

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 368 740**

51 Int. Cl.:
C12N 15/00 (2006.01)
C12Q 1/48 (2006.01)
C12P 19/44 (2006.01)
C12P 1/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05713041 .1**
96 Fecha de presentación: **04.02.2005**
97 Número de publicación de la solicitud: **1718739**
97 Fecha de publicación de la solicitud: **08.11.2006**

54 Título: **MÉTODOS DE REPLEGAMIENTO DE GLICOSILTRANSFERASAS DE MAMÍFERO.**

30 Prioridad:
04.02.2004 US 542210 P
06.08.2004 US 599406 P
12.11.2004 US 627406 P

45 Fecha de publicación de la mención BOPI:
21.11.2011

45 Fecha de la publicación del folleto de la patente:
21.11.2011

73 Titular/es:
NOVO NORDISK A/S
NOVO ALLÉ
2880 BAGSVÄRD, DK

72 Inventor/es:
SARIBAS, Sami;
HAKES, David;
WILLETT, Scott;
JOHNSON, Karl, F.;
BEZILA, Daniel James y
DEFREES, Shawn

74 Agente: **Arias Sanz, Juan**

ES 2 368 740 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

Métodos de replegamiento de glicosiltransferasas de mamífero

Campo de la invención

- 5 La presente invención proporciona métodos de replegamiento de glicosiltransferasas de mamífero que se han producido en células bacterianas, incluyendo mutantes de glicosiltransferasa que tienen una capacidad mejorada para replegarse, y métodos para usar tales glicosiltransferasas replegadas. La invención también proporciona métodos de replegamiento de más de una glicosiltransferasa en un único recipiente, métodos para usar tales glicosiltransferasas replegadas, y mezclas de reacción que comprenden las glicosiltransferasas replegadas.

Antecedentes de la invención

- 10 Los organismos eucariotas sintetizan estructuras de oligosacárido o glicoconjugados, tales como glicolípidos o glicoproteínas, que son comercial y terapéuticamente útiles. La síntesis *in vitro* de oligosacáridos o glicoconjugados puede llevarse a cabo usando glicosiltransferasas eucariotas recombinantes. El método más eficaz para producir glicosiltransferasas eucariotas recombinantes para la síntesis de oligosacáridos es expresar la proteína en bacterias. Sin embargo, en bacterias, muchas glicosiltransferasas eucariotas se expresan como proteínas insolubles en
- 15 cuerpos de inclusión bacterianos, y los rendimientos de proteína activa a partir de los cuerpos de inclusión pueden ser muy bajos. Por tanto, existe la necesidad de métodos mejorados para producir glicosiltransferasas eucariotas en bacterias. La presente invención resuelve esta y otras necesidades.

Breve resumen de la invención

- 20 La presente invención proporciona un método de replegamiento de una glicosiltransferasa eucariota, recombinante, insoluble que comprende un dominio de proteína de unión a maltosa (MBD). La glicosiltransferasa eucariota, recombinante, insoluble se solubiliza en un tampón de solubilización y entonces se pone en contacto con un tampón de replegamiento que comprende un par redox, de manera que la glicosiltransferasa eucariota replegada cataliza la transferencia de un azúcar de un sustrato donador a un sustrato aceptor. En una realización, la glicosiltransferasa eucariota se trunca para eliminar la totalidad o una parte de una región de tallo de la proteína. En otra realización, se
- 25 elimina una cisteína desapareada en la glicosiltransferasa eucariota mediante sustitución por un aminoácido distinto de cisteína. En una realización adicional, se elimina una cisteína desapareada en la glicosiltransferasa eucariota mediante sustitución por un aminoácido distinto de cisteína y la glicosiltransferasa eucariota también se trunca para eliminar la totalidad o una parte de una región de tallo de la proteína.

- 30 En una realización, la glicosiltransferasa eucariota se selecciona del grupo que consiste en GnT1, GalT1, StIII Gal3, ST3Gall, St6 GalNAcTI, Core GalTI, GalNAcT2.

En una realización, la glicosiltransferasa eucariota comprende además un dominio de purificación, por ejemplo, un dominio de unión a almidón (SBD), un dominio de tiorredoxina, un dominio SUMO, un dominio de poli-His, un dominio de epítipo myc y un dominio de glutatión-S-transferasa.

En una realización, la glicosiltransferasa eucariota comprende además un dominio de autoescisión.

- 35 En una realización, la glicosiltransferasa eucariota se expresa en una célula huésped bacteriana como un cuerpo de inclusión insoluble.

- En una realización, una segunda glicosiltransferasa eucariota recombinante, insoluble se repliega con la primera glicosiltransferasa eucariota. En una realización adicional, una tercera glicosiltransferasa eucariota recombinante, insoluble se repliega con la primera glicosiltransferasa eucariota y la segunda glicosiltransferasa eucariota. Puede
- 40 añadirse una glicosiltransferasa eucariota recombinante, insoluble adicional y replegarse juntas dependiendo de las necesidades del usuario, por ejemplo, replegamiento de 4, 5, 6, 7, 8, 9 ó 10 glicosiltransferasas juntas.

En una realización, el par redox se selecciona del grupo que consiste en glutatión reducido/glutatión oxidado (GSH/GSSG) y cisteína/cistamina.

- 45 En una realización, el sustrato aceptor se selecciona de una proteína, un péptido, una glicoproteína y un glicopéptido.

En una realización, la glicosiltransferasa eucariota es una sialiltransferasa. Las sialiltransferasas pueden incluir, por ejemplo, StIII Gal3, ST3Gall, St6 GalNAcTI. En un aspecto adicional, el sustrato donador es una molécula de CMP-ácido siálico-PEG y el sustrato aceptor se selecciona de una proteína, un péptido, una glicoproteína y un glicopéptido.

- 50 La presente invención también proporciona una glicosiltransferasa eucariota, recombinante, en la que se han delecionado una región de anclaje de tallo y un dominio transmembrana de la proteína, y en el que la glicosiltransferasa se fusiona en el marco con un dominio de proteína de unión a maltosa (MBP). En una realización,

la fusión de la glicosiltransferasa eucariota, recombinante y el dominio de MBP se expresa como un cuerpo de inclusión insoluble en bacterias, por ejemplo, *E. coli*.

5 En una realización, se deleta la totalidad o una parte de la región de tallo de la glicosiltransferasa eucariota, recombinante. En otra realización, se elimina una cisteína desapareada en la glicosiltransferasa eucariota, recombinante mediante sustitución por un aminoácido distinto de cisteína.

En una realización adicional, la glicosiltransferasa eucariota, recombinante es una de las siguientes: una proteína GnT1, una proteína GalT1, una proteína StIIIGal3, una proteína ST3Gall, una proteína St6 GalNAcT1, una proteína Core GalTI o una proteína GalNAcT2.

10 En una realización, la glicosiltransferasa eucariota, recombinante es una proteína GnT1. En un aspecto, la proteína GnT1 es una proteína GnT1 humana truncada seleccionada de GnT1 Δ 35 y GnT1 Δ 103. En otro aspecto, la proteína GnT1 es una proteína GnT1 humana que comprende una sustitución de cisteína desapareada seleccionada del grupo que consiste en CYS121ALA, CYS121ASP y ARG120ALA, CYS121HIS. En un aspecto adicional, la proteína GnT1 tanto está truncada como se le ha eliminado un residuo de cisteína desapareado mediante una mutación por sustitución.

15 En una realización, la glicosiltransferasa eucariota, recombinante es una proteína GalT1. En un aspecto, la proteína GalT1 es una proteína GalT1 bovina truncada seleccionada de GalT1 Δ 70 y GalT1 Δ 129. En otro aspecto, la proteína GalT1 es una proteína GalT1 bovina que comprende una sustitución de cisteína desapareada de CYS342THR. En un aspecto adicional, la proteína GalT1 tanto está truncada como se le ha eliminado un residuo de cisteína desapareado mediante una mutación por sustitución.

20 En una realización, la glicosiltransferasa eucariota, recombinante es una proteína ST3GallIII. En un aspecto, la proteína ST3GallIII es una proteína ST3GallIII de rata truncada seleccionada de ST3GallIII Δ 28, ST3GallIII Δ 73, ST3GallIII Δ 85 y ST3GallIII Δ 86. En otro aspecto, la proteína ST3GallIII comprende una sustitución de aminoácido por un residuo de cisteína desapareado. En un aspecto adicional, la proteína ST3GallIII tanto está truncada como se le ha eliminado un residuo de cisteína desapareado mediante una mutación por sustitución.

25 En una realización, la glicosiltransferasa eucariota, recombinante de la reivindicación 15, en la que la glicosiltransferasa es una proteína Core1 GalT1. En un aspecto, la proteína Core1 GalT1 es una proteína de *Drosophila* truncada o una proteína humana truncada. En otro aspecto, una proteína Core1 GalT1 humana o de *Drosophila* comprende una sustitución de aminoácido por un residuo de cisteína desapareado. En un aspecto adicional, una proteína Core1 GalT1 humana o de *Drosophila* tanto está truncada como se le ha eliminado un residuo de cisteína desapareado mediante una mutación por sustitución.

30 En una realización, la glicosiltransferasa eucariota, recombinante es una proteína ST3Gal1. En un aspecto, la proteína ST3Gal1 es una proteína humana truncada seleccionada de ST3Gal1 Δ 29, ST3Gal1 Δ 45 y ST3Gal1 Δ 56. En otro aspecto, la proteína ST3Gal1 comprende una sustitución de aminoácido por un residuo de cisteína desapareado. En un aspecto adicional, la proteína ST3Gal1 tanto está truncada como se le ha eliminado un residuo de cisteína desapareado mediante una mutación por sustitución.

35 En una realización, la glicosiltransferasa eucariota, recombinante es una proteína ST6GalNAc1. En un aspecto, la proteína ST6GalNAc1 es una proteína de ratón truncada, una proteína de gallina truncada o una proteína humana truncada, por ejemplo, uno de los truncamientos enumerados en la tabla 14. En otro aspecto, una proteína ST6GalNAc1 de ratón, gallina o humana comprende una sustitución de aminoácido por un residuo de cisteína desapareado. En un aspecto adicional, una proteína ST6GalNAc1 de ratón, gallina o humana tanto está truncada como se le ha eliminado un residuo de cisteína desapareado mediante una mutación por sustitución.

40 En una realización, la glicosiltransferasa eucariota, recombinante de la reivindicación 15, en la que la glicosiltransferasa es una proteína GalNAcT2. En un aspecto, la proteína GalNAcT2 es una proteína humana truncada seleccionada de GalNAcT2 Δ 40, GalNAcT2 Δ 51, GalNAcT2 Δ 74 y GalNAcT2 Δ 95. En otro aspecto, la proteína GalNAcT2 humana comprende una sustitución de aminoácido por un residuo de cisteína desapareado. En un aspecto adicional, la proteína GalNAcT2 humana tanto está truncada como se le ha eliminado un residuo de cisteína desapareado mediante una mutación por sustitución.

45 La presente invención también proporciona un método de remodelación de una proteína, un péptido, una glicoproteína o un glicopéptido usando las glicosiltransferasas eucariotas, recombinantes enumeradas anteriormente, tras haberse replegado las proteínas y tienen actividad enzimática.

50 La presente invención proporciona métodos mejorados para replegar glicosiltransferasas eucariotas insolubles en una forma activa y también proporciona glicosiltransferasas, por ejemplo, enzimas N-acetilglucosaminiltransferasa I (GnT1) que tienen propiedades de replegamiento mejoradas.

55 En un aspecto, la invención proporciona una enzima N-acetilglucosaminiltransferasa I (GnT1) eucariota recombinante, que se ha mutado para reemplazar un residuo de cisteína desapareado por un aminoácido que

mejora el replegamiento de la enzima a partir de un precipitado insoluble, por ejemplo, cuerpos de inclusión bacterianos. La enzima GnT1 incluye al menos el dominio catalítico de la enzima GnT1. La enzima GnT1 es biológicamente activa, es decir, que puede catalizar la transferencia de un sustrato donador a un sustrato aceptor.

5 En una realización, la enzima GnT1 es una proteína humana. Algunas mutaciones del residuo CYS121 en GnT1 humana mejoran el replegamiento. Esos mutantes incluyen por ejemplo, mutación CYS121SER, una mutación CYS121ALA, mutación CYS121ASP, y un doble mutante, ARG120ALA, CYS121HIS. Se muestran secuencias representativas de los mutantes de GnT1 en las figuras 7-11. En otras eucariotas, por ejemplo, mutaciones similares de un residuo de cisteína desapareado, CYS123, mejoran el replegamiento de la enzima GnT1.

10 En otra realización, la enzima GnT1 también incluye una cola de aminoácidos, por ejemplo, una proteína de unión a maltosa (MBP), una cola de polihistidina, una glutatión S transferasa (GST), una proteína de unión a almidón (SBP) y un epítipo myc.

15 En otro aspecto, la invención proporciona ácidos nucleicos que codifican para una enzima GnT1 eucariota recombinante, que se ha mutado para reemplazar un residuo de cisteína desapareado por un aminoácido que mejora el replegamiento de la enzima a partir de un precipitado insoluble, por ejemplo, cuerpos de inclusión bacterianos. Como anteriormente, la enzima GnT1 codificada incluye al menos el dominio catalítico de la enzima GnT1, y es biológicamente activa, es decir, que puede catalizar la transferencia de un sustrato donador a un sustrato aceptor.

20 En una realización, los ácidos nucleicos codifican para una enzima GnT1 humana. Algunas mutaciones del residuo CYS121 en GnT1 humana mejoran el replegamiento. Esos mutantes incluyen por ejemplo, mutación CYS121SER, una mutación CYS121ALA, mutación CYS121ASP, y un doble mutante, ARG120ALA, CYS121HIS. Se muestran secuencias representativas de ácidos nucleicos de proteínas y ácidos nucleicos mutantes de GnT1 en las figuras 7-11. En otras eucariotas, por ejemplo, mutaciones similares de un residuo de cisteína desapareado, CYS123, mejoran el replegamiento la enzima GnT1.

25 En una realización adicional, la enzima GnT1 codificada también incluye una cola de aminoácidos, por ejemplo, una proteína de unión a maltosa (MBP), una cola de polihistidina, una glutatión S transferasa (GST), una proteína de unión a almidón (SBP) y un epítipo myc.

La invención también incluye vectores de expresión que incluyen los ácidos nucleicos de GnT1 mutados, células huésped que incluyen los vectores de expresión de GnT1, y métodos de producción de las enzimas GnT1 mutadas usando el sistema de vector de expresión/huésped.

30 En otra realización, la invención proporciona un método de adición de residuos de N-acetilglucosamina a una molécula aceptora con un residuo de manosa terminal, poniendo en contacto la molécula aceptora con una molécula de N-acetilglucosamina activada y una enzima GnT1 eucariota que se ha mutado para mejorar el replegamiento. La molécula aceptora puede ser, por ejemplo, un polisacárido, un oligosacárido, un glicolípido o una glicoproteína.

35 En otro aspecto, la invención proporciona un método de replegamiento de al menos dos proteínas glicosiltransferasas eucariotas recombinantes, insolubles en un único recipiente, poniendo en contacto las glicosiltransferasas con un tampón de replegamiento que incluye un par redox.

Tras el replegamiento, al menos dos de las glicosiltransferasas replegadas tienen actividad biológica, por ejemplo, pueden catalizar la transferencia de un sustrato donador a un sustrato aceptor.

40 El tampón de replegamiento también puede incluir un detergente, o un agente caotrópico, o arginina, o PEG. En algunas realizaciones, el pH del tampón de replegamiento es de entre 6,0 y 10,0. En una realización, el pH del tampón de replegamiento es de entre 6,5 y 8,0. En otra realización, el pH del tampón de replegamiento es de entre 8,0 y 9,0.

45 En otra realización, las glicosiltransferasas incluyen una cola de aminoácidos, por ejemplo, una proteína de unión a maltosa (MBP), una cola de polihistidina, una glutatión S transferasa (GST), una proteína de unión a almidón (SBP) y un epítipo myc

En una realización, se repliegan juntas más de una glicosiltransferasa de una ruta biosintética de glicano unido a N.

En una realización, se repliega una sialiltransferasa con otra glicosiltransferasa usando los métodos de la invención.

50 En una realización, se repliega una N-acetilglucosaminiltransferasa con otra glicosiltransferasa usando los métodos de la invención.

En una realización, se repliega una galactosiltransferasa con otra glicosiltransferasa usando los métodos de la invención.

En otra realización, se repliegan juntas una sialiltransferasa, una N-acetilglucosaminiltransferasa y una galactosiltransferasa en un único recipiente usando los métodos de la invención.

En una realización, se repliegan juntas más de una glicosiltransferasa de una ruta biosintética de glicano unidos a O. En una realización adicional, una primera enzima es una N-acetilgalactosaminiltransferasa. En una realización preferida, una primera enzima es una N-acetilglucosaminiltransferasa 2 (GalNAcT2).

La presente invención también proporciona una mezcla de reacción incluyendo una enzima GnT1 eucariota recombinante, que se ha mutado para reemplazar un residuo de cisteína desapareado por un aminoácido que mejora el repliegamiento de la enzima a partir de un precipitado insoluble, por ejemplo, cuerpos de inclusión bacterianos y al menos otra glicosiltransferasa que se han replegado en el mismo recipiente. La segunda glicosiltransferasa puede ser, por ejemplo, una sialiltransferasa o una galactosiltransferasa. En una realización, la mezcla de reacción incluye la enzima GnT1 eucariota mutada, una sialiltransferasa y una galactosiltransferasa. Las mezclas de reacción pueden usarse con una molécula aceptora con un azúcar donador, para producir por ejemplo, un polisacárido, un oligosacárido, un glicolípido o una glicoproteína.

En otro aspecto, la invención proporciona un método de repliegamiento de una sialiltransferasa eucariota recombinante insoluble, mediante (a) solubilizar la sialiltransferasa; y luego (b) poner en contacto la sialiltransferasa soluble con un tampón de repliegamiento que incluye un par redox. La sialiltransferasa replegada es biológicamente activa y cataliza la transferencia de ácido siálico de un sustrato donador a un sustrato aceptor. En una realización, se dializa o diafiltra la sialiltransferasa replegada.

El tampón de repliegamiento también puede incluir un detergente, o un agente caotrópico, o arginina. En algunas realizaciones el pH del tampón de repliegamiento es de entre 6,0 y 10,0. En una realización, el pH del tampón de repliegamiento es de entre 6,5 y 8,0. En otra realización, el pH del tampón de repliegamiento es de entre 8,0 y 9,0. En otra realización, el pH del tampón de repliegamiento es de entre 7,5 y 8,5.

En una realización, el par redox en el tampón de repliegamiento es glutatión reducido/glutatión oxidado (GSH/GSSG). En una realización adicional, la razón molar de GSH/GSSG es de entre 100:1 y 1:10. En una realización preferida, la razón molar de GSH/GSSG es 10:1. En una realización todavía adicional, el tampón de repliegamiento comprende GSH aproximadamente 0,02-10 mM, GSSG 0,005-10 mM, lauril-maltósido 0,005-10 mM, NaCl 50-250 mM, KCl 2-10 mM, PEG 3350 al 0,01-0,05% y L-arginina 150-550 mM.

En otra realización, la sialiltransferasa incluye una cola de aminoácidos por ejemplo, proteína de unión a maltosa (MBP), una cola de polihistidina, una glutatión S transferasa (GST), una proteína de unión a almidón (SBP) y un epítipo myc. En una realización adicional, se purifica la sialiltransferasa usando una molécula de unión a cola que se une a la cola de aminoácidos. Por ejemplo, la cola de aminoácidos puede ser MBP y la molécula de unión a cola puede ser amilosa, maltosa o una ciclodextrina.

En otra realización, la sialiltransferasa replegada cataliza la transferencia de ácido siálico de CMP-ácido siálico a una glicoproteína.

En una realización adicional, la sialiltransferasa replegada cataliza la transferencia de PEG de 10K o PEG de 20K de CMP-SA-PEG (10 kDa) o CMP-SA-PEG (20 kDa) a una glicoproteína.

En otra realización, la sialiltransferasa es ST3GalIII de hígado de rata.

En otro aspecto, la invención proporciona un método de adición de un resto sialilo a una glicoproteína, poniendo en contacto la glicoproteína con CMP-ácido siálico con una sialiltransferasa de mamífero replegada que se replegó usando los métodos dados a conocer en el presente documento.

En otro aspecto, la invención proporciona un método de adición de un resto de PEG a una glicoproteína, comprendiendo el método poner en contacto la glicoproteína con CMP-SA-PEG (10 kDa) o CMP-SA-PEG (20 kDa) y una sialiltransferasa de mamífero replegada que se replegó usando los métodos dados a conocer en el presente documento.

En un aspecto adicional la invención proporciona un método de repliegamiento de una N-acetilgalactosaminiltransferasa 2 (GalNAcT2) eucariota recombinante soluble solubilizando la GalNAcT2 en un tampón de solubilización; y luego poner en contacto la GalNAcT2 soluble con un tampón de repliegamiento que incluye un par redox para replegar la GalNAcT2. Tras el repliegamiento, la GalNAcT2 replegada cataliza la transferencia de N-acetilgalactosamina de un sustrato donador a un sustrato aceptor. El método puede incluir opcionalmente las etapas de dializar o diafiltrar la GalNAcT2 replegada o la purificación adicional de la GalNAcT2 replegada.

En algunas realizaciones el par redox del tampón de repliegamiento es glutatión reducido/glutatión oxidado (GSH/GSSG) o cisteína/cistamina. El tampón de repliegamiento también puede incluir lo siguiente: un detergente, un agente caotrópico o arginina. En algunas realizaciones, el pH del tampón de repliegamiento es de entre 6,0 y 10,0. En una realización preferida, el pH del tampón de repliegamiento es de aproximadamente 8,0.

En realizaciones preferidas, el pH del tampón de solubilización es de entre 6,0 y 10,0. En una realización más preferida, el pH del tampón de solubilización es de aproximadamente 8,0.

5 La GalNAcT2 expresada de manera recombinante puede incluir una cola de aminoácidos. La cola de aminoácidos puede ser, por ejemplo, una proteína de unión a maltosa (MBP), una cola de polihistidina, una glutatión S transferasa (GST), una proteína de unión a almidón (SBP) o un epítipo myc. Puede usarse una molécula de unión a cola para purificar la GalNAcT2 replegada. Cuando la cola de aminoácidos es MBP y la molécula de unión a cola es generalmente una de las siguientes: amilosa, maltosa o una ciclodextrina.

En una realización preferida, la GalNAcT2 replegada cataliza la transferencia de N-acetilgalactosamina de un sustrato donador a un péptido, una proteína, un glicopéptido o una glicoproteína.

10 Breve descripción de los dibujos

La figura 1 proporciona las condiciones tampón sometidas a prueba en el replegamiento de MBP-ST3GalIII a partir de cuerpos de inclusión bacterianos. También se proporciona la actividad de las enzimas replegadas.

La figura 2 proporciona un perfil de elución de MBP-ST3GalIII replegada y dializada de una columna de amilosa.

La figura 3 proporciona las actividades de ST3GalIII de las fracciones de elución de la columna de amilosa.

15 La figura 4 proporciona los resultados de un ensayo de glicopegilación de transferrina usando MBP-ST3GalIII replegada purificada. Los carriles son tal como sigue: (1) marcadores de PM [250, 148, 98, 64, 50 kDa]; (2) asialotransferrina de control sin enzima, indicado mediante fecha continua; (3) producción de transferrina-SA-PEG (20 kDa) con la fracción n.º 5, productos indicados por la punta de flecha; (4) producción de transferrina-SA-PEG (20 kDa) con la fracción n.º 6, productos indicados por la punta de flecha; (5) MBP-ST3GalIII replegada, purificada Fr. n.º 6, indicado por la flecha de puntos; (6) marcadores de PM; (7) igual que 2; (8) producción de transferrina-SA-PEG (10 kDa) con Fr. n.º 4, productos indicados por llaves; y (9) producción de transferrina-SA-PEG (10 kDa) con Fr. n.º 5, productos indicados por llaves.

25 La figura 5 proporciona los resultados de un ensayo de glicopegilación de EPO usando la SuperGlycoMix replegada. Los carriles son tal como sigue: (1) marcadores de PM, SeeBlue2 Invitrogen, (250, 148, 98, 64, 50, 36, 22, 16, 6 kDa); (2) control positivo con EPO, + GalT1 expresada en NSO, GnT1 BV, ST3GalIII de *Aspergillus* y nucleótidos de azúcar; (3) control negativo, igual que 2 sin UDP-GlcNAc; (4) EPO, MBP-GalT1 purificada y replegada por separado (Δ 129) C342T, MBP-GnT1 replegada (Δ 103) y ST3GalIII expresada en *Aspergillus niger*; (5) EPO, SuperGlycoMix (mezcla de MBP-ST3GalIII, MBP-GalT1 (Δ 129) C342T, MBP-GnT1(Δ 103) C123A y nucleótidos de azúcar).

30 La figura 6 proporciona una alineación de una secuencia de aminoácidos de GnT1 humana (línea superior, NP_002397; SEQ ID NO:1) y una secuencia de GnT1 de conejo (línea inferior, P27115; SEQ ID NO:2). Las cisteínas desapareadas conservadas están subrayadas y en negrita.

35 La figura 7 proporciona la secuencia de aminoácidos (SEQ ID NO:3) de un mutante de Cis121Ser de GnT1 y una secuencia de ácido nucleico (SEQ ID NO:4) que codifica para la proteína GnT1 mutante. La secuencia de aminoácidos representada comienza con el residuo de aminoácido 104 de la proteína humana de longitud completa y es representativa de proteínas GnT1 de mamífero con la siguiente mutación de cisteína desapareada: ... stvrrsl~~dkllh~~... (SEQ ID NO:5), en el que se ha mutado el residuo en negrita con respecto a la cisteína de tipo natural.

40 La figura 8 proporciona la secuencia de aminoácidos (SEQ ID NO:6) de un mutante de Cis121Asp de GnT1 y una secuencia de ácido nucleico (SEQ ID NO:7) que codifica para la proteína GnT1 mutante. La secuencia de aminoácidos representada comienza con el residuo de aminoácido 104 de la proteína humana de longitud completa y es representativa de proteínas GnT1 de mamífero con la siguiente mutación de cisteína desapareada: ... stvrrld~~dkllh~~... (SEQ ID NO:8), en el que se ha mutado el residuo en negrita con respecto a la cisteína de tipo natural.

45 La figura 9 proporciona la secuencia de aminoácidos (SEQ ID NO:9) de un mutante de Cis121Thr de GnT1 y una secuencia de ácido nucleico (SEQ ID NO:10) que codifica para la proteína GnT1 mutante. La secuencia de aminoácidos representada comienza con el residuo de aminoácido 104 de la proteína humana de longitud completa y es representativa de proteínas GnT1 de mamífero con la siguiente mutación de cisteína desapareada: ... stvrrtld~~dkllh~~... (SEQ ID NO:11), en el que se ha mutado el residuo en negrita con respecto a la cisteína de tipo natural.

50 La figura 10 proporciona la secuencia de aminoácidos (SEQ ID NO:12) de un mutante de Cis121Ala de GnT1 y una secuencia de ácido nucleico (SEQ ID NO:13) que codifica para la proteína GnT1 mutante. La secuencia de aminoácidos representada comienza con el residuo de aminoácido 104 de la proteína humana de longitud completa y es representativa de proteínas GnT1 de mamífero con la siguiente mutación de cisteína desapareada: ... stvrrald~~dkllh~~... (SEQ ID NO: 14), en el que se ha mutado el residuo en negrita con respecto a la cisteína de tipo natural.

La figura 11 proporciona la secuencia de aminoácidos (SEQ ID NO: 15) de un mutante de Arg120Ala, Cis121His de GnT1 y una secuencia de ácido nucleico (SEQ ID NO:16) que codifica para la proteína GnT1 mutante. La secuencia

de aminoácidos representada comienza con el residuo de aminoácido 104 de la proteína humana de longitud completa y es representativa de proteínas GnT1 de mamífero con la siguiente doble mutación: ... stvra**h**ldkllh... (SEQ ID NO:17), en el que se ha mutado el residuo en **negrita** con respecto a la cisteína de tipo natural.

5 La figura 12 proporciona la secuencia de aminoácidos de ST3GalIII de hígado de rata (SEQ ID NO:18). La secuencia subrayada y en cursiva se delecionó para realizar la delección Δ28.

Las figuras 13A y 13B proporcionan las secuencias de ácido nucleico (SEQ ID NO:20) y de aminoácidos (SEQ ID NO:19) de longitud completa de UDP-N-acetilgalactosaminiltransferasa 2 (GalNAcT2). El número de registro del ácido nucleico y la proteína es NM_004481.

10 Las figuras 14A y 14B proporcionan las secuencias de ácido nucleico (SEQ ID NO:22) y de aminoácidos (SEQ ID NO:21) de una Δ51GalNAcT2. La numeración se basa en las secuencias de aminoácidos y de ácido nucleico de longitud completa mostradas en las figuras 13A y B.

15 La figura 15 proporciona una demostración de la concentración de proteína de MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y replegamiento a pH 6,5 o pH 8,0. Los valores de pH sometidos a prueba se expresan como pH de solubilización-pH de replegamiento. Se midieron las concentraciones de proteína inmediatamente tras replegamiento (barras de color gris claro), tras diálisis (barras de color gris oscuro), y tras concentración (barras de color blanco).

20 La figura 16 proporciona una demostración de la actividad enzimática de MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y replegamiento a pH 6,5 o pH 8,0. Los valores de pH sometidos a prueba se expresan como pH de solubilización-pH de replegamiento. Se midió la actividad tras diálisis (barras de color gris claro) y tras concentración (barras de color gris oscuro).

La figura 17 proporciona una demostración de la actividad específica de MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y replegamiento a pH 6,5 o pH 8,0. Los valores de pH sometidos a prueba se expresan como pH de solubilización-pH de replegamiento. Se midió la actividad específica tras diálisis (barras de color blanco) y tras concentración (barras de color gris oscuro).

25 Las figuras 18A y 18B proporcionan resultados de la remodelación del factor estimulante de colonias de granulocitos (G-CSF) recombinantes usando MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y replegamiento a pH 6,5 o pH 8,0. La figura 18A muestra los resultados usando una MBP-GalNAcT2(D51) purificada de control, o un control negativo que carecía de un sustrato, o MBP-GalNAcT2(D51) expresada de forma bacteriana que se solubilizó a pH 6,5 y se replegó a pH 6,5. La figura 18B muestra los resultados experimentales.

30 La figura 19 proporciona un perfil de proteínas MBP-GalNAcT2(D51) replegadas tras elución desde una columna Q Sepharose XL (QXL) (Amersham Biosciences, Piscataway, NJ). La parte superior de la figura muestra un cromatograma que ilustra la elución de MBP-GalNAcT2(D51) desde la columna QXL. Se indican los números de fracción en el eje X y se indica la absorbancia relativa de cada fracción en el eje Y. La parte inferior muestra una imagen de dos geles electroforéticos usados para visualizar las fracciones eluidas. Se describe el contenido de cada carril en el gel en la figura.

35 La figura 20 proporciona la actividad de GalNAcT2 de fracciones de columna específicas de la columna QXL mostradas en la figura 19. Se aplicaron las fracciones más activas a una columna de hidroxiapatita tipo I (80 μm) (BioRad, Hercules, CA).

40 La figura 21 proporciona un perfil de proteínas MBP-GalNAcT2(D51) replegadas tras elución desde la columna de HA tipo I. La parte superior de la figura muestra un cromatograma que ilustra la elución de MBP-GalNAcT2(D51) desde la columna de HA tipo I. Se indican los números de fracción en el eje X y se indica la absorbancia relativa de cada fracción en el eje Y. La parte inferior muestra una imagen de un gel electroforético usado para visualizar las fracciones eluidas. Se describe el contenido de cada carril en el gel en la figura.

La figura 22 proporciona la actividad de fracciones de GalNAcT2 eluidas de HA tipo I.

45 La figura 23 proporciona una comparación de purificación y actividad de proteínas ST3Gal3 fusionadas o bien a una cola de MBP o bien a una cola de MBP-SBD.

La figura 24 proporciona las secuencias de aminoácidos de la proteína de fusión MBP-ST3Gal1 (SEQ ID NO:23) (A) y la proteína de fusión MBP-SBD-ST3Gal1 (SEQ ID NO: 24) (B).

50 La figura 25 proporciona la actividad sialiltransferasa de la proteína de fusión MBP-ST3Gal3 y la proteína de fusión MBP-SBD-ST3Gal3. También se muestran controles positivos y negativos.

La figura 26 proporciona la secuencia de aminoácidos de proteínas ST6GalNAcI de ratón y humana fusionadas a MBP. La parte A muestra la secuencia de una fusión con truncamiento de ratón: MBP-mST6GalNAcI S127 (SEQ ID

- NO:25). La parte B muestra la secuencia de una fusión con truncamiento humana: MBP-hST6GalNAcl K36 (SEQ ID NO:26).
- 5 La figura 27 proporciona geles de SDS-PAGE de (A) concentraciones de enzima glicosiltransferasa unida a O tras replegamiento conjunto y los (B) resultados de un ensayo enzimático tras replegamiento conjunto. MBP-GalNAcT2 y MBP-ST3Gal se replegaron conjuntamente. Se sometió a prueba la actividad enzimática tras la adición de enzima Core I Gal T1. Los sustratos fueron Sia-2b y 20K-Peg-CMP-NAN.
- La figura 28 proporciona un gel de SDS-PAGE que muestra la expresión de la proteína SiaA nativa en *E. coli* antes y tras inducción con IPTG.
- 10 La figura 29 proporciona un gel de SDS-PAGE que muestra la expresión de una proteína de fusión MBP-SiaA en *E. coli* antes y tras inducción con IPTG.
- La figura 30 proporciona la secuencia de aminoácidos de la proteína GalT1 bovina de longitud completa (SEQ ID NO:27).
- La figura 31 representa mutantes de GalT1 esquemáticamente, así como una proteína de control GalT1(40) (S96A+C342T).
- 15 La figura 32 proporciona los resultados de ensayos enzimáticos de la proteína MBP-GalT1(D70) replegada y purificada. El ensayo midió la conversión de LNT2 (lacto-N-triosa-2) en LNnT (lacto-N-neotetraosa) usando UDP-Gal (uridina 5'-difosfogalactosa) como sustrato donador.
- La figura 33 proporciona un ensayo de remodelación con ARNasa B de MBP-GalT1(D70) y una proteína de control GalT1(40) (S96A+C342T), también denominada GalT1 de Qasba.
- 20 La figura 34 proporciona la cinética de glicosilación de ARNasa B usando la proteína MBP-GalT1(D70) replegada y purificada o NSO GalT1, una forma soluble de la proteína GalT1 bovina que se expresó en un sistema de células de mamífero.
- La figura 35 proporciona un esquema de las proteínas de fusión MBP-GnT1, y representa los truncamientos, por ejemplo, Δ103 o Δ35, y la mutación Cis121Ser (parte superior). La parte inferior de la figura proporciona la proteína GnT1 humana de longitud completa (SEQ ID NO:1).
- 25 La figura 36 proporciona un gel de SDS-PAGE que muestra en el panel derecho las proteínas de fusión MBP-GnT1 replegadas: MBP-GnT1(D35) C121A, MBP-GnT1(D103) R120A + C121H, y MBP-GnT1(D103) C121A. El panel izquierdo muestra actividades de GnT1 de dos lotes diferentes (A1 y A2) de MBP-GnT1(D35) C121A replegada a diferentes puntos de tiempo.
- 30 La figura 37 proporciona una secuencia de longitud completa de ST3Gal1 porcina (SEQ ID NO:28).
- La figura 38 proporciona secuencias de aminoácidos de longitud completa para A) ST6GalNAcT1 humana (SEQ ID NO: 29) y para B) ST6GalNAcT1 de gallina (SEQ ID NO:30), y C) una secuencia de la proteína ST6GalNAcT1 de ratón que comienza en el residuo 32 de la proteína de ratón nativa (SEQ ID NO:31).
- La figura 39 proporciona un esquema de varios mutantes de truncamiento de ST6GalNAcl humana preferidos.
- 35 La figura 40 muestra un esquema de proteínas de fusión con MBP incluyendo los mutantes de truncamiento de ST6GalNAcl humana.
- La figura 41 proporciona la secuencia de longitud completa de la proteína Core 1 GalT1 humana (SEQ ID NO:32).
- La figura 42 proporciona las secuencias de dos proteínas Core 1 GalT1 de *Drosophila* (SEQ ID NO:33 y 34).
- 40 La figura 43 proporciona las secuencias de proteínas MBP bacterianas a modo de ejemplo que pueden fusionarse glicosiltransferasas para mejorar el replegamiento A. MBP de *Yersinia* (SEQ ID NO:35); B. MBP de *E. coli* (SEQ ID NO:36); C. MBP de *Pyrococcus furiosus* (SEQ ID NO:37); D. MBP de *Thermococcus litoralis* (SEQ ID NO:38); E. MBP de *Thermatoga maritime* (SEQ ID NO:39); y F. MBP de *Vibrio cholera* (SEQ ID NO:40).
- 45 La figura 44 proporciona una alineación de proteínas GalNAcT1 (SEQ ID NO:41) y GalNAcT2 (SEQ ID NO:19) humanas. Dado que los programas de alineación tienen en cuenta las inserciones o deleciones de secuencia, la numeración de residuos de cisteína no es la misma que la de las secuencias publicadas y el texto mencionado. En el caso de la cisteína 227 de hGalNAc-T2 (publicada) corresponde a la posición 235 en la alineación y la cisteína 229 (publicada) es 237 en la alineación. Las cisteínas de hGalNAc-T1 son 212 (publicada), que corresponde a la cisteína 235 (alineación) y 214 (publicada), que corresponde a la cisteína 237 (alineación). Los residuos de cisteína relevantes se indican mediante un tamaño de fuente mayor. Péptidos de consenso = SEQ ID NOS: 42-65.

La figura 45 muestra la posición de residuos de cisteína apareados y desapareados en la proteína ST6GalNAcI. También se muestran sustituciones de cisteína única y doble, por ejemplo, C280S, C362S, C362T, (C280S + C362S) y (C280S + C362T).

Definiciones

5 Las proteínas glicosiltransferasas recombinantes de la invención son útiles para transferir un sacárido de un sustrato donador a un sustrato aceptor. La adición tiene lugar generalmente en el extremo no reductor de un resto de oligosacárido o hidrato de carbono en una biomolécula. Las biomoléculas tal como se definen en el presente documento incluyen pero no se limitan a moléculas biológicamente significativas tales como hidratos de carbono, proteínas (por ejemplo, glicoproteínas), y lípidos (por ejemplo, glicolípidos, fosfolípidos, esfingolípidos y gangliósidos).

10

Se usan las siguientes abreviaturas en el presente documento:

Ara = arabinosilo;

Fru = fructosilo;

Fuc = fucosilo;

15 Gal = galactosilo;

GalNAc = N-acetilgalactosilamino;

Glc = glucosilo;

GlcNAc = N-acetilglucosilamino;

Man = manosilo; y

20 NeuAc = sialilo(N-acetilneuraminilo)

FT o FucT = fucosiltransferasa*

ST = sialiltransferasa*

GalT= galactosiltransferasa*

25 Se usan números arábigos o romanos de manera intercambiable en el presente documento según la convención de denominación utilizada en la técnica para indicar la identidad de una glicosiltransferasa específica (por ejemplo, FTVII y FT7 se refieren a la misma fucosiltransferasa).

Se considera que los oligosacáridos tienen un extremo reductor y un extremo no reductor, ya sea o no el sacárido en el extremo reductor de hecho un azúcar reductor. Según la nomenclatura aceptada, los oligosacáridos se representan en el presente documento con el extremo no reductor a la izquierda y el extremo reductor a la derecha.

30 Todos los oligosacáridos descritos en el presente documento se describen con el nombre o la abreviatura para el sacárido no reductor (por ejemplo, Gal), seguido por la configuración del enlace glucosídico (α o β), el enlace de anillo, la posición de anillo del sacárido reductor implicado en el enlace, y luego el nombre o la abreviatura del sacárido reductor (por ejemplo, GlcNAc). La unión entre dos azúcares puede expresarse, por ejemplo, como 2,3, 2 \rightarrow 3 o (2,3). Cada sacárido es una piranosa o furanosa.

35 La expresión "ácido siálico" se refiere a cualquier miembro de una familia de azúcares carboxilados de nueve carbonos. El elemento más común de la familia del ácido es el ácido N-acetil-neuramínico ácido (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 se hidroxila el grupo N-acetilo de NeuAc. 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 incluidos los ácidos siálicos 9-sustituidos tales como un 9-O-acil C₁-C₆-Neu5Ac como 9-O-lactil-Neu5Ac o 9-O-acetil-Neu5Ac, 9-desoxi-9-fluoro-Neu5Ac y 9-azido-9-desoxi-Neu5Ac. Para una revisión de la familia de ácido siálico véanse, 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 el uso de compuestos de ácido siálico en un procedimiento de sialilación se dan a conocer en la solicitud internacional WO 45 92/16640, publicada el 1 de octubre de 1992.

Un "sustrato aceptor" para una glicosiltransferasa es un resto de oligosacárido que puede actuar como aceptor para una glicosiltransferasa particular. Cuando el sustrato aceptor se pone en contacto con la correspondiente glicosiltransferasa y el sustrato donador de azúcar, y otros componentes necesarios de la mezcla de reacción, y se

incuba la mezcla de reacción durante un periodo de tiempo suficiente, la glicosiltransferasa transfiere residuos de azúcar desde el sustrato donador de azúcar hasta el sustrato aceptor. El sustrato aceptor variará a menudo para diferentes tipos de una glicosiltransferasa particular. Por ejemplo, el sustrato aceptor para un galactósido de mamífero 2-L-fucosiltransferasa (α 1,2-fucosiltransferasa) incluirá un Gal β 1,4-GlcNAc-R en un extremo terminal no reductor de un oligosacárido; esta fucosiltransferasa une un residuo de fucosa a Gal mediante una unión α 1,2. Gal β 1,4-GlcNAc-R y Gal β 1,3-GlcNAc-R terminales y los análogos sialilados de los mismos son sustratos aceptores para α 1,3 y α 1,4-fucosiltransferasas, respectivamente. Estas enzimas, sin embargo, unen el residuo de fucosa al residuo de GlcNAc del sustrato aceptor. Por consiguiente, la expresión "sustrato aceptor" se toma en el contexto con la glicosiltransferasa particular de interés para una aplicación particular. Sustratos aceptores para glicosiltransferasas adicionales, se describen en el presente documento. Los sustratos aceptores también incluyen por ejemplo, péptidos, proteínas, glicopéptidos y glicoproteínas.

Un "sustrato donador" para glicosiltransferasas es un azúcar de nucleótido activado. Tales azúcares activados consisten generalmente en derivados de monofosfato de uridina, guanosina y 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 sirve como grupo saliente. Por ejemplo, un sustrato donador para fucosiltransferasas es GDP-fucosa. Sustratos donadores para sialiltransferasas, por ejemplo, son nucleótidos de azúcar activados que comprenden el ácido siálico deseado. Por ejemplo, en el caso de NeuAc, el azúcar activado es CMP-NeuAc. Otros sustratos donadores incluyen por ejemplo, GDP manosa, UDP-galactosa, UDP-N-acetilgalactosamina, CMP-NeuAc-PEG (también denominado CMP-ácido siálico-PEG), UDP-N-acetilglucosamina, UDP-glucosa, UDP-ácido glucuriónico y UDP-xilosa. Los azúcares incluyen, por ejemplo, NeuAc, manosa, galactosa, N-acetilgalactosamina, N-acetilglucosamina, glucosa, ácido glucuriónico y xilosa. Los sistemas bacterianos vegetales y fúngicos pueden usar a veces otros azúcares de nucleótido activados.

Un "método de remodelación de una proteína, un péptido, una glicoproteína o un glicopéptido" tal como se usa en el presente documento, se refiere a la adición de un residuo de azúcar a una proteína, un péptido, una glicoproteína o un glicopéptido usando una glicosiltransferasa. En una realización preferida, el residuo de azúcar está unido covalentemente a una molécula de PEG.

Una "glicosiltransferasa eucariota" tal como se usa en el presente documento se refiere a una enzima que se deriva de un organismo eucariota y que cataliza la transferencia de un residuo de azúcar de un sustrato donador, es decir, un azúcar de nucleótido activado a un sustrato aceptor, por ejemplo, un oligosacárido, un glicolípido, un péptido, una proteína, un glicopéptido o una glicoproteína. En realizaciones preferidas, una glicosiltransferasa eucariota transfiere un azúcar de un sustrato donador a un péptido, una proteína, un glicopéptido o una glicoproteína. En otra realización preferida, una glicosiltransferasa eucariota es una glicosiltransferasa transmembrana tipo II. Una glicosiltransferasa eucariota puede derivarse de un organismo eucariota, por ejemplo, un organismo eucariota multicelular, una planta, un animal invertebrado, tal como *Drosophila* o *C. elegans*, un animal vertebrado, un anfibio o reptil, un mamífero, un roedor, un primate, un ser humano, un conejo, una rata, un ratón, una vaca o un cerdo, etcétera.

Una "N-acetilglucosaminiltransferasa I (GnTI o GNTI) eucariota" tal como se usa en el presente documento, se refiere a una β -1,2-N-acetilglucosaminiltransferasa I aislada de un organismo eucariota. La enzima cataliza la transferencia de N-acetilglucosamina (GlcNAc) de un donador de UDP-GlcNAc a una molécula aceptora que comprende un azúcar de manosa. Como otras glicosiltransferasas eucariotas, GnTI tiene un dominio transmembrana, una región de tallo y un dominio catalítico. Las proteínas GnTI eucariotas incluyen, por ejemplo, humana, número de registro NP_002397; hámster chino, número de registro AAK61868; conejo, número de registro AAA31493; rata, número de registro NP_110488; hámster dorado, número de registro AAD04130; ratón, número de registro P27808; pez cebra, número de registro AAH58297; *Xenopus*, número de registro CAC51119; *Drosophila*, número de registro NP_525117; *Anopheles*, número de registro XP_315359; *C. elegans*, número de registro NP_497719; *Physcomitrella patens*, número de registro CAD22107; *Solarium tuberosum*, número de registro CAC80697; *Nicotiana tabacum*, número de registro CAC80702; *Oryza sativa*, número de registro CAD30022; *Nicotiana benthamiana*, número de registro CAC82507; y *Arabidopsis thaliana*, número de registro NP_195537.

Una "N-acetilgalactosaminiltransferasa (GalNAcT) eucariota" tal como se usa en el presente documento, se refiere a una N-acetilgalactosaminiltransferasa aislada de un organismo eucariota. La enzima cataliza la transferencia de N-acetilgalactosamina (GalNAc) de un donador UDP-GalNAc a una molécula aceptora. Como otras glicosiltransferasas eucariotas, las enzimas GalNAcT tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Se han aislado y caracterizado varias enzimas GalNAcT, por ejemplo, GalNAcT1, número de registro X85018; GalNAcT2, número de registro X85019 (ambas descritas en White *et al.*, J. Biol. Chem. 270: 24156-24165 (1995)); y GalNAcT3, número de registro X92689 (descrita en Bennett *et al.*, J. Biol. Chem. 271: 17006-17012 (1996)).

Una " β -1,4-galactosiltransferasa (GalT1) eucariota" tal como se usa en el presente documento, se refiere a una β -1,4-galactosiltransferasa aislada de un organismo eucariota. La enzima cataliza la transferencia de galactosa de un donador UDP-Gal a una molécula aceptora. Como otras glicosiltransferasas eucariotas, las enzimas GalT1 tienen un dominio transmembrana, una región de tallo, y un dominio catalítico. Se han aislado y caracterizado varias enzimas GalT1, por ejemplo, la secuencia bovina de longitud completa, D'Agostaro *et al.*, Eur. J. Biochem. 183: 211-217 (1989) y número de registro CAA32695.

Una "eucariota $\alpha(2,3)$ sialiltransferasa (ST3Gal3) eucariota" tal como se usa en el presente documento, se refiere a una $\alpha(2,3)$ sialiltransferasa aislada de un organismo eucariota. Esta enzima cataliza la transferencia de ácido siálico a la Gal de un glicósido Gal β 1,3GlcNAc, Gal β 1,3GalNAc o Gal β 1,4-GlcNAc (véanse, por ejemplo, Wen *et al.* (1992) J. Biol. Chem. 267: 21011; Van den Eijnden *et al.* (1991) J. Biol. Chem. 266: 3159). El ácido siálico está unido a una Gal con la formación de una unión α entre los dos sacáridos. El enlace (unión) entre los sacáridos es entre la posición 2 de NeuAc y la posición 3 de Gal. Como otras glicosiltransferasas eucariotas, las enzimas ST3GalIII tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Esta enzima particular puede aislarse de hígado de rata (Weinstein *et al.* (1982) J. Biol. Chem. 257: 13845); se conocen las secuencias de ADNc humano (Sasaki *et al.* (1993) J. Biol. Chem. 268: 22782-22787; Kitagawa & Paulson (1994) J. Biol. Chem. 269: 1394-1401) y AND genómico (Kitagawa *et al.* (1996) J. Biol. Chem. 271: 931-938), facilitando la producción de esta enzima mediante expresión recombinante. Se ha clonado ST3GalIII de rata y se conoce la secuencia. Véanse, por ejemplo, Wen *et al.*, J. Biol. Chem. 267: 21011-21019 (1992) y número de registro M97754.

Una " α -N-acetilgalactosaminida α -2,6-sialiltransferasa I (ST6GalNAcT1) eucariota" tal como se usa en el presente documento, se refiere a una α -(2,6)sialiltransferasa aislada de un organismo eucariota. La enzima cataliza la transferencia de ácido siálico de un donador CMP-ácido siálico a una molécula aceptora. La transferencia es una unión α 2,6 a N-acetilgalactosamina-O-Thr/Ser. Como otras glicosiltransferasas eucariotas, las enzimas ST6GalNAcT1 tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Se han aislado y caracterizado varias enzimas ST6GalNAcT1, por ejemplo, la secuencia de ratón de longitud completa, Kurosawa *et al.*, J. Biochem. 127: 845-854 (2000) y número de registro JC7248. Otras secuencias de aminoácidos de ST6GalNAcT1 a modo de ejemplo se encuentran en la figura 38.

Una "gal β 1,3GalNAc α 2,3-sialiltransferasa (ST3Gall) eucariota" tal como se usa en el presente documento, se refiere a una gal β 1,3GalNAc α 2,3-sialiltransferasa aislada de un organismo eucariota. La enzima cataliza la transferencia de ácido siálico de un donador CMP-ácido siálico a una molécula aceptora. La transferencia es una unión α 2,3 a N-acetilgalactosamina-O-Thr/Ser. Como otras glicosiltransferasas eucariotas, las enzimas ST3Gall tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Se han aislado y caracterizado varias enzimas ST3Gall, por ejemplo, la secuencia porcina de longitud completa porcina, Gillespie *et al.*, J. Biol. Chem. 267: 21004-21010 (1992) y número de registro A45073.

Una "core I galactosiltransferasa (Core 1 GalT1) eucariota" tal como se usa en el presente documento se refiere a una proteína con actividad Core 1 β 1,3-galactosiltransferasa. Como otras glicosiltransferasas eucariotas, las enzimas Core 1 GalT1 tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Se han aislado y caracterizado varias enzimas Core 1 GalT1, por ejemplo, las secuencias de *Drosophila* y humana de las figuras 41 y 42. La proteína humana se caracteriza en Ju *et al.*, J. Biol. Chem. 277 (1), 178-186 (2002),

Un "residuo de cisteína desapareado" tal como se usa en el presente documento, se refiere a un residuo de cisteína, que en una proteína plegada correctamente (es decir, una proteína con actividad biológica), no forma un enlace disulfuro con otro residuo de cisteína.

Una "glicosiltransferasa insoluble" se refiere a una glicosiltransferasa que se expresa en cuerpos de inclusión bacterianos. Las glicosiltransferasas insolubles normalmente se solubilizan o desnaturalizan usando por ejemplo, detergentes o agentes caotrópicos o alguna combinación. "Replegamiento" se refiere a un proceso de restauración de la estructura de una glicosiltransferasa biológicamente activa a una glicosiltransferasa que se ha solubilizado o desnaturalizado. Por tanto, un tampón de replegamiento, se refiere a un tampón que mejora o acelera el replegamiento de una glicosiltransferasa.

