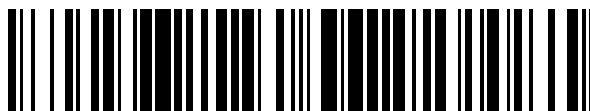


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 374 935**

51 Int. Cl.:
C12P 21/02 (2006.01)
C12N 9/10 (2006.01)
C12N 15/70 (2006.01)
C12N 15/78 (2006.01)
C12N 1/21 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06748724 .9**
96 Fecha de presentación: **24.03.2006**
97 Número de publicación de la solicitud: **1863921**
97 Fecha de publicación de la solicitud: **12.12.2007**

54 Título: **EXPRESIÓN DE GLICOSILTRANSFERASAS EUCARIOTAS ACTIVAS, SOLUBLES EN ORGANISMOS PROCARIOTAS.**

30 Prioridad:
24.03.2005 US 665396 P
05.04.2005 US 668899 P
31.10.2005 US 732409 P

45 Fecha de publicación de la mención BOPI:
23.02.2012

45 Fecha de la publicación del folleto de la patente:
23.02.2012

73 Titular/es:
BioGeneriX AG
Janderstrasse 3
68199 Mannheim, DE

72 Inventor/es:
SCHWARTZ, Marc, F. y
SOLIMAN, Tarik

74 Agente: **Arias Sanz, Juan**

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).

ES 2 374 935 T3

DESCRIPCIÓN

Expresión de glicosiltransferasas eucariotas activas, solubles en organismos procariotas.

Campo de la invención

La presente invención proporciona métodos mejorados para la producción de glicosiltransferasas eucariotas activas, solubles en microorganismos procariotas que tienen un entorno oxidante.

Antecedentes de la invención

Los organismos eucariotas sintetizan estructuras de oligosacáridos 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 muchas proteínas recombinantes es expresar la proteína en bacterias. Sin embargo, en bacterias, muchas glicosiltransferasas eucariotas se expresan como proteínas insolubles en cuerpos de inclusión bacterianos, y los rendimientos de proteína glicosiltransferasa eucariota activa a partir de los cuerpos de inclusión pueden ser muy bajos. Además, muchas glicosiltransferasas eucariotas se expresan como proteínas glicosiladas en sus células de origen. Por tanto, se cree que la expresión de las proteínas en bacterias no incluiría patrones de glicosilación nativos, disminuyendo además la esperanza de expresión de proteína glicosiltransferasa eucariota activa. Véase, por ejemplo, Breton *et al.*, Biochimie 83:713-718 (2001). Por tanto, existe una necesidad de métodos mejorados para producir glicosiltransferasas eucariotas enzimáticamente activas en organismos procariotas, tales como, por ejemplo, bacterias. La presente invención soluciona esta y otras necesidades.

Breve resumen de la invención

En un aspecto, la presente invención proporciona un método de producción de una glicosiltransferasa eucariota activa soluble según la reivindicación 1.

En una realización, la glicosiltransferasa eucariota es un miembro seleccionado de una N-acetilglucosaminiltransferasa I (GnT o GNT) eucariota, una N-acetilgalactosaminiltransferasa (GalNAcT) eucariota, una galactosiltransferasa (GalT) eucariota y una sialiltransferasa eucariota. Los ejemplos de cada una de estas clases de enzimas incluyen, por ejemplo, para una N-acetilglucosaminiltransferasa eucariota proteínas GnT1, BGnT-I, GnT-II, GnT-III, GnT-IV (por ejemplo, GnT-IVa y GnT-IVb), GnT-V, GnT-VI, GnT-IVH, MGNT1 y OGT; para una N-acetilgalactosaminiltransferasa eucariota proteínas GalNAc-T2, GalNAc-T1 y GalNAc-T3; para una galactosiltransferasa (GalT) eucariota una β -1,4-galactosiltransferasa (GalT1) eucariota o una core 1 galactosiltransferasa (core 1 GalT1) eucariota; para una sialiltransferasa eucariota, una α (2,3)sialiltransferasa (ST3Gal3) eucariota, una α -N-acetilgalactosaminida α -2,6-sialiltransferasa I (ST6GalNAcT1) eucariota o una gal β 1,3GalNAc α 2,3-sialiltransferasa (ST3Gal1) eucariota.

Algunos ejemplos preferidos de glicosiltransferasas eucariotas para su uso en la invención incluyen una N-acetilglucosaminiltransferasa I (GnT1 o GNT1) eucariota, una N-acetilgalactosaminiltransferasa (GalNAcT) eucariota, por ejemplo, una GalNAcT1, GalNAcT2 o GalNAcT3, una β -1,4-galactosiltransferasa (GalT1) eucariota, una α -2,3-sialiltransferasa (ST3Gal3) eucariota, una α -N-acetilgalactosaminida α -2,6-sialiltransferasa I (ST6GalNAc-1) eucariota, una gal β 1,3GalNAc α 2,3-sialiltransferasa (ST3Gal-1) eucariota y una core 1 galactosiltransferasa (Core-1-GalT-1) eucariota.

En una primera realización el microorganismo procariota es una bacteria de *E. coli* o una de *Pseudomonas* que tiene un entorno oxidante. Por ejemplo, la *E. coli* puede manipularse para inactivar ácidos nucleicos de reductasa genómicos o endógenos, por ejemplo, un gen *txrB* y un gen *gor*. Otras cepas de *E. coli* que pueden usarse incluyen, por ejemplo, una cepa mutante *txrB gor supp* o una cepa mutante *txrB gshA supp*, ambas de las cuales se dan a conocer en la patente estadounidense n.º 6.872.563. En una realización adicional, el microorganismo procariota, por ejemplo, células de *E. coli* o células de *Pseudomonas*, se hacen crecer a una temperatura entre 12-30°C mientras se expresa la glicosiltransferasa eucariota. El microorganismo procariota también puede expresar proteínas adicionales para potenciar la solubilidad de la glicosiltransferasa eucariota, por ejemplo una proteína disulfuro isomerasa (PDI) heteróloga o una proteína chaperona heteróloga, o tanto una PDI heteróloga como una proteína chaperona heteróloga.

En otra realización, el método comprende además la etapa de aislar la glicosiltransferasa eucariota. En realizaciones adicionales, la glicosiltransferasa eucariota comprende una etiqueta de purificación, por ejemplo, un dominio de proteína de unión a maltosa o un dominio de proteína de unión a almidón. En realizaciones adicionales, la glicosiltransferasa eucariota soluble se produce a escala comercial. La escala comercial incluye la preparación de una cantidad suficiente de enzima para producir un producto glicosilado a una escala comercial (microgramo, miligramo o gramo).

Breve descripción de los dibujos

La figura 1 demuestra el análisis de SDS-PAGE de la expresión y purificación parcial de ST3Gal-1 porcina truncada

etiquetada con MBP soluble. Se expresó MBP-ST3Gal-1 en las cepas mutantes JM109 (carriles 2-5) y *trxB* gor supp (carriles 6-9). Los carriles 2-3 y 6-7 son lisados clarificados antes y tras la incubación con resina de amilosa, respectivamente. Los carriles 4-5 y 8-9 son eluciones en serie de la resina de amilosa que contienen MBP-ST3Gal-1 parcialmente purificada. En el primer carril están los marcadores de peso molecular.

5 La figura 2a demuestra el análisis de SDS-PAGE de la expresión y purificación parcial de GalNAc-T2 humana truncada etiquetada con MBP soluble a partir de lisados mutantes de *trxB* gor supp (carriles 2-4). Los carriles 2-3 son lisados clarificados antes y tras la incubación con resina de amilosa, respectivamente. El carril 4 es la elución de la resina de amilosa que contiene MBP-GalNAc-T2 parcialmente purificada. En el primer carril están los marcadores de peso molecular. La figura 2b demuestra el análisis de SDS-PAGE de la expresión y solubilidad de GalNAc-T2 humana truncada en los extremos amino y carboxilo terminales expresada en *E. coli* mutante *trxB* gor supp. Se separaron las células lisadas mediante centrifugación en fracciones insolubles (carril 2) y solubles (carril 3) y se resolvieron mediante SDS-PAGE. En el primer carril están los marcadores de peso molecular.

10 La figura 3 demuestra el análisis de SDS-PAGE de la expresión y purificación parcial de Core-1-Gal-T1 de *Drosophila* truncada etiquetada con MBP soluble. Se expresó MBP-Core-1-Gal-T1 y se purificó a partir de células mutantes JM109 (carriles 2-3) y *trxB* gor supp (carriles 4-5). El carril 1 contiene marcadores de peso molecular. Los carriles 2 y 4 son lisados clarificados y se muestran las eluciones de resina de amilosa parcialmente purificada en los carriles 3 y 5.

15 La figura 4 demuestra el análisis de SDS-PAGE de la expresión y purificación parcial de ST6GalNAc-1 humana truncada etiquetada con MBP soluble a partir de células mutantes *trxB* gor supp. El carril 1 contiene marcadores de peso molecular. Los carriles 2 y 3 son lisados clarificados antes y tras la incubación con resina de amilosa, respectivamente. El carril 4 contiene MBP-ST6GalNAc-1 parcialmente purificada eluida de la resina de amilosa.

20 La figura 5 proporciona los rendimientos de actividad expresada a partir de experimentos de producción de aumento a escala para MBP-GalNAc-T2, MBP-Core-1-Gal-T1 y MBP-ST3Gal-1. Se indujeron recipientes de fermentación de diez litros sembrados con células mutantes *trxB* gor supp que expresaban el constructo indicado durante 48 horas. Se monitorizó la actividad glicosiltransferasa (U/litro de medio) a partir de alícuotas tomadas en los tiempos indicados tras la inducción.

25 La figura 6 demuestra la glicosilación de interferón-alfa-2b usando glicosiltransferasas eucariotas producidas en *E. coli* mutante *trxB* gor supp. Se analizaron los productos de reacción mediante espectrometría de masas MALDI TOF. La figura 6A muestra interferón-alfa-2b no modificado. La figura 6B muestra el resultado de la incubación con MBP-GalNAc-T2 y UDP-GalNAc. La figura 6C muestra el resultado de la incubación con MBP-GalNAc-T2, MBP-Core-1-Gal-T1, UDP-GalNAc y UDP-Gal. Se observó la masa esperada debido a la adición de GalNAc (esperada +203,2, observada +209,1) o GalNAc-Gal (esperada +365,6, observada +365,3) a interferón-alfa-2b.

30 La figura 7 demuestra el análisis de SDS-PAGE de una reacción de glicoPEGilación usando glicosiltransferasas eucariotas producidas en *E. coli* mutante *trxB* gor supp. En primer lugar, se incubó interferón-alfa-2b (carril 2) con MBP-GalNAc-T2, MBP-Core-1-Gal-T1, UDP-GalNAc y UDP-Gal durante seis horas a 32°C (carril 3). Entonces, se añadieron MBP-ST3Gal-1 y CMPSA-40kDa-PEG purificadas (carril 4, punto de tiempo de hora cero), y se incubaron durante la noche a 32°C para permitir que se produjera la reacción de glicoPEGilación (carril 5). Los marcadores de peso molecular están en el primer carril.

35 La figura 8 proporciona el análisis de SDS-PAGE de la solubilidad y purificación parcial de GnT1 humana truncada etiquetada con MBP (carriles 2-4, 8-10) y MBP-GnT1 C121S (carriles 5-7, 11-13) expresadas en cepas de *E. coli* mutantes JM109 (carriles 2-7) y *trxB* gor supp (carriles 8-13). Se lisaron las células a partir de cultivos inducidos y se separó el material insoluble mediante centrifugación (carriles 2, 5, 8, 11). Se incubaron los sobrenadantes (carriles 3, 6, 9, 12) con resina de amilosa. Entonces se eluyeron las proteínas de fusión de MBP parcialmente purificadas de la resina (carriles 4, 7, 10, 13). En el primer carril están los marcadores de peso molecular.

40 La figura 9 proporciona el análisis de SDS-PAGE de la solubilidad y purificación parcial de GalT1 bovina truncada etiquetada con MBP (carriles 2-4, 8-10) y MBP-GalT1 C342T (carriles 5-7, 11-13) expresadas en cepas de *E. coli* mutantes JM109 (carriles 2-7) y *trxB* gor supp (carriles 8-13). Se lisaron los cultivos inducidos y se separó el material insoluble mediante centrifugación (carriles 2, 5, 8, 11). Se incubaron los sobrenadantes (carriles 3, 6, 9, 12) con resina de amilosa. Entonces, se eluyeron las proteínas de fusión de MBP parcialmente purificadas de la resina (carriles 4, 7, 10, 13). En el primer carril están los marcadores de peso molecular.

45 La figura 10 proporciona el análisis de SDS-PAGE de la expresión soluble y purificación parcial de ST3Gal3 de rata truncada etiquetada con MBP (carriles 2-4) y ST3Gal3 etiquetada con SBD-MBP (carriles 5-7) expresadas en una cepa de *E. coli* mutante *trxB* gor supp. Los carriles 2-3 y 5-6 son lisados clarificados antes y tras la incubación con resina de amilosa, respectivamente. Los carriles 4 y 7 son las eluciones de la resina de amilosa que contienen proteínas de fusión de ST3Gal3 parcialmente purificadas. En el primer carril están los marcadores de peso molecular.

Definiciones

Las proteínas glicosiltransferasas recombinantes producidas mediante los métodos 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 oligosacárido o resto de 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).

Las siguientes abreviaturas se usan en el presente documento:

Ara = arabinosilo;

Fru = fructosilo;

Fuc = fucosilo;

Gal = galactosilo;

GalNAc = N-acetilgalactosilamino;

Glc = glucosilo;

GlcNAc = N-acetilglucosilamino;

Man = manosilo; y

NeuAc = sialilo (N-acetilneuraminilo)

FT o FucT = fucosiltransferasa*

ST = sialiltransferasa*

GalT = galactosiltransferasa*

Se usan número arábigos o romanos de manera intercambiable en el presente documento según la convención de nomenclatura usada 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, si el sacárido en el extremo reductor es o no 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.

El término "ácido siálico" se refiere a cualquier miembro de una familia de azúcares carboxilados de nueve carbonos. El miembro lo más común de la familia del ácido siálico 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-glicolilneuramínico (Neu5Gc o NeuGc), en el que el grupo N-acetil de NeuAc está hidroxilado. Un tercer miembro de la familia del ácido siálico es el ácido 2-ceto-3-desoxi-nonulosónico (KDN) (Nadano *et al.* (1986) J. Biol. Chem. 261: 11550-11557; Kanamori *et al.*, J. Biol. Chem. 265: 21811-21819 (1990)). También se incluyen ácidos siálicos 9-sustituidos tales como a 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 del 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 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 el sustrato donador de azúcar y la glicosiltransferasa correspondiente, y otros componentes de la mezcla de reacción necesarios, y la mezcla de reacción se incuba durante un periodo de tiempo suficiente, la glicosiltransferasa transfiere residuos de azúcar del sustrato donador de azúcar al sustrato aceptor. El sustrato aceptor a menudo variará para diferentes tipos de una glicosiltransferasa particular. Por ejemplo, el sustrato aceptor para una galactosida 2-L-fucosiltransferasa (α 1,2-fucosiltransferasa) de mamífero incluirá una Gal β 1,4-GlcNAc-R en un extremo terminal no reductor de un oligosacárido; esta fucosiltransferasa une un residuo de fucosa a la Gal mediante un enlace α 1,2. Gal β 1,4-GlcNAc-R y Gal β 1,3-GlcNAc-R terminales y 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, el término "sustrato aceptor" se toma en contexto con la glicosiltransferasa particular de interés para una aplicación particular. Se describen en el presente documento sustratos aceptores para glicosiltransferasas adicionales. Los sustratos aceptores también incluyen por ejemplo, glicolípidos, péptidos,

proteínas, glicopéptidos, glicoproteínas y proteínas terapéuticas.

Un “sustrato donador” para glicosiltransferasas es un azúcar de nucleótido activado. Tales azúcares activados consisten generalmente en derivados de uridina, guanosina y citidina monofosfato (UMP, GMP y CMP, respectivamente) de los azúcares o derivados de difosfato (UDP, GDP y CDP, respectivamente) de los azúcares en los que el nucleósido monofosfato o difosfato sirve como grupo saliente. Por ejemplo, un sustrato donador para fucosiltransferasas es GDP-fucosa. Los sustratos donadores para sialiltransferasas, por ejemplo, son nucleótidos de azúcar activado 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 denominada CMP-ácido siálico-PEG), UDP-N-acetilglucosamina, UDP-glucosa, UDP-ácido glucurónico y UDP-xilosa. Los azúcares incluyen, por ejemplo, NeuAc, manosa, galactosa, N-acetilgalactosamina, N-acetilglucosamina, glucosa, ácido glucurónico 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 covalentemente unido 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, de 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, es decir, un azúcar de nucleótido, a un péptido, una proteína, un glicopéptido o una glicoproteína. En otra realización preferida, una glicosiltransferasa eucariota es una glicosiltransferasa transmembrana tipo II. Las glicosiltransferasas transmembrana tipo II no modificadas normalmente incluyen un dominio citoplasmático aminoterminal, un dominio transmembrana o de señal-anclaje, una región de tallo y un dominio catalítico. Véase, por ejemplo, Paulson y Colley, J. Biol. Chem. 264:17615-17618 (1989). Una glicosiltransferasa eucariota puede derivarse de un organismo eucariota, por ejemplo, un organismo eucariota unicelular o 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 y etc. Siguen ejemplos de glicosiltransferasas eucariotas y también se encuentran en la lista de secuencias adjunta.

Una “sialiltransferasa eucariota” tal como se usa en el presente documento se refiere a una sialiltransferasa derivada de un organismo eucariota. La enzima cataliza la transferencia de un resto ácido siálico de un donador de CMP-ácido siálico a una molécula aceptora. Las sialiltransferasas eucariotas también pueden reconocerse por la presencia de motivos estructurales conservados, por ejemplo, un motivo L de sialilo y un motivo S de sialilo tal como se describe en Tsuji, J. Biochem. 120:1-13 (1996).

Motivos de sialiltransferasa adicionales, por ejemplo, el motivo muy pequeño (VS) y el motivo III, se describen en Patel y Balaji, Glycobiology, 16:108-116 (2006), publicado electrónicamente el 5 de octubre de 2005. Las sialiltransferasas eucariotas incluyen enzimas que forman una variedad de enlaces incluyendo $\alpha 2 \rightarrow 3$, $\alpha 2 \rightarrow 6$, $\alpha 2 \rightarrow 8$. Las sialiltransferasas eucariotas transfieren el resto ácido siálico a diferentes azúcares aceptores en una molécula aceptora, por ejemplo, galactosa, GalNAc y otra molécula de ácido siálico. Las sialiltransferasas eucariotas que catalizan la reacción específica, es decir, que son miembros de las familias de ST3Gal, ST6Gal, ST6GalNAc o ST8Sia pueden identificarse por la presencia de residuos de aminoácido conservados dentro de esas familias. Tales residuos de aminoácido conservados basados en la familia se dan a conocer en Patel y Balaji, Glycobiology, 16:108-116 (2006), publicado electrónicamente el 5 de octubre de 2005, que se incorpora en el presente documento como referencia para todos los fines. Siguen ejemplos de sialiltransferasas eucariotas y también se encuentran en la lista de secuencias adjunta.

Una “ $\alpha(2,3)$ 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 de Gal β 1,3GlcNAc, Gal β 1,3GalNAc o Gal β 1,4GlcNAc (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 se une a una Gal con la formación de un enlace α entre los dos sacáridos. La unión (enlace) entre los sacáridos es entre la posición 2 de NeuAc y la posición 3 de Gal. Como otras glicosiltransferasas eucariotas, las enzimas ST3Gal3 tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Esta enzima particular puede aislarse del hígado de la rata (Weinstein *et al.* (1982) J. Biol. Chem. 257: 13845); se conocen las secuencias de ADNc (Sasaki *et al.* (1993) J. Biol. Chem. 268: 22782-22787; Kitagawa y Paulson (1994) J. Biol. Chem. 269: 1394-1401) y ADN genómico humanas (Kitagawa *et al.* (1996) J. Biol. Chem. 271: 931-938), facilitando la producción de esta enzima mediante expresión recombinante. Se ha clonado ST3Gal3 de rata y la secuencia se conoce. Véase, por ejemplo, Wen *et al.*, J. Biol. Chem. 267:21011-21019 (1992) y número de registro M97754. Se dan a conocer proteínas ST3Gal3 a modo de ejemplo en, por ejemplo, SEQ ID NOS:22-32.

