcgagc agcagatctc cctcaag 27  
  
 <210> 20  
 <211> 28  
 <212> ADN  
 40 <213> Artificial  
  
 <220>  
 <223> Cebador  
 45 <400> 20  
  
 cataggatcc ctacagccg gaggcgtc 28  
  
 <210>21  
 <211> 40  
 <212> ADN  
 50 <213> Artificial  
  
 <220>  
 <223> Cebador  
 55 <400> 21  
  
 taatagcact cactatagcc agaattcggc ttgtcgggc 40  
 60 <210> 22  
 <211> 28  
 <212> ADN  
 <213> Artificial  
 65 <220>

<223> Cebador

<400> 22

5 agtaacggcc gcgatatccg gtcaagc 28

<210> 23

<211> 26

<212> ADN

10 <213> Artificial

<220>

<223> Cebador

15 <400> 23

cgaagaggcc gtcctagcgg agggag 26

<210> 24

20 <211> 24

<212> ADN

<213> Artificial

<220>

25 <223> Cebador

<400> 24

ggtgccgaag ccctagcgg gctc 24

30

<210> 25

<211> 29

<212> ADN

<213> Artificial

35

<220>

<223> Cebador

<400> 25

40

ccggagatgg agtagtacta cacgccgga 29

**REIVINDICACIONES**

1. Una molécula de ácido nucleico que codifica una proteína con la actividad enzimática de una alternano sacarasa, en la que la molécula de ácido nucleico está seleccionada de entre
  - 5 a) una molécula de ácido nucleico con una secuencia desde el nucleótido de la posición 118 hasta el nucleótido de la posición 3945, inclusive, de la secuencia de ácidos nucleicos mostrada en la SEQ ID NO 1 o la SEQ ID NO 3,
  - b) una molécula de ácido nucleico que codifica una proteína cuya secuencia de aminoácidos tiene al menos el 70 % de identidad con la secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2, en la que dicha proteína tiene el mismo número de aminoácidos que la proteína con la secuencia desde el aminoácido de la posición 40 hasta el aminoácido de la posición 1315, inclusive, de la SEQ ID NO 2,
  - 10 c) una molécula de ácido nucleico cuya secuencia tiene al menos el 70 % de identidad con la secuencia de a), donde dicha secuencia tiene el mismo número de nucleótidos que la secuencia de a),
  - d) una moléculas de ácido nucleico cuya secuencia de ácidos nucleicos, como consecuencia de la degeneración del código genético, se desvía de la secuencia de las moléculas de ácido nucleico de a), b) o c),
  - 15 e) una molécula de ácido nucleico cuya secuencia de ácidos nucleicos es complementaria a la secuencia de longitud completa de una de las moléculas de ácido nucleico de a), b), c) o d), en la que la secuencia complementaria tiene el mismo número de nucleótidos que las moléculas de ácido nucleico de a), b), c) o d).
- 20 2. La molécula de ácido nucleico como se reivindica en la reivindicación 1, que tiene unido en su extremo 5' un triplete de ácido nucleico ATG y en su extremo 3' un triplete de ácido nucleico seleccionado de entre TAA, TAG y TGA.
3. Un vector que comprende una molécula de ácido nucleico como se reivindica en la reivindicación 1 o 2, en el que la molécula de ácido nucleico codifica una proteína como se define en la reivindicación 1b.
- 25 4. El vector como se reivindica en la reivindicación 3, que es un plásmido.
5. Una célula huésped transformada con una molécula de ácido nucleico como se reivindica en la reivindicación 1 o 2, un vector como se reivindica en la reivindicación 3 o un plásmido como se reivindica en la reivindicación 4 o una célula huésped que deriva de dicha célula transformada y que comprende una molécula de ácido nucleico como se reivindica en la reivindicación 1 o 2, o un vector como se reivindica en la reivindicación 3.
- 30 6. La célula huésped como se reivindica en la reivindicación 5, que es la célula de un microorganismo.
7. La célula huésped como se reivindica en la reivindicación 6, que es una célula de E. coli.
8. La célula huésped como se reivindica en la reivindicación 5, que es una célula vegetal, que comprende una molécula de ácido nucleico como se reivindica en la reivindicación 1 o 2, o un vector como se reivindica en la reivindicación 3.
- 35 9. Una planta que comprende células vegetales como se reivindica en la reivindicación 8 que comprenden una molécula de ácido nucleico como se reivindica en la reivindicación 1 o 2, o un vector como se reivindica en la reivindicación 3.
10. La planta como se reivindica en la reivindicación 9, que es una planta de almacenamiento de azúcar o almidón.
11. La planta como se reivindica en la reivindicación 9 o 10, que es patata, maíz, remolacha azucarera, sorgo azucarero o caña de azúcar.
- 40 12. Un material de propagación de una planta como se reivindica en cualquiera de las reivindicaciones 9-11, que comprende células vegetales como se reivindica en la reivindicación 8.
13. Una parte de una planta cosechable obtenida a partir de una planta como se reivindica en cualquiera de las reivindicaciones 9-11, que comprende células vegetales como se reivindica en la reivindicación 8.
- 45 14. Un procedimiento para generar una planta como se reivindica en cualquiera de las reivindicaciones 9-11, que comprende las etapas siguientes:
  - a) introducir en una célula vegetal una molécula de ácido nucleico exógena como se reivindica en la reivindicación 1 o 2 o un vector como se reivindica en la reivindicación 3 o 4 por transformación,
  - b) regenerar una planta a partir de células vegetales obtenidas en la etapa a), y