Un "par redox" se refiere a mezclas de reactivos de tiol reducido y oxidado e incluyen glutatión reducido y oxidado (GSH/GSSG), cisteína/cistina, cisteamina/cistamina, DTT/GSSG y DTE/GSSG. (Véase, por ejemplo, Clark, Cur. Op. Biotech. 12: 202-207 (2001)).

La expresión "poner en contacto" se usa en el presente documento de manera intercambiable con los siguientes: combinado con, añadido a, mezclado con, que se hace pasar por, incubado con, que se hace fluir sobre, etc.

El término "PEG" se refiere a poli(etilenglicol). PEG es un polímero a modo de ejemplo que se ha conjugado a péptidos. Se ha demostrado que el uso de PEG para derivatizar agentes terapéuticos peptídicos reduce la inmunogenicidad de los péptidos y prolonga el tiempo de eliminación de la circulación. Por ejemplo, la patente estadounidense n.º 4.179.337 (Davis *et al.*) se refiere a péptidos no inmunogénicos, tales como enzimas y hormonas peptídicas acopladas a polietilenglicol (PEG) o polipropilenglicol. Se usan entre 10 y 100 moles de polímero por mol de péptido y se mantiene al menos el 15% de la actividad fisiológica.

La expresión "actividad específica" tal como se usa en el presente documento se refiere a la actividad catalítica de una enzima, por ejemplo, una proteína de fusión de glicosiltransferasa recombinante de la presente invención, y puede expresarse en unidades de actividad. Tal como se usa en el presente documento, una unidad de actividad cataliza la formación de 1 μ mol de producto por minuto a una temperatura dada (por ejemplo, a 37°C) y valor de pH

(por ejemplo, a pH 7,5). Por tanto, 10 unidades de una enzima es una cantidad catalítica de esa enzima en la que 10 μmol de sustrato se convierten en 10 μmol de producto en un minuto a una temperatura de, por ejemplo, 37°C y un valor de pH de, por ejemplo, 7,5.

5 Los oligosacáridos “unidos a N” son aquellos oligosacáridos que se unen a una estructura principal peptídica a través de asparagina, por medio de una unión asparagina-N-acetilglucosamina. Los oligosacáridos
 10 unidos a N también se denominan “N-glicanos”. Todos los oligosacáridos unidos a N tienen un núcleo de pentasacárido común de $\text{Man}_3\text{GlcNAc}_2$. Difieren en la presencia de, y en el número de ramificaciones (también denominadas antenas) de azúcares periféricos tales como N-acetilglucosamina, galactosa, N-acetilgalactosamina, fucosa y ácido siálico. Opcionalmente, esta estructura también puede contener una molécula de fucosa y/o una molécula de xilosa de núcleo.

Los oligosacáridos “unidos a O” son aquellos oligosacáridos que se unen a una estructura principal peptídica a través de treonina, serina, hidroxiprolina, tirosina, u otros aminoácidos que contienen hidroxilo.

15 Una “glicofoma sustancialmente uniforme” o un “patrón de glicosilación sustancialmente uniforme”, cuando se refiere a una especie de glicoproteína, se refiere al porcentaje de sustratos aceptores que se glicosilan por la glicosiltransferasa de interés (por ejemplo, fucosiltransferasa). Un experto en la técnica entenderá, que el material de partida puede contener sustratos aceptores glicosilados. Por tanto, la cantidad calculada de glicosilación incluirá sustratos aceptores que se glicosilan mediante los métodos de la invención, así como aquellos sustratos aceptores ya glicosilados en el material de partida.

20 La expresión “actividad biológica” se refiere a una actividad enzimática de una proteína. Por ejemplo, actividad biológica de una sialiltransferasa se refiere a la actividad de transferir un resto de ácido siálico de una molécula donadora a una molécula aceptora. Actividad biológica de una GalNAcT2 se refiere a la actividad de transferir un resto de N-acetilgalactosamina de una molécula donadora a una molécula aceptora. Para proteínas GalNAcT2 , una molécula aceptora puede ser una proteína, un péptido, una glicoproteína o un glicopéptido. Actividad biológica de una proteína GnT1 se refiere a la actividad de transferir un resto de N-acetilglucosamina de una molécula donadora
 25 a una molécula aceptora. Actividad biológica de una galactosiltransferasa se refiere a la actividad de transferir un resto de galactosa de una molécula donadora a una molécula aceptora.

“Escala comercial” se refiere a la producción a escala de gramos de un producto sacárido en una única reacción. En realizaciones preferidas, escala comercial se refiere a la producción de más de aproximadamente 50, 75, 80, 90 ó 100, 125, 150, 175 ó 200 gramos.

30 El término “sustancialmente” en las definiciones anteriores de “sustancialmente uniforme” significa generalmente que se glicosilan al menos aproximadamente el 60%, al menos aproximadamente el 70%, al menos aproximadamente el 80%, o más preferiblemente al menos aproximadamente el 90%, y todavía más preferiblemente al menos aproximadamente el 95% de los sustratos aceptores para una glicosiltransferasa particular.

35 El término “aminoácido” se refiere a aminoácidos que se producen de manera natural y sintéticos, así como análogos de aminoácido y miméticos de aminoácido que funcionan de manera similar a los aminoácidos que se producen de manera natural. Los aminoácidos que se producen de manera natural son aquellos codificados por el código genético, así como aquellos aminoácidos que se modifican posteriormente, por ejemplo, hidroxiprolina, γ -carboxiglutamato y O-fosfoserina. Los análogos de aminoácido se refieren a compuestos que tienen la misma estructura química básica que un aminoácido que se produce de manera natural, es decir, un carbono α que está unido a un hidrógeno, un grupo carboxilo, un grupo amino y un grupo R, por ejemplo, homoserina, norleucina, metionina sulfóxido, metionina metil-sulfonio. Tales análogos tienen grupos R modificados (por ejemplo, norleucina)
 40 o estructuras principales peptídicas modificadas, pero conservan la misma estructura química básica que un aminoácido que se produce de manera natural. Miméticos de aminoácido se refiere a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que funciona de manera similar a un aminoácido que se produce de manera natural.
 45

“Proteína”, “polipéptido” o “péptido” se refieren a un polímero en el que los monómeros son aminoácidos y se unen entre sí a través de enlaces amida, alternativamente denominados un polipéptido. Cuando los aminoácidos son α -aminoácidos, puede usarse o bien el isómero óptico L o bien el isómero óptico D. Adicionalmente, también están incluidos los aminoácidos no naturales, por ejemplo, β -alanina, fenilglicina y homoarginina. También pueden usarse en la presente invención aminoácidos que no están codificados genéticamente. Además, también pueden usarse en la invención aminoácidos que se han modificado para incluir grupos reactivos. Todos los aminoácidos usados en la presente invención pueden ser o bien el isómero D o bien el L. Generalmente se prefieren los isómeros L. Además, también son útiles otros peptidomiméticos en la presente invención. Para una revisión general, véanse, Spatola, A. F., en CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES AND PROTEINS, B. Weinstein, eds., Marcel Dekker, Nueva York, pág. 267 (1983).
 50
 55

El término “recombinante” cuando se usa con referencia a una célula indica que la célula replica un ácido nucleico heterólogo, o expresa un péptido o proteína codificada por un ácido nucleico heterólogo. Las células recombinantes pueden contener genes que no se encuentran dentro de la forma nativa (no recombinante) de la célula. Las células

- recombinantes también pueden contener genes que se encuentran en la forma nativa de la célula en la que se modifican los genes y vuelven a introducirse en la célula mediante medios artificiales. El término también engloba células que contienen un ácido nucleico endógeno con respecto a la célula que se ha modificado sin eliminar el ácido nucleico de la célula; tales modificaciones incluyen las obtenidas mediante reemplazo génico, mutación específica del sitio, y técnicas relacionadas. Una “proteína recombinante” es una que se ha producido por una célula recombinante. En realizaciones preferidas, una glicosiltransferasa eucariota recombinante la produce una célula bacteriana recombinante.
- Una “proteína de fusión” se refiere a una proteína que comprende secuencias de aminoácidos que son además de, en lugar de, menos que, y/o diferentes de las secuencias de aminoácidos que codifican para la proteína de longitud completa original o nativa o subsecuencias de la misma.
- Los componentes de proteínas de fusión incluyen “enzimas auxiliares” y/o “colas de purificación”. “Una enzima auxiliar” tal como se hace referencia en el presente documento, es una enzima que está implicada en catalizar una reacción que, por ejemplo, forma un sustrato para una glicosiltransferasa. Una enzima auxiliar puede, por ejemplo, catalizar la formación de un azúcar de nucleótido que se usa como resto donador por una glicosiltransferasa. Una enzima auxiliar también puede ser una que se usa en la generación de un nucleótido trifosfato requerido para la formación de un azúcar de nucleótido, o en la generación del azúcar que se incorpora en el azúcar de nucleótido. La proteína de fusión recombinante de la invención puede construirse y expresarse como una proteína de fusión con una “cola de purificación” molecular en un extremo, que facilita la purificación de la proteína. Tales colas pueden usarse también para la inmovilización de una proteína de interés durante la reacción de glicosilación. Las colas adecuadas incluyen “colas de epítipo”, que son una secuencia de proteína que reconoce específicamente un anticuerpo. Las colas de epítipo se incorporan generalmente en proteínas de fusión para permitir el uso de un anticuerpo fácilmente disponible para aislar o detectar de manera inequívoca la proteína de fusión. Una “cola FLAG” es una cola de epítipo usada comúnmente, reconocida específicamente por un anticuerpo monoclonal anti-FLAG, que consiste en la secuencia AspTyrLysAspAspAsp Asplys o una variante sustancialmente idéntica de la misma. Los expertos en la técnica conocen otras colas adecuadas, e incluyen, por ejemplo, una cola de afinidad tal como un péptido de hexahistidina, que se unirá a iones metálicos tales como iones níquel o cobalto. Pueden purificarse las proteínas que comprenden colas de purificación usando una pareja de unión que se une a la cola de purificación, por ejemplo, anticuerpos contra la cola de purificación, iones níquel o cobalto o resinas, y amilosa, maltosa o una ciclodextrina. Las colas de purificación también incluyen dominios de unión a almidón, dominios de tiorredoxina de *E. coli* (vectores y anticuerpos disponibles comercialmente de, por ejemplo, Santa Cruz Biotechnology, Inc. y Alpha Diagnostic International, Inc.), y la mitad carboxi-terminal de la proteína SUMO (vectores y anticuerpos disponibles comercialmente de, por ejemplo, Life Sensors Inc.). Se usan preferiblemente dominios de unión a maltosa por su capacidad para mejorar el replegamiento de glicosiltransferasas eucariotas insolubles, pero también pueden usarse para ayudar en la purificación de una proteína de fusión. La purificación de proteínas con dominio de unión a maltosa la conocen los expertos en la técnica. Se describen dominios de unión a almidón en el documento WO 99/15636, incorporado al presente documento como referencia. La purificación por afinidad de una proteína de fusión que comprende un dominio de unión a almidón usando una resina derivatizada con betaciclodextrina (BCD) se describe en el documento USSN 60/468.374, presentado el 5 de mayo de 2003, incorporado al presente documento como referencia en su totalidad.
- La expresión “dominio funcional” con referencia a glicosiltransferasas, se refiere a un dominio de la glicosiltransferasa que confiere o modula una actividad de la enzima, por ejemplo, especificidad por el sustrato aceptor, actividad catalítica, afinidad de unión, localización dentro del aparato de Golgi, anclaje a una membrana celular, u otra actividad biológica o bioquímica. Los ejemplos de dominios funcionales de glicosiltransferasas incluyen, pero no se limitan a, el dominio catalítico, región de tallo y dominio de anclaje de señales.
- La expresión “nivel de expresión” con referencia a una proteína se refiere a la cantidad de una proteína producida por una célula. La cantidad de proteína producida por una célula puede medirse mediante los ensayos y unidades de actividad descritos en el presente documento o conocidos por un experto en la técnica. Un experto en la técnica sabría cómo medir y describir la cantidad de proteína producida por una célula usando una variedad de ensayos y unidades, respectivamente. Por tanto, la cuantificación y descripción cuantitativa del nivel de expresión de una proteína, por ejemplo, una glicosiltransferasa, no se limita a los ensayos usados para medir la actividad o las unidades usadas para describir la actividad, respectivamente. La cantidad de proteína producida por una célula puede determinarse mediante ensayos conocidos convencionales, por ejemplo, el ensayo de proteínas de Bradford (1976), el kit de ensayo de proteínas con ácido bicinonínico de Pierce (Rockford, Illinois), o tal como se describe en la patente estadounidense n.º 5.641.668.
- La expresión “actividad enzimática” se refiere a una actividad de una enzima y puede medirse mediante los ensayos y las unidades descritos en el presente documento o conocidos por un experto en la técnica. Los ejemplos de una actividad de una glicosiltransferasa incluyen, pero no se limitan a, aquellos asociados con los dominios funcionales de la enzima, por ejemplo, especificidad por el sustrato aceptor, actividad catalítica, afinidad de unión, localización dentro del aparato de Golgi, anclaje a una membrana celular, u otra actividad biológica o bioquímica.
- Una “región de tallo” con referencia a glicosiltransferasas se refiere a un dominio de proteína, o una subsecuencia de la misma, que en las glicosiltransferasas nativas se ubica adyacente al dominio transmembrana, y se ha notificado

- que funciona como una señal de retención para mantener la glicosiltransferasa en el aparato de Golgi y como sitio de escisión proteolítica. Las regiones de tallo generalmente comienzan con el primer aminoácido hidrófilo que sigue al dominio transmembrana hidrófobo y termina en el dominio catalítico, o en algunos casos el primer residuo de cisteína que sigue al dominio transmembrana. Las regiones de tallo a modo de ejemplo incluyen, pero no se limitan a, la región de tallo de fucosiltransferasa VI, residuos de aminoácido 40-54; la región de tallo de GnT1 de mamífero, residuos de aminoácido desde aproximadamente el 36 hasta aproximadamente el 103 (véase, por ejemplo, la enzima humana); la región de tallo de GalT1 de mamífero, residuos de aminoácido desde aproximadamente el 71 hasta aproximadamente el 129 (véase por ejemplo, la enzima bovina); la región de tallo de ST3GalIII de mamífero, residuos de aminoácido desde aproximadamente el 29 hasta aproximadamente el 84 (véanse, por ejemplo, la enzima de rata); la región de tallo de Core I GalT1 de invertebrado, residuos de aminoácido desde aproximadamente el 36 hasta aproximadamente el 102 (véase por ejemplo, la enzima de *Drosophila*); la región de tallo de Core 1 GalT1 de mamífero, residuos de aminoácido desde aproximadamente el 32 hasta aproximadamente el 90 (véase por ejemplo, la enzima humana); la región de tallo de ST3Gal1 de mamífero, residuos de aminoácido desde aproximadamente el 28 hasta aproximadamente el 61 (véase por ejemplo, la enzima porcina) o para los residuos de aminoácido de la enzima humana de desde aproximadamente el 18 hasta aproximadamente el 58; la región de tallo de ST6GalNAcI de mamífero, residuos de aminoácido desde aproximadamente el 30 hasta aproximadamente el 207 (véase por ejemplo, la enzima murina), aminoácidos 35-278 para la enzima humana o los aminoácidos 37-253 para la enzima de gallina; la región de tallo de GalNAcT2 de mamífero, residuos de aminoácido desde aproximadamente el 71 hasta aproximadamente el 129 (véase por ejemplo, la enzima de rata).
- Un “dominio catalítico” se refiere a un dominio de proteína, o una subsecuencia del mismo, que cataliza una reacción enzimática realizada por la enzima. Por ejemplo, un dominio catalítico de una sialiltransferasa incluirá una subsecuencia de la sialiltransferasa suficiente para transferir un residuo de ácido siálico de un donador a un sacárido aceptor. Un dominio catalítico puede incluir una enzima completa, una subsecuencia de la misma, o puede incluir secuencias de aminoácidos adicionales que no están unidas a la enzima, o una subsecuencia de la misma, tal como se encuentra en la naturaleza. Una región catalítica a modo de ejemplo es, pero no se limita a, el dominio catalítico de fucosiltransferasa VII, residuos de aminoácido 39-342; el dominio catalítico de GnT1 de mamífero, residuos de aminoácido desde aproximadamente el 104 hasta aproximadamente el 445 (véase, por ejemplo, la enzima humana); el dominio catalítico de GalT1 de mamífero, residuos de aminoácido desde aproximadamente el 130 hasta aproximadamente el 402 (véase por ejemplo, la enzima bovina); y el dominio catalítico de ST3GalIII de mamífero, residuos de aminoácido desde aproximadamente el 85 hasta aproximadamente el 374 (véase, por ejemplo, la enzima de rata). Se describen dominios catalíticos y mutantes de truncamiento de proteínas GalNAcT2 en el documento USSN 60/576.530 presentado el 3 de junio de 2004; y en la solicitud de patente provisional estadounidense con número de expediente del apoderado 040853-01-5149-P1, presentada el 3 de agosto de 2004. También pueden identificarse dominios catalíticos mediante alineación con glicosiltransferasas conocidas.
- Una “subsecuencia” se refiere a una secuencia de ácidos nucleicos o aminoácidos que comprende una parte de una secuencia de ácidos nucleicos o aminoácidos más larga (por ejemplo, proteína) respectivamente.
- Un “truncamiento de glicosiltransferasa” o una “glicosiltransferasa truncada” o variantes gramaticales, se refieren a una glicosiltransferasa que tiene menos residuos de aminoácido que una glicosiltransferasa que se produce de manera natural, pero que conserva la actividad enzimática. Las glicosiltransferasas truncadas incluyen, por ejemplo, enzimas GnT1 truncadas, enzimas GalT1 truncadas, enzimas ST3GalIII truncadas, enzimas GalNAcT2 truncadas, enzimas Core 1 GalT1 truncadas, residuos de aminoácido desde aproximadamente el 32 hasta aproximadamente el 90 (véase por ejemplo, la enzima humana); enzimas ST3Gal1 truncadas, enzimas ST6GalNAcI truncadas y enzimas GalNAcT2 truncadas. Puede deleccionarse cualquier número de residuos de aminoácido siempre que la enzima conserve la actividad. En algunas realizaciones, pueden deleccionarse dominios o partes de dominios, por ejemplo, puede deleccionarse un dominio de anclaje de señales dejando un truncamiento que comprende una región de tallo y un dominio catalítico; puede deleccionarse un dominio de anclaje de señales y una parte de una región de tallo dejando un truncamiento que comprende la región de tallo restante y un dominio catalítico; o pueden deleccionarse un dominio de anclaje de señales y una región de tallo dejando un truncamiento que comprende un dominio catalítico.
- La expresión “ácido nucleico” se refiere a un polímero de desoxirribonucleótido o ribonucleótido en forma o bien monocatenaria o bien bicatenaria, y a menos que se limite de otro modo, engloba análogos conocidos de nucleótidos naturales que se hibridan con ácidos nucleicos de manera similar a nucleótidos que se producen de manera natural. A menos que se indique de otro modo, una secuencia de ácido nucleico particular incluye la secuencia complementaria de la misma.
- Un “casete de expresión recombinante” o simplemente un “casete de expresión” es un constructo de ácido nucleico, generado de manera recombinante o sintética, con elementos de ácido nucleico que pueden afectar la expresión de un gen estructural en huéspedes compatibles con tales secuencias. Los casetes de expresión incluyen al menos promotores y opcionalmente, señales de terminación de la transcripción. Normalmente, el casete de expresión recombinante incluye un ácido nucleico que va a transcribirse (por ejemplo, un ácido nucleico que codifica para un polipéptido deseado), y un promotor. También pueden usarse factores adicionales necesarios o útiles en la realización de la expresión tal como se describe en el presente documento. Por ejemplo, un casete de expresión

también puede incluir secuencias de nucleótidos que codifican para una secuencia señal que dirige la secreción de una proteína expresada de la célula huésped. Señales de terminación de la transcripción, potenciadores y otras secuencias de ácido nucleico que influyen en la expresión génica, también pueden estar incluidos en un casete de expresión. En realizaciones preferidas, un casete de expresión recombinante que codifica para una secuencia de aminoácidos que comprende una glicosiltransferasa eucariota se expresa en una célula huésped bacteriana.

Una "secuencia heteróloga" o un "ácido nucleico heterólogo", tal como se usa en el presente documento, es el que se origina a partir de una fuente foránea para la célula huésped particular, o, si es a partir de la misma fuente, se modifica con respecto a su forma original. Por tanto, un gen de glicoproteína heteróloga en una célula huésped eucariota incluye un gen que codifica para glicoproteína que es endógeno con respecto a la célula huésped particular que se ha modificado. La modificación de la secuencia heteróloga puede producirse, por ejemplo, tratando el ADN con una enzima de restricción para generar un fragmento de ADN que puede unirse operativamente al promotor. También son útiles técnicas tales como mutagénesis dirigida al sitio para modificar una secuencia heteróloga.

El término "aislado" se refiere a material que está sustancialmente o esencialmente libre de componentes que interfieren con la actividad de una enzima. Para un sacárido, una proteína o un ácido nucleico de la invención, el término "aislado" se refiere a material que está sustancialmente o esencialmente libre de componentes que normalmente acompañan al material tal como se encuentra en su estado nativo. Normalmente, un sacárido, una proteína o un ácido nucleico aislados de la invención es puro en al menos aproximadamente el 80%, habitualmente al menos aproximadamente el 90%, y preferiblemente puro en al menos aproximadamente el 95% tal como se mide mediante la intensidad de banda en un gel teñido con plata u otro método para determinar la pureza. La pureza u homogeneidad puede indicarse mediante varios medios bien conocidos en la técnica. Por ejemplo, puede resolverse una proteína o un ácido nucleico en una muestra mediante electroforesis en gel de poliacrilamida, y luego puede visualizarse la proteína o el ácido nucleico mediante tinción. Para determinados fines, puede ser deseable una alta resolución de la proteína o el ácido nucleico y puede utilizarse HPLC o un medio similar para purificación, por ejemplo.

La expresión "operativamente unido" se refiere a una unión funcional entre una secuencia de control de la expresión de ácido nucleico (tal como un promotor, una secuencia señal o una matriz de sitios de unión a 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 traducción del ácido nucleico correspondiente a la segunda secuencia.

Las expresiones "idéntica" o "identidad" en porcentaje, en el contexto de dos o más secuencias de ácidos nucleicos o proteína, se refieren a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de residuos de aminoácido o nucleótidos que son iguales, cuando se compara y alinean para una máxima correspondencia, tal como se mide usando uno de los siguientes algoritmos de comparación de secuencias o mediante inspección visual.

La frase "sustancialmente idéntica", en el contexto de dos ácidos nucleicos o proteínas, se refiere a dos o más secuencias o subsecuencias que tienen una identidad de secuencia de ácido nucleico o aminoácidos al menos superior a aproximadamente el 60%, 65%, 70%, 75%, 80%, 85%, 90%, preferiblemente una identidad de nucleótidos o residuos de aminoácido del 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99%, cuando se compara y alinea para una máxima correspondencia, tal como se mide usando uno de los siguientes algoritmos de comparación de secuencias o mediante inspección visual. Preferiblemente, la identidad sustancial existe en una región de las secuencias que es al menos de aproximadamente 50 residuos de longitud, más preferiblemente en una región de al menos aproximadamente 100 residuos, y lo más preferiblemente las secuencias son sustancialmente idénticas en al menos aproximadamente 150 residuos. En la realización más preferida, las secuencias son sustancialmente idénticas en la totalidad de la longitud de las regiones codificantes.

Para la comparación de secuencias, normalmente una secuencia actúa como secuencia de referencia, con la que se comparan secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, se introducen 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 secuencias. Luego, el algoritmo de comparación de secuencias calcula la identidad de secuencia en porcentaje para la(s) secuencia(s) de prueba con relación a la secuencia de referencia, basándose en los parámetros de programa designados.

Puede llevarse a cabo una alineación de secuencias óptima para comparación, por ejemplo, mediante el algoritmo de homología local de Smith & Waterman, *Adv. Appl. Math.* 2: 482 (1981), mediante el algoritmo de alineación por homología de Needleman & Wunsch, *J. Mol. Biol.* 48: 443 (1970), mediante el método de búsqueda por similitud de Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85: 2444 (1988), mediante implementaciones informatizadas 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 (véanse generalmente, *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., (1995 Suplemento) (Ausubel)).

Ejemplos de algoritmos que son adecuados para determinar la similitud de secuencia y la identidad de secuencia en porcentaje 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. Está disponible para el público el software para realizar análisis BLAST a través del Centro Nacional para la Información Biotecnológica (www.ncbi.nlm.nih.gov/). Este algoritmo implica identificar en primer lugar pares de secuencias de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia de consulta, que o bien coinciden o bien satisfacen cierta puntuación umbral de valor positivo T cuando se alinea con una palabra de la misma longitud en una secuencia de la base de datos. T se denomina el umbral de puntuación de palabra vecina (Altschul *et al.*, citado anteriormente). Estas coincidencias de palabras vecinas inicial actúan como simientes para iniciar búsquedas para hallar HSP más largas que las contienen. Las coincidencias de palabra se extienden luego en ambos sentidos a lo largo de cada secuencia en la medida en que puede aumentarse la puntuación de alineación acumulativa. Las puntuaciones acumulativas se calculan usando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre > 0) y N (puntuación de penalización para residuos no coincidentes; siempre < 0). Para secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de las coincidencias de palabra en cada sentido se detiene cuando: la puntuación de alineación acumulativa disminuye en la cantidad X desde su máximo valor alcanzado; la puntuación acumulativa llega a cero o menos, debido a la acumulación de una o más alineaciones de residuos de puntuación negativa; o se alcanza el extremo de cualquier secuencia. Los parámetros W, T y X del algoritmo BLAST determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa como valores por defecto una longitud de palabra (W) de 11, una esperanza (E) de 10, M=5, N=-4, y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP usa como valores por defecto una longitud de palabra (W) de 3, una esperanza (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89: 10915 (1989)).

Además de calcular la identidad de secuencia en porcentaje, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin & Altschul, *Proc. Nat'l. 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 de que una coincidencia entre dos secuencias de nucleótidos o aminoácidos se produzca al azar. Por ejemplo, se considera que un ácido nucleico es similar a una secuencia de referencia si la menor probabilidad de suma en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es menor que aproximadamente 0,1, más preferiblemente menor que aproximadamente 0,01, y lo más preferiblemente menor que aproximadamente 0,001.

Una indicación adicional de que dos secuencias de ácido nucleico o proteínas son sustancialmente idénticas es que la proteína codificada por el primer ácido nucleico da reacción cruzada de manera inmunológica con la proteína codificada por el segundo ácido nucleico, tal como se describe a continuación. Por tanto, una proteína normalmente es sustancialmente idéntica a una segunda proteína, por ejemplo, cuando los dos péptidos difieren sólo en sustituciones conservativas. Otra indicación de que dos secuencias de ácido nucleico son sustancialmente idénticas es que las dos moléculas se hibridan entre sí en condiciones rigurosas, tal como se describe a continuación.

La frase "que se hibrida específicamente con" se refiere a la unión, formación de dúplex o hibridación de una molécula sólo a una secuencia de nucleótidos particular en condiciones rigurosas cuando esa secuencia está presente en un ADN o ARN de mezcla compleja (por ejemplo, celular total).

La expresión "condiciones rigurosas" se refiere a condiciones en las que una sonda se hibridará con su subsecuencia diana, pero no con otras secuencias. Las condiciones rigurosas dependen de la secuencia y serán diferentes en diferentes circunstancias. Las secuencias más largas se hibridan específicamente a mayores temperaturas. Generalmente, se seleccionan las condiciones rigurosas para que sean aproximadamente 15°C menores que el punto de fusión térmico (Tf) para la secuencia específica a una fuerza iónica y pH definidos. La Tf es la temperatura (con fuerza iónica, pH y concentración de ácido nucleico definidos) a la que el 50% de las sondas complementarias con la secuencia diana se hibridará con la secuencia diana en el equilibrio. (Como las secuencias diana están presentes generalmente en exceso, a Tf, el 50% de las sondas está ocupado en el equilibrio). Normalmente, condiciones rigurosas serán aquellas en las que la concentración salina es menor que aproximadamente 1,0 M de concentración de ión Na, normalmente aproximadamente de 0,01 a 1,0 M de concentración de ión Na (u otras sales) a pH de 7,0 a 8,3 y la temperatura es al menos de aproximadamente 30°C para sondas cortas (por ejemplo, de 10 a 50 nucleótidos) y al menos de aproximadamente 60°C para sondas largas (por ejemplo, de más de 50 nucleótidos). También pueden lograrse condiciones rigurosas con la adición de agentes desestabilizantes tales como formamida. Para la hibridación específica o selectiva, una señal positiva es normalmente de al menos dos veces el fondo, preferiblemente 10 veces la hibridación de fondo. A modo de ejemplo condiciones de hibridación rigurosas pueden ser como las siguientes: formamida al 50%, SSC 5x, y SDS al 1%, incubación a 42°C, o, SSC 5x, SDS al 1%, incubación a 65°C, con lavado con SSC 0,2x, y SDS al 0,1% a 65°C. Para PCR, una temperatura de aproximadamente 36°C es típica para la amplificación de baja rigurosidad, aunque las temperaturas de apareamiento pueden variar entre aproximadamente 32-48°C dependiendo de la longitud de cebadores. Para amplificación por PCR de alta rigurosidad, una temperatura de aproximadamente 62°C es típica, aunque las temperaturas de apareamiento de alta rigurosidad pueden oscilar entre aproximadamente 50°C y aproximadamente 65°C, dependiendo de la especificidad y longitud de cebadores. Las condiciones de ciclo típicas

para amplificaciones tanto de alta como de baja rigurosidad incluyen una fase de desnaturalización de 90-95°C durante 30-120 s, una fase de apareamiento que dura 30-120 s y una fase de extensión de aproximadamente 72°C durante 1-2 min. Están disponibles protocolos y directrices para reacciones de amplificación de baja y alta rigurosidad, por ejemplo, en Innis, *et al.* (1990) PCR Protocols: A Guide to Methods and Applications Academic Press, N. Y.

Las frases “se une específicamente a una proteína” o “específicamente inmunoreactivo con”, cuando se hace referencia a un anticuerpo se refiere a una reacción de unión que es determinativa de la presencia de la proteína en presencia de una población heterogénea de proteínas y otros compuestos biológicos. Por tanto, en condiciones de inmunoensayo diseñadas, los anticuerpos especificados se unen preferentemente a una proteína particular y no se unen en una cantidad significativa a otras proteínas presentes en la muestra. La unión específica a una proteína en tales condiciones requiere de un anticuerpo que se selecciona por su especificidad por una proteína particular. Puede usarse una variedad de formatos de inmunoensayo para seleccionar anticuerpos específicamente inmunoreactivos con una proteína particular. Por ejemplo, se usan rutinariamente inmunoensayos ELISA en fase sólida para seleccionar anticuerpos monoclonales específicamente inmunoreactivos con una proteína. Véase Harlow y Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Publicaciones, Nueva York, para una descripción de formatos de inmunoensayo y condiciones que pueden usarse para determinar inmunoreactividad específica.

“Variaciones modificadas de manera conservativa” de una secuencia de polinucleótido particular se refiere a aquellos polinucleótidos que codifican para secuencias de aminoácidos idénticas o esencialmente idénticas, o en las que el polinucleótido no codifica para una secuencia de aminoácidos, con respecto a secuencias esencialmente idénticas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican para cualquier proteína dada. Por ejemplo, los codones CGU, CGC, CGA, CGG, AGA y AGG codifican todos para el aminoácido arginina. Por tanto, en cualquier posición en la que está especificada una arginina por un codón, el codón puede alterarse a cualquiera de los correspondientes codones descritos sin alterar la proteína codificada. Tales variaciones de ácido nucleico son “variaciones silenciosas”, que son una especie de “variaciones modificadas de manera conservativa”. Cada secuencia de polinucleótido descrita en el presente documento que codifica para una proteína también describe cada posible variación silenciosa, excepto que se indique de otro modo. Un experto reconocerá que cada codón en un ácido nucleico (excepto AUG, que es comúnmente el único codón para metionina, y UGG que es comúnmente el único codón para triptófano) puede modificarse para producir una molécula funcionalmente idéntica mediante técnicas convencionales. Por consiguiente, cada “variación silenciosa” de un ácido nucleico que codifica para una proteína está implícita en cada secuencia descrita.

Además, un experto reconocerá que las sustituciones, deleciones o adiciones individuales que alteran, añaden o delecionan un solo aminoácido o un pequeño porcentaje de aminoácidos (normalmente menor del 5%, más normalmente menos del 1%) en una secuencia codificada son “variaciones modificadas de manera conservativa” en las que las alteraciones dan como resultado la sustitución de un aminoácido por un aminoácido químicamente similar. Las tablas de sustituciones conservativas que proporcionan aminoácidos funcionalmente similares se conocen bien en la técnica.

Un experto apreciará que muchas variaciones conservativas de proteínas, por ejemplo, glicosiltransferasas, y ácido nucleico que codifican para proteínas producen productos esencialmente idénticos. Por ejemplo, debido a la degeneración del código genético, las “sustituciones silenciosas” (es decir, sustituciones de una secuencia de ácido nucleico que no dan como resultado una alteración en una proteína codificada) son una característica implícita de cada secuencia de ácido nucleico que codifica para un aminoácido. Tal como se describe en el presente documento, las secuencias se optimizan preferiblemente para la expresión en una célula huésped particular usada para producir las glicosiltransferasas quiméricas (por ejemplo, levadura, humana, y similares). De manera similar, las “sustituciones de aminoácidos conservativas”, en uno o unos cuantos aminoácidos en una secuencia de aminoácidos se sustituyen por diferentes aminoácidos con propiedades altamente similares (véase, la sección de definiciones, citada anteriormente), también se identifican fácilmente como que son altamente similares a una secuencia de aminoácidos particular, o a una secuencia de ácido nucleico particular que codifica para un aminoácido. Tales variaciones por sustitución conservativa de cualquier secuencia particular son una característica de la presente invención. Véase también, Creighton (1984) Proteins, W. H. Freeman and Company. Además, las sustituciones, deleciones o adiciones individuales que alteran, añaden o delecionan un solo aminoácido o un pequeño porcentaje de aminoácidos en una secuencia codificada son también “variaciones modificadas de manera conservativa”.

La práctica de esta invención puede implicar la construcción de ácidos nucleicos recombinantes y la expresión de genes en células huésped, preferiblemente células huésped bacterianas. En la técnica, se conocen técnicas de clonación molecular para lograr estos fines. Los expertos conocen bien una amplia variedad de métodos de clonación y amplificación *in vitro* adecuados para la construcción de ácidos nucleicos recombinantes tales como vectores de expresión. Se encuentran ejemplos de estas técnicas e instrucciones suficientes para dirigir a los expertos a través de muchos ejercicios de clonación en Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volumen 152 Academic Press, Inc., San Diego, CA (Berger); 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., (1999, Suplemento) (Ausubel). Los expertos en la técnica conocen células huésped adecuadas para la expresión de los polipéptidos recombinantes, e incluyen, por ejemplo, células procariontas, tales como *E. coli*, y células eucariotas incluyendo células de insecto, de mamífero y fúngicas (por ejemplo, *Aspergillus niger*)

Se encuentran ejemplos de protocolos suficientes para dirigir a los expertos a través de los métodos de amplificación *in vitro*, incluyendo la reacción en cadena de la polimerasa (PCR) la reacción en cadena de la ligasa (LCR), amplificación de la replicasa Q β y otras técnicas mediadas por ARN polimerasa en Berger, Sambrook y Ausubel, así como en Mullis *et al.* (1987) patente estadounidense n.º 4.683.202; PCR Protocols A Guide to Methods and Applications (knis *et al.* eds) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (1 de octubre de 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3: 81-94; (Kwoh *et al.* (1989) Proc. Natl. Acad. Sci. USA 86: 1173; Guatelli *et al.* (1990) Proc. Natl. Acad. Sci. USA 87: 1874; Lomell *et al.* (1989) J. Clin. Chem. 35: 1826; Landegren *et al.* (1988) Science 241: 1077-1080; Van Brunt (1990) Biotechnology 8: 291-294; Wu y Wallace (1989) Gen 4: 560; y Barringer *et al.* (1990) Gen 89: 117. Se describen métodos de clonación *in vitro* mejorados de ácidos nucleicos amplificados en Wallace *et al.*, patente estadounidense n.º 5.426.039.

Descripción detallada de la invención

I. Introducción

La presente invención proporciona condiciones para el replegamiento de glicosiltransferasas eucariotas que se expresan como proteínas insolubles en cuerpos de inclusión bacterianos. Se usan tampones de replegamiento que comprenden pares redox para mejorar el replegamiento de glicosiltransferasas eucariotas insolubles. También puede mejorarse el replegamiento fusionando un dominio de unión a maltosa a la glicosiltransferasa insoluble eucariota. Para algunas glicosiltransferasas eucariotas insolubles, también puede mejorarse el replegamiento mediante mutagénesis dirigida al sitio para eliminar cisteínas desapareadas. Puede proporcionarse una mejora adicional del replegamiento mediante el truncamiento de una glicosiltransferasa eucariota para eliminar, por ejemplo, un dominio de anclaje de señales, un dominio transmembrana, y/o la totalidad o una parte de una región de tallo de la proteína. La invención también proporciona métodos para replegar más de una glicosiltransferasa en un único recipiente, mejorando de ese modo el replegamiento de las proteínas y aumentando la eficacia de la producción de proteínas. Las glicosiltransferasas eucariotas replegadas pueden usarse para producir o para remodelar polisacáridos, oligosacáridos, glicolípidos, proteínas, péptidos, glicopéptidos y glicoproteínas. Las glicosiltransferasas eucariotas replegadas también pueden usarse para glicopegilar proteínas, péptidos, glicopéptidos o glicoproteínas tal como se describe en el documento PCT/US02/32263.

II. Replegamiento de glicosiltransferasas insolubles

Muchas proteínas recombinantes expresadas en bacterias se expresan como agregados insolubles en cuerpos de inclusión bacterianos. Los cuerpos de inclusión son depósitos de proteína que se encuentran tanto en el espacio citoplasmático como periplasmático de bacterias. (Véase, por ejemplo, Clark, Cur. Op. Biotech. 12: 202-207 (2001)). Las glicosiltransferasas eucariotas se expresan frecuentemente en cuerpos de inclusión bacterianos. Algunas glicosiltransferasas eucariotas son solubles en bacterias, es decir, no se producen en cuerpos de inclusión, cuando sólo se expresa el dominio catalítico de la proteína. Sin embargo, muchas glicosiltransferasas eucariotas siguen siendo insolubles y se expresan en cuerpos de inclusión bacterianos, aunque sólo se exprese el dominio catalítico, y se proporcionan métodos para el replegamiento de estas proteínas para producir glicosiltransferasas activas en el presente documento.

A. Condiciones para el replegamiento de glicosiltransferasas activas

Para producir glicosiltransferasas eucariotas activas a partir de células bacterianas, se expresan glicosiltransferasas eucariotas en cuerpos de inclusión bacterianos, se recogen las bacterias, se rompen y los cuerpos de inclusión se aíslan y se lavan. Luego se solubilizan las proteínas dentro de los cuerpos de inclusión. Puede realizarse la solubilización usando desnaturizantes, por ejemplo, cloruro de guanidinio o urea; extremos de pH, tales como condiciones ácidas o alcalinas; o detergentes.

Tras la solubilización, se eliminan los desnaturizantes de la mezcla de glicosiltransferasas. La eliminación de los desnaturizantes puede realizarse mediante una variedad de métodos, incluyendo dilución en un tampón de replegamiento o métodos de intercambio de tampón. Los métodos de intercambio de tampón incluyen diálisis, diafiltración, filtración en gel e inmovilización de la proteína sobre un soporte sólido. (Véase, por ejemplo, Clark, Cur. Op. Biotech. 12: 202-207 (2001)). Cualquiera de los métodos anteriores puede combinarse para eliminar desnaturizantes.

La formación de enlaces disulfuro en la glicosiltransferasa eucariota está promovida por la adición de un tampón de replegamiento que comprende un par redox. Los pares redox incluyen glutatión reducido y oxidado (GSH/GSSG), cisteína/cistina, cisteamina/cistamina, DTT/GSSG y DTE/GSSG. (Véase, por ejemplo, Clark, Cur. Op. Biotech. 12: 202-207 (2001), que se incorpora al presente documento como referencia para todos los fines). En algunas

realizaciones, se añaden pares redox a una razón particular de componente reducido con respecto al oxidado, por ejemplo, de 1/20, 20/1, 1/4, 4/1, 1/10, 10/1, 1/2, 2/1, 1/5, 5/1 ó 5/5.

5 Puede realizarse el replegamiento en tampones a pH que oscilan entre, por ejemplo, 6,0 y 10,0. Los tampones de replegamiento pueden incluir otros aditivos para mejorar el replegamiento, por ejemplo, L-arginina (0,4-1 M); PEG; bajas concentraciones de desnaturalizantes, tales como urea (1-2 M) y cloruro de guanidinio (0,5-1,5 M); y detergentes (por ejemplo, Chaps, SDS, CTAB, lauril-maltósido y Triton X-100).

El replegamiento puede ser durante un periodo de tiempo dado, por ejemplo, durante 1-48 horas, o durante la noche. El replegamiento puede realizarse desde aproximadamente 4°C hasta aproximadamente 40°C, incluyendo temperaturas ambientales.

10 Una proteína glicosiltransferasa eucariota que comprende un dominio catalítico se expresa en cuerpos de inclusión bacterianos y luego se repliega usando los métodos anteriores. También pueden usarse glicosiltransferasas eucariotas que comprenden la totalidad o una parte de una región de tallo y un dominio catalítico en los métodos descritos en el presente documento, como pueden usarse glicosiltransferasas eucariotas que comprenden un dominio catalítico fusionado a una proteína MBP.

15 Los expertos reconocerán que una proteína se ha replegado correctamente cuando la proteína replegada tiene actividad biológica detectable. Para una glicosiltransferasa, actividad biológica es la capacidad para catalizar la transferencia de un sustrato donador a un sustrato aceptor, por ejemplo, una ST3GalIII replegada puede transferir ácido siálico a un sustrato aceptor. La actividad biológica incluye, por ejemplo, actividades específicas de al menos 1, 2, 5, 7 ó 10 unidades de actividad. Unidad se define tal como sigue: una unidad de actividad cataliza la formación de 1 μ mol de producto por minuto a una temperatura dada (por ejemplo, a 37°C) y valor de pH (por ejemplo, a pH 7,5). Por tanto, 10 unidades de una enzima es una cantidad catalítica de esa enzima en la que 10 μ mol de sustrato se convierten en 10 μ mol de producto en un minuto a una temperatura de, por ejemplo, 37°C y un valor de pH de, por ejemplo, 7,5.

25 En una realización, se expresa ST3GalIII eucariota en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

En una realización, se expresa GnT1 eucariota en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

En una realización, se expresa GalT1 eucariota en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

30 En una realización, se expresa ST3Gal eucariota en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

En una realización, se expresa St6 GalNAcT1 eucariota en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

35 En una realización, se expresa Core GalT1 eucariota en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

En una realización, se expresa GalNAcT2 eucariota en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

B.Fusión de glicosiltransferasas eucariotas a dominios de proteína de unión a maltosa para mejorar el replegamiento

40 Normalmente se fusionan dominios de proteína de unión a maltosa (MBP) a proteínas para mejorar la solubilidad de la proteína con una célula. Véase, por ejemplo, Kapust y Waugh Pro. Sci. 8: 1668-1674 (1999). Sin embargo, muchas glicosiltransferasas eucariotas, incluyendo glicosiltransferasas eucariotas truncadas, siguen siendo insolubles cuando se expresan en bacterias, incluso tras fusión a un dominio de MBP. Sin embargo, esta solicitud da a conocer que los dominios de MBP pueden mejorar el replegamiento de glicosiltransferasas eucariotas insolubles tras la solubilización de las proteínas a partir de, por ejemplo, un cuerpo de inclusión. Pueden usarse dominios de MBP de una variedad de fuentes bacterianas en la invención, por ejemplo *Yersinia E. coli*, *Pyrococcus furiosus*, *Thermococcus litoralis*, *Thermatoga maritime* y *Vibrio cholerae*, véase, por ejemplo, la figura 43. En una realización preferida se fusiona una proteína MBP de *E. coli* a una glicosiltransferasa eucariota. Pueden situarse ligadores de aminoácido entre el dominio de MBP y la glicosiltransferasa. En otra realización preferida, el dominio de MBP se fusiona en el extremo amino-terminal de la proteína glicosiltransferasa. Los métodos descritos anteriormente pueden usarse para replegar las proteínas de fusión de MBP-glicosiltransferasa.

50 En una realización, se fusiona una proteína ST3GalIII eucariota a un dominio de MBP, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

En una realización, se fusiona una proteína GnT1 eucariota a un dominio de MBP, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

5 En una realización, se fusiona una proteína GalT1 eucariota a un dominio de MBP, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

En una realización, se fusiona una proteína ST3Gall eucariota a un dominio de MBP, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

10 En una realización, se fusiona una proteína St6 GalNAcTI eucariota a un dominio de MBP, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

15 En una realización, se fusiona una proteína Core GalTI eucariota a un dominio de MBP, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

En una realización, se fusiona una proteína GalNAcT2 eucariota a un dominio de MBP, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

20 Pueden añadirse colas de aminoácidos adicionales a una fusión MBP-glicosiltransferasa. Por ejemplo, pueden añadirse colas de purificación para mejorar la purificación de la proteína replegada. Las colas de purificación incluyen, por ejemplo, una cola de polihistidina, una glutatión S transferasa (GST), una proteína de unión a almidón (SBP), un dominio de tiorredoxina de *E. coli*, una mitad carboxi-terminal de la proteína SUMO, un epítipo FLAG, y un epítipo myc. Las glicosiltransferasas replegadas pueden purificarse adicionalmente usando una pareja de unión que se une a la cola de purificación. En una realización preferida, se fusiona una cola de MBP a la glicosiltransferasa eucariota para mejorar el replegamiento. Pueden fusionarse colas de purificación a proteína de fusión de MBP-glicosiltransferasa, incluyendo por ejemplo, GnT1, GalT1, StIII Gal3, ST3Gall, St6 GalNAcTI, Core GalTI o GalNAcT2.

30 En otra realización, la adición de un dominio de MBP a una proteína puede aumentar la expresión de la proteína. Véase, por ejemplo, el ejemplo 12 en el que la fusión de la proteína SiaA a un dominio de MBP aumentó la expresión de la proteína. Otras proteínas con expresión mejorada con la fusión a MBP incluyen por ejemplo, GnT1, GalT1, StIII Gal3, ST3Gall, St6 GalNAcTI, Core GalTI o GalNAcT2.

En otra realización, una cola de proteína de autoescisión, tal como una inteína, está incluida en el dominio de MBP y la glicosiltransferasa para facilitar la eliminación del dominio de MBP tras replegarse la proteína de fusión. Las inteínas y kits para su uso están disponibles comercialmente, por ejemplo, de New England Biolabs.

35 C.Mutagénesis de glicosiltransferasas para mejorar el replegamiento

El replegamiento de glicosiltransferasas también puede mejorarse mediante mutagénesis de la secuencia de aminoácidos de la glicosiltransferasa. En una realización, se identifica y se muta un residuo de cisteína desapareado para mejorar el replegamiento de una glicosiltransferasa. En otra realización, se trunca el extremo amino-terminal de la glicosiltransferasa para eliminar un dominio transmembrana, o para eliminar un dominio transmembrana y la totalidad o una parte de la región de tallo de la proteína. En una realización adicional, se muta una glicosiltransferasa para eliminar al menos un residuo de cisteína desapareado y para truncar el extremo amino-terminal de la proteína, por ejemplo, para eliminar un dominio transmembrana, o para eliminar un dominio transmembrana y la totalidad o una parte de la región de tallo de la proteína. Una vez que se ha aislado una secuencia de ácido nucleico de la glicosiltransferasa, pueden usarse métodos de biología molecular convencionales para cambiar la secuencia de ácido nucleico y por tanto la secuencia de aminoácidos codificada de una manera descrita en el presente documento.

40 1.Mutagénesis de cisteínas desapareadas en glicosiltransferasas para mejorar el replegamiento

A medida que se produce el replegamiento, residuos de cisteína en una proteína desnaturalizada forman enlaces disulfuro que ayudan a reproducir la estructura de la proteína activa. El apareamiento incorrecto de residuos de cisteína puede conducir a un plegamiento incorrecto de la proteína. Las proteínas con residuos de cisteína desapareados son susceptibles de plegamiento incorrecto debido a que una cisteína desapareada normalmente puede formar un enlace disulfuro con una cisteína apareada normalmente haciendo que sean imposible un apareamiento de cisteínas y replegamiento de proteína correctos. Por tanto, un método para mejorar el replegamiento de una glicosiltransferasa particular es identificar residuos de cisteína desapareados y eliminarlos.

Pueden identificarse residuos de cisteína desapareados determinando la estructura de la glicosiltransferasa de interés. Puede determinarse la estructura de la proteína basándose en datos reales para la glicosiltransferasa de interés, por ejemplo, dicroísmo circular, RMN y cristalografía de rayos X. También puede determinarse la estructura de la proteína usando modelización por ordenador. La modelización por ordenador es una técnica que puede usarse para modelizar estructuras relacionadas basándose en estructuras tridimensionales conocidas de moléculas homólogas. Está disponible comercialmente el software convencional. (Véase por ejemplo, www.accelrys.com para la multitud de software disponible para realizar la modelización por ordenador). Una vez que se identifica un residuo de cisteína desapareado, puede mutarse el ADN que codifica para la glicosiltransferasa de interés usando técnicas de biología molecular convencionales para eliminar la cisteína desapareada, mediante delección o mediante sustitución por otro residuo de aminoácido. Se usa de nuevo la modelización por ordenador para seleccionar un aminoácido de tamaño, forma y carga apropiados para la sustitución. También pueden determinarse las cisteínas desapareadas mediante mapeo peptídico. Una vez que se muta la glicosiltransferasa de interés, la proteína se expresa en cuerpos de inclusión bacterianos y se determina la capacidad de replegamiento. Una glicosiltransferasa replegada correctamente tendrá actividad biológica.

En realizaciones preferidas, los siguientes residuos de aminoácido se sustituyen por un residuo de cisteína desapareado en una glicosiltransferasa eucariota para mejorar el replegamiento de: Ala, Ser, Thr, Asp, Ile o Val. Gly también puede usarse si la cisteína desapareada no está en una estructura helicoidal.

La N-acetilglucosaminiltransferasa I (GnTI, número de registro NP_002397) humana es un ejemplo de una glicosiltransferasa que mostró un replegamiento mejorado tras mutagénesis de una cisteína desapareada. (Véase, por ejemplo, el ejemplo 2, a continuación). La GNTI humana está estrechamente relacionada con varias proteínas GNTI eucariotas, por ejemplo, hámster chino, número de registro AAK61868; conejo número de registro AAA31493; rata número de registro NP_110488; hámster dorado, número de registro AAD04130; ratón, número de registro P27808; pez cebra, número de registro AAH58297; *Xenopus*, número de registro CAC51119; *Drosophila*, número de registro NP_525117; *Anopheles*, número de registro XP 315359; *C. elegans*, número de registro NP_497719; *Physcomitrella patens*, número de registro CAD22107; *Solanum tuberosum*, número de registro CAC80697; *Nicotiana tabacum*, número de registro CAC80702; *Oryza sativa*, número de registro CAD30022; *Nicotiana benthamiana*, número de registro CAC82507; y *Arabidopsis thaliana*, número de registro NP_195537.

Se había determinado estructura de la proteína N-acetilglucosaminiltransferasa I (GnTI) de conejo y mostró que CYS123 estaba desapareada. (Los números de residuos de aminoácido se refieren a la secuencia de proteína de longitud completa incluso cuando se ha truncado una proteína GNTI). Se usó modelización por ordenador basada en la GnTI de conejo para determinar la estructura de la proteína GnTI humana. Se muestra una alineación en la figura 6. En la proteína GnTI humana, CYS121 estaba desapareada. Se realizaron sustituciones para CYS121 en GnTI humana. Un mutante de CYS121SER y un mutante de CYS121ALA eran activos. En cambio, un mutante CYS121THR no tenía actividad detectable y un mutante CYS121ASP tenía baja actividad. Se construyó un doble mutante, ARG120ALA, CYS121HIS, basándose en la estructura predicha de la proteína GNT1 de *C. elegans*, y tenía actividad.