Una “ α -N-acetilgalactosaminida α -2,6-sialiltransferasa I (ST6GalNAcT1 o ST6GalNAc-1) eucariota” tal como se usa

en el presente documento se refiere a una $\alpha(2,6)$ sialiltransferasa aislada de un organismo eucariota. La enzima cataliza la transferencia de ácido siálico de un donador de CMP-ácido siálico a una molécula aceptora. La transferencia es un enlace $\alpha 2,6$ 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. Se dan a conocer proteínas ST6GalNAc-1 a modo de ejemplo en, por ejemplo, SEQ ID NOS:62-77.

Una "Gal $\beta 1,3$ GalNAc $\alpha 2,3$ -sialiltransferasa (ST3 Gall o ST3Gal-1) eucariota" tal como se usa en el presente documento se refiere a una Gal $\beta 1,3$ GalNAc $\alpha 2,3$ -sialiltransferasa aislada de un organismo eucariota. La enzima cataliza la transferencia de ácido siálico de un donador de CMP-ácido siálico a una molécula aceptora. La transferencia es un enlace $\alpha 2,3$ 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, Gillespie *et al.*, J. Biol. Chem. 267:21004-21010 (1992) y número de registro A45073. Se dan a conocer proteínas ST3Gal-1 a modo de ejemplo en, por ejemplo, SEQ ID NOS:53-61.

Otras sialiltransferasas que pueden usarse en la presente invención incluyen, por ejemplo, proteínas "beta galactósido alfa 2,6-sialiltransferasa (ST6Gal I) eucariotas". Se dan a conocer proteínas ST6Gal 1 a modo de ejemplo en, por ejemplo, SEQ ID NOS:78-82.

Las proteínas sialiltransferasa eucariotas usadas en la invención también incluyen proteínas alfa 2,8 sialiltransferasa, por ejemplo, ST8Sia I, ST8Sia II, ST8Sia III y ST8Sia IV. Se encuentran ejemplos de estas proteínas sialiltransferasa, por ejemplo, en SEQ ID NOS:83-97.

Una "N-acetilglucosaminiltransferasa eucariota" tal como se usa en el presente documento se refiere a una N-acetilglucosaminiltransferasa derivada de un organismo eucariota. La enzima cataliza la transferencia de N-acetilglucosamina (GlcNAc) de un donador de UDPGlcNAc a una molécula aceptora. Como otras glicosiltransferasas eucariotas, N-acetilglucosaminiltransferasa tiene un dominio transmembrana, una región de tallo y un dominio catalítico. Siguen ejemplos de N-acetilglucosaminiltransferasas eucariotas y también se encuentran en la lista de secuencias adjunta.

Una " β -1,2-N-N-acetilglucosaminiltransferasa eucariota I (GnTI o GNTI)" tal como se usa en el presente documento se refiere a β -1,2-N-acetilglucosaminiltransferasa I derivada de un organismo eucariota. Como otras glicosiltransferasas eucariotas, GnTI tiene un dominio transmembrana, una región de tallo y un dominio catalítico. Las proteínas GnT1 eucariotas incluyen, por ejemplo, humana, número de registro NP_002397; de hámster chino, número de registro AAK61868; de conejo, número de registro AAA31493; de rata, número de registro NP_110488; de hámster dorado, número de registro AAD04130; de ratón, número de registro P27808; de pez cebra, número de registro AAH58297; de *Xenopus*, número de registro CAC51119; de *Drosophila*, número de registro NP_525117; de *Anopheles*, número de registro XP_315359; de *C. elegans*, número de registro NP_497719; de *Physcomitrella patens*, número de registro CAD22107; de *Solanum tuberosum*, número de registro CAC80697; de *Nicotiana tabacum*, número de registro CAC80702; de *Oryza sativa*, número de registro CAD30022; de *Nicotiana benthamiana*, número de registro CAC82507; y de *Arabidopsis thaliana*, número de registro NP_195537.

Se dan a conocer proteínas GnT1 a modo de ejemplo en, por ejemplo, SEQ ID NOS:1-11. Otras proteínas de N-acetilglucosaminiltransferasa eucariota que pueden usarse en la presente invención incluyen, por ejemplo, BGnT-1, GnT-II, GnT-III, GnT-IV (por ejemplo, GnT-IVa y GnT-IVb), GnT-V, GnT-VI y GnT-IVH, que se muestran a modo de ejemplo en SEQ ID NOS:140-160.

Otras proteínas de N-acetilglucosaminiltransferasa eucariota pueden producirse usando los métodos de la presente invención e incluyen, por ejemplo, proteína fringe maniac, MGNT1 y proteínas OGT, mostradas a modo de ejemplo en por ejemplo, SEQ ID NOS:171-175.

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 (Gal-NAc) de un donador de 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)).

Siguen ejemplos de N-acetilgalactosaminiltransferasas eucariotas y también se encuentran en las figuras 12-36 y en la lista de secuencias adjunta. Se dan a conocer proteínas de GalNAc-T2, GalNAc-T1 y GalNAc-T3 a modo de ejemplo en, por ejemplo, SEQ ID NOS:33-40 y 192-197, 126-132 y 189-191, y 133-135, respectivamente.

Una "galactosiltransferasa eucariota" tal como se usa en el presente documento se refiere a una

galactosiltransferasa derivada de un organismo eucariota. La enzima cataliza la transferencia de galactosa de un donador de UDP-Gal a una molécula aceptora. Como otras glicosiltransferasas eucariotas, las galactosiltransferasas tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Siguen ejemplos de galactosiltransferasa eucariotas y también se encuentran en la lista de secuencias adjunta.

Una “ β -1,4-galactosiltransferasa (GalT1) eucariota” tal como se usa en el presente documento se refiere a una β -1,4-galactosiltransferasa derivada de un organismo eucariota. La enzima cataliza la transferencia de galactosa de un donador de 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. Se dan a conocer proteínas de GalT1 a modo de ejemplo en, por ejemplo, SEQ ID NOS:12-21. Otras galactosiltransferasas que pueden usarse en la presente invención incluyen, por ejemplo, SEQ ID NOS:136-139.

Una “core I galactosiltransferasa (core 1 GalT1 o core-1-Gal-T1) 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 ser humano. La proteína humana se caracteriza en Ju *et al.*, J. Biol. Chem. 277 (1), 178-186 (2002). Se dan a conocer proteínas de core 1 GalT1 a modo de ejemplo en, por ejemplo, SEQ ID NOS:41-52 y 198-199.

Una “fucosiltransferasa eucariota” tal como se usa en el presente documento se refiere a una fucosiltransferasa derivada de un organismo eucariota. La enzima cataliza la transferencia de fucosa de un donador de UDP-fucosa a una molécula aceptora. Como otras glicosiltransferasas eucariotas, las fucosiltransferasas tienen un dominio transmembrana, una región de tallo y un dominio catalítico. Siguen ejemplos de galactosiltransferasa eucariotas y también se encuentran en la lista de secuencias adjunta. Se dan a conocer proteínas de fucosiltransferasa eucariota a modo de ejemplo que pueden usarse en la presente invención en, por ejemplo, SEQ ID NOS:98-125. También se incluyen en los métodos de la invención proteínas de péptido-O-fucosiltransferasa, mostradas a modo de ejemplo en, por ejemplo, SEQ ID NOS:167-170.

Otras proteínas de glicosiltransferasa eucariota que pueden usarse en la presente invención incluyen, por ejemplo, dolicol-fosfato manosiltransferasa polipéptido 1, o Dpm1, mostrada a modo de ejemplo en SEQ ID NO:162; proteínas de alfa-1,6-manosiltransferasa, alfa-1,3-manosiltransferasa y beta-1,4-manosiltransferasa, mostradas a modo de ejemplo en SEQ ID NOS: 163-166.

Una “proteína terapéutica” tal como se usa en el presente documento se refiere a una proteína, péptido, glicoproteína o glicopéptido que se administra a un sujeto para tratar una enfermedad o disfunción o para mejorar la salud del sujeto. En una realización preferida el sujeto es un ser humano. En una realización preferida adicional, la proteína terapéutica es una proteína humana. En una realización adicional, la proteína terapéutica está glicosilada o modificada de otra forma mediante una o más glicosiltransferasas producidas en un microorganismo que tiene un entorno intracelular oxidante.

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

Una “pareja redox” se refiere a mezclas de reactivos de tiol reducidos y oxidados e incluyen glutatión (GSH/GSSG), cisteína/cistina, cisteamina/cistamina, DTT/GSSG y DTE/GSSG reducidos y oxidados. (Véase, por ejemplo, Clark, Cur. Op. Biotech. 12:202-207 (2001)).

El término “oxidante” o “agente oxidante” se refiere a un compuesto que oxida moléculas en su entorno, es decir, que cambia las moléculas en su entorno para que estén más oxidadas y sean más oxidantes. Un oxidante actúa aceptando electrones, reduciéndose de ese modo a sí mismo tras haber oxidado un sustrato. Por tanto, un oxidante es un agente que acepta electrones.

La expresión “condiciones oxidantes” o “entorno oxidante” se refiere a una condición o un entorno en el que un sustrato es más probable que se oxide que se reduzca. Por ejemplo, el periplasma de una célula de *E. coli* de tipo natural constituye un entorno oxidante, mientras que el citoplasma de una célula de *E. coli* de tipo natural es un entorno reductor.

Una enzima en un “estado oxidado” se refiere a una enzima que tiene menos electrones que su forma reducida.

El término “reductor” o “agente reductor” se refiere a un compuesto que reduce moléculas en su entorno, es decir, que cambia moléculas en su entorno para hacerlas más reducidas y más reductoras. Un agente reductor actúa donando electrones, oxidándose de ese modo a sí mismo tras haber reducido un sustrato. Por tanto, un agente reductor es un agente que dona electrones. Los ejemplos de agentes reductores incluyen ditiotreitól (DTT), mercaptoetanol, cisteína, tioglicolato, cisteamina, glutatión y borohidruro de sodio.

El término “reductasa” se refiere a una tiorredoxina reductasa, glutatión o glutatión reductasa (también denominadas “oxidoreductasas”) o cualquier otra enzima que pueda reducir miembros de los sistemas de tiorredoxina o glutarredoxina.

5 La expresión “rutas de reductasa” se refiere a los sistemas en las células que mantienen el entorno en condiciones reductoras, e incluyen el sistema de glutarredoxina y el sistema de tiorredoxina.

La expresión “condiciones reductoras” o “entorno reductor” se refiere a una condición o un entorno en el que un sustrato es más probable que se reduzca que se oxide. Por ejemplo, el citoplasma de una célula eucariota constituye un entorno reductor.

10 “Formación de enlaces disulfuro” u “oxidación de enlaces disulfuro”, usadas de manera intercambiable en el presente documento, se refiere al proceso de formación de un enlace covalente entre dos cisteínas presentes en un o dos polipéptidos. La oxidación de enlaces disulfuro está mediada por el intercambio tiol-disulfuro entre las cisteínas del sitio activo de enzimas y cisteínas en la proteína diana. La formación de enlaces disulfuros está catalizada por enzimas que se denominan catalizadores de formación de enlaces disulfuro.

Una enzima en un “estado reducido”, tiene más electrones que su forma oxidada.

15 La “reducción de enlaces disulfuro” se refiere al proceso de escisión de un enlace disulfuro, dando como resultado de ese modo dos grupos tiol. La reducción de enlaces disulfuro está mediada por el intercambio tiol-disulfuro entre las cisteínas del sitio activo de enzimas y cisteínas en la proteína diana.

20 La expresión “isomerización de enlaces disulfuro” se refiere a un intercambio de enlaces disulfuro entre diferentes cisteínas, es decir, la reorganización de los enlaces disulfuro. La isomerización de los enlaces disulfuro está mediada por el intercambio tiol-disulfuro entre las cisteínas del sitio activo de enzimas y cisteínas en la proteína diana y está catalizada por isomerasas. En *E. coli*, la isomerización está catalizada por DsbC o DsbG, una oxidoreductasa de enlaces disulfuro periplásmica.

Un “catalizador de la formación de enlaces disulfuro” es un agente que estimula la formación de enlaces disulfuro. Un agente de este tipo debe estar en un estado oxidado para ser activo.

25 Un “catalizador de la isomerización de enlaces disulfuro”, también denominado “enlace disulfuro isomerasa” es un agente que estimula la isomerización de enlaces disulfuro. Un agente de este tipo debe estar en una forma reducida para ser activo.

La expresión “poner en contacto” se usa en el presente documento de manera intercambiable con lo siguiente: combinado con, añadido a, mezclado con, pasado sobre, incubado con, hecho fluir sobre, etc.

30 Las “proteínas chaperonas” son proteínas que se sabe que promueven el plegamiento apropiado de proteínas recién sintetizadas. Las proteínas chaperonas incluyen, por ejemplo, factor desencadenante, miembros de la familia de chaperona Hsp70, por ejemplo DnaK; miembros de la familia de chaperona Hsp100, por ejemplo ClpB, y miembros de la familia de chaperona Hsp60, por ejemplo GroEL. Véase, por ejemplo, Sorensen y Mortensen, BioMed Central, www.microbialcellfactories.com/content/4/1/1. Los chaperones también se sabe que permiten el plegamiento de la proteína a 4°C, por ejemplo, Cpn60 y Cpn 10 de *Oleispira antartica* RB8^T. Véase, por ejemplo, Id. y Ferrer *et al.*, Nat. Biotechnol. 21:1266-1267 (2003).

40 “Proteína disulfuro isomerasas” o “proteínas PDI” pueden producir o reorganizar los enlaces disulfuro. Las proteínas PDI se describen, por ejemplo, en Georgiou *et al.* patente estadounidense n.º 6.027.888, que se incorpora en el presente documento como referencia para todos los fines. Las proteínas PDI se derivan de organismos procariotas y eucariotas. Las proteínas PDI eucariotas incluyen las de la familia Interpro IPR005792 proteína disulfuro isomerasa. Las proteínas PDI eucariotas incluyen proteínas PDI de por ejemplo, PDI de hígado de rata, proteínas Ero1p y Pdi1p de *Saccharomyces*. Las proteínas procariotas incluyen por ejemplo, DsbC de *E. coli*. Véase, por ejemplo, Frand *et al.*, Trends in Cell Biol. 10:203-210 (2000).

45 Otras proteínas procariotas que actúan para mantener el estado redox de los enlaces disulfuro de proteínas incluyen, por ejemplo, DsbB, DsbA, DsbC, DsbD y DsbG de *E. coli*. Estas proteínas se conocen bien en la técnica y se describen en, por ejemplo, Beckwith *et al.* patente estadounidense n.º 6.872.563.

50 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 prolongar el tiempo de aclaramiento 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.

El término “actividad específica” tal como se usa en el presente documento se refiere a la actividad catalítica de una enzima, por ejemplo, una 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 μmol de producto por minuto a una temperatura (por ejemplo, a 37°C) y valor de pH (por ejemplo, a pH 7,5) dados. 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.

Oligosacáridos “N-unidos” son los oligosacáridos que están unidos a una estructura principal de péptido a través de asparagina, por medio de un enlace asparagina-N-acetilglucosamina. Los oligosacáridos N-unidos también se denominan “N-glicanos”. Los oligosacáridos N-unidos que se producen de manera natural tienen un núcleo pentasacárido común de $\text{Man}_3\text{GlcNAc}_2$. Difieren en la presencia de y en el número de ramificaciones (también denominadas antena) de azúcares periféricos tales como N-acetilglucosamina, galactosa, N-acetilgalactosamina, fucosa y ácido siálico. Opcionalmente, esta estructura también puede contener una molécula de xilosa y/o molécula de fucosa de núcleo. Usando las glicosiltransferasas eucariotas solubles producidas mediante los métodos de la invención, pueden producirse oligosacáridos que imitan estructuras N-unidas naturales o que están diseñadas por el usuario. En una realización de la invención, las glicosiltransferasas eucariotas solubles que generan oligosacáridos N-unidos se expresan en uno o más microorganismos que tienen un entorno intracelular oxidante. Las glicosiltransferasas eucariotas solubles que generan oligosacáridos N-unidos incluyen, por ejemplo, las enzimas GnT1, GalT1 y ST3Gal3.

Los oligosacáridos “O-unidos” son los oligosacáridos que están unidos a una estructura principal de péptido a través de treonina, serina, hidroxiprolina, tirosina o otros aminoácidos que contiene hidroxilo. Usando las glicosiltransferasas eucariotas solubles producidas mediante los métodos de la invención, pueden producirse oligosacáridos que imitan estructuras O-unidas naturales o que están diseñadas por el usuario. En una realización de la invención, las glicosiltransferasas eucariotas solubles que generan oligosacáridos O-unidos se expresan en uno o más microorganismos que tienen un entorno intracelular oxidante. Las glicosiltransferasas eucariotas solubles que generan oligosacáridos O-unidos incluyen, por ejemplo, las enzimas GalNAc-T2, Cor-1-Gal-T1, ST6GalNAc-1 y ST3Gal-1.

Una “glicoforma sustancialmente uniforme” o un “patrón de glicosilación sustancialmente uniforme”, cuando se hace referencia a una especie de glicoproteína, se refiere al porcentaje de sustratos aceptores que se glicosilan mediante la glicosiltransferasa de interés (por ejemplo, fucosiltransferasa). El experto en la técnica entenderá que el material de partida puede contener sustratos aceptores glicosilados. Por tanto, la cantidad calculada de la 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.

El término “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 de 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 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 glicosilado en una única reacción. En realizaciones preferidas, escala comercial se refiere a la producción de al menos aproximadamente 0,2, 0,5, 1, 2, 5, 10, 15, 25, 50, 75, 80, 90 ó 100, 125, 150, 175, 200, 500 ó 1000 gramos de un producto glicosilado en una única reacción. Producción a escala comercial de un polipéptido de glicosiltransferasa eucariota se refiere a la producción a escala de gramos de un polipéptido de glicosiltransferasa eucariota. En realizaciones preferidas, escala comercial se refiere a la producción de entre 1 U/kg de proteína y 1000 U/kg de proteína.

El término “sustancialmente” en las definiciones anteriores de “sustancialmente uniforme” generalmente significa que 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 están glicosilados.

El término “aminoácido” se refiere a aminoácidos no naturales y que se producen de manera natural, por ejemplo, aminoácidos sintéticos, así como análogos de aminoácido y miméticos de aminoácido que funcionan de una manera similar a los aminoácidos que se producen de manera natural. Los aminoácidos que se producen de manera natural son los codificados por el código genético, así como los aminoácidos que se modifican posteriormente, por ejemplo, hidroxiprolina, γ -carboxiglutamato y O-fosfoserina. Análogos de aminoácido se refiere 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, sulfóxido de metionina, metilsulfonio de metionina. Tales análogos tienen grupos R modificados (por ejemplo, norleucina) 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 una

manera similar a un aminoácido que se produce de manera natural.

“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 denominado polipéptido. Cuando los aminoácidos son α -aminoácidos, puede usarse o bien el isómero óptico L o bien el isómero óptico D. Adicionalmente, también se incluyen aminoácidos no naturales, por ejemplo, β -alanina, fenilglicina y homoarginina. También pueden usarse aminoácidos que no están codificados por genes en la presente invención. Además, también pueden usarse en la invención aminoácidos que se han modificado para que incluyan grupos reactivos. Todos los aminoácidos usados en la presente invención pueden ser o bien el isómero D bien el L. Se prefieren generalmente los isómeros L. Además, otros peptidomiméticos son también útiles en la presente invención. Para una revisión general, véase, Spatola, A. F., en *Chemistry and Biochemistry of Amino Acids, Peptides and Proteins*, B. Weinstein, eds., Marcel Dekker, Nueva York, p. 267 (1983).

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 los genes se modifican y se reintroducen en la célula por medios artificiales. El término también abarca células que contienen un ácido nucleico endógeno para la célula que se ha modificado sin retirar el ácido nucleico de la célula; tales modificaciones incluyen las obtenidas mediante sustitución génica, mutación específica de sitio y técnicas relacionadas. Una “proteína recombinante” es una que se ha producido mediante una célula recombinante. En realizaciones preferidas, una glicosiltransferasa eucariota recombinante se produce mediante una célula bacteriana recombinante.

Una “proteína de fusión” se refiere a una proteína que comprende secuencias de aminoácidos que están además de, en lugar de, son menos que y/o son diferentes de las secuencias de aminoácidos que codifican para la proteína de longitud completa u original o subsecuencias de la misma. Puede añadirse más de un dominio adicional a una glicosiltransferasa tal como se describe en el presente documento, por ejemplo, un dominio auxiliar y una etiqueta de epítipo o etiqueta de purificación, o múltiples etiquetas de epítipo o etiquetas de purificación.

