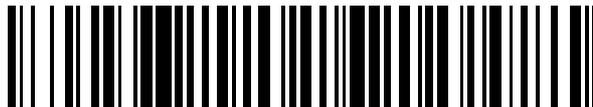


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 456 292**

51 Int. Cl.:

C12P 19/18 (2006.01)

C12P 19/26 (2006.01)

C12N 1/21 (2006.01)

C12N 9/04 (2006.01)

C12N 9/10 (2006.01)

C12N 9/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.03.2007 E 07726672 (4)**

97 Fecha y número de publicación de la concesión europea: **15.01.2014 EP 1991690**

54 Título: **Procedimiento de producción de oligosacáridos sialilados**

30 Prioridad:

09.03.2006 US 780350 P

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

21.04.2014

73 Titular/es:

**CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (100.0%)
3, RUE MICHEL-ANGE
75794 PARIS CEDEX 16, FR**

72 Inventor/es:

SAMAIN, ERIC

74 Agente/Representante:

CURELL AGUILÁ, Mireia

ES 2 456 292 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

Procedimiento de producción de oligosacáridos sialilados.

5 **Campo de la invención**

La presente invención se refiere a un procedimiento para la síntesis *in vivo* a gran escala de oligosacáridos sialilados, al cultivo de un microorganismo en un medio de cultivo, que comprende opcionalmente un precursor exógeno tal como la lactosa, en el que dicho microorganismo comprende unos genes heterólogos que codifican una
10 CMP-Neu5Ac sintetasa, una ácido siálico sintasa, una GlcNAc-6-fosfato 2-epimerasa y una sialiltransferasa, y en el que los genes endógenos que codifican la ácido siálico aldolasa (*NanA*) y ManNAc cinasa (*NanK*) se han suprimido o inactivado. La invención se refiere asimismo a este microorganismo que puede producir ácido siálico activado internamente.

15 **Antecedentes de la invención**

El ácido N-acetilneuramínico (Neu5Ac) es el miembro más corriente de la familia del ácido siálico de aminoazúcares. El Neu5Ac se encuentra con frecuencia como un azúcar terminal en los hidratos de carbono complejos de la superficie celular y ejerce una función importante en muchos procesos biológicos tales como la adhesión celular y la
20 unión de las toxinas y los virus de la (Varki, 1993). El Neu5Ac es también un componente importante de la fracción de hidrocarbonatada de los gangliósidos que son especialmente abundantes en el tejido cerebral y están involucrados en varias patologías (Zhang y Kiechle, 2004).

En razón de sus importantes funciones biológicas, los oligosacáridos que contienen ácido siálico han atraído considerable interés y se han desarrollado muchos procedimientos para sintetizar estas estructuras para la investigación fundamental y potenciales aplicaciones terapéuticas. Sin embargo, la producción a gran escala de oligosacáridos sialilados no se ha conseguido hasta la actualidad.

Las síntesis químicas no son prácticas debido a las múltiples etapas de protección y desprotección y se han dedicado muchos esfuerzos a los procedimientos enzimáticos y biotecnológicos. Las sialiltransferasas utilizan CMP-Neu5Ac como azúcar-nucleótido activado y el desarrollo de procesos eficientes para las síntesis enzimáticas de sialiloligosacáridos ha sido posible mediante la identificación de los genes de sialiltransferasa bacteriana que están bien expresados en *E. coli* y el diseño de múltiples sistemas enzimáticos que imitan la ruta natural de la biosíntesis de azúcar-nucleótidos (Gilbert *et al.*, 1998).
35

Una mejora significativa se produjo a partir de la utilización de células bacterianas vivas para producir sialiloligosacáridos (Priem *et al.*, 2002). En este enfoque, se produjo directamente sialil-lactosa cultivando células de cepas de *Escherichia coli* metabólicamente modificadas que sobreexpresan los genes de *Neisseria meningitidis* para la α -2,3-sialiltransferasa y para CMP-Neu5Ac sintasa. Las bacterias se cultivaron a alta densidad celular con glicerol como fuente de carbono y energía, mientras que se suministraron lactosa exógena y Neu5Ac como precursores para la síntesis de sialil-lactosa. Durante el cultivo, la lactosa y Neu5Ac fueron interiorizadas activamente por β -galactósido de *E. coli* y Neu5Ac permeasas. Para evitar el catabolismo de la lactosa y de Neu5Ac, se utilizaron cepas mutantes desprovistas de actividades de β -galactosidasa y Neu5Ac aldolasa. La lactosa y Neu5Ac acumuladas en el citoplasma donde Neu5Ac se convirtió después en CMP-Neu5Ac para ser transferido a continuación a la lactosa para formar sialil-lactosa (patente europea EP 1194584 de los inventores). Este sistema se aplicó a la producción del fragmento carbohidratado de los gangliósidos GM2 y GM1 mediante la expresión adicional de los genes de glucosiltransferasa apropiados (Antoine *et al.*, 2003). Los oligosacáridos polisialilados (azúcares de GD3 y GT3) también se produjeron por este procedimiento y con el gen *cstII* de *Campylobacter* que codifica una α -2,3- y α -2,8-sialiltransferasa bifuncional (solicitud US 60/690.837 de los inventores y Antoine *et al.*, 2005).
45
50

La producción a gran escala de sialiloligosacáridos por este procedimiento microbiológico requiere cantidad importante de ácido siálico como precursor. El ácido siálico se puede purificar de fuentes naturales tales como la leche y la yema de huevo, pero los rendimientos son bajos y el procedimiento no es adecuado para producción a gran escala. El ácido siálico se prepara generalmente por síntesis enzimática mediante la ácido siálico aldolasa utilizando N-acetilmanosamina (ManNAc) y piruvato como sustrato. Para reducir el coste, se prepara normalmente ManNAc por epimerización química o enzimática de N-acetilglucosamina, que es un sustrato más económico que ManNAc (Lee *et al.*, 2004; Maru *et al.*, 1998.). A pesar de estas mejoras, el coste de ácido siálico todavía es relativamente alto y este coste obstaculiza el desarrollo de un sistema económico para la producción de sialiloligosacáridos.
55
60

Antoine *et al.* (*Chembiochem.*, vol. 4, nº 5, 9 de mayo de 2003, págs. 406-412) y Priem *et al.* (*Glycobiology*, vol. 12, nº 4, oct. 2002, págs. 235-240) describen la producción de azúcares sialilados de lactosa y Neu5Ac como precursores externos. La patente EP 1 484 406 describe la producción de Neu5Ac en *E. coli* que sobreexpresa la GlcNAc-2-epimerasa y la Neu5Ac sintasa.
65

Hamamoto Tomoki *et al.* (*Biotechnology and Biochemistry* oct. 2005, vol. 69, nº 10, págs 1944-1950) dan a conocer

la síntesis enzimática de CMP-Neu5Ac a partir de GlcNAc y CMP utilizando *E. Coli* que comprende una GlcNAc 6-P 2-epimerasa heteróloga y una NeuAc sintasa heteróloga en combinación con células de levadura y CMP-NeuAc sintetasa de *H. influenza* purificada.

5 El documento WO 2005/090552 se refiere a la producción de glucoproteínas sialiladas en levaduras y Ends *et al.* (*Applied Microbiology and Biotechnology*, Springer Verlag, Berlín, vol. 53, nº 3, marzo de 2000, páginas 257-261) describe la producción de sialiloligosacáridos por combinación de dos cepas diferentes de *E. coli*.

10 También, cepas similares a K1 de *E. coli* y *N. meningitidis* son capaces de producir CMP-Neu5Ac, pero son patógenas y no se pueden utilizar en procesos biotecnológicos por razones de seguridad. La mayoría de las demás bacterias, como por ejemplo K12 de *E. coli*, no tienen la organización enzimática para la biosíntesis de CMP-Neu5Ac, y es un objetivo de la invención modificar genéticamente las cepas no patógenas que serían capaces de producir CMP-Neu5Ac a partir de UDP-GlcNAc endógena.

15 En relación con la presente invención, se ha diseñado un nuevo sistema microbiano para la producción rentable a gran escala de sialiloligosacáridos sin necesidad de un suministro exógeno de ácido siálico. Los microorganismos modificados metabólicamente de la invención son viables, no patógenos y se pueden utilizar en procesos de cultivo a gran escala e industriales. Han optimizado rutas modificadas y la supresión de los ciclos metabólicos inútiles y han conducido a la biosíntesis de CMP-Neu5Ac activada que actúa como donante de ácido siálico *in situ* para formar oligosacáridos sialilados.

Breve resumen de la invención

25 La presente invención proporciona un procedimiento de producción de oligosacáridos sialilados por cultivo fermentativo de microorganismos. En particular, la invención se refiere a un procedimiento de síntesis de oligosacáridos portadores de uno o varios resto(s) de ácido siálico, sin adición de ácido siálico exógeno al medio de cultivo, que comprenden de manera no limitativa:

- restos de oligosacáridos de los gangliósidos seleccionados de entre

30 **GM3 (3'sialil-lactosa**, Neu5Ac α -3Gal β -4Glc) y oligosacáridos que comprende el motivo GM3,
GD3 Neu5Ac α -8Neu5Ac α -3Gal β -4Glc,
GT3 (Neu5Ac α -8Neu5Ac α -8Neu5Ac α -3Gal β -4Glc);
GM2 GalNAc β -4 (Neu5Ac α -3) Gal β -4Glc,
35 **GM1** Gal β -3GalNAc β -4 (Neu5Ac α -3)Gal β - 4Glc,
GD1a Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc
GT1a Neu5Ac α -8Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc
GD2 GalNAc β -4(Neu5Ac α -8Neu5Ac α 3) Gal β -4Glc
GT2 GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α 3) Gal β -4Glc
40 **GD1b**, Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc
GT1b Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc
GQ1b Neu5Ac α -8Neu5Ac α -3Gal β - 3GalNAc β -4(Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc
GT1c Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc
GQ1c Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc
45 **GP1c** Neu5Ac α -8Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α 3) Gal β -4Glc
GD1 α Neu5Ac α -3Gal β -3 (Neu5Ac α -6) GalNAc β -4Gal β -4Glc
Fucosil-GM1 Fuc α -2Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc;

50 que se en su totalidad se pueden extender a la producción de los gangliósidos correspondientes por reacción de los restos de oligosacáridos anteriores con ceramida.

- Otros azúcares sialilados como por ejemplo:

55 - **6'sialil-lactosa** (Neu5Ac α -6Gal β -4Glc) y oligosacáridos que comprenden 6'sialil-lactosa
- **Hexasacárido SGG** (Neu5Ac α -3Gal β -3GalNAc β -3Gal α -4Gal β -4Gal)
- **tetrasacárido sialilado** (Neu5Ac α -3Gal β -4GlcNAc β -4GlcNAc)
- **pentasacárido LST_D** (Neu5Ac α -3Gal β -4GlcNAc β -3Gal β -4Glc)

60 En un aspecto particular, el procedimiento de la invención se basa en la absorción activa de un precursor exógeno, tal como por ejemplo un mono-, di- o tri-sacárido, más particularmente un precursor exógeno seleccionado de entre lactosa, galactosa, β -galactósido y α -galactosidasa tal como globotriosa (Gal α -4Gal β -4Glc), mientras que las células están creciendo en un sustrato de carbono alternativo, tal como glicerol o glucosa. La expresión "precursor exógeno" hace referencia a un compuesto implicado en la ruta biosintética de los oligosacáridos según la invención que es interiorizada por las células.

65 También proporciona microorganismos metabólicamente modificados que pueden producir específicamente los

oligosacáridos sialilados anteriores sin subproductos tales como GA1, GA2, GA3, GA4 y GA5 y la utilización del gen *cstIII* aislado de cepas *C. jejuni* que expresan estructuras de lipooligosacáridos que imitan al gangliósido GM1, tal como la cepa de NCTC de *C. jejuni* nº de registro 11168, para la producción específica de GM1.

5 Descripción de los dibujos

La figura 1 muestra la producción de sialil-lactosa por la cepa de *E. coli* metabólicamente modificada.

La figura 2 muestra la formación de subproductos durante la producción de azúcar de GM1 por cepas de *E. coli* metabólicamente modificadas.

La figura 3 muestra la formación de subproductos durante la biosíntesis del azúcar de GM2.

La figura 4 muestra la ruta biosintética UDP-Gal.

La figura 5 muestra la formación de subproductos durante la biosíntesis del azúcar de GD2.

La figura 6 muestra la estrategia para la producción de GT1a utilizando sialil transferasa.

La figura 7 muestra la ruta metabólicamente modificada para la producción de azúcar LST_D (Neu5Ac α -3Gal β -4GlcNac β -3 β -4Gal) a partir de lactosa exógena.

La figura 8 muestra la ruta metabólicamente modificada para la producción del hexasacárido SGG (Neu5Ac α -3Gal β -3GalNac β -3Gal α -4Gal β -4Gal) a partir de globotriosa exógena.

La figura 9 muestra la ruta metabólicamente modificada para la producción del tetrasacárido sialilado (Neu5Ac α -3Gal β -4GlcNac α -4GlcNAc).

La figura 10 es un análisis por TLC de fracción intracelular y extracelular de cultivo de alta densidad celular de la cepa DC7 con una alimentación continua de lactosa. Banda 1: Solución patrón (2 mg·ml⁻¹ cada una) de lactosa, lacto-*N*-neotetraosa (LNnT), lacto-*N*-neohexosa (LNnH). Bandas 2, 3, 4, 5, 6 y 7: fracciones intracelulares retiradas 0, 7, 23, 31, 47 y 54 horas después de la adición de lactosa. Bandas 8, 9, 10, 11, 12 y 13: fracciones extracelulares retiradas 0, 7, 23, 31, 47 y 54 horas después de la adición de lactosa. Sialilactosa (2) se ha demostrado previamente que migra como tetrasacárido LNnT.

La figura 11 muestra la producción de sialil-lactosa en cultivo de alta densidad celular de la cepa DC7 con una alimentación continua de la lactosa. (▲) cantidad acumulada de lactosa añadida, (□) sialil-lactosa intracelular, (■) sialil-lactosa extracelular, (-) cultivo bacteriano.

La figura 12 es un análisis por TLC de oligosacáridos producidos por cultivo de alta densidad celular de la cepa NF03 que contiene los plásmidos pUC18-cstII y pBBR3-SS. La concentración inicial de lactosa era de 3 g·l⁻¹. Banda 1: solución patrón (2 mg·ml⁻¹ cada una) de lactosa lacto-*N*-neotetraosa (LNnT), lacto-*N*-neohexosa (LNnH). Bandas 2 y 3: fracciones intracelulares retiradas 7 y 24 horas después de la adición de lactosa. Bandas 4 y 5: fracciones extracelulares retiradas 7 y 24 horas después de la adición de lactosa.

La figura 13 es un análisis por TLC de oligosacáridos producidos por cultivo de alta densidad celular de la cepa DC15 (pBS-cgtAII-nst, pBBR3-SS-wbpP, pSU-cgtB). La concentración inicial de lactosa fue de 5 g·l⁻¹. Banda 1: solución patrón (2 mg·ml⁻¹ cada una) de lactosa lacto-*N*-neotetraosa (LNnT), lacto-*N*-neohexosa (LNnH) Bandas 2, 3, 4, 5: fracciones intracelulares retiradas 7, 20, 30 y 44 horas después de la adición de lactosa. Bandas 6, 7, 8, 9: fracciones extracelulares retiradas 7, 20, 30 y 44 horas después de la adición de lactosa. Sialilactosa (1) y el azúcar de GM1 (3) migran como LNnT y LNnH respectivamente.

La figura 14 es un análisis por TLC de oligosacáridos producidos por cultivo de alta densidad celular de la cepa DC21 (pBS-cgtAII-nst, pBBR3-SS-gne, pSU-cgtB). La concentración inicial de lactosa fue de 5 g·l⁻¹. Bandas 1 y 10: solución patrón (2 mg·ml⁻¹ cada una) de lactosa lacto-*N*-neotetraosa (LNnT), lacto-*N*-neohexosa (LNnH). Bandas: 2, 3, 4, 5: fracciones intracelulares retiradas 5, 20, 28 y 44 horas después de la adición de lactosa. Bandas: 6, 7, 8, 9: fracciones extracelulares retiradas 5, 20, 28 y 44 horas después de la adición de lactosa. Sialilactosa (1) y el azúcar de GM1 (3) migran como LNnT y LNnH respectivamente.

La figura 15 es un análisis por TLC de oligosacáridos producidos por cultivo de alta densidad celular de la cepa DC22. (pBS-cstIII-cgtAII, pBBR3-SS-gne, PSU-cgtB) La concentración inicial de lactosa fue de 5 g·l⁻¹. Bandas 1 y 10: solución patrón (2 mg·ml⁻¹ cada una) de lactosa lacto-*N*-neotetraosa (LNnT), lacto-*N*-neohexosa (LNnH) Bandas 2, 3, 4, 5: fracciones intracelulares retiradas 7, 20, 30 y 44 horas después de la adición de lactosa. Bandas: 6, 7, 8, 9: fracciones extracelulares retiradas 7, 20, 30 y 44 horas después de la adición de lactosa. Sialilactosa (1) y el azúcar de GM1 (3) migran como LNnT y LNnH respectivamente.

La figura 16 es un análisis por TLC de oligosacáridos producidos por cultivo de alta densidad celular de la cepa ZWT (ZLKA, pBS-nst, pBBR3-SS-gne, pWKS-cgtAll). La concentración inicial de lactosa fue de $5 \text{ g}\cdot\text{l}^{-1}$. Banda 1: solución patrón ($2 \text{ mg}\cdot\text{ml}^{-1}$ cada una) de lactosa, lacto-*N*-neotetraosa (LNnT), lacto-*N*-neohexosa (LNnH). Bandas 2, 3, 4: fracciones intracelulares retiradas 0, 7 y 22 horas después de la adición de lactosa. Bandas 5, 6, 7: fracciones extracelulares retiradas 0, 7, y 22 horas después de la adición de lactosa.

La figura 17 es un análisis por TLC de oligosacáridos producidos por cultivo de alta densidad celular de la cepa ZWU (ZWU, pBS-nst, pBBR3-SS-gne, pWKS-cgtAll). La concentración de lactosa inicial fue de $5 \text{ g}\cdot\text{l}^{-1}$. Banda 1: Solución patrón ($2 \text{ mg}\cdot\text{ml}^{-1}$ cada una) de lactosa, lacto-*N*-neotetraosa (LNnT) lacto-*N*-neohexosa (LNnH). Bandas 2, 3, 4 y 5: fracciones intracelulares retiradas 0, 7, 22 y 30 horas después de la adición de lactosa. Bandas 6, 7, 8 y 9: fracciones extracelulares retiradas 0, 7, 22 y 30 horas después de la adición de lactosa.

La figura 18 muestra un espectro de masas ESI⁻ del compuesto de la fracción (2) purificado a partir de la fracción intracelular de la cepa de referencia ZWT (A) y el pico de la cepa ZWU2 con mutante galU (B) a m/z 835 corresponde al azúcar de GM2 y el pico a m/z 794 corresponde al análogo de galactosilada (azúcar de GA2).

La figura 19 es un análisis por TLC de fracciones obtenidas después de la separación de la fracción intracelular de un cultivo de un litro de la cepa ZWU1. La separación se llevó a cabo en Dowex 1 (forma HCO₃⁻) utilizando un gradiente de NaHCO₃ 0-1 M. El volumen de cada tubo era de 10 ml. Los rendimientos de las fracciones A, B, C y D fueron 0,5 g, 0,85, 0,6 y 1,2 g, respectivamente.

La figura 20 es un análisis por TLC de oligosacáridos producidos por cultivo de alta densidad celular de la cepa NF17 (ZLKA, pBS-cgtA-cstII, PSU-cgtB, pBBR-SS-gne). La concentración inicial de lactosa fue de $5 \text{ g}\cdot\text{l}^{-1}$ Bandas 1 y 10: solución patrón ($2 \text{ mg}\cdot\text{ml}^{-1}$ cada una) de lactosa, lacto-*N*-neotetraosa (LNnT) lacto-*N*-neohexosa (LNnH). Bandas 2, 3, 4 y 5: fracciones intracelulares retiradas 7, 22, 30 y 46 horas después de la adición de lactosa. Bandas 6, 7, 8 y 9: fracciones extracelulares retiradas 7, 22, 30 y 46 horas después de la adición de lactosa.

Descripción del origen del material genético

Tabla 1. Genes, plásmidos y cepas de *Escherichia coli* utilizados en presente invención

Genes	Descripción	Referencia o fuente
<i>nst</i>	α -2,3-sialiltransferasa de la cepa MC58 de <i>N. meningitidis</i> L3	U60660
<i>neuA</i>	CMP-Neu5Ac sintetasa de la cepa ATCC 43438 de <i>C. jejuni</i>	AF400048
<i>neuB</i>	Ácido siálico sintasa de la cepa ATCC 43438 de <i>C. jejuni</i>	AF400048
<i>neuC</i>	GlcNAc-6-fosfato 2-epimerasa de la cepa ATCC 43438 de <i>C. jejuni</i>	AF400048
<i>cgtA</i>	β -4 GalNAc transferasa de la cepa OH4384 de <i>C. jejuni</i> O:19	AF130984
<i>wbpP</i>	UDP-GlcNAc C4 epimerasa de <i>P. aeruginosa</i>	AF035937
<i>gne</i> (Cj1131c)	UDP-GlcNAc C4 epimerasa de la cepa NCTC 11168 de <i>C. jejuni</i> O:2	AL139077
<i>cstIII</i> (Cj1140)	α -2,3 sialiltransferasa de la cepa NCTC 11168 de <i>C. jejuni</i> O:2	AL139077
<i>cgtAll</i>	β -4 GalNAc transferasa de la cepa ATCC 43456 de <i>C. jejuni</i> O:36	AF401528
<i>cgtB</i> (Cj1139c)	β -3 Gal transferasa de la cepa NCTC 11168 de <i>C. jejuni</i> O:2	AL139077
<i>cstII</i>	α -2,3 α -2,8-sialiltransferasa de la cepa ATCC 43438 de <i>C. jejuni</i>	AF400048
<i>nodC</i>	N-acetilglucosaminiltransferasa de <i>Azorhizobium caulinodans</i>	AAB51164
<i>lgtB</i>	β 1,4-galactosiltransferasa de <i>Neisseria meningitidis</i>	AAC44085
<i>chiA</i>	Quitinasa de <i>Bacillus circulans</i>	AAA81528
Plásmidos		
pWKS130	Activador Plac del vector de clonación, Km ^r , bajo número de copias, replicón de pSC101	(Wang y Kushner, 1991)
pSU27-18	Activador P _{lac} de los vectores de clonación Cm ^r derivados de pACYC184	(Martínez <i>et al.</i> , 1988)
pBAD33	Activador P _{ara} de los vectores de clonación Cm ^r derivados de pACYC184	Guzmán <i>et al.</i> 1995
pBS-nst	Derivado de pBluescript II SK portador de <i>nst</i> (anteriormente denominado NST-01)	(Priem <i>et al.</i> , 2002)
pBBR1MCS-3	Activador P _{lac} del vector de clonación, Tc ^r , bajo número de copias,	(Kovach <i>et al.</i> , 1995)
pBBR3-SS	Derivado de pBBR1MCS-3 portador de <i>neuABC</i>	presente invención
pBBR3-SS-wbpP	Derivado de pBBR1MCS-3 que lleva <i>neuABC</i> y <i>wbpP</i>	presente invención
pBBR3-SS-gne	Derivado de pBBR1MCS-3 que lleva <i>neuABC</i> y <i>gne</i>	presente invención
pUC-cstII	Derivado de pUC18 que lleva <i>cstII</i>	presente invención
pBS-cgtAll-nst	Derivado de pBluescript II SK que lleva <i>cgtAll</i> y <i>nst</i>	presente invención
pBS-cgtA-cstII	Derivado de pBluescript II KS que lleva <i>cgtA</i> y <i>cstII</i>	presente invención

	Descripción	Referencia o fuente
pSU18-cgtB	Derivado de pSU27 18 que lleva <i>cgtB</i>	presente invención
pBS-cstIII-cgtAll	Derivado de pBluescript II KS que lleva <i>cstIII</i> y <i>cgtAll</i>	presente invención
pWKS-cgtAll	Derivado de pWKS130 que lleva <i>CgtAll</i> ,	presente invención
pBAD33-cgtAll	Derivado de pBAD33 que lleva <i>cgtAll</i> ,	presente invención
pWKS-igtB-chiA	Derivado de pWKS130 que lleva <i>nst</i> y <i>chiA</i> de la quitinasa	(Dumon <i>et al.</i> , 2005)
pBS-nst-nodC	Derivado de pBluescript II SK que lleva <i>nodC</i> y <i>nst</i> de <i>A. caulinodans</i>	presente invención
pBS-cstII-cgtB	Derivado de pBluescript II KS que lleva <i>cstII</i> y <i>cgtB</i>	presente invención
Cepas		
DC	DH1 <i>lacZ lacA</i>	(Dumon <i>et al.</i> , 2005)
GLK	DH1 <i>lacZ lacA galK</i>	(Dumon <i>et al.</i> , 2005)
ZLKA	DC Δ NanKETA	presente invención
AZL	DC <i>NanA</i>	presente invención
AZK	DC Δ NanK <i>NanA</i>	presente invención
ZWU	ZLKA <i>galU</i>	presente invención
GLKA	GLK Δ NanKETA	presente invención
ZW	ZLKA <i>melA wcaJ</i>	presente invención
DC6	DC (pBS-nst, pBBR3-SS)	presente invención
AW1	AZL (pBS-nst, pBBR3-SS)	presente invención
DC7	ZLKA (pBS-nst, pBBR3-SS)	presente invención
DC7	ZLKA (pBS-nst, pBBR3-SS)	presente invención
DC0	ZLKA (pBS-nst)	presente invención
AZK1	AZK (pBS-nst, pBBR3-SS)	presente invención
NF3	ZLKA (pUC-cstII, pBBR3-SS)	presente invención
DC15	ZLKA (pBS-cgtAll-nst, pBBR3-SS-wbpP, pSU-cgtB)	presente invención
DC21	ZLKA (pBS-cgtAll-nst, pBBR3-SS-gne, pSU-cgtB)	presente invención
DC22	ZLKA (pBS-cstIII-cgtAll, pBBR3-SS-gne, pSU-cgtB)	presente invención
ZWT	ZLKA (pBS-nst, pBBR-SS-gne, pWKS-cgtAll)	presente invención
ZWU2	ZWU (pBS-nst, pBBR-SS-gne, pWKS-cgtAll)	presente invención
NF08	ZLKA (pUC18-cstII, pBBR3-SS-gne, pWKS-cgtAll)	presente invención
NF09	ZLKA (pUC18-cstII, pBBR3-SS-gne, pBAD33-cgtAll)	presente invención
ZWU1	ZWU (pUC18-cstII, pBBR3-SS-gne, pBAD33-cgtAll)	presente invención
NF17	ZLKA (pBS-cgtA-cstII, pSU-cgtB, pBBR-SS-gne)	presente invención
GLK7	GLKA (pBS-nst, pBBR3-SS)	presente invención
SN4	ZLKA (pBS-nst-nodC, pBBR3-SS, pWKS-igtB-chiA)	presente invención
NF21	ZLKA (pBS-cstII-cgtB, pBBR3-SS-gne, pBAD33-cgtAll)	presente invención

Descripción detallada de la invención

5 En una primera forma de realización, la invención se refiere a un procedimiento para la producción de oligosacáridos que comprende al menos un resto de ácido siálico, denominado en la presente memoria oligosacáridos sialilados, comprendiendo el procedimiento la etapa que consiste en cultivar un microorganismo en un medio de cultivo, que comprende opcionalmente un precursor exógeno, en el que dicho microorganismo puede producir ácido siálico
10 activado internamente y comprende genes heterólogos que codifican una CMP-Neu5Ac sintetasa, una ácido siálico sintasa, una GlcNAc-6-fosfato 2 epimerasa y una sialiltransferasa, y en el que los genes endógenos que codifican la ácido siálico aldolasa (*NanA*) y la ManNAc cinasa (*NanK*) se han suprimido o inactivado.