Las secuencias de aminoácidos de las proteínas GnTI eucariotas enumeradas anteriormente pueden usarse para determinar la estructura de la proteína basándose en modelización por ordenador y la función conservada de CYS123 de conejo y CYS121 de ser humano. Basándose en ese análisis, el residuo 123 es una cisteína desapareada en las siguientes proteínas: GnTI de hámster chino, la GnTI de conejo, la GnTI de rata, la GnTI de hámster dorado y la GnTI de ratón. Por tanto, CYS123 puede mutarse en cada una de las enzimas GnTI a serina, alanina o arginina para producir una proteína activa con una actividad de replegamiento mejorada. El siguiente doble mutante en las proteínas anteriores, ARG122ALA CYS123HIS, también mostrará replegamiento mejorado.

En una realización, se muta cualquiera de las proteínas GnTI eucariotas enumeradas anteriormente para eliminar un residuo de cisteína desapareado, por ejemplo, CYS121SER, CYS121ALA, CYS121ASP o el doble mutante ARG120ALA CYS121HIS, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

Otra glicosiltransferasa que muestra un replegamiento mejorado con la mutación de una cisteína desapareada es Gal T1. Se mutó la cisteína 342 a un residuo de treonina en la Gal T1 bovina y la enzima mutada mostró un replegamiento mejorado tras la solubilización. Véase, por ejemplo, Ramakrishnan *et al.* J. Biol. Chem. 276: 37666-37671 (2001). De manera interesante, la mutación de una cisteína desapareada a treonina en la enzima GnTI, suprimió la actividad.

En una realización, se muta una proteína Gal T1 para eliminar un residuo de cisteína desapareado, por ejemplo, CYS342THR, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.

Otra glicosiltransferasa que muestra un replegamiento mejorado con la mutación de una cisteína desapareada es GalNAc T2. Muchos residuos de aminoácido se comparten entre la proteína GalNAc T2 y la proteína GalNAc T1. Tras la mutación de dos residuos de cisteína, CYS212 y CYS214, la proteína GalNAc T1 humana siguió siendo activa cuando se expresó a partir de células COS. Véase, por ejemplo, Tenno *et al.*, Eur. J. Biochem. 269: 4308-

- 4316 (2002). Las mutaciones activas incluyeron CYS212ALA, CYS214ALA, CYS212SER, CYS214SER, y un doble mutante CYS212SER CYS214SER. Los residuos de cisteína correspondientes a los residuos CYS212 y CYS214 de GalNAc T1 humana se conservan en la proteína GalNAc T2 humana, es decir, los residuos CYS227 y CYS229. Véase, por ejemplo, la figura 44. Por tanto, puede usarse una proteína GalNAc T2 que comprende una de las siguientes mutaciones para mejorar el repliegamiento de la proteína insoluble: CYS227ALA, CYS229ALA, CYS227SER, CYS229SER, y un doble mutante CYS227SER CYS229SER. La numeración de residuos se refiere a proteínas GalNAc T2 humanas, pero pueden identificarse residuos de cisteína conservados que corresponden a CYS227 y CYS229 en otras proteínas GalNAc T2 eucariotas, por ejemplo, de ratón, rata, conejo, cerdo, y mutados para mejorar el repliegamiento.
- 5
- 10 En una realización, se muta una proteína GalNAc T2 para eliminar un residuo de cisteína desapareado, por ejemplo, CYS227ALA, CYS229ALA, CYS227SER, CYS229SER o un doble mutante CYS227SER CYS229SER, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína.
- 15 Otra glicosiltransferasa que muestra un repliegamiento mejorado en mutación de una cisteína desapareada es Core 1 Gal T1. En una realización, se muta la Core 1 Gal T1 de *Drosophila*. En otra realización, se muta la Core 1 Gal T1 humana. La proteína Core 1 Gal T1 de *Drosophila* tiene siete residuos de cisteína. Cada residuo de cisteína se muta individualmente o bien a serina o bien a alanina. Las proteínas Core 1 Gal T1 de *Drosophila* mutadas se expresan en cuerpos de inclusión de *E. coli*, se solubilizan y se repliegan. Se somete a ensayo la actividad enzimática de la Core 1 Gal T1 de *Drosophila* mutante repliegada y se compara con la actividad de la Core 1 Gal T1 de *Drosophila* repliegada de tipo natural. Un repliegamiento mejorado está indicado por un aumento de la actividad enzimática en una Core 1 Gal T1 de *Drosophila* mutante en comparación con la proteína de tipo natural.
- 20
- 25 En una realización, se muta una proteína Core 1 Gal T1 para eliminar un residuo de cisteína desapareado, por ejemplo, o bien la proteína de *Drosophila* o bien la proteína humana, se expresa en cuerpos de inclusión bacterianos, se solubiliza y se repliega en un tampón que comprende un par redox, por ejemplo, GSH/GSSG o cistamina/cisteína. Residuos de cisteína preferidos para la sustitución en la Core 1 Gal T1 de *Drosophila* son C103, C127, C208, C246, C261, C315 y C316.
2. Truncamiento de glicosiltransferasas para mejorar el repliegamiento
- 30 Las glicosiltransferasas eucariotas incluyen generalmente los siguientes dominios: un dominio catalítico, una región de tallo, un dominio transmembrana y un dominio de anclaje de señales. Cuando se expresa en bacterias, normalmente se delecionan el dominio de anclaje de señales y los dominios transmembrana. Las glicosiltransferasas eucariotas usadas en los métodos de la invención pueden incluir la totalidad o una parte de la región de tallo y el dominio catalítico. En algunas realizaciones, las glicosiltransferasas eucariotas comprenden sólo el dominio catalítico.
- 35 Pueden identificarse dominios de glicosiltransferasa para mutagénesis de delección. Por ejemplo, los expertos en la técnica pueden identificar una región de tallo en una glicosiltransferasa eucariota y deleccionar los aminoácidos de la región de tallo uno a uno para identificar proteínas glicosiltransferasas eucariotas truncadas con alta actividad en el repliegamiento.
- 40 En esta solicitud se hace referencia a los mutantes de delección de dos maneras: Δ o D seguido por el número de residuos deleccionados del extremo amino-terminal de la secuencia de aminoácidos de longitud completa nativa, o mediante el símbolo y el número de residuo del primer residuo de aminoácido traducido de la secuencia de aminoácidos de longitud completa nativa. Por ejemplo, ST6GalNAc Δ o D35 y ST6GalNAc K36, se refieren ambos al mismo truncamiento de la proteína ST6GalNAc.
- 45 Por ejemplo, la proteína ST3GalIII de rata incluye una región de tallo desde aproximadamente los residuos de aminoácido 29-84. El dominio catalítico de la proteína comprende aminoácidos desde aproximadamente el residuo 85-374. Por tanto, una proteína ST3GalIII de rata truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84 ó 85 residuos.
- 50 También pueden realizarse mutaciones de delección en una proteína GnT1. Por ejemplo, la proteína GnT1 humana incluye una región de tallo desde aproximadamente los residuos de aminoácido 31-112. Por tanto, una proteína GnT1 humana truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110 ó 111 residuos.
- 55 También pueden realizarse mutaciones de delección en una proteína Gal T1. Por ejemplo, la proteína GalT1 bovina incluye una región de tallo desde aproximadamente los residuos de aminoácido 71-129. Por tanto, una proteína GalT1 bovina truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 70,

71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 017, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127 ó 128 residuos.

5 También pueden realizarse mutaciones de delección en una proteína Core1 GalT1. Por ejemplo, la proteína Core1 GalT1 de *Drosophila* incluye una región de tallo desde aproximadamente los residuos de aminoácido 36-102. Por tanto, una proteína Core1 GalT1 de *Drosophila* truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101 ó 102 residuos. Como otro ejemplo, la proteína Core1 GalT1 humana incluye una región de tallo desde aproximadamente los residuos de aminoácido 32-90. Por tanto, una proteína Core1 GalT1 humana truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89 ó 90 residuos.

15 También pueden realizarse mutaciones de delección en una proteína ST3Gal1. Por ejemplo, la proteína ST3Gal1 humana incluye una región de tallo desde aproximadamente los residuos de aminoácido 18-58. Por tanto, una proteína ST3Gal1 humana truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57 ó 58 residuos. Como otro ejemplo, la proteína ST3Gal1 porcina incluye una región de tallo desde aproximadamente los residuos de aminoácido 28-61. Por tanto, una proteína ST3Gal1 porcina truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60 ó 61 residuos.

20 También pueden realizarse mutaciones de delección en una proteína GalNAcT2. Por ejemplo, la proteína GalNAcT2 de rata incluye una región de tallo desde aproximadamente los residuos de aminoácido 40-95. Por tanto, una proteína GalNAcT2 de rata truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94 ó 95 residuos.

25 También pueden realizarse mutaciones de delección en una proteína ST6GalNAcI. Por ejemplo, la proteína ST6GalNAcI de ratón incluye una región de tallo desde aproximadamente los residuos de aminoácido 30-207. Por tanto, una proteína ST6GalNAcI de ratón truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 017, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206 ó 207 residuos. Como otro ejemplo, la proteína ST6GalNAcI incluye una región de tallo desde aproximadamente los residuos de aminoácido 35-278. Por tanto, una proteína ST6GalNAcI humana truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 017, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277 ó 278 residuos. Todavía como otro ejemplo, la proteína ST6GalNAcI de gallina incluye una región de tallo desde aproximadamente los residuos de aminoácido 37-253. Por tanto, una proteína ST6GalNAcI de gallina truncada puede tener delecciones en el extremo amino-terminal de aproximadamente por ejemplo, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 017, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 156, 157, 158, 159; 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252 ó 253 residuos.

D. Replegamiento en un solo recipiente de glicosiltransferasas

Estas realizaciones de la invención se basan en la sorprendente observación de que múltiples glicosiltransferasas eucariotas expresadas en cuerpos de inclusión bacterianos pueden replegarse en un único recipiente, es decir, un método en un solo recipiente. Usando este método pueden replegarse al menos dos glicosiltransferasas juntas dando como resultado ahorros de tiempo y materiales. Las condiciones de replegamiento se describieron anteriormente. Las condiciones de replegamiento se optimizan para la mezcla de glicosiltransferasas, por tanto, las condiciones pueden no ser óptimas para alguna enzima en particular en la mezcla. Sin embargo, dado que se optimiza el replegamiento para la combinación de glicosiltransferasas, cada una de las glicosiltransferasas replegadas en el producto final tiene actividad biológica detectable. Actividad biológica se refiere a la actividad enzimática de las enzimas replegadas y puede expresarse como actividad específica. La actividad biológica incluye por ejemplo, actividades específicas de al menos 0,1, 0,5, 1,2, 5,7 ó 10 unidades de actividad. Unidad se define tal como sigue: una unidad de actividad cataliza la formación de 1 μmol de producto por minuto a una temperatura dada (por ejemplo, a 37°C) y valor de pH (por ejemplo, a pH 7,5). Por tanto, 10 unidades de una enzima es una cantidad catalítica de esa enzima en la que 10 μmol de sustrato se convierten en 10 μmol de producto en un minuto a una temperatura de, por ejemplo, 37°C y un valor de pH de, por ejemplo, 7,5. La mezcla de reacción que comprende las glicosiltransferasas replegadas puede usarse entonces por ejemplo, para sintetizar oligosacáridos, para sintetizar glicolípidos, para remodelar glicoproteínas y para glicopegar glicoproteínas.

En algunas realizaciones, las glicosiltransferasas pueden solubilizarse individualmente de los cuerpos de inclusión y luego combinarse en condiciones apropiadas para el replegamiento. En otras realizaciones, se combinan cuerpos de inclusión que contienen glicosiltransferasas, se solubilizan, y luego se repliegan en condiciones apropiadas.

Los tampones de replegamiento incluyen normalmente un par redox. Puede realizarse el replegamiento a pH que oscilan entre, por ejemplo, 6,0 y 10,0. Los tampones de replegamiento pueden incluir otros aditivos para mejorar el replegamiento, por ejemplo, L-arginina (0,4-1 M); PEG; bajas concentraciones de desnaturalizantes, tales como urea (1-2 M) y cloruro de guanidinio (0,5-1,5 M); y detergentes (por ejemplo, Chaps, SDS, CTAB y Triton X-100).

En algunas realizaciones, se realiza el replegamiento en un recipiente estacionario, es decir, sin mezclar, agitar, remover o mover de otro modo la mezcla de reacción.

La combinación de enzimas replegadas puede incluir enzimas para construir una estructura de oligosacárido particular. Los expertos podrán identificar glicosiltransferasas apropiadas para la inclusión en la mezcla una vez que se identifica un producto final deseado.

Las mezclas de reacción de enzimas replegadas pueden incluir glicosiltransferasas que se han mutado para mejorar el replegamiento, por ejemplo, la enzima GnTI descrita anteriormente.

En una realización preferida, se repliegan juntas enzimas que realizan etapas de glicosilación unida a N en un único recipiente. Por ejemplo, pueden expresarse N-acetilglucosaminiltransferasa I (GnTI), β -1,4-galactosiltransferasa I (Gal TI) y N-acetil-lactosaminida α -2,3-sialiltransferasa (ST3GalIII) en cuerpos de inclusión bacterianos, solubilizarse y replegarse juntas en un único recipiente. El producto final mostró actividad de las tres proteínas, indicando que estaban todas correctamente replegadas. El replegamiento también se produjo cuando se replegaron juntas GnTI y Gal TI sin ST3GalIII. Los experimentos se describen en detalle en el ejemplo 3.

En otra realización preferida, se logra la glicosilación unida a O de un péptido o proteína usando las glicosiltransferasas replegadas y expresadas de forma bacteriana de esta descripción. Por ejemplo, puede usarse una enzima MBP-GalNAcT2(D51) replegada para añadir GalNAc a polipéptidos. Por ejemplo, el ejemplo 4 proporciona una demostración de que MBP-GalNAcT2(D51) replegada puede usarse para añadir GalNAc a la proteína GCSF. Pueden usarse combinaciones de glicosiltransferasas unidas a O para remodelar por ejemplo, proteínas, péptidos, glicoproteínas o glicopéptidos. Esas combinaciones incluyen por ejemplo, GalNAc-T2 y ST6GalNA1; o GalNAc-T2, core 1 GalT1 y ST3Gal1 o ST3GalT2

III. Glicosiltransferasas

Las glicosiltransferasas de uso en la práctica de la presente invención son glicosiltransferasas eucariotas. Los ejemplos de tales glicosiltransferasas incluyen las descritas en Staudacher, E. (1996) Trends in Glycoscience and Glicotechnology, 8: 391-408, afmb.cnrs-mrs.fr/~pedro/CAZY/gtf.html y www.vei.co.uk/TGN/gt_guide.htm, pero no se limitan a las mismas.

Glicosiltransferasas eucariotas

Algunas glicosiltransferasas eucariotas tienen dominios topológicos en su extremo amino-terminal que no se requieren para la actividad catalítica (véase, la patente estadounidense n.º 5.032.519). De las glicosiltransferasas caracterizadas hasta la fecha, el "dominio citoplasmático", lo más comúnmente es de entre aproximadamente 1 y aproximadamente 10 aminoácidos de longitud, y es el dominio más amino-terminal; el dominio adyacente, denominado el "dominio de anclaje de señales", es generalmente de entre aproximadamente 10-26 aminoácidos de

longitud; adyacente al dominio de anclaje de señales hay una "región de tallo", que es generalmente de entre aproximadamente 20 y aproximadamente 60 aminoácidos de longitud, y se conoce por su función como señal de retención para mantener la glicosiltransferasa en el aparato de Golgi; y en el lado de carboxilo de la región de tallo está el dominio catalítico.

5 Muchas glicosiltransferasas de mamífero se han clonado y expresado y las proteínas recombinantes se han caracterizado en cuanto a la especificidad del sustrato aceptor y donador y también se han investigado a través de mutagénesis dirigida al sitio en intentos por definir residuos o dominios implicados en la especificidad de o bien sustrato aceptor o bien donador (Aoki *et al.* (1990) EMBO. J. 9: 3171-3178; Harduin-Lepers *et al.* (1995) Glycobiology 5 (8): 741-758; Natsuka y Lowe (1994) Current Opinion in Structural Biology 4: 683-691; Zu *et al.* (1995) Biochem. Biophys. Res. Comm. 206 (1): 362-369; Seto *et al.* (1995) Eur. J. Biochem. 234: 323-328; Seto *et al.* (1997) J. Biol. Chem. 272: 14133-141388).

15 En un grupo de realizaciones, se obtiene un dominio funcional de las proteínas glicosiltransferasas recombinantes de la presente invención a partir de una sialiltransferasa conocida. Los ejemplos de sialiltransferasas que son adecuados para su uso en la presente invención incluyen, pero no se limitan a, ST3GalIII, ST3Gal IV, ST3Gal I, ST6Gal I, ST3Gal V, ST6Gal II, ST6GalNAc I, ST6GalNAc II y ST6GalNAc III (la nomenclatura para sialiltransferasas usada en el presente documento es tal como se describe en Tsuji *et al.* (1996) Glycobiology 6: v-xiv). Una α 2,3-sialiltransferasa (EC 2.4.99.6) a modo de ejemplo transfiere ácido siálico a la Gal terminal no reductora de un disacárido o glicósido de Gal β 1 \rightarrow 4GlcNAc. Véanse, Van den Eijnden *et al.*, J. Biol. Chem., 256: 3159 (1981), Weinstein *et al.*, J. Biol. Chem., 257: 13845 (1982) y Wen *et al.*, J. Biol. Chem., 267: 21011 (1992). Otra α 2,3-sialiltransferasa (EC 2.4.99.4) a modo de ejemplo transfiere ácido siálico a la Gal terminal no reductora de un disacárido o glicósido Gal β 1 \rightarrow 3GalNAc. Véanse, Rearick *et al.*, J. Biol. Chem., 254: 4444 (1979) y Gillespie *et al.*, J. Biol. Chem., 267: 21004 (1992). Las enzimas a modo de ejemplo adicionales incluyen Gal- β -1,4-GlcNAc α -2,6-sialiltransferasa (Véase, Kurosawa *et al.* Eur. J. Biochem. 219: 375-381 (1994)). La nomenclatura de sialiltransferasas se describe en Tsuji, S. *et al.* (1996) Glycobiology 6: v-vii.

25 Un ejemplo de una sialiltransferasa que es útil en los métodos reivindicados es ST3GalIII, que también se denomina α (2,3)sialiltransferasa (EC 2.4.99.6). Esta enzima cataliza la transferencia de ácido siálico a la Gal de un glicósido Gal β 1,3-GlcNAc, Gal β 1,3-GalNAc o Gal β 1,4-GlcNAc (véanse, por ejemplo, Wen *et al.* (1992) J. Biol. Chem. 267: 21011; Van den Eijnden *et al.* (1991) J. Biol. Chem. 256: 3159). El ácido siálico se une a una Gal con la formación de una unión α entre los dos sacáridos. Los enlaces (unión) entre los sacáridos son entre la posición 2 de NeuAc y la posición 3 de Gal. Esta enzima particular puede aislarse de hígado de rata (Weinstein *et al.* (1982) J. Biol. Chem. 257: 13845); se conocen las secuencias de ADNc humano (Sasaki *et al.* (1993) J. Biol. Chem. 268: 22782-22787; Kitagawa & Paulson (1994) J. Biol. Chem. 269: 1394-1401) y genómico (Kitagawa *et al.* (1996) J. Biol. Chem. 271: 931-938), facilitando la producción de esta enzima mediante expresión recombinante. En una realización preferida, los métodos de sialilación reivindicados usan una ST3GalIII de rata. Se ha clonado ST3GalIII de rata y se conoce la secuencia. Véase, por ejemplo, Wen *et al.*, J. Biol. Chem. 267: 21011-21019 (1992) y número de registro M97754.

40 En otro grupo de realizaciones, se obtienen un dominio funcional de las proteínas glicosiltransferasas recombinantes de la presente invención a partir de una fucosiltransferasa. Los expertos en la técnica conocen varias fucosiltransferasas. Brevemente, las fucosiltransferasas incluyen cualquiera de las enzimas que transfieren L-fucosa de GDP-fucosa a una posición de hidroxilo de un azúcar aceptor. En algunas realizaciones, por ejemplo, el azúcar aceptor es un grupo Gal β (1 \rightarrow 4)GlcNAc en un glicósido de oligosacárido. Las fucosiltransferasas adecuadas para esta reacción incluyen la Gal β (1 \rightarrow 3,4)GlcNAc α (1 \rightarrow 3,4)fucosiltransferasa conocida (FTIII, E.C. n.º 2.4.1.65) que se obtiene a partir de leche humana (véanse, Palcic, *et al.*, Carbohydrate Res. 190: 1-11 (1989); Prieels, *et al.*, J. Biol. Chem. 256: 10456-10463 (1981); y Nunez, *et al.*, Can. J. Chem. 59: 2086-2095 (1981)) y la Gal β (1 \rightarrow 4)GlcNAc α (1 \rightarrow 3)fucosiltransferasas (FTIV, FTV y FTVI, E.C. n.º 2.4.1.65) y NeuAc α (2,3) β Gal(1 \rightarrow 4) β GlcNAc α (1 \rightarrow 3) fucosiltransferasas (FTVII) que se encuentran en suero humano. También, está disponible la α 1,3-fucosiltransferasa IX (secuencias de nucleótidos de FTIX humana y de ratón) tal como se describe en Kaneko *et al.* (1999) FEBS Lett. 452: 237-242. Además, está disponible una forma recombinante de Gal β (1 \rightarrow 3,4)GlcNAc α (1 \rightarrow 3,4)fucosiltransferasa (véanse, Dumas, *et al.*, Bioorg Med. Letters 1: 425-428 (1991) y Kukowska-Latallo, *et al.*, Genes and Development 4: 1288-1303 (1990)). Otras fucosiltransferasas a modo de ejemplo incluyen α 1,2-fucosiltransferasa (E.C. n.º 2.4.1.69). Puede llevarse a cabo la fucosilación enzimática mediante los métodos descritos en Mollicone, *et al.*, Eur. J. Biochem. 191: 169-176 (1990) o la patente estadounidense n.º 5.374.655.

55 En otro grupo de realizaciones, se obtiene un dominio funcional de las proteínas glicosiltransferasas recombinantes de la presente invención a partir de galactosiltransferasas conocidas. Las galactosiltransferasas a modo de ejemplo incluyen β -1,4-galactosiltransferasa I, α 1,3-galactosiltransferasas (E.C. n.º 2.4.1.151, véanse, por ejemplo, Dabkowski *et al.*, Transplant Proc. 25: 2921 (1993) y Yamamoto *et al.* Nature 345: 229-233 (1990), bovina (GenBank j04989, Joziassse *et al.* (1989) J. Biol. Chem. 264: 14290-14297), murina (GenBank m26925; Larsen *et al.* (1989) Proc. Nat'l. Acad. Sci. USA 86: 8227-8231), porcina (GenBank L36152; Strahan *et al.* (1995) Immunogenetics 41: 101-105)). Otra α 1,3-galactosiltransferasa adecuada es la que está implicada en la síntesis del antígeno B de grupo sanguíneo (EC 2.4.1.37, Yamamoto *et al.* (1990) J. Biol. Chem. 265: 1146-1151 (humana)). También son adecuadas para su uso en las proteínas de fusión de la invención las α 1,4-galactosiltransferasas, que incluyen, por ejemplo, EC

60

2.4.1.90 (LacNAc sintetasa) y EC 2.4.1.22 (lactosa sintetasa) (bovina (D'Agostaro *et al.* (1989) Eur. J. Biochem. 183: 211-217), humana (Masri *et al.* (1988) Biochem. Biophys. Res. Commun. 157: 657-663), murina (Nakazawa *et al.* (1988) J. Biochem. 104: 165-168), así como E.C.2.4.1.38 y la ceramida galactosiltransferasa (EC 2.4.1.45, Stahl *et al.* (1994) J. Neurosci. Res. 38: 234-242). Otras galactosiltransferasas adecuadas incluyen, por ejemplo, α 1,2-galactosiltransferasas (de por ejemplo, *Schizosaccharomyces pombe*, Chapell *et al.* (1994) Mol. Biol. Cell 5: 519-528).

Otras glicosiltransferasas que son útiles en las proteínas recombinantes de fusión de la presente invención se han descrito en detalle, como para la sialiltransferasas, galactosiltransferasas y fucosiltransferasas. En particular, la glicosiltransferasa también puede ser, por ejemplo, una glucosiltransferasa, por ejemplo, Alg8 (Stagljov *et al.*, Proc. Natl. Acad. Sci. USA 91: 5977 (1994)) o Alg5 (Heesen *et al.* Eur. J. Biochem. 224: 71 (1994)), N-acetilgalactosaminiltransferasas tales como, por ejemplo, β (1,3)-N-acetilgalactosaminiltransferasa, β (1,4)-N-acetilgalactosaminiltransferasas (patente estadounidense n.º 5.691.180, Nagata *et al.* J. Biol. Chem. 267: 12082-12089 (1992), y Smith *et al.* J. Biol. Chem. 269: 15162 (1994)) y proteína N-acetilgalactosaminiltransferasa (Homa *et al.* J. Biol. Chem. 268: 12609 (1993)). Las N-acetilglucosaminiltransferasas adecuadas incluyen GnTI (2.4.1.101, Hull *et al.*, BBRC 176: 608 (1991)), GnTII, y GnTIII (Ihara *et al.* J. Biochem. 113: 692 (1993)), GnTV (Shoreiban *et al.* J. Biol. Chem. 268: 15381 (1993)), N-acetilglucosaminiltransferasa unidas a O (Bierhuizen *et al.* Proc. Natl. Acad. Sci. USA 89: 9326 (1992)), N-acetilglucosamina-1-fosfato transferasa (Rajput *et al.* Biochem J. 285: 985 (1992) y hialuronano sintasa. También son de interés las enzimas implicadas en la síntesis de proteoglicanos, tales como, por ejemplo, N-acetilgalactosaminiltransferasa I (EC 2.4.1.174), y enzimas implicadas en la síntesis de sulfato de condroitina, tales como N-acetilgalactosaminiltransferasa II (EC 2.4.1.175). Las manosiltransferasas adecuadas incluyen α (1,2)manosiltransferasa, α (1, 3) manosiltransferasa, β (1, 4) manosiltransferasa, Dol-P-Man sintasa, OChI y PmtI. Las xilosiltransferasas incluyen, por ejemplo, proteína xilosiltransferasa (EC 2.4.2.26).

En algunas realizaciones, se expresan N-acetilgalactosaminiltransferasas eucariotas en bacterias y se repliegan usando los métodos de esta descripción. Se han aislado y caracterizado varias enzimas GalNAcT, por ejemplo, GalNAcT1, número de registro X85018; GalNAcT2, número de registro X85019 (ambas descritas en White *et al.*, J. Biol. Chem. 270: 24156-24165 (1995)); y GalNAcT3, número de registro X92689 (descrita en Bennett *et al.*, J. Biol. Chem. 271: 17006-17012 (1996)).

IV. Ácidos nucleicos

Los expertos en la técnica conocen ácidos nucleicos que codifican para glicosiltransferasas, y métodos de obtención de tales ácidos nucleicos. Pueden clonarse ácidos nucleicos adecuados (por ejemplo, ADNc, genómico o subsecuencias (sondas)), o amplificarse mediante métodos *in vitro* tales como la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR), el sistema de amplificación basado en transcripción (TAS), o el sistema de replicación de secuencia autosostenida (SSR). Los expertos conocen bien una amplia variedad de metodologías de clonación y amplificación *in vitro*. Se encuentran ejemplos de estas técnicas e instrucciones suficientes para dirigir a los expertos a través de muchos ejercicios de clonación en Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook *et al.* (1989) Molecular Cloning - A Laboratory Manual (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, (Sambrook *et al.*); 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., (1994, Complemento) (Ausubel); Cashion *et al.*, patente estadounidense número 5,017,478; y Carr, patente europea n.º 0.246.864.

Puede prepararse ADN que codifica para una glicosiltransferasa, o una subsecuencia de la misma, mediante cualquier método adecuado descrito anteriormente, incluyendo, por ejemplo, clonación y restricción de secuencias apropiadas con enzimas de restricción. En una realización preferida, se aíslan ácidos nucleicos que codifican para glicosiltransferasas mediante métodos de clonación rutinarios. Puede usarse una secuencia de nucleótidos de una glicosiltransferasa tal como se proporciona en, por ejemplo, GenBank u otra base de datos de secuencias (véase anteriormente) para proporcionar sondas que se hibridan específicamente con un gen de glicosiltransferasa en una muestra de ADN genómico, o con un ARNm, que codifica para una glucosiltransferasa, en una muestra de ARN total (por ejemplo, en una transferencia de tipo Southern o Northern). Una vez que se identifica el ácido nucleico diana que codifica para una glicosiltransferasa, puede aislarse según métodos convencionales conocidos por los expertos en la técnica (véanse, por ejemplo, Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, 2ª Ed., Vols. 1-3, Cold Spring Harbor Laboratory; Berger y Kimmel (1987) Methods in Enzymology, Vol. 152: Guide to Molecular Cloning Techniques, San Diego: Academic Press, Inc.; o Ausubel *et al.* (1987) Current Protocols in Molecular Biology, Greene Publishing and Wiley-Interscience, Nueva York). Además, pueden escindirse los ácidos nucleicos aislados con enzimas de restricción para crear ácidos nucleicos que codifican para la glicosiltransferasa de longitud completa, o subsecuencias de la misma, por ejemplo, que contienen subsecuencias que codifican para al menos una subsecuencia de una región de tallo o dominio catalítico de una glicosiltransferasa. Estos fragmentos de enzima de restricción, que codifican para una glicosiltransferasa o subsecuencias de la misma, pueden ligarse luego, por ejemplo, para producir un ácido nucleico que codifica para una proteína de fusión de glicosiltransferasa recombinante.

Puede caracterizarse un ácido nucleico que codifica para una glicosiltransferasa, o una subsecuencia de la misma, sometiendo a ensayo el producto expresado. Pueden usarse ensayos basados en la detección de las propiedades físicas, químicas o inmunológicas de la proteína expresada. Por ejemplo, puede identificarse una glicosiltransferasa clonada, incluyendo una proteína de fusión de glicosiltransferasa, mediante la capacidad de una proteína codificada por el ácido nucleico para catalizar la transferencia de un sacárido de un sustrato donador a un sustrato aceptor. En un método preferido, se emplea electroforesis capilar para detectar los productos de reacción. Este ensayo sumamente sensible implica usar derivados de aminofenilo de o bien un sacárido o bien un disacárido que se marcan con fluoresceína tal como se describe en Wakarchuk *et al.* (1996) *J. Biol. Chem.* 271 (45): 28271-276. Por ejemplo, para someter a ensayo una enzima IgtC de *Neisseria* IgtC, pueden usarse o bien FCHASE-AP-Lac o bien FCHASE-AP-Gal, mientras que para la enzima IgtB de *Neisseria* un reactivo apropiado es FCHASE-AP-GlcNAc (íd.).

Además, puede sintetizarse químicamente un ácido nucleico que codifica para una glicosiltransferasa, o una subsecuencia de la misma. Los métodos adecuados incluyen el método del fosfotriéster de Narang *et al.* (1979) *Meth. Enzymol.* 68: 90-99; el método del fosfodiéster de Brown *et al.* (1979) *Meula. Enzymol.* 68: 109-151; el método de la dietilfosoramidita de Beaucage *et al.* (1991) *Tetra. Lett.*, 22: 1859-1862; y el método de soporte sólido de la patente estadounidense n.º 4.458.066. La síntesis química produce un oligonucleótido de cadena sencilla. Éste puede convertirse en ADN bicatenario mediante hibridación con una secuencia complementaria, o mediante polimerización con una ADN polimerasa usando la cadena sencilla como molde. Un experto reconoce que aunque la síntesis química de ADN se limita a menudo a secuencias de aproximadamente 100 bases, pueden obtenerse secuencias más largas mediante el ligamiento de secuencias más cortas.

Pueden clonarse ácidos nucleicos que codifican para glicosiltransferasas, o subsecuencias de las mismas, usando métodos de amplificación de ADN tales como la reacción en cadena de la polimerasa (PCR). Por tanto, por ejemplo, la secuencia de ácido nucleico o subsecuencia se amplifica por PCR, usando un cebador homosenrido que contiene un sitio de enzima de restricción (por ejemplo, NdeI) y un cebador antisentido que contiene otro sitio de enzima de restricción (por ejemplo, HindIII). Esto producirá un ácido nucleico que codifica para la glicosiltransferasa o subsecuencia deseada y que tiene sitios de enzima de restricción terminales. Este ácido nucleico puede ligarse luego fácilmente en un vector que contiene un ácido nucleico que codifica para la segunda molécula y que tiene los sitios de enzima de restricción correspondientes apropiados. Un experto en la técnica puede determinar cebadores de PCR adecuados usando la información de secuencia proporcionada en GenBank u otras fuentes. También pueden añadirse sitios de enzima de restricción apropiados al ácido nucleico que codifica para la proteína glicosiltransferasa o subsecuencia de proteína mediante mutagénesis dirigida al sitio. El plásmido que contiene la secuencia de nucleótidos o subsecuencia que codifica para la glicosiltransferasa se escinde con la endonucleasa de restricción apropiada y luego se liga en un vector apropiado para amplificación y/o expresión según métodos convencionales. Se encuentran ejemplos de técnicas suficientes para dirigir a los expertos a través de los métodos de amplificación *in vitro* en Berger, Sambrook, y Ausubel, así como Mullis *et al.*, (1987) la patente estadounidense n.º 4,683, 202; PCR Protocols A Guide to Methods and Applications (Innis *et al.*, eds) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (el 1 de octubre de 1990) *C&EN* 36-47; *The Journal Of NIH Research* (1991) 3: 81-94; (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173; Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874; Lomell *et al.* (1989) *J. Clin. Chem.*, 35: 1826; Landegren *et al.*, (1988) *Science* 241: 1077-1080; Van Brunt (1990) *Biotechnology* 8: 291-294; Wu y Wallace (1989) *Gen* 4: 560; y Barringer *et al.* (1990) *Gen* 89: 117.

Pueden compararse otras propiedades físicas de una proteína glicosiltransferasa clonada, incluyendo proteína de fusión de glicosiltransferasa, expresada a partir de un ácido nucleico particular, con las propiedades de glicosiltransferasas conocidas para proporcionar otro método de identificación de secuencias o dominios adecuados de la glicosiltransferasa que son determinantes de especificidad de sustrato aceptor y/o actividad catalítica. Alternativamente, puede mutarse un supuesto gen de glicosiltransferasa o gen de glicosiltransferasa recombinante, y determinarse su papel como glicosiltransferasa, su capacidad para replegarse o el papel de secuencias o dominios particulares, detectando una variación en la estructura de un hidrato de carbono producido normalmente por la glicosiltransferasa no mutada, que se produce de manera natural o control.

Pueden identificarse dominios funcionales de glicosiltransferasas clonadas usando métodos convencionales para mutar o modificar las glicosiltransferasas y someter a prueba las proteínas modificadas o mutadas para determinar actividades tales como actividad de sustrato aceptor y/o actividad catalítica, tal como se describe en el presente documento. Los dominios funcionales de las diversas glicosiltransferasas pueden usarse para construir ácidos nucleicos que codifican para proteínas de fusión de glicosiltransferasas recombinantes que comprende los dominios funcionales de una o más glicosiltransferasas. Estas proteínas de fusión luego pueden someterse a prueba para determinar la actividad catalítica o de sustrato aceptor deseada.

En un enfoque a modo de ejemplo para la clonación de proteínas de fusión de glicosiltransferasas recombinantes, se alinean y se comparan las secuencias de ácido nucleico o de aminoácidos conocidas de glicosiltransferasas clonadas para determinar la cantidad de identidad de secuencia entre diversas glicosiltransferasas. Esta información puede usarse para identificar y seleccionar dominios de proteína que confieren o modulan actividades de glicosiltransferasa, por ejemplo, actividad de sustrato aceptor y/o actividad catalítica basándose en la cantidad de identidad de secuencia entre las glicosiltransferasas de interés. Por ejemplo, pueden usarse dominios que tienen identidad de secuencia entre las glicosiltransferasas de interés, y que se asocian con una actividad conocida, para

construir proteínas de fusión de glicosiltransferasas recombinantes que contienen ese dominio, y que tienen la actividad asociada con ese dominio (por ejemplo, especificidad de sustrato aceptor y/o actividad catalítica).

V. Expresión de glicosiltransferasas recombinantes

5 Pueden expresarse glicosiltransferasas eucariotas recombinantes en una variedad de células huésped, incluyendo *E. coli*, otros huéspedes bacterianos, levadura, y diversas células eucariotas superiores tales como las líneas celulares COS, CHO y HeLa y las líneas celulares de mieloma. Las células huésped pueden ser células de mamífero, células vegetales o microorganismos, tales como, por ejemplo, células de levadura, células bacterianas o células fúngicas filamentosas. Los ejemplos de células huésped adecuadas incluyen, por ejemplo, *Azotobacter sp.* (por ejemplo, *A. vinelandii*), *Pseudomonas sp.*, *Rhizobium sp.*, *Erwinia sp.*, *Escherichia sp.* (por ejemplo, *E. coli*), *Bacillus*, *Pseudomonas*, *Proteus*, *Salmonella*, *Serratia*, *Shigella*, *Rhizobia*, *Vitreoscilla*, *Paracoccus* y *Klebsiella sp.*, entre muchos otros. Las células pueden ser de cualquier de varios genera, incluyendo *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. xilinus*, *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 ejemplos de útiles bacterias incluyen, pero no se limitan a, *Escherichia*, *Enterobacter*, *Azotobacter*, *Erwinia*, *Klebsiella*.

20 Normalmente, el polinucleótido que codifica para la proteína de fusión se sitúa bajo el control de un promotor que es funcional en la célula huésped deseada. Se conoce bien una variedad extremadamente alta de promotores, y pueden usarse en los vectores de expresión de la invención, dependiendo de la aplicación particular. Habitualmente, el promotor seleccionado depende de la célula en que el promotor va a ser activo. Otras secuencias de control de la expresión tales como sitios de unión al ribosoma, sitios de terminación de la transcripción y similares también están incluidos opcionalmente. Los constructos que incluyen una o más de estas secuencias de control se denominan "casetes de expresión". Por consiguiente, la invención proporciona casetes de expresión en los que se incorporan los ácidos nucleicos que codifican para proteínas de fusión para la expresión de alto nivel en una célula huésped deseada.

30 A menudo se obtienen secuencias de control de la expresión que son adecuados para su uso en una célula huésped particular mediante la clonación de un gen que se expresa en esa célula. Las secuencias de control procariontas usadas comúnmente, que se definen en el presente documento que incluyen promotores para la iniciación de la transcripción, opcionalmente con un operador, junto con secuencias de sitio de unión al ribosoma, incluyen promotores usados comúnmente tales como los sistemas promotores de beta-lactamasa (penicilinas) y lactosa (*lac*) (Change *et al.*, *Nature* (1977) 198: 1056), el sistema promotor de triptófano (*trp*) (Goeddel *et al.*, *Nucleic Acids Res.* (1980) 8: 4057), el promotor *tac* (DeBoer, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* (1983) 80: 21-25); y el promotor P_L derivado de *lambda* y el sitio de unión al ribosoma del gen *N* (Shimatake *et al.*, *Nature* (1981) 292: 128). El sistema promotor particular no es crítico para la invención, puede usarse cualquier promotor disponible que funcione en procariontas.

40 Para la expresión de glicosiltransferasas eucariotas recombinantes en células procariontas distintas de *E. coli*, se requiere un promotor que funciona en la especie procarionta particular. Tales promotores pueden obtenerse de genes que se han clonado a partir de los promotores de la especie, o pueden usarse promotores heterólogos. Por ejemplo, el promotor híbrido *trp-lac* funciona en *Bacillus* además de *E. coli*.

45 Un sitio de unión al ribosoma (RBS) está incluido de manera conveniente en los casetes de expresión de la invención. Un RBS en *E. coli*, por ejemplo, consiste en una secuencia de nucleótidos de 3-9 nucleótidos de longitud ubicados 3-11 nucleótidos en sentido 5' del codón de iniciación (Shine y Dalgarno, *Nature* (1975) 254: 34; Steitz, *In Biological regulation and development: Gene expression* (ed. R. F. Goldberger), vol. 1, pág. 349, 1979, Plenum Publishing, NY).

50 Para la expresión de las glicosiltransferasas eucariotas recombinantes en levadura, los promotores convenientes incluyen GAL1-10 (Johnson y Davies (1984) *Mol. Cell. Biol.* 4: 1440-1448) ADH2 (Russell *et al.* (1983) *J. Biol. Chem.* 258: 2674-2682), PH05 (EMBO J. (1982) 6: 675-680) y MF α (Herskowitz y Oshima (1982) *In 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 promotor adecuado para su uso en levadura es el promotor híbrido ADH2/GAPDH tal como se describe en Cousens *et al.*, *Gen* 61: 265-275 (1987). Para hongos filamentosos tales como, por ejemplo, cepas de los hongos *Aspergillus* (McKnight *et al.*, patente estadounidense n.º 4.935-349), los ejemplos de útiles promotores incluyen aquellos derivados de genes glicolíticos de *Aspergillus nidulans*, tales como el promotor de ADH3 (McKnight *et al.*, *EMBO J.* 4: 2093-2099 (1985)) y el promotor *tpiA*. Un ejemplo de un terminador adecuado es el terminador de ADH3 (McKnight *et al.*).

Los promotores constitutivos adecuados para su uso en plantas incluyen, por ejemplo, promotores de la región VI y la región de iniciación de la transcripción 35S del virus del mosaico de la coliflor (VMCo), el promotor en 1' o 2' derivado de ADN-T de *Agrobacterium tumefaciens*, y otros promotores activos en células vegetales que conocen los

5 expertos en la técnica. Otros promotores adecuados incluyen el promotor de transcrito de longitud completa del virus del mosaico de la escrofularia, promotores de actina, promotores de histona, promotores de tubulina, o el promotor de la manopina sintasa (MAS). Otros promotores vegetales constitutivos incluyen diversos promotores de ubiquitina o poliubiquitina derivados de, entre otros, *Arabidopsis* (Sun y Callis, *Plant J.*, 11 (5): 1017-1027 (1997)), los
 10 promotores mas, Mac o DoubleMac (descritos en la patente estadounidense n.º 5.106.739 y por Comai *et al.*, *Plant Mol. Biol.* 15: 373-381 (1990)) y otras regiones de iniciación de la transcripción de diversos genes de plantas conocidos por los expertos en la técnica. Los promotores útiles para plantas también incluyen aquéllos obtenidos a partir de plásmidos Ti o Ri, de células vegetales, virus de plantas u otros huéspedes en los que se encuentra que los promotores son funcionales en plantas. Los promotores bacterianos que funcionan en plantas, y por tanto son
 15 adecuados para su uso en los métodos de la invención incluyen el promotor de octopina sintetasa, el promotor de nopalina sintasa, y el promotor de manopina sintetasa. Los promotores de plantas endógenos adecuados incluyen el promotor de la subunidad pequeña (ssu) de la ribulosa-1,6-bisfosfato (RUBP) carboxilasa, el promotor de (α -congliginina, el promotor de faseolina promotor, el promotor de ADH y promotores de choque térmico.

15 Pueden usarse promotores o bien constitutivos o bien regulados en la presente invención. Los promotores regulados pueden ser ventajosos porque las células huésped se pueden hacer crecer hasta altas densidades antes de que se induzca la expresión de las proteínas de fusión. La expresión de alto nivel de proteínas heterólogas ralentiza el crecimiento celular en algunas situaciones. Un promotor inducible es un promotor que dirige la expresión de un gen en la que el nivel de expresión puede alterarse por factores ambientales o de desarrollo tales como, por ejemplo, temperatura, pH, condiciones anaerobias o aerobias, luz, factores de transcripción y productos químicos. Tales
 20 promotores se denominan en el presente documento "promotores inducibles", que permiten que se controle el momento de la expresión de la glicosiltransferasa o enzima implicada en la síntesis de azúcares de nucleótidos. Para *E. coli* y otras células huésped bacterianas, los expertos en la técnica conocen promotores inducibles. Éstos incluyen, por ejemplo, el promotor lac, el promotor de P_L del bacteriófago lambda, el promotor híbrido trp-lac (Amann *et al.* (1983) *Gen* 25: 167; de Boer *et al.* (1983) *Proc. Nat'l. Acad. Sci. USA* 80: 21), y el promotor T7 de bacteriófago (Studier *et al.* (1986) *J. Mol. Biol.*; Tabor *et al.* (1985) *Proc. Nat'l. Acad. Sci. USA* 82: 1074-8). Estos promotores y su
 25 uso se tratan en Sambrook *et al.*, citado anteriormente. Un promotor inducible particularmente preferido para la expresión en procariontes es un promotor dual que incluye un componente de promotor tac unido a un componente de promotor obtenido a partir de un gen o genes que codifican para enzimas implicadas en el metabolismo de la galactosa (por ejemplo, un promotor de un gen de UDPgalactosa 4-epimerasa (galE)). El promotor dual tac-gal, que se describe en la publicación de solicitud de patente PCT n.º W098/20111, proporciona un nivel de expresión que es mayor que el proporcionado por cualquier promotor solo.

30 Los expertos en la técnica conocen promotores inducibles para su uso en plantas (véanse, por ejemplo, las referencias citadas en Kuhlemeier *et al* (1987) *Ann. Rev. Plant Physiol.* 38: 221), e incluyen aquéllos de los genes de la subunidad pequeña de la 1,5-ribulosa bisfosfato carboxilasa de *Arabidopsis thaliana* (el promotor "ssu"), que son inducibles por la luz y activos sólo en tejido fotosintético.

35 Los expertos en la técnica también conocen bien promotores inducibles para otros organismos. Éstos incluyen, por ejemplo, el promotor de arabinosa, el promotor lacZ, el promotor de metalotioneína y el promotor de choque térmico, así como muchos otros.

40 Un constructo que incluye un polinucleótido de interés operativamente unido a señales de control de la expresión génica que, cuando se ponen en una célula huésped apropiada, dirigen la expresión del polinucleótido, se denomina un "casete de expresión". Los casetes de expresión que codifican para las proteínas de fusión de la invención se ponen a menudo en vectores de expresión para su introducción en la célula huésped. Los vectores incluyen normalmente, además de un casete de expresión, una secuencia de ácido nucleico que permite que el vector se replique independientemente en una o más células huésped seleccionadas. Generalmente, esta secuencia es una
 45 que permite que el vector se replique independientemente del ADN cromosómico del huésped, e incluye orígenes de replicación o secuencias de replicación autónoma. Se conocen bien secuencias de este tipo para una variedad de bacterias. Por ejemplo, el origen de replicación del plásmido pBR322 es adecuado para la mayoría de bacterias Gram-negativas. Alternativamente, el vector puede replicarse al integrarse en el complemento genómico de la célula huésped y que se replica cuando la célula experimenta la replicación de ADN. Un vector de expresión preferido para la expresión de las enzimas es en células bacterianas es pTGK, que incluye un promotor dual tac-gal y se describe en la publicación de solicitud de patente PCT n.º W098/20111.

50 También puede ser deseable añadir secuencias reguladoras que pueden permitir la regulación de la expresión del polipéptido con relación al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son aquéllos que provocan que se active o desactive la expresión del gen en respuesta a un estímulo químico o físico, incluyendo la presencia de un compuesto regulador. Los sistemas reguladores en sistemas procariontes incluyen los sistemas de operador lac, tac y trp. En levadura, puede usarse el sistema de ADH2 o el sistema de GAL1. En hongos filamentosos, pueden usarse el promotor de α -amilasa TAKA, el promotor de glucoamilasa de *Aspergillus niger* y el promotor de glucoamilasa de *Aspergillus oryzae* como secuencias reguladoras.

60 La construcción de constructos de polinucleótido generalmente requiere el uso de vectores que puedan replicarse en bacterias. Está disponible comercialmente una gran cantidad de kits para la purificación de plásmidos a partir de bacterias (véanse, por ejemplo, EasyPrepJ, FlexiPrepJ, ambos de Pharmacia Biotech; StrataCleanJ, de Stratagene;

y el sistema de expresión QIAexpress, Qiagen). Los plásmidos aislados y purificados pueden manipularse luego adicionalmente para producir otros plásmidos, y usarse para transfectar células. También es posible la clonación en *Streptomyces* o *Bacillus*.

5 A menudo se incorporan marcadores seleccionables en los vectores de expresión usados para expresar los polinucleótidos de la invención. Estos genes pueden codificar para un producto génico, tal como una proteína, necesaria para la supervivencia o el crecimiento de células huésped transformadas hechas crecer en un medio de cultivo selectivo. Las células huésped no transformadas con el vector que contiene el gen de selección no sobrevivirán en el medio de cultivo. Los genes de selección típicos codifican para proteínas que confieren resistencia a antibióticos u otras toxinas, tales como ampicilina, neomicina, kanamicina, cloranfenicol o tetraciclina. Alternativamente, los marcadores seleccionables pueden codificar para proteínas que complementan deficiencias auxotróficas o suministrar nutrientes críticos no disponibles a partir de medios complejos, por ejemplo, el gen que codifica para D-alanina racemasa para bacilos. 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 huésped. Los expertos en la técnica conocen varios marcadores seleccionables y se describen por ejemplo en Sambrook *et al.*, citado anteriormente. Un marcador seleccionable preferido para su uso en células bacterianas es un marcador de resistencia a kanamicina (Vieira y Messing, Gen 19: 259 (1982)). El uso de la selección de kanamicina es ventajosa con respecto, por ejemplo, a la selección de ampicilina debido a que la ampicilina se degrada rápidamente por P-lactamasa en medio de cultivo, eliminando por tanto la presión selectiva y permitiendo que el cultivo llegue a tener un sobrecrecimiento con células que no contienen el vector.

20 La construcción de vectores adecuados que contienen uno o más de los componentes enumerados anteriormente emplea técnicas de ligamiento convencionales tal como se describe en las referencias citadas anteriormente. Se escinden los plásmidos o fragmentos de ADN aislados, se adaptan y se vuelven a ligar en la forma deseada para generar los plásmidos requeridos. Para confirmar las secuencias correctas en los plásmidos construidos, pueden analizarse los plásmidos mediante técnicas convencionales tales como mediante digestión con endonucleasa de restricción, y/o secuenciación según métodos conocidos. Se conocen en la técnica, técnicas de clonación molecular para lograr estos fines. Los expertos conocen bien una amplia variedad de métodos de clonación y amplificación *in vitro* adecuados para la construcción de ácidos nucleicos recombinantes. Se encuentran ejemplos de estas técnicas e instrucciones suficientes para dirigir a los expertos a través de muchos ejercicios de clonación en Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Volumen 152, Academic Press, Inc., San Diego, CA (Berger); 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., (1998 Complemento) (Ausubel).

35 Se conoce bien en la técnica una variedad de vectores comunes adecuados para su uso como materiales de partida para construir vectores de expresión de la invención. Para la clonación en bacterias, los vectores comunes incluyen vectores derivados de pBR322 tales como PBLUESCRIPT™ y vectores derivados de fago λ . En levadura, los vectores incluyen plásmidos de integración en levadura (por ejemplo, Ylp5) y plásmidos de replicación en levadura (los plásmidos de la serie YRp) y pGPD-2. Puede lograrse la expresión en células de mamífero usando una variedad de plásmidos disponibles comúnmente, incluyendo pSV2, pBC12BI y p91023, así como vectores virales líticos (por ejemplo, virus vaccinia, adenovirus y baculovirus), vectores virales episómicos (por ejemplo, virus del papiloma bovino) y vectores retrovirales (por ejemplo, retrovirus murinos).

40 Los métodos para introducir vectores de expresión en una célula huésped elegida no son particularmente críticos, y tales métodos los conocen los expertos en la técnica. Por ejemplo, los vectores de expresión pueden introducirse en células procariotas, incluyendo *E. coli*, mediante transformación con cloruro de calcio, y en células eucariotas mediante tratamiento con fosfato de calcio o electroporación. También son adecuados otros métodos de transformación.

45 Puede usarse acoplamiento traduccional para mejorar la expresión. La estrategia usa un marco de lectura abierto corto en sentido 5' derivado de un gen altamente expresado nativo para el sistema de traducción, que se pone en sentido 3' del promotor, y un sitio de unión al ribosoma seguido después de unos pocos codones de aminoácido por un codón de terminación. Justo antes del codón de terminación hay un segundo sitio de unión al ribosoma, y tras el codón de terminación hay un codón de inicio para la iniciación de la traducción. El sistema disuelve la estructura secundaria en el ARN, permitiendo la iniciación eficaz de la traducción. Véase Squires, et. al. (1988), J. Biol. Chem. 263: 16297-16302.