- c) en caso apropiado, generar más plantas con ayuda de las plantas obtenidas en la etapa b).
15. Una proteína con la actividad enzimática de una alternano sacarasa, proteína que está seleccionada de entre
- 5 a) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 200 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1295 hasta el aminoácido de la posición 1345 de la SEQ ID NO 2,
- b) una proteína con una secuencia de aminoácidos que tiene al menos el 70 % de identidad con una secuencia de aminoácidos de la proteína de a), en la que dicha proteína tiene el mismo número de aminoácidos que la proteína de a).
16. La proteína como se reivindica en la reivindicación 15, que tiene una metionina unida a su extremo N.
- 10 17. Un procedimiento de producción de una proteína con la actividad enzimática de una alternano sacarasa como se reivindica en la reivindicación 15 o 16, en el que
- 1) se transforma una célula huésped con
- 15 a) una molécula de ácido nucleico cuya secuencia codifica la secuencia de aminoácidos de la proteína como se reivindica en la reivindicación 15 o 16, en el que la molécula de ácido nucleico se puede obtener a partir de una molécula de ácido nucleico con una secuencia mostrada en la SEQ ID NO 1 o 3, o
- b) un vector o plásmido que comprende una molécula de ácido nucleico como se describe en a),
- 2) se cultiva la célula huésped en medio de cultivo y
- 3) se aísla la proteína a partir de las células cultivadas y/o el medio de cultivo.
18. Un procedimiento de producción de una proteína con la actividad enzimática de una alternano sacarasa como se reivindica en la reivindicación 15 o 16, en el que
- 20 1) se transforma una célula vegetal con
- a) una molécula de ácido nucleico cuya secuencia codifica la secuencia de aminoácidos de la proteína como se reivindica en la reivindicación 15 o 16, en el que la molécula de ácido nucleico se puede obtener a partir de una molécula de ácido nucleico con una secuencia mostrada en la SEQ ID NO 1 o 3, o
- 25 b) un vector o plásmido que comprende una molécula de ácido nucleico como se describe en a),
- 2) se regenera una planta a partir de las células vegetales obtenidas en la etapa 1),
- 3) se expresa la proteína en la planta y posteriormente se extrae y se aísla.
19. Un procedimiento de producción de alternano con un promedio en peso del peso molecular Pm en el intervalo de desde 12 000 000 hasta 30 000 000 g/mol, en el que
- 30 1) se pone en contacto una solución que comprende sacarosa con una proteína con la actividad enzimática de una alternano sacarasa, y
- 2) se aísla alternano a partir de la solución, en el que la proteína es
- 35 a) una proteína con una secuencia de aminoácidos que comienza en la región desde el aminoácido de la posición 1 hasta el aminoácido de la posición 200 de la SEQ ID NO 2 y que termina en la región desde el aminoácido de la posición 1295 hasta el aminoácido de la posición 1345 de la SEQ ID NO 2,
- b) una proteína con una secuencia de aminoácidos que tiene al menos el 70 % de identidad con una secuencia de aminoácidos de la proteína de a), en la que dicha proteína tiene el mismo número de aminoácidos que la proteína de a).
20. El procedimiento como se reivindica en la reivindicación 19, en el que proteína tiene una metionina unida a su extremo N.
- 40 21. El procedimiento como se reivindica en cualquiera de las reivindicaciones 19 o 20, en el que la solución que comprende sacarosa es una solución acuosa con una concentración en sacarosa de al menos el 15 % en base al peso de agua.
22. El procedimiento como se reivindica en cualquiera de las reivindicaciones 19-21, en el que la etapa 1) del
- 45 procedimiento se lleva a cabo a una temperatura de 35-45 °C.