En el procedimiento anterior, y en función del punto final, el gen de la sialiltransferasa heteróloga puede seleccionarse de entre α -2,3-sialiltransferasa, por ejemplo, codificarse por *nst*, α -2,3 α -2,8-sialiltransferasa (*cstII*), y α -2,3-sialiltransferasa (*cstIII*) o α -2,6-sialil transferasa. La CMP-Neu5Ac sintetasa heteróloga puede ser *neuA*, la ácido siálico sintasa heteróloga puede ser *neuB*, y la GlcNAc-6-fosfato 2-epimerasa heteróloga puede ser *neuC*. Los genes *neuA*, *neuB* y *neuC* se pueden aislar de cepas bacterianas que contienen estructura sialilada en su membrana celular, tales como la cepa de *C. jejuni* n° de registro 43438 de la ATCC.
15

Los genes *nanT*, *NanA*, *NanK* y *nanE* forman parte del mismo operón, que está regulado por la proteína *nanR* de unión al ADN e inducido por Neu5Ac (Kalivoda *et al.*, 2003). Por lo tanto, los microorganismos de la invención también pueden ser *NanKEAT*-. La producción de Neu5Ac como producto intermedio durante la síntesis de CMP-Neu5Ac por la cepa genéticamente modificada que sobreexpresa los genes *neuBCA* puede inducir por lo tanto la ruta del catabolismo de ácido siálico y crear dos ciclos inútiles que reduzcan la capacidad de biosíntesis de CMP-Neu5Ac de las bacterias. Un primer ciclo inútil puede ser resultado de la actividad combinada de la ácido siálico sintasa *NeuB* con la ácido siálico aldolasa *NanA*. Un segundo ciclo inútil puede ser resultado de la acción combinada
20
25

de la UDP-GlcNAc 2 epimerasa NeuC con las cuatro enzimas NanK, NanE, NagA GlmM y GlmU que catalizan la formación de UDP-GlcNAc de ManNAc. Según el procedimiento propuesto en la presente memoria, se evita la degradación de Neu5Ac y ManNAc. Esto se puede hacer mejor destruyendo los genes *NanA* y *NanK* en las cepas que se utilizarán para la producción de sialiloligosacáridos. En una forma de realización específica, se suprimen o inactivan los genes *nanT*, *NanA*, *NanK* y *nanE*. Esto puede llevarse a la práctica, por ejemplo, extrayendo todo el operón.

En una forma de realización preferida, el microorganismo anterior codifica una proteína que facilita la absorción de la lactosa y carece de enzimas que metabolizan la lactosa. Por ejemplo, en *E. coli*, la célula es preferentemente *LacY+* (β -galactósido permeasa), *LacZ-* (β galactosidasa), y opcionalmente *MeIA-* (α -galactosidasa).

En otra forma de realización preferida, el medio comprende un precursor exógeno que se selecciona por ejemplo de entre lactosa, galactosa, β -galactósido y α -galactósido, tal como globotriosa (Gal α -4Gal β -4Glc).

La invención también se refiere al microorganismo anterior y a un medio de cultivo celular que comprende el microorganismo anterior y un precursor exógeno seleccionado de entre lactosa, galactosa, β -galactósido y α -galactósido, tal como globotriosa (Gal α -4Gal β -4Glc).

Definiciones

La expresión "ácido siálico" se refiere a cualquier miembro de una familia de azúcares carboxilados de nueve carbonos. El miembro más frecuente de la familia del ácido siálico es N-acetil-neuramínico (2-ceto-5-acetamido-3,5-didesoxi-D-glicero-D-galactononulopiranos-1-ónico (a menudo abreviado como Neu5Ac, NeuAc o NANA). Un segundo miembro de la familia es el ácido N-glicolil-neuramínico (Neu5Gc o NeuGc), en el que el grupo N-acetilo de Neu5Ac está hidroxilado. Un tercer miembro de la familia del ácido siálico es el ácido 2-ceto-3-desoxi-nonulosónico (KDN) (Nadano *et al.* (1986) *J. Biol. Chem.* 261: 11550-11557; Kanamori *et al.*, *J. Biol. Chem.* 265: 21811-21819 (1990)). También están comprendidos los ácidos siálicos sustituidos en 9 como O-9-acil C₁-C₆-Neu5Ac como 9-O-lactil-Neu5Ac o 9-O-acetil-Neu5Ac, 9-desoxi-9-fluoro-Neu5Ac y 9-azido-9-desoxi-Neu5. Para examen de la familia del ácido siálico, véase, por ejemplo, Varki, *Glycobiology* 2: 25-40 (1992); *Sialic Acids: Chemistry, Metabolism and Function*, R. Schauer, Ed. (Springer-Verlag., Nueva York (1992)). La síntesis y utilización de compuestos de ácido siálico en un procedimiento de sialilación se da a conocer en la solicitud internacional WO 92/16640, publicada el 1 de octubre de 1992.

La expresión "sialiltransferasa CstII bifuncional de *Campylobacter jejuni*" se refiere a una sialiltransferasa que presenta actividades tanto de α -2,3 como de α -2,8-sialiltransferasa. En algunas formas de realización se utiliza la sialiltransferasa CstII con el nº de registro 43438 de la ATCC.

Un "sustrato aceptor" o un "sacárido aceptor" para una glucosiltransferasa es un resto de oligosacárido que puede actuar como un aceptor para una determinada glucosiltransferasa. Cuando el sustrato aceptor se pone en contacto con la glucosiltransferasa correspondiente y el sustrato donante de azúcar, y otros componentes necesarios de la mezcla de reacción, y la mezcla de reacción se incuba durante un período de tiempo suficiente, la glucosiltransferasa transfiere restos de azúcar desde el sustrato donante de azúcar al sustrato aceptor. Por ejemplo, un sustrato aceptor para las sialiltransferasas utilizadas en los procedimientos de la invención es la lactosa Gal β 1,4-Glc.

Un "sustrato donante" para glucosiltransferasas es un azúcar de nucleótido activado. Dichos azúcares activados constan generalmente de derivados de monofosfato de uridina, guanosina, citidina de los azúcares (UMP, GMP y CMP, respectivamente) o derivados de difosfato de los azúcares (UDP, GDP y CDP, respectivamente) en los que el monofosfato o difosfato de nucleósido actúa como grupo saliente. Por ejemplo, un sustrato donante para sialiltransferasas utilizado en los procedimientos de la invención es CMP-Neu5Ac.

Un "medio de cultivo" se refiere a cualquier medio líquido, semisólido o sólido que se puede utilizar para soportar el cultivo de un microorganismo utilizado en los procedimientos de la invención. En algunas formas de realización, el microorganismo es una bacteria, por ejemplo, *E. coli*. Los medios para el cultivo de microorganismos son bien conocidos, véase, por ejemplo, Sambrook *et al.* y *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., *Current Protocols*, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (suplemento de 1998) (Ausubel). Los medios pueden ser medios ricos, por ejemplo, caldo de cultivo Luria o caldo Terrific, o medio sintético o semisintético, por ejemplo, medio M9. En algunas formas de realización preferidas, el medio de cultivo comprende lactosa y ácido siálico.

"Escala comercial" se refiere a producción en balanza gram de un sacárido producto sialilado en una sola reacción. En las formas de realización preferidas, a escala comercial se refiere a la producción superior a aproximadamente 50, 75, 80, 90 o 100, 125, 150, 175 o 200 gramos.

La expresión "operativamente unido" se refiere al enlace funcional entre una secuencia de control de expresión de ácido nucleico (tal como un activador, secuencia señal o una matriz de puntos de fijación del factor de transcripción)

y una segunda secuencia de ácido nucleico, en el que la secuencia de control de la expresión afecta a la transcripción y/o la traducción del ácido nucleico correspondiente a la segunda secuencia.

Un "polinucleótido heterólogo" o un "gen heterólogo", como se utiliza en la presente memoria, es el que se origina a partir de una fuente extraña a la célula anfitriona específica, o, si de la misma fuente, se modifica a partir de su forma original. Por lo tanto, un gen de sialiltransferasa heterólogo en una célula comprende un gen que es endógeno a la célula anfitriona específica pero que ha sido modificado. La modificación de la secuencia heteróloga se puede producir, por ejemplo, por tratamiento del ADN con una enzima de restricción para generar un fragmento de ADN que puede estar operativamente unido a un activador. Técnicas tales como la mutagenia dirigida también son útiles para modificar una secuencia heteróloga.

Un "casete de expresión recombinante" o simplemente un "casete de expresión" es un montaje de ácido nucleico, generado por recombinación o síntesis, con elementos de ácido nucleico que son capaces de afectar la expresión de un gen estructural en anfitriones compatibles con dichas secuencias. Los casetes de expresión comprenden por lo menos activadores y opcionalmente, señales de terminación de la transcripción. Por lo general, el casete de expresión recombinante comprende un ácido nucleico que se transcribe (por ejemplo, un ácido nucleico que codifica un polipéptido deseado), y un activador. También pueden utilizarse factores adicionales necesarios o útiles para efectuar la expresión. Las señales de terminación de la transcripción, potenciadores, y otras secuencias de ácidos nucleicos que influyen en la expresión génica, también pueden incluirse en un casete de expresión. Cuando en un microorganismo se expresa más de una proteína heteróloga, los genes que codifican las proteínas se pueden expresar en un solo casete de expresión o en múltiples casetes de expresión que son compatibles y se pueden mantener en la misma célula. Como se utiliza en la presente memoria, casete de expresión también comprende montajes de ácido nucleico que se insertan en el cromosoma del microorganismo anfitrión. Los expertos en la materia apreciarán que se puede producir la inserción de un ácido nucleico en un cromosoma, por ejemplo, por recombinación homóloga. Un casete de expresión se puede construir para la producción de más de una proteína. Las proteínas pueden ser reguladas por una única secuencia activadora, como por ejemplo, un operón. O varias proteínas pueden ser codificadas por ácidos nucleicos con activadores y puntos de fijación a ribosomas individuales.

El término "aislado" se refiere a material que está sustancial o esencialmente exento de componentes que interfieren con la actividad biológica de la molécula. Para las células, sacáridos, ácidos nucleicos y polipéptidos, el término "aislado" se refiere a material que está sustancial o esencialmente exento de componentes que normalmente acompañan al material tal como se encuentra en su estado natural. Por lo general, los sacáridos, oligosacáridos, proteínas o ácidos nucleicos aislados de la invención tienen al menos aproximadamente 50%, 55%, 60%, 65%, 70%, 75%, 80% o 85% de pureza, habitualmente por lo menos aproximadamente 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de pureza medida por intensidad de la banda en un gel teñido con plata u otro método para determinar la pureza. La pureza o la homogeneidad puede indicarse por numerosos medios bien conocidos en la técnica, tales como electroforesis en gel de poliacrilamida de una proteína o muestra de ácido nucleico, seguido por la visualización en la tinción. Para ciertos fines se necesitará alta resolución y HPLC o un medio similar para la purificación utilizada. Para oligosacáridos, por ejemplo, productos sialilados, la pureza se puede determinar utilizando, por ejemplo, cromatografía en capa fina, HPLC, o espectroscopia de masas.

Los términos "idéntica" o "identidad" por ciento en el contexto de dos o más secuencias de ácidos nucleicos o de polipéptidos, se refieren a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de restos de aminoácidos o nucleótidos que son iguales, cuando se comparan y se alinean para máxima correspondencia, medidos utilizando uno de los siguientes algoritmos de comparación de secuencias o por inspección visual.

La expresión "sustancialmente idénticos", en el contexto de dos ácidos nucleicos o polipéptidos, se refiere a dos o más secuencias o subsecuencias que tienen por lo menos 60%, preferentemente 80% u 85%, más preferentemente al menos 90%, 91%, 92 %, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de restos de nucleótidos o de aminoácidos, cuando se comparan y alinean para una correspondencia máxima, medida utilizando uno de los siguientes algoritmos de comparación de secuencia o por inspección visual. Preferentemente, la identidad sustancial se encuentra en una región de las secuencias que tiene por lo menos aproximadamente 50 restos de longitud, más preferentemente en una región de por lo menos aproximadamente 100 restos, y aún más preferentemente las secuencias son sustancialmente idénticas en por lo menos aproximadamente 150 restos. En una forma de realización más preferida, las secuencias son sustancialmente idénticas en toda la longitud de las regiones codificadoras.

Para la comparación de secuencias, por lo general una secuencia actúa como secuencia de referencia, con la que se comparan las secuencias de prueba. Cuando se utiliza un algoritmo de comparación de secuencias, se introducen las secuencias de prueba y de referencia en un ordenador, se designan coordenadas de subsecuencia, si es necesario, y se designan los parámetros del programa de algoritmo de secuencia. El algoritmo de comparación de secuencias calcula entonces el porcentaje de identidad de secuencias para la(s) secuencia(s) de ensayo con relación a la secuencia de referencia, basándose en los parámetros del programa designados.

La alineación óptima de secuencias para comparación puede llevarse a cabo, por ejemplo, mediante el algoritmo de

homología local de Smith y Waterman, *Adv. Appl. Math.* 2:482 (1981), mediante el algoritmo de alineación de homologías de Needleman y Wunsch, *J. Mol. Biol.* 48:443 (1970), por la búsqueda para el procedimiento de similitud de Pearson y Lipman, *Proc. Nat'l. Acad. Sci. USA.* 85:2444 (1988), mediante ejecuciones digitales de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el paquete de software de Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante inspección visual (véase, en general, *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (Suplemento de 1995) (Ausubel)).

Los ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-410 y Altschuel *et al.* (1977) *Nucleic Acids Res.* 25: 3389-3402, respectivamente. El software para realizar análisis BLAST está disponible al público a través del National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/). Este algoritmo implica en lugar primer identificar pares de secuencias con alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia problema, que igualan o satisfacen alguna puntuación T umbral con valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de bases de datos. Se hace referencia a T como el umbral de puntuación de palabra vecina (Altschul *et al.*, anteriormente). Estos aciertos iniciales de palabras vecinas actúan como semillas para iniciar búsquedas para encontrar HSP más largos que los contienen. Los aciertos de palabras después se extienden en ambas direcciones a lo largo de cada secuencia tanto como la puntuación de alineación acumulativa se puede aumentar. Las puntuaciones acumulativas se calculan utilizando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de restos coincidentes; siempre > 0) y N (puntuación de penalización para restos no coincidentes, siempre < 0). Para secuencias de aminoácidos, se utiliza una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabras en cada dirección se detiene cuando: la puntuación de alineación acumulativa disminuye en la cantidad X desde su valor máximo alcanzado; la puntuación acumulativa llega a cero o menos, debido a la acumulación de una o más alineaciones de restos con puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros W , T y X del algoritmo BLAST determinan la sensibilidad y velocidad de alineación. El programa BLASTN (para secuencias de nucleótidos) utiliza por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, $M = 5$, $N = -4$, y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP utiliza por defecto una longitud de palabra (W) de 3, una expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, *Proc. Natl. Acad. Sci. USA.* 89:10915 (1989)).

Además de calcular el porcentaje de identidad de secuencia, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, *Proc. Natl. Acad. Sci. USA.* 90:5873-5787 (1993)). Una medida de similitud proporcionada por el algoritmo BLAST es la menor probabilidad de suma ($P(N)$), que proporciona una indicación de la probabilidad por la que una coincidencia entre dos secuencias de nucleótidos o aminoácidos se produciría por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la menor probabilidad de suma en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es inferior a aproximadamente 0,1, más preferentemente inferior a aproximadamente 0,01, y todavía más preferentemente inferior a aproximadamente 0,001.

"Variaciones modificadas de manera conservadora" de una secuencia de polinucleótidos concreta, se refiere a aquellos polinucleótidos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o cuando el polinucleótido no codifica una secuencia de aminoácidos, a secuencias esencialmente idénticas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier polipéptido dado. Por ejemplo, los codones CGU, CGC, CGA, CGG, AGA y AGG codifican todos el aminoácido arginina. Por lo tanto, en cada posición donde una arginina es especificada por un codón, el codón puede ser alterado a cualquiera de los codones correspondientes descritos sin alterar el polipéptido codificado. Dichas variaciones de ácido nucleico son "sustituciones imperceptibles" o "variaciones imperceptibles", que son una especie de "variaciones modificadas de manera conservadora." Cada secuencia de polinucleótidos descrita en la presente memoria que codifica un polipéptido también describe cada posible variación imperceptible, excepto donde se indique lo contrario. Por lo tanto, las sustituciones imperceptibles son una característica implícita de cada secuencia de ácido nucleico que codifica un aminoácido. Un experto en la materia apreciará que cada codón en un ácido nucleico (excepto AUG, que es ordinariamente el único codón para metionina) puede ser modificado para producir una molécula funcionalmente idéntica mediante técnicas normalizadas. En algunas formas de realización, las secuencias de nucleótidos que codifican las enzimas se optimizan preferentemente para la expresión en un determinado

Del mismo modo, "sustituciones conservadoras de aminoácidos", en uno o unos pocos aminoácidos en una secuencia de aminoácidos se sustituyen con diferentes aminoácidos con propiedades muy similares también se identifican fácilmente como muy similares a una secuencia de aminoácidos específica, o a una secuencia de ácido nucleico específica que codifica un aminoácido. "Sustituciones, eliminaciones o adiciones individuales que alteran, añaden o suprimen un solo aminoácido o un pequeño porcentaje de aminoácidos (por lo general menos del 5%, más por lo general menos de 1%) en una secuencia codificada son "variaciones modificadas de manera conservadora" donde las alteraciones producen la sustitución de un aminoácido por un aminoácido químicamente similar. Las tablas de sustitución conservadora que proporcionan aminoácidos funcionalmente similares son bien conocidas en la

técnica. Véase, por ejemplo, Creighton (1984) *Proteins*, W.H. Freeman and Company.

Sialiltransferasas bifuncionales

5 Como se expuso anteriormente, en los procedimientos de la invención se utilizan sialiltransferasas bifuncionales. Los ácidos nucleicos que codifican dichas enzimas se han aislado de *C. jejuni* y se dan a conocer en las patentes US nº 6.699.705, nº 6.503.744 y el documento WO/02074942. Los ejemplos de cepas de *C. jejuni* que se pueden utilizar como fuentes de sialiltransferasas bifuncionales comprenden OH4384 (las secuencias de ácidos nucleicos se encuentran en los registros nº AR271700 y nº AX934425 del GenBank), OH4382, O:10 (las secuencias de ácidos nucleicos se encuentran en los registros nº AR271701 (SEC. ID. nº 1), nº AX934427 (SEC. ID. nº 2) del GenBank, O:23, y O:41 (las secuencias de ácidos nucleicos se encuentran en los registros nº AR271702 (SEC. ID. nº 3) y nº AX934429 (SEC. ID. nº 4) del GenBank). Debe apreciarse que las variaciones modificadas de manera conservadora tal como se definen anteriormente de las SEC. ID. nº 1, 2, 3 y 4 se pueden aplicar en la presente memoria.

15 Células anfitrionas

Las células recombinantes de la invención se preparan generalmente creando u obteniendo un polinucleótido que codifica la(S) enzima(s) concretas de interés, colocando el polinucleótido en un casete de expresión bajo el control de un activador y otras señales de control apropiadas, e introduciendo el casete de expresión en una célula. Más de una de las enzimas se pueden expresar en las mismas células anfitrionas utilizando una variedad de procedimientos. Por ejemplo, un único vector extracromosómico puede comprender múltiples casetes de expresión o más de un vector extracromosómico compatible puede utilizarse para mantener un casete de expresión en una célula anfitriona. Los casetes de expresión también pueden insertarse en un cromosoma de la célula anfitriona, utilizando procedimientos conocidos por los expertos en la materia. Los expertos en la materia apreciarán que también se pueden utilizar combinaciones de casetes de expresión en vectores extracromosómicos y casetes de expresión insertados en un cromosoma de célula anfitriona. Otros modificación de la célula anfitriona, descrita con detalle a continuación, se puede realizar para aumentar la producción del oligosacárido deseado. Por ejemplo, el microorganismo puede ser LacY+ que permite el transporte activo de la lactosa. Las células anfitrionas no necesitan ser nanT+ puesto que el ácido sialílico activado se produce internamente con el procedimiento según la invención.

Las células recombinantes de la invención son generalmente microorganismos, tales como, por ejemplo, células de levadura, células bacterianas o células fúngicas. Los ejemplos de células adecuadas comprenden, por ejemplo, *Azotobacter* sp. (por ejemplo, *A. vinelandii*), *Pseudomonas* sp., *Rhizobium* sp., *Erwinia* sp., *Bacillus* sp., *Streptomyces* sp., *Escherichia* sp. (por ejemplo, *E. coli*) y *Klebsiella* sp., entre muchos otros. Las células pueden ser de cualquiera de varios géneros, como por ejemplo *Saccharomyces* (por ejemplo, *S. cerevisiae*), *Candida* (por ejemplo, *C. utilis*, *C. parapsilosis*, *C. krusei*, *C. versatilis*, *C. lipolytica*, *C. zeylanoides*, *C. guilliermondii*, *C. albicans* y *C. Humicola*), *Pichia* (por ejemplo, *P. farinosa* y *P. ohmeri*), *Torulopsis* (por ejemplo, *T. Candida*, *T. sphaerica*, *T. xylinus*, *T. famata* y *T. versatilis*), *Debaryomyces* (por ejemplo, *D. subglobosus*, *D. cantarellii*, *D. globosus*, *D. hansenii* y *D. japonicus*), *Zygosaccharomyces* (por ejemplo, *Z. rouxii* y *Z. bailii*), *Kluyveromyces* (por ejemplo, *K. marxianus*), *Hansenula* (por ejemplo, *H. anomala* y *H. jadinii*) y *Brettanomyces* (por ejemplo; *B. lambicus* y *B. anomalus*).

Los activadores para su utilización en *E. coli* incluyen los activadores T7, trp o lambda. Además se dispone de un punto de fijación al ribosoma y preferentemente una señal de terminación de la transcripción. Para la expresión de proteínas heterólogas en otras células procarionas distintas de *E. coli*, se requiere un activador que funciona en las especies procarionas específicas. Dichos activadores pueden obtenerse de genes que han sido clonados a partir de las especies, o pueden utilizarse activadores heterólogos. Por ejemplo, el activador híbrido trp-lac funciona en *Bacillus* además de en *E. coli*. Los procedimientos de transformación de otros procarionas aparte de *E. coli* son bien conocidos. Por ejemplo, los procedimientos de transformación de especies de *Bacillus* y activadores que se pueden utilizar para expresar proteínas se dan a conocer en la patente US nº 6.255.076 y en la patente US nº 6.770.475.