55 Las glicosiltransferasas eucariotas recombinantes de la invención también pueden unirse además a otras proteínas bacterianas. Este enfoque a menudo da como resultado altos rendimientos, porque secuencias de control procariotas normales dirigen la transcripción y la traducción. En *E. coli*, se usan fusiones lacZ a menudo para expresar proteínas heterólogas. Están disponibles fácilmente vectores adecuados, tales como el pUR, pEX y la serie pMR100 (véase, por ejemplo, Sambrook *et al.*, citado anteriormente.). Para determinadas aplicaciones, puede ser deseable escindir los aminoácidos de la enzima auxiliar y/o distintos a de glicosiltransferasa de la proteína de fusión tras purificación. Esto puede lograrse mediante cualquiera de varios métodos conocidos en la técnica, incluyendo escisión mediante bromuro de cianógeno, una proteasa, o mediante el factor X_a (véanse, por ejemplo, Sambrook *et al.*, citado anteriormente.; Itakura *et al.*, Science (1977) 198: 1056; Goeddel *et al.*, Proc. Natl. Acad. Sci. USA (1979) 76: 106; Nagai *et al.*, Nature (1984) 309: 810; Sung *et al.*, Proc. Natl. Acad. Sci. USA (1986) 83: 561). Los sitios de

escisión pueden modificarse mediante ingeniería genética en el gen para la proteína de fusión en el punto de escisión deseado.

Puede expresarse más de una glicosiltransferasa eucariota recombinante en una única célula huésped situando múltiples casetes de traducción en un único vector de expresión, o utilizando diferentes marcadores seleccionables para cada uno de los vectores de expresión que se emplean en la estrategia de clonación.

Un sistema adecuado para obtener proteínas recombinantes a partir de *E. coli* que mantiene la integridad de sus extremos N-terminales lo han descrito Miller *et al.* *Biotechnology* 7: 698-704 (1989). En este sistema, el gen de interés se produce como una fusión C-terminal con los primeros 76 residuos del gen de ubiquitina de levadura que contiene un sitio de escisión de peptidasa. La escisión en la unión de los dos restos da como resultado la producción de una proteína que tiene un residuo N-terminal auténtico intacto.

Los vectores de expresión de la invención pueden transferirse en la célula huésped elegida mediante métodos bien conocidos tales como transformación con cloruro de calcio para *E. coli* y tratamiento con fosfato de calcio o electroporación para células de mamífero. Pueden seleccionarse las células transformadas por los plásmidos mediante la resistencia a antibióticos conferida por los genes contenidos en los plásmidos, tales como los genes *amp*, *gpt*, *neo* y *hyg*.

VI. Proteínas y purificación de proteínas

Las proteínas glicosiltransferasas eucariotas recombinantes pueden purificarse según procedimientos convencionales de la técnica, incluyendo precipitación con sulfato de amonio, columnas de afinidad, cromatografía en columna, electroforesis en gel y similares (véanse, generalmente, R. Scopes, *Protein Purification*, Springer-Verlag, N. Y. (1982), Deutscher, *Methods in Enzymology* Vol. 182: *Guide to Protein Purification.*, Academic Press, Inc. N. Y. (1990)). En realizaciones preferidas, la purificación de las proteínas glicosiltransferasas eucariotas recombinantes se produce tras el repliegamiento de la proteína. Se prefieren composiciones sustancialmente puras de una homogeneidad de al menos aproximadamente el 70 al 90%; más preferiblemente al menos el 91%, 92%, 93%, 94%, 95%, 96% o el 97%; y lo más preferido es una homogeneidad del 98 al 99% o más. También pueden usarse las proteínas purificadas, por ejemplo, como inmunógenos para la producción de anticuerpos.

Para facilitar la purificación de las proteínas glicosiltransferasas eucariotas recombinantes de la invención, los ácidos nucleicos que codifican para las proteínas glicosiltransferasas eucariotas recombinantes también pueden incluir una secuencia codificante para un epítipo o "cola" para el que está disponible un reactivo de unión por afinidad, es decir, una cola de purificación. Los ejemplos de epítipos adecuados incluyen los genes indicadores *myc* y *V-5*; están disponibles comercialmente vectores de expresión útiles para la producción recombinante de proteínas de fusión que tienen estos epítipos (por ejemplo, los vectores *pcDNA3.1/Myc-His* y *pcDNA3.1/V5-His* de Invitrogen (Carlsbad CA) son adecuados para la expresión en células de mamífero). Los expertos en la técnica conocen vectores de expresión adecuados adicionales para unir una cola a las proteínas de fusión de la invención, y los correspondientes sistemas de detección, y varios están disponibles comercialmente (por ejemplo, "FLAG" (Kodak, Rochester NY). Otro ejemplo de una cola adecuada es una secuencia de polihistidina, que puede unirse a ligandos de afinidad de quelatos metálicos. Normalmente, se usan seis histidinas adyacentes (SEQ ID NO:67), aunque puede usarse más o menos de seis. Los ligandos de afinidad de quelatos metálicos que pueden servir como resto de unión para una cola de polihistidina incluyen ácido nitrilo-tri-acético (NTA) (Hochuli, E. (1990) "Purification of recombinant proteins with metal chelating adsorbents" in *Genetic Engineering: Principles and Methods*, J. K. Setlow, Ed., Plenum Press, NY; disponibles comercialmente de Qiagen (Santa Clarita, CA)).

Las colas de purificación también incluyen dominios de unión a maltosa y dominios de unión a almidón. Los expertos en la técnica conocen la purificación de proteínas con dominio de unión a maltosa. Se describen dominios de unión a almidón en el documento WO 99/15636, incorporado al presente documento como referencia. La purificación por afinidad de una proteína de fusión que comprende un dominio de unión a almidón usando una resina derivatizada con betaciclodextrina (BCD) se describe en el documento USSN 60/468.374, presentado el 5 de mayo de 2003, incorporado al presente documento como referencia en su totalidad.

Los expertos en la técnica conocen otros haptenos que son adecuados para su uso como colas y se describen, por ejemplo, en the *Handbook of Fluorescent Probes and Research Chemicals* (6ª Ed., Molecular Probes, Inc., Eugene O). Por ejemplo, dinitrofenol (DNP), digoxigenina, barbitúricos (véase, por ejemplo, la patente estadounidense n.º 5.414.085), y varios tipos de fluoróforos son útiles como haptenos, como lo son los derivados de estos compuestos. Están disponibles comercialmente kits para la unión de haptenos y otros restos a proteínas y otras moléculas. Por ejemplo, cuando el hapteno incluye un tiol, puede usarse un ligador heterobifuncional tal como SMCC para unir la cola a residuos de lisina presentes en el reactivo de captura.

Un experto reconocería que pueden realizarse modificaciones a los dominios catalíticos o funcionales de la glicosiltransferasa y/o dominios catalíticos auxiliares de la enzima sin disminuir su actividad biológica. Pueden realizarse algunas modificaciones para facilitar la clonación, expresión o incorporación del dominio catalítico en una proteína de fusión. Los expertos en la técnica conocen bien tales modificaciones e incluyen, por ejemplo, la adición de codones en cualquier extremo terminal del polinucleótido que codifica para el dominio catalítico para

proporcionar, por ejemplo, una metionina añadida en el extremo amino-terminal para proporcionar un sitio de iniciación, o aminoácidos adicionales (por ejemplo, poli-His) situados en cualquier extremo terminal para crear sitios de enzima de restricción o codones de terminación o secuencias de purificación ubicados de manera conveniente.

VII. Usos de las glicosiltransferasas replegadas

5 La invención proporciona proteínas glicosiltransferasas eucariotas recombinantes y métodos de uso de las proteínas glicosiltransferasas eucariotas recombinantes para sintetizar enzimáticamente glicoproteínas, glicolípidos y restos de oligosacárido, y para glicopegilar glicoproteínas. Las reacciones con glicosiltransferasas de la invención tienen lugar en un medio de reacción que comprende al menos una glicosiltransferasa, sustrato aceptor y sustrato donador, y normalmente un catión metálico divalente soluble. En algunas realizaciones, también están presentes enzimas auxiliares y sustratos para el resto catalítico de las enzimas auxiliares, de modo que las enzimas auxiliares pueden sintetizar el sustrato donador para la glicosiltransferasa. Las proteínas glicosiltransferasas eucariotas recombinantes y los métodos de la presente invención se basan en el uso de las proteínas glicosiltransferasas eucariotas recombinantes para catalizar la adición de un sacárido a un sustrato aceptor.

15 Se conocen varios métodos de uso de glicosiltransferasas para sintetizar glicoproteínas y glicolípidos que tienen restos de oligosacárido deseados. Se describen métodos a modo de ejemplo en, por ejemplo, el documento WO 96/32491, Ito *et al.* (1993) Pure Appl. Chem. 65: 753 y las patentes estadounidenses 5.352.670, 5.374.541 y 5.545.553.

20 Las proteínas glicosiltransferasas eucariotas recombinantes preparadas tal como se describe en el presente documento pueden usarse en combinación con glicosiltransferasas adicionales, que pueden tener o no un replegamiento requerido para su actividad. Por ejemplo, puede usarse una combinación de una proteína glicosiltransferasa eucariota recombinante replegada y una glicosiltransferasa bacteriana, que puede replegarse o no tras el aislamiento de una célula huésped. De manera similar, la glicosiltransferasa eucariota recombinante puede usarse con enzimas auxiliares recombinantes, que pueden ser parte o no de la proteína de fusión.

25 Los productos producidos mediante los procedimientos anteriores pueden usarse sin purificación. En algunas realizaciones, se producen oligosacáridos. Pueden usarse técnicas bien conocidas, convencionales, por ejemplo, cromatografía en capa fina o gruesa, cromatografía de intercambio iónico o filtración con membranas para la recuperación de sacáridos glicosilados. También, por ejemplo, puede usarse la filtración con membranas, utilizando una membrana de nanofiltración u ósmosis inversa tal como se describe en la patente australiana cedida legalmente n.º 735695. Como ejemplo adicional, puede usarse la filtración con membranas en la que las membranas tienen un corte de peso molecular de aproximadamente 1000 hasta aproximadamente 10.000 para eliminar proteínas. Como otro ejemplo, puede usarse entonces la nanofiltración u ósmosis inversa para eliminar sales. Las membranas de nanofiltración son una clase de membranas de ósmosis inversa por las que pasan sales monovalentes pero que retienen las sales polivalentes y solutos sin carga mayores que de aproximadamente 200 a aproximadamente 1000 Daltons, dependiendo de la membrana usada. Por tanto, por ejemplo, los oligosacáridos producidos mediante las composiciones y los métodos de la presente invención pueden retenerse en la membrana y contaminar las sales que pasarán a su través.

VIII. Sustrato donador/sustratos aceptores

40 Los sustratos donadores adecuados usados por las proteínas de fusión de glicosiltransferasas recombinantes y los métodos de la invención incluyen, pero no se limitan a, UDP-Glc, UDP-GlcNAc, UDP-Gal, UDP-GalNAc, GDP-Man, GDP-Fuc, UDP-GlcUA y CMP-ácido siálico. Guo *et al.*, Applied Biochem. and Biotech. 68: 1-20 (1997)

45 Los sustratos aceptores adecuados usados por las proteínas de fusión de glicosiltransferasas recombinantes y los métodos de la invención incluyen, pero no se limitan a, polisacáridos, oligosacáridos, proteínas, lípidos, gangliósidos y otras estructuras biológicas (por ejemplo, células completas) que pueden modificarse mediante los métodos de la invención. Las estructuras a modo de ejemplo, que pueden modificarse mediante los métodos de la invención incluyen cualquiera de varios glicolípidos, glicoproteínas y estructuras de hidrato de carbono en células conocidas por los expertos en la técnica tal como se expone en la tabla 1.

Tabla 1

<p><u>Hormonas y factores de crecimiento</u></p> <ul style="list-style-type: none"> • G-CSF • GM-CSF • TPO • EPO • Variantes de EPO • α-TNF • Leptina <p><u>Enzimas e inhibidores</u></p> <ul style="list-style-type: none"> • t-PA • Variantes de t-PA • Urocinasa • Factores VII, VIII, IX, X • ADNasa • Glucocerebrosidasa • Hirudina • α1 antitripsina • Antitrombina III <p><u>Citocinas y citocinas quiméricas</u></p> <ul style="list-style-type: none"> • Interleucina-1 (IL-1), 1B, 2,3,4 • Interferón-α (IFN-α) • IFN-α-2b • IFN-β • IFN-γ • IL-2 de toxina diftérica quimérica 	<p><u>Receptores y receptores quiméricos</u></p> <ul style="list-style-type: none"> • CD4 • Receptor del factor de necrosis tumoral (TNF) • Alfa-CD20 • AcM-CD20 • AcM-alfa-CD3 • AcM-receptor de TNF • AcM-CD4 • PSGL-1 • AcM-PSGL-1 • Complemento • GlyCAM o su quimera • N-CAM o su quimera • LFA-3 • CTLA-IV <p><u>Anticuerpos monoclonales (inmunoglobulinas)</u></p> <ul style="list-style-type: none"> • AcM-anti-VSR • AcM-anti-receptor de IL-2 • AcM-anti-CEA • AcM-anti-receptor IIb/IIIa de plaquetas • AcM-anti-EGF • AcM-anti-receptor de Her-2 <p><u>Células</u></p> <ul style="list-style-type: none"> • Glóbulos rojos • Glóbulos blancos (por ejemplo, células T, células B, células dendríticas, macrófagos, linfocitos citolíticos, neutrófilos, monocitos y similares) • Células madre
--	--

5 Se describen ejemplos de sustratos aceptores adecuados usados en reacciones catalizadas por fucosiltransferasas, y ejemplos de sustratos aceptores adecuados usados en reacciones catalizadas por sialiltransferasas en Guo *et al.*, Applied Biochem. and Biotech. 68: 1-20 (1997), pero no se limitan a los mismos.

IX. Reacciones con glicosiltransferasas

10 Las proteínas glicosiltransferasas eucariotas recombinantes, sustratos aceptores, sustratos donadores y otros componentes de la mezcla de reacción se combinan mediante mezclado en un medio de reacción acuoso. El medio generalmente tiene un valor de pH de aproximadamente 5 hasta aproximadamente 8.5. La selección de un medio se basa en la capacidad del medio para mantener el valor de pH en el nivel deseado. Por tanto, en algunas realizaciones, el medio se tampona hasta un valor de pH de aproximadamente 7,5. Si no se usa un tampón, debe mantenerse el pH del medio en aproximadamente de 5 a 8,5, dependiendo de la glicosiltransferasa particular usada. Para fucosiltransferasas, se mantiene preferiblemente el intervalo de pH en desde aproximadamente 6,0 hasta 8,0. Para sialiltransferasas, el intervalo es preferiblemente de desde aproximadamente 5,5 hasta aproximadamente 7,5.

15 Las cantidades o concentraciones de enzima se expresan en unidades de actividad, que es una medida de la tasa de catálisis inicial. Una unidad de actividad cataliza la formación de 1 μ mol de producto por minuto a una temperatura dada (normalmente 37°C) y valor de pH (normalmente 7,5). Por tanto, 10 unidades de una enzima es una cantidad catalítica de esa enzima en la que se convierten 10 μ mol de sustrato en 10 μ mol de producto en un minuto a una temperatura de 37°C y un valor de pH de 7,5.

20 La mezcla de reacción puede incluir cationes metálicos divalentes (Mg^{2+} , Mn^{2+}). El medio de reacción también puede comprender detergentes de solubilización (por ejemplo, Triton o SDS) y disolventes orgánicos tales como metanol o etanol, si es necesario. Las enzimas pueden utilizarse libres en disolución o pueden unirse a un soporte tal como un polímero. La mezcla de reacción es por tanto sustancialmente homogénea al comienzo, aunque puede formarse algún precipitado durante la reacción.

25 La temperatura a la que se lleva a cabo el procedimiento anterior puede oscilar entre justo por encima de la temperatura de congelación hasta la temperatura a la que se desnaturaliza la enzima más sensible. Ese intervalo de temperatura es preferiblemente de aproximadamente 0°C a aproximadamente 45°C, y más preferiblemente de aproximadamente 20°C a aproximadamente 37°C.

La mezcla de reacción así formada se mantiene durante un periodo de tiempo suficiente para obtener el alto rendimiento deseado de los determinantes de oligosacárido deseados presente en los grupos de oligosacárido unidos a la glicoproteína que va a glicosilarse. Para preparaciones a gran escala, a menudo se permitirá que avance la reacción durante entre aproximadamente 0,5-240 horas, y más normalmente entre aproximadamente 1-18 horas.

5 Puede llevarse a cabo una o más de las reacciones con glicosiltransferasas como parte de un ciclo de glicosiltransferasa. Se han descritos condiciones y descripciones preferidas de ciclos de glicosiltransferasas. Se describen varios ciclos de glicosiltransferasas (por ejemplo, ciclos de sialiltransferasas, ciclos de galactosiltransferasas y ciclos de fucosiltransferasas) en la patente estadounidense n.º 5.374.541 y el documento WO 9425615 A. Se describen otros ciclos de glicosiltransferasas en Ichikawa *et al.* J. Am. Chem. Soc. 114: 9283
10 (1992), Wong *et al.* J. Org. Chem. 57: 4343 (1992), DeLuca, *et al.*, J. Am. Chem. Soc. 117: 5869-5870 (1995), y Ichikawa *et al.* In Carbohydrates and Carbohydrate Polymers. Yaltami, ed. (ATL Press, 1993).

Pueden sustituirse otras glicosiltransferasas en ciclos de transferasas similares tal como se ha descritos en detalle para las fucosiltransferasas y sialiltransferasas. En particular, la glicosiltransferasa también puede ser, por ejemplo, glucosiltransferasas, por ejemplo, Alg8 (Stagljev *et al.*, Proc. Natl. Acad. Sci. USA 91: 5977 (1994)) o Alg5 (Heesen
15 *et al.* Eur. J. Biochem. 224: 71 (1994)), N-acetilgalactosaminiltransferasas tales como, por ejemplo, $\alpha(1,3)$ N-acetilgalactosaminiltransferasa, $\beta(1,4)$ N-acetilgalactosaminiltransferasas (Nagata *et al.* J. Biol. Chem. 267: 12082-12089 (1992) y Smith *et al.* J. Biol. Chem. 269: 15162 (1994)) y polipéptido N-acetilgalactosaminiltransferasa (Homa
20 *et al.* J. Biol. Chem. 268: 12609 (1993)). Las N-acetilglucosaminiltransferasas adecuadas incluyen GnTI (2.4.1.101, Hull *et al.*, BBRC 176: 608 (1991)), GnTII y GnTIII (Ihara *et al.* J. Biochem. 113: 692 (1993)), GnTV (Shoreiban *et al.*
J. Biol. Chem. 268: 15381 (1993)), N-acetilglucosaminiltransferasa unida a O (Bierhuizen *et al.* Proc. Natl. Acad. Sci. USA 89: 9326 (1992)), N-acetilglucosamina-1-fosfato transferasa (Rajput *et al.* Biochem J. 285: 985 (1992) y hialuronano sintasa. Las manosiltransferasas adecuadas incluyen $\alpha(1,2)$ manosiltransferasa, α
25 (1,3)manosiltransferasa, $\beta(1,4)$ manosiltransferasa, Dol-P-Man sintasa, OChI y PmtI.

Para los ciclos de glicosiltransferasas anteriores, las concentraciones o cantidades de los diversos reactantes
30 usados en los procedimientos dependen de numerosos factores incluyendo condiciones de reacción tales como temperatura y valor de pH, y la elección y cantidad de sacáridos aceptores que van a glicosilarse. Dado que el proceso de glicosilación permite la regeneración de nucleótidos activantes, azúcares donadores activados y la eliminación del PPI producido en presencia de cantidades catalíticas de las enzimas, el proceso está limitado por las concentraciones o cantidades de los sustratos estequiométricos tratados anteriormente. El límite superior para las
35 concentraciones de reactantes que pueden usarse según el método de la presente invención está determinado por la solubilidad de tales reactantes.

Preferiblemente, se seleccionan las concentraciones de nucleótidos activantes, donador de fosfato, el azúcar donador y las enzimas de manera que la glicosilación avanza hasta que se consume el aceptor. Las consideraciones
40 tratadas a continuación, aunque en el contexto de una sialiltransferasa, pueden aplicarse generalmente a otros ciclos de glicosiltransferasas.

Cada una de las enzimas está presente en una cantidad catalítica. La cantidad catalítica de una enzima particular
45 varía según la concentración del sustrato de esa enzima así como con respecto a condiciones de reacción tales como temperatura, tiempo y valor de pH. Los expertos en la técnica conocen bien medios para determinar la cantidad catalítica para una enzima dada en condiciones de reacción y a concentraciones de sustrato preseleccionadas.

X. Síntesis de oligosacáridos con múltiples enzimas

Tal como se trató anteriormente, en algunas realizaciones, pueden usarse dos o más enzimas para formar un
50 determinante de oligosacárido deseado en una glicoproteína o un glicolípido. Por ejemplo, un determinante de oligosacárido particular podría requerir la adición de una galactosa, un ácido siálico y una fucosa para presentar una actividad deseada. Por consiguiente, la invención proporciona métodos en los que se usan dos o más enzimas, por
55 ejemplo, glicosiltransferasas, trans-sialidasas o sulfotransferasas, para obtener la síntesis de alto rendimiento de un determinante de oligosacárido deseado.

En una realización particularmente preferida, una de las enzimas usadas es una sulfotransferasa que sulfona el
60 sacárido o el péptido. Incluso es más preferido el uso de una sulfotransferasa para preparar un ligando para una selectina (Kimura *et al.*, Proc Natl Acad Sci USA 96 (8): 4530-5 (1999)).

En algunos casos, un oligosacárido unido a glicoproteína o glicolípido incluirá un sustrato aceptor para la
65 glicosiltransferasa particular de interés con la biosíntesis *in vivo* de la glicoproteína o el glicolípido. Pueden glicosilarse tales glicoproteínas o glicolípidos usando las proteínas de fusión de glicosiltransferasas recombinantes y los métodos de la invención sin modificación previa del patrón de glicosilación de la glicoproteína o el glicolípido,
70 respectivamente. En otros casos, sin embargo, una glicoproteína o un glicolípido de interés carecerán de un sustrato aceptor adecuado. En tales casos, pueden usarse los métodos de la invención para alterar el patrón de glicosilación de la glicoproteína o el glicolípido de modo que los oligosacáridos unidos a glicoproteína o glicolípido incluyan

entonces un sustrato aceptor para la unión catalizada por glicosiltransferasa de una unidad de sacárido preseleccionada de interés para formar un resto de oligosacárido deseado.

5 Los oligosacáridos unidos a glicoproteína o glicolípido opcionalmente puede "cortarse" en primer lugar o bien en su totalidad o bien en parte, para exponer o bien un sustrato aceptor para la glicosiltransferasa o bien un resto al que pueden añadirse uno o más residuos apropiados para obtener un sustrato aceptor adecuado. Enzimas tales como glicosiltransferasas y endoglicosidasas son útiles para las reacciones de unión y corte. Por ejemplo, una glicoproteína que presenta oligosacáridos de tipo "con alto contenido en manosa" puede someterse a corte mediante una manosidasa para obtener un sustrato aceptor que, con la unión de una o más unidades de sacárido preseleccionadas, forma el determinante de oligosacárido deseado.

10 Los métodos también son útiles para sintetizar un resto de oligosacárido deseado en una proteína o un lípido que está sin glicosilar en su forma nativa. Puede unirse un sustrato aceptor adecuado para la correspondiente glicosiltransferasa a tales proteínas o lípidos antes de la glicosilación usando los métodos de la presente invención. Véase, por ejemplo, la patente estadounidense n.º 5.272.066 para métodos de obtención de polipéptidos que tienen aceptores adecuados para glicosilación.

15 Por tanto, en algunas realizaciones, la invención proporciona métodos para la sialilación *in vitro* de grupos de sacárido presentes en un glicoconjugado que implican en primer lugar modificar el glicoconjugado para crear un aceptor adecuado.

XI. Conjugación de azúcares modificados con péptidos

20 Se conjugan los azúcares modificados con un péptido o una proteína glicosilados o no glicosilados usando una enzima apropiada para mediar la conjugación. Preferiblemente, se seleccionan las concentraciones del/de los azúcar(es) modificado(s), enzima(s) y péptido(s) o proteína(s) aceptor(es) de manera que la glicosilación avanza hasta que se consume el aceptor. Las consideraciones tratadas a continuación, aunque en el contexto de una sialiltransferasa, pueden aplicarse generalmente a otras reacciones con glicosiltransferasas.

25 Se conocen varios métodos de uso de glicosiltransferasas para sintetizar estructuras de oligosacárido deseadas y pueden aplicarse generalmente a la presente invención. Se describen métodos a modo de ejemplo en, por ejemplo, el documento WO 96/32491, Ito *et al.*, Pure Appl. Chem. 65: 753 (1993), y las patentes estadounidenses n.ºs 5.352.670, 5.374.541 y 5.545.553.

30 En algunas realizaciones, se usa una endoglicosidasa en la reacción en combinación con glicosiltransferasas. Las enzimas se usan para alterar una estructura de un sacárido en el péptido en cualquier punto o bien antes o bien después de la adición del azúcar modificado al péptido.

35 En otra realización, el método hace uso de una o más exo- o endoglicosidasas. La glicosidasa es normalmente un mutante, que se modifica mediante ingeniería genética para formar enlaces glicosilo en vez de romperlos. La glicanasa mutante incluye normalmente una sustitución de un residuo de aminoácido por un residuo de aminoácido ácido de sitio activo. Por ejemplo, cuando la endoglicanasa es endo-H, los residuos de sitio activo sustituidos serán normalmente Asp en la posición 130, Glu en la posición 132 o una combinación de los mismos. Los aminoácidos se reemplazan generalmente por serina, alanina, asparagina o glutamina.

40 La enzima mutante cataliza la reacción, habitualmente mediante una etapa de síntesis que es análoga a la reacción inversa de la etapa de hidrólisis de la endoglicanasa. En estas realizaciones, la molécula donadora de glicosilo (por ejemplo, una estructura de oligo- o mono-sacárido deseada) contiene un grupo saliente y la reacción avanza con la adición de la molécula donadora a un residuo de GlcNAc en la proteína. Por ejemplo, el grupo saliente puede ser un halógeno, tal como fluoruro. En otras realizaciones, el grupo saliente es una Asn o un resto peptídico de Asn. Aún en realizaciones adicionales, se modifica el residuo de GlcNAc en la molécula donadora de glicosilo. Por ejemplo, el residuo de GlcNAc puede comprender un resto de 1,2-oxazolina.

45 En una realización preferida, cada una de las enzimas utilizada para producir un conjugado de la invención está presente en una cantidad catalítica. La cantidad catalítica de una enzima particular varía según la concentración del sustrato de esa enzima así como con respecto a condiciones de reacción tales como temperatura, tiempo y valor de pH. Los expertos en la técnica conocen bien medios para determinar la cantidad catalítica para una enzima dada en condiciones de reacción y a concentraciones de sustrato preseleccionadas.

50 La temperatura a la que se lleva a cabo el procedimiento anterior puede oscilar entre justo por encima de la temperatura de congelación hasta la temperatura a la que se desnaturaliza la enzima más sensible. Intervalos de temperatura preferidos son de aproximadamente 0°C a aproximadamente 55°C, y más preferiblemente de aproximadamente 20°C a aproximadamente 30°C. En otra realización a modo de ejemplo, uno o más componentes del presente método se llevan a cabo a una temperatura elevada usando una enzima termófila.

55 Se mantiene la mezcla de reacción durante un periodo de tiempo suficiente para que se glicosile el aceptor, formando de ese modo el conjugado deseado. Puede detectarse a menudo parte del conjugado tras unas pocas horas, obteniéndose habitualmente cantidades recuperables en el plazo de 24 horas o menos. Los expertos en la

técnica entienden que la velocidad de reacción depende de varios factores variables (por ejemplo, concentración de enzima, concentración de donador, concentración de aceptor, temperatura, volumen de disolvente), que se optimizan para un sistema seleccionado.

5 La presente invención también prevé la producción a escala industrial de péptidos modificados. Tal como se usa en el presente documento, una escala industrial generalmente produce al menos un gramo de conjugado terminado, purificado.

10 En el análisis que sigue, se ponen ejemplos de la invención mediante la conjugación de restos de ácido siálico modificados con un péptido glicosilado. Se marca con PEG el ácido siálico modificado a modo de ejemplo. El centro del siguiente análisis sobre el uso de ácido siálico modificado con PEG y péptidos glicosilados es para la claridad de ilustración y no pretende implicar que la invención se limita a la conjugación de estas dos parejas. Un experto entiende que el análisis puede aplicarse generalmente a las adiciones de restos de glicosilo modificados distintos del ácido siálico. Además, puede aplicarse igualmente el análisis a la modificación de una unidad de glicosilo con agentes distintos de PEG incluyendo otros polímeros solubles en agua, restos terapéuticos y biomoléculas.

15 Puede usarse un enfoque enzimático para la introducción selectiva de hidratos de carbono PEGilados o PPGilados sobre un péptido o glicopéptido. El método utiliza azúcares modificados que contienen PEG, PPG o un grupo funcional reactivo enmascarado, y se combina con la glicosiltransferasa o glicosintasa apropiada. Seleccionando la glicosiltransferasa que realizará la unión con el hidrato de carbono deseado y utilizando el azúcar modificado como sustrato donador, puede introducirse el PEG o PPG directamente sobre la estructura principal del péptido, sobre residuos de azúcar existentes de un glicopéptido o sobre residuos de azúcar que se han añadido a un péptido.

20 Está presente un aceptor para la sialiltransferasa en el péptido que va a modificarse mediante los métodos de la presente invención o bien como una estructura que se producen de manera natural o bien uno situado en el mismo de manera recombinante, de manera enzimática o de manera química. Los aceptores adecuados, incluyen, por ejemplo, aceptores de galactosilo tales como Gal β 1,4-GlcNAc, Gal β 1,4-GalNAc, Gal β 1,3-GalNAc, lacto-N-tetraosa, Gal β 1,3-GlcNAc, Gal β 1,3-Ara, Gal β 1,6GlcNAc, Gal β 1,4-Glc (lactosa), y otros aceptores conocidos por los expertos en la técnica (véase, por ejemplo, Paulson *et al.*, J. Biol. Chem. 253: 5617-5624 (1978)).

25 En una realización, está presente un aceptor para la sialiltransferasa en el glicopéptido que va a modificarse con la síntesis *in vivo* del glicopéptido. Pueden sialilarse tales glicopéptidos usando los métodos reivindicados sin modificación previa del patrón de glicosilación del glicopéptido. Alternativamente, pueden usarse los métodos de la invención para sialilar un péptido que no incluye un aceptor adecuado; se modifica en primer lugar el péptido para que incluya un aceptor mediante métodos conocidos por los expertos en la técnica. En una realización a modo de ejemplo, se añade un residuo de GalNAc mediante la acción de una GalNAc transferasa.

30 En una realización a modo de ejemplo, se ensambla un aceptor de galactosilo uniendo un residuo de galactosa a un aceptor apropiado unido al péptido, por ejemplo, una GlcNAc. El método incluye incubar el péptido que va a modificarse con una mezcla de reacción que contiene una cantidad adecuada de una galactosiltransferasa (por ejemplo, gal β 1,3 o gal β 1,4), y un donador de galactosilo adecuado (por ejemplo, UDP-galactosa). Se permite que avance la reacción sustancialmente hasta que se completa o, alternativamente, se termina la reacción cuando se añade una cantidad preseleccionada del residuo de galactosa. Otros métodos de ensamblaje de un aceptor de sacárido seleccionado resultarán evidentes para los expertos en la técnica.

35 Aun en otra realización, en primer lugar se "cortan" los oligosacáridos unidos a glicopéptido, o bien en su totalidad o bien en parte, para exponer o bien un aceptor para la sialiltransferasa o bien un resto al que puede añadirse uno o más residuos apropiados para obtener un aceptor adecuado. Enzimas tales como glicosiltransferasas y endoglicosidasas (véanse, por ejemplo la patente estadounidense n.º 5.716.812) son útiles para las reacciones de unión y corte.

40 Se encuentran métodos para la conjugación de azúcares modificados con péptidos o proteínas, por ejemplo, en el documento USSN 60/328.523 presentado el 10 de octubre de 2001; el documento USSN 60/387.292, presentado el 7 de junio de 2002; el documento USSN 60/391.777 presentado el 25 de junio de 2002; el documento USSN 60/404.249 presentado el 16 de agosto de 2002; y el documento PCT/US02/32263; incorporándose cada uno de ellos al presente documento como referencia para todos los fines.

Ejemplos

50 Ejemplo 1: Replegamiento de ST3GalIII de hígado de rata expresada en bacterias.

Replegamiento de la proteína de fusión GST-ST3GalIII de hígado de rata

55 Se clonó N-acetil-lactosaminida α -2, 3-sialiltransferasa (ST3GalIII) de hígado de rata en el vector pGEX-KT-Ext y se expresó como cuerpos de inclusión de GST-ST3-Gal III en células BL21 de *E. coli*. Se replegaron los cuerpos de inclusión usando un sistema redox de GSH/GSSG. La enzima replegada, GST-ST3-GalIII, era activa y transfirió ácido siálico a un sustrato de azúcar LNNt y a glicoproteínas asialiladas, por ejemplo, transferrina y factor IX.

Clonación de ST3GalIII en vector pGEX-Xt-KT

Se clonó el gen de ST3-GalIII de hígado de rata en sitios *Bam*HI y *Eco*RI del vector pGEX-KT-Ext tras amplificación por PCR usando los siguientes cebadores:

Homosentido Sial 5'Tm 5'-TTTGGATCCAAGCTACACTTACTCCAATGG (SEQ ID NO:68)

5 Antisentido: Sial 3'completo 5'-TTTGAATTCCTCAGATACCACTGCTTAAGTC (SEQ ID NO:69)

Expresión de GST-ST3GalIII en células BL21 de *E. coli*

Se transformó pGEX-ST3GalIII, un vector de expresión que comprende la fusión ST3GalIII-GST, en células BL21 de *E. coli* químicamente competentes. Se recogieron colonias individuales, se inocularon en cinco ml de medios LB con carbenicilina 100 µg/ml, y se hicieron crecer durante la noche a 37°C con agitación. Al día siguiente, se transfirió un ml de cultivo durante la noche a un litro de medios LB con carbenicilina 100 µg/ml. Se hicieron crecer las bacterias hasta una DO₆₂₀ de 0,7, luego se añadió IPTG 150 µM (final) al medio. Se hicieron crecer las bacterias a 37°C durante de una a dos horas más, luego se cambió a temperatura ambiente y se hicieron crecer durante la noche con agitación. Se recogieron las células mediante centrifugación; se resuspendieron los sedimentos bacterianos en tampón PBS y se lisaron usando una prensa francesa. Se separaron las fracciones solubles e insolubles mediante centrifugación durante treinta minutos a 10.000 RPM en un rotor SS 34, de Sorvall a 4°C.

Purificación de los cuerpos de inclusión

Se añadieron cincuenta ml de tampón de lavado de Novagen (Tris HCl 20 mM, pH 7,5, EDTA 10 mM, Triton X-100 al 1%) a la fracción insoluble, es decir, los cuerpos de inclusión (CI). Se agitó en vórtex la fracción insoluble para resuspender el sedimento. Se centrifugaron los CI y se lavaron al menos dos veces mediante resuspensión en tampón de lavado como anteriormente. Se recuperaron precipitados limpios (CI) y se almacenaron a -20°C hasta su uso.

Replegamiento de cuerpos de inclusión

Se pesaron los CI (144 mg) y se disolvieron en tampón IBS de Genotech (1,44 ml). Se incubaron los CI resuspendidos a 4°C durante una hora en un tubo para centrifuga Eppendorf. Se retiró el material insoluble mediante centrifugación a máxima velocidad en una centrifuga Eppendorf. Se diluyeron los CI solubilizados hasta un volumen final de 4 ml. Se sometió a prueba el replegamiento de GST-ST3GalIII en disoluciones tampón de replegamiento que contenían ciclodextrina, polietilenglicol (PEG), ND SB-201 o un sistema redox de GSH/GSSG. Se diluyó un ml de los CI solubilizados rápidamente mediante pipeteo en la disolución de replegamiento, se mezcló vigorosamente durante 30-40 segundos, y luego se agitó suavemente durante dos horas a 4°C. Se dializaron alícuotas de tres ml de las disoluciones de GST-ST3GalIII replegada frente a tampón PBS frío o un tampón que contenía Tris.HCl 50 mM, pH 7,0; NaCl 100 mM; y glicerol al 1% usando casetes Slide-A-lyzers de Pierce (MWCO: 3,5 kDa,). Tras la diálisis, se concentraron las disoluciones de GST-ST3GalIII 3,6 y 12 veces usando concentradores Vivaspin 5 K (VivaScience) en centrifuga Jouan a 4.000 rpm a 4°C.

Tras replegamiento y diálisis, se analizaron las proteínas GST-ST3GalIII replegadas mediante electroforesis en gel de SDS-poliacrilamida. La fusión GST-ST3GalIII, con un peso molecular de aproximadamente 63-64 kDa, estaba presente en todas las condiciones de replegamiento. (Datos no mostrados).

Sialilación de oligosacáridos usando GST-ST3 Gal III replegada

Se llevaron a cabo ensayos enzimáticos usando sustratos de oligosacárido, usando EC-FIL (electroforesis capilar-fluorescencia inducida por láser). Se sometieron a ensayo las enzimas ST3 Gal III replegadas para determinar su capacidad para transferir ácido siálico de CMP-NAN (citidina 5-monofosfato-β-D-ácido siálico) a LNnT-APTS (lacto-N-neotetraosa-ácido 9-aminopireno-1-4,6-trisulfónico) para formar LSTd-APTS (Lactosialico-tetrasacárido-d-APTS). Se realizaron las reacciones en placas de microtitulación de 96 pocillos en 100 µl de un tampón que contenía MOPS 20 mM, pH 6,5; CMP-NAN 0,8 mM; LNnT 22,1 mM; LNnT-APTS 25 µM; MnCl₂ 2,5 mM. Se iniciaron las reacciones mediante la adición de 20 µl de ST3 Gal III replegada a 30°C durante treinta minutos. Se extinguieron las reacciones con una dilución 1 a 25 con agua. Se analizó la reacción diluida mediante EC-FIL usando un capilar recubierto con N-CHO según la guía del fabricante. Se calcularon las actividades como la razón de las áreas de pico normalizadas de LNnT-APTS con respecto a LSTd-APTS. En la tabla 2, se muestran resultados que comparan diferentes condiciones de replegamiento. En la tabla 3, se muestran dos experimentos adicionales usando el sistema GSH/GSSG.

Tabla 2. Actividades de GST-ST3-Gal III tras el examen de diferentes sistemas de plegamiento. Se sometieron a ensayo las proteínas directamente sin concentración.

Ciclodextrina	PEG	ND SB-201	GSH/GSSG
0	0	0	7,8 U/l*

*Se notifican las actividades en este caso como unidades por l de enzima replegada.

Tabla 3. Actividades de GST-ST3GalIII tras dos experimentos de plegamiento separados usando el sistema GSH/GSSG.

GSH/GSSG	Conc.	Actividad
Ensayo de replegamiento 1	12x	182 U/l*
Ensayo de replegamiento 2	40x	531 U/l*

*Se notifican las actividades en este caso como unidades por l de enzima replegada.

5 Sialilación de glicoproteínas usando GST-ST3 Gal III replegada

Se añadieron veinte μ l de transferrina asialilada (2 μ g/ μ L) o factor IX asialilado (2 μ g/l), a cincuenta μ l de un tampón que contenía Tris 50 mM, pH 8,0; y NaCl 150 mM, con 10 μ l de MnCl₂ 100 mM; 10 μ l de CMP-NAN 200 mM; y azida de sodio al 0,05%. Se incubó la mezcla de reacción con 30 μ l de GST-ST3GalIII replegada a 30°C durante la noche o más con agitación a 250 rpm. Tras detenerse las reacciones, se separaron las proteínas sialiladas en IEF de pH 7-3 (gel de isoelectroenfoco, Invitrogen) y se tiñeron con azul de Coomassie según las directrices del fabricante. Se sialilaron tanto transferrina como factor IX por GST-ST3GalIII. (Datos no mostrados).

Replegamiento de una ST3GalIII de hígado de rata fusionada con una cola de MBP

Se clonó ST3GalIII de hígado de rata en vector pMAL-c2x y se expresó como fusión con proteína de unión a maltosa (MBP), MBP-ST3GalIII, en cuerpos de inclusión de células TB1 de *E. coli*. La MBP-ST3GalIII replegada era activa y transfirió ácido siálico a LNnT, un sustrato de azúcar, y a glicoproteínas asialiladas, por ejemplo asialo-transferrina.

Clonación de ST3GalIII en vector pMAL-c2x

Se clonó el ácido nucleico de ST3-GalIII de hígado de rata en sitios *Bam*H1 y *Xba*I del vector pMAL-c2x tras amplificación por PCR usando los siguientes cebadores:

Homosentido ST3BAMH1 5'-TAATGGATTCAAGCTACACTTACTCCAATGG (SEQ ID NO:70)

Antisentido: ST3XBA1 5'-GCGCTCTAGATCAGATACCACTGCTTAAGT (SEQ ID NO:71)

Se fusionaron nucleótidos que codificaban para los aminoácidos 28-374, por ejemplo, la región de tallo y el dominio catalítico de ST3GalIII, a la cola de aminoácidos de MBP.

Se construyeron tres truncamientos de ST3GalIII y se fusionaron a MBP. Se aislaron los tres insertos de ST3Gal III (Δ 73, Δ 85, Δ 86) mediante PCR usando los siguientes cebadores 5' (ST3 BamHI Δ 73) TGTATCGGATCCCTGGCCACCAAGTACGCTAACTT (SEQ ID NO:72); (ST3 BamHI Δ 85) TGTATCGGATCCTGCAAACCCGGCTACGCTTCAGCCAT (SEQ ID NO:73); y (ST3 BamHI Δ 86) TGTATCGGATCCAAACCCGGCTACGCTTCAGCCAT (SEQ ID NO:74) respectivamente, por parejas con el cebador 3' común (ST3-XhoI)-GGTCTCCTCGAGTCAGATACCACTGCTTAA (SEQ ID NO:75). Se digirió cada producto de PCR con BamHI y XhoI, se subclonó en vector pCWin2-MBP Kanr digerido con BamHI-XhoI, se transformó en células TB1, y se examinó para determinar el constructor correcto.

Se llevaron a cabo las reacciones de PCR en las siguientes condiciones. Un ciclo a 95°C durante 1 minuto. Se añadió un μ l de polimerasa Vent. Se realizaron diez de los siguientes ciclos: 94°C durante 1 minuto; 65°C durante 1 minuto; y 72°C durante 1 minuto. Tras diez minutos finales a 72°C, se enfrió la reacción hasta 4°C.

Todos los truncamientos de ST3GalIII tenían actividad tras el replegamiento. Se realizaron los experimentos descritos a continuación usando el truncamiento MBP Δ 73ST3GalIII.

Expresión de MBP-ST3GalIII en células TB1 de *E. coli*

Se transformó el plásmido pMAL-ST3GalIII en células TB1 de *E. coli* químicamente competentes. Se recogieron tres colonias aisladas que contenían el constructo TB1/pMAL-ST3GalIII de las placas de LB-agar. Se hicieron crecer las colonias en cinco ml de medios LB complementados con carbenicilina 60 μ g/ml a 37°C con agitación hasta que los cultivos líquidos alcanzaron una DO₆₂₀ de 0,7. Se retiraron dos alícuotas de un ml de cada cultivo y se usaron para inocular medios nuevos con o sin IPTG 500 μ M (final). Se hicieron crecer los cultivos a 37°C durante dos horas. se recogieron las células bacterianas mediante centrifugación. Se prepararon lisados celulares totales calentando los sedimentos celulares en presencia de SDS y DTT. IPTG indujo la expresión de MBP-ST3GalIII. (Datos no mostrados).

Expresión de MBP-ST3GalIII y purificación de los cuerpos de inclusión:

Se inoculó una alícuota de un ml de cultivo durante la noche de TB1/pMAL-ST3GalIII en 0,5 litros de medios LB con carbenicilina 50 µg/ml y se hicieron crecer hasta una DO₆₂₀ de 0,7. Se indujo la expresión de MBP-ST3GalIII mediante la adición de IPTG 0,5 mM, seguido por incubación durante la noche a temperatura ambiente. Al día siguiente, se recogieron las células bacterianas mediante centrifugación. Se resuspendieron los sedimentos celulares en un tampón que contenía TrisHCl 75 mM, pH 7,4; NaCl 100 mM; y glicerol al 1%. Se lisaron las células bacterianas usando una prensa francesa. Se separaron las fracciones solubles e insolubles mediante centrifugación durante treinta minutos, 4°C, 10.000 rpm, rotor SS 34, de Sorvall). Se separaron las fracciones solubles e insolubles mediante centrifugación durante treinta minutos a 10.000 RPM en un rotor SS 34, de Sorvall a 4°C.

Purificación de los cuerpos de inclusión y replegamiento de MBP-ST3GalIII usando GSH/GSSG

Se purificaron y se suspendieron los cuerpos de inclusión de MBP-ST3GalIII usando los mismos métodos y tampones usados para las proteínas de fusión GST-ST3GalIII descritas anteriormente. Se replegaron las MBP-ST3GalIII usando el sistema GSH/GSSG descrito anteriormente. Se dializaron las enzimas MBP-ST3GalIII frente a Tris.HCl 65 mM frío, pH 7,5, NaCl 100 mM, glicerol al 1% usando la bolsa de diálisis SnakeSkin de Pierce (MWCO: 7 kDa). Se concentraron de 3-14 veces las MBP-ST3GalIII replegadas y dializadas usando concentradores Vivaspin 5 K (VivaScience) en centrífuga Jouan a 4.000 rpm a 4°C. Se analizaron las proteínas MBP-ST3GalIII replegadas mediante electroforesis en gel de SDS-poliacrilamida. Se detectó una MBP-ST3GalIII de 81 kDa. (Datos no mostrados).

Ensayos de actividad enzimática de MBP-ST3 Gal III

Se sometieron a ensayo enzimas MBP-ST3 Gal III replegadas para determinar su capacidad para transferir ácido siálico de CMP-NAN a LNnT-APTS para formar LSTd-APTS, tal como se describió anteriormente. Las enzimas MBP-ST3 Gal III replegadas eran activas y transfirieron ácido siálico a LNnT-APTS para formar LSTd-APTS. (Datos no mostrados).

Se sometieron a ensayo enzimas MBP-ST3 Gal III replegadas para determinar su capacidad para transferir ácido siálico de CMP-NAN a glicoproteínas. Se sometió a ensayo la transferencia de ácido siálico a asialo-transferrina tal como se describió anteriormente, para enzimas GST-ST3-GalIII. Las enzimas MBP-ST3 Gal III replegadas eran activas y transfirieron ácido siálico a asialo-transferrina. (Datos no mostrados). Aunque las enzimas GST-ST3 Gal III y MBP-ST3 Gal III replegadas tenían actividades similares para la transferencia de ácido siálico a una molécula aceptora de oligosacárido soluble, las enzimas MBP-ST3 Gal III replegadas eran más activas en la transferencia de ácido siálico a una molécula aceptora de glicoproteína.

Ensayos adicionales de condiciones para el replegamiento de MBP-ST3GalIII

Se replegó MBP-ST3GalIII usando las condiciones mostradas en la figura 1. Se mezclaron el tampón, par redox y detergente (si se usó) antes de la adición de CI solubilizados para iniciar la reacción de replegamiento. Se diluyeron 1/20 los CI. El replegamiento de MBP-ST3GalIII también fue satisfactorios usándola con diferentes pares redox, por ejemplo cistamina2 HCl/cisteína a razones molares de 1/4, 4/1, 1/10 ó 5/5. (Datos no mostrados).

Ensayos de actividad enzimática de ST3 Gal III

Se sometieron a ensayo enzimas MBP-ST3 Gal III replegadas para determinar su capacidad para transferir ácido siálico de CMP-NAN a LNnT-APTS para formar LSTd-APTS, tal como se describió anteriormente. Los resultados se muestran en la figura 1. Se observaron las mayores actividades de MBP-ST3 Gal III replegadas usando las condiciones, 8, 11, 13 y 16. Cuando se amplió a escala el replegamiento hasta cinco ml, las proteínas MBP-ST3 Gal III replegadas usando las condiciones 8 y 16, tenían la mayor actividad. (Véase, por ejemplo, la tabla 4).

Tabla 4.

Condición	Proteína plegada U/l	CI U/g
8	70	37,0
6	50	40,5

Purificación de MBP-ST3GalIII en columna de amilosa

Se combinaron proteínas MBP-ST3 GalIII replegadas de la preparación de replegamiento de 5 ml y se dializaron frente a TrisHCl 100 mM, pH 7,4, NaCl 100 mM y glicerol al 1%. Se aplicaron las proteínas MBP-ST3GalIII replegadas a una columna de amilosa. La mayor parte de la proteína MBP-ST3GalIII replegada se unió a la columna de amilosa y se eluyó con maltosa 10 mM. Se muestra un perfil de elución en la figura 2. Se determinó la actividad enzimática de las fracciones de MBP-ST3GalIII usando el ensayo de LnNT y se muestra en la figura 3.

Glicopegilación de asialotransferrina con MBP-ST3GalIII replegada:

5 Se incubó asialo-transferina (2 mg/ml) con fracciones purificadas de 100 μ l de MBP-ST3GalIII replegada en presencia de CMP-SA-PEG (10 kDa, 1,6 mM) o CMP-SA-PEG (20 kDa, 1,06 mM) en 230 μ l de reacción. Se llevaron a cabo las reacciones de glicopegilación a 30°C durante la noche o durante tres días. Se retiraron alícuotas de las reacciones y se analizaron en un gel de SDS-poliacrilamida al 4-20%. Los resultados se muestran en la figura 4. MBP-ST3GalIII replegada, purificada transfiere ácido siálicos pegilados de 10 o 20 K a asialo-transferina.

Replegamiento de MBP-ST3GalIII a gran escala

Se usó el siguiente método para preparar MBP-ST3GalIII replegada a gran escala.

10 Se disolvieron CI húmedos (470 mg) en tampón de solubilización de CI (13 ml) en tubo de cultivo de 15 ml. El tampón de solubilización de CI incluye lo siguiente: guanidina HCl 4 M; TrisHCl 100 mM, pH 9; y NaCl 100 mM. Se incubaron los CI en tampón de solubilización de CI a 4°C durante aproximadamente 1 hora con agitación suave. Se eliminó cualquier material insoluble mediante centrifugación en tubos Eppendorf de 1,5 ml, a 4°C a velocidad máx., durante 30 minutos. Se transfirieron los CI solubilizados a tubos limpios y se determinó la concentración de proteína usando la absorbancia a 280 nm.

15 Se preparó la siguiente disolución de replegamiento y se mantuvo a 4°C: tampón MES 55 mM, pH 6,5; NaCl 264 mM; KCl 11 mM; PEG 550 al 0,055%; arginina 550 mM. Se complementó el tampón con lauril-maltósido (LM) 0,3 mM; glutatión oxidado (GSSG) 0,1 mM; glutatión reducido (GSH) 1 mM inmediatamente antes la adición de CI solubilizados. Se añadieron dos ml de CI solubilizados en 43 ml de tampón de replegamiento en tubo de cultivo estéril de 50 ml. Se puso el tubo en un agitador de balanceo y se agitaron suavemente durante 24 horas a 4°C.

20 Se dializó la proteína replegada en tubos de diálisis (MWCO: 7 kD) frente a tampón de diálisis (Tris HCl 100 mM, pH 7,5; NaCl 100 mM; y glicerol al 5%) dos veces (en tampón en exceso 10-20 volúmenes).

Se analizó la MBP-Gal replegada, dializada a gran escala para determinar la actividad de ST3GalIII, y mostró aproximadamente 53,6 U/g de CI.

