



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 342 125**

51 Int. Cl.:

**C12N 5/10** (2006.01)

**C12N 15/63** (2006.01)

**C12N 9/38** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07732142 .0**

96 Fecha de presentación : **27.03.2007**

97 Número de publicación de la solicitud: **1999254**

97 Fecha de publicación de la solicitud: **10.12.2008**

54 Título:  **$\beta$ -galactosidasa con actividad de transgalactosilación.**

30 Prioridad: **28.03.2006 GB 0606112**

45 Fecha de publicación de la mención BOPI:  
**01.07.2010**

45 Fecha de la publicación del folleto de la patente:  
**01.07.2010**

73 Titular/es: **Clasado Inc.**  
**Edificio Interseco, Piso 2**  
**Calle Elvira Mendez, Apartado 0816-01560**  
**Panama, PA**

72 Inventor/es: **Tzortzis, Georgios;**  
**Goulas, Athanasios, K. y**  
**Goulas, Theodoros**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 342 125 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

$\beta$ -galactosidasa con actividad de transgalactosilación.

5 La presente invención se refiere a una nueva  $\beta$ -galactosidasa con actividad de trans-galactosilación capaz de convertir lactosa en una mezcla de galacto-oligosacáridos. En particular se refiere a una  $\beta$ -galactosidasa aislada a partir de una cepa recientemente descubierta de *Bifidobacterium bifidum*.

10 La invención se refiere particularmente a secuencias de ADN que codifican la nueva enzima  $\beta$ -galactosidasa aislada, a la enzima codificada por dicha secuencia de ADN y a la célula hospedadora que comprende la secuencia de ADN o que contiene un vector recombinante que incorpora la secuencia de ADN. La invención se refiere también al uso de la enzima codificada por una secuencia de ADN, o la célula hospedadora que contiene una secuencia de ADN o el vector recombinante, para producir los galacto-oligosacáridos.

15 La bifidobacterias colonizan de forma natural el tracto intestinal inferior, un medio que es pobre en mono y disacáridos ya que tales azúcares se consumen con preferencia por los huéspedes y microbios presentes en el tracto intestinal superior. Para sobrevivir en el tracto intestinal inferior las bifidobacterias producen varios tipos de exo y endoglicosidasas unidas a la superficie y/o en formas extracelulares, mediante las cuales pueden utilizar diversos hidratos de carbono.

20 Además de la actividad hidrolasa, algunas enzimas de las bifidobacterias muestran actividad transferasa. Esta actividad de trans-galactosilación de las glicosidasas se usa ampliamente para la síntesis enzimática de diversos oligosacáridos, que han demostrado que actúan como factores que promueven el crecimiento de las bifidobacterias.

25 Se sabe que algunas bifidobacterias producen enzimas  $\beta$ -galactosidasas que están involucradas en el metabolismo bacteriano de la lactosa. En Møller, P. L. *et al* in *Appl & Environ. Microbial.*, (2001), 62, (5), 2276-2283 se describe el aislamiento y caracterización de tres genes de  $\beta$ -galactosidasa de una cepa de *Bifidobacterium bifidum*. Se encontró que los tres tipos de  $\beta$ -galactosidasas eran capaces de catalizar la formación de beta galacto-oligosacáridos por transgalactosilación.

30 En Durmortier *et al* en *Carbohydrate Research*, 201, (1990), 115-123 se describió la formación de beta oligosacáridos por una reacción de trans-galactosilación durante la hidrólisis de lactosa con *Bifidobacterium bifidum* DSM 20456. Su análisis de la estructura de la mezcla de los oligosacáridos producidos mostró que las uniones eran enlaces  $\beta$ -(1→3),  $\beta$ -(1→6) y  $\beta$ -(1→4)-D-galactosil. Durmortier sugirió que los compuestos producidos por *Bifidobacterium bifidum* están involucrados en la adherencia de la bacteria en el intestino grueso.

35 La patente WO01/90317 describe una nueva  $\beta$ -galactosidasa de *Bifidobacterium bifidum*, en particular una versión truncada de la enzima que tiene una actividad de trans-galactosilación alta.

40 Se ha encontrado una cepa de *Bifidobacterium bifidum* que es capaz de producir una actividad de enzima galactosidasa que convierte la lactosa en una nueva mezcla de galacto-oligosacáridos que contiene inesperadamente hasta un 35% de disacáridos que incluyen galabiosa (Gal ( $\alpha$  1-6)- Gal). Este disacárido es conocido (véase Paton, J C. Paton, A W (1998), *Clin. Microbiol. Revs.*, 11, 450-479; Carlsson, K A (1989), *Ann. Reviews Biochem.*, 58, 309-350) por ser un antiadhesivo capaz de evitar la adhesión de toxinas como la toxina Shiga y patógenos como *E. coli*, a la pared del

45 La cepa de *B bifidum* fue depositada con el numero de acceso NCIMB 41171 en la Colección Nacional de Bacterias Industriales y Marinas (Nacional Collection of Industrial & and Marine Bacteria), Aberdeen, Reino Unido el 31 de Marzo de 2003. También está descrita en la patente de Reino Unido GB2412380.

50 Actualmente se ha encontrado que esta cepa de *B bifidum* produce varias  $\beta$ -galactosidasas, incluyendo una  $\beta$ -galactosidasa nueva. Esta enzima produce una serie de oligosacáridos diferentes que tienen uniones  $\beta$ .

55 Según la invención se proporciona una secuencia de ADN que codifica una proteína con una secuencia de aminoácidos según la secuencia SEQ.ID NO: 2. La secuencia de ADN viene dada por la SEQ. ID NO: 1 o puede comprender un fragmento o degeneración de la misma.

Por "degeneración" se entiende una secuencia de ADN que presenta de 75 a 95% de homología.

60 Esta secuencia de ADN puede comprender sustituciones, adiciones o deleciones de nucleótidos que den como resultado un cambio en la secuencia de aminoácidos según la SEQ. ID NO: 2 inferior a un 25%. Las sustituciones de nucleótidos pueden producir como resultado sustituciones de aminoácidos conservadoras.

65 Según un segundo aspecto de la invención se proporciona una enzima codificada por una secuencia de ADN definida anteriormente. Tal enzima puede comprender la secuencia de aminoácidos dada en SEQ. ID NO: 2 o un fragmento de la misma.

Según un tercer aspecto de la invención se proporciona un vector recombinante, preferiblemente un vector de expresión, que comprende una secuencia de ADN definida anteriormente. Este vector se puede incorporar en una célula hospedadora, tal como una célula bacteriana, levadura u hongo. Alternativamente, la secuencia de ADN puede ser incorporada en dicha célula hospedadora. Se puede seleccionar una célula hospedadora adecuada del grupo que comprende *Bifidobacterium*, *Lactococcus*, *Lactobacillus*, *Bacillus* por ejemplo *Bacillus subtilis* o *Bacillus circulans*, *Escherichia* y *Aspergillus* por ejemplo *Aspergillus niger*.

Utilizando lactosa como sustrato, la enzima codificada por la secuencia de ADN definida anteriormente produce una mezcla de oligosacáridos, que comprende disacáridos, como Gal ( $\beta$ 1-3) Glc, Gal ( $\beta$ 1-3) Gal, Gal ( $\beta$ 1-6) Gal y Gal ( $\alpha$ 1-6) Gal, trisacáridos y tetrasacáridos como Gal ( $\beta$ 1-6) Gal ( $\beta$ 1-4) Glc, Gal ( $\beta$ 1-3) Gal ( $\beta$ 1-4) Glc y Gal ( $\beta$ 1-6) Gal ( $\beta$ 1-6) Gal ( $\beta$ 1-4) Glc.

La enzima o la célula hospedadora descrita anteriormente puede usarse para producir una mezcla de galacto-oligosacáridos que puede formar parte de un producto para mejorar la salud intestinal. Dicho producto puede seleccionarse del grupo que consiste en productos lácteos (por ejemplo leche líquida, leche en polvo seca tal como la leche entera en polvo, leche desnatada en polvo, leche desnatada enriquecida en grasa en polvo, suero de leche en polvo, leches infantiles, leche maternizada, helados, yogur, queso, productos de leche fermentada), bebidas tales como el zumo de fruta, los alimentos infantiles, cereales, pan, galletas, dulces, pasteles, suplementos alimenticios, suplementos dietéticos, alimentos para animales, alimentos para aves o por supuesto cualquier otro alimento o bebida. La presencia de galacto-oligosacáridos en dichos productos tiene la ventaja de incrementar el crecimiento de las *Bifidobacterium* promotoras de la salud en el producto o en la flora intestinal del consumidor después de la ingesta del producto o en ambos.

Alternativamente, los oligosacáridos producidos así pueden utilizarse para la preparación de un medicamento, por ejemplo en forma de comprimido o de cápsula, para evitar la adhesión de los patógenos o de las toxinas producidas por los patógenos a la pared del intestino. El medicamento puede ser administrado al paciente, por ejemplo después de un tratamiento con antibióticos, que a menudo altera o destruye la flora intestinal normal sana.