Los componentes de las proteínas de fusión incluyen “enzimas auxiliares” y/o “etiquetas de purificación”. Una “enzima auxiliar” tal como se hace referencia en el presente documento es una enzima que está implicada en la catalización de 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 un resto donador por una glicosiltransferasa. Una enzima auxiliar también puede ser una que se usa en la generación de un trifosfato de nucleótido requerido para la formación de un azúcar de nucleótido, o en la generación del azúcar que se incorpora al azúcar de nucleótido. Se dan a conocer ejemplos de enzimas auxiliares y fusión de enzimas auxiliares, por ejemplo, en la solicitud PCT CA98/01180, presentada el 15 de diciembre de 1998.

Las glicosiltransferasas recombinantes de la invención pueden construirse y expresarse como una proteína de fusión con un “etiqueta de purificación” molecular en un extremo, que facilita la purificación de la proteína. Tales etiquetas también pueden usarse para la inmovilización de una proteína de interés durante la reacción de glicosilación. Las etiquetas adecuadas incluyen “etiquetas de epítipo”, que son una secuencia de proteína que se reconoce específicamente por un anticuerpo. Se incorporan generalmente etiquetas de epítipo en las proteínas de fusión para permitir el uso de un anticuerpo fácilmente disponible para detectar o aislar inequívocamente la proteína de fusión. Una “etiqueta FLAG” es una etiqueta de epítipo comúnmente usada, que se reconoce específicamente por un anticuerpo anti-FLAG monoclonal, que consiste en la secuencia AspTyrLysAspAspAspAspLys (SEQ ID NO:201) o una variante sustancialmente idéntica del mismo. Otras etiquetas de epítipo que pueden usarse en la invención incluyen, por ejemplo, etiqueta myc, AU1, AU5, DDDDK (SEQ ID NO:202) (EC5), etiqueta E, etiqueta E2, Glu-Glu, un péptido de 6 residuos, EYMPME (SEQ ID NO:203), derivado de la proteína T media de Polioma, HA, HSV, IRS, KT3, etiqueta S, etiqueta S1, etiqueta T7, etiqueta V5, VSV-G, β -galactosidasa, Gal4, proteína fluorescente verde (GFP), luciferasa, proteína C, proteína A, proteína de unión a celulosa, GST (glutathión S-transferasa), una etiqueta step; Nus-S, PPI-asas, Pfg 27, proteína de unión a calmodulina, dsb A y fragmentos de los mismos, y granzima B. Están disponibles comercialmente anticuerpos y péptidos de epítipo que se unen específicamente a secuencias de epítipo de, por ejemplo, Covance Research Products, Inc.; Bethyl Laboratories, Inc.; Abcam Ltd.; y Novus Biologicals, Inc.

Los expertos en la técnica conocen otras etiquetas de purificación adecuadas, e incluyen, por ejemplo, una etiqueta de afinidad tal como un péptido de hexahistidina (SEQ ID NO:204) u otros péptidos de polihistidina, que se unirán a iones metálicos tales como iones níquel o cobalto. Las proteínas que comprenden etiquetas de purificación pueden purificarse usando una pareja de unión que une a la etiqueta de purificación, por ejemplo, anticuerpos frente a la etiqueta de purificación, resinas o iones níquel o cobalto y amilosa, maltosa o una ciclodextrina. Las etiquetas 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 Alfa Diagnostic International, Inc.), y la mitad carboxilo terminal de la proteína SUMO (vectores y anticuerpos disponibles comercialmente de por ejemplo, Life Sensors Inc.). Los dominios de unión a almidón, tales como un dominio de unión a maltosa de *E. coli* y SBD (dominio de unión a almidón) de una amilasa de *A. niger*, se describen en el

documento WO 99/15636, incorporado en el 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 se describe en el documento USSN 60/468.374, presentado el 5 de mayo de 2003.

Las glicosiltransferasas también pueden incluir una etiqueta proteica de autoescisión, tal como una "inteína". Las inteínas facilitan el retiro de, por ejemplo, una etiqueta de epítipo o de purificación. Las inteínas y los kits para su uso están disponibles comercialmente, por ejemplo, de New England Biolabs.

La expresión "dominio funcional" con referencia a las glicosiltransferasas se refiere a un dominio de la glicosiltransferasa que confiere o modula una actividad de la enzima, por ejemplo, especificidad de sustrato aceptor, actividad catalítica, afinidad de unión, localización dentro del aparato de Golgi, anclaje a una membrana celular u otra actividad bioquímica o biológica. Los ejemplos de dominios funcionales de glicosiltransferasas incluyen, pero no se limitan a, el dominio catalítico, la región de tallo, el dominio transmembrana o de señal-anclaje y la cola citoplasmática amino terminal.

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 las unidades de actividad descritas en el presente documento o conocidas por el 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 descripción cuantitativa y la cuantificación 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. Otro método de determinación de la expresión de proteínas es analizar un lisado u otra muestra que contiene la proteína usando electroforesis en gel, por ejemplo, SDS-PAGE, seguido por una etapa de visualización. Las etapas de visualización incluyen tintes y colorantes de proteína, por ejemplo, tinte de plata o Comassie, o inmunoensayos, tales como análisis de inmunotransferencia de tipo Western usando un anticuerpo que se unirá específicamente a la proteína de interés. Los anticuerpos pueden dirigirse contra la glicosiltransferasa o contra una etiqueta de epítipo o de purificación covalentemente unida a la proteína.

El término "actividad enzimática" se refiere a una actividad de una enzima y puede medirse mediante los ensayos y las unidades descritas en el presente documento o conocidas por un experto en la técnica. Los ejemplos de una actividad de una glicosiltransferasa incluyen, pero sin limitarse a, las asociadas con los dominios funcionales de la enzima, por ejemplo, especificidad de sustrato aceptor, actividad catalítica, afinidad de unión, localización dentro del aparato de Golgi, anclaje a una membrana celular, u otra actividad bioquímica o biológica.

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 está ubicado adyacente al dominio transmembrana o de señal-anclaje, entre la región de membrana y el dominio catalítico más corto, y se ha notificado que funciona como una señal de retención para mantener la glicosiltransferasa en el aparato de Golgi y como un sitio de escisión proteolítica. Las regiones de tallo generalmente comienzan con el primer aminoácido hidrófilo tras el dominio transmembrana hidrófobo y terminan en el dominio catalítico, o en algunos casos el primer residuo de cisteína tras el dominio transmembrana. Las regiones de tallo a modo de ejemplo incluyen, pero sin limitarse 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 36 hasta aproximadamente 103 (véase, por ejemplo, la enzima humana); la región de tallo de GalT1 de mamífero, residuos de aminoácido desde aproximadamente 71 hasta aproximadamente 129 (véase por ejemplo, la enzima bovina); la región de tallo de ST3GalIII de mamífero, residuos de aminoácido desde aproximadamente 29 hasta aproximadamente 84 (véase, por ejemplo, la enzima de rata); la región de tallo de Core-1-Gal-T1 de invertebrados, residuos de aminoácido desde aproximadamente 36 hasta aproximadamente 102 (véase por ejemplo, la enzima de *Drosophila*); la región de tallo de Core-1-Gal-T1 de mamífero, residuos de aminoácido desde aproximadamente 32 hasta aproximadamente 90 (véase por ejemplo, la enzima humana); la región de tallo de ST3Gal1 de mamífero, residuos de aminoácido desde aproximadamente 28 hasta aproximadamente 61 (véase por ejemplo, la enzima porcina) o para la enzima humana los residuos de aminoácido desde aproximadamente 18 hasta aproximadamente 58; la región de tallo de ST6GalNAc-1 mamífero, residuos de aminoácido desde aproximadamente 30 hasta aproximadamente 207 (véase por ejemplo, la enzima murina), aminoácidos 35-278 para la enzima humana o aminoácidos 37-253 para la enzima del pollo; la región de tallo de GalNAc-T2 de mamífero, residuos de aminoácido desde aproximadamente 71 hasta aproximadamente 129 (véase por ejemplo, la enzima de rata).

Un "dominio catalítico" se refiere a un dominio de proteína, o una subsecuencia de la misma, 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 entera, 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 sin limitarse a, el dominio catalítico de fucosiltransferasa VII, residuos de aminoácido 39-342; el dominio catalítico de GnT1 de mamífero, residuos de

aminoácido desde aproximadamente 104 hasta aproximadamente 445 (véase, por ejemplo, la enzima humana); el dominio catalítico de GalT1 de mamífero, residuos de aminoácido desde aproximadamente 130 hasta aproximadamente 402 (véase por ejemplo, la enzima bovina); y el dominio catalítico de ST3Gal3 de mamífero, residuos de aminoácido desde aproximadamente 85 hasta aproximadamente 374 (véase, por ejemplo, la enzima de la rata). Se describen dominios catalíticos y mutantes de truncamiento de proteínas de GalNAc-T2 en el documento USSN 60/576.530 presentado el 3 de junio de 2004; y solicitud de patente provisional estadounidense con número de expediente del apoderado 040853-01-5149-P1, presentada el 3 de agosto de 2004. Los dominios catalíticos también pueden identificarse mediante alineación con glicosiltransferasas conocidas.

Una "subsecuencia" se refiere a una secuencia de ácidos nucleicos o aminoácidos que comprende una parte de una secuencia más larga de ácidos nucleicos o aminoácidos (por ejemplo, proteína) respectivamente.

Un "truncamiento de glicosiltransferasa" o una "glicosiltransferasa truncada" o sus 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 GnTI truncadas, enzimas GalT1 truncadas, enzimas ST3GalIII truncadas, enzimas GalNAc-T2 truncadas, enzimas core 1 GalT1 truncadas, residuos de aminoácido desde aproximadamente 32 hasta aproximadamente 90 (véase por ejemplo, la enzima humana); enzimas ST3Gal1 truncadas, enzimas ST6GalNAc-1 truncadas y enzimas GalNAc-T2 truncadas. Cualquier número de residuos de aminoácido puede delecionarse siempre y que la enzima conserve su actividad. En algunas realizaciones, pueden delecionarse dominios o partes de dominios, por ejemplo, un dominio de señal-anclaje puede delecionarse dejando un truncamiento que comprende una región de tallo y un dominio catalítico; un dominio de señal-anclaje y una parte de una región de tallo pueden delecionarse dejando un truncamiento que comprende la región de tallo restante y un dominio catalítico; o un dominio de señal-anclaje y una región de tallo pueden delecionarse dejando un truncamiento que comprende un dominio catalítico. También pueden producirse truncamientos de glicosiltransferasa en el extremo C-terminal de la proteína. Por ejemplo, algunas enzimas GalNAcT tienen un dominio de lectina C-terminal que puede delecionarse sin disminuir la actividad enzimática.

El término "ácido nucleico" se refiere a un polímero de ribonucleótidos o desoxirribonucleótidos en forma o bien mono o bien bicatenaria, y a menos que se limite de otra manera, abarca análogos conocidos de nucleótidos naturales que se hibridan con ácidos nucleicos de una manera similar a nucleótidos que se producen de manera natural. A menos que se indique de lo contrario, 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 a 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 necesarios o útiles adicionales para efectuar 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 a partir de la célula huésped. También pueden incluirse en un casete de expresión señales de terminación de la transcripción, potenciadores y otras secuencias de ácido nucleico que influyen en la expresión génica. 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 una que se origina a partir de una fuente foránea a 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 una glicoproteína que es endógeno para 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 estar operativamente unido al promotor. Técnicas tales como mutagénesis dirigida al sitio son también útiles para modificar una secuencia heteróloga.

El término "aislado" se refiere a material que está sustancial o esencialmente libre de componentes que interfieren con la actividad de una enzima. Para un sacárido, proteína o ácido nucleico de la invención, el término "aislado" se refiere a material que está sustancial o esencialmente libre de componentes que normalmente acompañan al material tal como se encuentra en su estado nativo. Normalmente, un sacárido, proteína o ácido nucleico aislado de la invención es al menos aproximadamente el 80% puro, habitualmente al menos aproximadamente el 90% y preferiblemente al menos aproximadamente el 95% puro tal como se mide mediante la intensidad de bandas 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, una proteína o ácido nucleico en una muestra puede resolverse mediante electroforesis en gel de poliacrilamida, y entonces la proteína o ácido nucleico puede visualizarse mediante tinción. Para ciertos fines puede ser deseable alta resolución de la proteína o ácido nucleico y puede utilizarse HPLC o un medio similar para la purificación, por ejemplo.

La expresión “operativamente unido” se refiere al enlace funcional entre una secuencia de control de la expresión de ácido nucleico (tal como un promotor, secuencia señal o alineamiento de sitios de unión de factores 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.

- 5 Los términos “idéntico” o porcentaje de “identidad,” en el contexto de dos o más ácidos nucleicos o secuencias de proteínas, se refieren a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de residuos de aminoácido o nucleótidos que iguales, cuando se comparan y alinean para una correspondencia máxima, tal como se mide usando uno de los siguientes algoritmos de comparación de secuencias o mediante inspección visual.
- 10 La frase “sustancialmente idéntico/a”, en el contexto de dos ácidos nucleicos o proteínas, se refiere a dos o más secuencias o subsecuencias que tienen al menos más de aproximadamente el 60% de identidad de secuencia de aminoácidos o ácido nucleico, el 65%, el 70%, el 75%, el 80%, el 85%, el 90%, preferiblemente el 91%, el 92%, el 93%, el 94%, el 95%, el 96%, el 97%, el 98% o el 99% de identidad de residuos de aminoácido o nucleótido, cuando se comparan y alinean para una correspondencia máxima, tal como se mide usando uno de los siguientes algoritmos de comparación de secuencias o mediante inspección visual. Preferiblemente, existe identidad sustancial a lo largo de una región de las secuencias que tiene al menos aproximadamente de 50 residuos de longitud, más preferiblemente a lo largo de una región de al menos aproximadamente 100 residuos y lo más preferiblemente las secuencias son sustancialmente idénticas a lo largo de al menos aproximadamente 150 residuos. En la realización más preferida, las secuencias son sustancialmente idénticas a lo largo de toda la longitud de las regiones codificantes.
- 20

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, las secuencias de referencia y de prueba se introducen en un ordenador, se designan coordenadas de subsecuencia, si es necesario, y se designan parámetros del programa de algoritmo de secuencia. El algoritmo de comparación de secuencias calcula entonces el porcentaje de identidad de secuencia para la(s) secuencia(s) de prueba con respecto a la secuencia de referencia, basándose en los parámetros de programa designados.

25

Puede realizarse una alineación de secuencias óptima para la comparación, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman, Adv. Appl. Math. 2:482 (1981), mediante el algoritmo de alineación de homología de Needleman y Wunsch, J. Mol. Biol. 48:443 (1970), mediante el método de búsqueda de similitud de Pearson y Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), mediante implementaciones informatizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante inspección visual (véase 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 and Sons, Inc., (suplemento de 1995) (Ausubel)).

30

Ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (1990) J. Mol. Biol. 215: 403-410 y Altschul *et al.* (1977) Nucleic Acids Res. 25: 3389-3402, respectivamente. El software para realizar análisis BLAST está disponible públicamente a través del National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/). Este algoritmo implica en primer lugar identificar 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 alguna puntuación T umbral de valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. T se denomina umbral de puntuación de palabra vecina (Altschul *et al.*, citado anteriormente). Estos aciertos de palabras vecinas iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largas que los contengan. Los aciertos de palabras se extienden entonces en ambas direcciones a lo largo de cada secuencia tanto como pueda 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 las secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabra en cada dirección se detiene cuando: la puntuación de alineación acumulativa desciende en la cantidad X desde su valor máximo logrado; la puntuación acumulativa baja hasta 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 del algoritmo BLAST W, T y X determinan la sensibilidad y la velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa como valores por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, M=5, N=-4, y una comparación de ambas hebras. Para secuencias de aminoácidos, el programa BLASTP usa como valores por defecto una longitud de palabra (W) de 3, una expectativa (E) de 10 y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)).

35

40

45

50

55

Además, para calcular el porcentaje de identidad de secuencia, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)). Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la cual una coincidencia entre dos

60

secuencias de aminoácidos o nucleótidos se produciría de casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de prueba con respecto al ácido nucleico de referencia es inferior a aproximadamente 0,1, más preferiblemente inferior a aproximadamente 0,01 y lo más preferiblemente inferior a 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 produce reacciones cruzadas inmunológicamente con la proteína codificada por el segundo ácido nucleico, tal como se describe a continuación. Por tanto, normalmente una proteína es sustancialmente idéntica a una segunda proteína, por ejemplo, cuando los dos péptidos difieren sólo por sustituciones conservativas. Otra indicación de que dos secuencias de ácido nucleico son sustancialmente idénticas es que las dos moléculas hibridan entre sí en condiciones de rigurosidad, tal como se describe a continuación.

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

La expresión “condiciones de rigurosidad” se refiere a condiciones en las que una sonda hibridará con su subsecuencia diana, pero no con otras secuencias. Las condiciones de rigurosidad dependen de la secuencia y serán diferentes en diferentes circunstancias. Secuencias más largas hibridan específicamente a temperaturas superiores. Generalmente, las condiciones de rigurosidad se seleccionan para que sean aproximadamente 15°C inferiores al punto de fusión térmico (T_m) para la secuencia específica a un pH y fuerza iónica definidos. La T_m es la temperatura (en concentración de ácido nucleico, pH y fuerza iónica definidos) a la que el 50% de las sondas complementarias a la secuencia diana hibridan con la secuencia diana en equilibrio. (Como las secuencias dianas están generalmente presentes en exceso, a la T_m , el 50% de las sondas están ocupadas en equilibrio). Normalmente, condiciones de rigurosidad serán aquellas en las que la concentración salina es inferior a aproximadamente ion Na 1,0 M, normalmente concentración de ion Na de aproximadamente 0,01 a 1,0 M (u otras sales) a pH de 7,0 a 8,3 y la temperatura es al menos aproximadamente de 30°C para sondas cortas (por ejemplo, de 10 a 50 nucleótidos) y al menos aproximadamente de 60°C para sondas largas (por ejemplo, más de 50 nucleótidos). También pueden lograrse condiciones rigurosas con la adición de agentes desestabilizantes tales como formamida. Para hibridación específica o selectiva, una señal positiva es normalmente al menos dos veces la hibridación de fondo, preferiblemente, 10 veces la hibridación de fondo. Condiciones de hibridación rigurosas a modo de ejemplo pueden ser tal como sigue: formamida al 50%, 5x SSC y SDS al 1%, incubación a 42°C, o, 5x SSC, SDS al 1%, incubación a 65°C, con lavado en 0,2x SSC y SDS al 0,1% a 65°C. Para la PCR, una temperatura de aproximadamente 36°C es típica para amplificación de rigurosidad baja, aunque las temperaturas de apareamiento pueden variar entre aproximadamente 32-48°C dependiendo de la longitud del cebador. Para amplificación por PCR de rigurosidad alta, una temperatura de aproximadamente 62°C es típica, aunque las temperaturas de apareamiento de rigurosidad alta pueden oscilar desde aproximadamente 50°C hasta aproximadamente 65°C, dependiendo de la longitud del cebador y la especificidad. Condiciones de ciclo típicas para amplificaciones de rigurosidad tanto alta como baja 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 rigurosidad baja y alta, 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 “inmunorreacciona específicamente con”, cuando se refieren a un anticuerpo se refieren a una reacción de unión que es determinante de la presencia de la proteína en presencia de una población heterogénea de proteínas y otros compuestos biológicos. Por tanto, en condiciones de inmunoensayo designadas, 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 un anticuerpo que se selecciona por su especificidad para una proteína particular. Pueden usarse una variedad de formatos de inmunoensayo para seleccionar anticuerpos específicamente inmunorreactivos con una proteína particular. Por ejemplo, se usan inmunoensayos de ELISA en fase sólida de manera rutinaria para seleccionar anticuerpos monoclonales específicamente inmunorreactivos con una proteína. Véase Harlow y Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, Nueva York, para una descripción de formatos de inmunoensayos y condiciones que pueden usarse para determinar inmunorreactividad específica.

“Variantes modificadas de manera conservativa” de una secuencia de polinucleótidos particular se refiere a aquellos polinucleótidos que codifican para secuencias de aminoácidos idénticas o esencialmente idénticas, o cuando el polinucleótido no codifica para una secuencia de aminoácidos, 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 cada posición en la que se especifica una arginina por un codón, el codón puede alterarse a cualquiera de los codones correspondientes descritos sin alterar la proteína codificada. Tales variaciones de ácido nucleico son “variaciones silenciosas”, que son una especie de “variantes 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 cuando se indique lo contrario. Un experto reconocerá que cada codón en un ácido nucleico (excepto AUG, que es habitualmente el único codón para metionina y UGG, que es habitualmente 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 adiciones, deleciones o sustituciones individuales que alteran, añaden o delecionan un aminoácido individual o un pequeño porcentaje de aminoácidos (normalmente menos del 5%, más normalmente menos del 1%) en una secuencia codificada son "variantes modificadas de manera conservativa" en las que las alteraciones dan como resultado la sustitución de un aminoácido por un aminoácido químicamente similar. Se conocen bien en la técnica tablas de sustituciones conservativas que proporcionan aminoácidos funcionalmente similares.