Ejemplo 2: Mutagénesis dirigida al sitio de GnT1 humana para mejorar el replegamiento

25 Se expresó una N-acetilglucosaminiltransferasa I humana truncada (103 aminoácidos amino-terminales delecionados) en *E. coli* como proteína de fusión de unión a maltosa (GnT1/MBP). La proteína de fusión era insoluble y se expresó en cuerpos de inclusión. Tras la solubilización y el replegamiento, la proteína de fusión GnT1/MBP tenía baja actividad. La estructura cristalina de una forma truncada de GnTI de conejo (105 aminoácidos amino-terminales delecionados) muestra un residuo de cisteína desapareado (CYS123) cerca del sitio activo. (Véase, por ejemplo, Unligil *et al.*, EMBO J. 19: 5269-5280 (2000)). Se identificó la correspondiente cisteína desapareada en la GnTI humana como CYS121 y se substituyó por una serie de aminoácidos que son similares en tamaño y características químicas. Los aminoácidos usados incluyen serina (Ser), treonina (Thr), alanina (Ala) y ácido aspártico (Asp). Además, también se preparó un doble mutante, ARG120ALA, CYS121HIS. Se expresaron las proteínas de fusión GnTI/MBP mutantes en *E. coli*, se replegaron y se sometieron a ensayo para determinar la actividad de GnTI hacia las glicoproteínas.

35 Se realizó la mutagénesis usando un kit de mutagénesis dirigida al sitio Quick Change de Stratagene. Se introdujeron sitios de restricción adicionales con algunas de las mutaciones de GnTI. Por ejemplo se introdujo un sitio *Apal* (subrayado, **GGGCCAC**) en el mutante de GnTI ARG120ALA, CYS121HIS, es decir, CGC CTG → **GCC CAC** (cambios en negrita). Se usaron los siguientes oligonucleótidos mutagénicos para preparar el doble mutante: GnT1 R120A, C121H+, 5'CCGCAGCACTGTTCCGG**GCCAC**CTGGACAAGCTGCTG 3' (SEQ ID NO:76); y GnT1 R120A, C121H-5'CAGCAGCTTGCCAG**GTGGGCC**CGAACAGTGCTGCGG 3' (SEQ ID NO:77) (cambios mostrados en negrita). Se introdujo un sitio *Ascl* (subrayado, **GGCGGCC**) en el mutante de GnT1 CYS121ALA, es decir, CTG → **GCC** (cambios en negrita). Se usaron los siguientes oligonucleótidos mutagénicos para preparar el mutante de GnT1 CYS121ALA: GnT1C123A+ 5'AGCACTGTTCCGG**CGCC**CTGGACAAGCTGCTG 3' (SEQ ID NO:78); y GnT1C123A-5'CAGCAGCTTGCCAG**GGCGC**CGAACAGTGCT 3' (SEQ ID NO:79).

45 Se comparó la actividad de las proteínas mutantes expresadas en *E. coli* con la actividad de GnT1 de tipo natural expresada en baculovirus. Un mutante de GNT1 de CYS121SER era activo en un ensayo basado en CCF. En cambio, un mutante CYS121THR no tenía actividad detectable y un mutante CYS121ASP tenía baja actividad. Un mutante de CYS121ALA era muy activo, y un doble mutante, ARG120ALA, CYS121HIS, basado en la secuencia de aminoácidos de la proteína GnTI de *C. elegans* (Gly14), también mostró actividad, incluyendo la transferencia de GlcNAc a glicoproteínas. Se proporcionan las secuencias de aminoácidos y de ácido nucleico codificantes de los mutantes de GnT1 en las figuras 7-11.

55 Se preparó un segundo truncamiento de GnT1 y se fusionó a MBP: MBP-GnT1(D35). La figura 35 proporciona un esquema de las proteínas de fusión MBP-GnT1, y representa los truncamientos, por ejemplo, Δ 103 o Δ 35, y la mutación Cys121Ser (parte superior). La parte inferior de la figura proporciona la proteína GnT1 humana de longitud completa. También se realizaron mutaciones de Cys121 en la proteína MBP-GnT1(D35).

Se expresaron ambas proteínas de fusión en *E. coli* y ambas tenían actividad para la remodelación de la glicoproteína ARNasa B. La figura 36 proporciona un gel de SDS-PAGE que muestra en el panel derecho las proteínas de fusión MBP-GnT1 replegadas: MBP-GnT1(D35) C121A, MBP-GnT1(D103) R120A + C121H, y MBP-GnT1(D103) C121A. El panel izquierdo muestra las actividades para remodelación de la glicoproteína ARNasa B de dos lotes diferentes (A1 y A2) de MBP-GnT1(D35) C121A replegada en diferentes puntos de tiempo. La MBP-GnT1(D103) C121A también remodeló la glicoproteína ARNasa B. Datos no mostrados.

Ejemplo 3: Fusiones con MPB a GalT1.

Se construyeron las siguientes fusiones entre GalT1 bovina truncada y MBP: MBP-GalT1 (D129) wt, (D70) wt o (D129 C342T). (Para la secuencia bovina de longitud completa, véanse, por ejemplo, D'Agostaro *et al.*, Eur. J. Biochem. 183: 211-217 (1989) y número de registro CAA32695). Cada constructo tenía actividad tras el replegamiento. La secuencia de aminoácidos de la proteína GalT1 bovina de longitud completa se proporciona en la figura 30. Se representan esquemáticamente los mutantes en la figura 31 con una proteína de control GalT1(40) (S96A+C342T). Véase, por ejemplo, Ramakrishnan *et al.*, J. Biol. Chem. 276: 37666-37671 (2001).

Se expresó MBP-GalT1(D70) en la cepa JM109 de *E. coli*. Tras la inducción durante la noche con IPTG, se aislaron los cuerpos de inclusión del sedimento insoluble tras lisarse las células usando una prensa francesa. Se lavaron los CI dos veces y luego se solubilizaron en GndHCl 4 M, NaCl 100 mM, TrisHCl 0,1 M, pH 9.0. El replegamiento se realizó a pH 6,5 con GSSH/GSH (10/1) mediante dilución con tampón de replegamiento (1/20, proteína 0,1-0,2 mg/ml), seguido por incubación durante la noche a 4°C, sin agitación. Se dializaron las proteínas replegadas frente a TrisHCl 50 mM pH 8,0 dos veces (MWCO: 7 kD). Se cargaron las proteínas MBP-GalT1(D70) dializadas en una columna de amilosa; se lavaron; y luego se eluyeron con maltosa 10 mM.

Se sometió a ensayo la actividad GalT1 usando oligosacáridos como aceptor. Se llevaron a cabo los ensayos enzimáticos usando HPLC/PAD (cromatografía de líquidos de alta resolución con detección amperométrica pulsada). Se realizó la conversión de LNT2 (lacto-N-triosa-2) en LNnT (lacto-N-neotetraosa) usando UDP-Gal (uridina 5'-difosfogalactosa) por la enzima GalT1 tal como sigue: se llevó a cabo la reacción en 100 µl de tampón Hepes 50 mM, pH 7 que contenía UDP-Gal 6 mM, LNT-2 5 mM, MnCl₂ 5 mM y 100 µl de enzima replegada a 37°C durante 60 minutos. Se extinguió la reacción (dilución 1 a 10) con agua y se centrifugó a través de un filtro de centrifugación de MWCO de 10.000. Luego se diluyó el filtrado 1 a 10. Se analizó mediante HPLC esta reacción diluida usando un sistema Dionex DX-500 y una columna CarboPac PA1 con tampón hidróxido de sodio. Se comparó el área de pico del producto de la muestra con una curva de calibración de LNnT, y se calculó la actividad basándose en la cantidad de LNnT producida por min. por µl de enzima en la reacción.

Las proteínas MBP-GalT1(D70) purificadas tenían actividad usando tanto glicoproteínas como oligosacáridos solubles (por ejemplo, ARNasaB) como moléculas aceptoras. Los resultados se muestran en las figuras 32 y 33. En el ensayo de remodelación de ARNasa B, se comparó MBP-GalT1(D70) con una proteína de control GalT1(40) (S96A+C342T), que es un truncamiento no fusionado de la proteína GalT1 bovina que también se expresó en *E. coli* y se replegó. La proteína MBP-GalT1(D70) tenía más actividad hacia la glicoproteína ARNasa B que la que tenía GalT1(40) (S96A+C342T). El truncamiento MBP-GalT1 (D129) también tenía más actividad hacia la glicoproteína ARNasa B que la que tenía la proteína GalT1(40) (S96A+C342T). (Datos no mostrados).

Se determinó la cinética de la proteína MBP-GalT1(D70) replegada y purificada para la glicosilación de ARNasa B y se comparó con NSO GalT1, una forma soluble de la proteína GalT1 bovina que se expresó en un sistema de células de mamífero. Tal como se muestra en la figura 34, la MBP-GalT1(D70) replegada y purificada tenía una cinética mejorada en comparación con la proteína NSO GalT1.

Ejemplo 4: Método de replegamiento en un único recipiente de múltiples glicosiltransferasas.

Las enzimas ST3GalIII, GalT1 y GnT1 eucariotas construyen cadenas de N-glicano en glicoproteínas. Pueden realizarse modificaciones adicionales, por ejemplo glicopegilación, usando CMP-NAN-PEG como sustrato donador. Las enzimas ST3GalIII, GalT1 y GnT1 eucariotas se expresan normalmente en sistemas de expresión eucariotas, por ejemplo células de mamífero o fúngicas.

Se solubilizaron las enzimas ST3GalIII, GalT1 y GnT1 eucariotas fusionadas cada una a un dominio de proteína de unión a maltosa (MBP), se combinaron y se replegaron juntas en un único recipiente.

Las enzimas fusionadas a MBP y replegadas eran activas y se usaron para añadir N-glicanos a glicoproteínas o para glicopegilar glicoproteínas. El tampón de replegamiento incluía un par redox, por ejemplo, glutatión oxidado/reducido (GSH/GSSG). Se mejoró el replegamiento mediante la adición de arginina y polietilenglicol 3350 (PEG). Pueden solubilizarse los CI individualmente y añadirse al tampón de replegamiento en diferentes proporciones o solubilizarse juntos de los CI y añadirse al tampón de replegamiento directamente. La purificación en una etapa o inmovilización de estas enzimas también puede realizarse usando la cola de fusión de MBP.

Preparación de una mezcla de glicosiltransferasas replegadas (SuperGlycoMix)

Preparación de los CI de glicosiltransferasas

Se muestran las cepas bacterianas usadas para producir las enzimas ST3GalIII, GalT1 y GnT1 eucariotas en la tabla 5. La tabla también muestra el peso molecular estimado de las proteínas de fusión de MBP. (PM basado en la composición de aminoácidos, software Vector NTI). Se expresaron todos los ácidos nucleicos que codificaban para las enzimas eucariotas a partir de vectores de expresión inducibles por IPTG.

5 Tabla 5

Cepa/construido	Proteína expresada (IBS)	PM (kD)
JM109/pCWori-MBP-GalT1 (Δ 129) C342T	MBP-GalT1(Δ 129) C342T	74,2
JM109/pCWIN2-MBP-GnT1 (Δ 103) C121A	MBP-GnT1(Δ 103) C121A	82,4
TB1/pMAL-ST3GalIII	MBP-ST3GalIII	82

10 Tras la inducción con IPTG de cultivos de *E. coli*, los CI que contenían las enzimas GnT1, GalT1 y ST3GalIII aisladas lisando las células usando una prensa francesa o lisis con detergente (reactivo Bugbuster de Novagen). Se recuperaron los sedimentos tras centrifugación y se procesaron para obtener CI, tal como se describió previamente. Se lavaron los CI al menos dos veces usando el tampón de lavado de CI de Novagen. Se almacenaron los CI lavados a -20°C hasta que están listos para usarse en los experimentos de replegamiento.

15 Se disolvieron por separado CI que contenían ST3GalIII, GalT1 o GnT1 en un tampón que contenía guanidina HCl 6 M, TrisHCl 50 mM pH 8,0, EDTA 5 mM, DTT 10 mM a 4°C durante una hora. Se obtuvieron sobrenadantes aclarados tras centrifugación (velocidad máxima en microcentrífuga Eppendorf). Se determinó el contenido en proteína de los CI solubilizados midiendo la absorbancia a 280 nm. Se determinaron los contenidos en proteína de la tabla 6 basándose en los coeficientes de extinción de cada MBP-glicosiltransferasa. Se calcularon los coeficientes de extinción usando el software Vector NTi (véase la tabla 5)

Tabla 6. Concentraciones de proteína en CI solubilizados.

Proteína	A280 a 1 mg/ml	mg/ml
MBP-ST3GalIII	1,49	4,23
MBP-GalT1 (Δ 129) C342T	1,39	6,80
MBP-GnT1(Δ 103) C121A	1,7	3,29

20 Replegamiento en un único recipiente de glicosiltransferasas

Se mezclaron los CI solubilizados en cantidades iguales, tal como se muestra en la tabla 7.

Tabla 7. Se mezclaron los CI solubilizados en las siguientes cantidades antes del replegamiento.

Proteína	V (ml)	mg	% de proteína total
MBP-ST3GalIII	0,8	3,4	36
MBP-GalT1 (Δ 129) C342T	0,5	3,4	36
MBP-GnT1(Δ 103) C121A	0,8	2,6	28
Total	2,1	9,4	100

25 La concentración de proteína de la mezcla de CI solubilizados totales era de 4,5 mg/ml. Se diluyó la mezcla aproximadamente 1/20 con el tampón de replegamiento haciendo que la concentración final de la mezcla de proteína total fuera de 0,22 mg/ml. El tampón de replegamiento contenía MES 55 mM, pH 6,5; arginina 550 mM; PEG3350 al 0,055%; NaCl 264 mM; KCl 11 mM; GSH 1 mM; y GSSG 0,1 mM. El replegamiento también puede realizarse en un tampón con Tris HCl, pH 8,2 y puede sustituirse un par redox de cisteína/cistamina por GSH/GSSG. Se diluyó la mezcla de CI con el tampón de replegamiento y se incubó a 4°C durante la noche (16-18 horas).
30 Concentraciones estimadas de las glicosiltransferasas en la reacción de replegamiento:

MBP-ST3GalIII	0,081 mg/ml
MBP-GalT1 (Δ 129) C342T	0,081 mg/ml
MBP-GnT1(Δ 103) C121A	0,062 mg/ml

35 Tras replegamiento durante la noche, se dializó la mezcla de glicosiltransferasas replegadas para eliminar el agente caotrópico (es decir, guanidina HCl). Se llevó a cabo diálisis dos veces frente a TrisHCl 50 mM pH 8,0 a 4°C (20 veces por diálisis) en una bolsa de diálisis (SnakeSkin, MWCO: 7 kD, Pierce). Se concentró seis veces la mezcla de glicosiltransferasas replegadas dializada (SuperGlycoMix, SGM) usando concentradores centrifugos VivaSpin de 6

ml (MWCO: 10 kD). Tras concentración, las tres glicoproteínas estaban presentes en la mezcla, tal como se determinó mediante análisis por SDS-PAGE. (Datos no mostrados). Tras concentrar la SGM, se determinaron las actividades enzimáticas de GnT1, GalT1 y ST3GalIII.

Actividades enzimáticas de SuperGlycoMix

- 5 SuperGlycoMix (SGM), la mezcla de glicosiltransferasas replegadas en un único recipiente contenía tres glicosiltransferasas: ST3GalIII, GalT1 y GnT1. Estas enzimas se sometieron a ensayo individualmente para determinar sus actividades enzimáticas y se analizaron usando los métodos indicados a continuación. Se enumeran las actividades enzimáticas en la tabla 8.

Ensayos de actividad enzimática de ST3 Gal III

- 10 Se llevaron a cabo ensayos de ST3GalIII usando HPLC/UV (cromatografía de líquidos de alta resolución con detección ultravioleta). Se realizó la conversión de LNnT (lacto-N-neotetraosa) en LSTd (lactosialico-tetrasacárido-d) usando CMP-NAN (citidina 5'-monofosfato-β-D-ácido siálico) por la enzima ST3GalIII tal como sigue. Se llevó a cabo la reacción en una placa de microtitulación de 96 pocillos en 100 μl de tampón MOPS 20 mM, pH 6,5 que contenía CMP-NAN 2 mM, LNnT 30 mM, MnCl₂ 10 mM y 20 μl de enzima replegada a 30°C durante 120 minutos. Se
15 extinguió la reacción calentando hasta 98°C durante 1 min. Se centrifugó la placa de microtitulación a 3600 rpm durante 10 min. para sedimentar cualquier precipitado. Se diluyeron 1:1 75 μl de sobrenadante con 75 μl de agua. Se analizó la reacción diluida mediante CL/UV usando una columna de poliamina II YMC-Pack con un gradiente de tampón fosfato de sodio/acetonitrilo y detección a 200 nm. Se comparó el área de pico del producto de la muestra con una curva de calibración de LSTd, y se calculó la actividad basándose en la cantidad de LSTd producida por
20 min. por μl de enzima en la reacción.

Ensayos de actividad enzimática de GalTI:

- Se llevaron a cabo los ensayos enzimáticos usando HPLC/PAD (cromatografía de líquidos de alta resolución con detección amperométrica pulsada). Se realizó la conversión de LNT2 (lacto-N-triosa-2) en LNnT (lacto-N-neotetraosa) usando UDP-Gal (uridina 5'-difosfogalactosa) por la enzima GalTI tal como sigue. Se llevó a cabo la
25 reacción en 100 μl de tampón Hepes 50 mM, pH 7 que contenía UDP-Gal 6 mM, LNT-2 5 mM, MnCl₂ 5 mM y 100 μl de enzima replegada a 37°C durante 60 minutos. Se extinguió la reacción (dilución 1 a 10) con agua y se centrifugó a través de un filtro de centrifugación de MWCO de 10.000. luego se diluyó el filtrado 1 a 10. Se analizó esta reacción diluida mediante HPLC usando un sistema Dionex DX-500 y una columna CarboPac PA1 con tampón hidróxido de sodio. Se comparó el área de pico del producto de la muestra con una curva de calibración de LNnT, y se calculó la actividad basándose en la cantidad de LNnT producida por min. por μl de enzima en la reacción.

- 30

Ensayos de actividad enzimática de GnTI:

- Se determina la actividad de GnTI midiendo la transferencia de un azúcar tritiado de UDP-3H-GlcNAc (uridina difosfato N-acetil-D-glucosamina [6-³H(N)]) a n-octil-3,6-di-O-(α-manopiranosil)β-D-manopiranosido (OM3), un núcleo de trimanosilo con una cola de octilo. Se llevó a cabo la reacción en 20 μl de tampón MES 100 mM, pH 6,0 que
35 contenía UDP-GlcNAc 3 mM, UDP-³H-GlcNAc 0,1 mM, OM3 0,5 mM, MnCl₂ 20 mM y 10 μl de enzima replegada a 37°C durante 60 minutos. Se extinguió la reacción (dilución 1 a 6) con agua y se aplicó a una resina de fase inversa polimérica en un formato de 96 pocillos que se acondicionó previamente según las recomendaciones del fabricante. Se lavó la resina dos veces con 200 μl de agua y se eluyó el producto con 50 μl de MeOH al 100% en una placa de captura. Se añadió líquido de centelleo (200 μl) a cada pocillo y se mezcló la placa y se contó usando un contador de centelleo de microplacas TopCount NXT de PerkinElmer. Se calculó la actividad basándose en la cantidad de ³H-GlcNAc incorporado en el producto por min. por μl de enzima en la reacción.

- 40

Tabla 8. Actividades enzimáticas de las glicosiltransferasas replegadas en SGM

Actividad enzimática	mU/ml
GnT1	1
GalT1	165
ST3GalIII	10

- 45 Las actividades notificadas en la tabla anterior son próximas o están en el intervalo en el que se replegaron estas enzimas por separado. Las actividades de GnT1 y GalT1 son próximas a las obtenidas usando sistemas de expresión de mamífero o baculovirus. Las actividades de ST3GalIII son algo menores que en la preparación de ST3GalIII obtenida tras el sistema de expresión fúngico. El ensayo de ST3GalIII usado en este caso se modifica del procedimiento y los valores notificados en este caso son aproximadamente 4-5 veces menores que los obtenidos con un método basado en EC-FIL (electroforesis capilar-fluorescencia inducida por láser).

- 50 Remodelación de ARNasaB-Man₅ usando SuperGlycoMix

5 Se remodeló una pequeña glicoproteína, ARNasaB con un azúcar Man₅ unido a N, mediante SGM en presencia de UDP-azúcares (UDP-GlcNAc y UDP-Gal). Se llevó a cabo la reacción de remodelación o bien usando UDP-GlcNAc o tanto UDP-GlcNAc como UDP-Gal para someter a pruebas las actividades tanto de GnT1 como de GalT1. Se añadieron ocho µl de SGM a tampón MES 10 mM pH 6,5 que contenía UDP-GlcNAc 5 mM o/y UDP-Gal 5 mM, 9 µg de ARNasaBMan₅, MnCl₂ 5 mM en 25 µl de ensayo incubado a 33°C para durante la noche hasta 48 horas. Al final de la reacción, se dializaron alícuotas de diez µl frente a H₂O y se pusieron en placas para MALDI-TOP muestras de 1,5 µl. Se analizaron las muestras en MALDI-TOF tras tratarse con TFA y ácido cinapínico.

10 Se realizó la remodelación de ARNasaBman₅ transfiriendo GlcNAc y Gal en Man₅ de la ARNasaB. Tras incubación durante 48 h a 33°C, se logró la mayoría de la transferencia de GlcNAc y Gal sobre ARNasaB tal como se indicó en los espectros de MALDI-TOF de la ARNasaBMan₅ remodelada. Los resultados se resumen en la tabla 9.

Tabla 9. Espectros de MALDI-TOF de la especie tras reacciones con SGM.

Reacción	Man ₅	m/z ARNasaB	
		Man ₅ -GlcNAc	Man ₅ GlcNAc-Gal
Sin enzima	14983	-	-
SGM+UDP-GlcNAc	14973	15177	-
SGM+ UDP-GlcNAc +UDP-Gal	14982	15170	15348

15 Se llevó a cabo la glicopegilación (20 K) en una reacción en un único recipiente compuesta por los siguientes componentes: MES 10 mM pH 6,5, MgCl₂ 5 mM, UDP-GlcNAc 5 mM, UDP-GalNAc 5 mM, CMP-SA-PEG 0,5 mM (20 kDa), 24 µg de EPO, 8 µl de SGM concentrada. En reacciones de control, se substituyó SGM por enzimas individuales o bien replegadas o bien expresadas en células de mamífero o células de insecto o *Aspergillus*. Tras incubaciones durante la noche, se analizaron las reacciones en gel de SDS-poliacrilamida. Los resultados se muestran en la figura 5. SGM añadió PEG de 20K a EPO.

Evaluación de las condiciones de replegamiento en un único recipiente para múltiples glicosiltransferasas

20 Se evaluaron las condiciones para el replegamiento de múltiples glicosiltransferasas, incluyendo pH y replegamiento de dos o tres enzimas a la vez.

Preparación de cuerpos de inclusión de glicosiltransferasas

25 Se describieron previamente cepas de *E. coli* transformadas con plásmidos de expresión de glicosiltransferasas, con una excepción. MBP-ST3GalIII se expresó en células JM109 a partir de un plásmido pCWori-ST3GalIII. Se aislaron los cuerpos de inclusión y se solubilizaron tal como se describió anteriormente. Se evaluaron los contenidos en proteína tal como se describió anteriormente y se muestran en la tabla 10.

Tabla 10. Se mezclaron los CI solubilizados en las siguientes cantidades antes del replegamiento.

Proteína	A280	A280 (a 1 mg/ml)	mg	% (de proteína sol.)
MBP-ST3GalIII	32,3	1,49	21,7	13,6
MBP-GalT1(Δ129) C342T	35,7	1,39	25,7	13,7
MBP-GnT1(Δ103) C121S	42,8	1,7	25,2	9,7

Replegamiento en un único recipiente de mezclas de CI de glicosiltransferasas

30 Tras determinar sus contenidos en proteína, se mezclaron los CI solubilizados en las cantidades mostradas antes de diluirse con los tampones de replegamiento (tabla 11). Se llevaron a cabo experimentos de replegamiento de las GT en un volumen de 44 ml a 4°C en fase estacionario usando tampón A o B (a continuación) y GSSG 0,1 mM y GSH 1 mM. Tampón A: MES 55 mM pH 6,5, arginina 550 mM, PEG3350 al 0,055%, NaCl 264 mM, KCl 11 mM, complementado con GSH 1 mM, GSSG 0,1 mM. Tampón B: TrisHCl 55 mM pH 8, arginina 550 mM, PEG3350 al 0,055%, NaCl 264 mM, KCl 11 mM, complementado con GSH 1 mM, GSSG 0,1 mM.

Tabla 11. Cantidades de mezclado de CI de GT solubilizados en 2 ml de IBSB

Replegamiento en tampón A

Replegamiento 1 (A-2x)			
	Conc. (mg/ml)	V (ml)	mg
MBP-GnT1 (Δ 103) C121S	25,2	0,2	5
MBP- GalT1 (Δ 129) C342T	25,7	0,2	5
IBSB	-	1,6	-
Replegamiento 2 (A-3x)			
	Conc. (mg/ml)	V (ml)	mg
MBP-GnT1 (Δ 103) C121S	25,2	0,2	5
MBP- GalT1 (Δ 129) C342T	25,7	0,2	5
MBP-ST3GalIII	21,7	0,4	8,7
IBSB	-	1,2	-

Replegamiento en tampón B

Replegamiento 3 (B-2x)			
	Conc. (mg/ml)	V (ml)	mg
MBP-GnT1 (Δ 103) C121S	25,2	0,2	5
MBP-GalT1 (Δ 129) C342T	25,7	0,2	5
IBSB	-	1,4	-
Replegamiento 4 (B-3x)			
	Conc. (mg/ml)	V (ml)	mg
MBP-GnT1 (Δ 103) C121S	25,2	0,2	5
MBP- GalT1 (Δ 129) C342T	25,7	0,2	5
MBP-ST3GalIII	21,7	0,4	8,7
IBSB	-	1,2	-

5 Para el doble replegamiento (2x, dos glicosiltransferasas) se añadieron 10 mg de proteína total en 2 ml, en 41 ml de tampón de replegamiento (anteriormente) 0,45 ml de GSH 100 mM, 0,45 ml de GSSG 10 mM, tras dilución, la proteína total era 0,44 mg/ml. Para el triple replegamiento de (3x, tres glicosiltransferasas) se añadieron 18,7 mg de proteína total en 2 ml, en 41 ml de tampón de replegamiento (anteriormente), 0,45 ml de GSH 100 mM, 0,45 ml de GSSG 10 mM. Tras dilución, la proteína total era 0,83 mg/ml. Las concentraciones de proteína eran mayores que en el experimento previo de triple replegamiento (0,22 mg/ml en SGM). Siguen las concentraciones estimadas de las glicosiltransferasas en la reacción de replegamiento:

10	MBP-ST3GalIII	0,39 mg/ml
	MBP-GalT1 (Δ 129) C342T	0,23 mg/ml
	MBP-GnT1(Δ 103) C121S	0,23 mg/ml

Tras el replegamiento durante la noche, se dializó la mezcla de glicosiltransferasas replegadas.

15 Se llevó a cabo la diálisis dos veces frente a TrisHCl 50 mM pH 8,0 a 4°C en una bolsa diálisis (SnakeSkin, MWCO: 7 kD, Pierce). Tras la diálisis, se concentró 9-12 veces la mezcla de glicosiltransferasas usando 6 ml de concentradores centrífugos VIVA-Spin (MWCO: 10 K).

El análisis por SDS-PAGE demostró que las proteínas estaban presentes tras el replegamiento, diálisis y concentración.

Ensayos enzimáticos de mezclas de glicosiltransferasas replegadas

20 Se realizaron ensayos enzimáticos tal como se describió anteriormente. Los resultados se muestran en la tabla 12.

Tabla 12. Actividades enzimáticas de las glicosiltransferasas replegadas tras experimentos de doble y triple replegamiento.

Replegamiento	Conc. en veces	Actividad enzimática	mU/ml
Tampón A (A-2x)		GnT1	0,84
		GalT1	598
Tampón A (A-3x)		GnT1	0,16
		GalT1	306
		ST3GalIII	4
Tampón B (B-2x)		GnT1	3,32
		GalT1	747

Tampón B (B-3x)

GnT1	0,47
GalT1	425
ST3GalIII	11

5 Se observó la mayor actividad al mezclar GnT1 y GalT1 fusionadas con MBP en cantidades iguales y replegadas en tampón B. La adición de una cantidad no equivalente de ST3GalIII fusionada con MBP afectó a la eficacia de replegamiento debido al alto contenido en proteína total. No obstante, pueden usarse dos tampón de replegamiento diferentes usando o bien dos GT o bien tres GT, para obtener proteínas solubles activas.

Ejemplo 5: Replegamiento de GalNAcT2 eucariota.

10 Se expresó una enzima GalNAcT2 humana truncada en *E. coli* y se usó para determinar condiciones óptimas para la solubilización y el replegamiento usando los métodos descritos anteriormente. Las secuencias de aminoácidos y ácido nucleico de GalNAcT2 humana de longitud completa se proporcionan en las figuras 13A y B. Las secuencias de la proteína mutante, GalNAcT2(D51), se muestran en las figuras 14A y B. Se expresó el mutante en *E. coli* como una proteína de fusión con MBP, MBP-GalNAcT2(D51). Se prepararon otros mutantes de GalNAcT2, se expresaron en *E. coli* y pudieron replegarse: MBP-GalNAcT2(D40), MBP-GalNAcT2(D73), y MBP-GalNAcT2(D94). Datos no mostrados. Se encuentran detalles de la construcción de los mutantes de delección adicionales en el documento USSN 60/576.530, presentado el 3 de junio de 2004 y el documento USSN 60/598.584, 3 de agosto de 2004, ambos de los cuales se incorporan al presente documento como referencia para todos los fines.

15 Se hicieron crecer cultivos de bacterias que expresaban MBP-GalNAcT2(D51) y se recogieron tal como se describió anteriormente. Se purificaron cuerpos de inclusión a partir bacterias tal como se describió anteriormente. Se realizó la solubilización de los cuerpos de inclusión a pH 6,5 o a pH 8,0. Tras la solubilización, se replegó la proteína MBP-GalNAcT2(D51) o bien a pH 6,5 o bien a pH 8,0 usando tampones A y B, es decir, tampón A: MES 55 mM pH 6,5, arginina 550 mM, PEG3350 al 0,055%, NaCl 264 mM, KCl 11 mM, complementado con GSH 1 mM, GSSG 0,1 mM; y tampón B: TrisHCl 55 mM pH 8,5, arginina 550 mM, PEG3350 al 0,055%, NaCl 264 mM, KCl 11 mM, complementado con GSH 1 mM, GSSG 0,1 mM. Tras el replegamiento, se dializó la proteína MBP-GalNAcT2(D51) y luego se concentró. La figura 15 proporciona una demostración de la concentración de proteína de MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y el replegamiento a pH 6,5 o pH 8,0.

20 Se realizó un ensayo de [³H]-UDP-GalNAc radiomarcado para determinar la actividad de la MBP-GalNAcT2(D51) replegada expresada en *E. coli* monitorizando la adición de GalNAc radiomarcado a un péptido aceptor. El aceptor fue un péptido similar a MuC-2 que tenía la secuencia MVTPTPTPTC (SEQ ID NO: 80). Se disolvió el péptido en Tris-HCl 1 M pH=8,0. Véanse, por ejemplo, el documento USSN 60/576.530 presentado el 3 de junio de 2004; y la solicitud de patente provisional estadounidense con número de expediente del apoderado 040853-01-5149-P1, presentada el 3 de agosto de 2004; ambos de los cuales se incorporan al presente documento como referencia para todos los fines. La figura 16 proporciona una demostración de la actividad enzimática de MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y el replegamiento a pH 6,5 o pH 8,0. La figura 17 proporciona una demostración de la actividad específica de MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y el replegamiento a pH 6,5 o pH 8,0. Los mayores niveles de actividad se observaron con MBP-GalNAcT2(D51) que se había solubilizado a pH 8,0 y replegado a pH 8,0. Los mayores niveles de actividad específica también se observaron con una solubilización a pH 8,0 y un replegamiento a pH 8,0.

25 Se sometió a ensayo MBP-GalNAcT2(D51) solubilizada y replegada para determinar su capacidad para añadir GalNAc a la proteína G-CSF. El ensayo consistió en una alícuota de enzima y un tampón de reacción (MES 27 mM, pH=7, NaCl 200 mM, MgCl₂ 20 mM, MnCl₂ 20 mM, y Tween 80 al 0,1%), proteína G-CSF (2 mg/ml en H₂O), y UDP-GalNAc 100 mM. Para cada muestra replegada, se añadieron 4,4 µl de muestra a 15 µl de disolución de reacción. Para el control positivo, se añadió 1 µl de patrón de GalNAcT2-baculovirus junto con 3,4 µl de H₂O en un tubo. Se incubaron las reacciones a 32°C en un agitador rotatorio durante varios días, tiempo durante el cual se sometió a ensayo un punto de tiempo durante la noche y un punto de tiempo de 5 días mediante MALDI. Véanse, por ejemplo, el documento USSN 60/576.530 presentado el 3 de junio de 2004; y la solicitud de patente provisional estadounidense con número de expediente del apoderado 040853-01-5149-P1, presentada el 3 de agosto de 2004; ambos de los cuales se incorporan al presente documento como referencia para todos los fines.

30 Las figuras 18A y 18B proporcionan resultados de la remodelación de factor estimulante de colonias de granulocitos (G-CSF) recombinante usando MBP-GalNAcT2(D51) replegada tras la solubilización a pH 6,5 o pH 8,0 y el replegamiento a pH 6,5 o pH 8,0. Se incluyeron un control positivo, es decir, MBP-GalNAcT2(D51) purificada que se había expresado en baculovirus, y un control negativo, es decir, mezcla de reacción que carecía de un sustrato. Los mayores niveles de actividad de remodelación de G-CSF se observaron usando MBP-GalNAcT2(D51) que se había solubilizado a pH 8,0 y replegado a pH 8,0.

Ejemplo 6: Replegamiento y purificación de GalNAcT2 eucariota.

35 Se hicieron crecer cuatro litros de bacterias que expresaban MBP-GalNAcT2(D51) recombinante y se recogieron. Se aislaron cuerpos de inclusión, se lavaron, y se solubilizó un peso en seco de dos gramos de cuerpos de inclusión a

4°C en 200 ml de tampón de solubilización (urea 7 M/Tris 50 mM/DTT 10 mM/EDTA 5 mM a pH 8,0). Tras la solubilización, luego se diluyó la mezcla en 4 l de tampón de replegamiento (Tris 50 mM/ L-arginina 550 mM/ NaCl 250 mM/ KCl 10 mM/ PEG 3350 al 0,05%/ L-cisteína 4 mM/ diclorhidrato de cistamina 1 mM a pH 8,0). Se llevó a cabo el replegamiento a 4-10°C durante aproximadamente 20 horas, con agitación. Luego se filtró la mezcla usando un filtro 10SP CUNO, se concentró 5 veces sobre una membrana de 4 ft², se diafiltró 4 veces con Tris 10 mM/NaCl 5 mM a pH 8,0. La conductividad de la disolución de MBP-GalNAcT2(D51) replegada final fue de 1,4 mS/cm. La proteína replegada se almacenó a 4°C durante varios días.

Se aplicaron las proteínas replegadas a una columna Q Sepharose XL (QXL) (Amersham Biosciences, Piscataway, NJ). Se muestra un perfil de elución en la figura 19 y la actividad enzimática de fracciones de columna específicas se muestra en la figura 20. Se combinaron las fracciones activas y se aplicaron a una columna de hidroxiapatita tipo I (80 µm) (BioRad, Hercules, CA). Se muestra un perfil de elución en la figura 21 y se muestra la actividad de fracciones eluidas con HA tipo I en la figura 22. La combinación de cromatografía en QXL y HA tipo I dio como resultado MBP-GalNAcT2(D51) activa altamente purificada.

Ejemplo 7: Purificación de MBP-SBD ST3Gal3 eucariota.

Se fusionó en marco el truncamiento Δ73 ST3GalIII con la proteína de unión a maltosa y un dominio de unión a almidón para formar una proteína con doble cola: MBP-SBD-ST3Gal3 (Δ73). La MBP estaba en el extremo amino-terminal, seguido por la SBD, y luego la proteína ST3GalIII truncada. Se compararon el replegamiento y la purificación de la proteína MBP-SBD-ST3Gal3 (Δ73) con una proteína de una única cola: MBP-ST3GalIII (Δ73). Se expresaron ambas proteínas en *E. coli* como cuerpos de inclusión insolubles, y se solubilizaron y replegaron tal como se describe en el presente documento. Luego se dializaron las proteínas y se sometieron a purificación por afinidad usando una columna de ciclodextrina que se une a las colas tanto de MBP como de SBD. Los resultados se muestran en la figura 23. La proteína MBP-SBD-ST3Gal3 (Δ73) tenía una actividad específica superior tras la diálisis y conservó más actividad específica tras la elución de la columna con ciclodextrina.

Ejemplo 8: Replegamiento de proteínas MBP-ST3Gal1 y MBPSBD-ST3Gal1

Se fusionó ST3Gal1 eucariota con MPB o MBP y SBD. Se usó la secuencia de ADN del gen porcino de ST3Gal1 como molde para el diseño de los constructos pcWINMBP-pST3Gal1 y pcWINMBP/SBD-pST3Gal1 descritos en el presente documento. La secuencia porcina de ST3Gal1 de longitud completa se proporciona en la figura 37. Para la expresión en *E. coli*, se usó una versión truncada y de codón optimizado de pST3Gal1, es decir, pST3Gal1 Δ45. Las secuencias de aminoácidos codificadas se proporcionan en la figura 24. A continuación se digirió la secuencia codificante del gen de ST3Gal1 y se transfirió a vectores pcWIN2-MBP y pcWINMBP/SBD usando los sitios de clonación BamHI y XhoI. Se confirmó que estos constructos eran correctos mediante análisis de restricción y de secuencia y luego se usaron para transformar la cepa JM109 de *E. coli* usando selección con kanamicina 50 µg/ml. Se usó una colonia individual de cada uno para inocular un cultivo de 2 ml de Maritone-kanamicina 10 µg/ml que se incubó durante 16 h a 37°C. Se mezcló cada cultivo por separado 1:1 con glicerol al 50% y se congeló a -80°C y se denominó vial de disolución madre. Se usó una pequeña cantidad de cada vial de disolución madre para trazar una línea en una placa de Maritone-Kan. Tras una incubación de 16 h a 37°C, se usó una única colonia de cada uno para inocular un cultivo de 25 ml de Maritone-kanamicina 10 µg/ml que se incubó durante 16 h a 37°C. Entonces se usó el cultivo de 25 ml para inocular un cultivo de 1 l de Maritone-kanamicina 10 µg/ml que se incubó a 37°C y se monitorizó para determinar la DO600. Cuando la DO600 alcanzó 0,8, se añadió IPTG a 1 mM y se incubaron las células durante 16 h adicionales. Luego se recogieron las células mediante centrifugación a 7000xG durante 15 min.

Luego se aislaron los cuerpos de inclusión, se solubilizaron las proteínas de fusión ST3Gal1 y se replegaron. Se resuspendieron los sedimentos de células bacterianas a una razón de 1 g de sedimento celular en húmedo por 10 ml de Tris 20 mM pH 8, EDTA 5 mM y se sometieron a lisis mediante alteración mecánica con dos pasos a través de un microfluidizador. Se sedimentó el material insoluble, es decir, los cuerpos de inclusión o CI, mediante centrifugación a 7000xg a 4°C en el instrumento Sorvall RC3 durante 30 minutos, y se desechó el sobrenadante. Un ciclo de lavado típico consistió en resuspender completamente el sedimento en el tampón de lavado, y repetir la centrifugación durante 15 minutos. Se lavó el sedimento una vez en un volumen en exceso (al menos 10 ml (hasta 20) por g de sedimento celular original) de tampón con alto contenido en sal (lavado I: Tris 10 mM pH 7,4, NaCl 1 M, EDTA 5 mM), una vez en tampón detergente (lavado II: Tris 25 mM pH 8, NaCl 100 mM, Triton X100 al 1%, Nadesoxicolato al 1%, EDTA 5 mM), y (para eliminar restos de detergente además del lavado de CI) tres veces en tampón de lavado (lavado III: Tris 10 mM pH 8, EDTA 5 mM).

Se resuspendieron los CI lavados en tampón de solubilización (urea 8 M, BisTris 50 mM pH 6,5, EDTA 5 mM, DTT 10 mM) y se ajustaron hasta una concentración de proteína de 2 mg/ml. Se realizó el replegamiento como una dilución rápida de 20 veces en tampón de replegamiento (Tris 55 mM pH 8,2, NaCl 10,56 mM, KCl 0,44 mM, MgCl₂ 2,2 mM, CaCl₂ 2,2 mM, Peg3350 al 0,055%, L-arginina 550 mM, GSH 1 mM, GSSG 100 mcM) y se agitó durante 16 h a 4°C. Luego se cambió el tampón por desalación usando G50 Sephadex a BisTris 50 mM pH 6,5, NaCl 75 mM, Tween 80 al 0,05%.

Para determinar si las enzimas replegadas eran activas, se realizó un ensayo de sialiltransferasa. Se monitorizó la transferencia de ácido siálico al donador (mucina submaxilar asialo-bovina) usando CMP-NAN radiomarcada. Se usó ST6GalNAcl de gallina expresada en un sistema de baculovirus como control positivo.

- 5 En resumen, se añadieron 40 μl de la mezcla de reacción a 10 μl de muestra de enzima y se incubaron a 37°C durante 1 hora. Se precipitó la glicoproteína añadiendo 100 μl de ácido fosfotungstácico/TCA al 15% a la reacción con mezclado. Tras la centrifugación, se aspiraron y desecharon los sobrenadantes. Se añadieron quinientos μl de TCA al 5% para lavar CMP-ácido siálico ^{14}C no incorporado del sedimento. Volvieron a centrifugarse las reacciones y se aspiraron y desecharon los sobrenadantes. Se resuspendieron los sedimentos usando 100 μl de NaOH 10 N. Se añadió un ml de tampón Tris 1 M, pH 7,5 al sedimento resuspendido y luego se transfirió la mezcla a un vial de centelleo. Se añadieron cinco ml de fluido de centelleo (Ecolume, ICN Biomedicals) y se mezclaron bien. Se determinaron las cuentas totales añadidas a una reacción añadiendo 40 μl de mezcla de reacción en un vial de centelleo y añadiendo 100 μl de NaOH 10 N, 1 ml de agua, y 5 ml de fluido de centelleo, y mezclando bien. Se contaron los viales durante 1 minuto. Las condiciones de reacción se proporcionan en la tabla 13.

TABLA 13

Reactivo	Fabricante	N.º cat	Cantidad por tubo de ensayo	Concentración final en el ensayo
CMP-ácido siálico [^{14}C]	Amersham	CFB-165	4 μl	100.000 CPM
CMP-ácido siálico 7 mM	Neose	AES 533pg.42	1,4 μl	0,2 mM
Mucina submaxilar asialo-bovina 25 mg/ml	Sigma, hidrolizado en Neose	M-3895	20 μl	250-500 μg
Tampón Bis/Tris 1M, pH 6,5	Sigma	B-9754	2,5 μl	50 mM
NaCl 5M	Sigma	S-7653	1 μl	100 mM
dH ₂ O	N/A	N/A	11,1 μl	

- 15 Volumen de mezcla de reacción total por tubo de ensayo: 40 μl
- Los resultados se muestran en la figura 25. Ambas proteínas de fusión replegadas, MBP-pST3Gal1 y MBP-SBD-pST3Gal1, tenían actividad de sialiltransferasa detectable.

Ejemplo 9: Replegamiento de proteínas MBP-ST6GalNAcl

- 20 Se fusionó ST6GalNAcl I eucariota con MPB. En resumen, se generaron cinco constructos de ST6GalNAcl I de ratón: D32, E52, S127, S186, y S201. Cada constructo se expresó detrás de la cola de MBP a partir del vector pcWin2-MBP, y se diferencian en la extensión de la región de 'tallo' incluida en el constructo. D32 es la forma más larga, comenzando inmediatamente en sentido 3' del dominio transmembrana amino-terminal predicho. S201 es la más corta, que comienza poco antes del inicio predicho del dominio catalítico conservado.
- 25 Además de los constructos de ratón, también se expresó ST6GalNAcl I K36 humana como fusión con MBP. El constructo humano comienza justo tras el dominio transmembrana. Se aisló ADN que codifica para ST6GalNAcl1 humana desde K36 hasta su extremo C-terminal mediante PCR usando el vector de expresión de baculovirus existente como molde, y se clonó en los sitios de BamHI-XhoI dentro de pcWin2-MBP.
- 30 Para referencia, se incluyen las secuencias para MBP-mST6GalNAcl S127 y MBP-hST6GalNAcl K36 en la figura 26. Además, la figura 38 proporciona secuencias de aminoácidos de longitud completa para ST6GalNAcl humana y para ST6GalNAcl de gallina, y una secuencia de la proteína ST6GalNAcl de ratón que comienza en el residuo 32 de la proteína de ratón nativa.
- 35 Se han preparado mutantes de delección adicionales a los descritos anteriormente y se encuentra una lista completa de ST6GalNAcl preferidas para su uso en la invención en la tabla 14. La figura 39 proporciona un esquema de varios mutantes de truncamiento de ST6GalNAcl humana preferidos. La figura 40 muestra un esquema de proteínas de fusión de MBP incluyendo los mutantes de truncamiento de ST6GalNAcl humana.

Tabla 14: Mutantes de ST6GalNAcl

	Sitio de truncamiento	Mutación
HUMANA	Δ35	K36
	Δ124	K125
	Δ257	S258
	Δ35	K36
	Δ72	T73
	Δ109	E110
	Δ133	M134
	Δ170	T171
	Δ232	A233
	Δ272	G273
GALLINA		
	Δ48	Q49
	Δ152	V153
	Δ225	L226
	Δ232	T233
RATÓN	Δ31	D32
	Δ51	E52
	Δ126	S127
	Δ185	S186
	Δ200	S201

La figura 45 muestra la posición de residuos de cisteína apareados y desapareados en la proteína ST6GalNAcI humana. También se muestran sustituciones simples y dobles de cisteína, por ejemplo, C280S, C362S, C362T, (C280S + C362S), y (C280S + C362T).

5 Estudios de expresión iniciales mostraron que las proteínas de fusión ST6GalNAcI se expresaban como proteínas insolubles. Para recuperar la enzima recombinante activa, se aislaron las proteínas insolubles inactivas y se replegaron tal como se describe:

10 Se indujeron cultivos de 0,5 l en crecimiento logarítmico de células JM109 que llevaban o bien pcWin2-MBP-mST6GalNAcI D32, E52, S127, S186 o bien pcWin2-MBP-hST6GalNAcI K36 con IPTG 1 mM durante la noche a 37°C. Se recogieron células mediante centrifugación, y se lisaron mediante alteración mecánica en un microfluidizador en 100 ml de Tris 20 mM pH 8, EDTA 5 mM. Se recogió la materia insoluble mediante centrifugación a 7000 x g durante 20 minutos. Se desecharon los sobrenadantes, y se lavaron los sedimentos con un tampón con alto contenido en sal (Tris 20 mM pH 7,4, NaCl 1 M, EDTA 5 mM), tampón detergente (Tris 25 mM pH 8, Na-desoxicolato al 1%, Triton x100 al 1%, NaCl 100 mM, EDTA 5 mM), y TE (Tris 10 mM pH 8, EDTA 1 mM).
 15 Cada lavado fue en 100 ml, y se recogió el sedimento mediante centrifugación tal como se describió anteriormente. Tras el lavado, se tomaron alícuotas de sedimentos de cuerpos de inclusión y se almacenaron a -80°C.

- Para seleccionar condiciones que permitieran un replegamiento apropiado, y por tanto una recuperación de la actividad de ST6GalNAc I, se solubilizaron alícuotas de los cuerpos de inclusión de la proteína de fusión ST6GalNAcI de ratón y humana en guanidina 6 M, DTT 10 mM, 1x TBS. Se normalizó la concentración de proteína mediante un ensayo de Bradford, y se transfirieron las proteínas solubilizadas a una serie de tampones de replegamiento de proteínas comercialmente disponibles. Se llevaron a cabo replegamientos en 0,25 ml a 0,2 mg/ml durante la noche a 4°C en una placa de 96 pocillos con agitación. Se transfirieron los replegamientos a una palca de diálisis de 96 pocillos (25000 MWCO) y se sometieron a diálisis frente a 1x TBS, Tween-80 al 0,05% durante cuatro horas a 4°C, seguido por una diálisis durante la noche frente a BisTris 10 mM pH 7,1, NaCl 100 mM, Tween-80 al 0,05% a 4°C.
- Se sometieron a prueba proteínas de fusión ST6GalNAcI replegadas recombinantes para determinar la actividad en un ensayo de actividad en fase sólida de 384 pocillos. En resumen, el ensayo de actividad detecta la transferencia mediada por ST6GalNAcI de un ácido siálico biotinilado desde CMP-NAN biotinilado hacia la superficie de un pocillo recubierto con mucina submaxilar asialo-bovina en una placa de 384 pocillos. Cada reacción (13,5 µl de replegamiento + 1,5 µl de 10x tampón de reacción) se realizó por cuadruplicado. 10x tampón de reacción fue BisTris 0,2 M pH 6,7, MgCl₂ 25 mM, MnCl₂ 25 mM, Tween-80 al 0,5%, y donador 1 mM. Tras la incubación durante la noche a 37°C, se lavó la placa con un exceso de 1x TBS, Tween-20 al 0,05%, y se detectó la biotina con estreptavidina marcada con europio según las instrucciones del fabricante (Perkin Elmer). Se documentaron niveles de fluorescencia de europio conservados en la placa, indicativo de actividad de ST6GalNAcI, con un lector de placas Victor3V de Perkin Elmer, y los resultados de expresión y actividad se resumen en la tabla 15. Tres de las proteínas de fusión ST6GalNAcI replegadas tenían actividad detectable.

Tabla 15: Resumen de proteínas de fusión ST6GalNAcI replegadas sometidas a prueba para determinar la actividad mediante ensayo en fase sólida.

Constructo	Proteína replegada detectada mediante SDS-PAGE	Actividad de proteína replegada detectada mediante ensayo en fase sólida
MBP-mST6GalNAcID32	+	-
MBP-mST6GalNAcIE52	++	-
MBP-mST6GalNAcIS127	+++	+
MBP-mST6GalNAcIS186	+/-	+/-
MBP-hST6GalNAcIK36	+/-	+

Ejemplo 10: Replegamiento de proteínas Core 1 GalT1

- Se fusiona Core 1 GalT1 eucariota con MPB o con la doble cola, MBPSPD. Se usan las proteínas Core 1 GalT1 de *Drosophila* y humana. La figura 41 proporciona la secuencia de longitud completa de proteína Core 1 GalT1 humana. La figura 42 proporciona las secuencias de dos proteínas Core 1 GalT1 de *Drosophila*. Se realizan truncamientos de cada enzima a lo largo de la región de tallo, es decir, comenzando con la región de tallo de longitud completa y delecionando un aminoácido cada vez de tal manera que el truncamiento más pequeño sólo comprende un dominio catalítico de Core 1 GalT1. También se mutan residuos de cisteína a lo largo del dominio catalítico de las proteínas uno cada vez por residuos o bien serina o bien alanina. Se realizan fusiones con MBP usando las proteínas truncadas, los mutantes de cisteína o combinaciones de los dos. Se expresan proteínas en *E. coli* como cuerpos de inclusión, se solubilizan, y luego se repliegan usando los métodos descritos en el presente documento. Se determina el replegamiento midiendo la actividad enzimática. Las enzimas activas se han replegado correctamente. Se mide la actividad enzimática de Core 1 GalT1 tal como se da a conocer en Ju *et al.*, J. Biol. Chem. 277: 178-186 (2002), que se incorpora al presente documento como referencia para todos los fines.

Ejemplo 11: Replegamiento en una sola etapa de glicosiltransferasas unidas a O

- Pueden usarse colectivamente las glicosiltransferasas unidas a O GalNAcT2, Core 1 y ST3Gal1 para añadir una estructura core 1 que incluye un ácido siálico terminal o ácido siálico-PEG en un residuo de serina o treonina de proteínas seleccionadas incluyendo proteínas terapéuticas. La expresión de estas enzimas en *E. coli* y recuperación de enzimas activas a partir del replegamiento de cuerpos de inclusión son útiles para el desarrollo de un procedimiento ampliable a escala rentable. En este caso se describe el replegamiento conjunto de dos de estas enzimas (MBP-GalNAcT2 y MBP-ST3Gal1) a partir de cuerpos de inclusión de *E. coli*. Las ventajas del replegamiento conjunto incluyen disminución del uso de reactivos y aumento de la eficacia de replegamiento.
- Se seleccionaron dos cepas de JM109 por resistencia a kanamicina y se determinó que llevaban los plásmidos de expresión apropiados que acumulan cuerpos de inclusión de MBP-GalNAcT2 y MBP-ST3Gal1 tras la inducción con IPTG.