Según aún otro aspecto de la invención se proporciona un proceso para producir una enzima tal y como se define anteriormente que comprende cultivar una célula hospedadora definida anteriormente en un medio de cultivo adecuado en condiciones que permitan la expresión de la enzima y recuperar la enzima resultante o los productos enzimáticos del cultivo.

La invención también está dirigida a un proceso para producir la mezcla de galacto-oligosacáridos que comprende poner en contacto la enzima definida anteriormente con un material que contenga lactosa en condiciones que conduzcan a la formación de la mezcla de galacto-oligosacáridos.

El material adecuado que contenga lactosa puede seleccionarse entre lactosa comercial, leche entera, leche semidesnatada, leche desnatada, suero de leche, leche desnatada enriquecida en grasa y suero de leche filtrado. Estos productos lácteos pueden obtenerse de vacas, búfalas, ovejas o cabras. La leche desnatada enriquecida en grasa se define como la leche entera que ha sido desnatada para eliminar la grasa de la leche, la cual se reemplaza posteriormente añadiendo aceite o grasa vegetal.

### Breve descripción de los dibujos

La Figura 1 muestra la secuencia de nucleótidos (SEQ. ID NO: 1) de  $\beta$ -galactosidasa de *Bifidobacterium bifidum* de la invención; y

La Figura 2 muestra la secuencia de aminoácidos (SEQ. ID NO: 2) correspondiente a la secuencia de nucleótidos de la Figura 1.

La Figura 3 es un gráfico que muestra el proceso temporal de la reacción durante la síntesis del galacto-oligosacárido con  $\beta$ -galactosidasa y lactosa al 40% (porcentaje en peso) en tampón fosfato 0,1 M a pH 6,0 como sustrato; y

La Figura 4 muestra un cromatograma de intercambio aniónico de alto rendimiento de la mezcla de galacto-oligosacáridos sintetizada por  $\beta$ -galactosidasa de *B. Bifidum* NCIMB 41171 utilizando lactosa al 40% (porcentaje en peso) en tampón fosfato 0,1 M a pH 6,0 como sustrato (Glc = Glucosa, Gal = Galactosa, Lac = Lactosa,  $\alpha$  (1-6) = galactobiosa, DP = grado de polimerización del inglés "Degree of polymerization").

El ADN genómico fue aislado de la cepa (NCIMB 41171) de *Bifidobacterium bifidum* utilizando el método de Lawson *et al.* (1989) *Fems Microbiol Letters*, 65, (1-2), 41-45. El ADN fue digerido con enzimas de restricción y los fragmentos con un tamaño de 15 kbp máximo se unieron con el vector pSP72 que había sido digerido con las mismas enzimas de restricción. Las células de *E. coli* se transformaron con un vector que contenía las inserciones del ADN cromosómico digerido de *B. bifidum* PstI, EcoRI, Bam HI, KpnI, SmaI o HindIII. Se seleccionaron los clones con actividad  $\beta$ -galactosidasa en placas de agar Luria Bertani que contienen p-nitrofenil, X-  $\beta$ -Gal (5-bromo-4-cloro-indol-3-indolil- $\beta$ -D-galactósido) e isopropil-  $\beta$ -D- tiogalactósido (IPTG). Las mezclas de ligación con el ADN cromosómico Pst I dio lugar a trece clones positivos de  $\beta$ -galactosidasa, uno de los cuales se identificó como pP2.

La secuenciación de ADN del fragmento de ADN insertado P2 se llevó a cabo utilizando el método de Sanger de terminación de cadenas por dideoxinucleótidos (Russel P., 2002 iGenetics, Pearson Education, Inc. San Francisco, 187-189) utilizando el kit de secuenciación por ciclo BigDye Terminador V.3.0 (Applied Biosystems EE.UU). La secuencia de ADN de P2 se muestra en la Figura 1 (SEQ.ID NO: 1).

El marco de lectura abierto (ORF del inglés Open Reading Frame) se localizó utilizando un buscador de ORF del NCBI (Centro Nacional de Información Biotecnológica). La secuencia de nucleótidos de la Figura 1 se tradujo en todos los seis posibles marcos de lectura y se identificó un marco de lectura abierto de 738 aminoácidos que codifica una posible  $\beta$ -galactosidasa. La traducción se muestra en la Figura 2 (SEQ.ID NO: 2).

La presente invención se describirá con más detalle en referencia al siguiente ejemplo.

#### Ejemplo 1

##### Materiales y métodos

Todos los productos químicos y medios de preparación utilizados a lo largo de este estudio se obtuvieron de Sigma (Dorset, Reino Unido), Invitrogen (Paisley, Reino Unido), Oxoid (Basingstoke, Reino Unido), Qiagen (West Sussex, Reino Unido) y Promega (Southampton, Reino Unido).

##### *Cepas bacterianas*

La cepa de *Bifidobacterium bifidum* (NCIMN 41171) se mantuvo en perlas criogénicas en tubos Microbank a  $-70^{\circ}\text{C}$ . Para experimentos posteriores, la cepa se revivió en un agar de Wilkinson Chalgren (WC) (Oxoid, Reino Unido) y en un medio TPY (medio de extracto de levadura tripticasa phytona del inglés Trypticase Phytone Yeast) y se cultivó anaeróbicamente (composición de  $\text{CO}_2$  y  $\text{N}_2$  al 80% y 20% respectivamente) a  $37^{\circ}\text{C}$  durante 48 horas. La morfología de la colonia y la ausencia de contaminación se probaron con tinción de gram.

##### *Cepas de E. coli*

La cepa de *Escherichia coli* DH5a utilizada en este estudio se incubó en la forma habitual en condiciones aeróbicas a  $37^{\circ}\text{C}$  en un agar de Luria Bertani (LB) o en un caldo (Sambrook J. And Russell W. D) (2001). Molecular cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Nueva York) y cuando era necesario se suplementó con antibióticos (100  $\mu\text{g/ml}$  ampicilina y/o 15  $\mu\text{g/ml}$  cloranfenicol) y 40  $\mu\text{l}$  de X- $\beta$ - Gal al 2%, 7  $\mu\text{l}$  IPTG (isopropil- $\beta$ - D-tiogalctosido) al 20% que se aplicaron en la superficie de una placa de agar de 90 mm previamente preparada.

La cepa DH5a de *E. coli* (Invitrogen, Paisley, Reino Unido) (genotipo: F  $\phi 80\text{lacZ}\Delta\text{M}\Delta(\text{lacZYA-argF})\text{U169 recA1 endA1 hsdR17}(\text{r}_k^-, \text{m}_k^-) \text{phoA supE44 thi-1 gyrA96 relA1}\lambda$ ) es una cepa  $\alpha$ -galactosidasa positiva y fue utilizada en experimentos de expresión y para otras manipulaciones genéticas.

##### *Extracción del ADN genómico de Bifidobacterium bifidum*

Se aisló el ADN genómico de la cepa (NCIMB 41171) de *Bifidobacterium bifidum* utilizando el siguiente método en el que se preparó el ADN cromosómico a partir de gránulos de células recogidos de 100 ml de caldo anaeróbico de WC. La células se volvieron a suspender en 10 ml de tampón TES (Tris-HCl 10 mM, EDTA 10 mM, NaCl 10 mM, pH=8) y se trataron con 200  $\mu\text{l}$  de una mezcla de lisozima/mutanolisina (4:1, lisozima 10 mg/ml mutanolisina 1 mg/ml) durante 30 minutos a  $37^{\circ}\text{C}$ . A continuación las células se trataron con 200  $\mu\text{l}$  de proteinasa K (a 20 mg/ml) y 200  $\mu\text{l}$  de ARNasa (ambos 10 mg/ml) y se incubaron durante una hora a  $65^{\circ}\text{C}$ . Finalmente las células se trataron con 2 ml de SDS al 10% y se incubaron durante 15 minutos a  $65^{\circ}\text{C}$ . Se añadieron 12 ml de fenol/cloroformo y se repitió la extracción hasta que la fase acuosa se pudo separar fácilmente de la interfase. El ADN genómico se precipitó con isopropanol y se volvió a suspender en Tris-HCl 10 mM- EDTA 1 mM (pH=8). El ADN genómico fue digerido con enzimas de restricción, se unió a pSP72 digerido después con las mismas enzimas y se trató con fosfatasa alcalina. La digestión del ADN genómico de *B. Bifidum* se llevó a cabo utilizando *EcoRI*, *PstI*, *BamHI*, *SmaI* y *KpnI*. Las mezclas de ligación se utilizaron para transformar *E. coli* DH5a y se identificaron clones positivos de  $\beta$ -galactosidasa como colonias azules en las placas que contenían X-Gal.