Un experto apreciará que muchas variaciones conservativas de proteínas, por ejemplo, glicosiltransferasas, y ácido nucleico que codifica 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 da 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 su expresión en una célula huésped particular usada para producir las glicosiltransferasas quiméricas (por ejemplo, de levadura, humana y similares). De manera similar, las "sustituciones de aminoácido conservativas", en uno o algunos aminoácidos en una secuencia de aminoácidos se sustituyen por aminoácidos diferentes con propiedades sumamente similares (véase la sección de definiciones; citada anteriormente), también se identifican fácilmente como sumamente similares a una secuencia de aminoácidos particular, o a una secuencia de ácido nucleico particular que codifica para un aminoácido. Tales variantes sustituidas de manera conservativa de cualquier secuencia particular son una característica de la presente invención. Véase también, Creighton (1984) *Proteins*, W.H. Freeman y Company. Además, adiciones, deleciones o sustituciones individuales que alteran, añaden o delecionan un aminoácido individual o un pequeño porcentaje de aminoácidos en una secuencia codificada son también "variantes modificadas de manera conservativa".

La práctica de esta invención puede involucrar la construcción de ácidos nucleicos recombinantes y la expresión de genes en células huésped, preferiblemente células huésped bacterianas. Se conocen bien en la técnica técnicas de clonación molecular para lograr estos resultados. 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 and Sons, Inc., (suplemento de 1999) (Ausubel). Los expertos en la técnica conocen bien células huésped adecuadas para la expresión de los polipéptidos recombinantes, e incluyen, por ejemplo, células procariotas, tales como *E. coli*, y células eucariotas incluyendo células de insectos, de mamífero y fúngicas (por ejemplo, *Aspergillus niger*)

Se encuentran ejemplos de protocolos suficientes para dirigir a los expertos a través de métodos de amplificación *in vitro*, incluyendo la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR), la amplificación por Q β -replicasa y otras técnicas mediadas por ARN polimerasa en Berger, Sambrook y Ausubel, así como Mullis *et al.* (1987) 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 y Levinson (1 de octubre de 1990) CyEN 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) Gene 4: 560; y Barringer *et al.* (1990) Gene 89: 117. Se describen métodos mejorados de clonación de ácidos nucleicos amplificados *in vitro* en Wallace *et al.*, patente estadounidense n.º 5.426.039.

Descripción detallada de la invención

I. INTRODUCCIÓN

Esta invención proporciona por primera vez métodos de mejora de la producción de glicosiltransferasas eucariotas activas, solubles en microorganismos usando organismos procariotas, por ejemplo, bacterias que tienen entornos intracelulares oxidantes. La invención abarca mejorar la expresión de glicosiltransferasas a través del uso de organismos procariotas que tienen de manera natural entornos intracelulares oxidantes, tales como *Pseudomonas*. La invención también abarca el uso de organismos procariotas que tienen entornos intracelulares reductores de manera natural, pero que se manipulan para que tengan un entorno intracelular oxidante. Por ejemplo, *E. coli* tiene generalmente un entorno intracelular reductor. La expresión de proteínas heterólogas en *E. coli* frecuentemente puede ser difícil o poco práctica debido a que los enlaces disulfuro no se oxidan apropiadamente, conduciendo a un mal plegamiento de la proteína y a la expresión de proteínas en cuerpos de inclusión. *E. coli* y otros organismos procariotas que tienen un entorno intracelular reductor de manera natural pueden manipularse, sin embargo, para generar un entorno intracelular que favorece la oxidación de enlaces disulfuro. Por ejemplo, puede manipularse *E. coli* para reducir la actividad de proteínas reductasa endógenas o bien mediante mutación de los ácidos nucleicos que codifican para esas proteínas o bien mediante manipulación de otras actividades de proteínas en un ciclo de

oxidación-reducción intracelular. Por ejemplo, en *E. coli*, la inactivación de mutaciones en la proteína tiorredoxina reductasa (trxB), la proteína glutatión reductasa (gor) o en ambas proteínas da como resultado células que tienen un entorno oxidante. Están disponibles comercialmente células de *E. coli* que tienen mutaciones en trxB y gor, por ejemplo, de EMD Biosciences, Inc.

- 5 En realizaciones preferidas, las glicosiltransferasas eucariotas activas solubles se expresan intracelularmente dentro del microorganismo procariota que tiene un entorno oxidante, es decir, dentro de la membrana celular y no en el espacio periplasmático o se secretan de la célula.

10 En otra realización preferida, las glicosiltransferasas eucariotas solubles expresadas en un microorganismo procariota que tiene un entorno oxidante mantienen su actividad enzimática y están sin glicosilar, glicosiladas de manera diferente o mínimamente glicosiladas en comparación con la misma glicosiltransferasa eucariota cuando se expresa en una célula eucariota, por ejemplo, una levadura, una célula de mamífero o una célula de origen, es decir, una célula humana para una glicosiltransferasa humana.

15 En una realización, la producción de glicosiltransferasas eucariotas activas, solubles en microorganismos procariotas que tienen entornos intracelulares oxidantes se potencia además haciendo crecer las células en condiciones que reducen el nivel de producción de proteína recombinante, es decir, la glicosiltransferasa eucariota, por debajo del de un nivel máximo.

20 En otra realización preferida, la producción de glicosiltransferasas eucariotas activas, solubles en organismos procariotas que tienen entornos intracelulares oxidantes haciendo crecer las células en condiciones que reducen el nivel de producción de proteína recombinante, por ejemplo, a una temperatura de crecimiento inferior a la óptima, da como resultado tanto un aumento de la expresión como un aumento de la actividad de la glicosiltransferasa eucariota soluble, en comparación con su expresión en una célula huésped del tipo natural o célula con un entorno reductor, por ejemplo, *E. coli*.

II. GLICOSILTRANSFERASAS EUCARIOTAS ACTIVAS SOLUBLES Y SU EXPRESIÓN EN MICROORGANISMOS PROCARIOTAS CON ENTORNOS INTRACELULARES OXIDANTES

25 Cualquier glicosiltransferasa que sea predominantemente insoluble cuando se expresa en un entorno reductor, por ejemplo, *E. coli* de tipo natural, puede expresarse en un organismo procariota que tiene un entorno oxidante intracelular para facilitar la expresión de una proteína soluble, activa. Entonces, las glicosiltransferasas se usan para la síntesis o remodelación de oligosacáridos, glicoproteínas, glicopéptidos o glicolípidos. Las glicosiltransferasas preferidas incluyen glicosiltransferasas eucariotas, tal como se describe en el presente documento.

30 A. Glicosiltransferasas eucariotas

Puede usarse cualquier glicosiltransferasa eucariota en los métodos de la presente invención. Las glicosiltransferasas eucariotas pueden ser las proteínas no modificadas, que se producen de manera natural o puede ser glicosiltransferasas que se han modificado para potenciar la actividad catalítica, o estabilidad, u otras características de las proteínas. La modificación de glicosiltransferasas eucariotas incluyen por ejemplo, 35 truncamiento de la proteína para eliminar por ejemplo, la región de tallo, el dominio de señal-anclaje, o una parte de la región de tallo o el dominio de señal-anclaje, o la eliminación de tanto la región de tallo como el dominio de señal-anclaje; o la eliminación de un residuo de cisteína desapareado mediante sustitución por otro residuo de aminoácido. Una glicosiltransferasa también puede truncarse en el extremo C-terminal para eliminar un dominio o dominios no catalíticos. Por ejemplo, un dominio de lectina C-terminal puede eliminarse de las enzimas GalNAcT sin disminuir la actividad enzimática. Se describen glicosiltransferasas modificadas, por ejemplo, en el documento USSN 40 60/542.210, presentado el 4 de febrero de 2004; el documento USSN 60/599.406, presentado el 6 de agosto de 2004; el documento USSN 60/627.406, presentado el 12 de noviembre de 2004; el documento USSN 60/576.433, presentado el 3 de junio de 2004; el documento USSN 60/650.011, presentado el 4 de febrero de 2005; el documento PCT/US05/19583, presentado el 3 de junio de 2005; el documento USSN 60/576.530, presentado el 3 de junio de 2004; el documento USSN 60/598.584, presentado el 3 de agosto de 2004; el documento PCT/US05/19442, 45 presentado el 3 de junio de 2005; el documento PCT/US05/03856, presentado el 4 de febrero de 2005; y el documento WO 2004/063344;

Las realizaciones preferidas de la invención incluyen métodos de producción de, por ejemplo, una N-acetilglucosaminiltransferasa eucariota (GnTI o GNTI, GnTII o GNTII, GnTIII o GNTIII, GnTIV o GNTIV, GnTV o GNTV, GnTIV o GNTIV); una N-acetilgalactosaminiltransferasa eucariota (GalNAcT, por ejemplo, GalNAcT1, GalNAcT2, o GalNAcT3); cualquier galactosiltransferasa, por ejemplo, una β -1,4-galactosiltransferasa (GalT1) eucariota o una core 1 galactosiltransferasa (core 1 GalT1) eucariota (core-1-Gal-T1); cualquier sialiltransferasa eucariota, por ejemplo, una α (2,3)sialiltransferasa (ST3Ga13) eucariota, o una α -Nacetilgalactosaminida- α -2,6-sialiltransferasa I (ST6GalNAc-1) eucariota, o una gal β 1,3GalNAc- α 2,3-sialiltransferasa (ST3Gal-1) eucariota; y 55 cualquier fucosiltransferasa eucariota. Se conocen muchos ejemplos de proteínas que tienen las actividades enumeradas anteriormente, véase, por ejemplo, afmb.cnrs-mrs.fr/CAZY/, en particular las familias de glicosiltransferasa 2, 4, 6, 7, 10, 1, 12, 13, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26, 28, 29, 31, 32, 34, 37, 38, 41, 42, 49, 52, 54, 65 ó 68. Otras glicosiltransferasas que pueden producirse usando los métodos descritos se encuentran

en la lista de secuencias informal adjunta.

Tal como se indica anteriormente, las glicosiltransferasas pueden modificarse antes de la producción usando los métodos descritos. Las modificaciones incluyen, por ejemplo, el truncamiento de la glicosiltransferasa para eliminar todo o una parte de un dominio no catalítico, tal como un dominio citoplasmático, un dominio de señal-ancraje, una región de tallo y o un dominio de lectina. Las glicosiltransferasas truncadas a modo de ejemplo que pueden producirse en la presente invención incluyen, por ejemplo, ST3Ga1 III ($\Delta 27$, $\Delta 28$, $\Delta 73$, $\Delta 85$, $\Delta 86$), GnT1 humana ($\Delta 103$), GalT1 bovina ($\Delta 40$, $\Delta 129$, $\Delta 70$), GalNAcT2 humana ($\Delta 51$, $\Delta 40$, $\Delta 73$, $\Delta 94$, $\Delta 51\Delta 445$, $\Delta 53$, $\Delta 53\Delta 445$), ST3Ga11 ($\Delta 45$), Core-1-Gal- T1 de *Drosophila* ($\Delta 31$, $\Delta 50$) y mutantes de ST6GalNAc1 humana mostrados en la tabla 1.

Tabla 1: mutantes de ST6GalNAc1

	Sitio de truncamiento	Mutación (primer aminoácido)
SER HUMANO	$\Delta 35$	K36
	$\Delta 124$	K125
	$\Delta 257$	S258
	$\Delta 35$	K36
	$\Delta 72$	T73
	$\Delta 109$	E110
	$\Delta 133$	M134
	$\Delta 170$	T171
	$\Delta 232$	A233
	$\Delta 272$	G273
POLLO	$\Delta 48$	Q49
	$\Delta 152$	V153
	$\Delta 225$	L226
	$\Delta 226$	R227
	$\Delta 232$	T233
	$\Delta 231$	K232
RATÓN	$\Delta 30$	K31
	$\Delta 31$	D32
	$\Delta 51$	E52
	$\Delta 126$	S127
	$\Delta 185$	S186
	$\Delta 200$	S201

III. ENTORNOS OXIDANTES, INTRACELULARES

En realizaciones preferidas, las glicosiltransferasas eucariotas activas, solubles se expresan en organismos procariotas que tienen entornos citoplasmáticos oxidantes.

A. Identificación de entornos oxidantes, intracelulares

El replegamiento de proteínas y la actividad de proteínas dependen frecuentemente de la correcta formación de enlaces disulfuro. Los enlaces disulfuro son reacciones de intercambio tiol-disulfuro (SH-SS) reversibles que se ven influidas en gran medida por el estado redox del entorno que rodea a la proteína. En muchas células, incluyendo *E. coli* y otros organismos procariotas, el glutatión, un tripéptido que contiene cisteína, es un importante tampón redox de tiol-disulfuro. El estado redox de microorganismos procariotas también se ve afectado por otras proteínas, tales como tiorredoxinas. Las proteínas reductasas, a su vez, regulan el estado redox del glutatión, las glutarredoxinas y tiorredoxinas. En *E. coli* los glutationes, codificados por *gshA* y *gshB*, regulan el estado redox de las glutarredoxinas. Las proteínas reductasas incluyen, por ejemplo, tiorredoxina reductasa y glutatión oxidorreductasa. *E. coli* tiene tiorredoxinas codificadas por los genes *trxA* y *trxC*, glutarredoxina 1, glutarredoxina 2 y glutarredoxina 3, codificadas por los genes *grxA*, *grxB* y *grxC*. Muchas de las proteínas que regulan el estado de oxidación de una célula, por ejemplo, tiorredoxina, glutatión, tiorredoxina reductasa y glutatión oxidorreductasa, comprende un motivo CX_1X_2C de sitio activo. Las proteínas también comprenden un motivo estructural de proteína conocido como pliegue de tiorredoxina.

Un método para identificar organismos procariotas que tienen un entorno intracelular oxidante es medir la razón de glutatión reducido (GSH) con respecto a glutatión oxidado (GSSG). Se han determinado las razones óptimas de GSH/GSSG para el plegamiento de proteínas. *In vitro*, se producen rendimientos máximos de proteína plegada

apropiadamente a razones de GSH/GSSG de menos de 50, preferiblemente menos de 40, más preferiblemente menos de 30, aún más preferiblemente menos de 20 y lo más preferiblemente menos de 10. En células de mamífero, las razones de GSH/GSSG citoplasmáticas oscilaban entre 30/1 y 100/1, mientras las razones de GSH/GSSG de la ruta secretora (en la que se produce la mayor parte del repliegamiento de proteínas) oscilaban entre 1/1 y 3/1. Hwang *et al.*, Science 257:1496-1502 (1992). *E. coli* expresa muy pocas proteínas intracelulares con enlaces disulfuro. Las proteínas de *E. coli* que tienen enlaces disulfuro se secretan en el espacio periplasmático, que tiene un entorno oxidante. Las razones de GSH/GSSG de *E. coli* intracelulares de tipo natural oscilaban entre 50/1 y 200/1. Hwang *et al.* citado anteriormente.

Los métodos de la invención pueden usarse para producir glicosiltransferasas eucariotas solubles en organismos procariotas que tienen un entorno intracelular oxidante. Los microorganismos con un entorno intracelular oxidante tienen normalmente razones de GSH/GSSG de menos de 50, preferiblemente menos de 40, más preferiblemente menos de 30, aún más preferiblemente menos de 20 y lo más preferiblemente menos de 10. Por tanto, en algunas realizaciones, los microorganismos de la invención tendrán razones de GSH/GSSG que oscilan, por ejemplo, entre 0 y 50, o entre 0,1 y 25, o entre 0,5 y 10.

Pueden identificarse organismos procariotas con entornos intracelulares, por ejemplo, determinando la razón de GSH/GSSG intracelular de los organismos procariotas. Están disponibles comercialmente ensayos para la concentración de glutatión total de, por ejemplo, Sigma. Se describen ensayos para la determinación de una razón de GSH/GSSG, por ejemplo, en Hwang *et al.*, Science 257: 1496-1502 (1992). Se describen métodos para cuantificar el contenido intracelular de GSH y GSSG mediante derivatización con N-(1-pirenil)maleimida (NPM) seguida por cuantificación usando HPLC en Ostergaard, *et al.*, J. Cell Biol. 166:337-345 (2004).

Están disponibles varios ensayos adicionales para los expertos para determinar si un organismo procariota tiene un entorno oxidante, intracelular. Esos ensayos incluyen la medición de la actividad glutatión reductasa y el estado redox del conjunto de glutatión (Tuggle y Fuchs, J. Bacter. 162:448-450 (1985)), la sensibilidad a oxidantes específicos de tiol en el medio de cultivo (Prinz *et al.*, J. Biol Chem. 272:15661-15667 (1997)), la activación transcripcional del gen OxyR en *E. coli* tras la exposición a peróxido de hidrógeno o diamida (Bessette *et al.*, PNAS 96:13703-13708 (1999)), la medición del estado redox de un gen indicador, tal como una proteína fluorescente verde sensible a redox, (rxYFP) (Ostergaard *et al.*, J. Cell Biol. 166:337-345 (2004)), la detección de glutatión usando colorantes sensibles a glutatión tales como monoclorobimano, CellTracker Green CMFDA, o-ftaldialdehído y naftaleno-2,3-dicarboxaldehído de por ejemplo, Molecular Probes, y la oxidación del residuo de cisteína en proteínas tras la exposición de las células a un reactivo alquilante de sulfhidrilo, tal como ácido 4-acetamido-4'-maleimidistibeno-2,2-disulfónico (Jurado *et al.*, J. Mol. Biol. 320:1-10 (2002)).

B. Microorganismos procariotas que tienen entornos intracelulares oxidantes.

Los métodos de la invención se llevan a cabo usando microorganismos procariotas que tienen entornos intracelulares oxidantes. Tales microorganismos incluyen microorganismos procariotas que tienen entornos oxidantes intracelulares, endógenos y microorganismos procariotas que se manipulan genéticamente para que tengan un entorno oxidante intracelular.

Algunos organismos procariotas tienen entornos oxidantes intracelulares, endógenos y, por tanto, promueven la formación de enlaces disulfuros de proteína dentro de la célula. Compartimientos intracelulares oxidantes en organismos procariotas excluyen específicamente un espacio periplasmático bacteriano. Pueden usarse organismos procariotas que tienen entornos oxidantes intracelulares, endógenos para producir glicosiltransferasas eucariotas activas, solubles en un compartimiento intracelular. Los organismos procariotas con entornos oxidantes intracelulares, endógenos incluyen miembros de, por ejemplo, especies de *Pseudomonas*, incluyendo *testosteroni*, *putida*, *aeruginosa*, *syringae* y *fluorescens*; algunas bacterias gram positivas; y algunas bacterias gram negativas. Se describen cepas y especies de *Pseudomonas* en, por ejemplo, la publicación de solicitud de patente estadounidense n.º US 2005/0186666, publicada el 25 de agosto de 2005, que se incorpora en el presente documento como referencia para todos los fines. Las bacterias gram positivas incluyen, por ejemplo, especies de *Bacillus*, *Listeria*, *Staphylococcus*, *Streptococcus*, *Enterococcus* y *Clostridium*.

También pueden usarse organismos procariotas con modificación de una ruta redox en los métodos de la invención para producir glicosiltransferasas eucariotas activas, solubles o proteínas terapéuticas activas, solubles. Las modificaciones pueden realizarse en organismos procariotas que tienen un entorno reductor, por ejemplo, *E. coli* u otras bacterias gram negativas o algunas bacterias gram positivas. Los microorganismos procariotas se modifican para promover un entorno intracelular oxidante, mejorando de ese modo el repliegamiento de proteínas y la formación de enlaces disulfuro intracelulares de por ejemplo, las glicosiltransferasas eucariotas.

Muchos organismos procariotas usan dos rutas para reducir los enlaces disulfuro que se forman en algunas proteínas citoplasmáticas, incluyendo proteínas expresadas de manera recombinante. Los componentes de estas rutas pueden manipularse para promover la formación de un entorno oxidante intracelular. La primera ruta es el sistema de tiorredoxina, que generalmente incluye una tiorredoxina reductasa y tiorredoxina. La tiorredoxina reductasa mantiene a la tiorredoxina en un estado reducido. La segunda ruta es el sistema de glutarredoxina, que generalmente incluye una glutatión oxidorreductasa, glutatión y glutarredoxinas. Las mutaciones inactivantes de

algunos componentes de estas rutas redox pueden aumentar en última instancia la formación de enlaces disulfuro en proteínas expresadas, y en el caso de proteínas heterólogas expresadas en el organismo procariota, pueden aumentar la solubilidad y actividad de las proteínas heterólogas expresadas. Por ejemplo, la eliminación en *E. coli* de la actividad tiorredoxina reductasa da como resultado una acumulación de tiorredoxina oxidada que actúa como oxidasa en el compartimiento intracelular.