Tras inducciones durante la noche separadas con IPTG 1 mM, se resuspendieron sedimentos de células bacterianas a una razón de 1 g de sedimento celular en húmedo por 10 ml de Tris 20 mM pH 8, EDTA 5 mM y se sometieron a lisis mediante alteración mecánica con dos pasos a través de un microfluidizador. Se sedimentó el material insoluble, es decir, los cuerpos de inclusión o CI, mediante centrifugación a 7000xG a 4°C en el instrumento Sorvall RC3 durante 30 minutos, y se desechó el sobrenadante. Un ciclo de lavado típico consistió en resuspender completamente el sedimento en el tampón de lavado, y repetir la centrifugación durante 15 minutos. Se lavó el sedimento una vez en un volumen en exceso (al menos 10 ml (hasta 20) por g de sedimento celular original) de tampón con alto contenido en sal (lavado I: Tris 10 mM pH 7,4, NaCl 1 M, EDTA 5 mM), una vez en tampón detergente (lavado II: Tris 25 mM pH 8, NaCl 100 mM, Triton X 100 al 1%, Na-desoxicolato al 1%, EDTA 5 mM), y (para eliminar restos de detergente además de lavar los CI) tres veces en tampón de lavado (lavado III: Tris 10 mM pH 8, EDTA 5 mM).

Se resuspendieron los CI lavados en tampón de solubilización (urea 8 M, BisTris 50 mM pH 6,5, EDTA 5 mM, DTT 10 mM) y se ajustaron hasta una concentración de proteína de 4 mg/ml. Se aclaró la disolución de proteína de urea mediante centrifugación a 14000xG durante 5 minutos. Se realizó el replegamiento como una dilución rápida hasta 0,1 mg/ml en tampón de replegamiento (Tris 55 mM pH 8,2, NaCl 10,56 mM, KCl 0,44 mM, MgCl₂ 2,2 mM, CaCl₂ 2,2 mM, Peg3350 al 0,055%, L-arginina 550 mM, GSH 1 mM, GSSG 100 mM) y se agitó durante 16 h a 4°C.

Se configuró el sistema experimental tal como sigue: se mantuvieron constantes las condiciones de replegamiento durante el replegamiento o bien MBP-GalNAcT2 y MBP-ST3Gal1 por separado o bien el replegamiento conjunto en el mismo recipiente. Tras el replegamiento, se desalaron las mezclas mediante centrifugación a baja velocidad usando G50 Sephadex equilibrada en BisTris 50 mM pH 6,5, NaCl 75 mM, Tween 80 al 0,05%. Se recuperó material soluble tras la centrifugación a una velocidad máxima en una microcentrífuga.

Para determinar el rendimiento de enzima soluble tras el replegamiento, se sometió cada reacción a análisis de SDS-PAGE seguido por tinción con azul de Coomassie para visualizar los polipéptidos (figura 27A). Los resultados muestran que el replegamiento conjunto aumenta el rendimiento de enzimas solubles (carriles para el replegamiento conjunto frente a separado).

En el segundo experimento, se sometieron 6 µg de la proteína terapéutica IF α -2b a una reacción de pegilación de tres proteínas en una etapa. La concentración de enzimas usada para esta reacción son 1,4 µg de MBP-GalNAcT2, 0,5 µg de BV Core 1, y 1,4 µg de MBP-ST3Gal1, los azúcares usados son 1,2 µg de UDP-GalNAc y 1,2 µg de UDP-Gal y 125 µg de 20K-Peg-CMP-NAN en una reacción de 20 µl durante o bien 0 h o bien 16 h usando los siguientes tampones de reacción (MES 50 mM pH 6,2, NaCl 150 mM, MnCl₂ 10 mM). Se visualizó el grado de pegilación de IF α -2b en la mitad de la reacción mediante tinción con azul de Coomassie tras SDS-PAGE (figura 27B). El experimento se diseñó para comparar directamente la pegilación en una etapa usando o bien la combinación de enzimas replegadas individualmente o bien las enzimas replegadas conjuntamente manteniendo todos los demás componentes y concentraciones constantes. Los resultados demuestran un aumento en el producto pegilado en la reacción en la que las enzimas se replegaron conjuntamente en comparación con replegarse por separado (carril 5 frente a 7).

Ejemplo 12: Expresión mejorada de la proteína de fusión MBP-SiaA

Se optimizó el codón del gen de SiaA a partir de *Haemophilus influenzae* no tipificable para la expresión en *E. coli*. Se digirió el gen con NdeI y EcoRI, se purificó en gel de agarosa y se ligó en pcWin2 digerido con NdeI/EcoRI, antes de la transformación en la cepa JM109 o W3110 de *E. coli*. Se aisló ADN de plásmido a partir de los recombinantes y se examinó con NdeI y EcoRI. Se sembraron una colonia de JM109 y una de W3110 que contenían cada una el constructo correcto en 2 ml de LB libre de animal que contenía sulfato de kanamicina 10 µg/ml y se hicieron crecer durante 6 h a 37°C a 250 RPM de agitación. Se retiró una alícuota de 100 µl, se centrifugó 2 minutos a 10.000 x g y se desechó el sobrenadante. Se congeló este sedimento a -20°C y representa la célula no inducida. Se añadió IPTG al resto del cultivo a una concentración final de IPTG 1 mM y se incubó durante 2h a 37°C y 250 RPM de agitación. Se retiró una alícuota de 100 µl y se procesó tal como se describe (representa célula inducida).

Se cambiaron los sitios de restricción en los extremos del gen de SiaA por 5'BamHI y el extremo 3' se mantuvo como EcoRI usando PCR. Se digirió el producto de PCR con BamHI y EcoRI, se purificó mediante electroforesis en gel de agarosa y se ligó en pcWin2MBP digerido con BamHI/EcoRI. Se cultivaron células JM109 transformadas con esta reacción de ligación y se aisló ADN de plásmido. Se examinó el ADN de plásmido usando BamHI y EcoRI. Se sembraron tres colonias que contenían la estructura correcta en 2 ml de LB libre de animal que contenía sulfato de kanamicina 10 µg/ml y se hicieron crecer durante 6 h a 37°C a 250RPM de agitación. Se retiró una alícuota de 100 µl, se centrifugó 2 minutos a 10.000 x g y se desechó el sobrenadante. Se congeló este sedimento a -20°C y representa la célula no inducida. Se añadió IPTG al resto del cultivo a una concentración final de IPTG 1 mM y se incubó durante 2 h a 37°C y 250 RPM de agitación. Se retiró una alícuota de 100 µl y se procesó tal como se describe (representa célula inducida).

Se sometió a ebullición cada alícuota de 100 µl 5 minutos en 100 µl de tampón de muestra de SDS-PAGE que contenía DTT 50 mM. Se cargaron las muestras sobre gel de acrilamida Tris-glicina al 4-20% (Invitrogen) y se

sometieron a electroforesis durante 2 h, se tñeron con Simply Blue Safestain de Invitrogen, se destiñeron con agua y se examinaron. La figura X muestra niveles de expresión no detectables de SiaA nativa tras la inducción, mientras que la figura Y muestra altos niveles de expresión de SiaA fusionada con MBP tras la inducción con IPTG. Este resultado demuestra que MBP proporcionado a partir del vector pcWin2MBP impulsa un alto nivel de expresión de proteína.

5

10

15

20

25

30

35

40

45

50

55

REIVINDICACIONES

1. Una proteína GalNAcT2 recombinante, en la que se delecionan una región de anclaje de tallo y un dominio transmembrana de la proteína GalNAcT2 recombinante y en la que se fusiona la proteína GalNAcT2 en el marco con un dominio de unión a maltosa.
- 5 2. La proteína GalNAcT2 recombinante de la reivindicación 1, en la que se deleciona la totalidad o una parte de región de tallo.
3. La proteína GalNAcT2 recombinante de la reivindicación 1, en la que se elimina una cisteína desapareada en la proteína GalNAcT2 recombinante mediante sustitución por aminoácido distinto a cisteína.
- 10 4. La proteína GalNAcT2 recombinante de la reivindicación 1, siendo la proteína GalNAcT2 recombinante una proteína GalNAcT2 humana truncada seleccionada de GalNAcT2 $\Delta 40$, GalNAcT2 $\Delta 51$, GalNAcT2 $\Delta 74$ y GalNAcT2 $\Delta 95$.
5. Un método de remodelación de una proteína, un péptido, una glicoproteína o un glicopéptido usando la proteína GalNAcT2 recombinante de la reivindicación 1.
- 15 6. Un método de replegamiento de una primera glicosiltransferasa eucariota, recombinante, insoluble, en el que la glicosiltransferasa comprende un dominio de proteína de unión a maltosa (MBD), comprendiendo el método las etapas de
 - (a) solubilizar la glicosiltransferasa eucariota, recombinante, insoluble en un tampón de solubilización; y
 - (b) poner en contacto la glicosiltransferasa eucariota soluble con un tampón de replegamiento que comprende un par redox para replegar la glicosiltransferasa eucariota, en el que la glicosiltransferasa eucariota replegada cataliza la transferencia de un azúcar de un sustrato donador a un sustrato aceptor; en el que la primera glicosiltransferasa eucariota es GalNAcT2.
- 20 7. El método de la reivindicación 6, en el que la primera glicosiltransferasa eucariota se trunca para eliminar la totalidad o una parte de una región de tallo.
- 25 8. El método de la reivindicación 6, en el que se elimina una cisteína desapareada en la primera glicosiltransferasa eucariota mediante sustitución por un aminoácido distinto de cisteína.
9. El método de la reivindicación 6, en el que la primera glicosiltransferasa eucariota comprende además un dominio de purificación seleccionado del grupo que consiste en un dominio de unión a almidón, un dominio de tiorredoxina, un dominio SUMO, un dominio de poli-His, un dominio de epítipo myc y un dominio de glutatión-S-transferasa.
- 30 10. El método de la reivindicación 6, en el que la primera glicosiltransferasa eucariota comprende además un dominio de autoescisión.
11. El método de la reivindicación 6, en el que la primera glicosiltransferasa eucariota se expresa en una célula huésped bacteriana como un cuerpo de inclusión insoluble.
- 35 12. El método de la reivindicación 6, en el que una segunda glicosiltransferasa eucariota recombinante, insoluble se repliega con la primera glicosiltransferasa eucariota.
13. El método de la reivindicación 12, en el que una tercera glicosiltransferasa eucariota recombinante, insoluble se repliega con la primera glicosiltransferasa eucariota y la segunda glicosiltransferasa eucariota.
14. El método de la reivindicación 6, en el que el par redox se selecciona del grupo que consiste en glutatión reducido/glutatión oxidado (GSH/GSSG) y cisteína/cistamina.
- 40 15. El método de la reivindicación 6, en el que el sustrato aceptor se selecciona de una proteína, un péptido, una glicoproteína y un glicopéptido.

n.º	1 mM	0,1 mM	0,3 mM	mM	mM	0,055%	550 mM	1,1 mM	2,2 mM	2,2 mM	440 mM	550 mM	Actividad
	GSH	GSSG	LM	NaCl	KCl	PEG 3350	GndHCl	EDTA	MgCl ₂	CaCl ₂	Sacarosa	L- Arg	U/g IB
2 (MES 55 nM pH 6,5)	+	+	+	10,56	0,44	0	+	0	+	+	0	0	0
3 (MES 55 nM pH 6,5)	+	+	0	10,56	0,44	+	+	+	0	0	+	+	0
n.º 5 (MES 55 nM pH 6,5)	+	+	0	264	11	0	0	0	+	+	+	0	0
n.º 6 (MES 55 nM pH 6,5)	+	+	+	264	11	+	0	+	0	0	0	+	40,00
n.º 10 (MES 55 nM pH 8,2)	+	+	+	10,56	0,44	0	0	+	0	0	+	0	0
n.º 11 (Tris 55 nM pH 8,2)	+	+	0	10,56	0,44	+	0	0	+	+	0	+	105,26
n.º 13 (Tris 55 nM pH 8,2)	+	+	0	264	11	0	+	+	0	0	0	0	15,65
n.º 16 (Tris 55 nM pH 8,2)	+	+	+	264	11	+	+	0	+	+	+	+	48,70

FIG. 1

Elución de MBP-ST3Gal III de una columna de amilosa

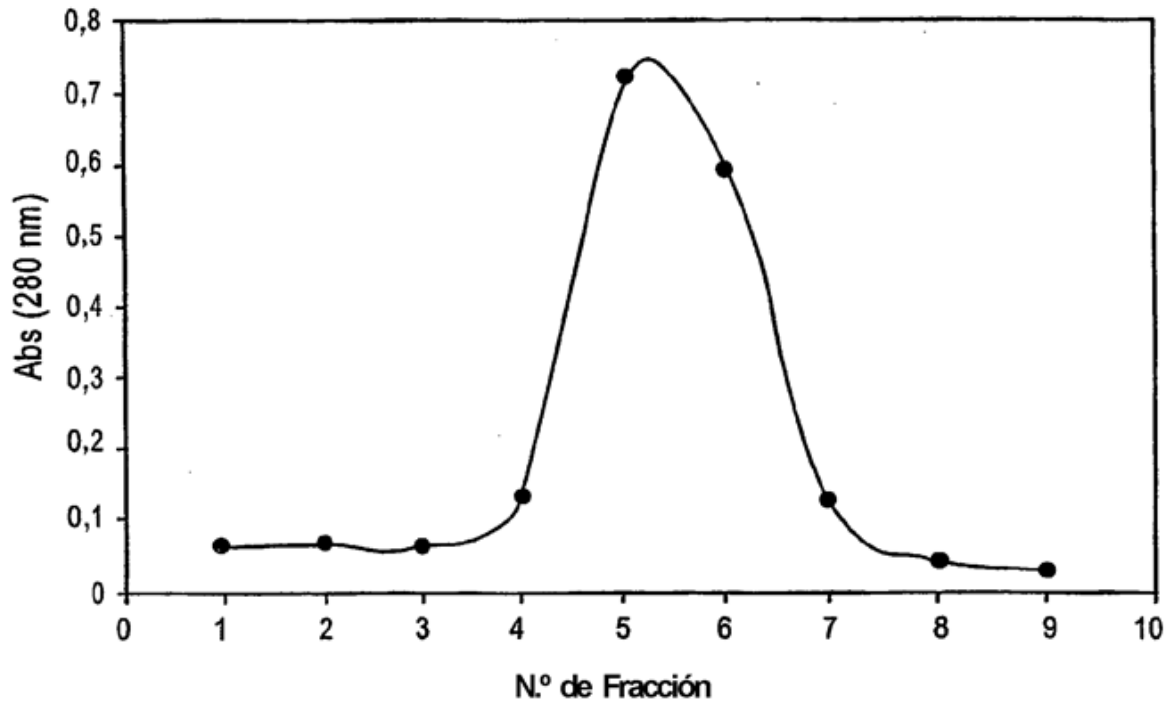


FIG. 2

Actividades de ST3 Gal III de las fracciones de MBP-ST3GalIII replegada
purificada con amilosa

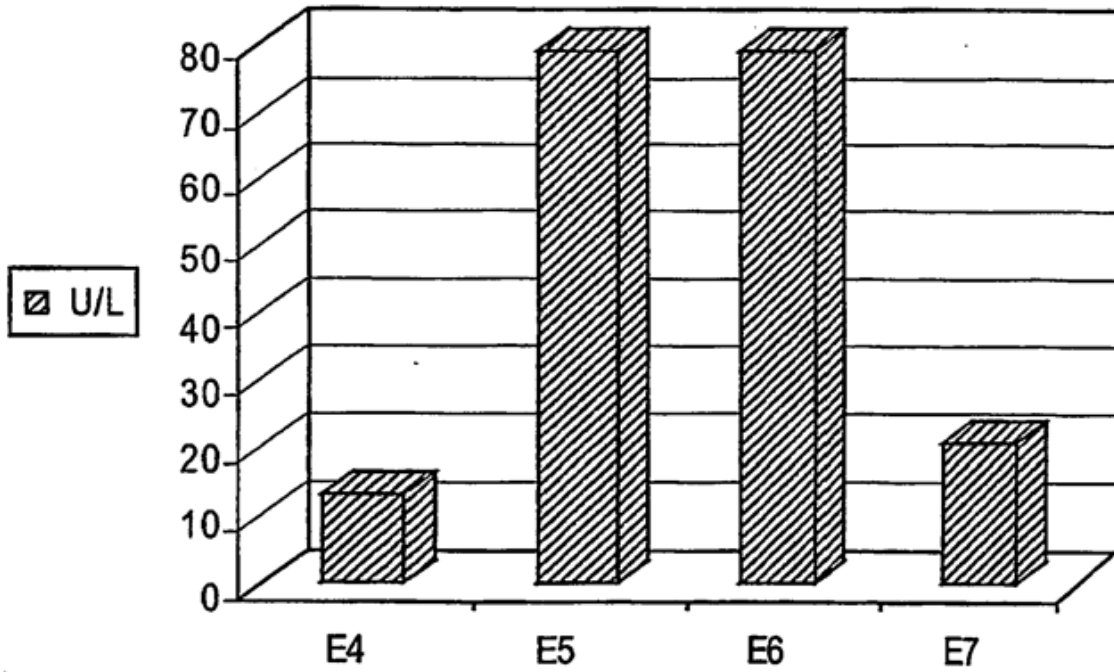


FIG. 3

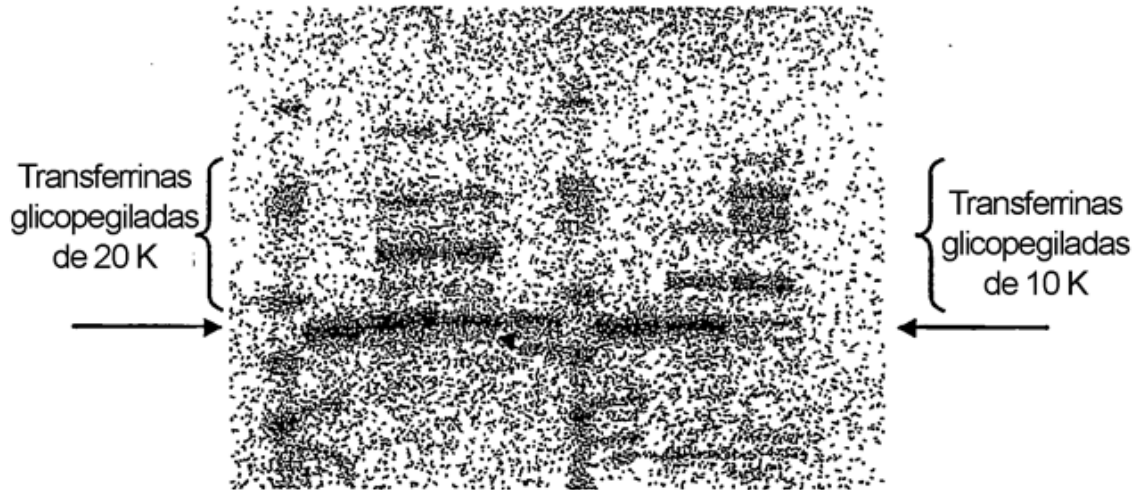
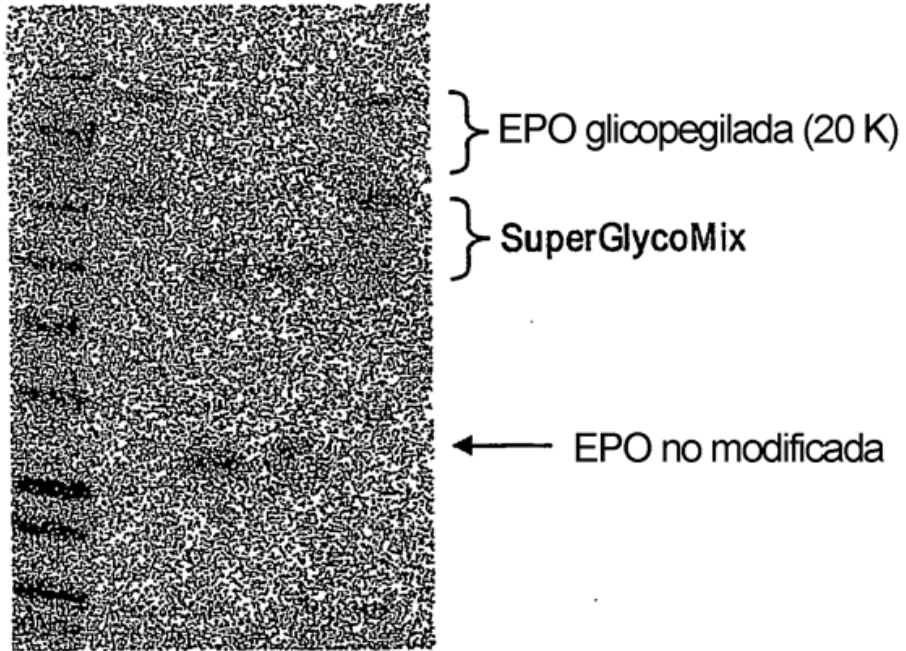


FIG. 4



Glicopegilación (20 K) de EPO

FIG. 5

```

                10      20      30      40      50      60
/usr/t MLKKQSAGLVLWGAILFVAWNALLLFFWTRPAPGRPPSVSALDGDPASLTREVIRLAQD
      ::::::::::::::::::::::::::::::: :: .::: ::::::::::::::
P27115 MLKKQSAGLVLWGAILFVAWNALLLFFWTRPVPSRLPSDNALDDDPASLTREVIRLAQD
                10      20      30      40      50      60

                70      80      90      100     110
/usr/t AEVELERQRGLLQOIGD--ALSSQGRVPTAAPPAQPRVFVPAPAVIPILVIACDRSTV
      :::::::::::::: . :: :: .: ::::::::::::::: ::::::::::::::
P27115 AEVELERQRGLLQOIREHHALWSQRWKVPTAAPPAQPHVFVTPPPAVIPILVIACDRSTV
                70      80      90      100     110     120

                120     130     140     150     160     170
/usr/t RRCLDKLLHYRPSAELEFPIIVSQDCGHEETAQAIASYGSAVTHIRQPDLSSIAVPPDHRK
      ::::::::::::::::::::::: ::::::::::::::: ::::::::::::::
P27115 RRCLDKLLHYRPSAELEFPIIVSQDCGHEETAQVIASYGSAVTHIRQPDLSNIAVQPDHRK
                130     140     150     160     170     180

                180     190     200     210     220     230
/usr/t FQGYKIARHYRWALGQVERQFRFPAAVVEDDLEVAPDFFEYFRATYPLLKADPSLWCV
      ::::::::::::::: .: ::::::::::::::: :::::::::::::::
P27115 FQGYKIARHYRWALGQIFHNENYPAAVVEDDLEVAPDFFEYFQATYPLLKADPSLWCV
                190     200     210     220     230     240

                240     250     260     270     280     290
/usr/t SAWNDNGKEQMDASRPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQR
      ::::::::::::::: .: ::::::::::::::: :::::::::::::::
P27115 SAWNDNGKEQMDSSKPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQR
                250     260     270     280     290     300

                300     310     320     330     340     350
/usr/t QGRACIRPEISRMITFGRKGVSHGQFFDQHLKFKIKLNQFVHFTQLDLSYLQREAYDRDF
      ::::::::::::::: .: ::::::::::::::: :::::::::::::::
P27115 KGRACVRPEISRMITFGRKGVSHGQFFDQHLKFKIKLNQFVPFTQLDLSYLQQEAYDRDF
                310     320     330     340     350     360

                360     370     380     390     400     410
/usr/t LARVYGAPQLQVEKVRITNDRKELGEVRVQYTGRDSEFKAFKALGVMDLDKSGVPRAGYRG
      ::::::::::::::: .: ::::::::::::::: :::::::::::::::
P27115 LARVYGAPQLQVEKVRITNDRKELGEVRVQYTGRDSEFKAFKALGVMDLDKSGVPRAGYRG
                370     380     390     400     410     420

                420     430     440
/usr/t IVTFQFRGRRVHLAPPPTWEGYDPSWN
      ::: ::::::::::: :::::::::::
P27115 IVTFLFRGRRVHLAPPQIWDGYDPSWI
                430     440

```

FIG. 6

Mutante de Cys121Ser de GnT1

avipilviacdrstrvrsldkllhhyrpsaelpfiivsqdcgheetaqaiasygsavthirqpdlssiavppdhrkfqggyykiarhyrwa
 lgqvfrqfrfpaavveddle vapdffeyfratypllkadpslwcvsawndngkeqmvdasipellytrtdffpfglgwlllaelwae
 lepkipkafwddwmrpeqrqgracirpeisrtmtfgrkqshgqffdqhlkfklnqfvhftqldlsylqreaydrdflarvyg
 apqlvekvrtndrkelgevrvqytrdsfkafakalgvmdlksvpragyrgivtfqfpgrrvhlappptwegydpwsn*

Gcgggtattccatcctggtcacgcctgtgaccgagcactgttcggcgctctctagacaagctgctgcattatggccctcggctga
 gctcttcccacatcgttagccaggactcggggcacgagagacggcccaggccatcgctctctacggcagcgcggtcacgcaca
 tccggcagcccagcctgagcagcattcgggtgcccggaccaccgcaagttccaggggtactactacaagatcgcgccactaccg
 ctggcgctggccaggctctccggcagtttcgcttcccggccgctggtggtgagatgacctggaggtggccccggacttctt
 cgagtactttcggccacctatccgctgctgaaaggccgacccctccctgtgtgctcgcctggcctggaatgacaacggcaaggagca
 gatggtagcggccagcggcctgagctgctctaccgacccgacttttcccctggcctgggctgctgctgtgtggccgagctcgggct
 gagctggagcccagaagtggccaaaggcctctgggacgactggatgcggcggccggagcagcgaggcgggcctgcatacg
 ccctgagatctcaagaacgatgaccttggccgcaagggtgtgagccacggcgagttctttgaccagcacctcaagttatcaagctga
 accagcagtttgtgcacttcacccagctggacctgtcttaacctgcagcggagggcctatgaccggagatttcctcggcccgtctacggg
 gctccccagctgcaggtggagaaaagtgaaggaccatgaccggaaaggactggggggaggtgcgggtgcagtatacggggcagggga
 cagcttcaaggcttcgccaaggctctgggtgtcatggatgaccttaagtcgggggtccggagagctgggtaccgggggtattgtcacctt
 ccagttcccggggcccgctgtccacctggcgccccaccacggcagctgggaggggctatgatcctagctgggaattag

FIG. 7

GnT1 Cys121Asp

avipilviacdrstvrrldkllhyrpsaelpiivsvdqgheetaqaiasygsavthirqpdlssiaavppdhrkfqgyykiarhyrwa
 lgvfrqfrfpaavvveddleavapdffeyfratypllkadpslwcvswndngkeqmvdasrpellyrtdffpfglgwillaelwae
 lepkwpkafwddwmrrpeqrqgracirpeisrmtfgrkgvshgqffdkhlfklnqfvhfiqldlsylqreaydrdrflarvys
 apqlqvekvrtndrkelgevrvytytgrdsfkafakalgvmddllksgvpragyrvtfqfpgrrvhlappptwegydpswn*

Gcgggtgattccatcctggtcatcgcctgtgaccgcagcactgttcggcgcatctagacaagctgctgattatcgccctcggcctg
 agctcttcccatacctgtagccagactggggcacgagggagacggcccagggccatcgccctctactacggcagcggctcaccgac
 atccggcagccccgacctgagcagcattgcgggtgcggccggaccaccgcaagttccaggsgctactacaagatcggcggccactacc
 gctggcgctgggcccagggtctccggcagtttcgcttccccggccgctggtggtggagatgacctgggaggtggccccggacttct
 tcgagtaetttcggggccacctatccgctgctgaaaggccgacccccctcctgtgtgctgctcggcctggcaatgacaacggcaaggagc
 agatgggtggacggcccagggcctgagctgctctaccgcaaccgacttttccctgggctggtggctgctggtggcagctctggggc
 tggagctggagcccagaagtggccaaaaggcctctgggacgactgcatgctggcggccggagcaggcaggsggscctgcatatac
 gcccctgagatctcaagaaagatgaccttggccgcaaggsgtggagccacggsgcagttctttgaccagcacctcaagtttatcaagctg
 aaccagcagtttggcacttcaccacagctggactgcttaccctgcaaggcgggagggcctatgaccaggatttctctcggccgctctacggg
 tgcctcccagctgcaagggtggagaaaagtggagaccaatgaccggaaaggagctggggggagggctggcgggtgcaagtatacggsgcaggga
 cagcttcaaggccttcccaaggctctggsgtggcatgagaccitaaagtcgggggttccggagagcggctacccggsggtattgtcacctt
 ccagttccccggggccggctgtccacctggcgcccccaaccgacgtggggaggggctatgacatcctagctgggaattag

FIG. 8

GnT1 Cys121Thr

avipilviacdrstrvrrtldkllhyrpsaelpiivsqdcgheetaqaiasygsvthirqpdlssia vppdhrkfqgyykiarhyrwal
gqvfrfrfpaavvveddlevapdffeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffpglgwillaelwael
epkwpkafwdwmrpeqrqgracirpeisrmtfgrkvgshgqffdqhlkflknqqfvhftqlslslylqreaydrdflarvys
apqlqvekvrtndrkelgevrvqytrgrdsfkafakalgvmdllksgvpragyrgivtfqfpgrrvhlappptwegydpwswn*

Gcgggtattcccatcctgggtcatcgcctgtgaccgagcactgttcggcgcactctagacaagctgctgattatcggccctcggcctg
agctcttccccatcatcgtttagccaggactgcggggcaagaggagacggccccaggccatcgcctctacggcagcggcggfcaacggcac
atccggcagcccgaccctgagcagcattgctggccggccggaccaccgcaagttccaggggctactacaaagatcggcggcgcactacc
gctggggcgcctggggcaggctctccggcagtttcgcttccccggccgctggccgctggctgagatgacctggagctggccccggactctc
tcgagtactttcggggccacctatccgctgctgaaaggccgacccccctcctgctggccctgggaatgacaacggcaaggagagc
agatggtgagccagcaggccctgagctgctctaccggcaccgaccttttccccggccctggcctggcctgctggcctgctggc
tgagctggagccccaaagtggccaaaggccctctggggacgactggagcggactggctggcggccggagcagcggcaggggggcggcctgcatatc
ggcctgagatctcaagaacgatgaccctggccgcaagggtgtgagccacggggcagttctttgacccagcaacctcaagtttatcaagctg
aaaccagcagttgtgacttcaccacagctggaccctgtcttaccctgacagcggggagggcctatgaccgagatttctcggccccgctctacggg
tgctccccagctgcaagggtggagaaagtggagaccacaatgaccgggaaaggagctgggggagggctggcggctgacgtatatacggggcagggga
cagcttcaaggccttccgcaaggctctgggggtgctcatgtagacccttaagctcgggggggttccggagagcctggcctaccgggggtattgtcacctt
ccagttccccggggccccgctgtccaccctggcggccccccccaccggacgtggggaggggctatgattccttagctgggaattag

FIG. 9

GnT1 Cys121Ala

avipilviacdstrvrralldklhlyrpsaelfpaiivsqdcgheetaqaiasygsavthirqpdlssiavpdhrkfqyykiarhyrwa
 lgvfrqfrfpaavveddlevapdffeyfratypilkadpslwcvsawndngkeqmvdasrpeillyrtidffpfglgwillaelwae
 lepkipkafwddwmrrpeqrqgracirpeisrmtfgrkqvshgqffdqhlkfiklnqqfwhftqlslslylqreaydrdflarvyg
 apqlqvekvrtndrkelgevrvqytrgdsfkafakalgvmdllksgvpragyrvitvtfqfgrrvhlapppptwegydpswn*

Ggggtattcecatcctggtcatcgcctgtgaccgagcactgttgccgagcgccttagacaagctgctgattatggccccgggctg
 agctcttccccatcattagccaggactggggcacgagggagacggccccaggccatcgcctctacggcagcggcgtcacggcac
 atccggcagccccgaccigagcagcattgggtgccccgggaccaccgcaagttccagggtactacaaagatcgcgcgccaactacc
 gctggggcgtgggccaaggcttccggcagtttcgcttccccggccgctggtggtgagggatgacctggaggtggccccgggacttct
 tcgagtaacttccgggcaacctatccgctgctgaaaggccccgacccccctgfggtgctctcggccctgggaatgacaacggcaaggaggc
 agatggtggagcggcagcggcctgagctgctctaccgaccactttccccggcctggggctggtgctgtgtggccgagcttgggg
 tggactggagccccaaagtggccaaaggcccttctggggacgactggatgcccggcggccggagcagcggcagggcgctgcatac
 gcccctgagatcacaagaaacgatgaccttggccgcaagggtgtgagccacggsgcagttctttgaccagcaccicaagttatcaagctg
 aaccagcagtttgtgcacttaccaccagctggacctgtcttacctgcaagcggggagggccctatgaccagattcctcggccccgtctacggg
 tgcctcccagctgcaaggtgagaaaagtgagaccaatgaccggaaaggagctggggggaggtgcccgggtgcagtataccggcagggga
 cagcttcaaggcttcccaaggctcctgggtgtcatggtgaccttaagtcgggggttccgagagctggctaccgggggtattgtcacctt
 ccaagttccccgggcccggctgttccacctggcggccccaccagcagctgggaggggctatgctctagctgggaattag

FIG. 10

Secuencia de aminoácidos de ST3Gal III de hígado de rata:

MGLLVFVRNLLALCLFLVLFGLYYSAWKLHLLQWEDSNLILSLDSAGQTLGTEYDRL
 GFLKLD SKLPAELATKYANFSEGACKPGYASAMMTAIFRFSKPAPMFLDDDFRKW
 ARIRFVPPFGIKGQDNLIKAILSVTKEYRLTPALDSLHCRRCIIVGNGGVLANKSLGS
 RIDDYDIVIRLNSAPVKGFEDVGSKTTLRITYPEGAMQRPEQYERDSLFLAGFKW
 QDFKWLKYIVYKERVSA SDGFWKS VATRVPKEPPEIRILNPYFIQEAFTLIGLPFNN
 GLMGRGNIPTLGSVAVT~~M~~ALDGCDEVAVAGFGYDMNTPNAPLHYHYYETVRMAAIKE
 SWTHNIQREKEFLRKLVKARVITDLSSGI

FIG. 12

Secuencias de ácido nucleico y de aminoácidos de UDP-N-acetilgalactosaminil-transferasa 2 (GalNAcT2) de longitud completa

Secuencia de aminoácidos

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

FIG. 13A

Gln	Arg	Arg	Ser	Arg	Gln	Gly	Asn	Pro	Val	Ala	Pro	Ile	Lys	Thr	Pro
	290					295					300				
Met	Ile	Ala	Gly	Gly	Leu	Phe	Val	Met	Asp	Lys	Phe	Tyr	Phe	Glu	Glu
305					310					315					320
Leu	Gly	Lys	Tyr	Asp	Met	Met	Met	Asp	Val	Trp	Gly	Gly	Glu	Asn	Leu
				325					330					335	
Glu	Ile	Ser	Phe	Arg	Val	Trp	Gln	Cys	Gly	Gly	Ser	Leu	Glu	Ile	Ile
			340					345					350		
Pro	Cys	Ser	Arg	Val	Gly	His	Val	Phe	Arg	Lys	Gln	His	Pro	Tyr	Thr
		355					360					365			
Phe	Pro	Gly	Gly	Ser	Gly	Thr	Val	Phe	Ala	Arg	Asn	Thr	Arg	Arg	Ala
	370					375					380				
Ala	Glu	Val	Trp	Met	Asp	Glu	Tyr	Lys	Asn	Phe	Tyr	Tyr	Ala	Ala	Val
385					390					395					400
Pro	Ser	Ala	Arg	Asn	Val	Pro	Tyr	Gly	Asn	Ile	Gln	Ser	Arg	Leu	Glu
				405					410					415	
Leu	Arg	Lys	Lys	Leu	Ser	Cys	Lys	Pro	Phe	Lys	Trp	Tyr	Leu	Glu	Asn
			420					425					430		
Val	Tyr	Pro	Glu	Leu	Arg	Val	Pro	Asp	His	Gln	Asp	Ile	Ala	Phe	Gly
		435					440					445			
Ala	Leu	Gln	Gln	Gly	Thr	Asn	Cys	Leu	Asp	Thr	Leu	Gly	His	Phe	Ala
	450					455					460				
Asp	Gly	Val	Val	Gly	Val	Tyr	Glu	Cys	His	Asn	Ala	Gly	Gly	Asn	Gln
465					470					475					480
Glu	Trp	Ala	Leu	Thr	Lys	Glu	Lys	Ser	Val	Lys	His	Met	Asp	Leu	Cys
				485					490					495	
Leu	Thr	Val	Val	Asp	Arg	Ala	Pro	Gly	Ser	Leu	Ile	Lys	Leu	Gln	Gly
			500					505					510		
Cys	Arg	Glu	Asn	Asp	Ser	Arg	Gln	Lys	Trp	Glu	Gln	Ile	Glu	Gly	Asn
		515					520					525			
Ser	Lys	Leu	Arg	His	Val	Gly	Ser	Asn	Leu	Cys	Leu	Asp	Ser	Arg	Thr
	530					535					540				
Ala	Lys	Ser	Gly	Gly	Leu	Ser	Val	Glu	Val	Cys	Gly	Pro	Ala	Leu	Ser
545					550					555					560
Gln	Gln	Trp	Lys	Phe	Thr	Leu	Asn	Leu	Gln	Gln					
				565					570						

FIG. 13A (CONT.)

Secuencia de ácido nucleico						
atgagggggc	gctoggggat	gctgctctgc	ttogpcttoc	tgtgggtgct	gggcatogoc	60
tactacatgt	actogggggg	cgctctgog	ctggocgggg	gogggggogg	ogggoggggc	120
aggaaggagg	actggaatga	aattgacoc	attaaaaaga	aagaocctca	tcacagcaat	180
ggagaagaga	aagcacaag	catggagacc	ctooctocag	ggaaagtaog	gtggocagac	240
tttaaccagg	aagcttatgt	tggagggag	atggtoogct	oogggcagga	ooccttaogoc	300
ogcaacaagt	tcaaccagg	ggagagtgat	aagcttogaa	tggacagagc	catooctgac	360
acoooggcatg	accagtgta	goggaagcag	tggoggggtgg	atctgocggc	caocagogtg	420
gtgatcacgt	ttcaaatga	agocaggto	gocctactca	ggacogtgg	cagogtgctt	480
aagaaaagoc	ogooocatct	cataaaaga	atcatcttgg	tggatgacta	cagcaatgat	540
oetgaggag	gggctctott	ggggaaaatt	gagaaagtgc	gagttcttag	aaatgatoga	600
ogagaaggoc	tcatgogctc	acgggttog	ggggocgatg	ctgoccaagc	caaggtoctg	660
aocttoctgg	acagtcactg	ogagtgtaat	gagcactggc	tggagococ	oetggaaagg	720
gtggogggagg	acaggactog	ggttgtgtca	occatcatog	atgtcattaa	tatggacaac	780
tttcagtatg	tgggggcatc	tgctgacttg	aagggoggtt	ttgattggaa	cttggttatc	840
aagtgggatt	acatgaogoc	tgagcagaga	aggtocoggc	aggggaoc	agtoogococ	900
ataaaaaoc	ocatgattgc	tggggggctg	tttgtgatgg	ataagttcta	ttttgaagaa	960
ctggggaagt	acgacatgat	gatggatgtg	tggggaggag	agaocctaga	gatctogttc	1020
ogogtgtggc	agtggtggg	cagocggag	atcatocogt	gcagocgtgt	gggacaogtg	1080
ttocoggaagc	agcaococ	caogttocog	ggtggcagtg	gcactgtctt	tgocogaaac	1140
acooogoggg	cagcagaggt	ctggatggat	gaatacaaaa	atttctatta	tgacagcagtg	1200
oettctgcta	gaaogttoc	ttatggaaat	attcagagca	gattggagct	taggaagaaa	1260
ctcagctgca	agoccttcaa	atggtacott	gaaaatgtct	atocagagtt	aagggttoca	1320
gacocatcagg	atatagcttt	tggggocctg	cagcagggaa	ctaactgoc	ogacactttg	1380
ggacactttg	ctgatgggtg	ggttggagtt	tatgaatgtc	acaatgctgg	gggaaocag	1440
gaatgggoc	tgacgaagga	gaagtoggtg	aagcocatgg	atttgtgoc	tactgtgggtg	1500
gaocgggcac	ogggctctct	tataaagctg	cagggctgoc	gagaaaatga	cagcagacag	1560
aatgggaac	agatogaggg	caactocaag	ctgagggcag	tgggcagcaa	octgtgocog	1620
gacagtoaca	ogoccaagag	ogggggoc	agcgtggagg	tgtgtggoc	ggocctttog	1680
cagcagtgga	agttcacgct	caocctgag	cag			1713

FIG. 13B

Secuencias de ácido nucleico y de aminoácidos de una $\Delta 51$ UDP-N-acetilgalactosaminiltransferasa 2,
GalNAct2

Secuencia de aminoácidos

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

FIG. 14A

290						295										300
Arg	Val	Gly	His	Val	Phe	Arg	Lys	Gln	His	Pro	Tyr	Thr	Phe	Pro	Gly	
305					310					315					320	
Gly	Ser	Gly	Thr	Val	Phe	Ala	Arg	Asn	Thr	Arg	Arg	Ala	Ala	Glu	Val	
				325					330					335		
Trp	Met	Asp	Glu	Tyr	Lys	Asn	Phe	Tyr	Tyr	Ala	Ala	Val	Pro	Ser	Ala	
			340					345					350			
Arg	Asn	Val	Pro	Tyr	Gly	Asn	Ile	Gln	Ser	Arg	Leu	Glu	Leu	Arg	Lys	
		355					360					365				
Lys	Leu	Ser	Cys	Lys	Pro	Phe	Lys	Trp	Tyr	Leu	Glu	Asn	Val	Tyr	Pro	
	370					375					380					
Glu	Leu	Arg	Val	Pro	Asp	His	Gln	Asp	Ile	Ala	Phe	Gly	Ala	Leu	Gln	
385					390					395					400	
Gln	Gly	Thr	Asn	Cys	Leu	Asp	Thr	Leu	Gly	His	Phe	Ala	Asp	Gly	Val	
				405					410					415		
Val	Gly	Val	Tyr	Glu	Cys	His	Asn	Ala	Gly	Gly	Asn	Gln	Glu	Trp	Ala	
			420					425					430			
Leu	Thr	Lys	Glu	Lys	Ser	Val	Lys	His	Met	Asp	Leu	Cys	Leu	Thr	Val	
		435					440					445				
Val	Asp	Arg	Ala	Pro	Gly	Ser	Leu	Ile	Lys	Leu	Gln	Gly	Cys	Arg	Glu	
	450					455					460					
Asn	Asp	Ser	Arg	Gln	Lys	Trp	Glu	Gln	Ile	Glu	Gly	Asn	Ser	Lys	Leu	
465					470					475					480	
Arg	His	Val	Gly	Ser	Asn	Leu	Cys	Leu	Asp	Ser	Arg	Thr	Ala	Lys	Ser	
				485					490					495		
Gly	Gly	Leu	Ser	Val	Glu	Val	Cys	Gly	Pro	Ala	Leu	Ser	Gln	Gln	Trp	
			500					505					510			
Lys	Phe	Thr	Leu	Asn	Leu	Gln	Gln									
		515					520									

FIG. 14A (CONT.)

Secuencia de ácido nucleico

aaaaagaaag	aocctcatca	cagcaatgga	gaagagaaag	cacaaagcat	ggagaooctc	60
cctocagggg	aagtaoogtg	gocagacttt	aaccaggaag	cttatgttgg	agggacgatg	120
gtocgctoo	ggcaggaooc	ttaogooogc	aacaagttca	accagggtgga	gagtgataag	180
cttogaatgg	acagagocat	coctgacaoc	oggcattgacc	agtgtcagcg	gaagcagtgg	240
oggtggatc	tgooogooac	cagcgtggtg	atcacgtttc	acaatgaagc	caggtogggc	300
ctactcagga	oogtggtcag	cgtgcttaag	aaaagooogc	ccatctcat	aaaagaatc	360
atcttgggtg	atgactacag	caatgatoot	gaggaooggg	ctctcttggg	gaaaattgag	420
aaagtgoag	ttcttagaaa	tgatogaoga	gaaggooctc	tgogctcaog	ggttoggggg	480
gocgatgctg	ooaaagooaa	ggtcoctgacc	ttcoctggaca	gtcactgoga	gtgtaatgag	540
cactggctgg	agooocooct	ggaaaggggtg	goggaggaca	ggactoggg	tggttcaoo	600
atcatogaty	tcattaatat	ggacaacttt	cagtatgttg	gggcatctgc	tgacttgaag	660
ggoggttttg	attggaactt	ggtattcaag	tggtattaca	tgacgocoga	gcagagaagg	720
toooogcag	ggaooocagt	ogooocata	aaaaoocoo	tgattgctgg	tggtctgttt	780
gtgatggata	agttctattt	tgaagaactg	gggaagtacg	acatgatgat	ggatgtgtgg	840
ggaggagaga	aocagagat	ctogttcoog	gtgtggcagt	gtgggtggcag	oocggagatc	900
atooogtga	gocgtgtggg	acaogtgttc	oggaagcagc	aoocctacac	gttcooggg	960
ggcagtggca	ctgtctttgc	ooaaacaoc	ogoooggcag	cagaggtctg	gatggatgaa	1020
tacaaaaatt	tctattatgc	agcagtgoct	tctgctagaa	acgttcocta	tggaatatt	1080
cagagcagat	tgagccttag	gaagaaactc	agctgcaagc	cttcaaaty	gtacottgaa	1140
aatgtctatc	cagagttaag	ggttocagac	catcaggata	tagcttttgg	ggocottgag	1200
cagggaacta	actgocoga	cactttggga	cactttgctg	atgggtgtgt	tgagtttat	1260
gaatgtcaca	atgctggggg	aaocaggaa	tgggocctga	ogaaggagaa	gtoggtgaag	1320
cacatggatt	tggtgocctac	tggtgtggac	ogggcaooog	gctctcttat	aaagctgag	1380
ggctgocog	aaaatgacag	cagacagaaa	tggaacaga	togagggcaa	ctooaagctg	1440
aggcaogtgg	gcagcaaoct	gtgocctggac	agtogpaoog	ooaagagocg	ggocctaagc	1500
gtggaggtgt	gtggoooggc	ootttogcag	cagtggaaat	tcaogctcaa	oocgocagc	1560

FIG. 14B

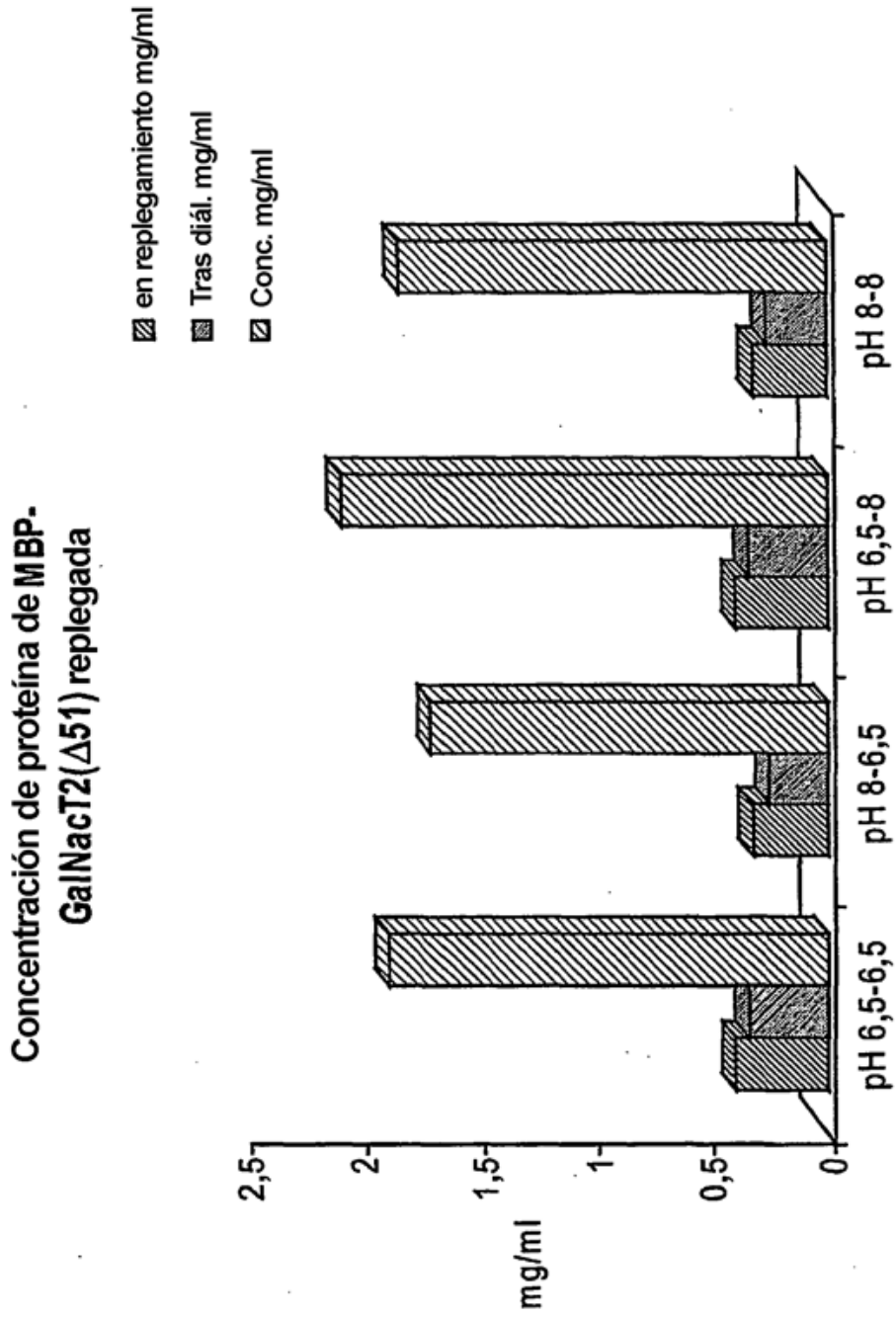


FIG. 15

Actividades de GalNAcT2 de MBP-GalNAcT2 (D51) replegada

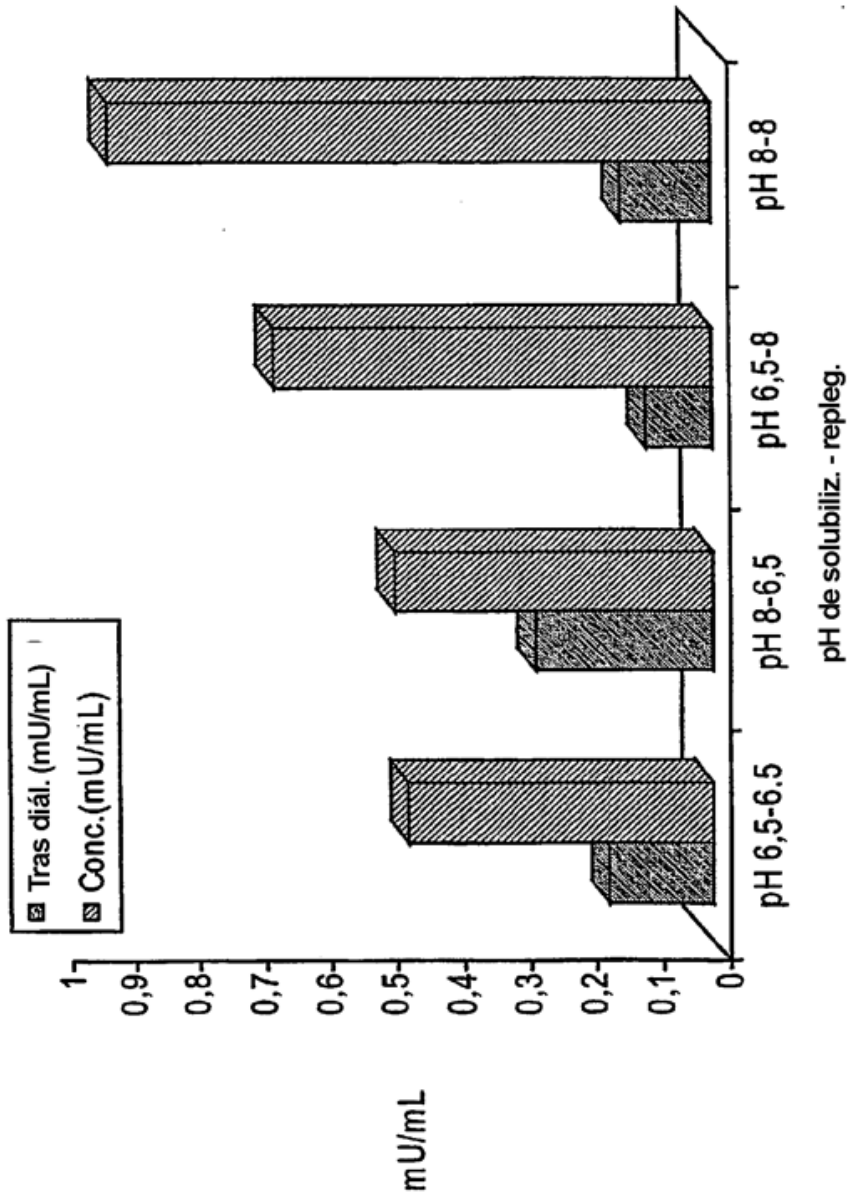


FIG. 16

Efecto del pH sobre las actividades específicas de MBP-GalNAcT2(D51)