##### *Preparación del Vector de ADN*

El vector utilizado en este estudio para el clonado y la expresión fue pSP72 (Promega, Reino Unido) (Krieg, P.A. y Melton, D.A. 1978). *In vitro* RNA synthesis with SP6 RNA polymerase. (*Methods in Enzymology*, **155**: 397-415).

Se eligió este vector por la falta de actividad de complementación del fragmento  $\alpha$  de  $\beta$ -galactosidasa que no está codificado en pSP72. Este vector no lleva el segmento corto del ADN de *E. coli* que contiene la secuencia reguladora y la información codificante para los 146 aminoácidos primeros de la  $\beta$ -galactosidasa que en combinación con las cepas de *E. coli* (por ejemplo DH5a) que expresan el extremo carboxi-terminal de esta  $\beta$ -galactosidasa dan una  $\beta$ -galactosidasa activa (complementación  $\alpha$ ).

## ES 2 342 125 T3

El vector fue digerido con las enzimas de restricción siguientes: PstI, Bam HI, HindIII, SmaI, KpnI, y, EcoRI según las instrucciones del fabricante utilizando diez veces mas de enzima que de ADN (unidades enzimáticas:  $\mu$ gr ADN igual a diez unidades de enzima por un  $\mu$ gr de plásmido de ADN o diez unidades enzimáticas por 0,5 pmol de plásmido de ADN). Después de la inactivación por calor de la enzima (20 min a 65°C) se analizaron los patrones de restricción por análisis de electroforesis en gel horizontal. La presencia de un solo fragmento en el gel indicaba la digestión completa del vector y únicamente la digestión de restricción del mismo.

Se comprobó también la digestión suficiente del vector mediante la transformación de moléculas no ligadas a células *E. coli* DH5a competentes. El número de colonias formadas en las placas de agar LB suplementadas con ampicilina (100  $\mu$ g/ml) fue un indicador de las moléculas no digeridas y del fondo esperado durante los experimentos siguientes.

A continuación los vectores se defosforilaron con fosfatasa alcalina de intestino de ternera CIAP (Promega, Southampton, Reino Unido) siguiendo las instrucciones del fabricante. La eficacia del tratamiento se comprobó por ligación (con ligasa de ADN del Bacteriófago T4 siguiendo las instrucciones del fabricante) después de la transformación en las células DH5a. El número de colonias formado representaba el número de moléculas recircularizadas (vectores no clonados) y al restar de las mismas las colonias formadas sin tratamiento con el vector CIAP se obtuvo el número de vectores no defosforilados.

### *Construcción de la Biblioteca de ADN Genómico*

El ADN genómico se digirió parcialmente con seis enzimas de restricción que reconocen las secuencias de hexanucleótidos frecuentes en los ADN procariotas. *EcoRI*, *BamHI*, *PstI*, *KpnI*, *SmaI* y *HindIII* son endonucleasas de restricción de tipo II que reconocen específicamente las secuencias 5'G/AATTC'3', 5'G/GATCC'3', 5'CTGCA/G'3', 5'GGTAC/C3', 5'CCC/GGG3' y 5'A/AGCTT3' respectivamente, y realizan rupturas en la doble hebra dentro de estas secuencias generando protuberancias 5' de cuatro nucleótidos, AATT, GATC, AGCT en el caso de *EcoRI*, *BamHI* y *Hind III* respectivamente, y protuberancias 3', ACGT, GTAC en el caso de *PstI* y *KpnI* respectivamente y extremos romos en el caso de *SmaI*.

Todas estas enzimas eran activas y capaces de escindir ADN solamente en presencia de iones de magnesio divalentes. Estos iones fueron el único cofactor requerido.

### *Digestión de restricción del ADN*

Todas las digestiones de restricción de las muestras de ADN genómico se incubaron durante 2 horas a 37°C y finalmente se inactivaron por calor a 65°C durante 20 minutos. Después las reacciones se enfriaron a temperatura ambiente y se añadió la cantidad adecuada de tampón de carga, seguido de mezclado suave en un capilar de vidrio sellado. A continuación se cargaron las disoluciones en pocillos de un gel de agarosa al 0,8% (fuente de alimentación 4-5 volts/cm durante 14-16 horas) y el tamaño del ADN digerido se estimó con el de estándares de ADN de 1kb (Promega, Reino Unido) (Sambrook J, Molecular Cloning: A Laboratory Manual (2002)).

### *Purificación de los fragmentos generados después de la digestión de restricción*

La purificación de los fragmentos de las mezclas de reacción y de los geles de agarosa se realizó utilizando el kit de extracción en gel QIAEX de Qiagen (West Sussex, Reino Unido). Los protocolos están descritos detalladamente en el manual del fabricante.

### *Transformación y ligación del ADN*

Después de la purificación de los fragmentos de ADN con el kit de extracción en gel QIAEX, se ligaron con el vector pSP72 tratado con CIAP. Para la ligación, se transfectaron las cantidades adecuadas de ADN a tubos de microfuga estériles de 0,5 ml como se muestra en la Tabla 1.

## ES 2 342 125 T3

TABLA 1  
Mezclas de ligación

| Tubo                                                                                                                                                                                              | ADN                                                                                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| A                                                                                                                                                                                                 | Vector (15 fmoles [ $\sim 29,7$ ng])                                                    |
| B                                                                                                                                                                                                 | Vector (15 fmoles $\sim 29,7$ ng ADN) mas inserto (15 fmoles exteriores $\sim 69,3$ ng) |
| C                                                                                                                                                                                                 | Control pUC (0,056 fmoles [ $\sim 100$ pg])                                             |
| La proporción molar de vector plásmido de ADN a inserto de fragmento de ADN debería ser de $\sim 1:1$ en la reacción de ligación. La concentración final de ADN debería ser $\sim 10$ ng/ $\mu$ l |                                                                                         |

El tubo A muestra el número de vectores de ADN auto-ligados que deben ser sustraídos del número total de transformantes después de la transformación. El tubo B muestra la ligación del vector con los fragmentos de ADN y el tubo C muestra el control para poder calcular la eficacia de la transformación.

Antes de cada ligación los fragmentos de ADN se calentaron a  $45^{\circ}\text{C}$  durante 5 minutos para fundir cualquier extremo cohesivo que anillarse durante la preparación de los fragmentos. Se eligió una proporción molar vector: inserto de ADN de 1:1 para todas las reacciones de ligación y la reacción de ensamblado se realizó según las instrucciones de Promega.

Se añadió a los tubos A y B  $1,0 \mu\text{l}$  de tampón de ligación x 10 y 0,5 unidades Weiss de ligasa de ADN T4 (Pormega, Reino Unido) y el volumen de ligación se ajustó a  $10 \mu\text{l}$  con agua grado para biología molecular. Se añadió a los tubos C  $1,0 \mu\text{l}$  de tampón de ligación x 10 y el volumen de ligación se ajustó a  $10 \mu\text{l}$  con agua grado para biología molecular.

Se añadieron los fragmentos de ADN a los tubos junto con el agua y después se calentaron a  $45^{\circ}\text{C}$  durante 5 minutos para fundir cualquier extremo cohesivo que anillase durante la preparación. El ADN se enfrió a  $0^{\circ}\text{C}$  antes de añadir el resto de los reactivos de ligación y las mezclas de reacción se incubaron durante la noche a  $16^{\circ}\text{C}$  (Sambrook and Russell, 2001).

Después de la precipitación con etanol y la purificación de los fragmentos ligados (para eliminar la mezcla de ligación que causa una reducción de la eficacia de la transformación) se llevaron a cabo las transformaciones según las instrucciones de Hanahan. Se añadieron  $\sim 50$  ng de ADN ligado en  $5 \mu\text{l}$  de disolución a  $100 \mu\text{l}$  de células DH5a competentes. Después del tratamiento de calentamiento y de la expresión del gen de resistencia a la ampicilina las células se extendieron sobre la superficie de las placas LB que contenían ampicilina ( $100 \mu\text{g/ml}$ ), X- $\beta$ -Gal ( $40 \mu\text{l}$  de X- $\beta$ -Gal al 2%) y IPTG ( $7 \mu\text{l}$  de IPTG al 20%).

Se midió el número de transformantes de cada reacción de ligación. El número de transformantes obtenido normalmente en el tubo C fue de  $2 \times 10^5$  -  $1 \times 10^6$  cfu/ $\mu\text{g}$  mientras que en el tubo A fue de 500-600 cfu/ $\mu\text{g}$ . El número de transformantes del tubo B era una indicación de la eficacia del tratamiento del vector de ADN. El número de transformantes del tubo B estaba en un intervalo a partir de  $2-4 \times 10^4$  cfu/ $\mu\text{g}$ .