Algunos ejemplos preferidos son microorganismos procariotas que tienen actividad reductasa ausente o reducida. Por ejemplo, la actividad de una tiorredoxina reductasa y/o una glutatión oxidoreductasa puede reducirse o eliminarse para modificar el entorno intracelular, produciendo de ese modo un entorno intracelular oxidante que favorece la formación de enlaces disulfuro.

Por ejemplo, cepas de *E. coli* que tienen mutaciones tanto en el gen de tiorredoxina reductasa (*trxB*) como el gen de glutatión oxidoreductasa (*gor*) pueden expresar proteínas con niveles superiores de formación de enlaces disulfuro. Véase, por ejemplo, Prinz *et al.*, J. Biol Chem. 272:15661-15667 (1997). Estos mutantes dobles *trxB gor* crecen muy lentamente en la mayoría de medios de cultivo, aunque el crecimiento puede potenciarse mediante la adición de un agente reductor, tal como DTT. Sin embargo, las cepas mutantes dobles frecuentemente dan lugar a cepas mutantes supresoras que conservan las mutaciones *trxB gor* y que crecen más rápido en medio que carece de DTT. Un ejemplo de una mutación supresora *trxB gor* en *E. coli* es una mutación del gen *ahpC*, que codifica para una subunidad catalítica de la alquil hidroperoxidasa, AhpCF. Esta mutación supresora añade un triplete al ADN que codifica para el sitio catalítico de la enzima AhpCF. Se dan a conocer cepas de *E. coli* mutantes dobles de crecimiento rápido, por ejemplo, las cepas *trxB, gor, supp* y *trxB, gshA, supp* en, por ejemplo, la patente estadounidense n.º 6.872.563, que se incorpora en el presente documento como referencia para todos los fines. Están disponibles comercialmente cepas de *E. coli* así manipuladas, por ejemplo, las cepas *trxB, gor, supp*, por ejemplo, con los nombres comerciales ORIGAMI TM, ORIGAMI 2 TM y ROSETTA-GAMI TM, de por ejemplo, EMD Biosciences, Inc. Otras mutaciones de *E. coli* pueden dar como resultado un entorno intracelular oxidante, por ejemplo, las cepas *trxB, gshA* y *trxB, gshA supp*.

Pueden usarse otras manipulaciones de componentes de una ruta redox en un microorganismo para potenciar la formación de enlaces disulfuro en una proteína, por ejemplo, una glicosiltransferasa eucariota. Por ejemplo, pueden sobreexpresarse proteínas con actividad oxidante, por ejemplo, proteínas tiorredoxinas de *E. coli* en cepas mutantes *trxB, gor*, en el microorganismo procariota. Otro ejemplo es la expresión o sobreexpresión de mutantes de tiorredoxina que tienen actividad oxidante potenciada. Ejemplos de tales mutantes se describen en, por ejemplo, Bessette, *et al.* PNAS 96:13703-13708 (1999). La expresión citoplasmática dirigida de ciertas enzimas oxidantes también puede usarse para potenciar la formación de enlaces disulfuro intracelulares. Por ejemplo, proteínas oxidantes que se expresan normalmente en el espacio periplasmático, por ejemplo, DsbC, pueden expresarse en un citoplasma bacteriano, por ejemplo, delecionando una secuencia de direccionamiento periplasmático o incluyendo una secuencia de retención citoplasmática. Otras proteínas periplasmáticas oxidantes pueden expresarse en el citoplasma bacteriano para potenciar la oxidación de proteínas citoplasmáticas, por ejemplo, delecionando una secuencia de direccionamiento periplasmático o incluyendo una secuencia de retención citoplasmática.

Pueden identificarse ácidos nucleicos de tiorredoxina reductasa, ácidos nucleicos de glutatión oxidoreductasa, ácidos nucleicos de tiorredoxina, ácidos nucleicos de glutatión y ácidos nucleicos que codifican para otras proteínas implicadas en el mantenimiento de un entorno redox intracelular en otras bacterias, 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 muchas otras. Tales genes pueden identificarse mediante análisis y comparación de secuencias con genes de tiorredoxina reductasa, genes de glutatión oxidoreductasa y genes conocidos que codifican para otras proteínas implicadas en el mantenimiento de un entorno redox intracelular o con la secuencia de aminoácidos de los productos codificados. Las proteínas codificadas pueden identificarse además funcionalmente mediante ensayos enzimáticos o mediante ensayos de complementación genética de mutantes de *E. coli* de una función génica apropiada. Los genes de tiorredoxina reductasa y glutatión oxidoreductasa endógenos pueden mutarse, por ejemplo, para inactivar el producto génico usando técnicas de biología molecular convencionales y esas cepas mutadas también pueden usarse para expresar proteínas con niveles aumentados de formación de enlaces disulfuro, en comparación con cepas no mutadas.

IV. EXPRESIÓN DE GLICOSILTRANSFERASAS EUCARIOTAS ACTIVAS, SOLUBLES EN MICROORGANISMOS PROCARIOTAS QUE TIENEN ENTORNOS OXIDANTES

Los polipéptidos de glicosiltransferasa eucariota activa, soluble de la invención pueden expresarse en una variedad de microorganismos procariotas con entornos intracelulares oxidantes, incluyendo *E. coli* y otros huéspedes bacterianos, tal como se describió anteriormente.

Una vez expresados en un organismo procariota que tiene un entorno intracelular oxidante, los polipéptidos de glicosiltransferasa eucariota activa, soluble pueden usarse para producir productos glicosilados. Por ejemplo, los polipéptidos de glicosiltransferasa eucariota activa, soluble pueden aislarse usando técnicas de purificación de proteínas convencionales y usarse en reacciones *in vitro* descritas en el presente documento para preparar productos glicosilados. También pueden usarse polipéptidos de glicosiltransferasa eucariota activa, soluble

parcialmente purificados en reacciones *in vitro* para preparar productos glicosilados, tal como pueden los organismos procariotas permeabilizados que expresan los polipéptidos de glicosiltransferasa eucariota activa, soluble.

Normalmente, el polinucleótido que codifica para el polipéptido de glicosiltransferasa eucariota se coloca bajo el control de un promotor que es funcional en los organismos procariotas deseados que tienen un entorno oxidante. Se conoce una variedad extremadamente amplia 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 la que el promotor va a ser activo. Otras secuencias de control de la expresión tales como sitios de unión a ribosomas, sitios de terminación de la transcripción y similares también se incluyen opcionalmente. Los constructos que incluyen uno o más de estas secuencias de control se denominan "casetes de expresión". Por consiguiente, la invención proporciona casetes de expresión dentro de los que se incorporan los ácidos nucleicos que codifican para las proteínas de fusión para la expresión a alto nivel en un microorganismo deseado que tiene un entorno oxidante.

Las secuencias de control de la expresión que son adecuadas para su uso en una célula huésped particular se obtienen a menudo clonando un gen que se expresa en esa célula. Las secuencias de control procariotas comúnmente usadas, que se define en el presente documento que incluyen promotores para la iniciación de la transcripción, opcionalmente con un operador, junto con secuencias de sitio de unión a ribosomas, incluyen promotores comúnmente usados tales como los sistemas promotor de lactosa (*lac*) y beta-lactamasa (penicilinas) (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 sitio de unión a 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 organismos procariotas.

Para la expresión de polipéptidos de glicosiltransferasa eucariota activa, soluble en células procariotas distintas de *E. coli*, se requiere un promotor que funcione en la especie procariota particular. Tales promotores pueden obtenerse de genes que se han clonado de la especie, o pueden usarse promotores heterólogos. Por ejemplo, el promotor *trp-lac* híbrido funciona en *Bacillus* además de *E. coli*. Se conocen promotores para otras especies bacterianas, por ejemplo *Pseudomonas*. Véase, por ejemplo, la publicación de solicitud de patente estadounidense n.º US 2005/0186666, publicada el 25 de agosto de 2005.

Se incluye convenientemente un sitio de unión a ribosomas (RBS) 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 ubicada 3-11 nucleótidos hacia el sentido de 5' del codón de iniciación (Shine y Dalgarno, Nature (1975) 254: 34; Steitz, en Biological regulation and development: Gene Expression (ed. R.F. Goldberger), vol. 1, p. 349, 1979, Plenum Publishing, NY).

Pueden usarse promotores o bien constitutivos o bien regulados en la presente invención. Los promotores regulados pueden ser ventajosos debido a que las células huésped pueden hacerse crecer en altas densidades antes de inducir la expresión de las proteínas de fusión. La expresión de alto nivel de proteínas heterólogas ralentiza el crecimiento celular en algunas situaciones y puede no desearse en todas las situaciones, véase a continuación. Un promotor inducible es un promotor que dirige la expresión de un gen cuando el nivel de expresión puede alterarse mediante factores del desarrollo o medioambientales tales como, por ejemplo, temperatura, pH, condiciones aerobias o anaerobias, luz, factores de transcripción y químicos. Tales promotores se denominan en el presente documento promotores "inducibles", que permiten controlar el momento de expresión de la glicosiltransferasa o enzima implicada en la síntesis del azúcar de nucleótido. Los expertos en la técnica conocen promotores inducibles para *E. coli* y otras células huésped bacterianas. Éstas incluyen, por ejemplo, el promotor *lac*, el promotor P_L del bacteriófago lambda, el promotor *trp-lac* híbrido (Amann *et al.* (1983) Gene 25: 167; de Boer *et al.* (1983) Proc. Nat'l. Acad. Sci. USA 80: 21) y el promotor del bacteriófago T7 (Studier *et al.* (1986) J. Mol. Biol.; Tabor *et al.* (1985) Proc. Nat'l. Acad. Sci. USA 82: 1074-8). Estos promotores y su uso se discuten en Sambrook *et al.*, citado anteriormente. Un promotor inducible particularmente preferido para la expresión en organismos procariotas es un promotor doble que incluye un componente de promotor *tac* ligado a un componente de promotor obtenido 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 doble *tac-gal*, que se describe en la publicación de solicitud de patente PCT n.º WO98/20111.

Otro promotor inducible es el promotor *cspUn*, que se induce altamente a bajas temperaturas en *E. coli*. Véase, por ejemplo, Sorensen y Mortensen, BioMed Central, www.microbialcellfactories.com/content/4/1/1 y Mujacic *et al.* Gene 238:325-332 (1999).

Un constructo que incluye un polinucleótido de interés operativamente unido a señales de control de la expresión génica que, cuando se coloca en una célula huésped apropiada, dirige la expresión del polinucleótido se denomina "casete de expresión". Se colocan a menudo casetes de expresión que codifican para las proteínas de fusión de la invención en vectores de expresión para su introducción en la célula huésped. Los vectores normalmente incluyen, además de un casete de expresión, una secuencia de ácido nucleico que permite al vector replicarse independientemente en una o más células huésped seleccionadas. Generalmente, esta secuencia es una que

permite al vector replicarse independientemente del ADN cromosómico del huésped, e incluye orígenes de replicación o secuencias de replicación autónoma. Tales secuencias se conocen bien para una variedad de bacterias. Por ejemplo, el origen de replicación del plásmido pBR322 es adecuado para la mayoría de las bacterias Gram-negativas. Alternativamente, el vector puede replicarse integrándose en el complemento genómico de la célula huésped y replicándose a medida que la célula experimenta replicación del ADN. Un vector de expresión preferido para la expresión de las enzimas en células bacterianas es pTGK, que incluye un promotor *tac-gal* doble y se describe en la publicación de solicitud de patente PCT n.º WO98/20111. Otro vector clonación útil es pCWin2-MBP o una versión de pCWin2 con una UTR en 5' modificada. Véase, por ejemplo, el documento PCT/USOS/00302, presentado el 6 de enero de 2005.

La construcción de constructos de polinucleótidos requiere generalmente el uso de vectores que pueden replicarse en bacterias. Está disponibles comercialmente una plétora de kits para la purificación de plásmidos a partir de bacterias (véanse, por ejemplo, EasyPrepJ, FlexiPrepJ, ambos de Pharmacia Biotech; StrataCleanJ, de Stratagene; y QIAexpress Expression System, Qiagen). Los plásmidos aislados y purificados pueden manipularse entonces adicionalmente para producir otros plásmidos, y usarse para transfectar células. Es posible la clonación en, por ejemplo, *E. coli*, *Streptomyces* o *Bacillus*.

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 las 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, cloramfenicol o tetraciclina. Alternativamente, los marcadores seleccionables pueden codificar para proteínas que complementen deficiencias auxotróficas o suministren nutrientes críticos no disponibles 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. Se conoce un sistema de expresión auxotrófo para especies de *Pseudomonas*. Véase, por ejemplo, la publicación de solicitud de patente estadounidense n.º US 2005/0186666, publicada el 25 de agosto de 2005.

La construcción de vectores adecuados que contienen uno o más de los componentes enumerados anteriormente emplea técnicas de ligamiento convencionales tal como se describen en las referencias citadas anteriormente. Los fragmentos de ADN o plásmidos aislados se escinden, adaptan y vuelven a ligarse en la forma deseada para generar los plásmidos requeridos. Para confirmar las secuencias correctas en plásmidos construidos, los plásmidos pueden analizarse mediante técnicas convencionales tal como mediante digestión con endonucleasas de restricción, y/o análisis de secuencia según métodos conocidos. Las técnicas de clonación molecular para lograr estos fines se conocen en la técnica. Los expertos conocen bien una amplia variedad de métodos de amplificación *in vitro* y clonación 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 and Sons, Inc., (suplemento de 1998) (Ausubel).

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 pBLTSCRIPT™, y vectores derivados de fago λ.

Los métodos para introducir los vectores de expresión en un microorganismo procariota elegido no son particularmente críticos, y los expertos en la técnica conocen tales métodos. 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. Otros métodos de transformación también son adecuados.

Puede usarse acoplamiento de la traducción para potenciar la expresión. La estrategia usa un marco de lectura abierto en el sentido de 5' corto derivado de un gen nativo altamente expresado para el sistema de traducción, que se coloca en el sentido 3' del promotor, y un sitio de unión a ribosomas seguido tras algunos codones de aminoácidos por un codón de terminación. Justo antes del codón de terminación está un segundo sitio de unión a ribosomas, y tras el codón de terminación está un codón de iniciación para la iniciación de la traducción. El sistema disuelve la estructura secundaria del ARN, permitiendo la iniciación eficaz de la traducción. Véase Squires, *et al.* (1988), *J. Biol. Chem.* 263: 16297-16302.

Los polipéptidos de glicosiltransferasa eucariota activa, soluble se expresan preferiblemente de manera intracelular. La expresión intracelular a menudo da como resultado rendimientos sorprendentemente altos. Si es necesario, la cantidad de proteína activa, soluble puede aumentarse realizando procedimientos de replegamiento (véase, por

ejemplo, Sambrook *et al.*, citado anteriormente.; Marston *et al.*, Bio/Technology (1984) 2: 800; Schoner *et al.*, Bio/Technology (1985) 3: 151). En otra realización, las proteínas de glicosiltransferasa eucariota activa, soluble se fusionan con una subsecuencia de proteína A, una proteína de unión a maltosa, una proteína de unión a almidón o albúmina sérica bovina (BSA), por ejemplo, para facilitar la purificación, la secreción o la estabilidad.

Los polipéptidos de glicosiltransferasa eucariota activa, soluble de la invención también pueden unirse además a otras proteínas bacterianas. Este enfoque a menudo da como resultado altos rendimientos, debido a que secuencias de control procariotas normales dirigen la transcripción y traducción. En *E. coli*, se usan a menudo fusiones de *lacZ* para expresar proteínas heterólogas. Otros ejemplos se discuten a continuación. Están disponibles fácilmente vectores adecuados, tales como las series pUR, pEX y pMR100 (véase, por ejemplo, Sambrook *et al.*, citado anteriormente). Para ciertas aplicaciones, puede ser deseable escindir los aminoácidos distintos de glicosiltransferasa de la proteína de fusión tras la 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). Pueden diseñarse por ingeniería genética sitios de escisión en el gen para la proteína de fusión en el punto de escisión deseado.

Puede expresarse más de una proteína recombinante en una única célula huésped colocando múltiples casetes transcripcionales 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. Por ejemplo, pueden expresarse múltiples glicosiltransferasas en una única célula, por ejemplo, glicosiltransferasas que dirigen la glicosilación N-unida o glicosiltransferasas que dirigen la glicosilación O-unida.

Se ha descrito un sistema adecuado para obtener proteínas recombinantes a partir de *E. coli* que mantiene la integridad de sus extremos N-terminales por 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 levaduras 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.

V. PURIFICACIÓN DE GLICOSILTRANSFERASAS EUCARIOTAS ACTIVAS, SOLUBLES

Los polipéptidos de glicosiltransferasa eucariota activa, soluble de la presente invención se expresan preferiblemente como proteínas intracelulares y pueden usarse en esta forma, en los métodos de la presente invención. Por ejemplo, pueden usarse células permeabilizables o un extracto celular bruto que contiene el polipéptido de glicosiltransferasa eucariota activa, soluble intracelular expresado en los métodos de la presente invención.

Alternativamente, el polipéptido de glicosiltransferasa eucariota activa, soluble puede purificarse según procedimientos convencionales de la técnica, incluyendo precipitación con sulfato de amonio, columnas de afinidad, cromatografía en columna, electroforesis en gel y similares (véase, 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)). Se prefieren composiciones sustancialmente puras de al menos aproximadamente el 70, el 75, el 80, el 85, el 90% de homogeneidad, y lo más preferido es el 92, el 95, el 98 hasta el 99% o más de homogeneidad. Las proteínas purificadas también pueden usarse, por ejemplo, como inmunógenos para la producción de anticuerpos.

Para facilitar la purificación y la expresión de los polipéptidos de glicosiltransferasa eucariota activa, soluble de la invención, los ácidos nucleicos que codifican para las proteínas también pueden incluir una secuencia codificante para un epítipo o "etiqueta" para la que está disponible un reactivo de unión por afinidad, es decir, una etiqueta de purificación. Los ejemplos de epítopos adecuados incluyen los genes indicadores V-5 y myc; están disponibles comercialmente vectores de expresión útiles para la producción recombinante de proteínas de fusión que tienen estos epítopos (por ejemplo, los vectores de Invitrogen (Carlsbad CA) pcDNA3.1/Myc-I-His y pcDNA3.1/V5-His son adecuados para la expresión en células de mamífero). Los expertos en la técnica conocen vectores de expresión adicionales adecuados para unir una etiqueta a las glicosiltransferasas de la invención, y los correspondientes sistemas de detección, y varios están disponibles comercialmente (por ejemplo, FLAG" (Kodak, Rochester NY). Otro ejemplo de una etiqueta 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:204), aunque pueden usarse más o menos de seis. Los ligandos de afinidad de quelatos metálicos adecuados que pueden servir como resto de unión para una etiqueta de polihistidina incluyen ácido nitrilo-tri-acético (NTA) (Hochuli, E. (1990) "Purification of recombinant proteins with metal chelating adsorbents" en Genetic Engineering: Principles and Methods, J.K. Setlow, Ed., Plenum Press, NY; disponible comercialmente de Qiagen (Santa Clarita, CA)). Otras etiquetas de purificación o de epítipo incluyen, por ejemplo, AU1, AU5, DDDDK (SEQ ID NO:202) (EC5), etiqueta E2, Glu-Glu, un péptido de 6 residuos, EYMPME (SEQ ID NO:203), derivado de la proteína media T de Polioma, HA, HSV, IRS, KT3, etiqueta S, etiqueta S1, etiqueta T7, etiqueta V5, VSV-G, β -galactosidasa, Gal4, proteína fluorescente verde (GFP), luciferasa, proteína C, proteína A, proteína de unión a celulosa, GST (glutathión S-transferasa), una etiqueta step, Nus-S, PPI-asas, Pfg 27, proteína de unión a calmodulina, dsb A y fragmentos de los mismos, y granzima B. Están

disponibles comercialmente anticuerpos y péptidos de epítipo que se unen específicamente a secuencias de epítipo de, por ejemplo, Covance Research Products, Inc.; Bethyl Laboratories, Inc.; Abcam Ltd.; y Novus Biologicals, Inc.