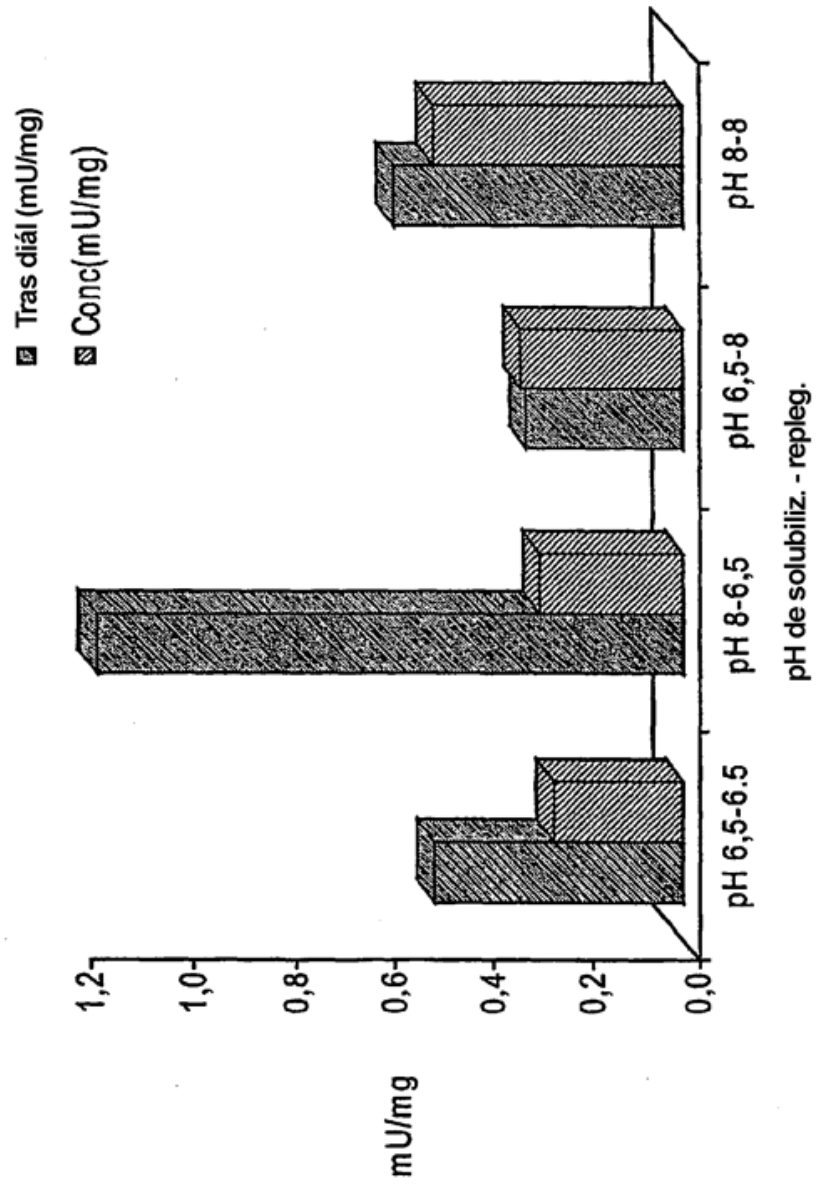


FIG. 17

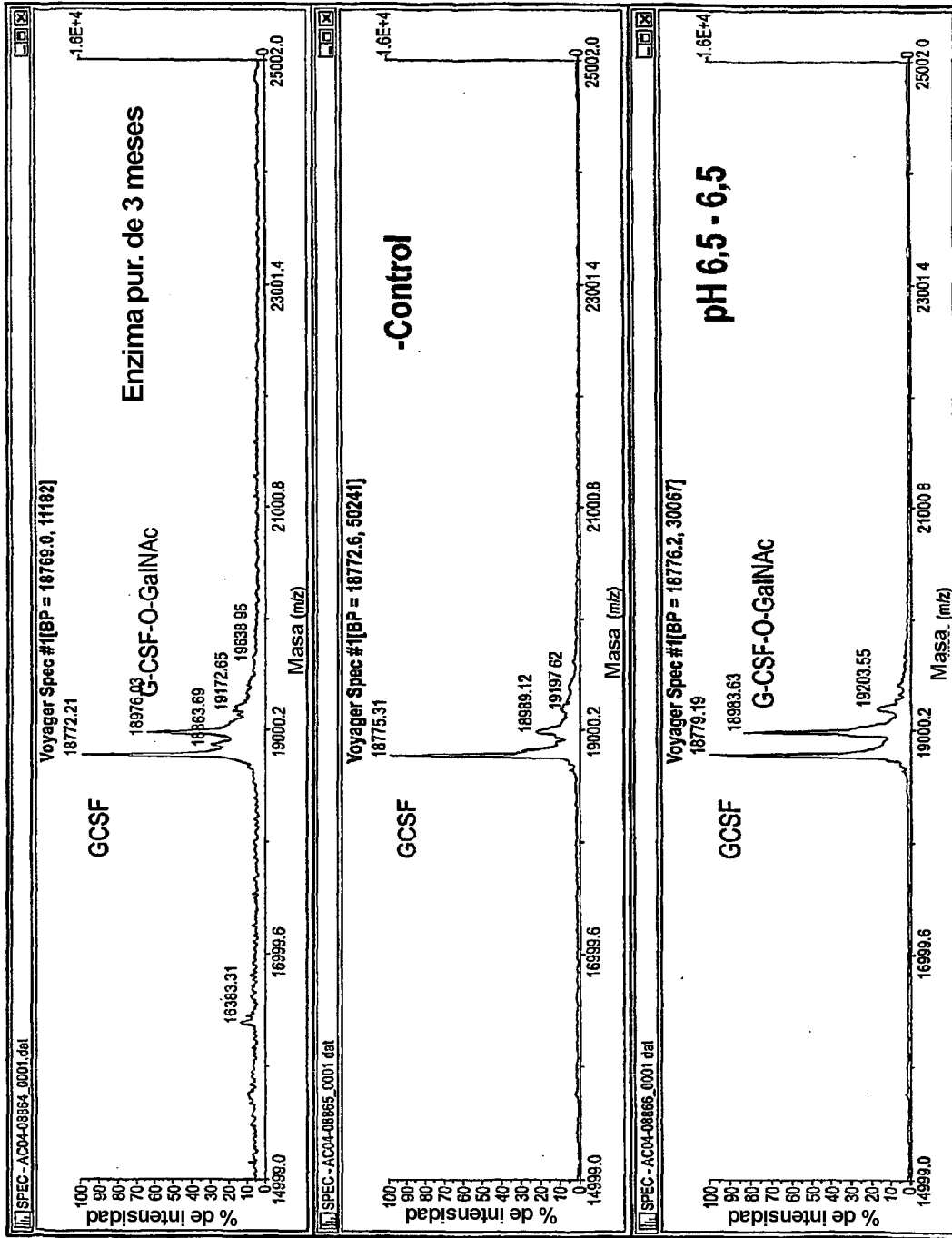


FIG. 18A

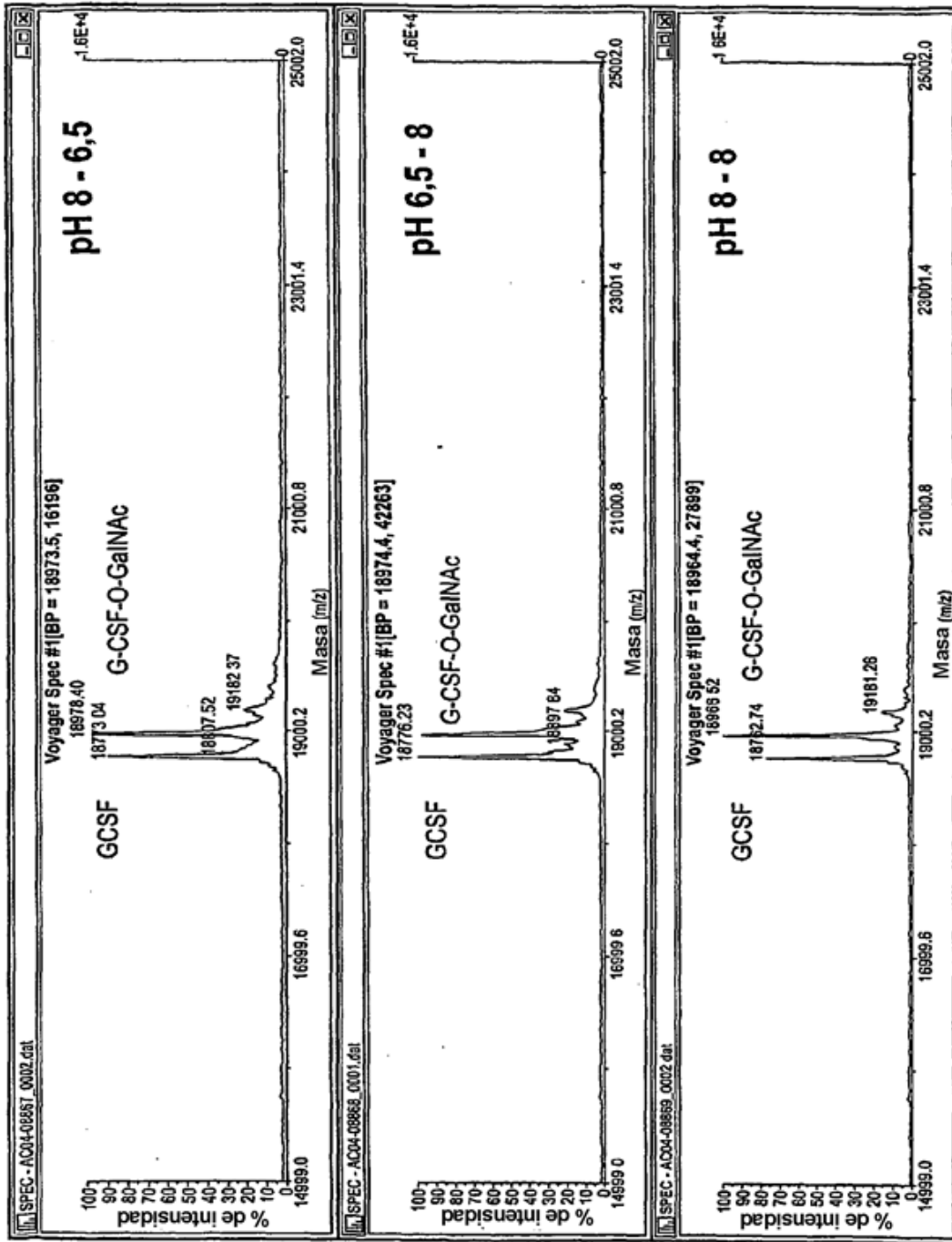


FIG. 18B

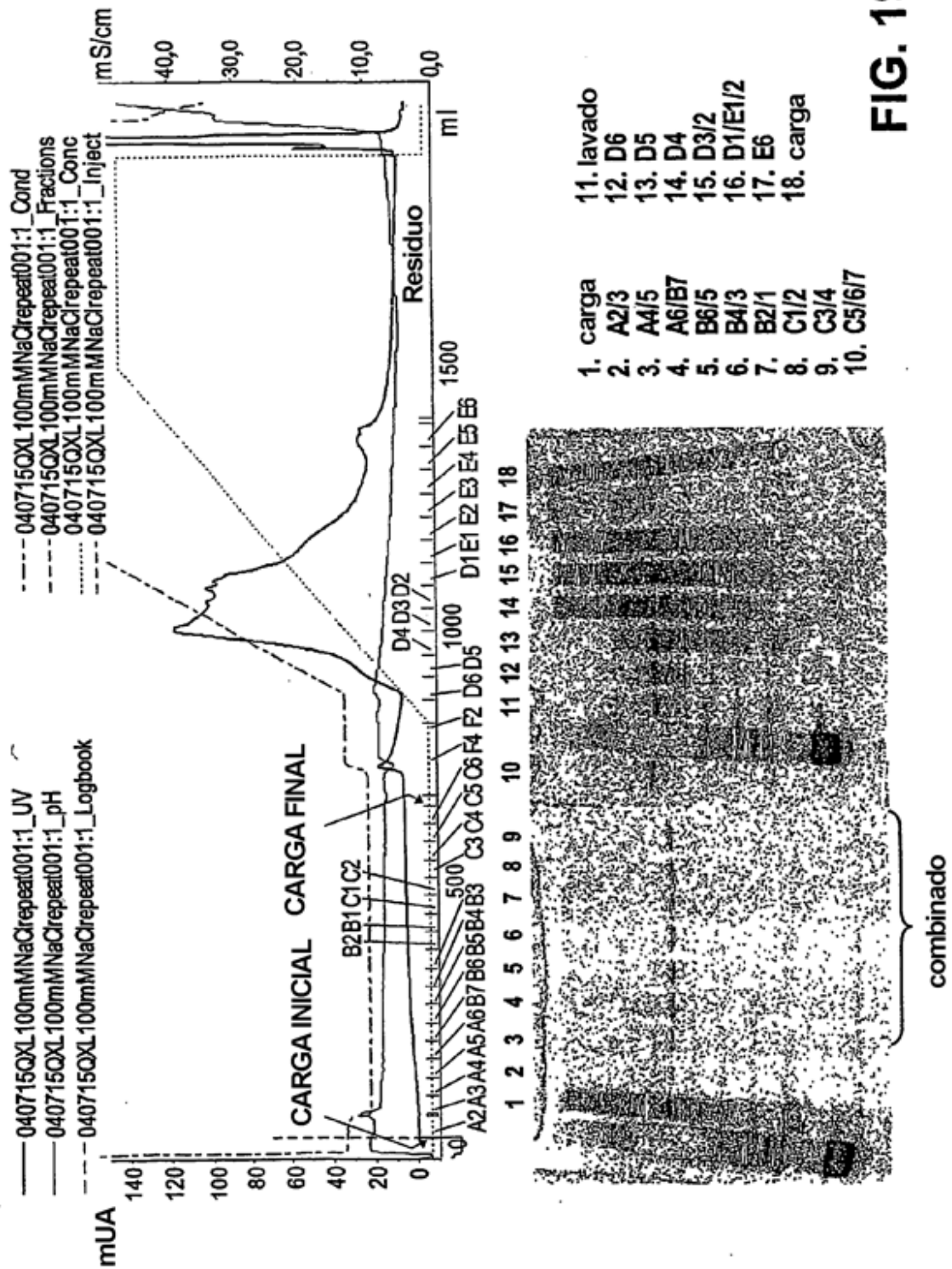


FIG. 19

	Volumen (ml)	Actividad (U/l)	A280
Carga	890	1,5	0,110
FT _{A4-C4}	670	9,2	ND
FT _{C5-C7}	120	1,0	0
Lavado	138	3,6	0,100
D6	45	4,5	0
D5	45	2,4	0,026
D4	45	2,0	0,108
D3/2	90	1,1	0,179
E6		0,0	0,017

FIG. 20

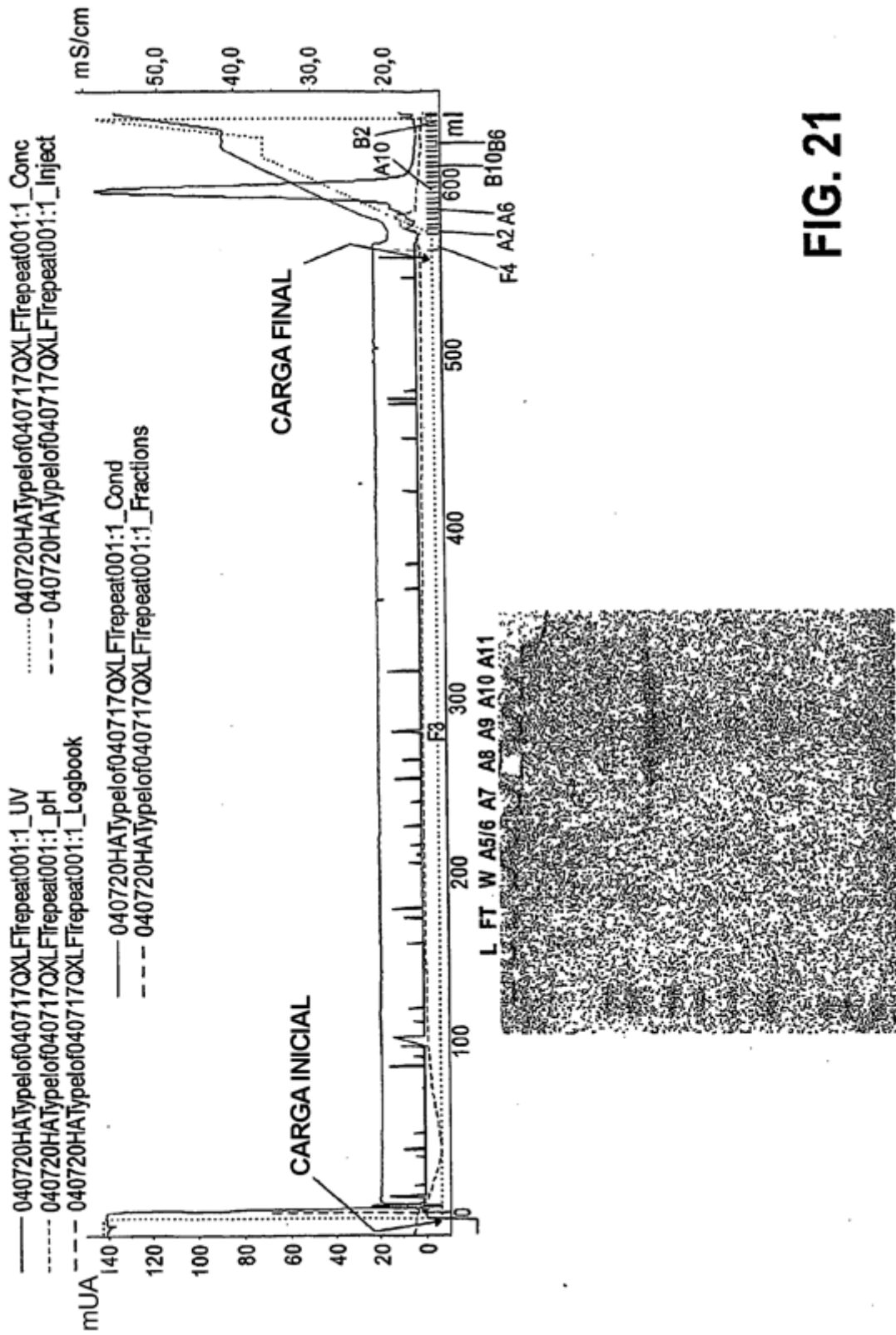


FIG. 21

A4-C4 combinada y pH ajustado a 7,0 usando carga de 670 mL de HCl 1 mM pH 7,0 16 mS/cm

	Volumen (ml)	Actividad (U/l) tras el diálisis	Actividad (U/l) tras el diálisis	A280	A280/1,51 (mg/ml)	Actividad (U)	Masa (mg)	Actividad específica (U/mg)
Carga	670	9,2	NA	NA				
FT	670	0,0	NA	0,122				
Lavado	9	2,9	NA	-0,013				
A5/6	6	1,1	NA	-0,005				
A7	3	0,1						
A8	3	1,3						
A9	3	4,6						
A10	3	2,4						
A11	3	0,4						
			19,3 (13 ml)	0,180	0,119	0,25	1,55	0,16

FIG. 22

Comparación de MBP con cola de MBP-SBD₃₉ en ST3Gal3 en purificación de columna de ciclodextrina

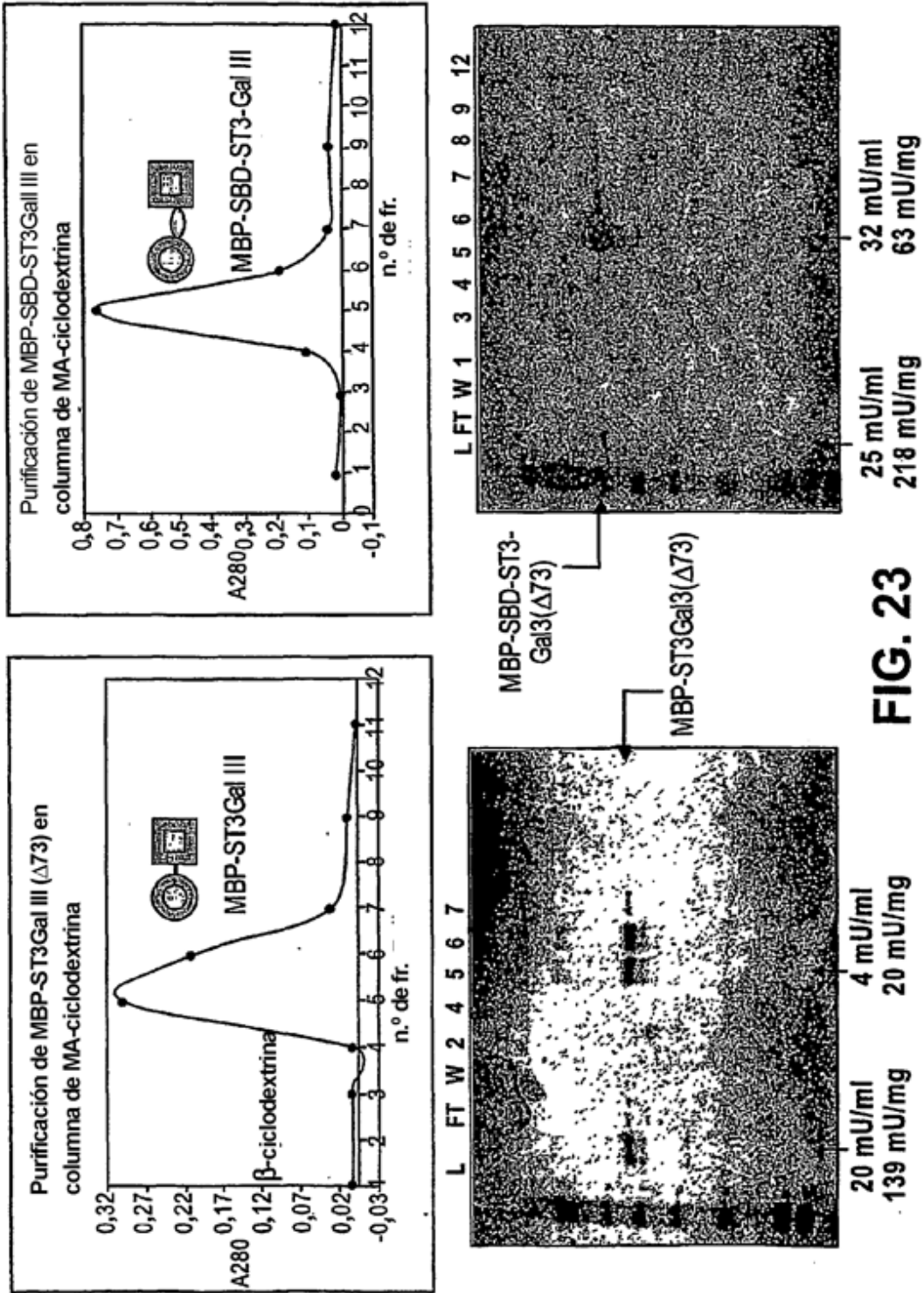


FIG. 23

Proteína de fusión MBP-pST3Gall

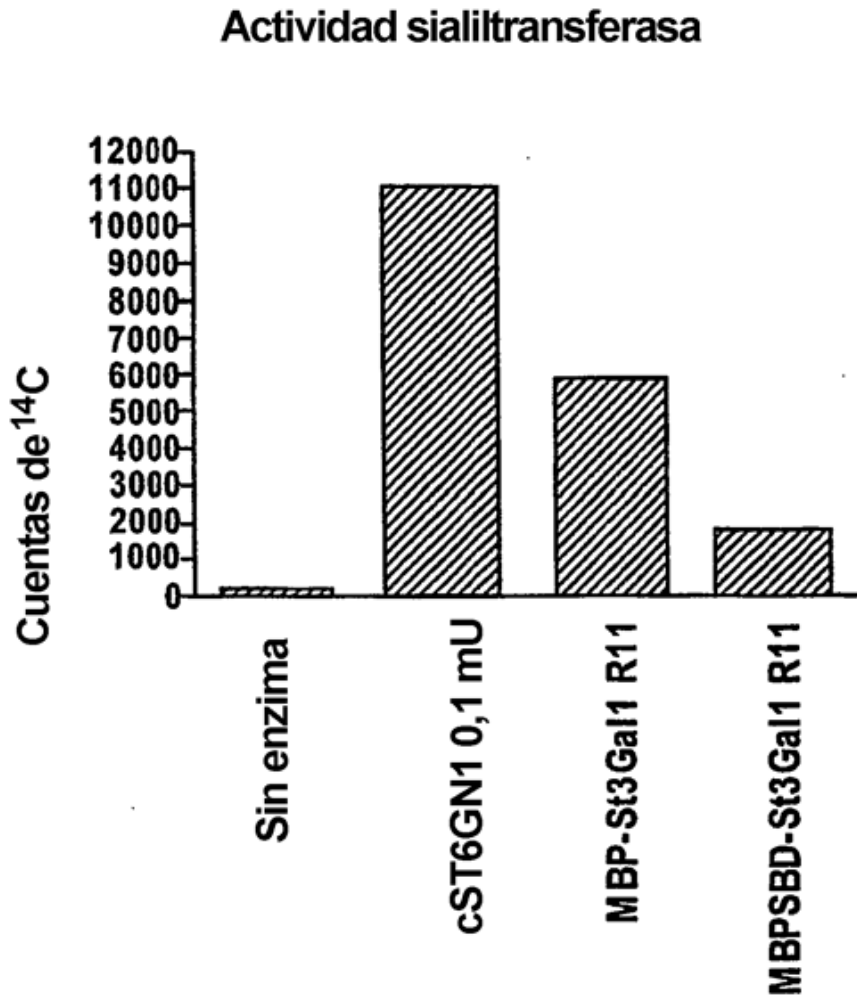
MKIEEGKLYIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDJIFWAHDRF
 GGYAQSGLLAEITPDKAFQDKLYPFTWDVRYNGKLIJAYPIAVEALSILYNNKDLLPNPPKTTWEEIPALD
 KELKAKGKSALMFNLQEPYFTWPLAADGGYAFKYENGGYDIKDVGVNDNAGAKAGLFLVLDLIKNNKH
 MNADTDYSLAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPFKGQPSKPFVGVLSAGINAASP
 NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNPQMSAF
 WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNLGIETRSEFGSELSNFKMLMKYPYR
 PCTCRICEEQRVSAWDFERNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLNDTIRELFQVVP
 GNVDPLEKRLVSCRRCAVVGNSGNLKEYYGPQIDSHDFVLRMNKAPTEGFEADVGSKTIHFFVYPE
 SFRELAQEVSMILVPPFKTTDLEWVISATTTGRISHTYVVPAAKIKVKKKELIYHPAFIKYVFDRWLQGH
 GRYPSTGILSVIFSLHICDEVLDLYGFGADSKGNWHHYWENNPNSAGAFRKTGVHDGDFESNVTTILASIN
 KIRIFKGR

FIG. 24A

Proteína de fusión MBPSBD-pST3Gall

MKIEEGKLYIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDJIFWAHDRF
 GGYAQSGLLAEITPDKAFQDKLYPFTWDVRYNGKLIJAYPIAVEALSILYNNKDLLPNPPKTTWEEIPALD
 KELKAKGKSALMFNLQEPYFTWPLAADGGYAFKYENGGYDIKDVGVNDNAGAKAGLFLVLDLIKNNKH
 MNADTDYSLAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPFKGQPSKPFVGVLSAGINAASP
 NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNPQMSAF
 WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNLGIETRSEFGSELSNFKMLMKYPYR
 SGSVTSTSKTTATAASKTSTSTSSCTTPTAVAVTFDLTATTYGENIYLVGSISQLGDWETSJDIALSAD
 KYTSSDPLWYVTVTLPAGESEFYEYKFIRESDDSVWEESDPNREYTVPOACGTSTATVTDIWRGSELSEN
 FKKLMKYPYRPPCTCRICEEQRVSAWDFERNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLN
 DTIRELFQVVPVGNVDPPLLEKRLVSCRRCAVVGNSGNLKEYYGPQIDSHDFVLRMNKAPTEGFEADV
 SKTIHFFVYPPESFRELAQEVSMILVPPFKTTDLEWVISATTTGRISHTYVVPAAKIKVKKKELIYHPAFIK
 YVFDRWLQGHGRYPSTGILSVIFSLHICDEVLDLYGFGADSKGNWHHYWENNPNSAGAFRKTGVHDGDF
 ESNVTTILASINKIRIFKGR

FIG. 24B



Prueba de pST3Gal1 para la actividad tras el
replegamiento de Hampton

FIG. 25

MKIEEGKLVWINGDKGYNGLAEVGKKFEKDTGKVTVEHPDKLEEKFPQVAATGDGPDIFWAHDRF
GGYAQSGLLAETPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLYNKDLLPNPPKTWEIPALD
KELKAKGKSALMFNLQEPYFTWPLAADGGYAFKYENKDYDKDYGVDNAGAKAGLFLVDLKNKH
MNADTDYSIAEAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQSPKPFVGLSAGINAASP
NKELAKEFLENYLITDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNNGIEGRISEFGSEHLDKVPRTPGAL
STRKTPMATGAVPAKVKVQATKSPASSPHPTIRRRQRKASEFKSEPRWDFEEYSLDMSSLQT
NCSASVKIKASKSPWLQNFILNITFLDSGRFTQSEWNRLHFAPFFGFMELNQSLVQKVVTTRFP
PVRQQQLLASLPTGYSKCIITCAVVGNGGILNDSRVGREIDSHDYVFRLSGAVIKGYEQDVGTRT
SFYGFATFSLTQSILGRRGFQHVPLGKDVRYLHFLLEGTRNRYEWLEAMFLNQTLAKTHLSWFR
HRPQEAFRNALDLDRYLLHHPDFLRVMKNRFLRSKTLDTAXWRIYRPTTGALLLTLALHLCDKV
SAYGFITEGHERFSDHYDTSWKRLIFYINHDFRLERMVWKRLEHDEGHWLYQRPOSDKAKN

FIG. 26A

MKIEEGKLVWINGDKGYNGLAEVGKKFEKDTGKVTVEHPDKLEEKFPQVAATGDGPDIFWAHDRF
GGYAQSGLLAETPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLYNKDLLPNPPKTWEIPALD
KELKAKGKSALMFNLQEPYFTWPLAADGGYAFKYENKDYDKDYGVDNAGAKAGLFLVDLKNKH
MNADTDYSIAEAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQSPKPFVGLSAGINAASP
NKELAKEFLENYLITDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNNGIEGRISEFGSKEPQTKPSRHQKIE
NIKERSLQSLAKPKSQAPTRARRITTYAEFVPENNALNTQTPKAHTIGDRGKEANQAPPEEQDK
VPHTAQRAAWKSPEKEKTMVNTLSPRGDAGMASGRTEAQSWSQDTKITTCGNGGGQTRKLT
SRTVSEKHQKAATTAKTLIPKSQHRMLAPTGAVSTRIRQKGVTTAVIPPEKPKPATPPAPFQ
SPTTQRNQRKKAANFKSEPRWDFEEKYSFEIGGLQITICPDSVKIKASKSLWLQKLFNLTLFLDS
RHFNQSEWDRLEHFAPPPFGFMELNYSLVQKVVTRFPVPPVQQQLLASLPAGSLRCITCAVVGNG
GILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRKTSFYGFTAFSLTQSLLILGNRFGKKNVPLGK
DVRYLHFLLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLHHPDFLRVM
KNRFLRSKTLDGAHWRIYRPTTGALLLTLALQLCDQVSAYGFITEGHERFSDHYDTSWKRLIFY
INHDFKLEREVWKRLEHDEGHRLYQRPGGTAKAKN

FIG. 26B

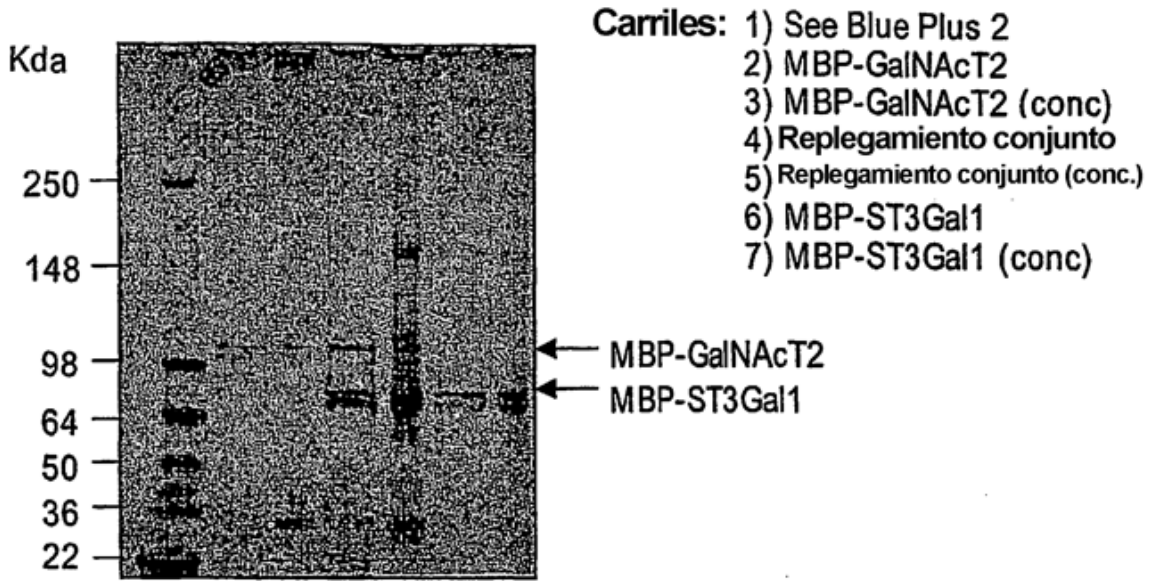


FIG. 27A

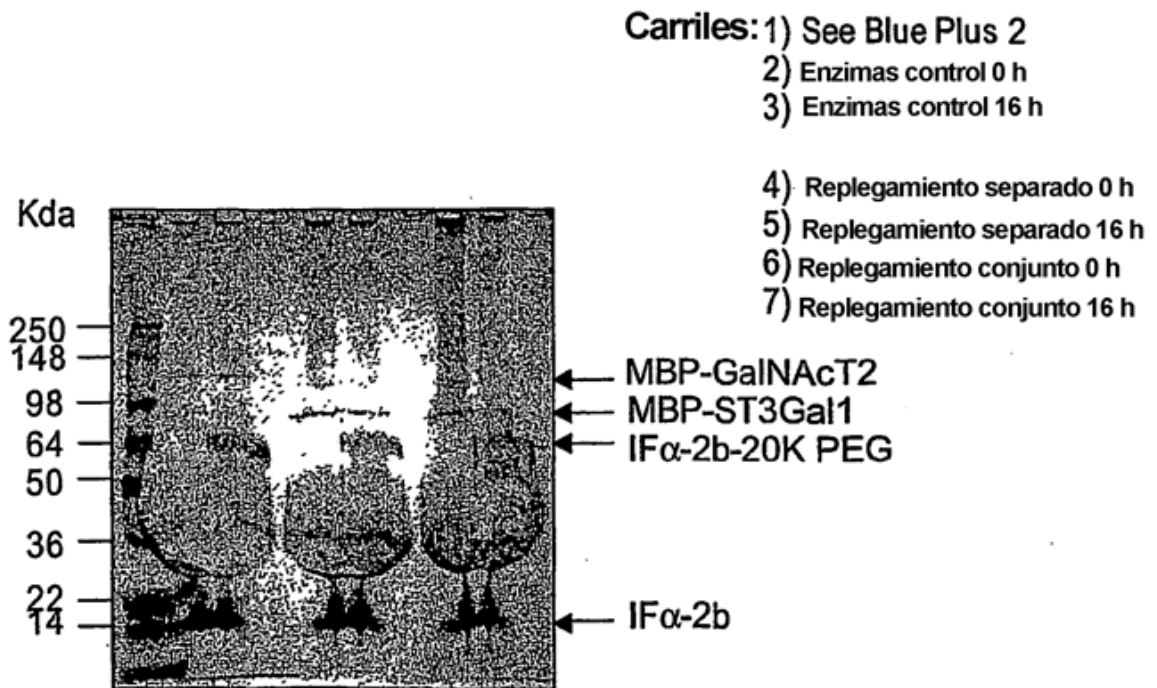
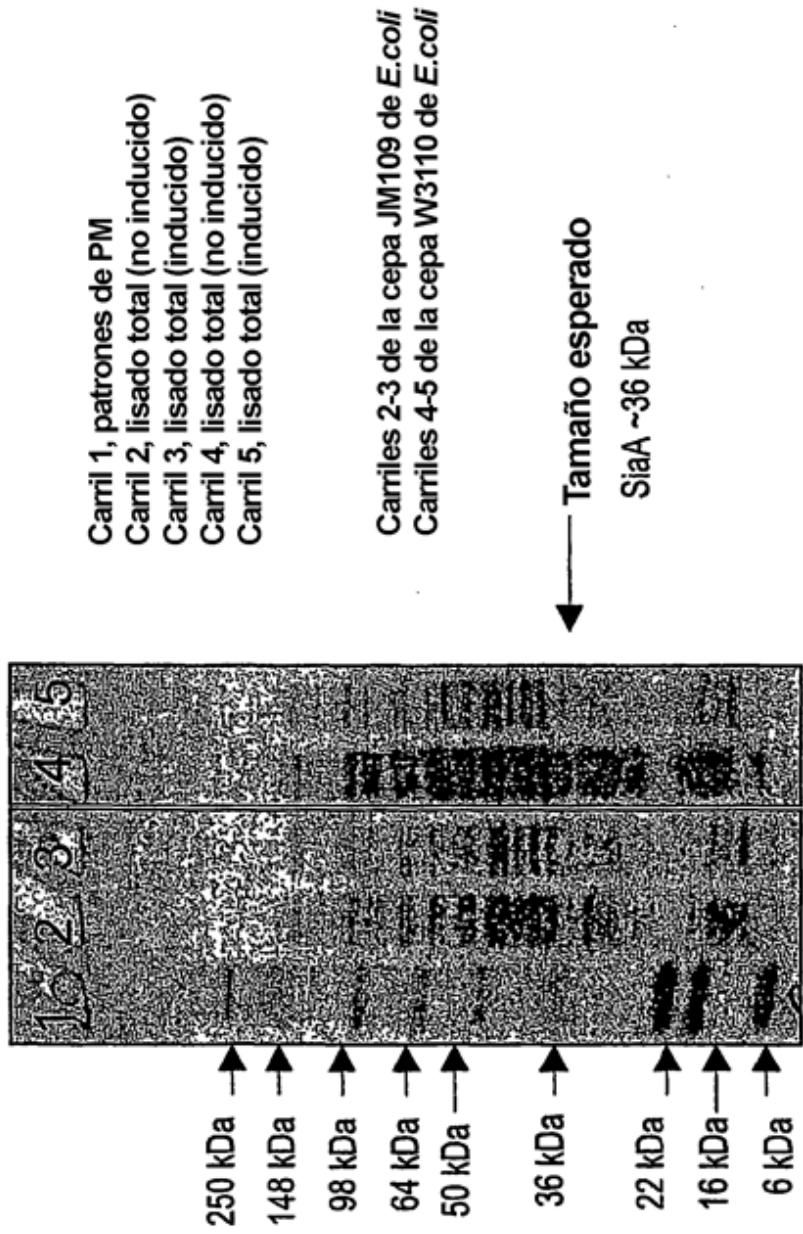


FIG. 27B

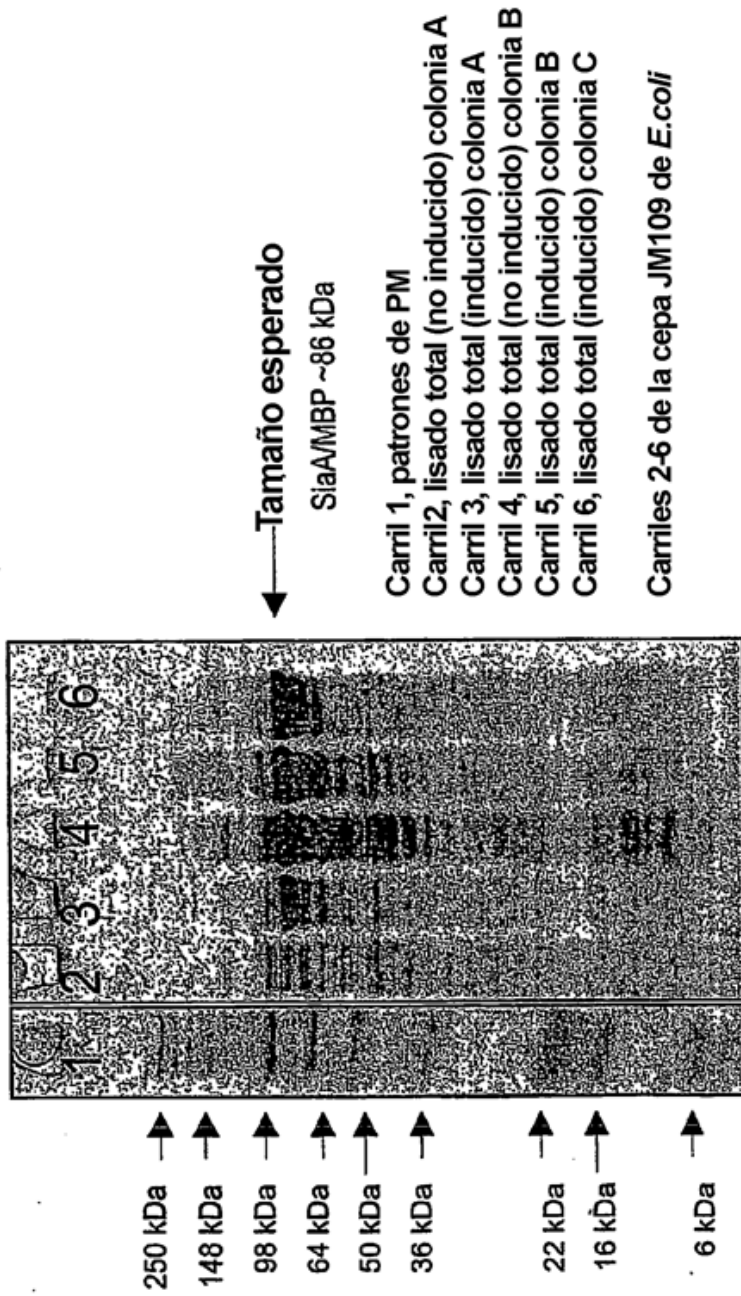
Expresión de SiaA inducida en *E. Coli*



No hay ninguna banda inducible obvia en la masa esperada de 36 kDa para la proteína SiaA nativa

FIG. 28

Expresión de SiaA/MBP inducida en *E.coli*



Producción de alto nivel de SiaA(MBP incluso en ausencia de inducción por IPTG. Compárese con la figura X en la que la producción de SiaA no es obvia. La presencia de la pareja de fusión (MBP) conduce a altos niveles de expresión.

FIG. 29

1 mkfrepllgg saampgaslq racrllvavc alhlgvtlvy ylagrdlrrl pqlvgvhhppl
 61 qgsshgaaai gqpsgelrlr gvappplqn sskprsraps nldayshpqp gpgpgsnlts
 121 apvpstttrs ltacpeespl lv9pmliefn ipvdlklieq qnpkvlggr ytpmdcisph
 181 kvaiiilfrn rqehlkywly ylhpmvqrqq ldygiyving agesmfnrak llnvgfkeal
 241 kdydyncfvf sdvdlipmnd hntyrctfsqp rhisvamdkf gfslpvvyf ggvsalskqq
 301 flsingfpnn ywgwggeddd iynrlafrgm svsrpnavig kcrmirhsrd kknepnpqrf
 361 driahtketm lsdglnsilty mvlevqrypl ytkitvdigt ps

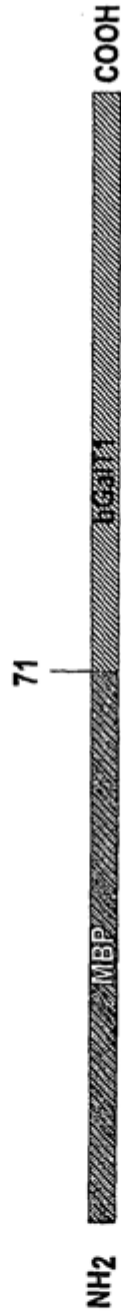
FIG. 30

Constructos de GalT1 (β 1,4GalT1) (b) bovina

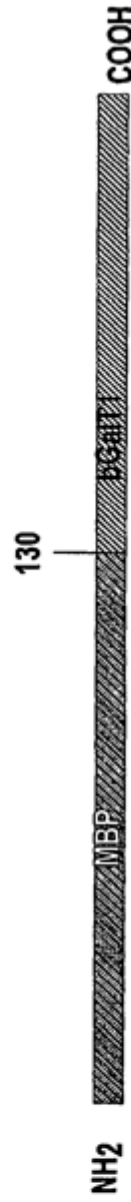
UDP-galactosa β -N-acetilglucosaminida β 4-galactosiltransferasa (EC 2.4.1.38)



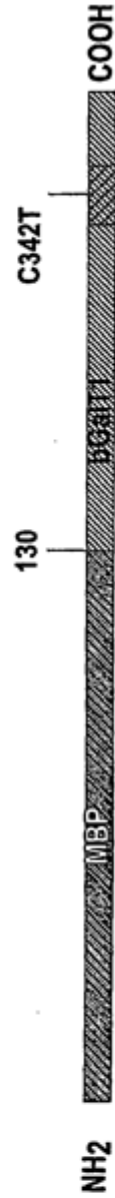
GalT1(40) (S96A+C342T)



MBP-GalT1(D70)



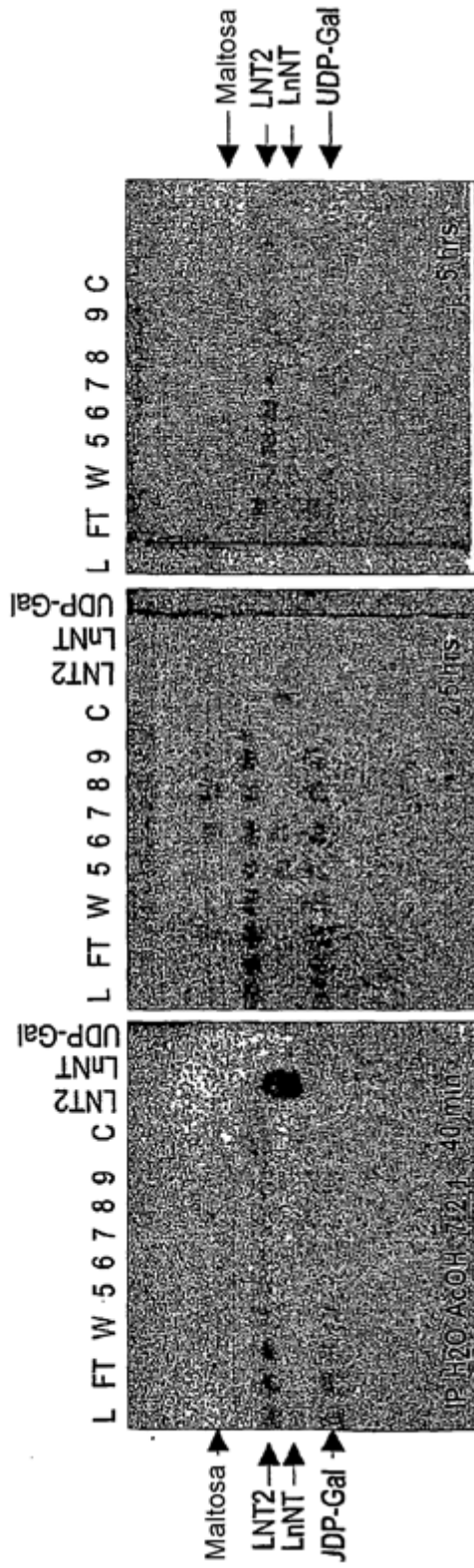
MBP-GalT1(D129)



MBP-GalT1(D129) C342T

FIG. 31

Ensayo de CCF de GalT1



Reacciones de LnNT con

L: muestra cargada (MBP-GalT1 (70) replegada, dializada en columna de amilosa)

FT: Fracción no retenida de la columna de amilosa

W: Lavado de la columna de amilosa

5: Fr. n.º 5 eluida con maltosa

6: Fr. n.º 6 eluida con maltosa

7: Fr. n.º 7 eluida con maltosa

8: Fr. n.º 8 eluida con maltosa

9: Fr. n.º 9 eluida con maltosa

C: Control con agua

FIG. 32

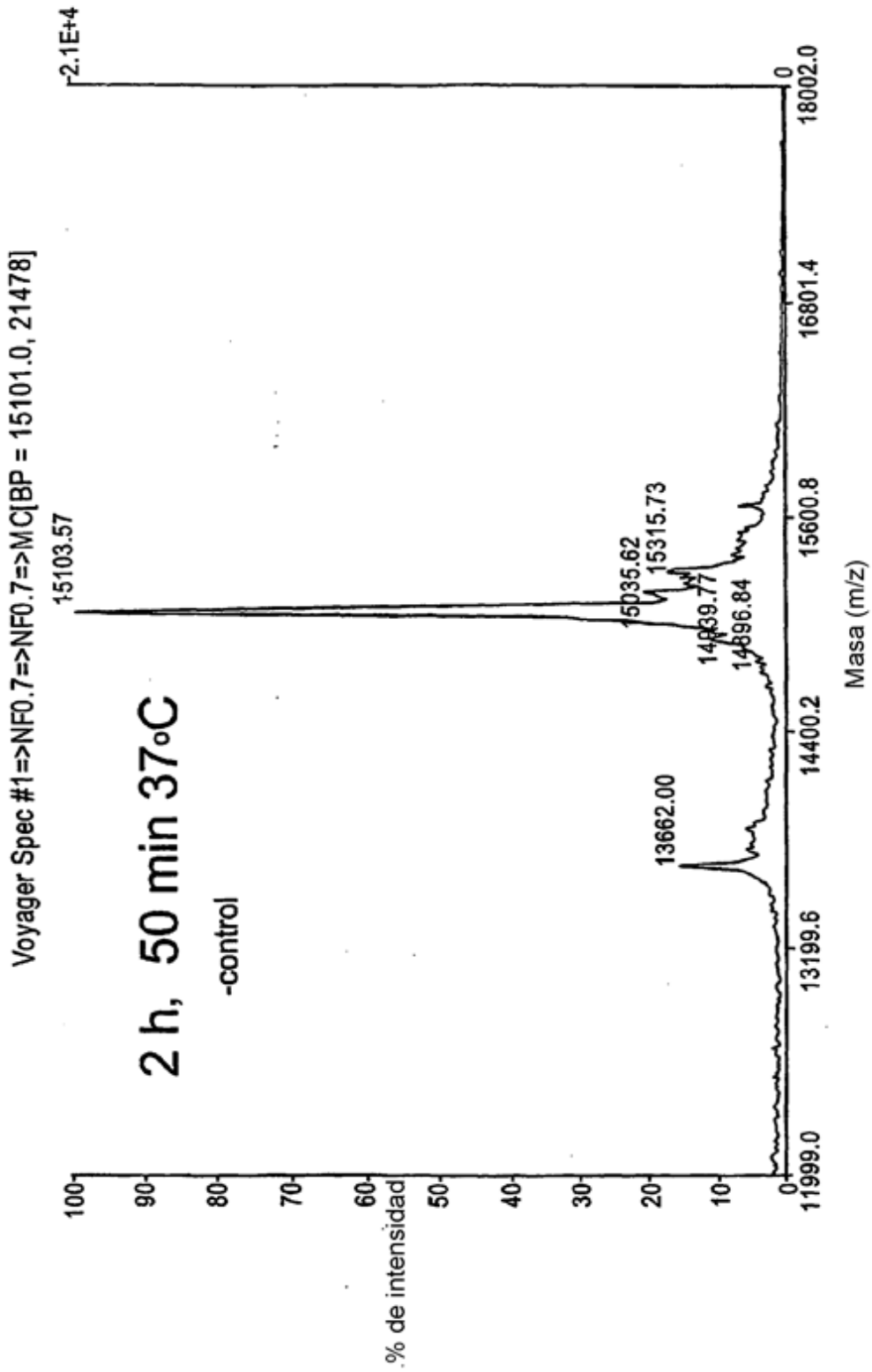


FIG. 33

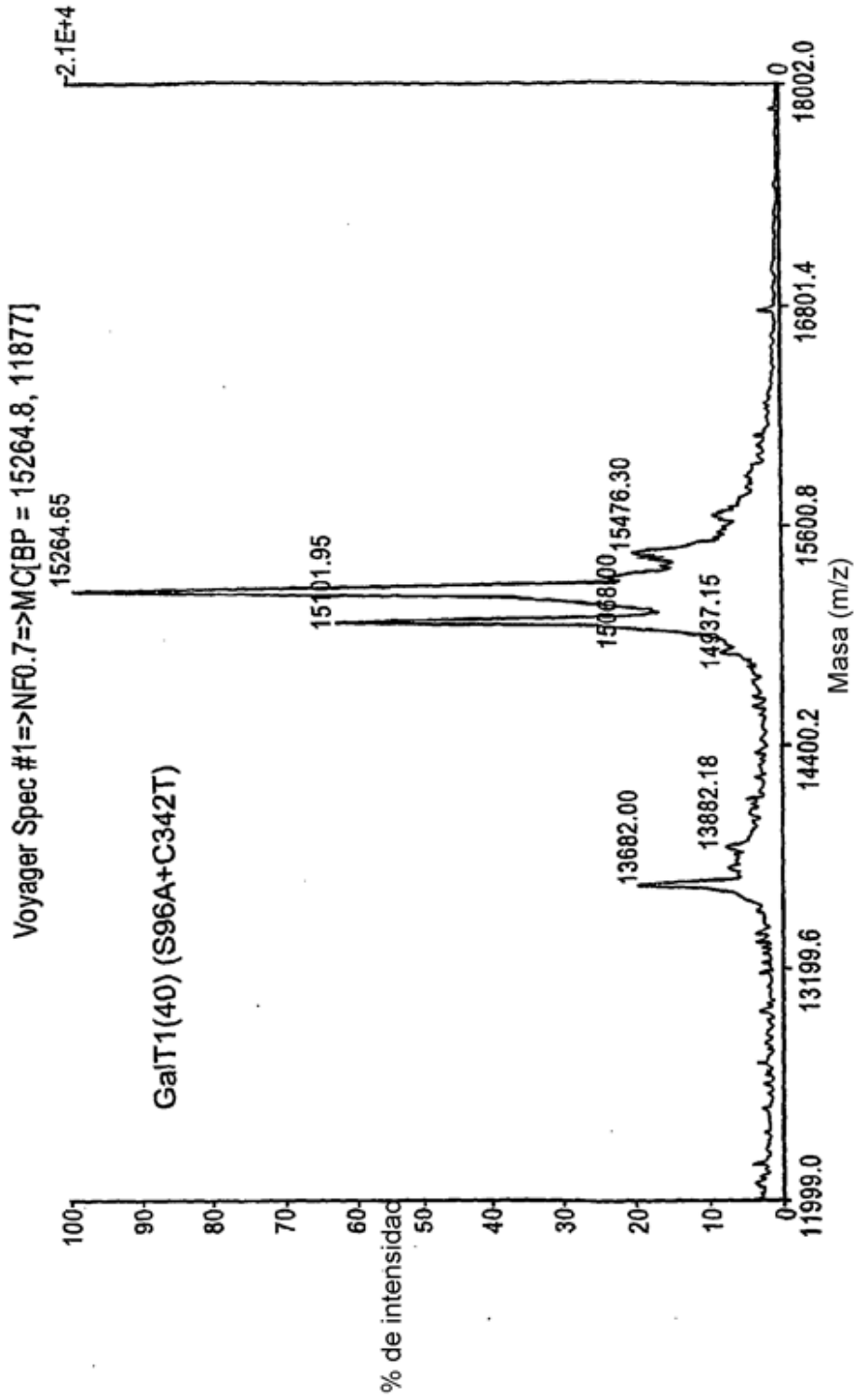


FIG. 33 (CONT.)