### Número de transformantes

Las mezclas de ligación con ADN cromosómico de *PstI* dieron lugar a 13 clones positivos de  $\beta$ -galactosidasa de entre los  $\sim 2500$  transformantes analizados mientras que con *BamHI* dieron lugar a 3 clones positivos ( $\sim 1300$  transformantes analizados), con *KpnI* dieron lugar a 7 clones positivos ( $\sim 2000$  transformantes analizados), con *SmaI* dieron lugar a 3 clones positivos ( $\sim 1600$  transformantes analizados) y con *HinIII* dieron lugar a 2 clones positivos ( $\sim 1200$  transformantes analizados).

### Digestión de clones positivos

Para identificar los genes de las diferentes  $\beta$ -galactosidasas, los plásmidos aislados de los clones positivos fueron digeridos según la siguiente tabla.

|                                                                                                                                                                                                       | Muestras                                                   | Enzimas                      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|------------------------------|
| 1ª Digestión                                                                                                                                                                                          | pB1, pB2, pB3, pB4, pB5, pB6, pB7                          | <i>Bam</i> HI                |
| 2ª Digestión                                                                                                                                                                                          | pP1, pP2, pP3, pP4, pP5, pP6, pP7, pP8,<br>pP9, pP10, pP11 | <i>Pst</i> I                 |
| 3ª Digestión                                                                                                                                                                                          | pP12, pP13, pP14                                           | <i>Pst</i> I                 |
| 4ª Digestión                                                                                                                                                                                          | pE1, pE2, pE3                                              | <i>Eco</i> RI                |
| 5ª Digestión                                                                                                                                                                                          | pP1, pP12, pB1, pP2, pE1, pE2, pE3.....                    | <i>Pst</i> I y <i>Eco</i> RI |
| 6ª Digestión                                                                                                                                                                                          | pS1, pS2, pS3                                              | <i>Sma</i> I                 |
| 7ª Digestión                                                                                                                                                                                          | pP1, pP12, pB1, pP2, pS1, pS2, pS3                         | <i>Pst</i> I y <i>Sma</i> I  |
| 8ª Digestión                                                                                                                                                                                          | pK1, pK2, pK3, pK4, pK5, pK6, pK7                          | <i>Kpn</i> I                 |
| 9ª Digestión                                                                                                                                                                                          | pP1, pP12, pB1, pP2, pK1, pK2, pK3, pK4,<br>pK5, pK6, pK7  | <i>Pst</i> I y <i>Kpn</i> I  |
| La primera letra (p) indica el plásmido y el gen insertado mientras que la segunda letra (P,B,E,S,K) indica la encima de restricción que se usó para aislar el clon correspondiente del ADN genómico. |                                                            |                              |

Los análisis de la electroforesis en gel de los fragmentos generados después de la digestión mostraron que cada uno de los plásmidos pB1, pP1, pP2 y pP11 tiene un inserto que codifica una  $\beta$ -galactosidasa diferente. Los clones que contienen P2 se utilizaron para análisis posteriores.

#### Secuenciación de ADN

La secuenciación del ADN se efectuó con el método de Sanger de terminación de cadenas por dideoxinucleótidos utilizando el kit de secuenciación por ciclo BigDye Terminador v.3.0 (Applied Biosystems EE.UU) y se analizó con el ABI Prism 3100, un sistema de análisis de ADN basado en la fluorescencia que incorpora electroforesis capilar.

Los extremos 5' y 3' de los fragmentos de insertos de ADN se secuenciaron con cebadores específicos de los vectores. A continuación los insertos se secuenciaron utilizando el Sistema de Cebadores Genómico (GPS-1) (New England Biolabs, Reino Unido). GPS-1 es un sistema *in vitro* basado en el transposón TN7 que utiliza la Transposasa TnsABC para insertar el Transposón aleatoriamente en el ADN diana. Se utilizó una proporción de masa donante: ADN diana de 1:4 siguiendo las instrucciones del fabricante. El número de plásmidos aislados para la secuenciación después de la inserción del Transcebador en el plásmido diana fue de 25. Este número se calculó siguiendo las instrucciones del fabricante y supone una profundidad de cobertura de cinco veces.

Debido a la larga secuencia de nucleótidos del plásmido pP2 que codifica la proteína P2, se seleccionó para ser secuenciada únicamente la parte que contiene el gen de la  $\beta$ -galactosidasa. La enzima inactivada después de la inserción del inserto de transposasa en una posición aproximada de 172 pb del fragmento secuenciado indicaba que el codón de iniciación se encontraba aguas arriba de esta posición. De manera similar, la inserción del inserto en la posición 2882bp eliminó completamente la actividad de la enzima indicando que el codón de finalización se encontraba aguas abajo de esta posición. Además la actividad de la enzima se eliminó completamente con la inserción de los insertos en las posiciones 262pb, 331pb, 375pb, 621pb, 866pb, 1348pb, 1358pb, 1394pb, 1513pb, 1704pb, 2128pb, 2519pb en relación con el primer nucleótido que había sido secuenciado.

El análisis del dominio N-terminal con el software SignalIP y PSORT. No mostró la señal de ningún péptido, indicando que P2 no se secretaba extracelularmente.

La mezcla de reacción de secuenciación contenía aproximadamente 400-600 ng de plásmido de ADN, 3,2 pmol de disolución de cebador y 4  $\mu$ l de disolución Terminator BigDye.

## 5 Identificación del marco de lectura abierto

El marco de lectura abierto (ORF) de P2 se localizó utilizando el buscador ORF de NCBI en la dirección web <http://www.ncbi.nlm.nih.gov/gorf.html>. Se utilizó el código genético de la bacteria y se determinó que la longitud del marco era de 100bp. La secuencia del nucleótido se tradujo en todos los seis posibles marcos de lectura y se identificó un marco de lectura abierto de 738 aminoácidos que codifica una posible  $\beta$ -galactosidasa (la traducción se muestra en la figura 2).

## Ejemplo 2

*Síntesis con la enzima  $\beta$ -galactosidasa clonada aislada de Bifidobacterium bifidum NCIMB 41171 en el hospedador E. coli (cepa DH5a)*

La síntesis descrita a continuación, a no ser que indique lo contrario, se efectuó con la totalidad de las células hospedadoras de *E. coli* DH5a después del tratamiento de la biomasa de *E. coli* (recogida por centrifugación a 10.000 g) con tolueno a una concentración de 2000 ppm para incrementar la permeabilidad celular y también para hacer las células no viables destruyendo la membrana citoplásmica. La biomasa de *E. coli* se preparó como se ha descrito en ejemplo 1 bajo la denominación de “Cepas de *E. coli*”.

## 25 Síntesis con la enzima clonada

La síntesis con la  $\beta$ -galactosidasa se efectuó con una concentración de sustrato del 40% (proporción en peso) de concentración inicial de lactosa. La disolución de síntesis se preparó con un tampón fosfato 0,1 M a pH = 6,0 que contiene 1 g/l de Tween 80 (monooleato de polioxietileno (20) sorbitán) adicional. La síntesis se efectuó a 40°C en un baño de agua con agitación a 150 rpm. El pH óptimo para la enzima específica se eligió en base a las medidas de actividad (utilizando como sustrato O-nitrofenil- $\beta$ -D-galactopiranosido) de una preparación enzimática específica a varios valores de pH.

Para la síntesis de galacto-oligosacáridos se utilizaron 2 ml del sobrenadante del lisado celular (después de la ruptura de las células de *E. coli* con una Prensa Francesa) con 8 g de lactosa al 50% (porcentaje en peso) para obtener una concentración de sustrato final del 40% (porcentaje en peso). Esta preparación enzimática tenía una actividad de 735 U/ml.

Las concentraciones de los distintos azúcares presentes en la mezcla durante la síntesis se muestran en la Figura 3. En la Figura 4 se muestra la cromatografía de intercambio aniónico de alto rendimiento acoplada a un cromatogramas de detección amperométrica pulsada (HPAEC-PAD) de las mezclas de galacto-oligosacáridos sintetizados por la  $\beta$ -galactosidasa clonada de *B. Bifidum* NCIMB 41171. Las concentraciones de azúcar de la mezcla de galacto-oligosacáridos en el momento óptimo de síntesis se muestran en la Tabla 1.

TABLA 1

*Composición de hidratos de carbono de la síntesis de galacto-oligosacáridos a una concentración de lactosa inicial del 40% (proporción en peso) en el momento en el que se observó la concentración máxima de oligosacáridos*

50

| Síntesis inicial de sustrato | GOS DP $\geq$ 3                       | GOS DP=2 | Lac.  | Glc.  | Gal.  |
|------------------------------|---------------------------------------|----------|-------|-------|-------|
| %(en peso)                   | Concentración (% de azúcares totales) |          |       |       |       |
| 40                           | 8,82                                  | 16,25    | 39,40 | 20,76 | 14,85 |

60

Lac: Lactosa, Glc: Glucosa, Gal: galactosa, DP: grado de polimerización

65

# REIVINDICACIONES

1. Una secuencia de ADN que codifica una  $\beta$ -galactosidasa, donde la secuencia se selecciona entre la secuencia dada en SEQ. ID NO: 1 o de un fragmento o una degeneración de la misma, teniendo dicha degeneración una homología de 75% a 95%.