En levaduras, activadores convenientes comprenden GAL1-10 (Johnson y Davies (1984). *Mol. Cell. Biol.* 4:1440-1448) ADH2 (Russell *et al.* (1983) *J. Biol. Chem.* 258:2674-2682), PHO5 (*EMBO J.* (1982) 6:675-680), y MFα (Herskowitz y Oshima (1982) en *The Molecular Biology of the Yeast Saccharomyces* (eds. Strathern, Jones y Broach) Cold Spring Harbor Lab., Cold Spring Harbor, N.Y., págs. 181-209). Otro activador adecuado para su utilización en levaduras es el activador híbrido ADH2/GAPDH según se describe en Cousens *et al.*, *Gene* 61:265-275 (1987). Para hongos filamentosos tales como, por ejemplo, cepas de los hongos *Aspergillus* (McKnight *et al.*, patente US nº 4.935.349), los ejemplos de activadores útiles comprenden los derivados de los genes glucolíticos de *Aspergillus nidulans*, tales como el activador ADH3 (McKnight *et al.*; *EMBO J.* 4: 2093 2099 (1985)) y el activador *tpiA*. Un ejemplo de un terminador adecuado es el terminador ADH3 (McKnight *et al.*).

En algunas formas de realización, los polinucleótidos se colocan bajo el control de un activador inducible, que es un activador que dirige la expresión de un gen en el que el nivel de expresión es alterable por factores ambientales o de desarrollo tales como, por ejemplo, la temperatura, el pH, condiciones anaerobias o aerobias, la luz, los factores de transcripción y las sustancias químicas. Dichos activadores se denominan en la presente memoria "activadores inducibles", que permiten a uno para controlar el momento de la expresión de la glucosiltransferasa o enzima

implicada en la síntesis de nucleótidos de azúcar. Para *E. coli* y otras células anfitrionas bacterianas, los activadores inducibles son conocidos por los expertos en la materia. Estos comprenden, por ejemplo, el activador *lac*. Un activador inducible particularmente preferido para la expresión en procariontes es un activador doble que comprende un componente del activador *tac* unido a un componente del activador obtenido de un gen o genes que codifican enzimas implicadas en el metabolismo de la galactosa (por ejemplo, un activador de un gen de UDPgalactosa 4-epimerasa (*galE*)).

Los activadores inducibles de otros organismos también son bien conocidos por los expertos en la materia. Estos comprenden, por ejemplo, el activador de arabinosa, el activador *lacZ*, el activador de metalotioneína y el activador de choque térmico, así como muchos otros.

La construcción de montajes polinucleotídicos requiere generalmente el uso de vectores capaces de replicarse en bacterias. Una gran cantidad de kits se encuentran comercializados para la purificación de plásmidos a partir de bacterias. Para su correcta utilización, siga las instrucciones del fabricante (véase, por ejemplo, EasyPrepJ, FlexiPrepJ, ambos de Pharmacia Biotech; StrataCleanJ, de Stratagene; y, QIAexpress Expression System, Qiagen). Los plásmidos aislados y purificados se pueden manipular más para producir otros plásmidos, y utilizarse para transfectar células. También es posible la clonación en *Streptomyces* o *Bacillus*.

En los vectores de expresión utilizados para construir las células de la invención se incorporan a menudo marcadores seleccionables. Estos genes pueden codificar un producto génico, tal como una proteína, necesaria para la supervivencia o crecimiento de células anfitrionas transformadas cultivadas en un medio de cultivo selectivo. Las células anfitrionas no transformadas con el vector que contiene el gen de selección no sobrevivirán en el medio de cultivo. Los genes de selección típicos codifican proteínas que confieren resistencia a antibióticos u otras toxinas, tales como ampicilina, neomicina, kanamicina, cloranfenicol o tetraciclina. Alternativamente, los marcadores seleccionables pueden codificar proteínas que complementan deficiencias auxotróficas o suministrar nutrientes críticos no disponibles en medios complejos, por ejemplo, el gen que codifica la D-alanina racemasa para *Bacilli*. A menudo, el vector tendrá un marcador seleccionable que es funcional en, por ejemplo, *E. coli*, u otras células en las que el vector se replica antes de introducirse en la célula diana. Numerosos marcadores seleccionables son conocidos por los expertos en la materia y se describen por ejemplo en Sambrook *et al.*; más arriba. Un marcador seleccionable preferido para su utilización en células bacterianas es un marcador de resistencia a kanamicina (Vieira y Messing, *Gene* 19: 259 (1982)). La utilización de la selección con kanamicina presenta ventajas con respecto a, por ejemplo, la selección de la ampicilina debido a que la ampicilina es degradada rápidamente por la β -lactamasa en medio de cultivo, eliminando así la presión selectiva y permitiendo que el cultivo llegue a estar desbordado con células que no contienen el vector.

La construcción de vectores adecuados que contienen uno o más de los componentes mencionados anteriormente emplea técnicas de ligadura normalizadas como se describe en las referencias citadas anteriormente. Los plásmidos o fragmentos de ADN aislados se escinden, adaptan y vuelven a ligar en la forma deseada para generar los plásmidos requeridos. Para confirmar las secuencias correctas en los plásmidos construidos, los plásmidos se pueden analizar por técnicas normalizadas, tales como por digestión con endonucleasas de restricción, y/o secuenciación según procedimientos conocidos. Las técnicas de clonación molecular para conseguir estos fines son conocidas en la técnica. Una amplia variedad de procedimientos de clonación y ampliación *in vitro* adecuados para la construcción de ácidos nucleicos recombinantes son bien conocidos por los expertos en la materia.

Una variedad de vectores comunes adecuados para la construcción de las células recombinantes de la invención son bien conocidos en la técnica. Para la clonación en bacterias, los vectores frecuentes comprenden vectores derivados de pBR322 tal como pBluescript™, y de vectores derivados del fago λ . En la levadura, los vectores comprenden plásmidos de integración de levadura (por ejemplo, YIp5) y plásmidos que replican levaduras (los plásmidos de la serie YRp) y pGPD-2.

Los procedimientos para introducir vectores de expresión en una célula anfitriona seleccionada no son particularmente críticos, y dichos procedimientos son conocidos por los expertos en la técnica. Por ejemplo, los vectores de expresión pueden introducirse en células procariontes, como por ejemplo *E. coli*, por transformación con cloruro de calcio, y en células eucariotas por tratamiento con fosfato de calcio o electroporación. Otros procedimientos de transformación también son adecuados.

Formas de realización preferidas

Producción de 3'-sialil-lactosa.

La producción de 3'-sialil-lactosa puede llevarse a cabo en una cepa metabólicamente modificada definida anteriormente que expresa el gen A que codifica para una enzima que comprende una actividad de α -2,3-sialiltransferasa, tales como una α -2, 3-sialiltransferasa o una α -2,8-sialiltransferasa α -2,3 bifuncionales que utilizan lactosa como aceptador. Como se indica, esta cepa está desprovista de actividad de ácido siálico aldolasa, ManNAc cinasa y β -galactosidasa y expresa los genes heterólogos que codifican la CMP-Neu5Ac sintetasa, una ácido siálico sintasa, y una GlcNAc-6-fosfato 2 epimerasa. Para la producción a gran escala de sialil-lactosa, esta cepa puede

cultivarse a alta densidad celular en sustrato de bajo coste tales como glucosa o glicerol y se alimenta con lactosa que será interiorizada por la lactosa permeasa y sialilada por la sialiltransferasa recombinante utilizando el CMP-Neu5Ac endógenamente generado a partir de UDP-GlcNAc como se muestra en la figura 1.

5 Producción de 6'-sialil-lactosa

En la presente memoria, el gen de la sialiltransferasa es un gen que codifica α -2,6-sialiltransferasa tal como el gen de *Photobacterium damsela* (Yamamoto *et al.*, 1998) que da como resultado la producción de 6'-sialil-lactosa (Neu5Ac α -6Gal β -4Glc).

10

Producción de azúcar de GD3 y GT3

El GD3 (Neu5Ac α -8Neu5Ac α -3Gal β -4GlcCer) es un gangliósido menor que se encuentra en la mayoría de los tejidos normales en los vertebrados superiores, como por ejemplo los seres humanos. El nivel de GD3 se ha demostrado que aumenta durante algunas situaciones patológicas, tales como cánceres (glioma, melanoma) y que ejerce una función importante en la angiogenia tumoral Zeng, *et al. Cancer Res.*, 60:6670 (2000). Por lo tanto, la producción de GD3 a gran escala y rentable es de particular interés. Para llegar a este punto final, el gen heterólogo de sialiltransferasa mencionado anteriormente se selecciona de entre un gen que codifica una α -2,3 y α -2,8-sialiltransferasa bifuncional, por ejemplo tal como el gen *cstII* de *Campylobacter jejuni* (Gilbert *et al.*, 2002 y depositado bajo el número de registro ATCC 43438) y da como resultado la producción de azúcar de GD3 y GT3. El polipéptido de la sialiltransferasa bifuncional cataliza la transferencia de un resto de sialilo de una molécula de ácido siálico activado producida internamente para la Neu5Ac α -3Gal β -4Glc (GM3) para formar Neu5Ac α -8Neu5Ac α -3Gal β -4Glc. Esta reacción puede extenderse más para producir GT3, que es el precursor de los gangliósidos de la serie C, que son los principales componentes en el cerebro de los peces adultos y se encuentran en abundancia en los cerebros fetales de los vertebrados superiores (Letinic, *et al. Neuroscience*, 86, 1 (1998)). También se encuentran en diversos tumores neuroectodérmicos y hay por lo tanto potencialmente un gran interés en tener fácil acceso al oligosacárido GT3. Para este fin, el procedimiento en la presente memoria comprende además el cultivo del microorganismo de tal modo que el polipéptido bifuncional de sialiltransferasa de *Campylobacter jejuni* cataliza la transferencia de un resto de sialilo procedente de una molécula de ácido siálico activado producido internamente a la Neu5Ac α -8Neu5Ac α -3Gal β -4Glc (GD3) para formar Neu5Ac α -8Neu5Ac α -3Gal β -4Glc (GT3).

15

20

25

30

Producción de azúcar de GM1

Como se ilustra en la figura 2, el sistema para la producción de sialil-lactosa mostrado anteriormente (véase también la figura 1) puede extenderse a la producción de la porción de hidrato de carbono de la gangliósido GM1 mediante la expresión de los genes adicionales para una β -1,4GalNAc transferasa y para una β -1,3-galactosiltransferasa. En la presente memoria, el microorganismo de la invención es como se mostró anteriormente en la figura 1 y comprende además secuencias heterólogas que codifican β -1,4GalNAc transferasa así como β -1,3-galactosiltransferasa. En esta forma de realización, la β -1,4GalNAc transferasa transfiere un resto de UDP-GalNAc a sialil-lactosa (GM3) para formar GM2 y la β -1,3-galactosiltransferasa transfiere un resto galactosil a GM2 para formar GM1. Las cepas pueden no ser capaces de producir de forma natural UDP-GalNAc, tal como las cepas K12 de *E. coli*. En este caso, la cepa de la invención puede complementarse por un gen que codifica una UDP-GlcNAc 4 epimerasa, tal como por ejemplo el gen *wbpP* de *P. aeruginosa* (Creuzenet *et al.*, 2000) y el gen *gne* de *C. jejuni* (Bernatchez *et al.*, 2005). El azúcar de GM1 presenta una galactosa terminal no reductora y esta estructura se puede utilizar como aceptadora por α -2,3-sialiltransferasa para producir el azúcar de GD1a. El azúcar de GD1a tiene la misma estructura de disacárido no reductor terminal que la sialilactosa y se puede utilizar como aceptor por la β -1,4GalNAc transferasa para formar un producto heptasacárido que puede ser galactosilado en el octasacárido representado como GA1 (Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc) en la figura 2.

35

40

45

50

55

60

La formación de la GD1a y sus derivados mayores reducen el rendimiento de la producción del azúcar de GM1 y es una forma de realización específica de la invención para reducir o suprimir la formación de estos subproductos. Para este fin, los inventores han descubierto una α -2,3-sialiltransferasa que no puede utilizar el azúcar de GM1 como aceptor. La cepa NCTC 111168 de *C. jejuni* expresa estructuras de lipooligosacárido que imitan al gangliósido GM1 (Linton *et al.*, 2000). Esta estructura de núcleo externo de lipooligosacáridos fue denominada por Linton *et al.* (2000) núcleo externo "VI". NCTC 11168 de *C. jejuni* contiene un gen denominado *cstIII* que codifica una proteína que presenta una identidad de secuencia del 53% con sialiltransferasa (*CstII*) a partir de otras cepas de *C. jejuni* que expresan al imitador de GD1a (Gilbert *et al.*, 2002). La actividad de sialiltransferasa de la proteína *CstIII* permite ventajosamente la producción del azúcar de GM1 como el único producto de oligosacárido. Por lo tanto, en una forma de realización preferida, la invención comprende el procedimiento anterior para la producción específica de GM1, en el que la sialiltransferasa heteróloga es una α -2,3-sialiltransferasa codificada por el gen *cstIII* aislado de las cepas de *C. jejuni* que expresan las estructuras de lipooligosacárido que imitan el gangliósido GM1, tales como la cepa NCTC de *C. jejuni* nº de registro 111168.

Producción de azúcar de GM2

65

El sistema para la producción de sialil-lactosa descrito anteriormente se puede extender a la producción de la

fracción hidrato de carbono del gangliósido GM2 mediante la expresión de los genes adicionales para una β -1,4-GalNAc transferasa y para UDP-GlcNAc 4 epimerasa. Se ha mencionado anteriormente que la CgtA β -1,4-GalNAc transferasa de *C. jejuni* presenta una actividad del lado β -1,4-galactosiltransferasa que dio como resultado la producción del análogo de azúcar de GM2 designado como GA2 en la figura 3 (Antoine *et al.*, 2003). El azúcar de GA2 contiene una galactosa no reductora terminal y se ha constatado en el contexto de la presente invención que esta galactosa puede servir como aceptor de la sialiltransferasa para producir un tetrasacárido disialilado (GA5, figura 3) que a su vez la CgtAll β -1,4-GalNAc transferasa muy activa puede convertir en el pentasacárido GA4 o GA3 (figura 3).

La formación de los subproductos GA2, GA3, GA4 y GA5 reduce el rendimiento de la producción de azúcar de GM2 y dificulta su purificación. Por lo tanto, está comprendido dentro del alcance de la invención suprimir la formación de estos subproductos. La actividad del lado de la β -1,4-galactosiltransferasa se puede suprimir convenientemente utilizando mutantes incapaces de producir UDP-Gal. Como se ilustra en la figura 4, dicho mutante puede obtenerse destruyendo uno de los tres genes siguientes: el gen *galE* que codifica la UDP-glucosa epimerasa, el gen *galU* que codifica la UDP-Glc pirofosforilasa, y el gen *pgm* que codifica la fosfoglucomutasa. Por lo tanto, la invención se refiere a un procedimiento definido anteriormente para la producción de oligosacáridos sialilados, que se extiende a la producción del fragmento hidrato de carbono del gangliósido GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GM2), en el que el microorganismo comprende además secuencias heterólogas que codifican una β -1,4-GalNAc transferasa, tal como el gen *CgtAll* de la cepa O:36 de *C. jejuni* n^o de registro 43456 en la ATCC, y una UDP-GlcNAc 4-epimerasa y en el que el microorganismo ha suprimido o inactivado al menos uno de los tres genes siguientes para interrumpir la producción endógena de UDP-Gal: el gen *galE* que codifica la UDP-glucosa epimerasa, el gen *galU* que codifica la UDP-Glc pirofosforilasa, y el gen *pgm* que codifica la fosfoglucomutasa, evitando dicha interrupción la producción de subproductos tales como GA2, GA3, GA4 y GA5.

Producción de azúcar de GD2

El sistema para la producción de azúcar de GD3 se puede extender a la producción de la fracción hidrato de carbono de la GD2 gangliósido mediante la expresión de los genes heterólogos adicionales que codifican una UDP-GlcNAc 4-epimerasa y una β -1,4-GalNAc transferasa que utilizan el azúcar de GD3 como aceptor, tal como la proteína CgtAll procedente de la cepa O:36 de *C. jejuni* n^o de registro 43456 de la ATCC.

En este sistema, tanto la CstII sialiltransferasa como la CgtAll GalNAc transferasa compiten por la utilización de sialil-lactosa como aceptor (figura 5). Para favorecer la producción de azúcar de GD2, se redujo en el contexto de la presente invención la expresión de *cgtAll* en la primera fase del cultivo (para permitir que la sialil-lactosa se convierta principalmente en azúcar de GD3) y a continuación se aumentó la expresión de *cgtAll* en una segunda fase para convertir GD3 en GD2. Esto se puede hacer colocando el gen *cgtAll* bajo el control de un activador que se regula de forma independiente a partir de los activadores que controlan los otros genes. La actividad del lado de la β -1,4-galactosiltransferasa de la proteína CgtAll puede dar como resultado la producción de análogos galactosilados tales como los compuestos GA2, GA5 y GA6 representados en la figura 5. Mientras que estos oligosacáridos pueden ser de interés, también está comprendido dentro del alcance de la invención evitar la formación de estos subproductos utilizando cepas mutantes incapaces de producir UDP-Gal como se representó anteriormente para conducir a la producción específica de GD2.

Producción de azúcares de GD1b, GT1c, GT1b, GQ1b, GQ1c y GP1c

La producción de azúcar de GD1b (Gal β -3GalNAc β -4 (Neu5Ac α -8Neu5Ac α -3)Gal β -4Glc) y azúcar de GT1c (Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α -3)Gal β -4Glc) se puede conseguir utilizando la misma combinación de genes que en el sistema de producción de azúcar de GD2 descrito anteriormente y, además, expresando un gen de β -3 Gal transferasa. En la presente memoria, el microorganismo comprende además secuencias heterólogas que codifican UDP-GlcNAc 4 epimerasa, una β -1,4-GalNAc transferasa que utilizar el azúcar de GD3 y GT3 como aceptor, tal como la proteína CgtAll de la cepa O:36 de *C. jejuni* n^o de registro 43456 de la ATCC y una β -1,3-galactosiltransferasa, tal como el gen *cgtB* de la cepa O:2 de *C. jejuni* n^o de registro 11168 de la NCTC.

El sistema para la producción de GD1b y GT1c puede extenderse a la producción de GT1b, GQ1c, GQ1b y la GP1c mediante la expresión de un gen que codifica una sialiltransferasa que puede utilizar GD1b y GT1c como aceptores.

Producción de azúcares de GD1a y GT1a

La estrategia para la producción de azúcar de GT1a se ilustra en la figura 6 y se basa en la coexpresión del gen *cstII* para la sialiltransferasa bifuncional con el gen *cgtA* que codifica una β -1,4GalNAc transferasa que no utiliza azúcar de GD3 como aceptor, siendo los productos finales los azúcares de GD3, GD1a y GT1a. En la presente memoria, la invención se refiere a un procedimiento según se describió anteriormente para la producción de oligosacáridos sialilados que se aplica para producir específicamente Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GD1a) y Neu5Ac α -8Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GT1a), en los que el microorganismo comprende secuencias heterólogas que codifican una α -2,3 α -2,8-sialiltransferasa bifuncional, tal como el gen *cstII* de la cepa de *C. jejuni* n^o de registro 43438 de la ATCC, una β -1,4-GalNAc transferasa, tal como el gen *cgtAll* de la cepa O:36 de

C. jejuni n° de registro 43456 de la ATCC, que no utiliza el azúcar de GD3 como aceptor, y una β -3 Gal transferasa, tal como el gen *cgtB* de la cepa O:2 de *C. jejuni* n° de registro 11168 de la NCTC.

Producción de oligosacáridos LST_D (sialil-LNnT) y sialil-lewis X

Ya se ha descrito que la cepa metabólicamente modificada que expresa los genes *lgtA* e *lgtB* que codifican β -1,3 GlcNAc transferasa y β -1,4-galactosiltransferasa respectivamente puede producir el tetrasacárido LNnT (Gal β -4GlcNAc β -3Gal β -4Glc) a partir de lactosa exógena (Priem *et al.*, 2002). El sistema anterior para la producción de oligosacáridos sialilados se puede extender a la producción del pentasacárido LST_D (Neu5Ac α -3Gal β -4GlcNAc β -3Gal β -4Glc) expresando el gen adicional *nst* y *neuBCA* en una cepa NanK-, NanA- como se ilustra en la figura 7. El sistema para la síntesis de LST_D se puede combinar con el sistema de fucosilación que se ha descrito para la producción de oligosacárido Lewis X (Dumon *et al.*, 2004) para producir oligosacáridos que llevan el motivo sialil-Lewis X (Neu5Ac α -3Gal β -4(Fuca-3)GlcNAc β -) en su extremo no reductor. En este sentido, el microorganismo comprende además una secuencia heteróloga que codifica una α -1,3-fucosil transferasa, tal como *futA* de *Helicobacter pylori* (por ejemplo, la SEC. ID. n° 18 - de *Helicobacter pylori* n° de registro 26695 de la ATCC) y *futB* (por ejemplo, SEC. ID. n° 19 - de *Helicobacter pylori* n° de registro 26695 de la ATCC).

Producción del hexasacárido sialosil galactosil globósido (SGG)

Se ha constatado que el sialosil galactosil globósido (Neu5Ac α -3Gal β -3GalNAc β -3Gal α -4Gal β -4Gal) es el receptor preferido para la unión de *Escherichia coli* uropatógena (Stapleton *et al.*, 1998) y podría utilizarse potencialmente como agente antiinfeccioso. La producción del hexasacárido SGG a partir de globotriosa se ha descrito recientemente utilizando ácido siálico añadido exógenamente (producción eficiente de azúcar de globósidos utilizando microorganismos modificados metabólicamente en la solicitud de patente US 60/711.406). La producción de hexasacárido SGG fue llevada a cabo por una cepa NanA- *melA*- que expresa (i) el gen *lgtD* de *Haemophilus influenzae* (cepa rd) que codifica tanto una transferasa GalNAc como unas actividades de galactosiltransferasa, (ii) la *nst* sialiltransferasa (α -2,3 sialiltransferasa, tal como por ejemplo de *N. meningitidis*, como la cepa MC58: número de registro U60660 en Genbank- SEC. ID. n°5, protein_id = AAC44541.1 - SEC.. ID. n° 6), que cataliza la transferencia desde un resto sialilo o de una molécula de ácido siálico activado a globopentosa para formar hexasacárido sialosil galactosil globósido (SGG), (iii) el gen *neuA* de CMP Neu5Ac sintasa y (iv) el gen *wbpP* para UDP-GlcNAc epimerasa.

Este sistema se ha modificado para trabajar sin adición de ácido siálico utilizando una cepa NanK- NanA- *melA*- y expresando además los genes *neuABC*, como se ilustra en la figura 8.

En algunas formas de realización, los microorganismos se manipulan para mejorar el transporte de un sacárido aceptor en la célula. En la presente memoria, donde la lactosa o globotriosa es el sacárido aceptor, se pueden utilizar células de *E. coli* que expresan o sobreexpresan la LacY permeasa.

Por lo tanto, se describe el procedimiento anterior en el que el microorganismo se cultiva en un medio con globotriosa y es *LacY+*, *MelA-*, *nanT+*, *NanA-*, *NanK-* y comprende *lgtD* heterólogos, genes para α -2,3-sialiltransferasa y UGP-GlcNAc C4 epimerasa, así como los genes *neuABC*.

Además en la presente memoria se describe un procedimiento de acoplamiento en el que se utiliza un primer microorganismo para preparar globósidos. Como se mencionó anteriormente, el medio de cultivo puede incluir lactosa o globotriosa pero no existe ninguna necesidad de suministrar ácido siálico en la configuración en la presente memoria ya que se produce internamente. Cuando se utiliza globotriosa, en la presente memoria se contempla un conjunto de dos microorganismos separados, cultivándose dicho primer microorganismo en un medio con lactosa y siendo *LacY+*, *LacZ-*, *MelA-* y comprendiendo un gen *lgtC* heterólogo para producir globotriosa (la enzima α -1,4-Gal transferasa puede ser codificada, por ejemplo, por los genes *LgtC* de *N. meningitidis*, *N. gonorrhoeae* o *Haemophilus influenzae*, más particularmente por el gen *LgtC* de *Neisseria meningitidis* L1 (126E) número de registro U65788 en GenBank - SEC. ID. n° 7, protein_id AAB48385 - SEC. ID. n° 8), cultivándose el segundo microorganismo en un medio con globotriosa y siendo *LacY+*, *MelA-*, *nanT+*, *NanA-*; *NanK-* y comprendiendo genes heterólogos *lgtD*, *wbpP* y *nst*, así como los genes *neuABC*. El gen *lgtD* codifica una β -3 GalNAc transferasa para catalizar la transferencia de un resto de galactosa desde la UDP-Gal a globotetraosa para formar globopentosa (actividad de β -3 Gal transferasa). Por ejemplo, se puede utilizar el gen *lgtD* de *Haemophilus influenzae* HI1578, número de registro de GenBank U32832 - SEC. ID. n° 9, protein_id = AAC23227 - SEC. ID. n° 10.