Las etiquetas de purificación también incluyen dominios de unión a maltosa y dominios de unión a almidón. Las proteínas que comprenden etiquetas de purificación pueden purificarse usando una pareja de unión que se une a la etiqueta de purificación, por ejemplo, anticuerpos frente a la etiqueta de purificación, iones níquel o cobalto o resinas, y amilosa, maltosa o una ciclodextrina. Las etiquetas de purificación también incluyen dominios de unión a almidón, dominios de tioredoxina de *E. coli* (vectores y anticuerpos disponibles comercialmente de por ejemplo, Santa Cruz Biotechnology, Inc. y Alpha Diagnostic International, Inc.), y la mitad carboxilo terminal de la proteína SUMO (vectores y anticuerpos disponibles comercialmente de por ejemplo, Life Sensors Inc.). Los dominios de unión a almidón, tales como un dominio de unión a maltosa de *E. coli* y SBD (dominio de unión a almidón) de una amilasa de *A. niger*, se describen en el documento WO 99/15636, incorporado en el 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 se describe en el documento WO 2005/014779, publicado el 17 de febrero de 2005, incorporado en el presente documento como referencia en su totalidad. En algunas realizaciones, un polipéptido de glicosiltransferasa eucariota activa, soluble comprende más de una etiqueta de purificación o de epítipo.

Los expertos en la técnica conocen otros haptenos que son adecuados para su uso como etiquetas y se describen, por ejemplo, en el Handbook of Fluorescent Probes and Research Chemicals (6ª Ed., Molecular Probes, Inc., Eugene OR). Por ejemplo, dinitrofenol (DNP), digoxigenina, barbituratos (véase, por ejemplo, la patente estadounidense n.º 5.414.085), y varios tipos de fluoróforos son útiles como haptenos, puesto que son derivados de estos compuestos. Están disponibles comercialmente kits para unir 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 etiqueta a los residuos de lisina presentes en el reactivo de captura.

Un experto reconocería que pueden realizarse modificaciones a los dominios funcional o catalítico del polipéptido de glicosiltransferasa eucariota activa, soluble sin disminuir su función biológica. Algunas modificaciones pueden hacerse para facilitar la clonación, la expresión o la 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) colocados en cualquier extremo terminal para crear secuencias de purificación o de codones de terminación o sitios de enzimas de restricción ubicados convenientemente ubicados.

En realizaciones preferidas, la purificación de las glicosiltransferasas eucariotas se simplifica mediante la expresión de las proteínas en microorganismos que tienen entornos oxidantes. Debido a que la solubilidad de las proteínas expresadas está potenciada, pueden omitirse etapas de purificación que requieren mucho tiempo, tales como solubilización, desnaturalización y repliegamiento, de un protocolo de purificación.

Las glicosiltransferasas eucariotas producidas mediante los métodos de la invención pueden usarse para producir productos peptídicos glicosilados y proteína glicosilada. Los productos peptídicos glicosilados y la proteína glicosilada también pueden purificarse, si lo desea el usuario, con cualquiera de los métodos de purificación de proteínas descritos en el presente documento.

VI. POTENCIACIÓN Y CARACTERIZACIÓN DE LA SOLUBILIZACIÓN DE PROTEÍNAS

La reducción de los enlaces disulfuros en proteínas expresadas de manera heteróloga, tales como los polipéptidos de glicosiltransferasa eucariota usados en los métodos de la invención, frecuentemente da como resultado un mal plegamiento de la proteína y precipitación de la solución. En células bacterianas tales como por ejemplo, *E. coli*, las proteínas mal plegadas se expresan como cuerpos de inclusión insolubles. La solubilización de una proteína está generalmente indicada por la presencia de la proteína en una fracción acuosa tras la centrifugación a una velocidad apropiada durante un periodo apropiado. Además, la expresión de las proteínas plegadas apropiadamente da como resultado niveles aumentados de actividad de la proteína. Por tanto, también pueden usarse ensayos de actividad enzimática para determinar si se ha producido un plegamiento apropiado de la proteína.

La solubilización de un polipéptido de glicosiltransferasa eucariota expresado en un microorganismo con un entorno oxidante puede compararse con la solubilización de un polipéptido de glicosiltransferasa eucariota expresado en un microorganismo con un entorno reductor, por ejemplo, una cepa de *E. coli* con un entorno reductor. En algunas realizaciones, un polipéptido de glicosiltransferasa eucariota expresado en un microorganismo con un entorno oxidante se expresa en una fracción soluble a niveles que son hasta 1,1, 1,2, 1,5, 2, 3, 5, 10, 15, 20, 50, 100, 500, 1000 o hasta 10.000 veces superiores a los niveles solubles del mismo polipéptido de glicosiltransferasa eucariota expresado en un microorganismo con un entorno reductor. En otras realizaciones, un polipéptido de glicosiltransferasa eucariota expresado en un microorganismo con un entorno oxidante tiene niveles de actividad, por ejemplo, U/célula o U/mg de proteína, hasta 1,1, 1,2, 1,5, 2, 3, 5, 10, 15, 20, 50, 100, 500, 1000 o hasta 10.000 veces superiores a los niveles de actividad del mismo polipéptido de glicosiltransferasa eucariota expresado en un

microorganismo con un entorno reductor.

A. Caracterización de la solubilidad de la proteína

En realizaciones preferidas, las glicosiltransferasas eucariotas se expresan como proteínas solubles intracelularmente dentro de un microorganismo procariota. La solubilidad de los polipéptidos de glicosiltransferasa eucariota puede determinarse tal como se dio a conocer anteriormente, determinando los niveles de proteína en una fracción acuosa tras la centrifugación a una velocidad apropiada durante un periodo apropiado. Los niveles de proteína pueden determinarse usando métodos conocidos por los expertos en la técnica, por ejemplo, inmunoensayos o comparación directa de proteínas separadas mediante, por ejemplo, SDS-PAGE. Pueden realizarse inmunoensayos usando anticuerpos específicos para el polipéptido de glicosiltransferasa eucariota de interés o usando anticuerpos específicos para una etiqueta de purificación o de epítipo que está covalentemente unida al polipéptido de glicosiltransferasa eucariota.

La solubilidad también puede determinarse sometiendo a ensayo la actividad enzimática de los polipéptidos de glicosiltransferasa eucariota en una fracción soluble a partir de un microorganismo procariota. En una realización preferida, la actividad glicosiltransferasa puede medirse en una fracción intracelular soluble a partir de un microorganismo procariota.

Los polipéptidos de glicosiltransferasa eucariota pueden usarse para preparar productos glicosilados en mezclas de reacción *in vitro*, incluyendo, por ejemplo, oligosacáridos, glicolípidos, glicoproteínas y glicopéptidos. Las mezclas de reacción *in vitro* pueden incluir microorganismos permeabilizados que comprenden los polipéptidos de glicosiltransferasa eucariota, polipéptidos de glicosiltransferasa eucariota parcialmente purificados o polipéptidos de glicosiltransferasa eucariota purificados; así como sustratos donadores, sustratos aceptores y tampones de reacción apropiados. Para reacciones *in vitro*, los polipéptidos de glicosiltransferasa eucariota, sustratos aceptores, sustratos donadores y otros componentes de la mezcla de reacción se combinan mediante mezcla en un medio de reacción acuoso. El medio tiene generalmente un valor de pH de aproximadamente 4,0 a aproximadamente 9,0. La selección de un medio está basada en la capacidad del medio para mantener el valor de pH al nivel deseado. Por tanto, en algunas realizaciones, el medio se tampona a un valor de pH de aproximadamente 7,5. Si no se usa tampón, el pH del medio debe mantenerse a de aproximadamente 5 a 8,5, dependiendo de la glicosiltransferasa particular usada. Para sialiltransferasas, por ejemplo, el intervalo es preferiblemente de desde aproximadamente 5,5 hasta aproximadamente 8,0.

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

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

La temperatura a la que se lleva a cabo el proceso anterior puede oscilar entre justo por encima de la temperatura de congelación hasta la temperatura en la que la enzima más sensible se desnaturaliza. Ese intervalo de temperatura es preferiblemente de aproximadamente 0°C a aproximadamente 45°C, y más preferiblemente a 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 determinantes de oligosacáridos deseados presentes en los grupos de oligosacárido unidos a la glicoproteína que va a glicosilarse. Para preparaciones a escala comercial, se dejará que la reacción se produzca a menudo durante entre aproximadamente 0,5-240 horas, y más normalmente entre aproximadamente 1-36 horas.

B. Potenciación de la solubilidad de la proteína

La potenciación adicional de la solubilidad de polipéptidos de glicosiltransferasa eucariota puede producirse, por ejemplo, reduciendo la tasa de expresión de proteínas o expresando la proteína en combinación con, por ejemplo una proteína chaperona.

La potenciación de la tasa de formación de enlaces disulfuro apropiados puede conducir a expresión superior de glicosiltransferasas solubles activas. Otro método para potenciar la expresión de glicosiltransferasas solubles activas es para reducir la tasa de expresión permitiendo de ese modo al polipéptido naciente más tiempo para lograr una conformación soluble, estable. La combinación de los dos métodos, tal como se describe en el presente documento, es una realización preferida de la invención. La expresión máxima de una proteína heteróloga se produce generalmente en condición de crecimiento óptimo para las células huésped. Un método para ralentizar la expresión

de proteínas es ralentizar la tasa de crecimiento de las células. En una realización preferida, se hacen crecer células huésped a una temperatura inferior a su temperatura de crecimiento óptima. Por ejemplo, la temperatura de crecimiento óptima de *E. coli* es de 37°C. Por tanto, una temperatura inferior a la temperatura de crecimiento óptima para *E. coli* es inferior a 37°C, por ejemplo, entre 4°C y 36°C, entre 8°C y 33°C, entre 12°C y 30°C, o entre 18°C y 26°C, o a aproximadamente 20°C, o a aproximadamente 24°C.

La temperatura usada para ralentizar la producción de proteína dependerá de la temperatura de crecimiento óptima de las células huésped. Como un ejemplo, *E. coli* y muchas otras bacterias tienen una temperatura de crecimiento óptima de 37°C. Por tanto, una temperatura inferior a una temperatura de crecimiento óptima para *E. coli* o para otras bacterias que crecen de manera óptima a 37°C podría ser de entre 4-35°C, entre 12-30°C, o entre 15-20°C. En una realización preferida la temperatura inferior a una temperatura de crecimiento óptima para *E. coli* o para otras bacterias que crecen de manera óptima a 37°C es de entre 18 y 23°C. Para células que crecen de manera óptima a 30°C, como lo hacen muchas levaduras, una temperatura inferior a una temperatura de crecimiento óptima podría ser de entre 10 y 25°C, entre 12 y 21°C, o entre 15 y 20°C.

Otro método para reducir la tasa de expresión de una proteína heteróloga es variar la concentración de una molécula que regula la expresión de un promotor inducible. Por ejemplo, algunas mutaciones *lacY* permiten controlar la expresión de la proteína variando la cantidad de IPTG, la molécula inductora, en el medio. En realizaciones preferidas la concentración de IPTG en el medio es inferior a la óptima para, por ejemplo, la expresión de una proteína que no forma cuerpos de inclusión cuando se sobreexpresa en un microorganismo procariota.

En algunas realizaciones, un polipéptido de glicosiltransferasa eucariota se expresa en un microorganismo que tiene un entorno oxidante y que comprende además una proteína chaperona heteróloga. Las proteínas chaperonas incluyen, por ejemplo, factor desencadenante; miembros de la familia de chaperona Hsp70, por ejemplo DnaK; miembros de la familia de chaperona Hsp100, por ejemplo ClpB, y miembros de la familia de chaperona Hsp60, por ejemplo GroEL. Véase, por ejemplo, Sorensen y Mortensen, BioMed Central, www.microbialcellfactories.com/content/4/1/1. También se conocen chaperonas que permiten el plegamiento de la proteína a 4°C, por ejemplo, Cpn60 y Cpn10 de *Oleispira antarctica* RB8^T. Véase, por ejemplo, Id. y Ferrer *et al.*, Nat. Biotechnol. 21:1266-1267 (2003). Las proteínas chaperonas a modo de ejemplo incluyen, pero no se limitan a, aquellas indicadas en la lista de secuencias adjunta.

En otras realizaciones, un polipéptido de glicosiltransferasa eucariota se expresa en un microorganismo que tiene un entorno oxidante que comprende además una proteína disulfuro isomerasa (PDI) heteróloga. Las proteínas de PDI pueden hacer o reorganizar enlaces disulfuro. Las proteínas de PDI se describen por ejemplo, en Georgiou *et al.* patente estadounidense n.º 6.027.888, que se incorpora en el presente documento como referencia para todos los fines. Las proteínas de PDI incluyen por ejemplo, PDI de hígado de rata, proteínas Ero1p y Pdi1p de *Saccharomyces*. Las proteínas procariotas incluyen por ejemplo, DsbC de *E. coli*. Véase, por ejemplo, Frand *et al.*, Trends in Cell Biol. 10: 203-210 (2000). En algunas realizaciones, se expresan DsbC en un citoplasma bacteriano por ejemplo, delecionando una secuencia de direccionamiento periplasmático o incluyendo una secuencia de retención citoplasmática.

Otras proteínas procariotas que actúan para mantener el estado redox de enlaces disulfuro de proteínas incluyen, por ejemplo, DsbB, DsbA, DsbC, DsbD y DsbG de *E. coli*. Estas proteínas se conocen bien en la técnica y se describen en, por ejemplo, Beckwith *et al.* patente estadounidense n.º 6.872.563, que se incorpora en el presente documento como referencia para todos los fines. En algunas realizaciones, DsbB, DsbA, DsbC, DsbD y DsbG se expresan en el citoplasma bacteriano para potenciar la oxidación de proteínas citoplasmáticas, por ejemplo, delecionando una secuencia de direccionamiento periplasmático o incluyendo una secuencia de retención citoplasmática.

En una realización adicional, un polipéptido de glicosiltransferasa eucariota se expresa en un microorganismo procariota que tiene un entorno oxidante y que también comprende una proteína chaperona heteróloga y/o una proteína de PDI heteróloga y/o una proteína tal como DsbB, DsbA, DsbC, DsbD y DsbG de *E. coli*.

C. Expresión de glicosiltransferasas eucariotas truncadas activas en microorganismos procariotas que tienen entornos reductores

Inesperadamente, algunas de las glicosiltransferasas eucariotas truncadas tenían actividad en organismos procariotas con entornos intracelulares reductores. La actividad en un entorno intracelular reductor era normalmente mucho menor que la actividad de la misma proteína en un organismo procariota que tiene un entorno intracelular oxidante. En algunas realizaciones las glicosiltransferasas eucariotas truncadas están fusionadas a, por ejemplo, un dominio de unión a maltosa o un dominio de unión a almidón, o combinaciones de los mismos.

Por ejemplo, la tabla 4 muestra que una proteína de GnT1 truncada del tipo natural fusionada a MBP tenía 1,7 U/litro de actividad. Por el contrario, la misma proteína tenía 45 U/litro de actividad cuando se expresaba en un mutante *trxB*, *gor*, *supp* de *E. coli*. Otras glicosiltransferasas sometidas a prueba parecían no tener actividad en el entorno reductor de las células JM109. Por tanto, en una realización, la invención incluye métodos para la producción de una glicosiltransferasa activa, soluble truncada en un microorganismo que tiene un entorno reductor, por ejemplo, *E. coli*.

En realizaciones preferidas, la glicosiltransferasa es GnT1.

VII. USOS DE GLICOSILTRANSFERASAS EUCARIOTAS SOLUBLES PRODUCIDAS POR MICROORGANISMOS PROCARIOTAS CON ENTORNOS OXIDANTES

La invención proporciona métodos para la producción de glicosiltransferasas eucariotas solubles en microorganismos procariotas, preferiblemente en una escala comercial. Las glicosiltransferasas eucariotas solubles se usan entonces para sintetizar enzimáticamente restos de oligosacárido, glicolípidos y glicoproteínas, y para glicoPEGilar glicoproteínas o glicopéptidos, incluyendo proteínas terapéuticas, también preferiblemente en una escala comercial. Las reacciones enzimáticas de la invención tienen lugar en un medio de reacción que comprende al menos una glicosiltransferasa eucariota, soluble, sustrato aceptor y sustrato donador, y normalmente un catión metálico divalente soluble. En algunas realizaciones, las enzimas auxiliares y sustratos para el resto catalítico de la enzima auxiliar también están presentes, de modo que las enzimas auxiliares pueden sintetizar el sustrato donador para la glicosiltransferasa. Las proteínas de glicosiltransferasa eucariota soluble catalizan la adición de un sacárido a un sustrato aceptor, por ejemplo, una proteína terapéutica activa soluble.

Se conocen varios métodos para usar glicosiltransferasas para sintetizar glicoproteínas y glicolípidos que tienen restos del oligosacárido deseado. Se describen métodos a modo de ejemplo, por ejemplo, en 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.

Las proteínas de glicosiltransferasa eucariota soluble preparadas tal como se describen en el presente documento pueden usarse en combinación con glicosiltransferasas adicionales. Por ejemplo, puede usarse una combinación de proteína de glicosiltransferasa eucariota soluble y una glicosiltransferasa bacteriana. De manera similar, las proteínas de glicosiltransferasa eucariota soluble pueden usarse con enzimas auxiliares recombinantes, que pueden o no fusionarse a proteínas de glicosiltransferasa eucariota soluble.

Los productos producidos mediante los procesos anteriores pueden usarse con o sin purificación adicional. 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 a través de membrana, para la recuperación de sacáridos glicosilados. También puede usarse, por ejemplo, la filtración a través de membrana, utilizando una membrana osmótica inversa o de nanofiltración tal como se describe en la patente australiana de titularidad compartida n.º 735695. Como un ejemplo adicional, puede usarse la filtración a través de membrana en la que las membranas tienen un punto de corte de peso molecular de aproximadamente 1000 a aproximadamente 10.000 para retirar proteínas. Como otro ejemplo, puede usarse entonces nanofiltración u ósmosis inversa para retirar sales. Las membranas de nanofiltro son una clase de membranas de ósmosis inversa que pasan sales monovalentes pero retienen sales polivalentes y solutos sin carga de más de aproximadamente 200 a aproximadamente 1000 Daltons, dependiendo de la membrana usada. Por tanto, por ejemplo, los oligosacáridos producidos mediante las composiciones y métodos de la presente invención pueden retenerse en la membrana y las sales contaminantes pasarán a través de ella.

Los productos que son por ejemplo, proteínas glicosiladas o péptidos glicosilados también pueden usarse sin purificación adicional. O si lo desea el usuario, pueden aislarse o purificarse proteínas glicosiladas o péptidos glicosilados usando métodos de purificación de proteínas convencionales.

VIII. SUSTRATO DONADOR/SUSTRATOS ACEPTORES

Sustratos donadores adecuados usados por las glicosiltransferasas eucariotas activas, solubles y métodos de la invención incluyen, pero no se limitan a, UDP-Glc, UDP-GlcNAc, UDP-Gal, UDP-GalNAc, GDP-Man, GDP-Fuc, UDPGlcUA, UDP-GlcNH₂, UDP-GalNH₂ y CMP-ácido siálico. Guo *et al.*, Applied Biochem. and Biotech. 68: 1-20 (1997).

Los sustratos aceptores adecuados usados por las proteínas de glicosiltransferasa eucariota, soluble y 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 varias estructuras de hidratos de carbono, glicoproteínas y glicolípidos conocidas por los expertos en la técnica.

Ejemplos de sustratos aceptores adecuados usados en reacciones catalizadas por proteína de glicosiltransferasa eucariota, soluble se describen en Guo *et al.*, Applied Biochem. and Biotech. 68: 1-20 (1997), pero no se limitan a ello.

En realizaciones preferidas, el sustrato aceptor es una proteína terapéutica. Proteínas terapéuticas preferidas para su modificación mediante la glicosiltransferasa producida usando microorganismos que tienen un entorno intracelular oxidativo se encuentran en la tabla 2.