2. La secuencia de ADN según la reivindicación 1, donde dicha secuencia comprende sustituciones, adiciones o deleciones de nucleótidos que tienen como resultado un cambio en la secuencia de aminoácidos según la SEQ. ID NO: 2 inferior al 25%.

3. Una secuencia de ADN según la reivindicación 1 o reivindicación 2, donde dicha secuencia comprende sustituciones de nucleótidos que tienen como resultado sustituciones de aminoácidos conservadoras.

4. Una enzima  $\beta$ -galactosidasa codificada por una secuencia de ADN de una cualquiera de las reivindicaciones 1 a 3.

5. Una enzima  $\beta$ -galactosidasa que comprende una secuencia de aminoácidos según SEQ. ID NO:2.

6. Un vector recombinante que comprende una secuencia de ADN de una cualquiera de las reivindicaciones 1 a 3.

7. El vector según la reivindicación 6, donde dicho vector es un vector de expresión.

8. Una célula hospedadora que es una célula de levadura, una célula de hongo o una célula bacteriana seleccionada del grupo que consiste en *Lactococcus*, *Lactobacillus*, *Escherichia*, *Bacillus* y *Aspergillus* que comprende una secuencia de ADN de una cualquiera de las reivindicaciones 1 a 3.

9. Una célula hospedadora que comprende el vector de la reivindicación 6 o de la reivindicación 7.

10. La célula hospedadora según la reivindicación 8 o reivindicación 9, donde la célula se selecciona del grupo que consiste en *Bacillus subtilis*, *Bacillus circulans* y *Aspergillus Níger*.

11. Uso de una enzima de la reivindicación 4 o reivindicación 5, o de una célula de una cualquiera de las reivindicaciones 8 a 10, para producir una mezcla de oligosacáridos.

12. Uso de una enzima de la reivindicación 4 o reivindicación 5, o una célula de una cualquiera de las reivindicaciones 8 a 10, para producir una mezcla de oligosacáridos que sea parte de un producto seleccionado del grupo que consiste en productos lácteos, tales como leche líquida, leche en polvo seca, leches infantiles, leche maternizada, helado, yogur, queso, productos de leche fermentada, bebidas tales como zumo de fruta, alimentos infantiles, cereales, pan, galletas, dulces, pasteles, suplementos alimenticios, suplementos dietéticos, productos comestibles probióticos, productos comestibles prebióticos, alimentos para animales, alimentos para aves y medicamentos.

13. Uso de una célula hospedadora de una cualquiera de las reivindicaciones 8 a 10, para producir un producto seleccionado del grupo que consiste en productos lácteos, tales como leche líquida, leche en polvo seca, leches infantiles, leche maternizada, helado, yogur, queso, productos de leche fermentada, bebidas tales como zumos de fruta, alimentos infantiles, cereales, pan, galletas, dulces, pasteles, suplementos alimenticios, suplementos dietéticos, productos comestibles probióticos, productos comestibles prebióticos, alimentos para animales, alimentos para aves y medicamentos.

14. Un proceso para producir una enzima de la reivindicación 4 o reivindicación 5, que comprende cultivar una célula hospedadora de una cualquiera de las reivindicaciones 8 a 10 en un medio de cultivo adecuado en condiciones que permita la expresión de dicha enzima y recuperar la enzima resultante del cultivo.

Figura 1

1 tccttagat tggattgaaa acttctcttt ttatgtaaa aggatattcg gggaaattcc  
 61 cttttgaaa atttcatgac caaaaalccc cttaacgtga gttttcgttc cactgagcgt  
 121 cagaccccggt agaaaagatc aaaggatctt cttagatcc tttttctg cgcgtaatct  
 181 gctgcttga acaaaaaaa cccccgtac cagcgggtgt tgtttgccg gatcaagagc  
 241 taccaactct tttccgaag gtaactggct tcagcagagc gcagatacca aalactgttc  
 301 ttctagtga gccgtagtta ggccaccact tcaagaatic ttagcaccg cctacatacc  
 361 tcgtctgct aatctgtta ccagtggctg cggcagtg cgataagtcg tgtcttaccg  
 421 ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggt  
 481 cgtgcacaca gcccagcttg gagcgaaacga cctacaccga actgagatac ctacagcgtg  
 541 agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggttaagcg  
 601 gcagggtcgg aacaggagag cgcacgaggc agcttcagg gggaaacgcc tggatcttt  
 661 alagictgt cgggttgcg cactctgac ttgagcgtcg attttgtga tgcctgtag  
 721 gggggcggag cctatgaaa aacgccagca acgcccctt ttacgggtc ctggccttt  
 781 gctggcctt tgcacatg ttcttctg cgttatcccc tgattctgt gataaccgta  
 841 ttaccgctt tgagttagct gataccgctc gccgcagccg aacgaccgag cgcagcaggt  
 901 cagtgcgga ggaagcggaa gagcgcccaa tacgaaacc gccctcccc gcgcgttggc  
 961 cgaticatta atgcaggta acctggctta tcgaaattaa tacgactcac tatagggaga  
 1021 ccggcagatc tgatatcatc gatgaattcg agctcggta cgggggatcc tctagatcg  
 1081 acctgcagaa gtactggcag tccggcctcg ccgcccgcgc cgtcaaggag tgatctcgcg  
 1141 ggagggccag tgcctgttg cagcatgct caggttcgtc tgacggctt gactgcgcc  
 1201 tagcggcttg cctccctac ggaacatga tatgatagg agggacgctc accgtgtgag  
 1261 atgcgggctg gcgcgatatg gcaccagccc gcacgcagc tcacagttga ttttctgc  
 1321 atgcgacggt tcggccgggg cgtcagacan tgaacaagta atggagcggg atatgagtaa  
 1381 acgcagaaag cacagtggc cgcagccgct gaaggcgccc gaatcccgtc tctggtatgg  
 1441 cgggtattac aatcccgacc aatggccgga ggaggtctgg gacgatgata tccgtctgat  
 1501 gaagaaggcc ggcgtcaacc ttgtatcagt cggcatttc icatgggcca agatcgagcc  
 1561 ggaggaaggc aagtacgatt tcgactggct ggaccgtgcc atcgataagc tcggcaaggc  
 1621 gggcatcgct gtagacctg cctccgccac cgcctcccc ccgatgtggc tgaccagggc  
 1681 ccatccggag gtgctgtgga aggacgagcg cggcgacacc gtgtggccgg gtgccgtga  
 1741 gcatggcgt ccgacgagcc ctgtgttccg cgagtacgc cgaacctgt gccgtctat  
 1801 ggccgagcac tataaggga acccatagct ggtggcctgg cagtgagca acgaatacgg  
 1861 ctgcacaac cgttcgact actcggagca cgcctgcgc gccctccaga agtgggtcaa  
 1921 gaagcgtac aagaccatcg acgcggtgaa cgaggcctgg ggaacagcgt totgggcca  
 1981 gcacatgaac gatttctccg agatcatccc gccctgctac atcggtagc gcaactcat  
 2041 gaaccggggc aagctgctcg actacaagcg gttcagctcc gacgcgctca aggagctta  
 2101 catcgccgaa cgtgacgtgc tggagtccat cagccgggt ctgccgtga ccaccaattt  
 2161 catggtgtct gccggcggt cgtatgtga ctatgacgac tgggtgctgg aggtcattt  
 2221 cgtctcaac gaccactact tcacgccggg cgaggacat ttgacgagg ttgcgtacgc  
 2281 cgcgtcgtc atggacggca ttacacgaa ggaaccgtgg ttccagatgg agcattccac  
 2341 ctccgggtg aattggcgtc cgtacaaata ccgtgcggag ccgggctcgg tcgtgcgga  
 2401 ctccctgcg caggtcgtta tgggcggcga tgcctatgc tatttcagt ggcgacagt  
 2461 caaggccggc gcggagaagt ggcactctc gatgtgccc catcgggcg aggtatcgca  
 2521 gatctccgt gatgtgtgc agctgggcgc cgtctgccc acgagggctt