Producción de sialilgalactosa

Se ha demostrado recientemente que el mutante *galk* que carece de actividad de galactocinasa puede utilizar galactosa exógena como aceptor para la síntesis de oligosacáridos con un terminal que reduce la galactosa (Dumon *et al.*, 2005). El procedimiento para la producción de oligosacáridos sialilados como se describió anteriormente puede ser utilizado convenientemente para producir el disacárido sialilgalactosa (Neu5Ac α -3Gal) por un *galk-*, *NanA-* y *NanK-* (o *NanKEAT-*) de microorganismo que expresa el gen de la sialiltransferasa y los genes *neuBCA*

cultivados en un medio con galactosa.

Producción de oligosacáridos sialilados con un terminal que reduce la galactosa

5 El procedimiento para la síntesis de sialilgalactosa se puede adaptar a la producción de análogos de toda la estructura sialilada mencionada anteriormente. El uso de galactosa como aceptor en lugar de lactosa dio como resultado la formación de análogos que carecen del resto de glucosa terminal.

10 Producción de oligosacáridos sialilados con una lactosa o galactosa terminal que presenta funciones químicas latentes

La amplia especificidad de la azúcar permeasa se puede utilizar para interiorizar derivados de lactosa o galactosa que presentan funciones químicas latentes para producir oligosacáridos conjugables. Esta estrategia se ha aplicado con éxito a la síntesis de las fracciones de oligosacáridos de los gangliósidos GM2 y GM3 con un alil o un propargil aglucón o (Fort *et al.*, 2005). La función alquino permite una adición de azido en condiciones acuosas y la función alqueno puede convertirse en un aldehído para unirse a proteínas por aminación reductora, o transformarse en un grupo amino versátil mediante la adición de cisteamina. Pueden utilizarse también otras funciones químicas tal como el grupo azida o amino. Todos estos derivados de lactosa o galactosa pueden utilizarse convenientemente para producir análogos conjugables de toda la estructura sialilada mencionada anteriormente.

20 Producción de oligosacáridos sialilados con estructura de quitina-oligosacárido en su extremo reductor

Se ha demostrado que cepas de *E. coli* que sobreexpresan el gen *nodC* de *Azorhizobium caulinodans* para quitina-oligosacárido sintasa producen más de $2 \text{ g} \cdot \text{l}^{-1}$ de quitinopentosa cuando se cultivaban a alta densidad celular (Samain *et al.*, 1997). Una vez producida en el citoplasma, la quitinopentosa puede servir como aceptor para las glucosiltransferasas que reconocen un resto terminal GlcNAc no reductor. Esta estrategia se utilizó para la síntesis del hexasacárido Gal β -4[GlcNAc β -4]₄GlcNAc por una cepa de *E. coli* que coexpresó el gen *nodC* de *Azorhizobium caulinodans* y el gen *IgtB* de *Neisseria meningitidis* para β -1,4-galactosiltransferasa (Bettler *et al.*, 1999). El motivo de *N*-acetil-lactosamina terminal de este hexasacárido es un aceptor para la sialiltransferasa. El heptasacárido sialilado Neu5Ac α -3Gal β -4[GlcNAc β -4]₄GlcNAc por lo tanto se puede producir convenientemente en cepas mutantes de *NanK*, *NanA* que coexpresan a *nodC*, *IgtB*, *nst* y *neuBCA*. Una estrategia recientemente desarrollada para reducir este tamaño consiste en hidrolizar con enzimas la quitinopentosa con una quitinasa en las bacterias vivas tan pronto como la produce NodC (Cottaz y Samain, 2005). Se ha constatado que en el procedimiento de la invención es posible evitar la formación del cebador de quitinopentosa de gran tamaño, que aumenta considerablemente el peso molecular de las estructuras diana utilizando el gen de quitinasa de *Bacillus circulans* por ejemplo.

Como se ilustra en la figura 9, el procedimiento de la invención permite la formación del tetrasacárido sialilado Neu5Ac α -3Gal β -4GlcNAc β -4GlcNAc por las cepas *NanK*-, *NanA*- que coexpresan *nodC*, *chiA*, *IgtB*, *nst* y *neuBCA*. Por lo tanto, la invención se refiere al procedimiento anterior; o alternativamente a un procedimiento en el que nada exógeno se añade al medio de cultivo; para la producción de oligosacáridos sialilados con estructura de quitina-oligosacárido en su extremo reductor, tal como el heptasacárido sialilado Neu5Ac α -3Gal β -4[GlcNAc β -4]₄GlcNAc, en el que la sialil-transferasa es un α -2,3-sialiltransferasa, tal como el gen *nst* de *Neisseria*, y el microorganismo comprende además una quitina-oligosacárido sintasa tales como el gen *nodC* de *Azorhizobium caulinodans* y un gen de β -1,4-galactosiltransferasa tal como el gen *IgtB* de *Neisseria meningitidis*. Se puede extender a la producción del tetrasacárido sialilado Neu5Ac α -3Gal β -4GlcNAc β -4GlcNAc, en el que el microorganismo comprende además una secuencia heteróloga que codifica una quitinasa, tal como el gen *chiA*.

En otro aspecto todavía, la invención se refiere a un microorganismo como se definió anteriormente, así como a un medio de cultivo celular que comprende un precursor exógeno seleccionado entre lactosa, galactosa, β -galactósido, y α -galactósido tal como globotriosa (Gal α -4Gal β -4Glc) y dicho microorganismo.

Ejemplo 1: Construcción de los mutantes *NanA*, *NanKA* y *NanKETA*

55 Todos los mutantes se construyeron a partir de la cepa DC (Dumon *et al.*, 2005) que era un derivado de la cepa DH1 que lleva las mutaciones *lacZ* y *lacA*. Puesto que todos los derivados de la cepa DH1 son mutantes de *recA*, se transformaron con el plásmido pEXT22 de bajo número de copias (Dykxhoorn *et al.*, 1996) que lleva un gen *recA* funcional y una resistencia a la kanamicina para recuperar un fenotipo de RecA⁺ transitorio para el procedimiento de inactivación de genes que implicaba recombinación del ADN. Una vez que el gen ha sido interrumpido, el plásmido se curó cultivando la célula sin kanamicina y detectando el fenotipo de RecA⁻.

60 La cepa AZL se construyó a partir de la cepa DC, mediante la inactivación de *NanA* usando el plásmido pMAK705 suicida (Hamilton *et al.*, 1989) como se describió anteriormente (Priem *et al.*, 2002).

65 Para construir la cepa ZLKA a partir de la cepa DC, se destruyeron los genes *NanKETA* suprimiendo un segmento de 3,339 kb en el ADN cromosómico utilizando el procedimiento de una etapa anteriormente descrito que utiliza cebadores de RCP para proporcionar la homología con la secuencia diana (Datsenko y Wanner, 2000). La

secuencia del cebador aguas arriba fue 5'GCAATTATTGATTCGGCGGATGGTTTGCCGATGGTGGTGTAGGCTG GAGCTGCTTC (SEC. ID. nº 11) y la secuencia del cebador aguas abajo fue 5'CTCGTACCCCTGCCGGCGCG CGTGAAAATAGTTTTGCGCATATGAATATCCTCCTTAG. ((SEC. ID. nº 12).

- 5 El mismo procedimiento se utilizó para inactivar el gen *NanK* en la cepa AZL para obtener la cepa AZK excepto que el tamaño del fragmento eliminado fue de 0,537 kb y que la secuencia del cebador aguas arriba fue

5'CACTGGCGATTGATATCGGCGGTAATAACTTGCCGCCGTGTAGGCTGGAGCTGC
TTC (SEC ID nº 13).

10 **Ejemplo 2: Clonación de los genes *neuBCA***

Un fragmento de 2.995 ADN que contiene la secuencia de los genes *neuBCA* se amplió por RCP utilizando el ADN genómico de la cepa 43438 de la ATCC de *Campylobacter jejuni* como plantilla.

- 15 Un secuencia de *KpnI* se añadió al cebador izquierdo:

5'GGTACCTAAGGAGGAAAATAAATGAAAGAAATAAAAATACAA (SEC ID nº 14)

y una secuencia *XhoI* se añadió al cebador derecho

- 20 5'CTCGAGTTAAGTCTCTAATCGATTGTTTTCCAATG (SEC ID nº 15).

El fragmento ampliado se clonó en primer lugar en el vector pCR4Blunt-TOPO (Invitrogen) y a continuación se subclonó en las secuencias *KpnI* y *XhoI* del vector pBBR1-MCS3 para formar pBBR3-SS.

25 **Ejemplo 3: Producción de sialil-lactosa por cepas de *E. coli* metabólicamente modificadas**

Se investigó la producción de sialil-lactosa con diferentes cepas mutantes contenidas en el gen *nst* de *N. meningitidis* para α -2, 3-sialiltransferasa y el plásmido pBBR3-SS que contenía los genes *neuC*, *NeuB* y *neuA* de *C. jejuni* que codifican la N-acetilglucosamina-6-fosfato-epimerasa, la ácido siálico sintasa y CMP-Neu5Ac sintetasa, respectivamente. La producción de sialil-lactosa se estimó por análisis cuantitativo colorimétrico del ácido siálico tanto en fracciones intracelulares como extracelulares (Tabla 2). Los resultados demostraron que el mutante AW1 de *NanA* y la cepa DC6, que no contenía ninguna mutación en el operón de ácido siálico, produjeron baja cantidad de sialil-lactosa con un rendimiento de producción similar. Los dos mutantes AZK1 y DC7 que llevaban las mutaciones *NanK* y *NanA* ambas producen cantidades cuatro veces mayores de sialil-lactosa. No pudo detectarse nada de ácido siálico en el cultivo de referencia de DC7 incubado sin lactosa, lo que indica que la alta concentración de ácido siálico total correspondía a la formación de sialil-lactosa.

La mejora de la producción de sialil-lactosa se confirmó también por análisis de TLC, que demostró que la banda correspondiente a sialil-lactosa era mucho más intensa en extractos de DC7 y AZK1 que en extractos de DC6 y AW1.

Tabla 2. Análisis cuantitativo colorimétrico de ácido siálico en fracciones intracelulares y extracelulares de cultivos de alta densidad celular de cepas genéticamente modificadas para la producción de sialil-lactosa

Cepa	Mutación	Genes heterólogos expresados	Aceptor	Concentración de Neu5Ac (g·l ⁻¹) intracelular extracelular	
DC		ninguno	lactosa	0	0
DC6		<i>neuBCA nst</i>	lactosa	0,94	0,43
AW1	<i>NanA</i>	<i>neuBCA nst</i>	lactosa	1,13	0,27
DC7	<i>NanKETA</i>	<i>neuBCA nst</i>	lactosa	2,32	3,25
DC7	<i>NanKETA</i>	<i>neuBCA nst</i>	ninguno	0,11	0
DC0	<i>NanKETA</i>	<i>neuBCA</i>	lactosa	0	0
AZK1	<i>NanK NanA</i>	<i>neuBCA nst</i>	lactosa	2,16	2,93

El ácido siálico total se analizó cuantitativamente por el método de la difenilamina (Werner y Odin, 1952). Los cultivos se incubaron 30 horas después de la adición de la lactosa suministrada a una concentración de 7,5 g/l, excepto para el cultivo de referencia de DC7 sin lactosa.

50 **Ejemplo 4: Producción a gran escala de sialil-lactosa con alimentación continua de lactosa**

Este rendimiento de la producción se aumentó ampliando el tiempo de cultivo a 71 horas. Se añadió lactosa a una concentración de 2 g·l⁻¹ al comienzo de las fases de alimentación discontinua. Se añadió en primer lugar de forma

continúa con un ritmo de aporte de $0,52 \text{ g}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$ durante 5 horas en la fase con un ritmo alto de alimentación de glicerol. El ritmo de aporte de lactosa se disminuyó a continuación a $0,3 \text{ g}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$ hasta el final del cultivo en la segunda fase con un ritmo de bajo alimentación de glicerol. El análisis por TLC demostró que sialil-lactosa (compuesto 2, figura 10) se produjo continuamente hasta el final del cultivo y que la producción sialil-lactosa no estuvo limitada por el aporte de lactosa (compuesto 1) que siempre pudo detectarse en pequeña cantidad a lo largo del cultivo en el fracción intracelular.

El análisis cuantitativo colorimétrico del ácido siálico indicó que la sialil-lactosa se acumulaba principalmente en la fracción intracelular en la primera parte del cultivo. La concentración de sialil-lactosa intracelular se estabilizó entonces a aproximadamente $10 \text{ g}\cdot\text{l}^{-1}$ y la sialil-lactosa, que se produjo además, se segregó en el medio extracelular donde se acumuló a una concentración final de $15,5 \text{ g}\cdot\text{l}^{-1}$ (figura 11).

Ejemplo 5: Purificación de sialil-lactosa

Se purificó sialil-lactosa procedente de un litro de cultivo de DC7 obtenida como se describe en el ejemplo 4. Al final del cultivo, la fracción extracelular se separó de las células por centrifugación. El pH de la fracción extracelular se redujo a 3,00 mediante la adición de una resina de intercambio catiónico fuerte (Amberlite IR120 forma H^+). Esto dio lugar a la precipitación de las proteínas que se separaron por centrifugación. El pH del sobrenadante claro se ajustó a 6,0 mediante la adición de un intercambiador aniónico débil (Dowex 66 forma de base libre) y después la mitad del sobrenadante se cargó en una columna (5 x 20 cm) Dowex 1 (forma HCO_3). La resina Dowex 1 retuvo la sialil-lactosa y, después de lavar con agua destilada, se eluyó con un gradiente continuo de NaHCO_3 0-500 mM. El mismo procedimiento se repitió con la otra mitad del sobrenadante. Las fracciones eluidas que contenían sialil-lactosa se agruparon y el NaHCO_3 se extrajo mediante un tratamiento con Amberlite IRL 20 (forma H^+) hasta pH 3,0. El pH se ajustó a 6,0 con NaOH y la sialil-lactosa se liofilizó.

Para la purificación de la fracción intracelular, las células se permeabilizaron por calentamiento (100°C , 45 min) y se volvieron a poner en suspensión en el mismo volumen que el medio de cultivo inicial. Los oligosacáridos se difundieron libremente fuera de las células y se recuperaron en el sobrenadante después de la centrifugación. La purificación de sialil-lactosa se llevó a cabo a continuación utilizando el mismo protocolo que para la fracción extracelular.

A partir de un cultivo de un litro de la cepa DC7, el rendimiento de la sialil-lactosa purificada fue de 9 gramos de la fracción extracelular y 6 gramos de la fracción intracelular. La identificación del producto purificado como sialil-lactosa se confirmó por análisis de espectrometría de masas.

Ejemplo 6: Producción de azúcares de GD3 y GT3

La producción de los azúcares de GD3 y GT3 procedentes de lactosa exógena y Neu5Ac se ha descrito anteriormente utilizando la cepa metabólicamente modificada que expresa el gen *cstII* de *Campylobacter jejuni* para la α -2,3 y α -2,8 sialiltransferasa bifuncional (Antoine *et al.*, 2005). En la presente memoria se ha investigado la producción de estos dos oligosacáridos sin suministro exógeno de ácido siálico utilizando el sistema descrito en el ejemplo 3. La cepa NF3 productora de GD3 era un mutante de *NanKEAT* que coexpresó *cstII* y *neuBCA* (Tabla I). La cepa NF3 se cultivó a alta densidad celular en presencia de $3 \text{ g}\cdot\text{l}^{-1}$ de lactosa. El análisis por TLC (figura 12) demostró que la lactosa se convirtió por completo en tres compuestos que se supone que son los azúcares de GM3 (1) GD3 (2) y GT3 (3).

La fracción intracelular del cultivo de la cepa NF03 se purificó por cromatografía de intercambio iónico sobre Dowex 1 como se describe en el ejemplo 4 y se separaron tres fracciones de oligosacáridos que contienen los azúcares de GM3, GD3 y GT3 respectivamente. Los rendimientos de las tres fracciones de azúcar fueron 0,16 g, 0,75 g y 1,26 g, respectivamente. La identificación se confirmó por análisis de espectrometría de masas de las fracciones purificadas.

Ejemplo 7: Producción de azúcar de GM1 utilizando la sialiltransferasa de *N. meningitidis* y la UDP-GlcNAc C4 epimerasa de *P. aeruginosa*.

La producción de azúcar de GM1 partir de lactosa exógena y de Neu5Ac se ha descrito anteriormente utilizando una cepa metabólicamente modificada que expresa: (i) el gen *nst* de *N. meningitidis* para α -2,3-sialiltransferasa; (ii) el gen *cgtA* de cepa OH4384 de *C. jejuni* O:19, que codifica una β -1,4-GalNAc transferasa; (iii) el gen *cgtB* de la cepa NCTC 11168 de *Campylobacter jejuni*, que codifican β -1,4-galactosiltransferasa, (iv) el gen *wbpP* de *P. aeruginosa* que codifica una UDP-GlcNAc C4 epimerasa (Antoine *et al.*, 2003). Los primeros intentos para producir azúcar de GM1 sin adición exógena de ácido siálico coexpresando *cgtA*, *cgtB* y *nst* con *neuBCA* indicaron que la etapa limitante era la conversión de sialil-lactosa en azúcar de GM2 por la CgtA GalNAc transferasa. Se ha demostrado que las GalNAc transferasas existen en diferentes versiones dependiendo de las cepas de *Campylobacter*. Se señaló que la versión clonada de *cgtA* procedente de la cepa OH4384 tiene una actividad específica mucho más baja que las de las versiones de la cepa ATCC 4356 o NTCC 11168 (Gilbert *et al.*, 2002; Varki, 1993). El gen *cgtAII* por lo tanto se clonó por RCP a partir del ADN genómico de la cepa ATCC 4356 y se subclonó con el gen *nst* en un

plásmido pBluescript, produciendo el plásmido nst-pBS-cgtAll (Tabla 1). El gen *wbpP* se clonó a partir del plásmido pBBRwbpP (Antoine *et al.*, 2003) aguas abajo del gen *neuBCA* en pBBR3-SS, produciendo pBBR-SS-wbpP. El gen *cgtB* se clonó a partir del plásmido pACT3cgtAB (Antoine *et al.*, 2003) en el plásmido pSU27-18, produciendo pSU18-cgtB. La cepa DC15 se construyó transformando la cepa ZLKA mutante de *NanKEAT* con los tres plásmidos pBS-cgtAll-nst, pBBR-SS-wbpP y pSU18-cgtB (tabla 1).

Como se muestra en la figura 13, la cepa DC15 no acumuló el azúcar de GM3 (1), lo que indica que la CgtAll procedente de la cepa ATCC 4356 era considerablemente más activo que CgtA procedente de la cepa OH4384. Al final del cultivo, los principales productos eran un compuesto (5) que migraba más lento que el azúcar de GM1 y un compuesto (2') que migraba como el azúcar de GM2. Después de la purificación en Dowex1, el análisis de espectrometría de masas indicó que el compuesto (2') era un análogo de azúcar de GM2 que presenta un resto de Gal en lugar de GalNAc y que además se designó como azúcar de GA2 (Tabla 3). La estructura era: Gal β -4(Neu5Ac α -3)Gal β 4Glc. El espectro de masas del compuesto (5) sugiere que era un octasacárido disialilado formado por la transferencia de dos restos de azúcar (una HexNAc y un Hex) en el azúcar de GD1a. Dado que estos dos azúcares son más probablemente añadidos por CgtAll y CgtB, la estructura del compuesto (5), que fue designado además como azúcar de Ga1, es probable que sea: Gal β -3GalNAc β -4Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (Tabla 3).

Tabla 3 Estructura de análogos de azúcar de gangliósido formados como subproductos durante la síntesis de azúcar de gangliósidos en razón de la actividad de β -1,4Galactosiltransferasa de la CgtA β -1,4-GalNActransferasa.

Estructura	Denominación en el texto
Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc	GA1
Gal β -4(Neu5Ac α -3)Gal β 4Glc	GA2
GalNAc β -4(Neu5Ac α -3)Gal β -4(Neu5Ac α -3)Gal β -4Glc	GA3
Gal β -4(Neu5Ac α -3)Gal β -4(Neu5Ac α -3)Gal β -4Glc	GA4
Neu5Ac α -3Gal β -4(Neu5Ac α -3)Gal β -4Glc	GA5
Gal β -4(Neu5Ac α -8Neu5Ac α -3)Gal β -4Glc	GA6

Ejemplo 8: Producción de azúcar de GM1 utilizando la sialiltransferasa de *N. meningitidis* y la UDP-Gal C4 epimerasa de *Campylobacter*(GNE).

El azúcar de GA2 se produjo por cultivo de DC15 porque la CgtAll es muy activa y se puede utilizar UDP-Gal como donante de azúcar en lugar de UDP-GalNAc en caso de escasez de UDP-GalNAc. Ya que esta escasez puede ser debida a una actividad insuficiente de la UDP-GlcNAc C4 epimerasa codificada por *wbpP*, se investigó la utilización de la epimerasa codificada por el gen *gne* de *C. jejuni*. Se ha demostrado recientemente que este gen es más activo que *WbpP* (Bernatchez *et al.*, 2005). El gen *gne* se clonó por RCP a partir del ADN genómico de la cepa NCTC 11168 de *C. jejuni* y se subclonó en el plásmido pBBR-SS aguas abajo de los genes *neuBCA*. El plásmido pBBR-SS-gne resultante se utilizó para construir la cepa DC21 que era similar a la cepa DC15, excepto que expresaba *gne* en lugar de *wbpP*. (Tabla 1).

Como se muestra en la figura 14, la expresión adicional de *gne* ha suprimido casi en su totalidad la acumulación de azúcar de isoGM2 y los dos productos principales fueron el azúcar de GM1 (3) y el azúcar de GA1 (5). El GM1 se acumuló temporalmente; su concentración intracelular alcanzó un máximo de 28 horas después de la adición de lactosa y a continuación disminuyó debido a la formación del compuesto (5).

Ejemplo 9: Producción de azúcar de GM1 utilizando el *C. jejuni* sialiltransferasa CstIII

El gen *cstIII* se clonó por RCP utilizando ADN genómico de la cepa NCTC 11168 de *C. jejuni* O:2 como plantilla. El gen *cstIII* se subclonó a continuación en un plásmido pBluescript aguas arriba de *cgtAll*. El plásmido resultante pBS-cstIII-cgtAll se utilizó para construir la cepa DC22 que también expresó *cgtB* y los genes *neuBCA* para la biosíntesis de CMP-Neu5.

El análisis por TLC (figura 15) demostró que el azúcar de GM1 (3) era casi el único oligosacárido encontrado en la fracción intracelular de la cepa DC22 al final del cultivo. La sialil-lactosa (1) y el azúcar de GM2 (2) apenas pudieron ser detectados como producto intermedio. Después de la purificación en Dowex1 como se describe en el ejemplo 5, el rendimiento de azúcar de GM1 fue de 6 g en un litro de cultivo.

Ejemplo 10: Producción de azúcar de GM2

La producción de azúcar de GM2 se investigó con la cepa ZWT que era similar a la cepa DC21 productora de GM1 excepto que no expresó el gen *cgtB*. Como se muestra en la figura 16, la cepa ZWT convirtió la lactosa muy rápidamente en los compuestos (2) que migraron como GM2. Inesperadamente se produjo también un compuesto (3) que migraba más lento que GM2.

Los compuestos (2) y (3) se purificaron a partir de la fracción intracelular del cultivo de la cepa ZWT por cromatografía en Dowex1. El análisis de espectrometría de masas indicó que el compuesto purificado (2) consistía en una mezcla de azúcar de GM2 y del análogo de azúcar de GA2.

5 El espectro de masas ESI⁻ de la fracción B mostró dos picos a m/z 1310 y 1269. El pico a m/z 1310 probablemente corresponde a los iones casi moleculares [(M+Na-H)-H]⁻ derivado del hexasacárido GalNAcβ-4(Neu5Acα-3)Galβ-4(Neu5Acα-3)Galβ-4Glc que fue designado además como azúcar de GA3 (Tabla 3). El segundo pico a m/z 1269 corresponde a los iones casi moleculares [(M+Na-H)-H]⁻ derivados de la estructura Galβ-4(Neu5Acα-3)Galβ-4(Neu5Acα-3)Galβ-4Glc (denominado azúcar de GA4).

10 La formación de los azúcares de GA3 y GA4 se explica como se ilustra en la figura 11 por una actividad secundaria de la Nst sialiltransferasa que parece ser capaz de sialilar el terminal Gal no reductor del azúcar de GA2. El pentasacárido disialilado resultante (azúcar de GA6) puede servir como aceptor para que la CgtAII se convierta en los azúcares de GA3 o GA4.