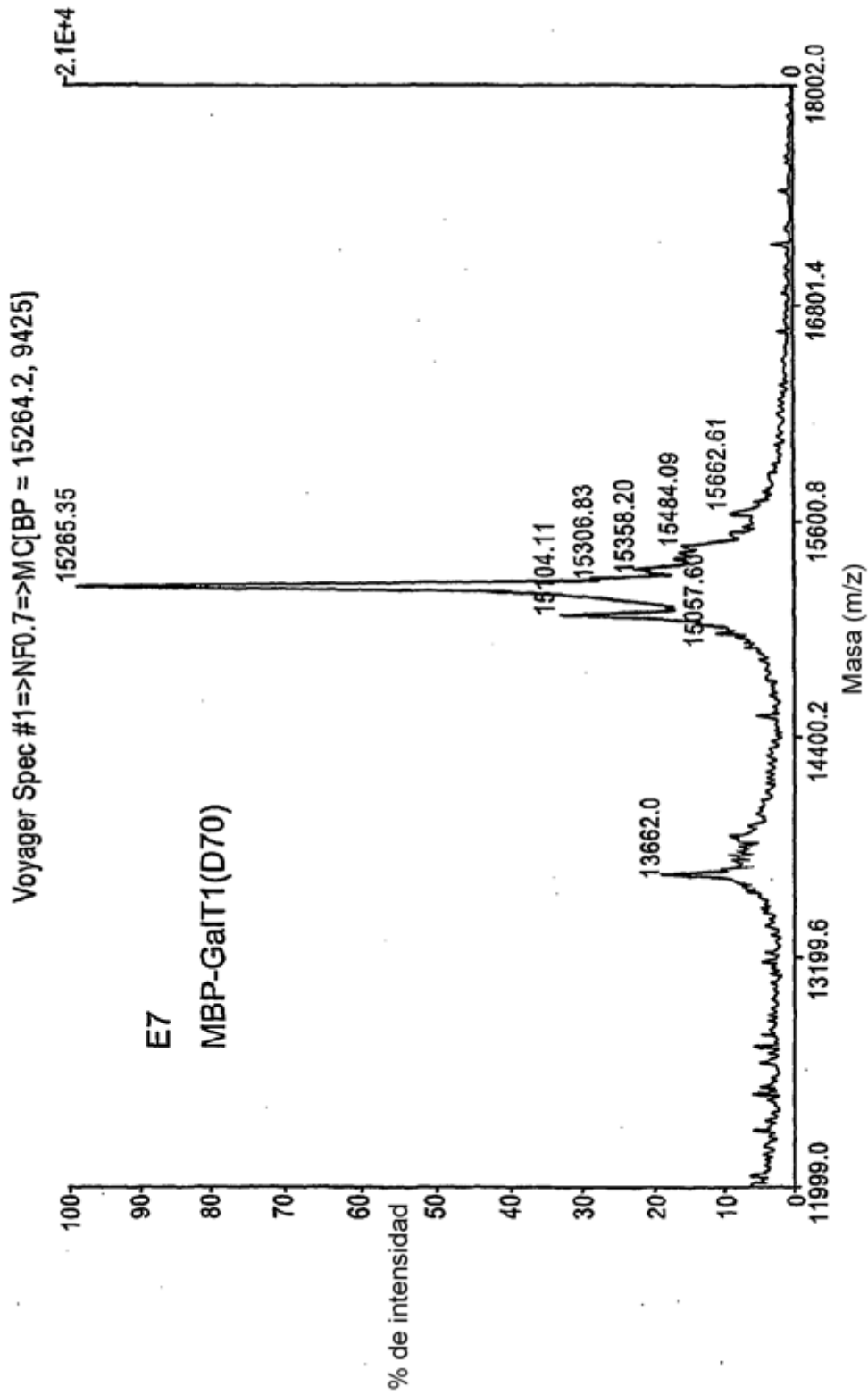


FIG. 33 (CONT.)

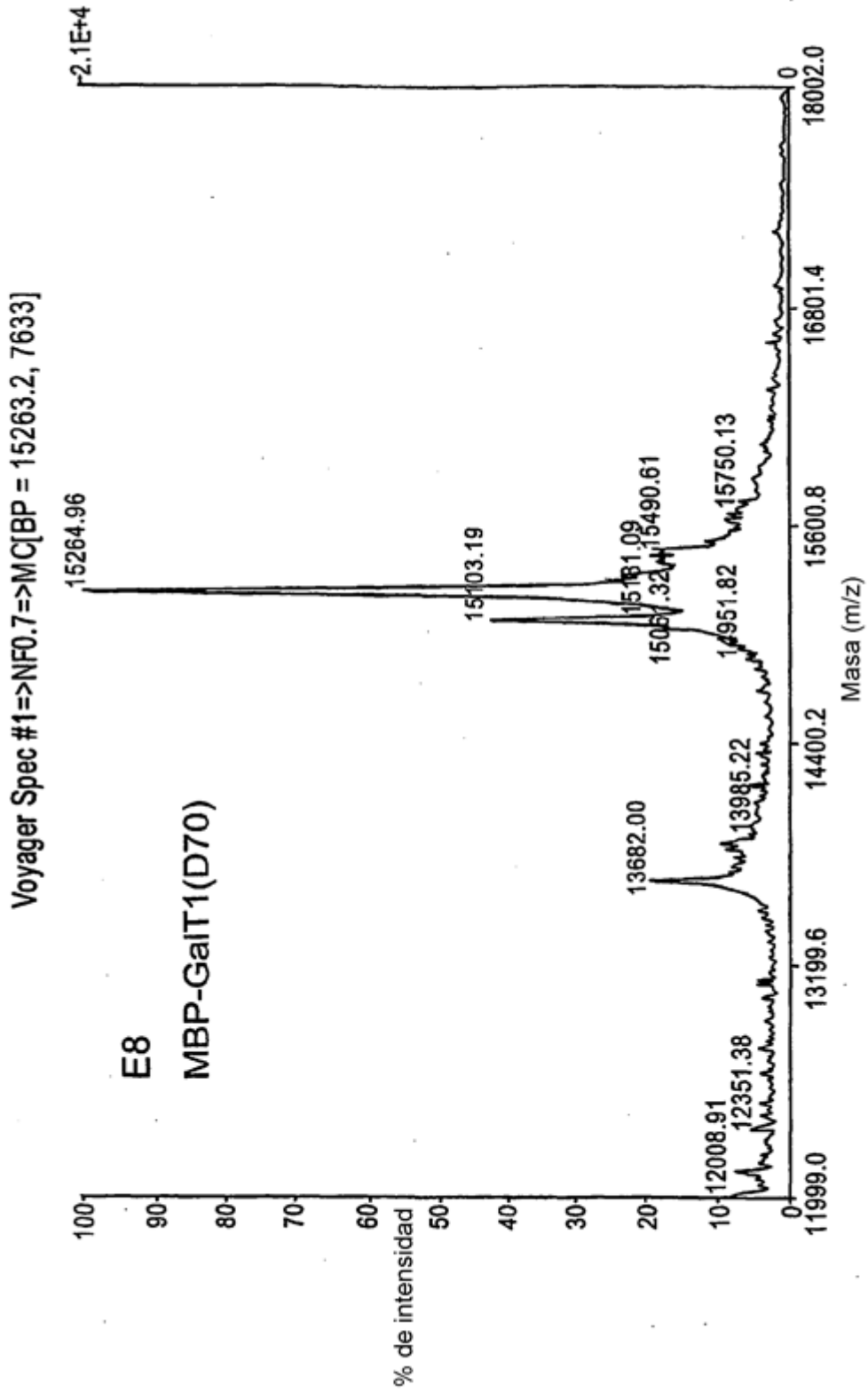


FIG. 33 (CONT.)

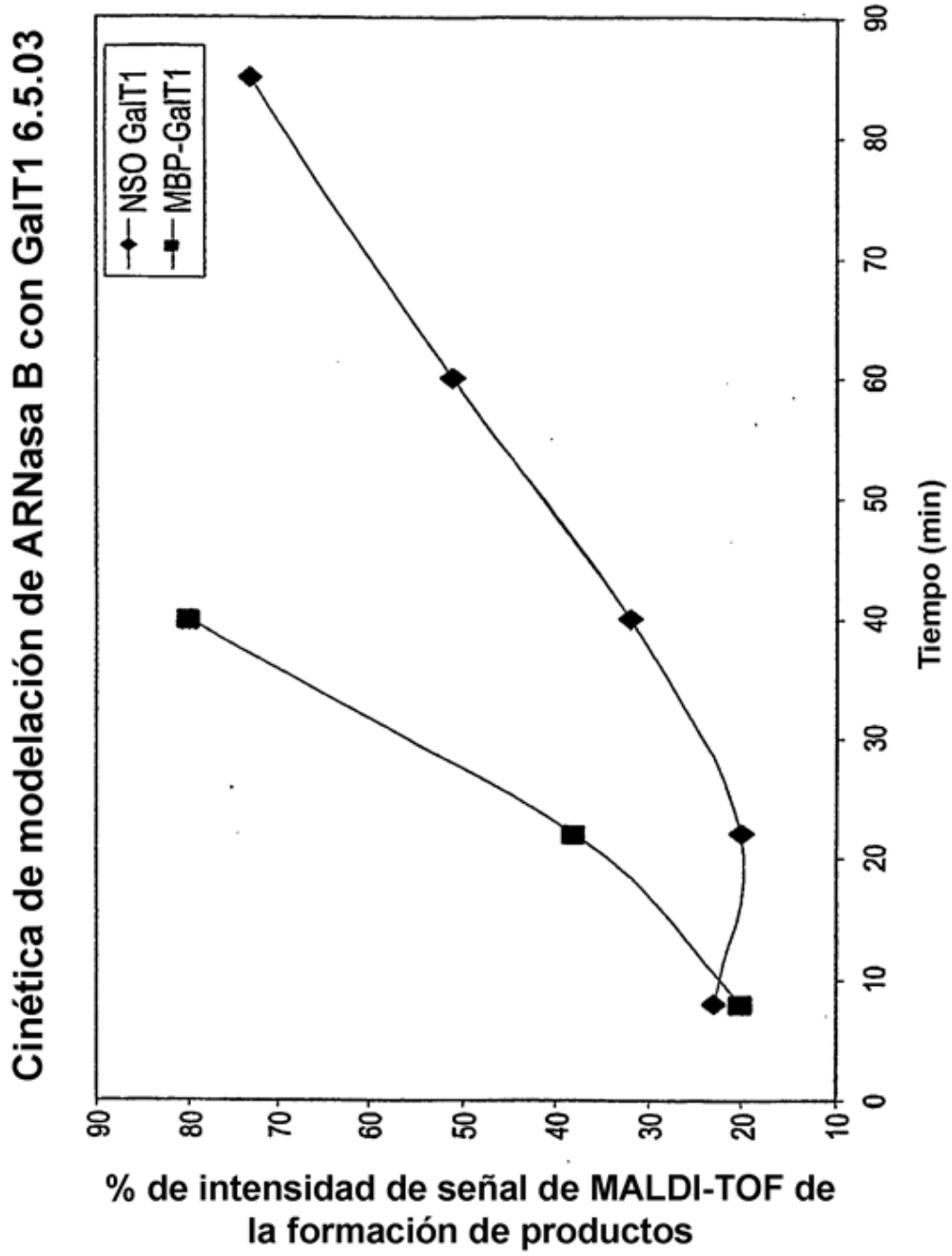
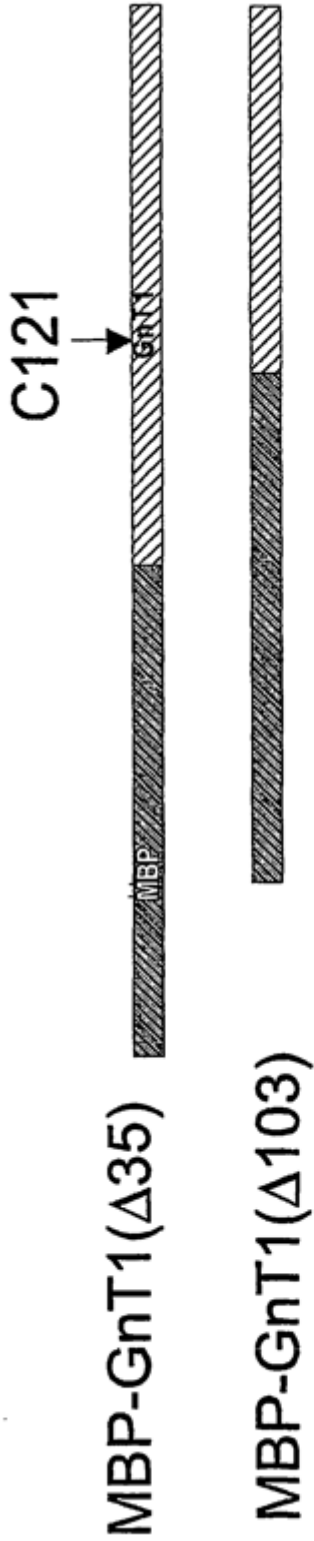


FIG. 34

Constructos de GnT1



1 mlkksaglv lwgailfvaw nallllffwt rpapgrppsv salgdpasl trevirlaqd
 61 aevelerqrg llqqigdals sgrgrvptaa ppaqprvpvt papavipily iacdrstvrr
 121 cidkllhyrp saelfpiivs qdcgheetaq aiasygsavt hirqpdllssi avppdhrkfq
 181 gyykiarhyr walgqvfrqf rfpaavvved dlevapdffe yfratypilk adpslwcvs
 241 wndngkeqmv dasrpellyr tdffpqlgw1 llaelwaele pkwpkafwdd wmrrpegrqg
 301 racirpeisr tmtfgrkgvs hgqffdqhlk fiklnqqfvh ftqldlsylq reaydrdfla
 361 rvygapqlqv ekvrtndrke lgevrvqytg rdsfkafaka lgvmdllksg vpragyrgiv
 421 tfqfgrrrvh lappptwegy dpswn

FIG. 35

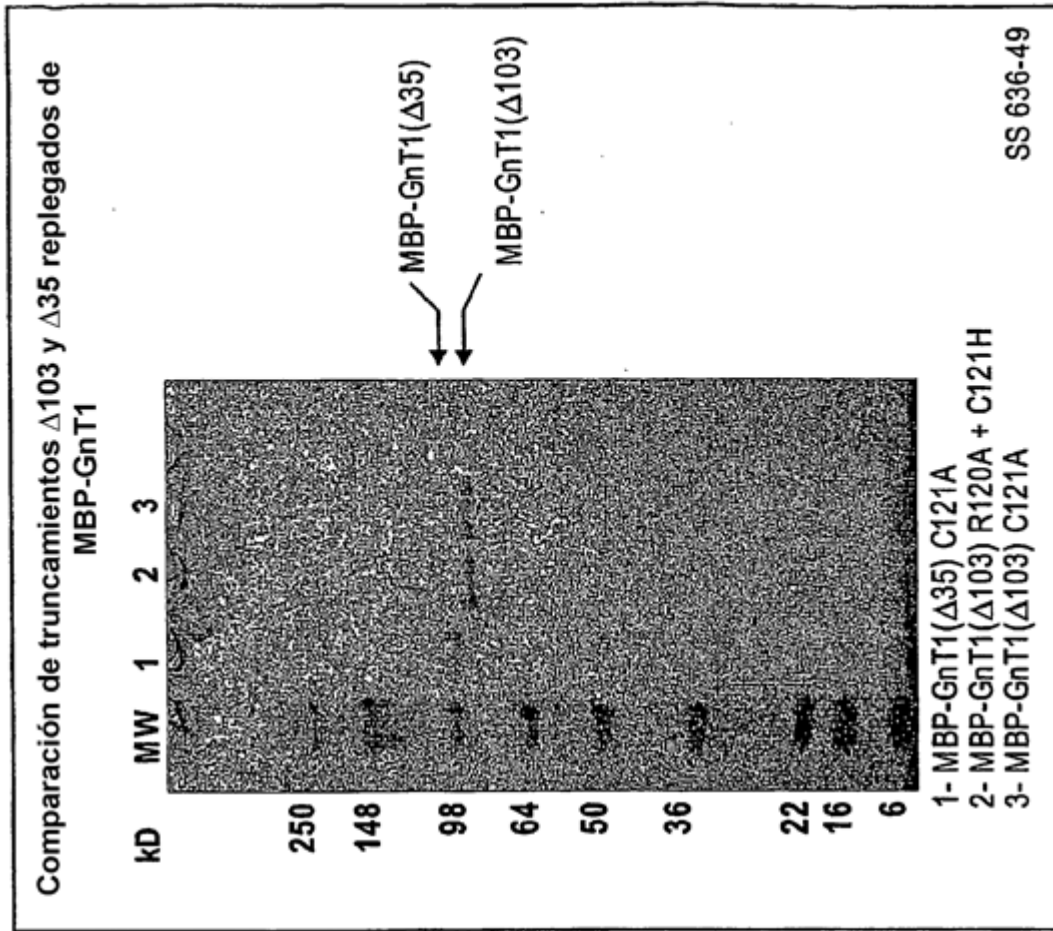
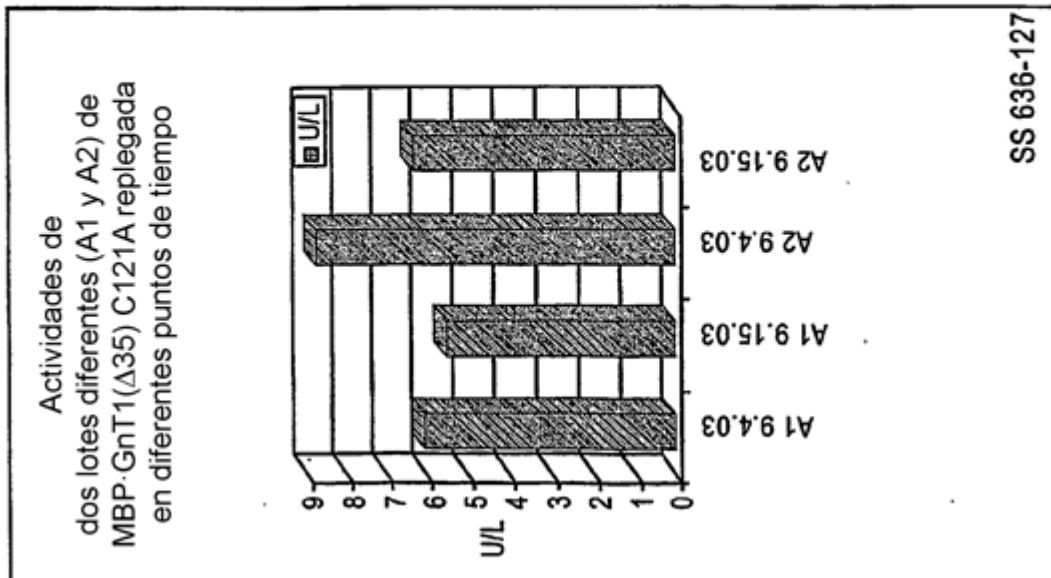


FIG. 36



1 mapmrkkstl klltllvlfi fltsfflnys htvtawfp kqmviesen fklmkypyr
 61 pctctrciee qrvsawfder fnrsmqpllt aknahleedt ykwrlrlqre kqpnnlndti
 121 relfqvppgn vpllekrlv scrrcavvgn sgnlkesyyg pqidshdfvl rmnkaptgef
 181 eadvgsktth hfvyfesfre laqevsmilv pfkttdleww isattgtis htyvvpvaki
 241 kvkkekilii hpafikyvfd rwlqghgryp stgilsvifs lhicdevdly gfgadskgnw
 301 hhywennpsa gafkrktgvhd gdfesnvtti lasinkirif kgr

FIG. 37

ST6GalNAcTI humana

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQS
LAKPKSQAPTRARRTTIYAEPVPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDK
VPHTAQRAAWKSPEKEKTMVNLSPRGQDAGMASGRTEAQSWSQDTKTTQGNQ
GQTRKLTASRTVSEKHQGKAATAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIP
PKEKQPQATPPPAPFQSPTTQRNQRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSV
KIKASKSLWLQKLFPLNLTLFLDSRHFNQSEWDRLEHFAPPFGMELNYSLVQKVVT
RFPPVPQQQLLLASLPAGSLRCITCAVVGNGGILNNSHMGQEIDSHDYVFRLSGALIK
GYEQDVGTRTSFYGFATAFSLTQSLILGNRQFKNVPLGKDVRYLHFLEGTRDYEWLE
ALLMNQTVMSKNLFWFRHRPQEAFFREALHMDRYLLLHPDFLRYMKNRFLRSKTLTLD
GAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDYDTSWKRLIFYINH
DFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

FIG. 38A

ST6GalNAcTI de gallina

MGFLIRRLPKDSRIFRWLLILTVFSFIITSFSAALFGMEKSIFRQLKIYQSIHMLQVDTQ
DQQGSNYSANGRISKVGLERDIAWLELNTAVSTPSGEGKEEQKKTVPVAKVEEAK
EKVTVKPFPEVMGITNTTASTASVVERTKEKTTARPVPGVGEADGKRTTIALPSMKE
DKEKATVKPSFGMKVAHANSTSKDKPKAEPPASVKAIRPVTQAATVTEKKKLRAA
DFKTEPQWDFDDEYILDSSSPVSTCSSESVRAKAAKSDWLRDLFLPNITLFDKSYFNV
SEWDRLEHFAPPYGFMEELNYSLVVEVMSRLPPNPHQQLLLANSSSNVSTCISCAVVG
NGGILNNSGMGQEIDSHDYVFRVSGAVIKGYEKDVGKTSFYGFATAFSLVSSLQNLG
HKGFKKIPQGKHRYIHFLEAVRDYEWLKALLLDKDIRKGFNYYGRRPRERFDEDF
TMNKYLVAHPDFLRYLKNRFLKSKNLQKPYWRLYRPTTGALLLLTALHLCDRVSAY
GYITEGHQKYS DHYYDKEWKRLVFYVNHDFNLEKQVWKRLHDENIMKLYQRS

FIG. 38B

Proteína ST6GalNAcTI de ratón que comienza en el residuo 32 de la proteína de ratón nativa

DPRAKDSRCQFIWKNDASAQENQQAEPQVPIMTLSPRVHNKESTSVSSKDLKQER
EAVQGEQAEGKEKRKLETIRPAPENPQSKAEPAAKTPVSEHLDKLPRTPGALSTRKTP
MATGAVPAKKKVVQATKSPASSPHPTTRRRQRLKASEFKSEPRWDFEEYSLDMSSL
QTNCSASVKIKASKSPWLQNIFLNLTLFLDSGRFTQSEWNRLEHFAPPFGMELNQS
VQKVVTFRFPVVRQQQLLLASLPTGYSKCITCAVVGNGGILNDSRVGREIDSHDYVFR
LSGAVIKGYEQDVGTRTSFYGFATAFSLTQSILILGRRGFQHVPLGKDVRYLHFLEGTR
NYEWLEAMFLNQTAKTHLSWFRHRPQEAFFRNALDLDRYLLLHPDFLRYMKNRFL
RSKTLDTAHWRIYRPTTGALLLLTALHLCDKVSAYGFITEGHQRFS DHYYDYDTSWKRL
IFYINHDFRLERMVWKRLHDEGIWLYQRPSDKAKN

FIG. 38C

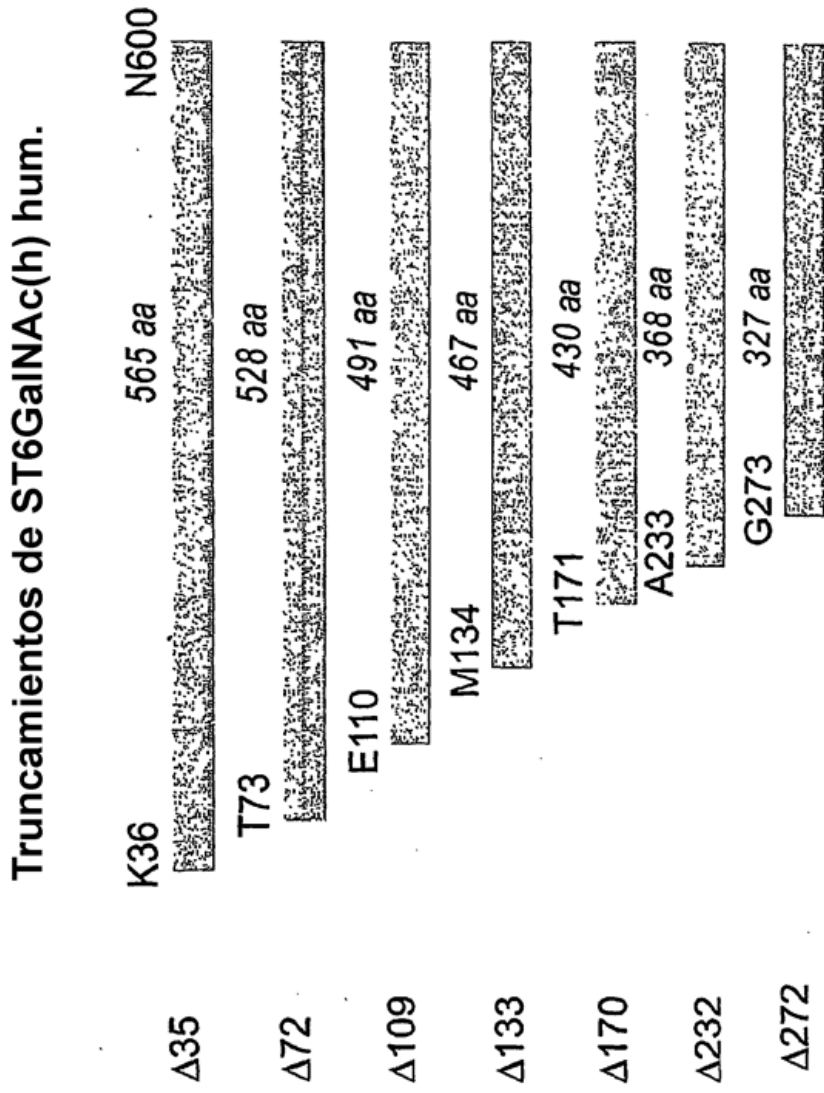


FIG. 39

Constructos de MBP-hST6GalNAcI

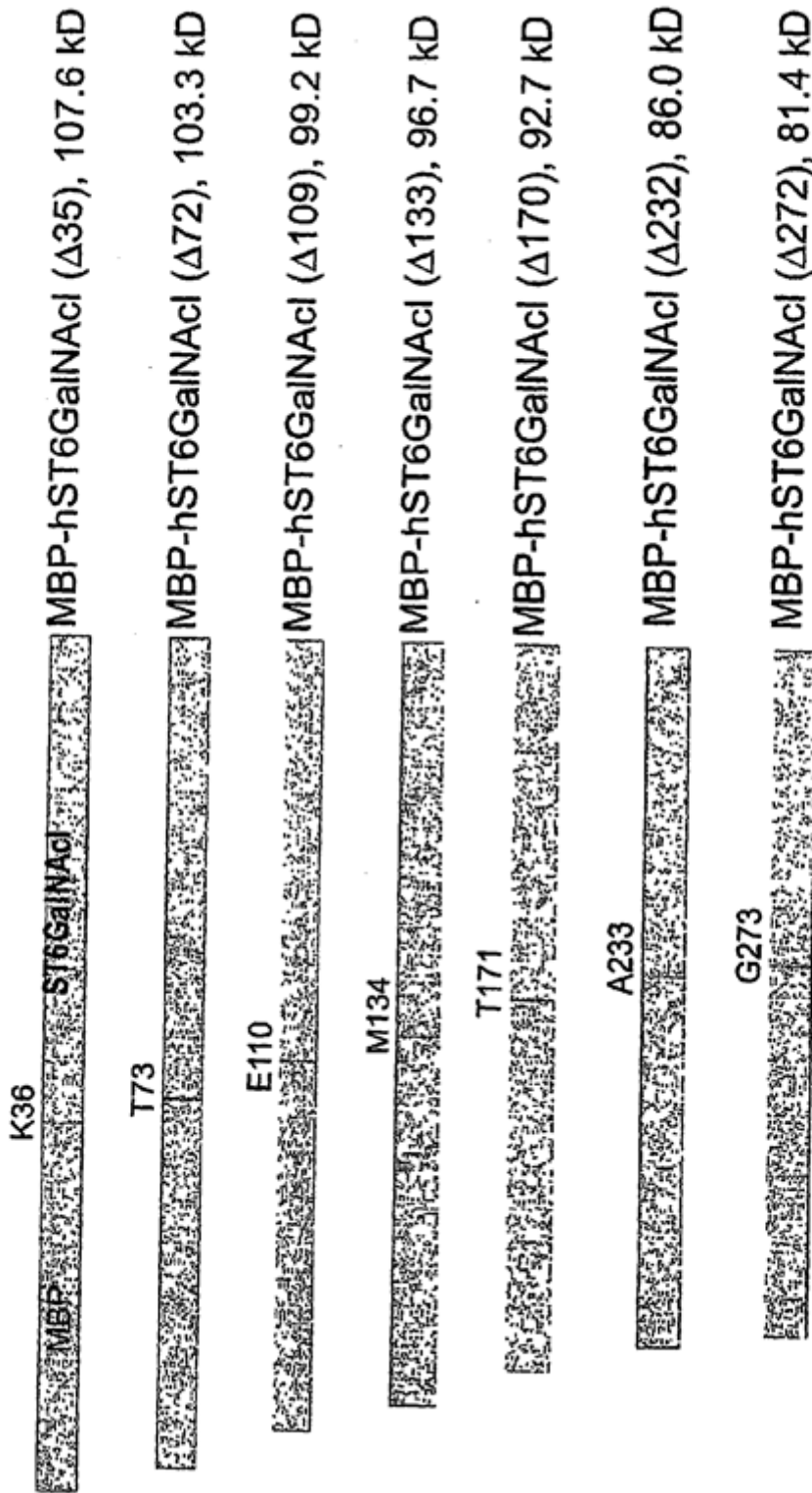


FIG. 40

MASKSWLNFLTFLCGSAIGFLLCSQLFSILLGEKVDTQPNVLHNDPHARHSDDNGQN
HLEGQMNFNADSSQHKDENTDIAENLYQKVRILCWVMTGPQNLEKKAKHVKATW
AQRCKNVLFMSSEENKDFPAVGLKTKEGRDQLYWKTIKAFQYVHEHYLEDADWFL
KADDDTYVILDNLRWLLSKYDPEEPIYFGRRFKPYVKQGYMSGGAGYVLSKEALKR
FVDAFKTDKCTHSSIEDLALGRCMEIMNVEAGDSRDTIGKETFHFPVPEHHLIKGYL
PRTFWYWNYNYPVVEGPGCCSDLAVSFHYVDSTTMYELEYLVYHLRPYGYLYRY
QPTLPERILKEISQANKNEDTKVCLGNP

FIG. 41

Secuencia	Tamaño	Intervalo	Modo
SC1GALT1.AMI	342	1-	342 Normal
SPTS122.AMI	342	1-	342 Normal
		10	20
SC1GALT1.AMI	1	E F M P Y D G H R H	D M M E M S G P E Q
SPTS122.AMI	1	E F M P Y D G H R H	D M M E M S G P E Q
		60	70
SC1GALT1.AMI	51	E V R V L C W I M T	H V K R T W G K R C
SPTS122.AMI	51	E V R V L C W I M T	H V K R T W G K R C
		110	120
SC1GALT1.AMI	101	V G E G R N N L W G	K T K E A Y K Y I Y
SPTS122.AMI	101	V G E G R N N L W G	K T K E A Y K Y I Y
		160	170
SC1GALT1.AMI	151	Y S P E T P V Y F G	Y M S G G A G Y V L
SPTS122.AMI	151	Y S P E T P V Y F G	Y M S G G A G Y V L
		210	220
SC1GALT1.AMI	201	S D N S G A E D V E	I G K C L Q N V N V
SPTS122.AMI	201	S D N S G A E D V E	I G K C L Q N V N V
		260	270
SC1GALT1.AMI	251	K F W Y W Q Y I F Y	K T D E G L D C C S
SPTS122.AMI	251	K F W Y W Q Y I F Y	K T D E G L D C C S
		310	320
SC1GALT1.AMI	301	N T P D A L P N K L	A V G E I M P E I K
SPTS122.AMI	301	N T P D A L P N K L	A V G E I M P E I K
		342	342
SC1GALT1.AMI	50	D V G G H E H V H E	N S T I A E R L Y S
SPTS122.AMI	50	D V G G H E H V H E	N S T I A E R L Y S
		90	100
SC1GALT1.AMI	100	N K L I F M S S A K	D D E L D A V A L P
SPTS122.AMI	100	N K L I F M S S A K	D D E L D A V A L P
		140	150
SC1GALT1.AMI	150	L K A D D D T Y T I	V E N M R Y M L Y P
SPTS122.AMI	150	L K A D D D T Y T I	V E N M R Y M L Y P
		190	200
SC1GALT1.AMI	200	S R E A V R R F V V	E A L P N P K L C K
SPTS122.AMI	200	S R E A V R R F V V	E A L P N P K L C K
		240	250
SC1GALT1.AMI	250	R G R F F F V P E	H H L I P S H T D K
SPTS122.AMI	250	R G R F F F V P E	H H L I P S H T D K
		290	300
SC1GALT1.AMI	300	P N Q M Y V L D Y L	I Y H L R P Y G I I
SPTS122.AMI	300	P N Q M Y V L D Y L	I Y H L R P Y G I I
		340	350
SC1GALT1.AMI	350	V S K R S T E T K I	Q *
SPTS122.AMI	350	V S K R S T E T K I	Q *

FIG. 42

MBP de *Yersinia*

MKIEEGKLV	WINGDKGYNG	LAEVGKKFEK	DTGIKVTIEH	PKLEEKFPQ
VAATGDGPD	IFWAHDRFGG	YAQSGLLAE	TPSKAFQEK	FPFTWDVRF
NGKLIYYP	VEALSIIYNK	DLVKEAPKTW	EEIPALDKTL	RANGKSAIMW
NLQEPYFTWP	VIAADGGYAF	KFENGVDYDAK	NVGVNNAGAQ	AGLQFIVDLV
KNKHINADTD	YSIAEAAFNK	GETAMTINGP	WAWSNIDKSK	INYGVTLLPT
FHGQPSKPFV	GVLTAGINAA	SPNKELATEF	LENYLITDQG	LAEVNKDKPL
GAVALKSFQE	QLAKDPRIAA	TMDNATNGEI	MPNIPQMAAF	WYATRSVAVLN
AITGRQTVEA	ALNDAATRIT	K		

FIG. 43AMBP de *E. coli*

MKIEEGKLV	WINGDKGYNG	LAEVGKKFEK	DTGIKVTVEH	PKLEEKFPQ
VAATGDGPD	IFWAHDRFGG	YAQSGLLAEI	TPDKAFQDKL	YPFTWDVAVRY
NGKLIAYPIA	VEALSIIYNK	DLLPNPPKTW	EEIPALDKEL	KAKGKSALMF
NLQEPYFTWP	LIAADGGYAF	KYENGYDKIK	DVGVDNAGAK	AGLTFVLVDLI
KNKHMNADTD	YSIAEAAFNK	GETAMTINGP	WAWSNIDTSK	VNYGVTVLPT
FKGQPSKPFV	GVLISAGINAA	SPNKELAKEF	LENYLLTDEG	LEAVNKDKPL
GAVALKSYEE	ELAKDPRIAA	TMENAQKGEI	MPNIPQMSAF	WYAVRTAVIN
AASGRQTVDE	ALKDAQTNS			

FIG. 43BMBP de *Pyrococcus furiosus*

MKIEEGKVV	WHAMQPNELE	VFQSLAE EYM	ALSPEVEIVF	EQKPNLEDAL
KAAIPTGQGP	DLFIWAHDWI	GKFAEAGLLE	PIDEYVTEDL	LNEFAPMAQD
AMQYKGHYYA	LPFAAETVAI	IYNKEMVSEP	PKTFDEMKA	MEKYYDPANE
KYGIAPWIPNA	YFISAIQAQAF	GGYYFDDKTE	QPGLDKPETI	EGFKFFFTEI
WPYMAPTGDY	NTQQSIFLEG	RAPMMVNGPW	SINDVKKAGI	NFGVVPLPPI
IKDGKEYWPR	PYGGVKLIYE	AAGIKNKDAA	WKFAKWLTT	EESIKTALALE
LGYIPVLTKV	LDDPEIKNDP	VIYGFQAVQ	HAYLMPKSPK	MSAVWGGVDG
AINEILQDPQ	NADIEGILKK	YQOEILNNMQ	G	

FIG. 43CMBP de *Thermococcus litoralis*

MKIEEGKIVF	AVGGAPNEIE	YWKGVIAEFE	KKYPGVTVEL	KRQATDTEQR
RDLVNLARG	KSSDPDVFLM	DVAWLQGFIA	SGWLEPLDDY	VQKDNVDLSV
FFQSVINLAD	KQGGKLYALP	VYIDAGLLYY	RKDLLEKYGY	SKPPETWQEL
VEMAQKIQSG	ERETNPFWG	FVWQGGQYEG	LVCDFVEYVY	SNGGSLGEFK
DGKWVPTLNK	PENVEALQFM	VDLIHKYKIS	PPNTYTEMTE	EPVRLMFQQG
NAAFERNWPY	AWGLHNADDS	PVKGKVGVP	LPHFPGHKS	ATLGGWHIGI
SKYSDNKALA	WEFVKFVESY	SVQKGFAMNL	GWNPGRVDVY	DDPAVVSISP
HLKELRAVFE	NAVPRPIVPY	YPQLSEIIQK	YVNSALAGKI	SPQEALDKAQ
KEAEELVKQY	SK			

FIG. 43DMBP de *Thermatoga maritime*

MKIEQTKLTI	WSSEKQVDIL	QKLGEEFKAK	YGIPVEVQYV	DFGSIKSKFL
TAAPQGGAD	IIVGAHDWVG	ELAVNGLIEP	IPNFSDLKNF	YDTALKAFSY
GKLYGVPYA	MEAVALIYNK	DYVDSVPKTM	DELIEKAKQI	DEEYGGVEVRG
FIYDVANFYF	SAPFILGYGG	YVFKETPQGL	DVTDIGLANE	GAVKGAKLIK
RMIDEGVLTP	GDNYGTMSM	FKEGLAAMI	NGLWAIKSYK	DAGINYGVAP
IPELEPGVPA	KPFVGVQGM	INAKSPNKVI	AMEFLTNFIA	RKETMYKIYL
ADPRLPARKD	VLELVKDNPD	VVAFTQSASM	GTPMPNVP	APVWSAMGDA
LSIIINGQAS	VEDALKEAVD	KIKAQIEK		

FIG. 43E

MBP de *Vibrio cholerae*

MKIEEGQLTI WINGDKGYNG LAEVGKKFEA DTGIKVTVAH PDALQDKFPQ
TAATGDGPDV VFWAHDREFG YAEAGLLVEI KPSAKIQEGI VDFAWDAVKY
NGKIIGYPIA VESLSLIYNK DLVPNPPKSW EEVAELDAKL KKEGKSAIMW
NLKEPYFTWP LMAADGGYAF KYGVDGYDVK DAGINNKGVK DAMNFVKGLV
DKGVISPDMD YSVSESAFNQ GNTAMTINGP WSWGNIKESG INYGVTTLPK
FNGQASKPFV GVLTAGISTA SPNKDLAVEF IENYLLTNDG LRMVNNDKPL
GAVALNSFQR ELDADARIAA TMDNAMNGEI MPNIPQMNAF WSSAKNAIIN
IVDGRQTVDA ALADAEKQMT KP

FIG. 43F

	1	50
HSGALNAT1 .pep	(1)MRKFAYCKVVLATSLIWLWLLDMFLLLYFS-----ECNKC	
HSGALNAT2 .pep	(1)MRRRS--RMLLCFAFLWVLG IAYYMYSGGGSALAGGAGGGAGRKEDWNEI	
Consenso	(1)MRK A KMLL A IWVL F L D N	
	51	100
HSGALNAT1 .pep	(35)DEKKERGLPAGDVLLEPVQKPHEGP-G-----EMGKPVVTPKEDQEKMKEM	
HSGALNAT2 .pep	(49)DPIKKKDLHHSNGEERKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSQQDP	
Consenso	(51)D K K L E Q P G D I D	
	101	150
HSGALNAT1 .pep	(79)FKINQFNLMASEMIALNRS LDPVRLGCKTKVYPDNLPTTSVVTVFHNEA	
HSGALNAT2 .pep	(99)YARNKFNQVESDKLRMDRALPDTRHDQQQRKQWRVDLPATSVVITFHNEA	
Consenso	(101)F N FN M SD I L RAIPD R D C K W LP TSVVI FHNEA	
	151	200
HSGALNAT1 .pep	(129)WSTLLRTVHSVINRSPRHMIETVLVDDASERDFLKRPLESYVKKLKV PV	
HSGALNAT2 .pep	(149)RSALLRTVWVSLKSPPHLIKEIILVDDYSN-----DPEDGALLGKIEKV	
Consenso	(151) S LLRTV SVI KSP HLI E I L V D D S P D L V	
	201	250
HSGALNAT1 .pep	(179)HVIRMEQRSLIRARLKGA AVSKGQVITFLDAHCECIVGWLEPLLARIKH	
HSGALNAT2 .pep	(194)RVLNRDRREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERRVAE	
Consenso	(201) VIR D R GLIRARLKGA A A VITFLDAHCEC WLEPLL RI	
	251	300
HSGALNAT1 .pep	(229)DRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKLNFRWYVVPQREMDRR	
HSGALNAT2 .pep	(244)DRTRVVSPIIDVINMDFQYVGASADLKGGFDWNLVFKWDYMTPEQRRSR	
Consenso	(251) DR VV PIIDVI D F YMAAS GGF W L FKW M R	
	301	350
HSGALNAT1 .pep	(279)KGDRTLPVVRTPTMAGGLFSIDRDYFQEI GTYDAGMDI WGGENLEISFRIW	
HSGALNAT2 .pep	(294)QGNPVAPIKTPMLIAGGLFVMDKFYFEELGKYDMMMDVWGGENLEISFRW	
Consenso	(301) G PIKTP IAGGLF IDK YF EIG YD MDI WGGENLEISFRIW	
	351	400
HSGALNAT1 .pep	(329)QCGGTLEIVTCSHVGHVFRKATPYTFFGGTGQI INKNNRRLAEVWDEFK	
HSGALNAT2 .pep	(344)QCGGSLEIIPCSRUGHVFRKQHPYTFFGGSGIVFARNTRRAAEVWDEYK	
Consenso	(351) QCGGSLEI I CS VGHVFRK PYTFFGGSG I KN RR AEVWDEFK	
	401	450
HSGALNAT1 .pep	(379)NFFYIISPQVTKVDYGDISSRVGLRHKLQCKPFSWYLENTYPDSQIPRHY	
HSGALNAT2 .pep	(394)NFFYA AVPSARNVPYGNIQSRLELRKKL SCKPFWYLENVYPELRVPDHQ	
Consenso	(401) NFFY P V YG I SRL LR KL CKPF WYLENTYPD IP H	
	451	500
HSGALNAT1 .pep	(429)FSLGEIRNVEINQCLDNMARKENEKVGI FNCHGMGGNQVFSYTANKEIRT	
HSGALNAT2 .pep	(444)DIAFGALQQGIN-CLDTLGHFADGVVGVYECHNAGGNQEWALITKEKSVKH	
Consenso	(451) N TN CLD LA VGIF CH GGNQ FA T K IK	
	501	550
HSGALNAT1 .pep	(479)EDLCLDVSKLN--GPVIMLKCHHLKGNQLWEYDVKLTLQHVNSNQCLDK	
HSGALNAT2 .pep	(493)MDLCLTVVDRAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCLDS	
Consenso	(501) DLCL V I L C Q WE L HV SN CLD	
	551	584
HSGALNAT1 .pep	(527)ATEEDSQVPSIRDONGSRSQWLLRNVTLPETIF-	
HSGALNAT2 .pep	(543)RTAK-SGGLSVEVCGPALSQQWKFTLNLQQ----	
Consenso	(551) T S SI C A SQW	

FIG. 44

INGENIERÍA GENÉTICA DE CISTEÍNAS EN ST6GalNAcI hum.

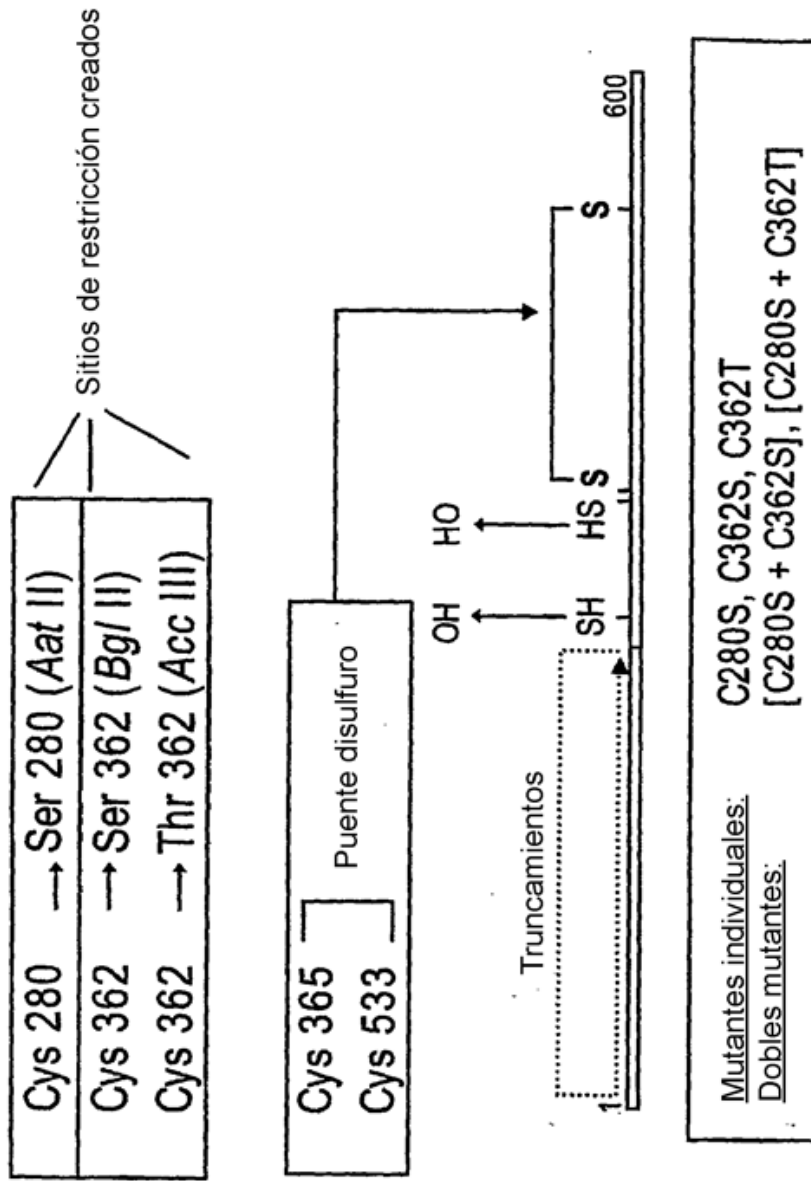


FIG. 45