2581 gatgggcacc aagacgggtca agtccaaagg cgcgcgtcgtg ttcgactacg agtcccagt  
 2641 ggccaccgag tacaccgca acccgaccca gcaggtcgat cattggaccg agccgttggg  
 2701 ttggttccgc gcgcttgccg acaacggcat caccgccgac gtggttccgg tgcggtcga  
 2761 ttgggaticc tacgagattg ccgtgcgcc gtgcgtgtat ttgctgtccg aagagacgag  
 2821 ccgcagggtt cgggagttcg tggcgaacgg cggcaagtgt ttcgtgacgt actacaccgg  
 2881 attgagcgac gagaacgacc acatcggct tggcggctac ccgggctcga ttcgtgacgt  
 2941 cgtcggcgtg cgcgtcgaag agtgcgcc gatgggcaac gacatgccgg gcgcattgga  
 3001 icacctcgac ctggacaatg gcacgggtggc gcatgacttt gccgacgtga tcacgtccac  
 3061 ggccgacacg tocacgggtac tggcctccta taaggcggaa cgtgggaccg gcatgaacga  
 3121 ggtgccggcc atcgtggcca acgggtatgg cgacgggccc accgtatatg tcggaigccg  
 3181 tctgggcccgt caggggcttg cgaagagcct gccggcgatg ctgggttcca tggggctgtc  
 3241 ggacctcgcc ggtgatggcc gcgtgctgcg tgcgaacgc gccgacgccg cggcggcgag  
 3301 ccgttcgag ttcgtgttca accgcacca tgaaccgggt accgtcgacg ttgaagggga  
 3361 ggccatcgcc gcttcgctcg cgcatttga cgacgggccc gccaccatcg atccgacggg  
 3421 tgtgtcgtg ctcaggcgat aatcgttggg aacactgggc tgaagggt taggaaaggc  
 3481 gtaigtgtg ggtgacacgc gacatacgcc ttatgggaaa gaaggcgtg gcgcttaccg  
 3541 ggctgcggcg atgctgttca gcgtcgtgc gtgcggcaac tcaagcagca gctccgtcc  
 3601 gaagcaggaa ggcgacgtca aggaatcac cgtgtgggt tggggcctac gctgactcag  
 3661 gtggccaagg acttcaaaaa aggagaccgg catcaatgtc aacctggtga acaccggcca  
 3721 gggcgacaag acctgggacg agttctatca ggacgccaag aagattcaca ccttggcga  
 3781 caactactac atcacgtccg acaccggtgt cgcgggcttc tacgactcga tgacctggt  
 3841 ggccagtgcg acgctgttct ccaccgaagg cgagacggtc accattaacc tgactggcg  
 3901 cccgaaggtc aaggcccgcg gtatcttcgg cgactacctt ggcaagtcct acaccggcaa  
 3961 ccagaagctg agcgatggcg tcgcgctti gggaacaggc tctgaaggac tacgcgaagg  
 4021 atcagggcta caccgtcaag taaccttcgc agtcaagcaa tctggcgtgg taatgaccgg  
 4081 aatacgggtga ccttcggica tcccttcctc gtgtgaaggc ccttccctc aacagggagg  
 4141 ggcttcaca tatctgcccc tgttgcaacg cgcgtgtaaa ctctacgat agcgaattct  
 4201 tcccgacaca tcgagcacgc taaggagatg acatgacgat atcggcacgg ttgtggcgcc  
 4261 tgcacctgca tatcttgtg ttgctcaaga tctgtgagat gacactggca gcacgcctcc  
 4321 agcgcgccgc cggacacgcg caccctcatc accgaatgga cgggggacca tatcatgacg  
 4381 actctgatcg ccaat

Figura 2

MYMIWRDAHRVRCGLARYGTSPHRMSQLIFRRCDGSAGASDNEQVMERNMSK  
RRKHSWPQPLKGAESRLWYGGDYNPDQWPPEVWDDDIRLMKKAGVNLVSV  
GIFSWAKIEPEEGKYDFDWLDRAIDKLGKAGIAVDLASATASPPMWLTQAFP  
EVLWKDERGDTVWPGAREHWRPTSPVFREYALNLCRRMAEHYKGNPYVVAWH  
VSNEYGCHNRFDYSDDAMRAFAQWCKKRYKTIDAVNEAWGTAFWAQHMNDFS  
EIIIPRYIGDGNFMNPGKLLDYKRFSSDALKELYIAERDVLESITPGLPLTT  
NFMVSAGGSMLDYDDWGAEVDFVSNHYFTPGEDHFDEVAYAASLMDGISRK  
EPWFQMEHSTSAVNWRPINYRAEPGSVVRDSLAQVAMGADAICYFQWRQSKA  
GAEKWHSSMVPHAGEDSQIFRDVCELGADLGRLSDEGLMGTKTVKSKVAVVF  
DYESQWATEYTANPTQQVDHWTEPLDWFRALADNGITADVVPVRSDWDSYEI  
AVLPCVYLLSEETSRRVREFVANGGKLFVTTYTGLSDENDHIWLGGYPGSIR  
DVVGVRVEEFAPMGNDMPGALDHLDLNGTVAHDFADVITSTADTSTVLASY  
KAERWTGMNEVPAIVANGYGDGRTVYVGCRLGRQGLAKSLPAMLGSMGLSDL  
AGDGRVLRVERADAAAASRFEFVFNRTHEPVTVDVEGEAIAASLAHVDDGRA  
TIDPTGVVVLRR

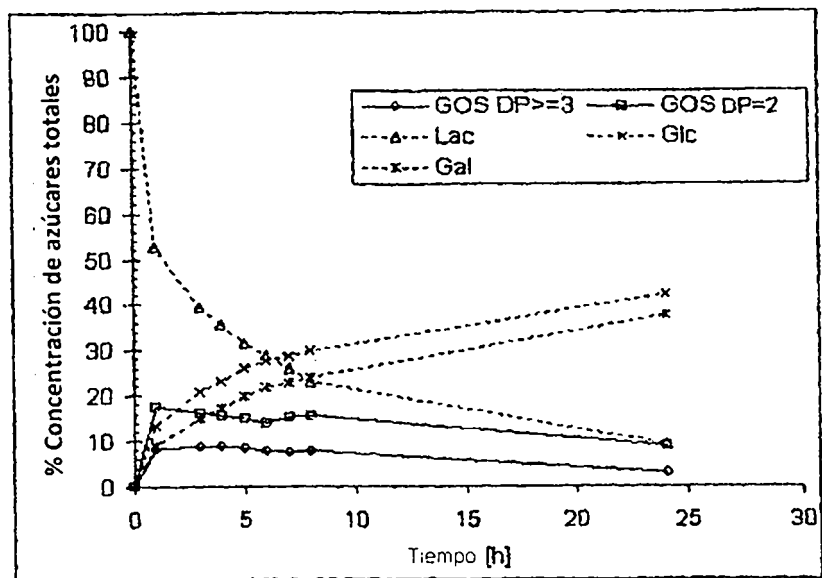


Figura 3

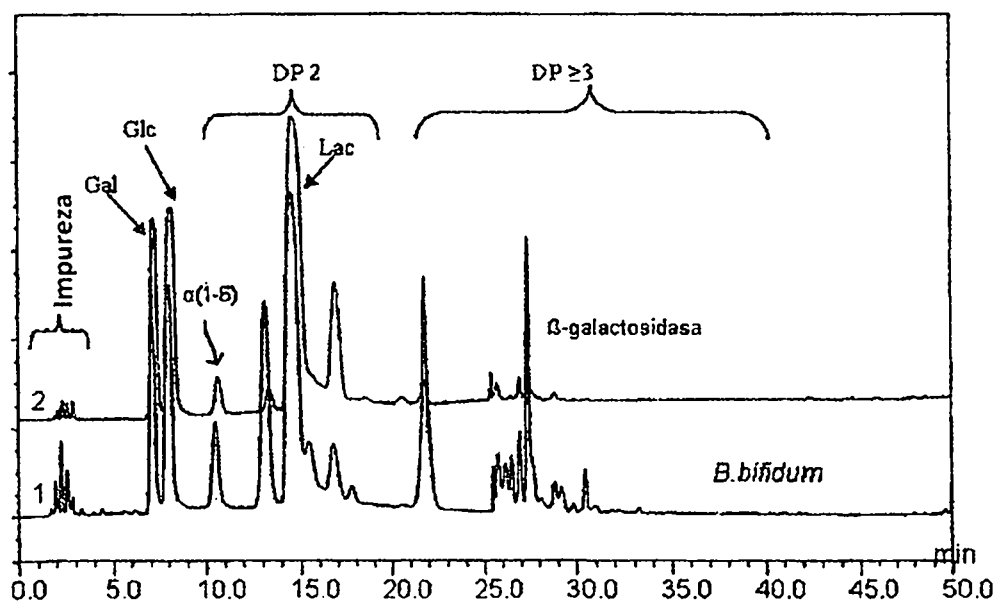


Figura 4

# ES 2 342 125 T3

## LISTA DE SECUENCIAS

<110> MCDONALD, Sarah C

5 <120> Producto y proceso

<130> 14011:KG

10 <150> GB 0606112.1

<151> 2006-03-28

<160> 2

15 <170> PatentIn version 3.4

<210> 1

20 <211> 4395

<212> ADN

<213> *Bifidobacterium bifidum*

25 <220>

<221> Beta-galactosidasa

<222> (1)..(4395)