15 **Ejemplo 11: Construcción de mutantes de *galU***

Para construir el mutante *galU*, un fragmento de ADN de 1,88 kb que contiene la secuencia *galU* se amplió por RCP utilizando el ADN genómico de *E. coli* K12 como plantilla y los dos cebadores siguientes:

20 5'CAATGCCAAATATGGGGAAC (SEC. ID. nº 16) y 5'GCGGCCGCGTCTTTTCTGGCTAA (SEC. ID. nº 17)

El fragmento ampliado se clonó en el vector pCR4blunt-TOPO (Invitrogen) y un fragmento de 0,244 kb situado entre las dos secuencias *NdeI* presentes en la secuencia *galU* se escindió por digestión con *NdeI*. El gen *galU* truncado se subclonó en las secuencias *Sall NotI* del vector pKO3. La ruptura de *galU* se llevó a cabo en la cepa ZLKA según el protocolo de sustitución de genes pKO3 (Link *et al.*, 1997), produciendo la cepa anfitriona ZWU.

25

Ejemplo 12: Producción de azúcar de GM2 por mutantes de *galU*

30 El cepa ZWU2 era similar a la cepa ZWT utilizado para la producción de azúcar de GM2 en el ejemplo 10, excepto que la cepa ZWU de *galU* se utilizó como cepa anfitriona en lugar de la cepa ZLKA (Tabla 1). Como se muestra en la figura 17, no se detectaron más subproductos (3) al final del cultivo de ZWU2. Además el análisis de espectrometría de masas demostró que el compuesto (2) sólo se componía de azúcar de GM2 y que el análogo de azúcar de GA2, que estaba presente en el compuesto (2) purificado a partir del cultivo ZWT, se pudo detectar (figura 18).

Ejemplo 13: Producción de azúcar de GD2

El primer intento de producir el azúcar de GD2 se realizó con la cepa NF08 que se construyó transformando la cepa anfitriona ZLKA con los tres plásmidos pUC18-cstII, pBBR3-SS-gne, pWKS-cgtAII (Tabla 1). La caracterización de oligosacáridos producidos por la cepa NF08 mostró que los principales productos eran el azúcar de GM2 y su azúcar de GA2 análogo galactosilado. También se formó una pequeña cantidad de azúcar de GD2 pero se purificó como una mezcla con los análogos GA5 y GA6 galactosilados (figura 5).

40

El gen *cgtAII* se subclonó en el plásmido pBAD 33 bajo el control del activador arabinosa (P_{ara}). De esta manera la expresión de *cgtAII* podría regularse independientemente de los demás genes que estaban bajo el control del activador P_{lac} . La cepa NF09 que contiene el plásmido pBAD33-cgtAII en lugar de pWKS-cgtAII se cultivó sin arabinosa durante las primeras 24 horas siguientes a la adición de lactosa ($3g \cdot l^{-1}$). Después se añadió arabinosa y la cepa se cultivó por un período adicional de 24 horas. Esta estrategia resultó eficaz en la disminución de la producción de azúcar de GM2 con un aumento simultáneo en el rendimiento de azúcar de GD2. Sin embargo, los análogos galactosilados GA5 o GA6 se producían todavía en cantidad significativa. Para evitar la formación de estos análogos, que son muy difíciles de separar de GD2, se investigó la producción de GD2 en la cepa ZWU1 mutante de *galU*.

45

50

La cepa ZWU1 era similar a la cepa NF09 excepto que la cepa ZWU de *galU* se utilizó como cepa anfitriona en lugar de la cepa ZLKA (Tabla 1). Se añadió arabinosa después de 25 horas de cultivo para aumentar la expresión de *cgtAII* después de que casi la totalidad de la sialil-lactosa ha desaparecido. La purificación de la fracción intracelular en Dowex1 condujo a la separación de cuatro fracciones (figura 19). El análisis de espectrometría de masas indicó que la fracción A estaba compuesta de azúcar de GM2 y que la fracción B consistía en azúcar de GD2. La fracción C contenía principalmente GD3 y la fracción D contenía azúcares de GT2. No pudieron detectarse análogos galactosilados y el azúcar de GD2 se obtuvo como producto principal.

55

60

Ejemplo 14: Producción de GD1b y GT1

Los dos genes *cstII* y *cgtB* se clonaron en el mismo plásmido pBluescript (pBS-cstII-cgtB). La cepa NF21 se construyó transformando la cepa anfitriona ZLKA con el plásmido pBS-cstII-cgtB y los dos plásmidos pBBR-SS-gne y pBAD33-cgtAII. La cepa NF21 se cultivó a alta densidad celular como se describe para la cepa NF09 en el ejemplo 13, en las condiciones que favorecieron la formación de GD2. Los oligosacáridos producidos a partir de un cultivo de

65

un litro de la cepa NF21 se purificaron en primer lugar en la resina Dowex 1 (HCO₃) y tres fracciones (A, B y C) se eluyeron con un gradiente de NaHCO₃ (0-1 M).

El análisis de espectrometría de masas indicó que la fracción A (490 mg) contenía azúcar de GD1, que la fracción B (870 mg) contenía una mezcla de azúcar de GD1 con análogo de GA5 o GA6, y que la fracción C (960 mg) contenía azúcar de GT1. Los azúcares de GD1 y GT1 se purificaron a partir de las fracciones A y C por cromatografía de exclusión por tamaño en una columna Toyopearl HW40S utilizando NaHCO₃ 100 mM como eluyente. Como se muestra en la Tabla 4, el análisis de ¹H RMN demuestra que los desplazamientos químicos de protones del H-1 de la galactosa terminal tanto de GD1 como GT1 estaban muy próximos al valor del H-1 del azúcar de GM1a en el que la galactosa terminal no está sustituida. Por el contrario la señal del H-1 de la galactosa terminal de GD1a y GT1a cambió a 4,62 ppm en razón de su sustitución con un ácido siálico. Estos resultados indicaban claramente que los azúcares de GD1 y GT1 producidas por la cepa NF21 tenían la estructura de GD1b y GT1c.

Tabla 4. Desplazamiento químico de la RMN de protón de azúcares de gangliósidos a 343°K

Cepa	Azúcar purificado	Galactosa terminal		Galactosa interna		GalNAc	Glucosa	
		H-1	H-3	H-1	H-3	H-1	α H-1	β H-1
DC22	GM1a	4,57	3,84	4,56	4,17	4,83	5,25	4,68
NF17	GD1a	4,62	4,11	4,54	4,17	4,83	5,25	4,68
NF17	GT1a	4,62	nd	4,53	nd	4,83	5,25	4,68
NF21	GD1b	4,55	nd	4,53	nd	4,83	5,25	4,68
NF21	GT1c	4,55	3,81	4,52	4,17	4,83	5,25	4,68

Ejemplo 15: Producción de azúcar de GT1a

Los dos genes *cstII* y *cgtA* se clonaron en el mismo plásmido pBluescript para coexpresarse con *cgtB* y los genes *neuABC* en una cepa anfitriona ZLKA mutante de *NanKEAT* (Tabla 1). El gen *cstII* se clonó aguas abajo de *cgtA* para presentar una expresión máxima de *cgtA* y menor expresión de *cstII* a fin de minimizar la formación de GD3. Como se muestra en la figura 20, el análisis por TLC del cultivo de NF17 indicaba que los compuestos que migran como sialil-lactosa (1) y azúcar de GM1 (3) se produjeron temporalmente. Una pequeña cantidad de compuesto (2) que migró como GD3 se recuperó al final del cultivo, pero los principales productos finales fueron los compuestos (4) y (5) que migraron más despacio que GM1. El análisis de espectrometría de masas de los compuestos purificados indicó que los compuestos (4) y (5) tienen un peso molecular correspondiente a GD1a y GT1a respectivamente.

Ejemplo 16: Producción de sialilgalactosa

La cepa GLKA se construyó suprimiendo los genes *NanKEAT* en el cromosoma de la cepa GLK mutante de GalK (Dumon *et al.*, 2005.). La cepa GLK7 se obtuvo transformando la cepa GLKA con los plásmidos pBS-nst y pBBR3-SS (Tabla 1). El cultivo de la cepa GLK7 a alta densidad celular en presencia de 3 g·l⁻¹ de galactosa dio lugar a la formación de un disacárido que se identificó como sialilgalactosa por análisis de espectrometría de masas. El rendimiento de la producción de sialilgalactosa se estimó en 6 g·l⁻¹ por análisis cuantitativo colorimétrico de ácido siálico total.

Ejemplo 17: Producción del tetrasacárido Neu5Acα-3Galβ-4GlcNAcβ-4GlcNAc

El plásmido pBS-nst-nodC se construyó clonando el gen *nodC* de *A.caulinodans* procedente del plásmido pBS-nodC (Cottaz y Samain, 2005) en las secuencias *EcoRV KpnI* de pBS-nst. La cepa SN4 se obtuvo transformando la cepa ZLKA con los tres plásmidos pBS-nst-nodC, pBBR3-SS, pWKS-IgtB-chiA (Tabla 1). El cultivo de la cepa SN4 a alta densidad celular dio lugar a la producción de un oligosacárido principal que se identificó como Neu5Acα-3Galβ-4GlcNAcβ-4GlcNAc por análisis de espectrometría de masas.

Referencias

Antoine, T., Priem, B., Heyraud, A., Greffe, L., Gilbert, M., Wakarchuk, W. W., Lam, J. S. & Samain, E. (2003). Large-scale in vivo synthesis of the carbohydrate moieties of gangliosides GM1 and GM2 by metabolically engineered *Escherichia coli*. *Chembiochem* 4, 406-412.

Antoine, T., Heyraud, A., Bosso, C. & Samain, E. (2005). Highly efficient biosynthesis of the oligosaccharide moiety of the GD3 ganglioside by using metabolically engineered *Escherichia coli*. *Angew Chem Int Ed Engl* 44, 1350-1352.

Bernatchez, S., Szymanski, C. M., Ishiyama, N., Li, J., Jarrell, H. C., Lau, P. C., Berghuis, A. M., Young, N. M. & Wakarchuk, W. W. (2005). A single bifunctional UDP-GlcNAc/Glc 4-epimerase supports the synthesis of three cell surface glycoconjugates in *Campylobacter jejuni*. *J Biol Chem* 280, 4792-4802.

- Bettler, E., Samain, E., Chazalet, V., Bosso, C., Heyraud, A., Joziassse, D. H., Wakarchuk, W. W., Imberty, A. & Geremia, A. R. (1999). The living factory: in vivo production of N-acetylglucosamine containing carbohydrates in *E. coli*. *Glycoconj J* 16, 205-212.
- 5 Cottaz, S. & Samain, E. (2005). Genetic engineering of *Escherichia coli* for the production of N(I),N(II)-diacetylchitobiose (chitinbiose) and its utilization as a primer for the synthesis of complex carbohydrates. *Metab Eng* 7, 311-317.
- 10 Creuzenet, C., Belanger, M., Wakarchuk, W. W. & Lam, J. S. (2000). Expression, purification, and biochemical characterization of WbpP, a new UDP-GlcNAc C4 epimerase from *Pseudomonas aeruginosa* serotype 06. *J Biol Chem* 275, 19060-19067.
- 15 Datsenko, K. A. & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-6645.
- Dumon, C., Samain, E. & Priem, B. (2004). Assessment of the two *Helicobacter pylori* alpha-1,3-fucosyltransferase ortholog genes for the large-scale synthesis of LewisX human milk oligosaccharides by metabolically engineered *Escherichia coli*. *Biotechnol Prog* 20, 412-419.
- 20 Dumon, C., Bosso, C., Utile, J. P., Heyraud, A. & Samain, E. (2005). Production of Lewis x Tetrasaccharides by Metabolically Engineered *Escherichia coli*. *Chembiochem* 7, 359-365.
- Dykxhoorn, D. M., St Pierre, R. & Linn, T. (1996). A set of compatible tac promoter expression vectors. *Gene* 177, 133-136.
- 25 Fort, S., Birikaki, L., Dubois, M. P., Antoine, T., Samain, E. & Driguez, H. (2005). Biosynthesis of conjugatable saccharidic moieties of GM2 and GM3 gangliosides by engineered *E. coli*. *Chem Commun (Camb)*, 2558-2560.
- 30 Ganguli, S., Zapata, G., Wallis, T., Reid, C., Boulnois, G., Vann, W. F. & Roberts, I. S. (1994). Molecular cloning and analysis of genes for sialic acid synthesis in *Neisseria meningitidis* group B and purification of the meningococcal CMP-NeuNAc synthetase enzyme. *J Bacteriol* 176, 4583-4589.
- 35 Gilbert, M., Bayer, R., Cunningham, A. M., DeFrees, S., Gao, Y., Watson, D. C., Young, N. M. & Wakarchuk, W. W. (1998). The synthesis of sialylated oligosaccharides using a CMP-Neu5Ac synthetase/sialyltransferase fusion. *Nat Biotechnol* 16, 769-772.
- 40 Gilbert, M., Karwaski, M. F., Bernatchez, S., Young, N. M., Taboada, E., Michniewicz, J., Cunningham, A. M. & Wakarchuk, W. W. (2002). The genetic bases for the variation in the lipo-oligosaccharide of the mucosal pathogen, *Campylobacter jejuni*. Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide. *J Biol Chem* 277, 327-337.
- Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. (1989). New method for generating deletions and gene replacements in *Escherichia coli*. *J Bacteriol* 171, 4617-4622.
- 45 Kalivoda, K. A., Steenbergen, S. M., Vimr, E. R. & Plumbridge, J. (2003). Regulation of sialic acid catabolism by the DNA binding protein NanR in *Escherichia coli*. *J Bacteriol* 185, 4806-4815.
- 50 Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., 2nd & Peterson, K. M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175-176.
- 55 Lee, J., Yi, J., Lee, S., Takahashi, S. & Kim, B. (2004). Production of N-acetylneuraminic acid from N acetylglucosamine and pyruvate using recombinant human renin binding protein and sialic aldolase in one pot. *Enzyme Microb Technol* 35, 121-125.
- Link, A. J., Phillips, D. & Church, G. M. (1997). Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* 179, 6228-6237.
- 60 Linton, D., Gilbert, M., Hitchen, P. G., Dell, A., Morris, H. R., Wakarchuk, W. W., Gregson, N. A. & Wren, B. W. (2000). Phase variation of a beta-1,3 galactosyltransferase involved in generation of the ganglioside GM1-like lipo-oligosaccharide of *Campylobacter jejuni*. *Mol Microbiol* 37, 501-514.
- 65 Martinez, E., Bartolome, B. & de la Cruz, F. (1988). pACYC184-derived cloning vectors containing the multiple cloning site and lacZ alpha reporter gene of pUC8/9 and pUC18/19 plasmids. *Gene* 68, 159-162.

- Maru, I., Ohnishi, J., Ohta, Y. & Tsukada, Y. (1998). Simple and large-scale production of N-acetylneuraminic acid from N-acetyl-D-glucosamine and pyruvate using N-acyl-D-glucosamine 2-epimerase and N-acetylneuraminidase. *Carbohydr Res* 306, 575-578.
- 5 Plumbridge, J. & Vimr, E. (1999). Convergent pathways for utilization of the amino sugars N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by *Escherichia coli*. *J Bacteriol* 181, 47-54.
- Priem, B., Gilbert, M., Wakarchuk, W. W., Heyraud, A. & Samain, E. (2002). A new fermentation process allows large-scale production of human milk oligosaccharides by metabolically engineered bacteria. *Glycobiology* 12, 235-240.
- 10 Samain, E., Drouillard, S., Heyraud, A., Driguez, H. & Geremia, R. A. (1997). Gram-scale synthesis of recombinant chitooligosaccharides in *Escherichia coli*. *Carbohydr Res* 302, 35-42.
- 15 Stapleton, A. E., Stroud, M. R., Hakomori, S. I. & Stamm, W. E. (1998). The globoseries glycosphingolipid sialosyl galactosyl globoside is found in urinary tract tissues and is a preferred binding receptor *In vitro* for uropathogenic *Escherichia coli* expressing pap-encoded adhesins. *Infect Immun* 66, 3856-3861.
- Vann, W. F., Tavares, J. J., Crowley, J., Vimr, E. & Silver, R. P. (1997). Purification and characterization of the *Escherichia coli* K1 neuB gene product N-acetylneuraminic acid synthetase. *Glycobiology* 7, 697-701.
- 20 Vann, W. F., Daines, D. A., Murkin, A. S., Tanner, M. E., Chaffin, D. O., Rubens, C. E., Vionnet, J. & Silver, R. P. (2004). The NeuC protein of *Escherichia coli* K1 is a UDP N-acetylglucosamine 2-epimerase. *J Bacteriol* 186, 706-712.
- 25 Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 3, 97-130.
- Vimr, E. R. & Troy, F. A. (1985). Identification of an inducible catabolic system for sialic acids (nan) in *Escherichia coli*. *J Bacteriol* 164, 845-853.
- 30 Wang, R. F. & Kushner, S. R. (1991). Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* 100, 195-199.
- Werner, I. & Odin, L. (1952). On the presence of sialic acid in certain glycoproteins and in gangliosides. *Acta Soc Med Ups* 57, 230-241.
- 35 Yamamoto, T., Nakashizuka, M. & Terada, I. (1998). Cloning and expression of a mannan bacterial beta-galactoside alpha2,6-sialyltransferase gene from *Photobacterium damsela* JT0160. *J Biochem (Tokyo)* 123, 94-100.
- 40 Zhang, X. & Kiechle, F. L. (2004). Review: Glycosphingolipids in health and disease. *Ann Clin Lab Sci* 34, 3-13.

Listado de secuencias

- 45 <110> Centre National de la Recherche Scientifique (CNRS) SAMAIN, Eric
- <120> Procedimiento de producción de oligosacáridos sialilados
- <130> D24149
- 50 <150> 60/780.350
- <151> 2006-03-09
- <160> 19
- 55 <170> PatentIn versión 3.3
- <210> 1
- <211> 876
- 60 <212> ADN
- <213> Secuencia artificial
- <220>
- <223> *Campylobacter* glucosiltransferasas, número de registro NCBI AR271701
- 65 <400> 1

ES 2 456 292 T3

atgaaaaaag	ttattattgc	tggaaatgga	ccaagtttaa	aagaaattga	ttattcaagg	60
ctaccaaatg	atthtgatgt	atthgatgac	aatcaattht	atthtgaaga	taaatactat	120
cttggtaaaa	aattcaaaagc	agtatthttac	aatcctgggc	ttttttttga	acaatactac	180
actttaaaac	atthaatcca	aatcaagaa	tatgagaccg	aactaattat	gtgttcta	240
tacaaccaag	ctcatctaga	aatgaaaat	tttgtaaaaa	ctttttacga	ttattttcct	300
gatgctcatt	tgggatatga	tttttttaaa	caacttaaaag	aatttaaatgc	ttatttttaa	360
tttcaogaaa	tttatctcaa	tcaaagaatt	acctcaggag	tctatatgtg	tgcagtagct	420
atagccctag	gatacaaaaga	aatttatctt	tctggaattg	atthttatca	aaatgggtca	480
tcttatgctt	ttgatacca	acaagaaaat	cttttaaaac	tggctcctga	ttttaaaaat	540
gatcgctcac	actatatcgg	acatagtaaa	aatacagata	taaaagcttt	agaatttcta	600
gaaaaaactt	acaaaataaa	actatattgc	ttatgtccta	acagtctttt	agcaaattht	660
atagaactag	cgccaaatth	aaattcaaat	tttatcatac	aagaaaaaaa	taactacact	720
aaagatatac	tcataccttc	tagtgaggct	tatggaaaaat	tttcaaaaaa	tattaattht	780
aaaaaaataa	aaattaaaga	aaatatttht	tacaagttga	taaaagatct	attaagatta	840
cctagtata	taaagcatta	tttcaaagga	aaataa			876

<210> 2

<211> 876

5 <212> ADN

<213> Campylobacter jejuni

<220>

10 <223> Alfa-2,3/alfa 2,8-sialiltransferasa Campylobacter sialiltransferasa II, número de registro NCBI AX934427

<400> 2

atgaaaaaag	ttattattgc	tggaaatgga	ccaagtttaa	aagaaattga	ttattcaagg	60
ctaccaaatg	atthtgatgt	atthgatgac	aatcaattht	atthtgaaga	taaatactat	120
cttggtaaaa	aattcaaaagc	agtatthttac	aatcctgggc	ttttttttga	acaatactac	180
actttaaaac	atthaatcca	aatcaagaa	tatgagaccg	aactaattat	gtgttcta	240
tacaaccaag	ctcatctaga	aatgaaaat	tttgtaaaaa	ctttttacga	ttattttcct	300
gatgctcatt	tgggatatga	tttttttaaa	caacttaaaag	aatttaaatgc	ttatttttaa	360
tttcaogaaa	tttatctcaa	tcaaagaatt	acctcaggag	tctatatgtg	tgcagtagct	420
atagccctag	gatacaaaaga	aatttatctt	tctggaattg	atthttatca	aaatgggtca	480
tcttatgctt	ttgatacca	acaagaaaat	cttttaaaac	tggctcctga	ttttaaaaat	540
gatcgctcac	actatatcgg	acatagtaaa	aatacagata	taaaagcttt	agaatttcta	600
gaaaaaactt	acaaaataaa	actatattgc	ttatgtccta	acagtctttt	agcaaattht	660
atagaactag	cgccaaatth	aaattcaaat	tttatcatac	aagaaaaaaa	taactacact	720
aaagatatac	tcataccttc	tagtgaggct	tatggaaaaat	tttcaaaaaa	tattaattht	780
aaaaaaataa	aaattaaaga	aaatatttht	tacaagttga	taaaagatct	attaagatta	840
cctagtata	taaagcatta	tttcaaagga	aaataa			876

15 <210> 3

<211> 876

<212> ADN

20 <213> Secuencia artificial

<220>

<223> Campylobacter glucosiltransferasas, número de registro NCBI AR271702

<400> 3

atgaaaaaag	ttattattgc	tggaaatgga	ccaagtttaa	aagaaattga	ttattcaaga	60
ctaccaaatg	atthtgatgt	atthgatgac	aatcaattht	atthtgaaga	taaatactat	120
cttggtaaaa	aattcaaaagc	agtatthttac	aatcctagtc	ttttttttga	acaatactac	180
actttaaaac	atthaatcca	aatcaagaa	tatgagaccg	aactaatcat	gtgttcta	240
tttaaccaag	ctcatctaga	aatcaaaaat	tttgtaaaaa	ctttttacga	ttattttcct	300
gatgctcatt	tgggatatga	ttttttcaaa	caacttaaaag	aattcaaatgc	ttatttttaa	360
tttcaogaaa	tttatthcaa	tcaaagaatt	acctcagggg	tctatatgtg	cacagtagcc	420
atagccctag	gatacaaaaga	aatttatctt	tctggaattg	atthttatca	aaatggatca	480
tcttatgctt	ttgatacca	acaaaaaaat	cttttaaaat	tggctcctaa	ttttaaaaat	540
gataattcac	actatatcgg	acatagtaaa	aatacagata	taaaagcttt	agaatttcta	600
gaaaaaactt	acgaaataaa	gctatattgt	ttatgtccta	acagtctttt	agcaaattht	660
atagaactag	cgccaaatth	aaattcaaat	tttatcatac	aagaaaaaaa	taactatact	720
aaagatatac	tcataccttc	tagtgaggct	tatggaaaaat	ttcaaaaaa	tattaattht	780
aaaaaaataa	aaattaaaga	aaatatttht	tacaagttga	taaaagatct	attaagatta	840
cctagtata	taaagcatta	tttcaaagga	aaataa			876

<210> 4

ES 2 456 292 T3

<211> 876
 <212> ADN
 <213> Campylobacter jejuni

5 <220>
 <223> Campylobacter glucosiltransferasas, número de registro NCBI AX934429

<400> 4

```

atgaaaaaag ttattattgc tgqaaatgga ccaagtttaa aagaaattga ttattcaaga      60
ctaccaaatg attttgatgt atttgatagc aatcaatttt attttgaaga taaatactat      120
cttggtaaaa aatgcaaagc agtattttac aatcctagtc ttttttttga acaatactac      180
actttaaaac atttaaatcca aaatcaagaa tatgagaccg aactaatcat gtgttctaata      240
tttaaccaag ctcatctaga aaatcaaaat tttgtaaaaa ctttttaaga ttattttcct      300
gatgctcatt tgggatatga ttttttcaaa caacttaaag aattcaatgc ttattttaaa      360
tttcacgaaa tttatttcaa tcaaaagaatt acctcagggg tctatatgtg cacagtagcc      420
atagccctag gatacaaaaga aatttatctt tcgggaattg atttttatca aaatggatca      480
tcttatgctt ttgataccaa acaaaaaaaal cttttaaaat tggctcctaa ttttaaaaaat      540
gataattcaq actatatcgg acatagtaaa aatacagata taaaagcttt agaattttta      600
gaaaaaactt acgaaataaa gctatattgt ttatgtccta acagtctttt agcaaatttt      660
atagaactay cgccaaatlt aaattcaaat tttatcatat aagaaaaaaa taactatact      720
aaagatatac tcataccttc taqtgaggct tatggaaaaa ttacaaaaaa tattaatttt      780
aaaaaaaaaa aaattaaaga aaatatttat tacaagttga taaaagatct attaagatta      840
cctagtgata taaagcatta tttcaaagga aaataa      876
  