Tabla 2. Proteínas terapéuticas preferidas

<u>Hormonas y factores de crecimiento</u>	<u>Receptores y receptores quiméricos</u>
Factor estimulante de colonia de granulocitos (G-CSF)	CD4
Factor estimulante de colonia de macrófagos-granulocitos (GM-CSF)	Receptor del factor de necrosis tumoral (TNF-R)
TPO	Fusión de TNF-R:IgG Fc
Eritropoietina (EPO)	Alfa-CD20
Variante EPO	PSGL-1
Hormona estimulante del folículo (FSH)	Complemento
Hormona del crecimiento humana (HGH)	GlyCAM o su quimera
Insulina	N-CAM o su quimera
alfa-TNF	<u>Anticuerpos monoclonales (inmunoglobulinas)</u>
Leptina	AcM anti-RSV
Gonadotropina coriónica humana	Receptor de AcM anti-IL-2
Factor-20 de crecimiento de fibroblastos (FGF-20)	AcM anti-CEA
Factor-20 de crecimiento de fibroblastos (FGF-21)	AcM glicoproteína IIb/IIIa (Reopro™)
<u>Enzimas e inhibidores</u>	AcM anti-EGF
Activador de plasminógeno de tipo tisular (TPA)	AcM Her-2 (Herceptin™)
Variante de TPA	AcM CD20 (Rituxan™)
Urocinasa	AcM alfa-CD3
Factor VII factor de coagulación	Fusión receptor de TNF -IgG Fc (Enbrel™)
Factor VIII	AcM TNF α (Remicade™)
Factor IX de coagulación	AcM CD4
Factor X	AcM PSGL-1
Factor XIII	Proteína AcM anti F del virus sincitial respiratorio
hrADNasa	Anti-trombina-III
Glucocerebrosidasa (Cerezyme™)	<u>Células</u>
Hirudin	Glóbulos rojos
α 1 antitripsina (inhibidor de proteasa α 1)	Glóbulos blancos (por ejemplo, células T, células B, células dendríticas, macrófagos, células NK, neutrófilos, monocitos y similares)
Antitrombina III	Células madre
α -Glucosidasa ácida (maltasa ácida)	Plaquetas
α -galactosidasa A	<u>Otros</u>
α -L-iduronidasa	Antígeno de superficie de hepatitis B (HbsAg)
Urocinasa	Toxina de difteria quimérica-IL-2
<u>Citocinas y citocinas quiméricas</u>	
Interleucina-1 (IL-1), 1B, 2, 3, 4	
Interferón-alfa (IFN-alfa)	
Interferón-alfa-2b	
Interferón-beta	
Interferón-gamma	
Interferón-omega	

Otras proteínas terapéuticas preferidas que pueden producirse en organismos procariotas que tienen entornos intracelulares oxidantes se dan a conocer en la solicitud n.º PCT/US02/32263, presentada el 9 de octubre de 2002; solicitud de patente provisional n.º 60/448.381, presentada el 19 de febrero de 2003, solicitud de patente provisional n.º 60/438.582, presentada el 6 de enero de 2003; solicitud de patente provisional n.º 60/407.527, presentada el 28 de agosto de 2002; solicitud de patente provisional n.º 60/404.249, presentada el 16 de agosto de 2002; solicitud de patente provisional n.º 60/396.594, presentada el 17 de julio de 2002; solicitud de patente provisional n.º 60/391.777, presentada el 25 de junio de 2002; solicitud de patente provisional n.º 60/387.292, presentada el 7 de junio de 2002; solicitud de patente provisional n.º 60/334.301, presentada el 28 de noviembre de 2001; solicitud de patente provisional n.º 60/334.233, presentada el 28 de noviembre de 2001; solicitud de patente provisional n.º 60/344.692, presentada el 19 de octubre de 2001; y solicitud de patente provisional n.º 60/328.523, presentada el 10 de octubre de 2001; y en las siguientes publicaciones de solicitud de patente estadounidense 20040142856, 20040137557, 20040132640, 20040126838, 20040115168, 20040082026, 20040077836, 20040063911, 20040043446. Las proteínas terapéuticas preferidas en las referencias anteriores también se denominan péptidos preferidos para remodelamiento.

En algunas realizaciones las proteínas terapéuticas incluyen un sitio de glicosilación O-unido. El sitio de glicosilación O-unido puede producirse de manera natural en una proteína o péptido del tipo natural, o puede estar en una proteína o péptido mutante, por ejemplo, una proteína o péptido mutante en la que se introduce un sitio de glicosilación O-unido que no se produce de manera natural, o una proteína o péptido mutante que comprende sitios de glicosilación O-unidos tanto que se producen de manera natural como que no se producen de manera natural. Proteínas a modo de ejemplo con sitios de glicosilación O-unidos incluyen, por ejemplo, factor estimulante de colonia de granulocitos (G-CSF), por ejemplo, tipos naturales de 175 y 178 aminoácidos (con o sin residuos de metionina N-

terminales), interferón (por ejemplo, interferón alfa, por ejemplo, interferón alfa 2b, o interferón alfa 2a), factor estimulante de colonia de macrófagos y granulocitos (GM-CSF), hormona de crecimiento humano, interleucina (por ejemplo, interleucina 2), y factor de crecimiento de fibroblastos (FGF). Ejemplos de proteínas y péptidos de tipo natural y mutantes se encuentran en, por ejemplo, el documento PCT/US2004/014254, presentado el 7 de mayo de 2004; solicitud de patente provisional estadounidense n.º 60/469.114, presentada el 9 de mayo de 2003; solicitud de patente provisional estadounidense n.º 60/494.751, presentada el 13 de agosto de 2003; solicitud de patente provisional estadounidense n.º 60/495.076, presentada el 14 de agosto de 2003; solicitud de patente provisional estadounidense n.º 60/535.290, presentada el 8 de enero de 2003; el documento PCT/US05/000799, presentado el 10 de enero de 2005; el documento USSN 11/033365 presentado el 10 de enero de 2005; solicitud de patente provisional estadounidense n.º 60/535284, presentada el 8 de enero de 2004; solicitud de patente provisional estadounidense n.º 601544411, presentada el 12 de febrero de 2004; solicitud de patente provisional estadounidense n.º 60/546631, presentada el 20 de febrero de 2004; solicitud de patente provisional estadounidense n.º 60/555813, presentada el 23 de marzo de 2004; solicitud de patente provisional estadounidense n.º 60/570891, presentada el 12 de mayo de 2004; el documento PCT/US05/39226, presentado el 31 de octubre de 2005; y solicitud de patente provisional estadounidense n.º 60/623342, presentada el 29 de octubre de 2004.

La invención también abarca proteínas terapéuticas que se han modificado para aumentar su resistencia a proteasas. En una realización, la proteína terapéutica resistente a proteasa es una proteína de hormona de crecimiento humana. Proteínas terapéuticas resistentes a proteasa a modo de ejemplo se encuentran en por ejemplo, la solicitud de patente provisional estadounidense n.º 60/669736, presentada el 8 de abril de 2005; solicitud de patente provisional estadounidense n.º 60/710401, presentada el 22 de agosto de 2005; y solicitud de patente provisional estadounidense n.º 60/720030, presentada el 23 de septiembre de 2005.

VII. Conjugación de azúcares modificados a péptidos

Los azúcares modificados se conjugan a un péptido o proteína glicosilada o no glicosilada usando una enzima apropiada para mediar la conjugación. Preferiblemente, las concentraciones del/de los azúcar(es) donador(es) modificado(s), enzima(s) y péptido(s) o proteína(s) aceptora(s) se seleccionan de tal manera que la glicosilación procede hasta que se consume el aceptor. Las consideraciones tratadas a continuación, aunque se exponen en el contexto de una sialiltransferasa, pueden aplicarse generalmente a otras reacciones de glicosiltransferasa.

La presente invención también proporciona la producción a escala industrial de péptidos modificados. Tal como se usa en el presente documento, una escala industrial generalmente produce por ejemplo, al menos un microgramo, un miligramo o un gramo de conjugado purificado, terminado.

En la siguiente explicación, la invención se muestra a modo de ejemplo mediante la conjugación de restos de ácido siálico modificados a un péptido glicosilado. El ácido siálico modificado a modo de ejemplo se marca con PEG. La atención de la siguiente explicación en el uso de péptidos glicosilados y ácido siálico modificado con PEG es por claridad de la ilustración y no pretende implicar que la invención se limita a la conjugación de estas dos parejas. Un experto entiende que la explicación puede aplicarse generalmente a las adiciones de restos de glicosilo modificado distintos del ácido siálico. Además, la explicación puede aplicarse de igual manera a la modificación de una unidad de glicosilo con agentes distintos de PEG incluyendo otros polímeros solubles en agua, restos terapéuticos y biomoléculas.

Puede usarse un enfoque enzimático para la introducción selectiva de hidratos de carbono PEGilados o PPGilados en 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 apropiada. Seleccionando la glicosiltransferasa que preparará el enlace de hidrato de carbono deseado y utilizando el azúcar modificado como el sustrato donador, el PEG o PPG puede introducirse directamente sobre la estructura principal del péptido, sobre residuos existentes de azúcar de un glicopéptido o sobre los residuos de azúcar que se han añadido a un péptido.

Un aceptor para la sialiltransferasa está presente en el péptido que va a modificarse mediante los métodos de la presente invención o bien como una estructura que se produce de manera natural o bien se coloca de manera recombinante, enzimática o química. Aceptores adecuados, incluyen, por ejemplo, aceptores de galactosilo tales como Gal β 1,4G1cNAc, Gal β 1,4GalNAc, Gal β 1,3GalNAc, lacto-N-tetraosa, Gal β 1,3GlcNAc, Gal β 1,3Ara, Gal β 1,6GlcNAc, Gal β 1,4Glc (lactosa), GalNAc y ácido siálico que contienen estructuras y otros aceptores conocidos por los expertos en la técnica (véase, por ejemplo, Paulson *et al.*, J. Biol. Chem. 253: 5617-5624 (1978)).

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

En una realización a modo de ejemplo, el aceptor de galactosilo se ensambla 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). A la reacción se le permite proceder sustancialmente hasta completarse 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.

En aún otra realización, en primer lugar se "recortan" los oligosacáridos unidos a glicopéptidos, o bien completamente o bien en partes, para exponer o bien a un aceptor para la sialiltransferasa o bien un resto al que pueden añadirse uno o más residuos apropiados para obtener un aceptor adecuado. Enzimas tales como glicosiltransferasas y endoglicosidasas (véase, por ejemplo la patente estadounidense n.º 5.716.812) son útiles para las reacciones de unión y recorte.

Métodos para la conjugación de azúcares modificados a péptidos o proteínas se encuentran 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.

Debe observarse que tal como se usa en el presente documento y en las reivindicaciones adjuntas, las formas singulares de "un", "y", y "el" incluyen referentes plurales a menos que el contexto claramente lo dicte de otra manera. Por tanto, por ejemplo, la referencia para "una célula" incluye una pluralidad de tales células y equivalentes de la misma conocidos por los expertos en la técnica, etc.

Las publicaciones comentadas en el presente documento se proporcionan solamente por su divulgación antes de la fecha de presentación de la presente solicitud. No debe interpretarse nada en el presente documento como una admisión de que la presente invención no tiene derecho a anteceder tal publicación por virtud de una invención previa. Además, las fechas de publicación proporcionadas pueden ser diferentes de las fechas de publicación reales que puede ser necesario confirmarlas independientemente.

Ejemplos

Ejemplo 1: Expresión de glicosiltransferasas de la ruta de oligosacáridos O-unidos.

Procedimientos generales

Se diseñaron constructos para expresar fusiones de proteína de unión a maltosa (MBP) a truncamientos aminoterminales de las glicosiltransferasas. Se diseñan constructos con un Δ (número) que se refiere a la posición del último aminoácido retirado (para truncamientos aminoterminales) o el primer aminoácido retirado (para truncamientos carboxiloterminales) de la proteína de longitud completa correspondiente. Se usaron los siguientes constructos: MBP-GalNAc-T2 humana (Δ 51), GalNAcT2 humana (Δ 51, Δ 445), MBP-Core-1-Gal-T1 de *Drosophila* (Δ 50), MBP-ST3Gal-1 porcina (Δ 45), y MBP-SBD-ST3Gal-1 porcina (Δ 45; SBD es la etiqueta de dominio de unión a almidón, insertada entre la MBP y los dominios catalíticos), y MBP-ST6GalNAc-1 humana (Δ 35). Se clonaron normalmente los ácidos nucleicos que codificaban para las enzimas en los sitios BamHI-XhoI o BamHI-EcoRI de pCWin2-MBP o una versión de pCWin2 con un 5' UTR modificado. Véase, por ejemplo, el documento PCT/US05/00302, presentado el 6 de enero de 2005, que se incorpora en el presente documento como referencia para todos los fines. Se realizó la clonación usando técnicas convencionales (por ejemplo Current Protocols in Molecular Biology, Ausubel, FM, *et al*, eds. John Wiley and Sons, Inc. 1998).

Para la expresión de la proteína, se usó el contenido de un asa de células de una placa reciente para inocular un cultivo de 100 ml de martone LB que contenía kanamicina 50 μ g/ml. Se incubaron los cultivos a 37°C con agitación a 180-200 rpm, y se monitorizaron para determinar la DO₆₂₀. Cuando la DO₆₂₀ alcanzó 0,4-0,6, se transfirieron los cultivos a un incubador con agitación a 20°C (180-200 rpm) durante 15-20 minutos. Entonces se añadió IPTG hasta una concentración final de 0,1 mM, y se continuó la incubación con agitación durante la noche a 20°C. Se recogieron las células mediante centrifugación a 4°C, 7000 x g durante 15 min.

Para el análisis de la solubilidad de proteína y la purificación parcial de proteínas de fusión, se resuspendieron sedimentos de células bacterianas de cultivos inducidos en 30 ml de TE (tris 20 mM pH 7,4, EDTA 1 mM), y se sometieron a lisis mediante ruptura mecánica con dos pases a través de un microfluidizador a 15.000 psi. Se sedimentó el material insoluble mediante centrifugación durante 10 minutos a 3000-5000 x g a 4°C. Se separaron los sobrenadantes de los sedimentos, y se tomó una muestra de cada sobrenadante para el análisis de ensayo de actividad. Se ajustaron los sobrenadantes restantes hasta una concentración final de NaCl 200 mM, y se incubaron mezclados a temperatura ambiente con resina de amilosa lavada. Tras 1-2 horas, se recogieron las perlas mediante centrifugación breve, se lavaron con 10 volúmenes de lecho de perlas de tampón de columna (tris 20 mM pH 7,4, NaCl 200 mM, EDTA 1 mM), y se recogieron en una columna desechable. Se eluyeron las proteínas de fusión purificadas con un volumen de lecho de perlas de tampón de columna que contenía maltosa 10 mM. Se analizaron las muestras de sedimento insoluble, sobrenadante, y elución de resina de amilosa mediante SDS-PAGE.

Para el ensayo de la actividad GalNAc-T2, se llevaron a cabo reacciones en una mezcla de muestra de enzima con Tris 20 mM pH 7, MnCl₂ 10 mM, UDP-GalNAc 1,5 mM, aceptor de péptidos sintéticos 1 mM. Tras una incubación de

30 minutos a 37°C, se extinguió la reacción con HCl 0,01 N, y se separó el aceptor de péptidos de la mezcla de reacción mediante centrifugación a través de un concentrador 10.000 MWCO. Se detectaron péptido y GalNAc-péptido y se cuantificaron mediante RP-HPLC.

Para el ensayo de la actividad Core-1-Gal-T1, se llevaron a cabo reacciones de galactosiltransferasa en una mezcla de muestra de enzima con mucina submaxilar asialo-bovina (asialo-BSM) 3,5 mg/ml, MES 50 mM pH 6,5, MnCl₂ 20 mM, y UDP-galactosa marcada radioactivamente. Tras una incubación de una hora a 37°C, se detuvo la reacción y se separaron los productos de reacción de proteína de la UDP-Gal mediante precipitación ácida. Se detectó posteriormente la galactosa radioactiva transferida a asialo-BSM y se cuantificó usando un contador de centelleo.

Para el ensayo de ST6GalNAc-1 y ST3Gal-1, se llevaron a cabo ensayos de sialiltransferasa en formato en fase sólida usando asialo-BSM como aceptor y formas biotiniladas del donador CMP-NAN. Brevemente, se incubó una placa de 96 pocillos recubierta con asialo-BSM a 37°C con muestras de MBP-ST6GalNAc-1 y CMP-NAN biotinilado en BisTris 20 mM pH 6,7, MgCl₂ 2,5 mM, MnCl₂ 2,5 mM, NaCl 50 mM, Tween-80 al 0,05% durante 2-4 horas. Entonces se lavó la microplaca, y se marcó el ácido siálico biotinilado transferido a la asialo-BSM unida a la placa con europio-estreptavidina, y se detectó mediante fluorescencia de resolución temporal.

ST3Gal-1

Se indujeron células JM109 y mutantes *trxB* gor supp que portaban el constructo MBP-ST3Gal-1 para la expresión durante la noche a 20°C. Tal como se muestra en la figura 1, se observó la expresión de MBP-ST3Gal-1 como proteína soluble en ambas cepas, con niveles de expresión más altos observados en la cepa mutante *trxB* gor supp. Se purificó la proteína de fusión soluble de ambos lisados sobre resina de amilosa (figura 1), y los ensayos de la actividad tanto en lisados como en muestras parcialmente purificadas indicaron que la proteína de fusión soluble era enzimáticamente activa (tabla 3), con niveles de actividad significativamente más altos recuperados de las células mutantes *trxB* gor supp.

GalNAc-T2

Cuando se expresan en células JM109, el MBP-GalNAcT2 es predominantemente insoluble, con sólo niveles traza de actividad detectados en la fracción soluble (tabla 3). Se introdujo el constructo GalNAc-T2 humano truncado marcado con MBP en una cepa mutante *trxB* gor supp, y se indujo para la expresión durante la noche a 20°C. Tal como se muestra en la figura 2, la MBP-GalNAc-T2 se expresó de manera soluble y se purificó rápidamente sobre resina de amilosa. Ensayos de la actividad tanto en lisados como en muestras parcialmente purificadas indicaron que la proteína de fusión soluble era una enzima activa expresada en niveles mucho más altos que los observados en lisados de células JM109 (tabla 3). Las GalNAcT2 truncadas que carecen tanto de secuencias aminoterminales como de su dominio lectina carboxiloterminales también eran solubles y activas cuando se expresan en células mutantes *trxB* gor supp (figura 2b y tabla 3).

Core-1-Gal-T1

Trabajos previos con el constructo Core-1-Gal-T1 de *Drosophila* truncado marcado con MBP se expresó de manera insoluble en células JM109 con ninguna actividad evidente en la fracción soluble (tabla 3). Se introdujo el constructo Core-1-Gal-T1 de *Drosophila* truncado marcado con MBP en una cepa mutante *trxB* gor supp, y se indujo para la expresión durante la noche a 20°C. Tal como se muestra en la figura 3, la MBP-GalNAc-T2 se expresó preferiblemente como una proteína soluble en la cepa mutante *trxB* gor supp, y se purificó parcialmente sobre resina de amilosa. Los ensayos de actividad en muestras de lisados indicaron que la proteína soluble era una enzima activa cuando se expresa o bien marcado con MBP o bien no marcado en el mutante *trxB* gor supp (tabla 3).

ST6GalNAc-1

Trabajos previos con constructos ST6GalNAc-1 marcados con MBP encontraron que la proteína de fusión se expresó como cuerpos de inclusión insolubles en células JM109. Se expresó el constructo ST6GalNAc-1 humano truncado marcado con MBP en una cepa mutante *trxB* gor supp como una proteína soluble, y se purificó parcialmente sobre resina de amilosa (figura 4). Los ensayos de actividad en la muestra de lisado detectaron actividad sialiltransferasa (tabla 3).

Tabla 3: Rendimientos basados en la actividad enzimática observada de las glicosiltransferasas de glicanos O-unidos

Un resumen de las actividades en muestras de lisados para las proteínas de fusión indicadas expresadas o bien en JM109 o bien en *E. coli* mutante *trxB* gor supp. nt, no sometido a prueba.

Glicosiltransferasa expresada	Expresada en células mutantes <i>trxB</i> gor supp Actividad enzimática (U/L)	Expresadas en células JM109 Actividad enzimática (U/L)
MBP-GalNAc-T2	6	nt
GalNAc-T2 Δ 445	5	nt
MBP-core1-Gal-T1	4,2	0
Core1-Gal-T1	2,5	nt
MBP-ST3Gal-1	100	1,1
MBP-SBD-ST3Gal-1	1,75	nt
MBP-ST6GalNAc-1	~0,03	0

Fermentación de 10 L de glicosiltransferasas expresadas en células mutantes *trxB* gor supp

Se sembraron en recipientes de fermentación de diez litros, cepas mutantes *trxB* gor supps de *E. coli* que expresaban una de las siguientes proteínas: MBP-ST3Gal-1, MBP-GalNAc-T2, y MBP-Core-1-Gal-T1. Tras el crecimiento a 37°C hasta una DO_{620} de aproximadamente 0,5 se desplazó la temperatura a 20°C. Después de que la temperatura de cultivo alcanzara 20°C, se añadió IPTG 0,1 mM para inducir la expresión de la proteína. Se extrajeron alícuotas de células en los puntos de tiempo indicados (figura 5), se procesaron para dar lisados clarificados, y se sometieron a ensayo para determinar la actividad enzimática. Tal como se muestra en la figura 5, se mantuvieron los niveles de enzima expresados o se aumentaron tras un periodo de inducción de 48 horas para cada proteína de fusión.