30 <400> 1

|    |                                                                     |      |
|----|---------------------------------------------------------------------|------|
|    | tcctttagat tggattgaaa acttcctttt tttatgtaaa aggatattcg gggaaattcc   | 60   |
|    | ctttttgaaa atttcattgac caaaaatccc cttaacgtga gttttcggtc cactgagcgt  | 120  |
| 35 | cagaccccggt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaattct | 180  |
|    | gctgcttgca aacaaaaaaa cccccgctac cagcgggtgt ttgtttgccg gatcaagagc   | 240  |
|    | taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc   | 300  |
| 40 | ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc   | 360  |
|    | tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg   | 420  |
|    | ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt  | 480  |
| 45 | cgtgcacaca gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg    | 540  |
|    | agctatgaga aagcggccag cttcccgaag ggagaaaggc ggacaggatc ccggtaaagcg  | 600  |
|    | gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatctt   | 660  |
| 50 | atagtcctgt cgggtttcgc caccctctgac ttgagcgtcg atttttgtga tgctcgtcag  | 720  |
|    | gggggcgagg cctatggaaa aacgccagca acgcggcctt ttacgggtc ctggcctttt    | 780  |
|    | gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta   | 840  |
| 55 | ttaccgcctt tgagttagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt   | 900  |
|    | cagttagcga ggaagcggaa gagcgccaa tacgcaaacc gcctctcccc gcgcgttggc    | 960  |
|    | cgattcatta atgcagggtta acctggctta tcgaaattaa tacgactcac tataggggaga | 1020 |
| 60 | ccggcagatc tgatatcatc gatgaattcg agctcggtag ccggggatcc tctagagtcg   | 1080 |
|    | acctgcagaa gtactggcag tccggcctcg ccgcccgcgc cgtcaaggag tgatctcgcg   | 1140 |
|    | ggagggccag tgtccgcttg cagcatgcgt cacttctgtc tgacggctctt gacgtgcgcc  | 1200 |
| 65 | tagcggctgg ccctccctac ggaacatgta tatgatattg agggacgctc accgtgtgcg   | 1260 |

# ES 2 342 125 T3

|    |                                                                   |      |
|----|-------------------------------------------------------------------|------|
|    | atgcgggctg gtcgcatatg gcaccagccc gcatcgcatg tcacagttga tttttcgtcg | 1320 |
| 5  | atgcgacggt tcggccgggg cgtcagacaa tgaacaagta atggagcggg atatgagtaa | 1380 |
|    | acgcagaaag cacagttggc cgcagccgct gaagggcgcc gaatcccgtc tctggtatgg | 1440 |
|    | cggtgattac aatcccgacc aatggccgga ggaggtctgg gacgatgata tccgtctgat | 1500 |
| 10 | gaagaaggcc ggcgtaacc ttgtatcagt cggcatcttc tcatgggcga agatcgagcc  | 1560 |
|    | ggaggaaggc aagtacgatt tcgactggct ggaccgtgcc atcgataagc tcggcaaggc | 1620 |
|    | gggcatcgct gtagacctcg cctccgccac cgcttccccg ccgatgtggc tgacccaggc | 1680 |
|    | ccatccggag gtgtgtgga aggacgagcg cggcgacacc gtgtggccgg gtgcccgtga  | 1740 |
| 15 | gcattggcgt ccgacgagcc ctgtgttccg cgagtacgcg ctgaacctgt gccgtcgtat | 1800 |
|    | ggccgagcac tataagggca acccatacgt ggtggcctgg cacgtgagca acgaatacgg | 1860 |
|    | ctgccacaac cgcttcgact actcggacga cgccatgcgc gccttcaga agtgggtgaa  | 1920 |
| 20 | gaagcgctac aagaccatcg acgcggtgaa cgaggcctgg ggaacagcgt tctgggccc  | 1980 |
|    | gcacatgaac gatttctccg agatcatccc gcctcgctac atcggtgacg gcaacctcat | 2040 |
|    | gaacccgggc aagctgctcg actacaagcg gttagctcc gacgcgctca aggagctcta  | 2100 |
| 25 | catcgccgaa cgtgacgtgc tggagtccat cacgccgggt ctgccgctga ccaccaattt | 2160 |
|    | catggtgtct gccggcggct cgatgcttga ctatgacgac tggggtgcgg aggtcgattt | 2220 |
|    | cgtctccaac gaccactact tcacgccggg cgaggaccat ttcgacgagg ttgcgtacgc | 2280 |
| 30 | cgctcgtcg atggacggca ttccacgcaa ggaaccgtgg ttccagatgg agcattccac  | 2340 |
|    | ctccgcgggt aattggcgtc cgatcaacta ccgtgcggag ccgggctcgg tcgtgcgcga | 2400 |
|    | ctccctcgcg caggctgcta tgggcgcgga tgccatctgc tatttcagt ggcgacagtc  | 2460 |
| 35 | caaggccggc gcggagaagt ggactcctc gatggtgccg catgcggcg aggattcgca   | 2520 |
|    | gatcttccgt gatgtgtgcg agctgggcgc gatctggga cgtctgtccg acgagggctt  | 2580 |
|    | gatgggcacc aagacggtca agtccaaggt cgccgtcgtg ttcgactacg agtcccagtg | 2640 |
| 40 | ggccaccgag tacaccgca acccgaccca gcaggctgat cattggaccg agccgttggg  | 2700 |
|    | ttggttccgc gcgcttgccg acaacggcat caccgcccgc gtggttccgg tgcggtccga | 2760 |
|    | ttgggattcc tacgagattg ccgtgctgcc gtgcgtgtat ttgctgtccg aagagacgag | 2820 |
|    | ccgcagggtt cgggagttcg tggcgaacgg cggcaagttg ttcgtgacgt actacaccgg | 2880 |
| 45 | attgagcgac gagaacgacc acatctggct tggcggctac ccgggctcga ttcgtgacgt | 2940 |
|    | cgtcggcgtg cgcgtcgaag agttcgcccc gatgggcaac gacatgccgg gcgcattgga | 3000 |
|    | tcacctcgac ctggacaatg gacgggtggc gcatgacttt gccgacgtga tcacgtccac | 3060 |
| 50 | ggccgacacg tccacggtag tggcctccta taaggcggaa cgctggaccg gcatgaacga | 3120 |
|    | ggtgccggcc atcgtggcca acgggtatgg cgacgggcgg accgtatatg tcggatgccg | 3180 |
|    | tctgggccgt caggggcttg cgaagagcct gccggcgatg ctgggttcca tggggctgtc | 3240 |
| 55 | ggacctcgct ggtgatggcc gcgtgctgcg tgtcgaacgc gccgacgccg cggcggcgag | 3300 |

# ES 2 342 125 T3

ccgcttcgag ttcgtgttca accgcaccca tgaaccggtg accgtcgacg ttgaagggga 3360  
 ggccatcgcg gcttcgctcg cgcgtgtcga cgacgggcgg gccaccatcg atccgacggg 3420  
 5 tgttgcgtg ctcaggcgat aatcgttggg aacactgggc tgaagggtc taggaaaggc 3480  
 gtatgtttgc ggtgacacgc gacatacgcc ttatgggaaa gaaggcgctg gcgcttaccg 3540  
 ggctgcggcg atgctgggtca gcgtcgctgc gtgcggcaac tcaagcagca gctccgctcc 3600  
 10 gaagcaggaa ggcgacgtca aggaatcac cgtgtgggct tggggcctac gctgactcag 3660  
 gtggccaagg acttcaaaaa aggagaccgg catcaatgtc aacctgggtga acaccggcca 3720  
 gggcgacaag acctgggacg agttctatca ggacgccaag aagattcaca cccttggcga 3780  
 15 caactactac atcacgtccg acaccggtgt cgccggcttc tacgactcga tgacctggct 3840  
 ggccagtgcg acgctgttct ccaccgaagg cgagacggtc accattaacc tgactggcgc 3900  
 cccgaaggtc aaggcccgcg gtatcttcgg cgactacctt ggcaagtctt acaccggcaa 3960  
 20 ccagaagctg agcgatggcg tcgcccgttt gggaacaggc tctgaaggac tacgcgaagg 4020  
 atcagggtca caccgtcaag taaccttcgc agtcaagcaa tctggcggtg taatgaccgg 4080  
 aatacgggtga ccttcgggtca tcccttcctc gtgtgaaggc ccctccccctc aacagggagg 4140  
 25 ggccttcaca tatctgcccc tgttgcaacg cgcgtgtaaa ctctacgatg agcgaattct 4200  
 tcccgcacaca tcgagcacgc taaggagatg acatgacgat atcggcacgg ttgtggcggc 4260  
 tgcacctgca tatctttgtg ttgtcaaga tctgtgagat gacactggca gcacgcctcc 4320  
 30 agcgcgccgc cggacacgcg caccctcatc accgaatgga cgggggacca tatcatgacg 4380  
 actctgatcg ccaat 4395