```

<210> 5
 <211> 1116

15 <212> ADN
 <213> Neisseria meningitidis

<220>
 <223> nst alfa-2,3-sialiltransferasa, número de registro GenBank U60660

20 <400> 5

```

atgggcttga aaaaggcttg tttgaccgtg ttgtgttga ttgtttttt tttcgggata      60
ttttatacat ttgaccgggt aaatcagggg gaaaggaatg cggtttccct qctgaaggag      120
aaacttttca atgaagaggg ggaaccggtc aatctgattt tctgttatac catattgcag      180
atgaaggtgg cggaaaggat tatggcgag catccgggcy agcgglttta tgygtgcty      240
atgtctgaaa acaggaatga aaaatacgat tattatttca atcagataaa qgataaggcg      300
gagcgggctg actttttcca cctgcctac ggtttgaaca aatcgtttaa tttcattccg      360
acgatggcgg agctgaaggt aaagtcgat ctgtgcgca aaytcaagcy gatttattty      420
gcaagttttg aaaaagtcag cattgccgcc ttttgagca cttaccggga tgcggaaatc      480
aaaacttttg acgacgggac aggcaattta attcaaagca gcagctattt ggcgatgag      540
ttttctgtaa acgggacgat caagcgggat tttgcccgga tgatgatcgg agattggagc      600
atcgccaaaa cccgcaatgc ttccgacgag cattacagca tattcaaggg tttgaaaaac      660
attatggagc acggcgcgg caagatgact tacctgcgcy tgttcgatgc gtcggaactg      720
aagaacgggg acgaaacggg cggcacgggt cggatacttt tgggttcgcy cyacaaagag      780
atgaagggaaa tttcggaaaa ggcggcaaaa aacttcaaaa tacaatatgt cgcgccgat      840
ccccgcctaaa cctacgggct ttccggcgta accacattaa attcgcctta tgtcatcgaa      900
gactatattt tgcgcgagat taagaaaaac ccgcatacga ggtatgaaat ttataccttt      960
ttcagcggcg cggcgttgac gatgaaggat tttcccaatg tgcacgttta cgcattgaaa      1020
ccggcttccc ttccggaaga ttattggctc aagccgggtg atgcctgtt tacccaatcc      1080
ggcatccoga ttttgacatt tgacgataaa aattaa      1116
  
```

25 <210> 6
 <211> 371
 <212> PRT
 <213> Neisseria meningitidis

30 <220>
 <223> Alfa-2,3-sialiltransferasa, Proteína id AAC44541.1

<400> 6

ES 2 456 292 T3

Met Gly Leu Lys Lys Ala Cys Leu Thr Val Leu Cys Leu Ile Val Phe
1 5 10 15
Cys Phe Gly Ile Phe Tyr Thr Phe Asp Arg Val Asn Gln Gly Glu Arg
20 25 30
Asn Ala Val Ser Leu Leu Lys Glu Lys Leu Phe Asn Glu Glu Gly Glu
35 40 45
Pro Val Asn Leu Ile Phe Cys Tyr Thr Ile Leu Gln Met Lys Val Ala
50 55 60
Glu Arg Ile Met Ala Gln His Pro Gly Glu Arg Phe Tyr Val Val Leu
65 70 75 80
Met Ser Glu Asn Arg Asn Glu Lys Tyr Asp Tyr Tyr Phe Asn Gln Ile
85 90 95
Lys Asp Lys Ala Glu Arg Ala Tyr Phe Phe His Leu Pro Tyr Gly Leu
100 105 110
Asn Lys Ser Phe Asn Phe Ile Pro Thr Met Ala Glu Leu Lys Val Lys
115 120 125
Ser Met Leu Leu Pro Lys Val Lys Arg Ile Tyr Leu Ala Ser Leu Glu
130 135 140
Lys Val Ser Ile Ala Ala Phe Leu Ser Thr Tyr Pro Asp Ala Glu Ile
145 150 155 160
Lys Thr Phe Asp Asp Gly Thr Gly Asn Leu Ile Gln Ser Ser Ser Tyr
165 170 175
Leu Gly Asp Glu Phe Ser Val Asn Gly Thr Ile Lys Arg Asn Phe Ala
180 185 190
Arg Met Met Ile Gly Asp Trp Ser Ile Ala Lys Thr Arg Asn Ala Ser
195 200 205
Asp Glu His Tyr Thr Ile Phe Lys Gly Leu Lys Asn Ile Met Asp Asp
210 215 220
Gly Arg Arg Lys Met Thr Tyr Leu Pro Leu Phe Asp Ala Ser Glu Leu
225 230 235 240
Lys Thr Gly Asp Glu Thr Gly Gly Thr Val Arg Ile Leu Leu Gly Ser
245 250 255
Pro Asp Lys Glu Met Lys Glu Ile Ser Glu Lys Ala Ala Lys Asn Phe
260 265 270
Lys Ile Gln Tyr Val Ala Pro His Pro Arg Gln Thr Tyr Gly Leu Ser
275 280 285
Gly Val Thr Thr Leu Asn Ser Pro Tyr Val Ile Glu Asp Tyr Ile Leu
290 295 300
Arg Glu Ile Lys Lys Asn Pro His Thr Arg Tyr Glu Ile Tyr Thr Phe
305 310 315 320
Phe Ser Gly Ala Ala Leu Thr Met Lys Asp Phe Pro Asn Val His Val
325 330 335
Tyr Ala Leu Lys Pro Ala Ser Leu Pro Glu Asp Tyr Trp Leu Lys Pro
340 345 350
Val Tyr Ala Leu Phe Thr Gln Ser Gly Ile Pro Ile Leu Thr Phe Asp
355 360 365
Asp Lys Asn
370

<210> 7

5 <211> 933

<212> ADN

<213> Neisseria meningitidis

<220>

10 <223> Gen LgtC, GenBank U65788

ES 2 456 292 T3

<400> 7
atggacatcg tatttgccgc agacgacaac tatgccgcct atctttgctg tgccggcaaaa 60
agcgtggaag cggcccatcc cgatacggaa atcaggttcc acgtcctcga tgccggcatc 120
agtgaggcaa accggggcggc ggttgcctgc aatttgccgg gggggggtaa tatccgcttt 180
atagacgtaa accccgaaga tttgccgcgc ttccccttaa acatcaggca cattccatc 240
acgacttatg cccgcttgaa atfgggcgaa tacattgccg attgcgataa agtccctgat 300
ctggatatag acgtattggt cagggacagc ctgaagccct tatgggatac cqatttgggc 360
gataactggc ttggcgcgtg catcgatttg tttgtcgaaa ggcagaatgc ttacaaacaa 420
aaaatcggta tggcggacgg agaataattt ttcaatgccg gcglattgct gatcaacctg 480
aaaaagtggc ggcagcacga tattttcaaa atggcctgcg aatgggtgga acaatacaag 540
gacgtgatgc aatatcagga tcaggatatt ttgaacgggc tgtttaaagg cggggtgtgt 600
tatgcgaaca gccgtttcaa ctttatgccg accaatgatg cctttatggc gaacaggttt 660
gcgtcccgcc ataccgaccc gctttaccgc qaccggactt atacggcgat gcctgtcggc 720
gtcagccatt attgcggccc ggcaaagccg tggcacaggg actgcaccgc gtggggtgcg 780
gaacgtttca cagaattggc gggcagcctg acgagcgttc ccgaagaatg gcgcggcaaa 840
cttgcgcgtc cgcaccgtgt gttccgaca aagcgtatgc ttcaaatgat gcgcagaaa 900
ctgtctgccca gattcttacg caagatttat tga 933

<210> 8

<211> 323

5

<212> PRT

<213> Neisseria meningitidis

<220>

<223> Proteína LgTC, proteína id AAB48385.

10

<400> 8

Met Glu Asn Cys Pro Leu Val Ser Val Ile Val Cys Ala Tyr Asn Ala
1 5 10 15
Glu Gln Tyr Ile Asp Glu Ser Ile Ser Ser Ile Ile Asn Gln Thr Tyr
20 25 30
Glu Asn Leu Glu Ile Ile Val Ile Asn Asp Gly Ser Thr Asp Leu Thr
35 40 45
Leu Ser His Leu Glu Glu Ile Ser Lys Leu Asp Lys Arg Ile Lys Ile
50 55 60
Ile Ser Asn Lys Tyr Asn Leu Gly Phe Ile Asn Ser Leu Asn Ile Gly
65 70 75 80
Leu Gly Cys Phe Ser Gly Lys Tyr Phe Ala Arg Met Asp Ala Asp Asp
85 90 95
Ile Ala Lys Pro Ser Trp Ile Glu Lys Ile Val Thr Tyr Leu Glu Lys
100 105 110
Asn Asp His Ile Thr Ala Met Gly Ser Tyr Leu Glu Ile Ile Val Glu
115 120 125
Lys Glu Cys Gly Ile Ile Gly Ser Gln Tyr Lys Thr Gly Asp Ile Trp
130 135 140
Lys Asn Pro Leu Leu His Asn Asp Ile Cys Glu Ala Met Leu Phe Tyr
145 150 155 160
Asn Pro Ile His Asn Asn Thr Met Ile Met Arg Ala Asn Val Tyr Arg
165 170 175
Glu His Lys Leu Ile Phe Asn Lys Asp Tyr Pro Tyr Ala Glu Asp Tyr
180 185 190
Lys Phe Trp Ser Glu Val Ser Arg Leu Gly Cys Leu Ala Asn Tyr Pro
195 200 205
Glu Ala Leu Val Lys Tyr Arg Leu His Gly Asn Gln Thr Ser Ser Val
210 215 220
Tyr Asn His Glu Gln Asn Glu Thr Ala Lys Lys Ile Lys Arg Glu Asn
225 230 235 240
Ile Thr Tyr Tyr Leu Asn Lys Ile Gly Ile Asp Ile Lys Val Ile Asn
245 250 255
Ser Val Ser Leu Leu Glu Ile Tyr His Val Asp Lys Ser Asn Lys Val
260 265 270
Leu Lys Ser Ile Leu Tyr Glu Met Tyr Met Ser Leu Asp Lys Tyr Thr
275 280 285
Ile Thr Ser Leu Leu His Phe Ile Lys Tyr His Leu Glu Leu Phe Asp
290 295 300
Leu Lys Gln Asn Leu Lys Ile Ile Lys Lys Phe Ile Arg Lys Ile Asn
305 310 315 320
Val Ile Phe

ES 2 456 292 T3

<210> 9
 <211> 972
 <212> ADN
 <213> Haemophilus influenzae

5

<220>
 <223> Gen lgtD, GenBank bajo el número de registro U32832

<400> 9

10

```

atggaaaatt gtccattagt atcggttatt gtttgtgctt ataacgctga gcaatatata      60
gatgaaagca tttcatccat tattaatcag acttatgaaa atctagaaat tatagttatc      120
aatgatgggt caacagattt gactttgtct catttagaag aaatatctaa attagataaa      180
aggataaaaa ttatcagtaa taaatataat ttagggttca taaattcttt gaatataggc      240
cttggttggg tttcaggtaa atattttgca agaatggatg ctgatgatag agctaaacca      300
tcgtggattg agaaaatagt tacctatctg gagaaaaatg atcatattac agcaatggga      360
tcatacttag agattattgt agaaaaagaa tgtggaatta tcggttctca atataaaact      420
ggagataatg gaaaaaatcc attgctacat aatgatattt gtgaagctat gcttttctat      480
aatccgatac ataacaacac tatgattatg agagcaaatg tatatagaga gcataaatta      540
atctttaata aaqattatcc gtatgcagaa gattataagt tttggtcaga ggtagtagg      600
cttggttggg tagctaatta tcctgaagca ttagtaaaat atagactaca tggaaaccaa      660
acatcatcag tttataatca tgaycaaaat gagacagcta aaaagataaa gagggaaaat      720
attacataat acccttaataa gataggataa gatataaaag taattaatag tgtgtcgttg      780
ctagaaatat atcatgtgga taaaagtaat aaagtgttga aaagtatact ttatgagatg      840
tatatgagct tagataaata tactataact tcactcttac attttattaa atatcatctt      900
gaattatttg atttaaagca aaatttaaag attataaaaa agttcataag aaaataaat      960
ggtatatttt ag                                     972
    
```

<210> 10
 <211> 323
 <212> PRT
 <213> Haemophilus influenzae

15

<220>
 <223> lgtD, proteína id AAC23227

20

<400> 10

```

Met Glu Asn Cys Pro Leu Val Ser Val Ile Val Cys Ala Tyr Asn Ala
1      5      10      15
Glu Gln Tyr Ile Asp Glu Ser Ile Ser Ser Ile Ile Asn Gln Thr Tyr
20      25      30
Glu Asn Leu Glu Ile Ile Val Ile Asn Asp Gly Ser Thr Asp Leu Thr
35      40      45
Leu Ser His Leu Glu Glu Ile Ser Lys Leu Asp Lys Arg Ile Lys Ile
50      55      60
Ile Ser Asn Lys Tyr Asn Leu Gly Phe Ile Asn Ser Leu Asn Ile Gly
65      70      75      80
Leu Gly Cys Phe Ser Gly Lys Tyr Phe Ala Arg Met Asp Ala Asp Asp
85      90      95
Ile Ala Lys Pro Ser Trp Ile Glu Lys Ile Val Thr Tyr Leu Glu Lys
100      105      110
Asn Asp His Ile Thr Ala Met Gly Ser Tyr Leu Glu Ile Ile Val Glu
115      120      125
Lys Glu Cys Gly Ile Ile Gly Ser Gln Tyr Lys Thr Gly Asp Ile Trp
130      135      140
Lys Asn Pro Leu Leu His Asn Asp Ile Cys Glu Ala Met Leu Phe Tyr
145      150      155      160
Asn Pro Ile His Asn Asn Thr Met Ile Met Arg Ala Asn Val Tyr Arg
165      170      175
Glu His Lys Leu Ile Phe Asn Lys Asp Tyr Pro Tyr Ala Glu Asp Tyr
180      185      190
Lys Phe Trp Ser Glu Val Ser Arg Leu Gly Cys Leu Ala Asn Tyr Pro
195      200      205
Glu Ala Leu Val Lys Tyr Arg Leu His Gly Asn Gln Thr Ser Ser Val
210      215      220
Tyr Asn His Glu Gln Asn Glu Thr Ala Lys Lys Ile Lys Arg Glu Asn
225      230      235      240
Ile Thr Tyr Tyr Leu Asn Lys Ile Gly Ile Asp Ile Lys Val Ile Asn
    
```

ES 2 456 292 T3

```

                245                250                255
Ser Val Ser Leu Leu Glu Ile Tyr His Val Asp Lys Ser Asn Lys Val
                260                265                270
Leu Lys Ser Ile Leu Tyr Glu Met Tyr Met Ser Leu Asp Lys Tyr Thr
                275                280                285
Ile Thr Ser Leu Leu His Phe Ile Lys Tyr His Leu Glu Leu Phe Asp
                290                295                300
Leu Lys Gln Asn Leu Lys Ile Ile Lys Lys Phe Ile Arg Lys Ile Asn
305                310                315                320
Val Ile Phe

```

- 5
<210> 11
<211> 56
<212> ADN
<213> Secuencia artificial
- 10
<220>
<223> Cebador 1 aguas arriba
- <400> 11
- gcaattatg attcggcgga tggttgccg atggtggtg aggctggagc tgcttc
56
- 15
<210> 12
<211> 58
<212> ADN
<213> Secuencia artificial
- 20
<220>
<223> Cebador aguas abajo
- <400> 12
- 25
ctcgtcacc tgccggcgc gcgtgaaaat agtttcgca tatgaatc ctccttag
58
- 30
<210> 13
<211> 58
<212> ADN
<213> Secuencia artificial
- <220>
<223> Cebador 2 aguas arriba
- 35
<400> 13
- cactggcgat tgatcgcg ggtactaac ttgccccgt gtaggctgga gctgcttc
58
- 40
<210> 14
<211> 42
<212> ADN
<213> Secuencia artificial
- 45
<220>
<223> Cebador izquierdo
- <400> 14
- 50
ggtacctaag gaggaaaata aatgaaagaa ataaaaatac aa
42
- 55
<210> 15
<211> 35
<212> ADN
<213> Secuencia artificial
- <220>
<223> Cebador derecho

ES 2 456 292 T3

<400> 15
ctcgagttaa gtctctaac gattgtttc caatg 35

5 <210> 16
<211> 20
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Cebador

<400> 16

15 caatgccaaa tatggggaac 20

<210> 17
<211> 23
<212> ADN
20 <213> Secuencia artificial

<220>
<223> Cebador

25 <400> 17

gcggccgcgt ctttctggc taa 23

<210> 18
30 <211> 425
<212> PRT
<213> Helicobacter pylori

<220>
35 <223> FutA, ATCC 26695

<400> 18

Met Phe Gln Pro Leu Leu Asp Ala Phe Ile Glu Ser Ala Ser Ile Glu
1 5 10 15
Lys Met Ala Ser Lys Ser Pro Pro Pro Pro Leu Lys Ile Ala Val Ala
20 25 30
Asn Trp Trp Gly Asp Glu Glu Ile Lys Glu Phe Lys Lys Ser Val Leu
35 40 45
Tyr Phe Ile Leu Ser Gln Arg Tyr Ala Ile Thr Leu His Gln Asn Pro
50 55 60
Asn Glu Phe Ser Asp Leu Val Phe Ser Asn Pro Leu Gly Ala Ala Arg
65 70 75 80
Lys Ile Leu Ser Tyr Gln Asn Thr Lys Arg Val Phe Tyr Thr Gly Glu
85 90 95
Asn Glu Ser Pro Asn Phe Asn Leu Phe Asp Tyr Ala Ile Gly Phe Asp
100 105 110
Glu Leu Asp Phe Asn Asp Arg Tyr Leu Arg Met Pro Leu Tyr Tyr Ala

ES 2 456 292 T3

115 120 125
 His Leu His Tyr Lys Ala Glu Leu Val Asn Asp Thr Thr Ala Pro Tyr
 130 135 140
 Lys Leu Lys Asp Asn Ser Leu Tyr Ala Leu Lys Lys Pro Ser His His
 145 150 155 160
 Phe Lys Glu Asn His Pro Asn Leu Cys Ala Val Val Asn Asp Glu Ser
 165 170 175
 Asp Leu Leu Lys Arg Gly Phe Ala Ser Phe Val Ala Ser Asn Ala Asn
 180 185 190
 Ala Pro Met Arg Asn Ala Phe Tyr Asp Ala Leu Asn Ser Ile Glu Pro
 195 200 205
 Val Thr Gly Gly Gly Ser Val Arg Asn Thr Leu Gly Tyr Lys Val Gly
 210 215 220
 Asn Lys Ser Glu Phe Leu Ser Gln Tyr Lys Phe Asn Leu Cys Phe Glu
 225 230 235 240
 Asn Ser Gln Gly Tyr Gly Tyr Val Thr Glu Lys Ile Leu Asp Ala Tyr
 245 250 255
 Phe Ser His Thr Ile Pro Ile Tyr Trp Gly Ser Pro Ser Val Ala Lys
 260 265 270
 Asp Phe Asn Pro Lys Ser Phe Val Asn Val His Asp Phe Asn Asn Phe
 275 280 285
 Asp Glu Ala Ile Asp Tyr Ile Lys Tyr Leu His Thr His Pro Asn Ala
 290 295 300
 Tyr Leu Asp Met Leu Tyr Glu Asn Pro Leu Asn Thr Leu Asp Gly Lys
 305 310 315 320
 Ala Tyr Phe Tyr Gln Asp Leu Ser Phe Lys Lys Ile Leu Asp Phe Phe
 325 330 335
 Lys Thr Ile Leu Glu Asn Asp Thr Ile Tyr His Lys Phe Ser Thr Ser
 340 345 350
 Phe Met Trp Glu Tyr Asp Leu His Lys Pro Leu Val Ser Ile Asp Asp
 355 360 365
 Leu Arg Val Asn Tyr Asp Asp Leu Arg Val Asn Tyr Asp Arg Leu Leu
 370 375 380
 Gln Asn Ala Ser Pro Leu Leu Glu Leu Ser Gln Asn Thr Thr Phe Lys
 385 390 395 400
 Ile Tyr Arg Lys Ala Tyr Gln Lys Ser Leu Pro Leu Leu Arg Ala Val
 405 410 415
 Arg Lys Leu Val Lys Lys Leu Gly Leu
 420 425

<210> 19

<211> 476

5 <212> PRT

<213> Helicobacter pylori

<220>

<223> FutB (gen HP 0651), ATCC 26695

10

<400> 19

Met Phe Gln Pro Leu Leu Asp Ala Phe Ile Glu Ser Ala Ser Ile Glu
 1 5 10 15
 Lys Met Val Ser Lys Ser Pro Pro Pro Leu Lys Ile Ala Val Ala
 20 25 30
 Asn Trp Trp Gly Asp Glu Glu Ile Lys Glu Phe Lys Lys Ser Val Leu
 35 40 45
 Tyr Phe Ile Leu Ser Gln Arg Tyr Ala Ile Thr Leu His Gln Asn Pro
 50 55 60
 Asn Glu Ser Ser Asp Leu Val Phe Ser Asn Pro Leu Gly Ala Ala Arg
 65 70 75 80

ES 2 456 292 T3

Lys Ile Leu Ser Tyr Gln Asn Thr Lys Arg Val Phe Tyr Thr Gly Glu
 85 90 95
 Asn Glu Ser Pro Asn Phe Asn Leu Phe Asp Tyr Ala Ile Gly Phe Asp
 100 105 110
 Glu Leu Asp Phe Asn Asp Arg Tyr Leu Arg Met Pro Leu Tyr Tyr Ala
 115 120 125
 His Leu His Tyr Glu Ala Glu Leu Val Asn Asp Thr Thr Ala Pro Tyr
 130 135 140
 Lys Leu Lys Asp Asn Ser Leu Tyr Ala Leu Lys Lys Pro Ser His His
 145 150 155 160
 Phe Lys Glu Asn His Pro Asn Leu Cys Ala Val Val Asn Asp Glu Ser
 165 170 175
 Asp Leu Leu Lys Arg Gly Phe Ala Ser Phe Val Ala Ser Asn Ala Asn
 180 185 190
 Ala Pro Met Arg Asn Ala Phe Tyr Asp Ala Leu Asn Ser Ile Glu Pro
 195 200 205
 Val Thr Gly Gly Gly Ser Val Arg Asn Thr Leu Gly Tyr Lys Val Gly
 210 215 220
 Asn Lys Ser Glu Phe Leu Ser Gln Tyr Lys Phe Asn Leu Cys Phe Glu
 225 230 235 240
 Asn Ser Gln Gly Tyr Gly Tyr Val Thr Glu Lys Ile Leu Asp Ala Tyr
 245 250 255
 Phe Ser His Thr Ile Pro Ile Tyr Trp Gly Ser Pro Ser Val Ala Lys
 260 265 270
 Asp Phe Asn Pro Lys Ser Phe Val Asn Val His Asp Phe Asn Asn Phe
 275 280 285
 Asp Glu Ala Ile Asp Tyr Ile Lys Tyr Leu His Thr His Pro Asn Ala
 290 295 300
 Tyr Leu Asp Met Leu Tyr Glu Asn Pro Leu Asn Thr Leu Asp Gly Lys
 305 310 315 320
 Ala Tyr Phe Tyr Gln Asp Leu Ser Phe Lys Lys Ile Leu Asp Phe Phe
 325 330 335
 Lys Thr Ile Leu Glu Asn Asp Thr Ile Tyr His Asn Asn Pro Phe Ile
 340 345 350
 Phe Tyr Arg Asp Leu His Glu Pro Leu Ile Ser Ile Asp Asp Leu Arg
 355 360 365
 Val Asn Tyr Asp Asp Leu Arg Val Asn Tyr Asp Asp Leu Arg Val Asn
 370 375 380
 Tyr Asp Asp Leu Arg Val Asn Tyr Asp Asp Leu Arg Val Asn Tyr Asp
 385 390 395 400
 Asp Leu Arg Val Asn Tyr Asp Asp Leu Arg Val Asn Tyr Asp Asp Leu
 405 410 415
 Arg Val Asn Tyr Asp Asp Leu Arg Val Asn Tyr Asp Asp Leu Arg Val
 420 425 430
 Asn Tyr Asp Arg Leu Leu Gln Asn Ala Ser Pro Leu Leu Glu Leu Ser
 435 440 445
 Gln Asn Thr Thr Phe Lys Ile Tyr Arg Lys Ala Tyr Gln Lys Ser Leu
 450 455 460
 Pro Leu Leu Arg Thr Ile Arg Arg Trp Val Lys Lys
 465 470 475