23. Un procedimiento de producción de alternano, que comprende las etapas de extraer y aislar el alternano a partir de una célula vegetal como se reivindica en la reivindicación 8 de una planta como se reivindica en cualquiera de las reivindicaciones 9-11, a partir de material de propagación como se reivindica en la reivindicación 12 o a partir de partes de plantas cosechables como se reivindica en la reivindicación 13.

Figura 1

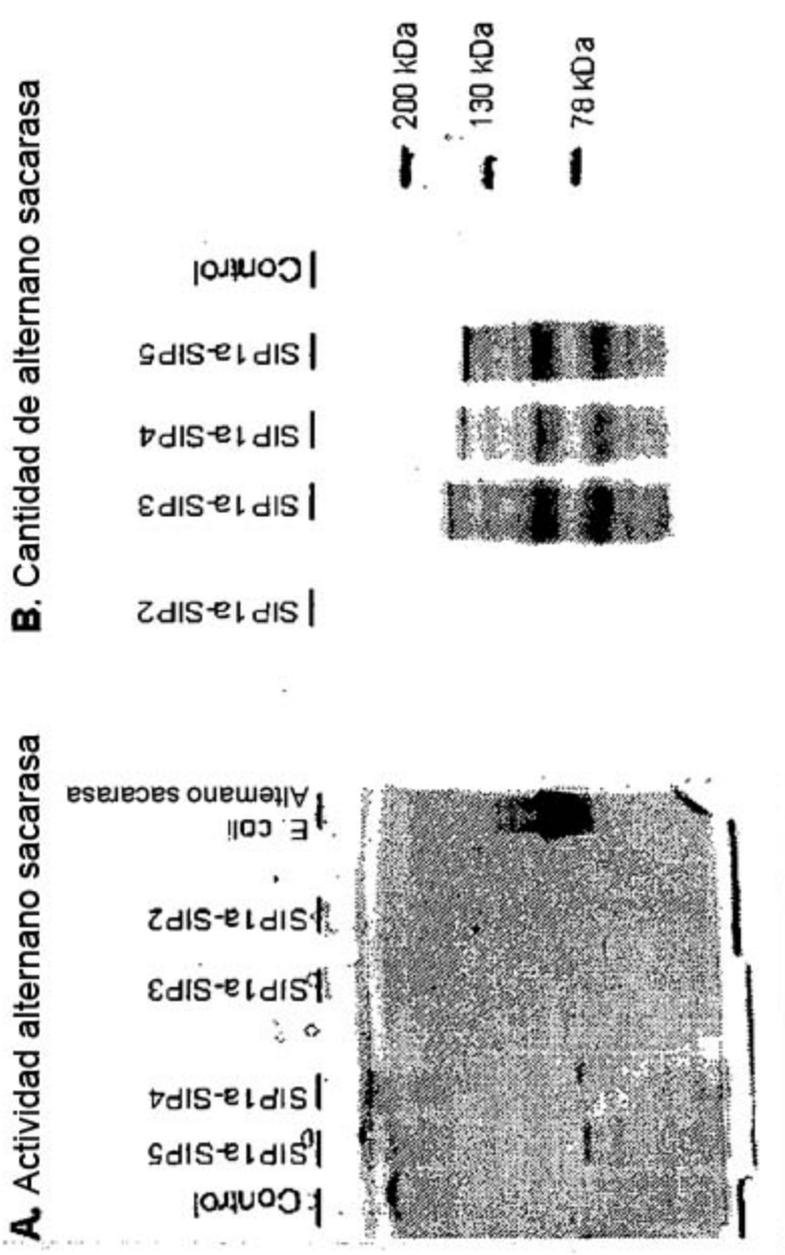


Figura 2

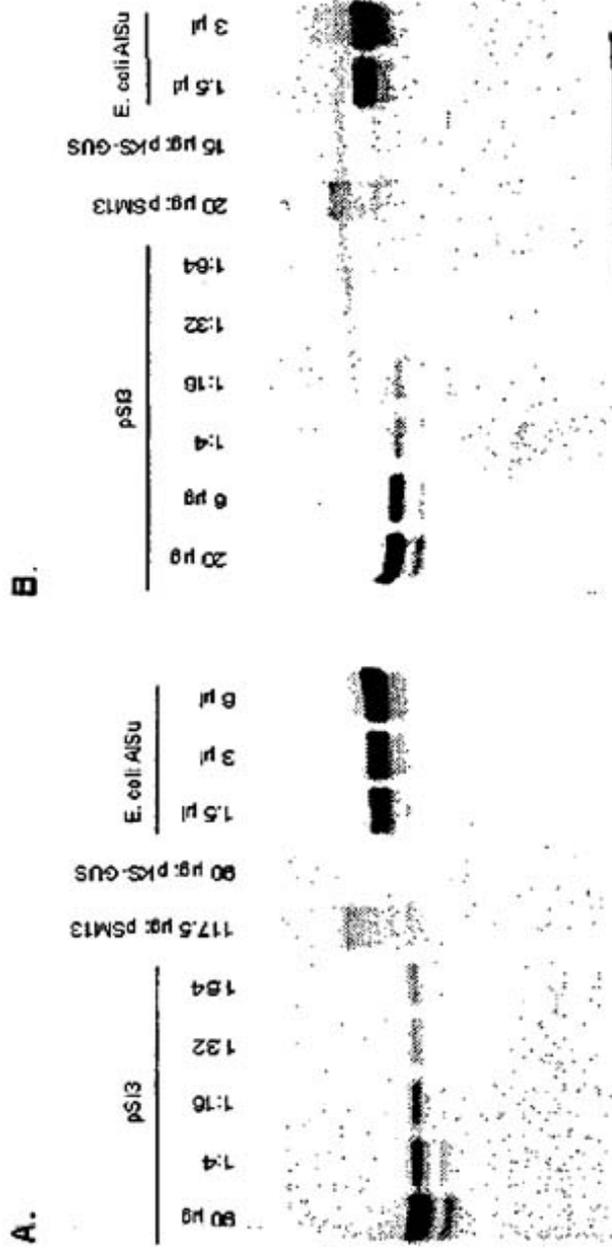
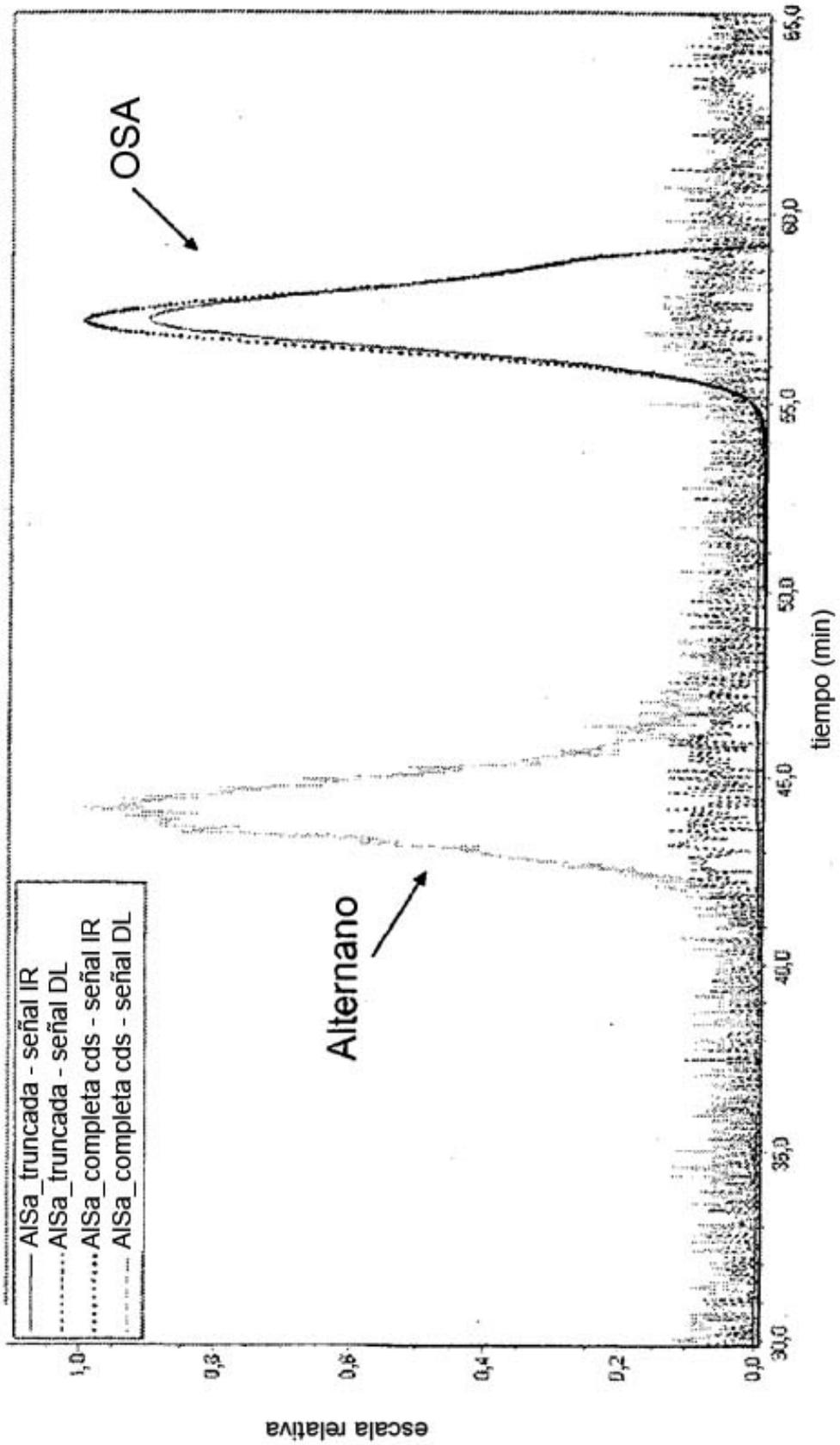


Figura 3

Biotransformación con maltosa al 10 %



**Figura 4**  
Biotransformación con isomaltulosa al 10 %