<210> 2

35 <211> 738

<212> PRT

<213> *Bifidobacterium bifidum*

40 <220>

<221> Beta-galactosidasa

<222> (1)..(738)

45 <400> 2

Met Tyr Met Ile Trp Arg Asp Ala His Arg Val Arg Cys Gly Leu Ala  
 1 5 10 15

50 Arg Tyr Gly Thr Ser Pro His Arg Met Ser Gln Leu Ile Phe Arg Arg  
 20 25 30

55 Cys Asp Gly Ser Ala Gly Ala Ser Asp Asn Glu Gln Val Met Glu Arg  
 35 40 45

Asn Met Ser Lys Arg Arg Lys His Ser Trp Pro Gln Pro Leu Lys Gly  
 50 55 60

60 Ala Glu Ser Arg Leu Trp Tyr Gly Gly Asp Tyr Asn Pro Asp Gln Trp

65

# ES 2 342 125 T3

|    | 65                          | 70                  | 75                          | 80          |
|----|-----------------------------|---------------------|-----------------------------|-------------|
| 5  | Pro Glu Glu Val             | Trp Asp Asp Asp Ile | Arg Leu Met Lys Lys         | Ala Gly     |
|    |                             | 85                  | 90                          | 95          |
| 10 | Val Asn Leu Val             | Ser Val Gly Ile Phe | Ser Trp Ala Lys             | Ile Glu Pro |
|    | 100                         | 105                 | 110                         |             |
|    | Glu Glu Gly Lys Tyr Asp Phe | Asp Trp Leu Asp Arg | Ala Ile Asp Lys             |             |
|    | 115                         | 120                 | 125                         |             |
| 15 | Leu Gly Lys Ala Gly Ile     | Ala Val Asp Leu Ala | Ser Ala Thr Ala Ser         |             |
|    | 130                         | 135                 | 140                         |             |
| 20 | Pro Pro Met Trp Leu Thr     | Gln Ala His Pro Glu | Val Leu Trp Lys Asp         |             |
|    | 145                         | 150                 | 155                         | 160         |
|    | Glu Arg Gly Asp Thr         | Val Trp Pro Gly     | Ala Arg Glu His Trp         | Arg Pro     |
|    | 165                         | 170                 | 175                         |             |
| 25 | Thr Ser Pro Val Phe         | Arg Glu Tyr Ala     | Leu Asn Leu Cys             | Arg Arg Met |
|    | 180                         | 185                 | 190                         |             |
|    | Ala Glu His Tyr Lys Gly     | Asn Pro Tyr Val     | Val Ala Trp His             | Val Ser     |
|    | 195                         | 200                 | 205                         |             |
| 30 | Asn Glu Tyr Gly Cys His     | Asn Arg Phe Asp Tyr | Ser Asp Asp Ala Met         |             |
|    | 210                         | 215                 | 220                         |             |
| 35 | Arg Ala Phe Gln Lys Trp     | Cys Lys Lys Arg Tyr | Lys Thr Ile Asp Ala         |             |
|    | 225                         | 230                 | 235                         | 240         |
|    | Val Asn Glu Ala Trp         | Gly Thr Ala Phe     | Trp Ala Gln His Met         | Asn Asp     |
|    | 245                         | 250                 | 255                         |             |
| 40 | Phe Ser Glu Ile Ile         | Pro Pro Arg Tyr     | Ile Gly Asp Gly             | Asn Phe Met |
|    | 260                         | 265                 | 270                         |             |
| 45 | Asn Pro Gly Lys Leu Leu     | Asp Tyr Lys Arg Phe | Ser Ser Asp Ala Leu         |             |
|    | 275                         | 280                 | 285                         |             |
|    | Lys Glu Leu Tyr Ile         | Ala Glu Arg Asp Val | Leu Glu Ser Ile Thr Pro     |             |
|    | 290                         | 295                 | 300                         |             |
| 50 | Gly Leu Pro Leu Thr         | Thr Asn Phe Met Val | Ser Ala Gly Gly Ser Met     |             |
|    | 305                         | 310                 | 315                         | 320         |
| 55 | Leu Asp Tyr Asp Asp         | Trp Gly Ala Glu Val | Asp Phe Val Ser             | Asn Asp     |
|    | 325                         | 330                 | 335                         |             |
|    | His Tyr Phe Thr Pro         | Gly Glu Asp His     | Phe Asp Glu Val Ala Tyr Ala |             |

# ES 2 342 125 T3

|    | 340                         | 345                 | 350                     |
|----|-----------------------------|---------------------|-------------------------|
| 5  | Ala Ser Leu Met Asp Gly Ile | Ser Arg Lys Glu Pro | Trp Phe Gln Met         |
|    | 355                         | 360                 | 365                     |
|    | Glu His Ser Thr Ser Ala Val | Asn Trp Arg Pro     | Ile Asn Tyr Arg Ala     |
|    | 370                         | 375                 | 380                     |
| 10 | Glu Pro Gly Ser Val Val     | Arg Asp Ser Leu Ala | Gln Val Ala Met Gly     |
|    | 385                         | 390                 | 395                     |
|    | Ala Asp Ala Ile Cys Tyr Phe | Gln Trp Arg Gln     | Ser Lys Ala Gly Ala     |
|    | 405                         | 410                 | 415                     |
|    | Glu Lys Trp His Ser Ser Met | Val Pro His Ala Gly | Glu Asp Ser Gln         |
|    | 420                         | 425                 | 430                     |
| 20 | Ile Phe Arg Asp Val Cys Glu | Leu Gly Ala Asp     | Leu Gly Arg Leu Ser     |
|    | 435                         | 440                 | 445                     |
|    | Asp Glu Gly Leu Met Gly Thr | Lys Thr Val Lys     | Ser Lys Val Ala Val     |
|    | 450                         | 455                 | 460                     |
| 25 | Val Phe Asp Tyr Glu Ser Gln | Trp Ala Thr Glu     | Tyr Thr Ala Asn Pro     |
|    | 465                         | 470                 | 475                     |
|    | Thr Gln Gln Val Asp His     | Trp Thr Glu Pro     | Leu Asp Trp Phe Arg Ala |
|    | 485                         | 490                 | 495                     |
|    | Leu Ala Asp Asn Gly Ile Thr | Ala Asp Val Val     | Pro Val Arg Ser Asp     |
|    | 500                         | 505                 | 510                     |
| 35 | Trp Asp Ser Tyr Glu Ile Ala | Val Leu Pro Cys Val | Tyr Leu Leu Ser         |
|    | 515                         | 520                 | 525                     |
|    | Glu Glu Thr Ser Arg Arg Val | Arg Glu Phe Val     | Ala Asn Gly Gly Lys     |
|    | 530                         | 535                 | 540                     |
|    | Leu Phe Val Thr Tyr Tyr Thr | Gly Leu Ser Asp     | Glu Asn Asp His Ile     |
|    | 545                         | 550                 | 555                     |
| 45 | Trp Leu Gly Gly Tyr Pro Gly | Ser Ile Arg Asp     | Val Val Gly Val Arg     |
|    | 565                         | 570                 | 575                     |
|    | Val Glu Glu Phe Ala Pro Met | Gly Asn Asp Met     | Pro Gly Ala Leu Asp     |
|    | 580                         | 585                 | 590                     |
| 50 | His Leu Asp Leu Asp Asn Gly | Thr Val Ala His     | Asp Phe Ala Asp Val     |
|    | 595                         | 600                 | 605                     |
| 55 | Ile Thr Ser Thr Ala Asp Thr | Ser Thr Val Leu     | Ala Ser Tyr Lys Ala     |

# ES 2 342 125 T3

|    | 610                                                                                | 615 | 620 |
|----|------------------------------------------------------------------------------------|-----|-----|
| 5  | Glu Arg Trp Thr Gly Met Asn Glu Val Pro Ala Ile Val Ala Asn Gly<br>625 630 635 640 |     |     |
|    | Tyr Gly Asp Gly Arg Thr Val Tyr Val Gly Cys Arg Leu Gly Arg Gln<br>645 650 655     |     |     |
| 10 | Gly Leu Ala Lys Ser Leu Pro Ala Met Leu Gly Ser Met Gly Leu Ser<br>660 665 670     |     |     |
| 15 | Asp Leu Ala Gly Asp Gly Arg Val Leu Arg Val Glu Arg Ala Asp Ala<br>675 680 685     |     |     |
|    | Ala Ala Ala Ser Arg Phe Glu Phe Val Phe Asn Arg Thr His Glu Pro<br>690 695 700     |     |     |
| 20 | Val Thr Val Asp Val Glu Gly Glu Ala Ile Ala Ala Ser Leu Ala His<br>705 710 715 720 |     |     |
| 25 | Val Asp Asp Gly Arg Ala Thr Ile Asp Pro Thr Gly Val Val Val Leu<br>725 730 735     |     |     |
|    | Arg Arg                                                                            |     |     |