REIVINDICACIONES

1. Procedimiento para la producción de oligosacáridos que comprenden al menos un resto de ácido siálico, a los que se hace referencia en la presente memoria como oligosacáridos sialilados, que comprende la etapa de cultivar un microorganismo en un medio de cultivo, que comprende opcionalmente un precursor exógeno, en el que dicho microorganismo comprende genes heterólogos que codifican una CMP-Neu5Ac sintetasa, una ácido siálico sintasa, una GlcNAc-6-fosfato 2-epimerasa y una sialiltransferasa, y en el que los genes endógenos que codifican la ácido siálico aldolasa (*NanA*) y la ManNac cinasa (*NanK*) se han suprimido o inactivado, pudiendo dicho microorganismo producir ácido siálico activado internamente.
2. Procedimiento según la reivindicación 1, en el que el gen heterólogo de sialiltransferasa se selecciona de entre α -2,3-sialiltransferasa, α -2,3- y α -2,8-sialiltransferasa (*cstII*), y α -2,6-sialiltransferasa.
3. Procedimiento según una de las reivindicaciones 1 a 2, en el que el gen heterólogo de CMP-Neu5Ac sintetasa es *neuA*, el gen heterólogo de ácido siálico sintasa es *neuB* y el gen heterólogo GlcNAc-6-fosfato 2-epimerasa es *neuC*.
4. Procedimiento según la reivindicación 3, en el que los genes *neuA*, *neuB* y *neuC* se aíslan de cepas bacterianas que contienen la estructura sialilada en sus membranas celulares, tales como la cepa de *C. jejuni* nº de registro 43438 de la ATCC.
5. Procedimiento según una de las reivindicaciones 1 a 4, en el que los genes *nanT*, *NanA*, *NanK* y *nanE* se suprimen o inactivan.
6. Procedimiento según una de las reivindicaciones 1 a 5, en el que dicho microorganismo codifica además una proteína que facilita la absorción de lactosa y carece de enzimas que metabolizan la lactosa.
7. Procedimiento según una de las reivindicaciones 1 a 6, en el que dicho microorganismo es una cepa de *E. coli*, que es *LacY+*, *LacZ-* y opcionalmente *MeIA-*.
8. Procedimiento según una de las reivindicaciones 1 a 7, en el que el precursor exógeno se selecciona de entre lactosa, galactosa, β -galactósido y α -galactósido, tal como globotriosa (Gal α -4Gal β -4Glc).
9. Microorganismo como se define en cualquiera de las reivindicaciones 1 a 7.
10. Medio de cultivo celular que comprende lactosa como precursor y el microorganismo de la reivindicación 9.
11. Procedimiento según una de las reivindicaciones 1 a 8, para la producción de 3'sialil-lactosa o 6'sialil-lactosa, en el que el microorganismo se cultiva a una alta densidad celular en un sustrato de carbono, tal como glucosa o glicerol, y se alimenta con lactosa que está interiorizada por la lactosa permeasa y sialilada por dicha sialiltransferasa recombinante utilizando el CMP-Neu5Ac endógenamente generado a partir de UDP-GlcNAc.
12. Procedimiento según cualquiera de las reivindicaciones 1 a 8 para la producción de GD3, en el que el gen heterólogo de sialiltransferasa es una α -2,3 y α -2,8 sialiltransferasa bifuncional, tal como el gen *cstII* de *Campylobacter jejuni* depositado en la ATCC con nº de registro 43438, que cataliza la transferencia de un resto sialilo o de una molécula de ácido siálico activado producida internamente a la Neu5Ac α -3Gal β -4Glc (GM3) para formar Neu5Ac α -8Neu5Ac α -3Gal β -4Glc (GD3).
13. Procedimiento según la reivindicación 12, que se extiende a la producción de GT3, en el que el microorganismo se cultiva además de tal manera que la sialiltransferasa bifuncional cataliza la transferencia de un resto sialilo de una molécula activada de ácido siálico producido internamente del Neu5Ac α -8Neu5Ac α -3Gal β -4Glc (GD3) para formar Neu5Ac5 α -8Neu5Ac α -8Neu5Ac α -3Gal β -4Glc (GT3).
14. Procedimiento según la reivindicación 11, que se extiende a la producción de la parte de hidratos de carbono del gangliósido Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GM1), en el que el microorganismo comprende además secuencias heterólogas que codifican β -1,4GalNAc transferasa y β -1,3-galactosiltransferasa, transfiriendo dicha β -1,4GalNAc transferasa un resto UDP-GalNAc a sialil-lactosa (GM3) para formar GalNAc β -4 (Neu5Ac α -3) Gal β -4Glc (GM2) y la β -1,3-galactosiltransferasa transfiere un resto galactosil a GM2 para formar GM1.
15. Procedimiento según la reivindicación 14, en el que el microorganismo comprende además una secuencia heteróloga que codifica una UDP-GlcNAc 4-epimerasa, tal como el gen *wbpP* de *P. aeruginosa* o el gen *gne* de *C. jejuni*.
16. Procedimiento según la reivindicación 11, que se extiende a la producción de Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GD1a) y Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GA1).
17. Procedimiento según la reivindicación 14 o 15, para la producción específica de GM1, en el que la sialil

transferasa heteróloga es una α -2, 3 sialiltransferasa codificada por el gen *cstIII* aislado de cepas de *C. jejuni* que expresan las estructuras de lipooligosacárido que imitan al gangliósido GM1, tal como la cepa de *C. jejuni* con nº de registro 11168 en la NCTC.

5 18. Procedimiento según la reivindicación 11, que se extiende a la producción de la parte de hidratos de carbono del gangliósido GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GM2), en el que el microorganismo comprende además secuencias heterólogas que codifican una β -1,4-GalNAc transferasa, tal como el gen *CgtAll* de la cepa de *C. jejuni* O:36 con nº de registro 43456 de la ATCC y una UDP-GlcNAc 4-epimerasa y en el que el microorganismo presenta por lo menos uno de los tres genes siguientes suprimidos o inactivados para interrumpir la producción endógena de UDP-Gal: el gen *galE* que codifica la UDP-glucosa epimerasa, el gen *GalU* que codifica la UDP-Glc pirofosforilasa y el gen *pgm* que codifica la fosfoglucomutasa.

15 19. Procedimiento según la reivindicación 11, que se extiende a la producción de la parte de hidratos de carbono del gangliósido de GD2 GalNAc β -4Gal β -4GlcNeu5Ac α -8Neu5Ac α 3 en el que el microorganismo comprende además unas secuencias heterólogas que codifican UDP-GlcNAc 4-epimerasa y una β -1,4-GalNAc transferasa que utilizan el azúcar de GD3 como aceptor, tal como la proteína *CgtAll* de la cepa de *C. jejuni* O:36 con el nº de registro 43456 de la ATCC.

20 20. Procedimiento según la reivindicación 19, que se extiende a la producción de azúcar de GD1b (Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -3)Gal β -4Glc) y azúcar de GT1c (Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc) en el que el microorganismo comprende además una secuencia heteróloga que codifica un gen de β -3 Gal transferasa, tal como el gen *cgtB* de la cepa O:2 de *C. jejuni* con el nº de registro 11168 de la NCTC.

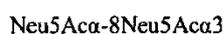
25 21. Procedimiento según la reivindicación 20, que se extiende a la producción de GT1b (Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc), GQ1c (Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc), GQ1b (Neu5Ac α -8Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc) y GP1c (Neu5Ac α -8Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -8Neu5Ac α -8Neu5Ac α 3)Gal β -4Glc) en el que el microorganismo comprende además una secuencia heteróloga que codifica un gen que codifica una sialiltransferasa utilizando GD1b y GT1c como aceptor.

30 22. Procedimiento según la reivindicación 19, en el que el microorganismo presenta por lo menos uno de los tres genes siguientes suprimidos o inactivados para interrumpir la producción endógena de UDP-Gal: el gen *galE* que codifica la UDP-glucosa epimerasa, el gen *galU* que codifica la UDP-Glc pirofosforilasa, y el gen *pgm*.

35 23. Procedimiento según la reivindicación 11, que se extiende a la producción de la parte de hidratos de carbono del gangliósido de GT1b:



|



40 en el que el microorganismo comprende además unas secuencias heterólogas que codifican UDP-GlcNAc 4-epimerasa, una β -1,4-GalNAc transferasa que utilizan el azúcar de GD3 como aceptor, tales como la proteína *CgtAll* de la cepa O:36 de *C. jejuni* con el nº de registro 43456 en la ATCC y una β -1,3-galactosiltransferasa, tal como el gen *cgtB* de la cepa O:2 de *C. jejuni* con el nº de registro 11168 en la NCTC.

45 24. Procedimiento según una de las reivindicaciones 1 a 8 para la producción de Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GD1a) y Neu5Ac α -8Neu5Ac α -3Gal β -3GalNAc β -4(Neu5Ac α -3)Gal β -4Glc (GT1a), en el que el microorganismo comprende unas secuencias heterólogas que codifican una α -2,3 α -2,8-sialiltransferasa bifuncional, tal como el gen *cstIII* de la cepa de *C. jejuni* con nº de registro 43438 en la ATCC, una β -1,4GalNAc transferasa, tal como el gen *cgtAll* de la cepa O:36 de *C. jejuni* con el nº de registro 43456 en la ATCC, que no utiliza el azúcar de GD3 como aceptor, y una β -1,3-galactosiltransferasa, tal como el gen *cgtB* de la cepa O:2 de *C. jejuni* con el nº de registro 11168 en la NCTC.

50 25. Procedimiento según una de las reivindicaciones 1 a 8 para producir el pentasacárido LST_D (Neu5Ac α -3Gal β -4GlcNAc β -3Gal β -4Glc), en el que la sialil transferasa es una α -2,3-sialil transferasa, tal como el gen *nst*, y en el que el microorganismo comprende además los genes heterólogos *lgtA* y *lgtB* que codifican β -1,3-GlcNAc transferasa y β -1,4-galactosiltransferasa respectivamente.

55 26. Procedimiento según la reivindicación 23 para la producción de oligosacáridos de sialil-lewis X, en el que el microorganismo comprende además una secuencia heteróloga que codifica una α -1,3-fucosil-transferasa tal como *futA* de *H. pylori* (SEC. ID. nº 18) o *futB* (SEC. ID. nº 19).

60 27. Procedimiento según una de las reivindicaciones 1 a 8 para la producción de sialosilgalactosil globósido (Neu5Ac α -3Gal β -3GalNAc β -3Gal α -4Gal β -4Gal), en el que el microorganismo se cultiva en un medio con globotriosa

y es *LacY+*, *MelA-*, *nanT+*, *NanA-* *NanK-* y comprende *IgtD* heterólogo, tal como el gen *IgtD* de HI1578 de *Haemophilus influenzae*, número de registro U32832 en Genbank, genes para la α -2,3-sialiltransferasa y UGP-GlcNAc C4-epimerasa, así como los genes *neuABC*.

- 5 28. Procedimiento según una de las reivindicaciones 1 a 8 para la producción de sialilgalactosa (Neu5Ac α -3 Gal) y oligosacáridos sialilados con una galactosa terminal reductora, en el que el microorganismo es *galk-*, *NanA-* y *NanK-* (o *NanKEAT-*) y expresa el gen de la sialiltransferasa y los genes *neuBCA* y se cultiva en un medio con galactosa.
- 10 29. Procedimiento según la reivindicación 1 para la producción de oligosacáridos sialilados con estructura de quitooligosacárido en su extremo reductor, tal como el heptasacárido sialilado Neu5Ac α -3Gal β -4[GlcNAc β -4]4GlcNAc, en el que la sialiltransferasa es una α -2,3-sialiltransferasa tal como el gen *nst* de *Neisseria* y el microorganismo comprende además una quitin-oligosacárido sintasa tal como el gen *nodC* de *Azorhizobium caulinodans* y un gen de β -1,4-galactosiltransferasa tal como el gen *IgtB* de *Neisseria meningitidis*.
- 15 30. Procedimiento según la reivindicación 29 para la producción del tetrasacárido sialilado Neu5Ac α -3Gal β -4GlcNAc β -4GlcNAc, en el que el microorganismo comprende además una secuencia heteróloga que codifica una quitinasa, tal como el gen *chiA*.
- 20 31. Procedimiento según la reivindicación 29 o 30, en el que no se añade ningún precursor exógeno al medio de cultivo.
32. Procedimiento según una de las reivindicaciones 11 a 31, en el que los microorganismos se están cultivando en glicerol o glucosa como sustrato de carbono.
- 25 33. Microorganismo tal como se define en cualquiera de las reivindicaciones 12, 14, 15, 17 y 18 a 30.

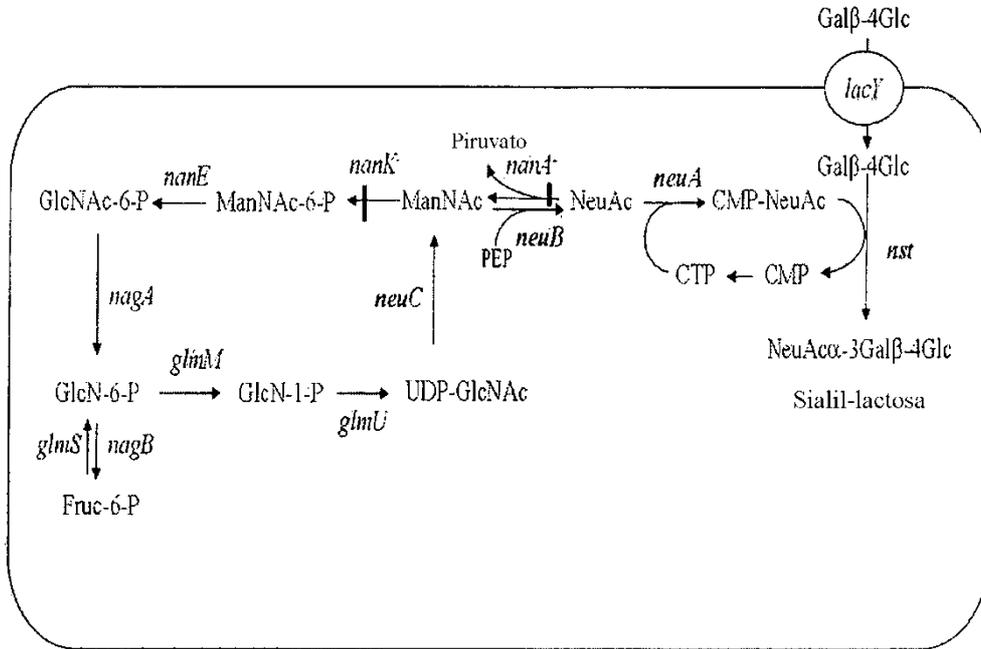


FIGURA 1

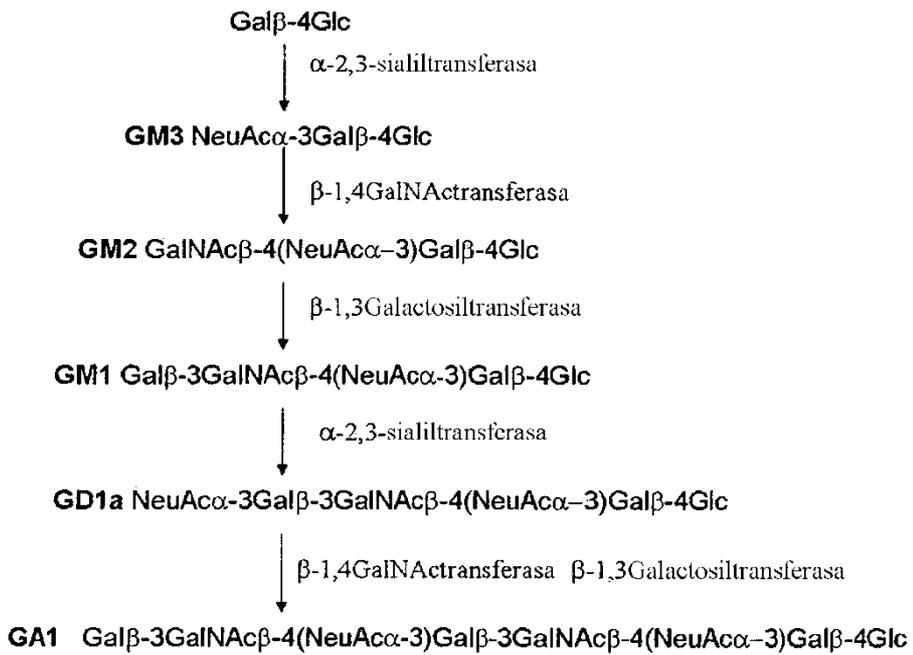


FIGURA 2

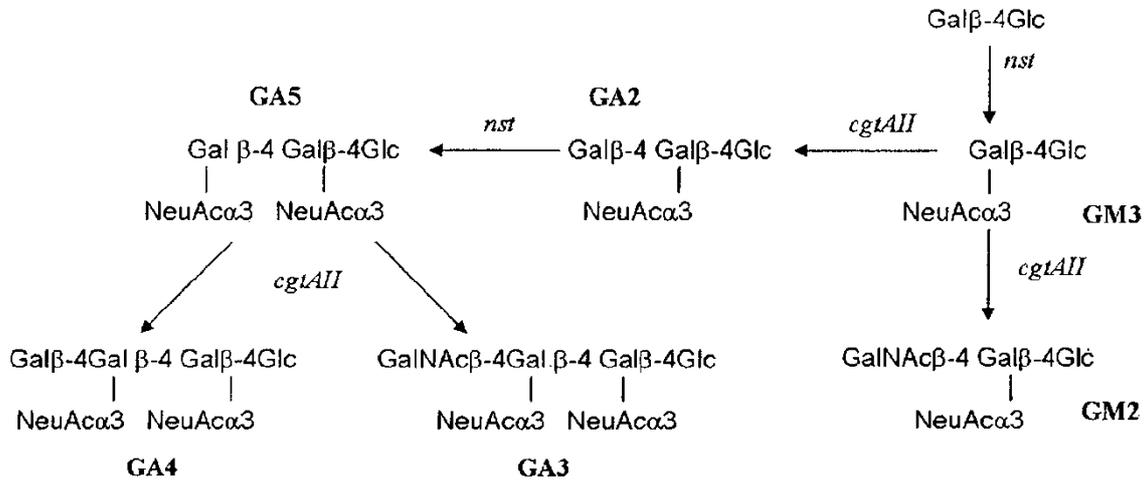


FIGURA 3

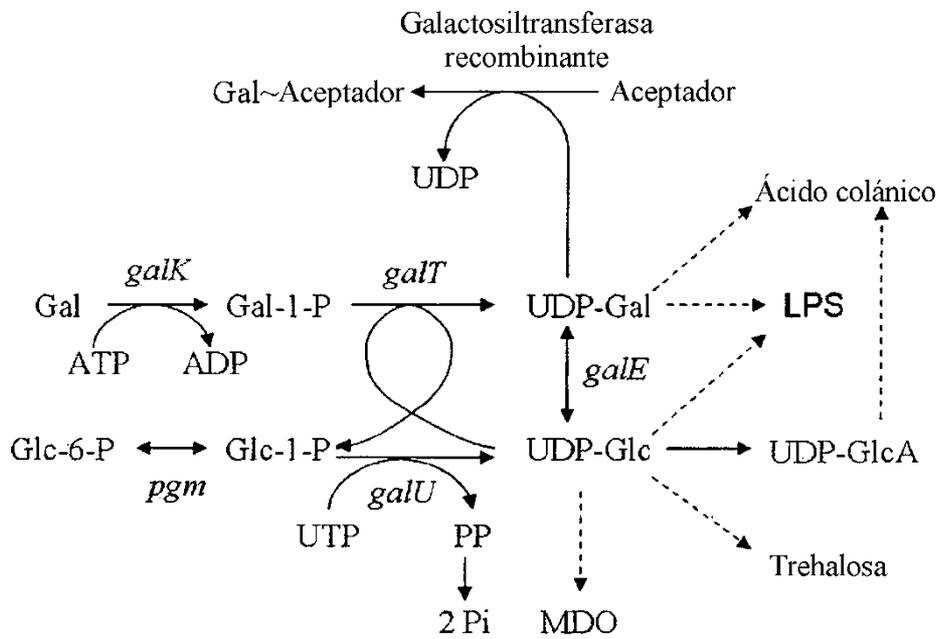


FIGURA 4

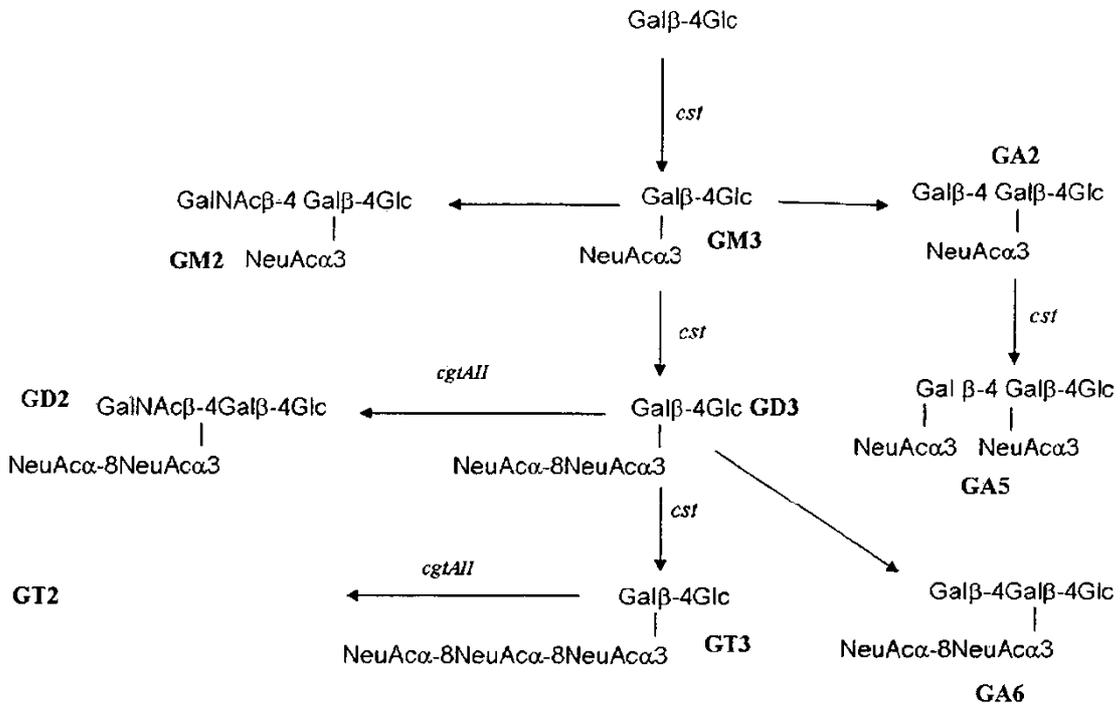


FIGURA 5

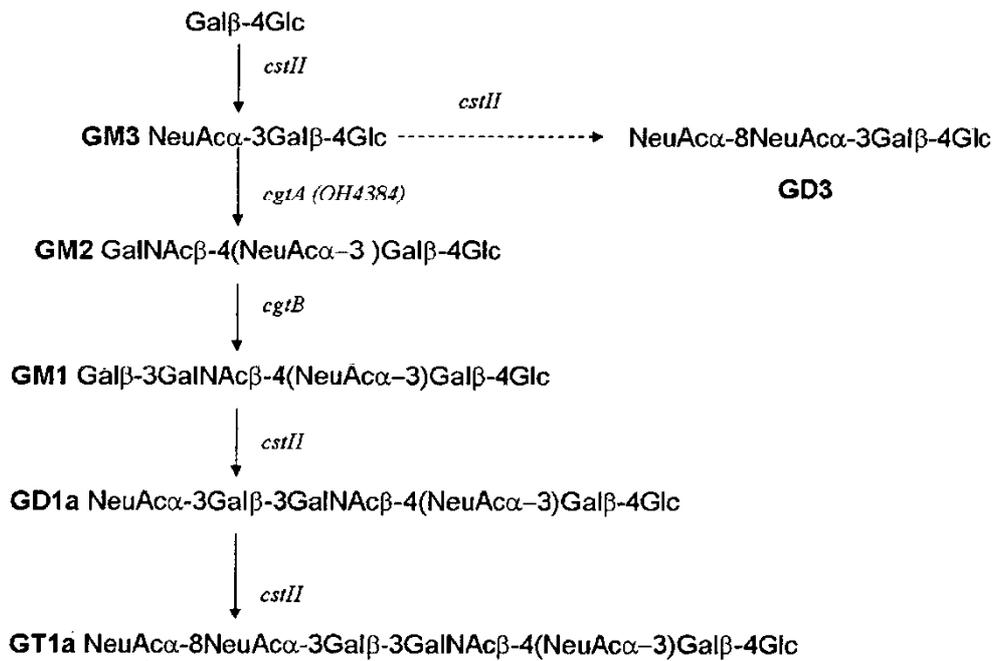


FIGURA 6

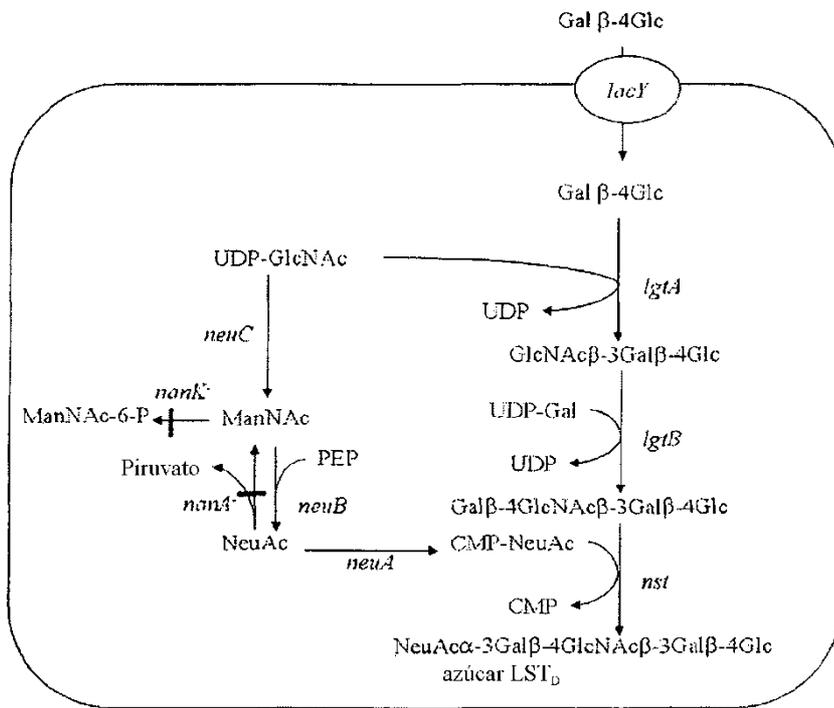


FIGURA 7

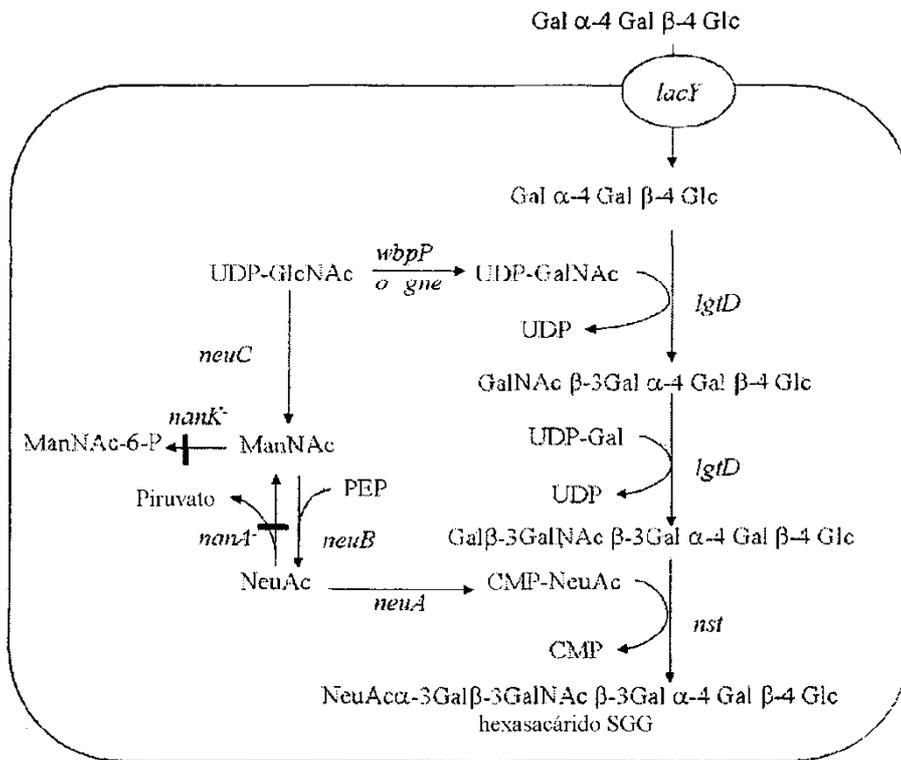


FIGURA 8

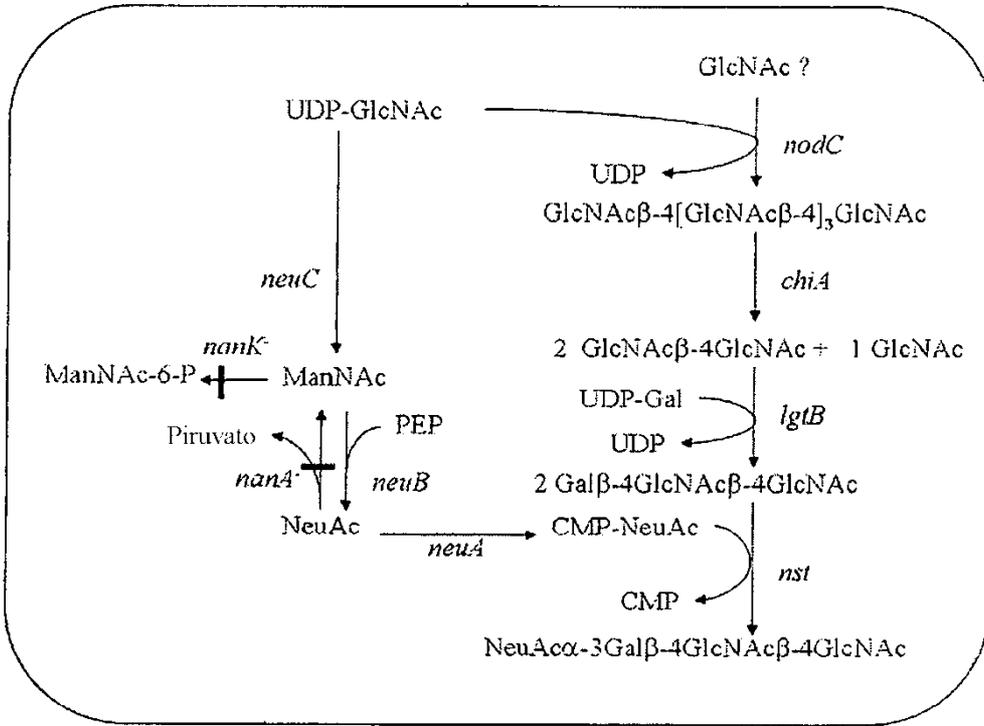


FIGURA 9

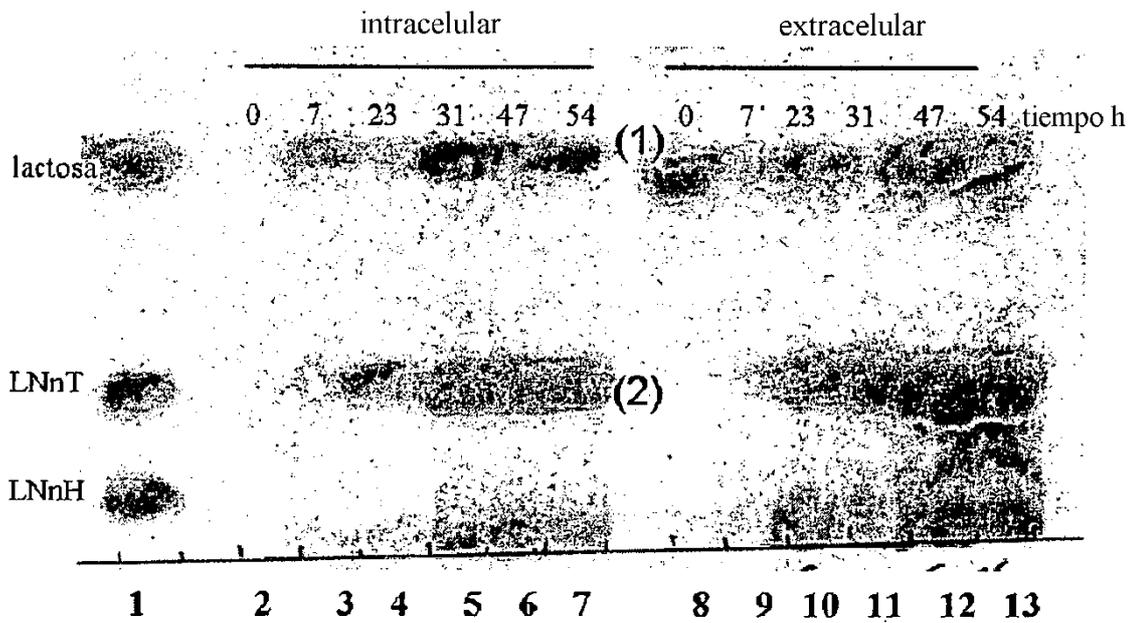


FIGURA 10

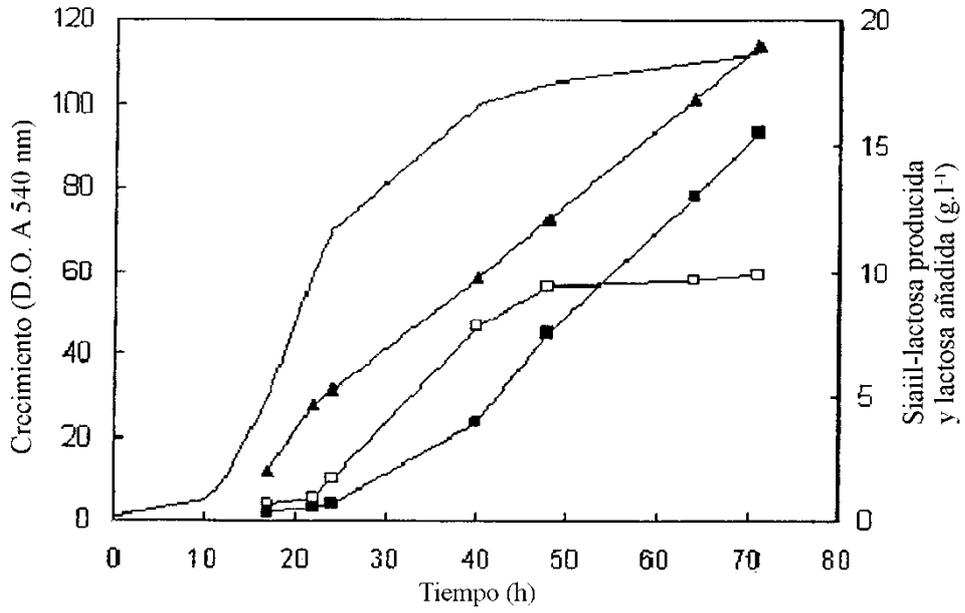


FIGURA 11

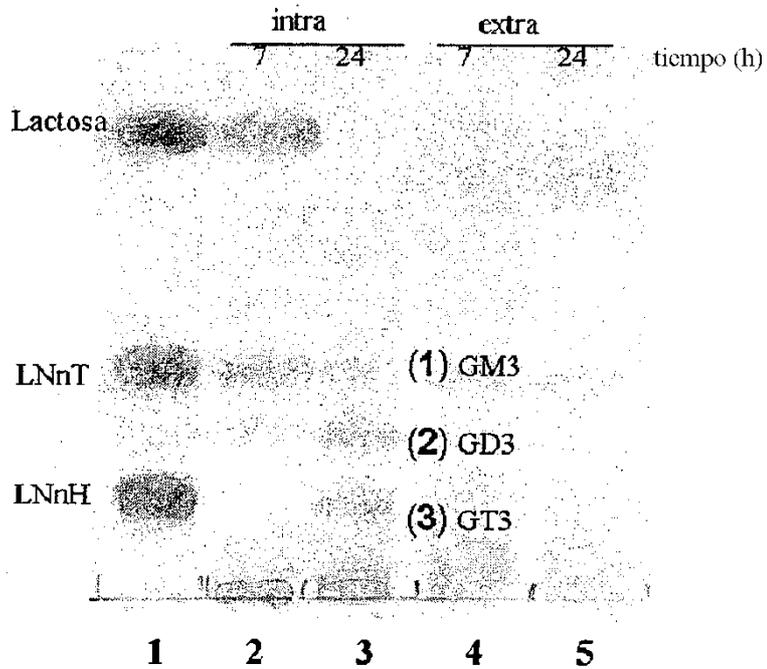


FIGURA 12

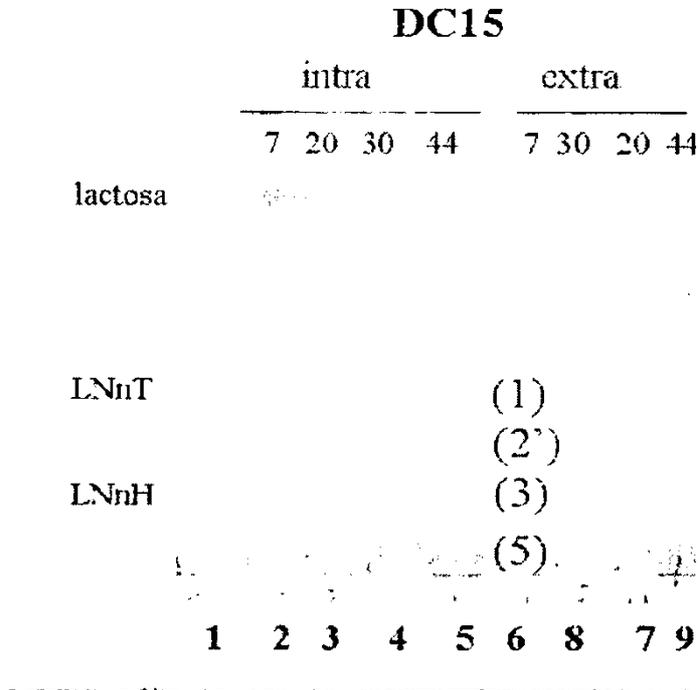


FIGURA 13

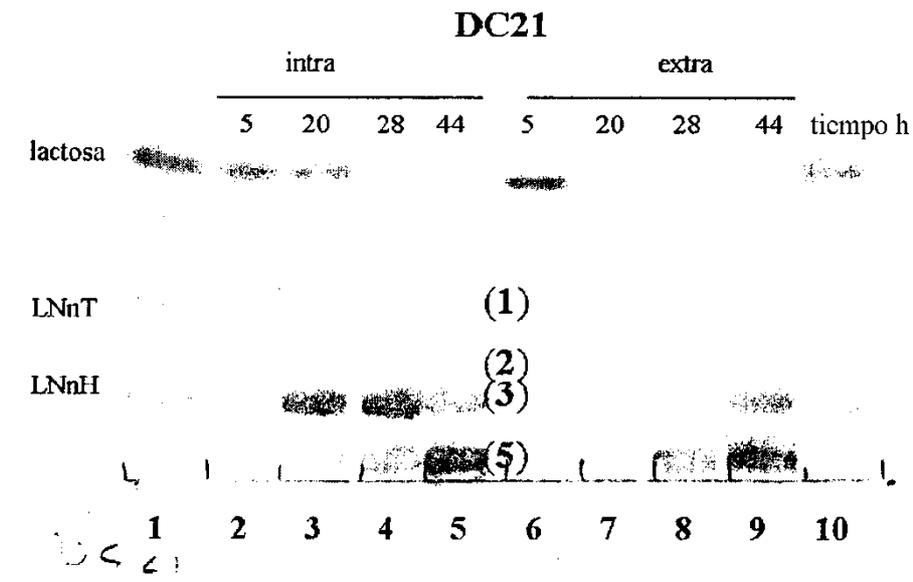


FIGURA 14

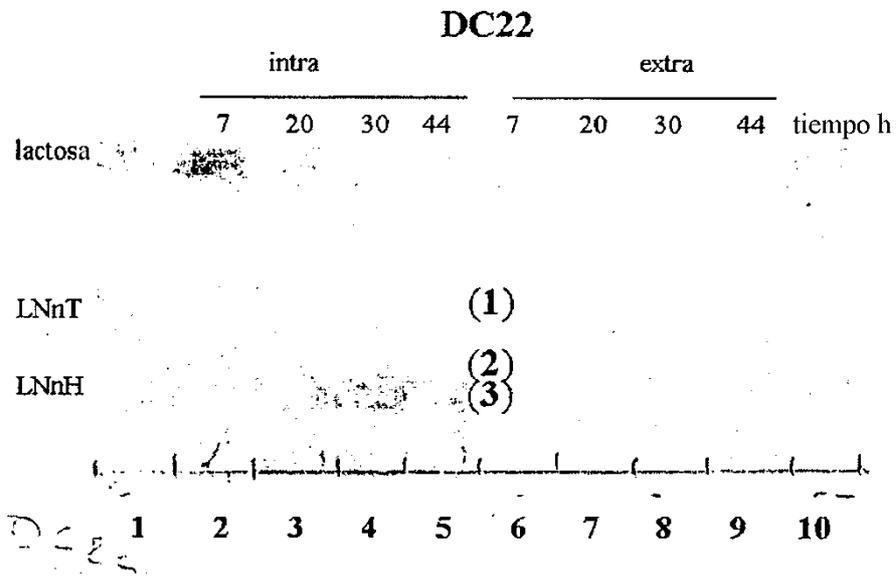


FIGURA 15

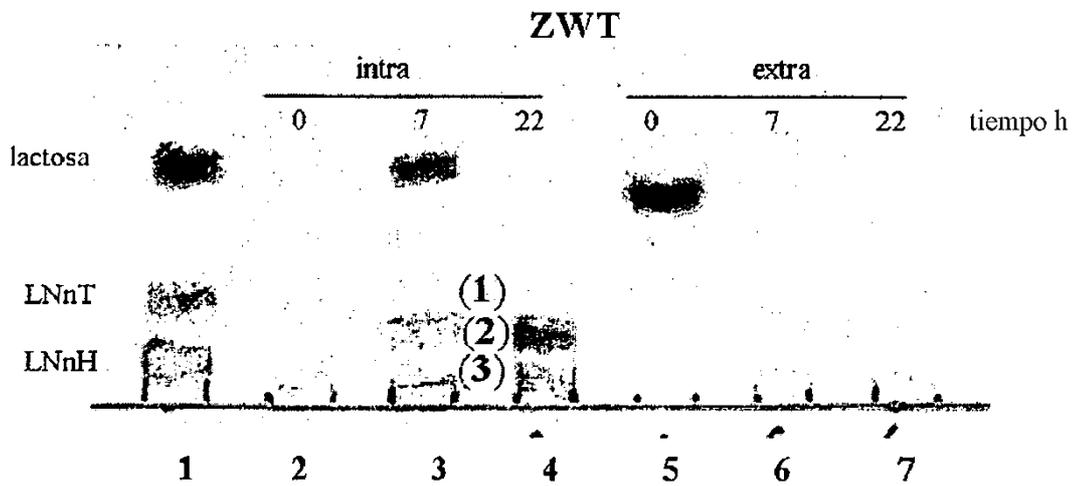


FIGURA 16

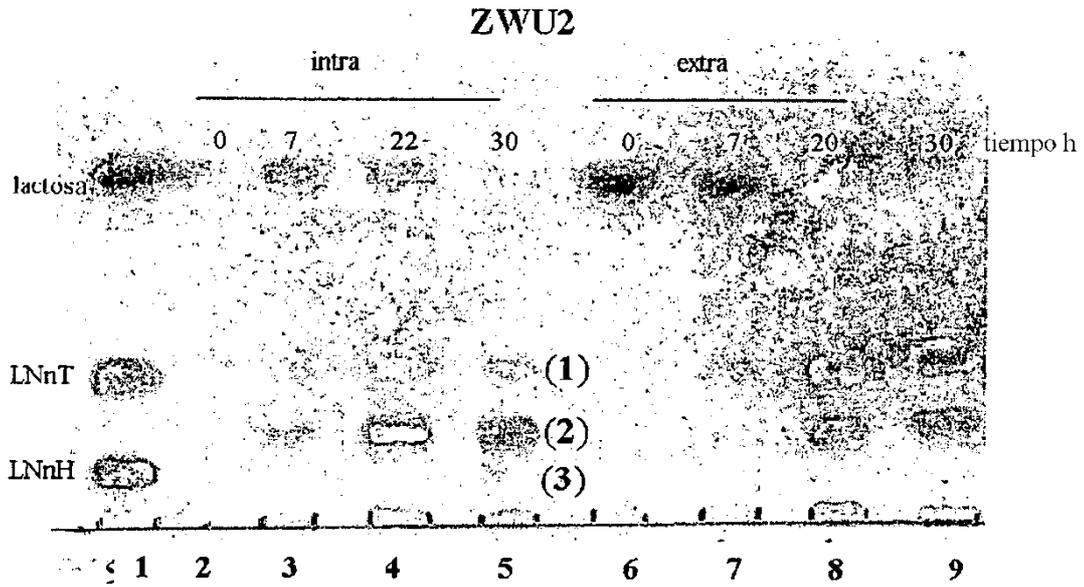
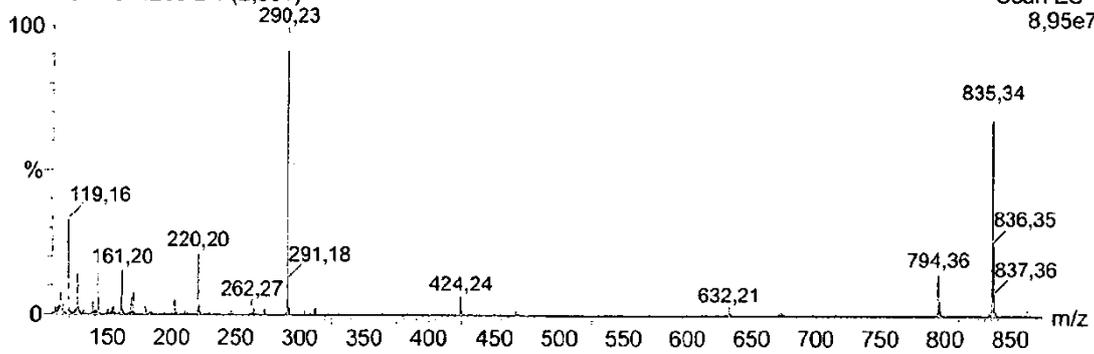


FIGURA 17

A

ref: ZWTA 07-Dec-2005
 FIERFORT071205-2 1 (2,001)

Scan ES-
 8,95e7



B

U2 01-Dec-2005
 FIERFORT011205-1 1 (2,001)

Scan ES-
 835,28 2,89e8

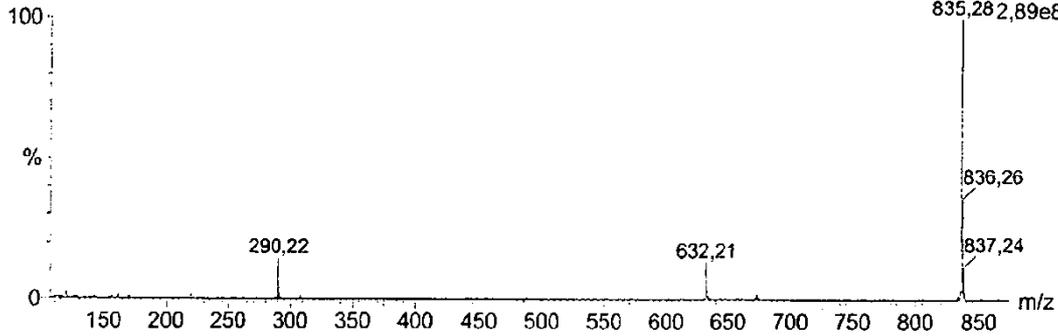


FIGURA 18

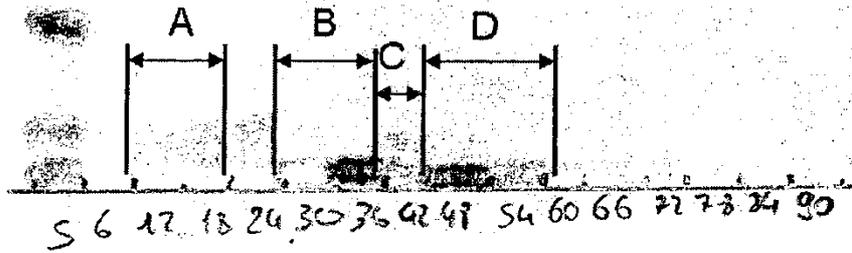


FIGURA 19

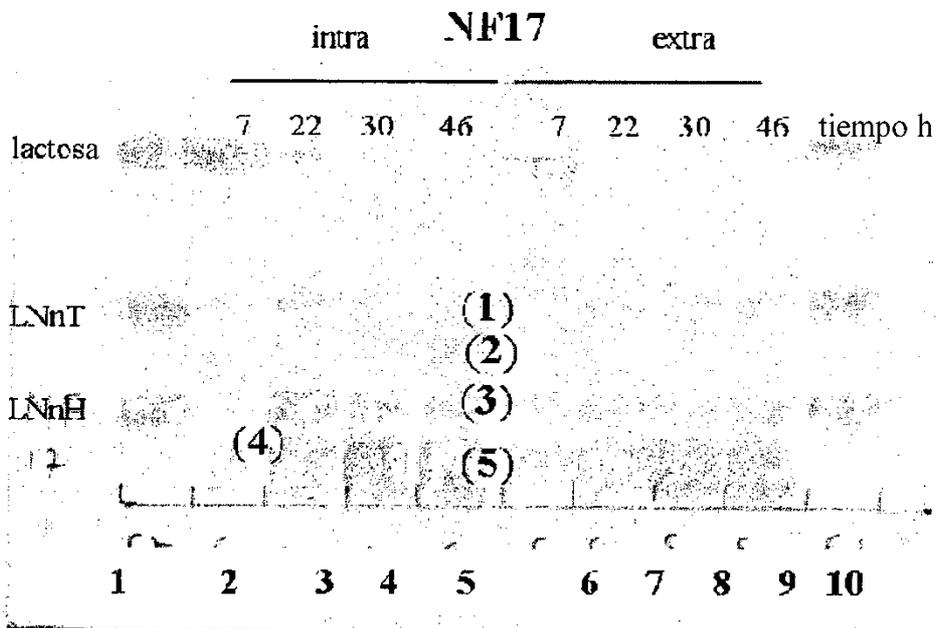


FIGURA 20