O-glicosilación y glicoPEGilación del interferón-alfa-2b

Se glicosiló interferón-alfa-2b usando glicosiltransferasas eucariotas que se habían producido en *E. coli* mutante *trxB* gor supp (figura 6). Primero se añadió GalNAc al interferón a través de la actividad de la enzima MBP-GalNAc-T2. Entonces se añadió galactosa a través de la actividad de la enzima MBP-Core1-Gal-T1. Se analizaron los productos de reacción mediante espectrometría de masas MALDI-TOF. La figura 6A es interferón-alfa-2b solo. La primera reacción (figura 6B) incluyó 40 μ g de interferón-alfa-2b, UDP-GalNAc 0,4 mM, y MBP-GalNAc-T2 20 mU/mg. La segunda reacción (figura 6C) incluyó interferón-alfa-2b (40 μ g), UDP-GalNAc 0,4 mM, MBP-GalNAc-T2 20 mU/mg, UDP-Gal 0,4 mM, y MBP-Core1-Gal-T1 20 mU/mg. Ambas reacciones se llevaron a cabo a 32°C durante seis horas en BisTris 20 mM pH 6,7, NaCl 50 mM, $MnCl_2$ 10 mM, y NaN_3 al 0,02%. Tal como se muestra en la figura 6A y 6B, se aumentó el pico de masa para el interferón en la primera reacción, lo que concuerda con la adición de GalNAc. De manera similar, en la figura 6C, la masa del interferón se aumenta adicionalmente en la segunda reacción lo que concuerda con la adición de GalNAc-Gal.

Tal como se muestra en la figura 7, también se usaron las glicosiltransferasas eucariotas producidas en *E. coli* mutante *trxB* gor supp para glicoPEGilar interferón-alfa-2b. Primero se llevó a cabo una reacción de glicosilación de interferón-alfa-2b usando MBP-GalNAcT2 y MBP-Core-1-Gal-T1 producidos a partir de lisados de mutante *trxB* gor supp. La primera reacción contenía 40 μ g de interferón-alfa-2b, UDP-Gal 0,4 mM, MBP-GalNAc-T2 20 mU/mg, UDP-Gal 0,4 mM, MBP-Core1-Gal-T1 20 mU/mg, BisTris 20 mM pH 6,7, NaCl 50 mM, $MnCl_2$ 10 mM, y NaN_3 al 0,02%. Tras una incubación de seis horas a 32°C, se añadieron CMP-NAN-40kDaPEG 0,08 mM, y MBP-ST3Gal-1 50 mU/mg, y la reacción continuó durante la noche a 32°C. Se monitorizó el progreso de la reacción mediante SDS-PAGE de alícuotas tomadas antes de la adición de MBP-GalNAc-T2 y MBP-core-1-Gal-T1, inmediatamente antes y después de la adición de MBP-ST3Gal-1, y en la finalización de la reacción. Tal como se muestra en la figura 7, carril 3, la adición de GalNAc-Gal provoca una reducción apreciable en la movilidad electroforética del interferón. La sialilPEGilación adicional del interferón mediante la MBP-ST3Gal-1 da como resultado una reducción drástica en la movilidad, lo que concuerda con la adición de ácido PEG-siálico 40 kDa (figura 7, carril 5)

Ejemplo 2: Expresión de glicosiltransferasas eucariotas a partir de la ruta de oligosacárido N-unido

Procedimientos generales

Se diseñaron constructos para expresar fusiones de proteína de unión a maltosa (MBP) a truncamientos aminoterminales de la glicosiltransferasas. Se designan los constructos con un Δ (número) que se refiere al número de aminoácidos retirados del extremo aminoterminal de la proteína nativa correspondiente. Se usaron los siguientes constructos: MBP-GnT1 humana (Δ 103), MBP-GalT1 bovina (Δ 129), y MBP-ST3Gal3 de rata (A72) y MBP-SBD-ST3Gal3 (A72; SBD es la etiqueta de dominio de unión a almidón, insertada entre la MBP y el dominio catalítico), y MBP-ST6GalNAc-1 humana (Δ 35). Para los GnT1 y GalT1, también se sometió a prueba una versión alternativa de cada enzima que portaba una simple mutación de cambio de sentido. Se clonaron normalmente los ácidos nucleicos que codificaban para las enzimas en los sitios BamHI-XhoI o BamHI-EcoRI de pCWin2-MBP. Véase, por ejemplo, el documento PGT/US05/00302, presentado el 6 de enero de 2005. Se realizó la clonación usando técnicas convencionales (por ejemplo Current Protocols in Molecular Biology, Ausubel, FM, *et al*, eds. John Wiley and Sons, Inc. 1998).

Se realizaron la expresión de la proteína, y los análisis de solubilidad y purificación tal como se describe en el ejemplo 1.

Para el ensayo de la actividad GnT1, se llevaron a cabo reacciones usando métodos desarrollados a partir de métodos comunes de la bibliografía (por ejemplo Schachter, Reck, y Paulson (2003) *Methods Enzymol* 363,459-475). Brevemente, se monitorizó la actividad n-acetilglucosaminiltransferasa como la transferencia de UDP-GlcNAc marcada radioactivamente a Man1,6(Man1,3)Man-O-octilo (OM3). Se purificaron el sustrato y el producto mediante fase inversa, y se detectaron y cuantificaron usando un contador de centelleo.

Para el ensayo de la actividad GalT1, se llevaron a cabo reacciones de galactosiltransferasa en una mezcla de muestra de enzima en HEPES 50 mM pH 7,5, lacto-n-triosa II 4,5 mM, UDP-Gal 6 mM, y MnSO₄ 5 mM. Tras una incubación de una hora a 37°C, se detuvo la reacción y se separaron productos de reacción de hidratos de carbono mediante centrifugación a través de un concentrador 10.000 MWCO. Se detectaron y cuantificaron el sustrato y el producto mediante HPLC.

Para el ensayo de ST3Gal3, se llevaron a cabo ensayos de sialiltransferasa en una mezcla de muestra de enzima en MOPS 20 mM pH 6,5, BSA 0,1 mg/ml, MnCl₂ 10 mM, CMP-NAN 2 mM, y lacto-n-neotetraosa 30 mM. Tras una incubación de dos horas a 30°C, se detuvo la reacción mediante inactivación térmica, y se detectaron y cuantificaron el sustrato de la reacción y el producto mediante HPLC.

GnT1

Se analizaron cultivos de JM109 y *E. coli* mutante trxB gor supp inducidos para la expresión de MBP-GnT1 y MBP-GnT1 C121S a 20°C para determinar la solubilidad y la actividad de las proteínas de fusión. Tal como se muestra en la figura 8, se observaron niveles más altos de expresión soluble de ambas proteínas de fusión MBP-GnT1 en la cepa mutante trxB gor supp, y se recuperaron todas de las proteínas de fusión solubles con resina de amilosa. Los ensayos de actividad en las muestras de MBP-GnT1 indicaron que se expresó la enzima a niveles de actividad más de 25 veces más altos en células mutantes trxB gor supp que en células JM109 (tabla 4). También se expresó la MBP-GnT1 C121S como una enzima activa soluble, aunque a niveles aproximadamente diez veces más bajos que en el constructo MBP-GnT1 en la cepa de *E. coli* equivalente (tabla 4).

GalT1

Se analizaron cultivos de JM109 y *E. coli* mutante trxB gor supp inducidos para la expresión de MBP-GalT1 y MBP-GalT1 C342T a 20°C para determinar la solubilidad y la actividad de las proteínas de fusión. Tal como se muestra en la figura 9, se observaron niveles más altos de la expresión soluble de ambas proteínas de fusión MBP-GalT1 en la cepa mutante trxB gor supp, y se recuperaron todas de las proteínas de fusión solubles usando resina de amilosa. Los ensayos de actividad en MBP-GalT1 expresada en JM109 fueron incapaces de detectar la actividad enzimática, mientras que se recuperó enzima MBP-GalT1 activa de la muestra de expresión del mutante trxB gor supp (tabla 4). Se recuperó el constructo MBP-GalT1 C342T como una enzima activa de ambas cepas de *E. coli*, observándose niveles de actividad más de 5 veces más altos en las muestras del mutante trxB gor supp.

ST3Gal3

Se analizaron cultivos de *E. coli* mutante trxB gor supp inducida para la expresión de MBP-ST3Gal3 y MBP-SBD-ST3Gal3 a 20°C para determinar la solubilidad y la actividad de las proteínas de fusión. Tal como se muestra en la figura 10, se expresaron de manera soluble ambas versiones de la proteína de fusión ST3Gal3 en las células mutantes trxB gor supp, y se recuperaron ambas sobre resina de amilosa. Ambas proteínas de fusión ST3Gal3 expresadas de manera soluble en células mutantes trxB gor supp eran enzimáticamente activas, con el constructo marcado con MBP más de cinco veces más activo que el constructo marcado con MBP-SBD (tabla 4). La MBP-SBD-ST3Gal3 expresada en células JM109 no tenía actividad enzimática detectable (tabla 4).

Tabla 4: Rendimientos basados en la actividad enzimática observada de las glicosiltransferasas glicanos N-unidos

Un resumen de las actividades en muestras de lisados para las proteínas de fusión indicadas expresadas o bien en JM109 o bien en *E. coli* mutante trxB gor supp. nt, no sometido a prueba

Glicosiltransferasa expresada	Expresada en células mutantes trxB gor supp Actividad enzimática (U/L)	Expresada en Células JM109 Actividad enzimática (U/L)
MBP-GnT1	45	1,7
MBP-GnT1 C121S	4,3	0,4
MBP-GalT1	175	0
MBP-GalT1 C342T	193	37
MBP-ST3Gal3	4,6	nt
MBP-SBD-ST3Gal3	0,9	0

Ejemplo 3: Expresión de glicosiltransferasas eucariotas en *Pseudomonas*.

Procedimientos generales

Se sometieron a prueba las actividades de las siguientes glicosiltransferasas eucariotas en un sistema de expresión en *Pseudomonas*: dos truncamientos N-terminales de ST3Gal-1 porcina, y ST6GalNAc-1 de pollo. Los constructos, refiriéndose (Δ número) al número de aminoácidos retirados del extremo aminoterminal de la proteína nativa correspondiente, fueron: ST3Gal-1 porcina (Δ 45), ST3Gal-1 porcina (Δ 56), y ST6GalNAc-1 de pollo (Δ 231). Se fusionaron las glicosiltransferasas a una secuencia de secreción en *Pseudomonas* en la que la expresión se dirige al periplasma, y/o se expresaron como proteínas no fusionadas dirigidas al citoplasma. Se dirigió la expresión por el promotor Ptac inducible por IPTG. Se realizó la clonación usando técnicas convencionales (por ejemplo Current Protocols in Molecular Biology, Ausubel, FM, *et al*, eds. John Wiley and Sons, Inc.1998).

Se transformaron plásmidos que comprendían ácidos nucleicos que expresaban las glicosiltransferasas en la cepa *P. fluorescens* DC206 (Δ pyrF, *lacI*^{R1}). Véase, por ejemplo, el documento US 2005/0186666, publicado el 25 de agosto de 2005. Se hicieron crecer las células transformadas a escala de un litro en un medio mínimo (M9) complementado con glucosa al 1% y oligoelementos. Tras una fase de crecimiento inicial, se indujo la expresión de glicosiltransferasa mediante la adición de IPTG al medio. Se hicieron crecer las células en medio de inducción durante entre 24 y 120 horas. Se recogieron las células mediante centrifugación, se desecharon los sobrenadantes y se congelaron los sedimentos celulares y se almacenaron a -20°C.

Para la preparación de lisado celular, se descongeló la pasta celular congelada (8-14 gramos de peso en húmedo) y se resuspendió en tampón de lisis MES (MES 50 mM, pH 6,5) que contenía NaCl 100 mM a una razón de aproximadamente 1 gramo de pasta celular en húmedo por 2 ml de tampón de lisis. Se perturbaron las suspensiones mediante dos pases a través de una celda de presión francesa a 10.000 y 20.000 psi, respectivamente. Se eliminó el material insoluble mediante centrifugación, y se clarificaron adicionalmente los sobrenadantes mediante el paso a través de filtros de jeringa de 0,45 μ m y 0,2 μ m. Se analizaron las muestras de las fracciones insoluble y soluble mediante SDS-PAGE.

Para el ensayo de actividad sialiltransferasa, se llevaron a cabo reacciones en una mezcla de muestra de enzima con asialo-fetina 0,5-2 mg/ml (para ST3Gal-1) o asialo-BSM (para ST6GalNAc-1), MES o BisTris 50 mM pH 6,5, NaCl 100 mM, y CMP-NAN marcado radioactivamente. Tras una incubación de 30-60 minutos a 37°C, se detuvo la reacción y se separaron los productos de reacción de la proteína de CMP-NAN mediante precipitación ácida. Se detectó posteriormente el ácido siálico radioactivo transferido a asialo-fetina o asialo-BSM y se cuantificó usando un contador de centelleo.

ST3Gal-1

Se sometieron a ensayo lisados celulares de cultivos de *Pseudomonas* que expresaban los constructos ST3Gal-1 periplásmicos o citoplásmicos para determinar la actividad sialiltransferasa. Tal como se resume en la tabla 5, se observó actividad sialiltransferasa en muestras de constructos Δ 45 y Δ 56 de ST3Gal-1 dirigidos tanto al periplasma como al citoplasma, con niveles más altos de actividad de las muestras Δ 56 de ST3Gal-1.

ST6GalNAc-1

Se sometieron a ensayo lisados celulares de cultivos de *Pseudomonas* que expresaban ST6GalNAc-1 de pollo citoplásmico para determinar la actividad sialiltransferasa. Tal como se resume en la tabla 5, se observó actividad sialiltransferasa en muestras de ST6GalNAc-1 de pollo dirigida al citoplasma.

Tabla 5: Rendimientos basados en la actividad enzimática observada de las glicosiltransferasas glicanos O-unidos expresadas en *Pseudomonas*

Un resumen de las pruebas de actividad y expresión de constructos ST6GalNAc-1 y ST3Gal-1 usando *Pseudomonas*

Sialiltransferasa expresada	Actividad enzimática (U/L)	
	Citoplasmática	Periplasmática
Δ 45 de ST3Gal-1 porcina	0,6	0,5
Δ 56 de ST3Gal-1 porcina	1,1	1,7
ST6GalNAc-1 de pollo	0,1	

Se entiende que los ejemplos y realizaciones descritos en el presente documento son sólo para fines ilustrativos y que a la luz de los mismos se sugerirán diversas modificaciones o cambios a los expertos en la técnica y deben incluirse dentro del alcance y las reivindicaciones adjuntas.

Lista de secuencias

SEQ ID NO: 1.

beta-1,2-N-acetilglucosaminiltransferasa I (GnTI, GnT1) humana

MLKKQSAGLVWLGAILFVAWNALLLLFFWTRPAPGRPPSVSALDGPASLTREVIRLAQDAEVELERQRGLLQQI
GDALSSQGRVPTAAPPAQPRVPVTPAPAVIPILVIACDRSTVRRCLDKLLHYRPSAELFPIIVSQDCGHEETAQ
AIASYGSVAVTHIRQPDLSIAVPPDHRKFQGYKILARHYRWALGQVFRQFRFPAAVVVEDDLEVAPDFFEYFRAT
YPLLKADPSLWCVSAWNDNGKEQMVDA SRPELLYRTDFFPGLGWLLLAE LWAELEPKWPKAFWDDWMRRPEQRQG
RACIRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVHFTQLDLSYLQREAYDRDFLARVYGAPQLQVEKVRT
NDRKELGEVRVQYTGRDSFKAKALGVMDLKGSGVPRAGYRGIVTFQFRGRRVHLAPPPTWEGYDPSWN

5 SEQ ID NO: 2.

beta-1,2-N-acetilglucosaminiltransferasa I (GnT1, GnT1) de conejo

MLKKQSAGLVWLGAILFVAWNALLLLFFWTRPVPSRLPSDNALDDDPASLTREVIRLAQDAEVELERQRGLLQQI
REHHALWSQRWKVPTAAPPAQPHVPVTPPPAVIPILVIACDRSTVRRCLDKLLHYRPSAELFPIIVSQDCGHEET
AQVIASYGSVAVTHIRQPDLSNIAVQPDHRKFQGYKILARHYRWALGQIFHNFNYPAAVVVEDDLEVAPDFFEYFQ
ATYPLLKADPSLWCVSAWNDNGKEQMVDS SKPELLYRTDFFPGLGWLLLAE LWAELEPKWPKAFWDDWMRRPEQR
KGRACVRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVFPFTQLDLSYLQQEAYDRDFLARVYGAPQLQVEKV
RTNDRKELGEVRVQYTGRDSFKAKALGVMDLKGSGVPRAGYRGIVTFLFRGRRVHLAPPQTWDGYDPSWT

SEQ ID NO: 3.

beta-1,2-N-acetilglucosaminiltransferasa I (GnTI, GnT1) de ratón

MLKKQTAGLVWLGAIIFVGWNALLLLFFWTRPAPGRPLPSDSALGDDPASLTREVIHLAEDAEAELEERQRGLLQQI
KEHYALWRQRWRVPTVAPPAPWRVPVTPSPVQIPILVIACDRSTVRRCLDKLLHYRPSAERFPIIVSQDCGHEET
AQVIASYGTAVTHIRQPDLSNIAVQPDHRKFQGYKILARHYRWALGQIFNKFKFPAAVVVEDDLEVAPDFFEYFQ
ATYPLLRTDPSLWCVSAWNDNGKEQMVDS SKPELLYRTDFFPGLGWLLLADLWAELEPKWPKAFWDDWMRRPEQR
KGRACIRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVFPFTQLDLSYLQQEAYDRDFLAQVYGAPQLQVEKV
RTNDQKELGEVRVQYTSRDSFKAKALGVMDLKGSGVPRAGYRGIVTFQFRGRRVHLAPPQTWTGYDPSWN

10

SEQ ID NO: 4.

beta-1,2-N-acetilglucosaminiltransferasa I (GnTI, GnT1) de rata

MLKKQSAGLVWLGAIIFVGWNALLLLFFWTRPAPGRPLPSDSALGDDPASLTREVIHLAEDAEAELEERQRGLLQQI
KEHYSLWRQRWRVPTVAPPAPWRVPVTPSPAVIPILVIACDRSTVRRCLDKLLHYRPSAEHFPIIVSQDCGHEET
AQVIASYGTAVTHIRQPDLSNIAVQPDHRKFQGYKILARHYRWALGQIFNKFKFPAAVVVEDDLEVAPDFFEYFQ
ATYPLLKADPSLWCVSAWNDNGKEQMVDS SKPELLYRTDFFPGLGWLLLADLWAELEPKWPKAFWDDWMRRPEQR
KGRACIRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVFPFTQLDLSYLQREAYDRDFLAQVYGAPQLQVEKV
RTNDRKELGEVRVQYTSRDSFKAKALGVMDLKGSGVPRAGYRGIVTFQFRGRRVHLAPPETWNGYDPSWN

SEQ ID NO: 5.

15 Mutante C121S de beta-1,2-N-acetilglucosaminiltransferasa I (GnT1 Δ103) humana truncada

AVIPILVIACDRSTVRRSLDKLLHYRPSAELFPIIVSQDCGHEETAQAIAASYGSAVTHIRQPDLSIAVPPDHRK
FQGYKIHARHYRWALGQVFRQFRPAAVVVEDDLEVAPDFFEYFRATYPLLKADPSLWCVSAWNDNGKEQMVDAS
RPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISRTMTFGRKGVSHGQFFDQH
LKFIKLNQQFVHFTQLDLSYLQREAYDRDRLARVYGAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAKALGV
MDDLKSGVPRAGYRGIVTFQFPGRRVHLAPPPTWEGYDPSWN

SEQ ID NO: 6.

Mutante C121D de beta-1,2-N-acetilglucosaminiltransferasa I (GnT1 Δ 103) humana truncada

AVIPILVIACDRSTVRRDLKLLHYRPSAELFPIIVSQDCGHEETAQAIAASYGSAVTHIRQPDLSIAVPPDHRK
FQGYKIHARHYRWALGQVFRQFRPAAVVVEDDLEVAPDFFEYFRATYPLLKADPSLWCVSAWNDNGKEQMVDAS
RPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISRTMTFGRKGVSHGQFFDQH
LKFIKLNQQFVHFTQLDLSYLQREAYDRDRLARVYGAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAKALGV
MDDLKSGVPRAGYRGIVTFQFPGRRVHLAPPPTWEGYDPSWN

5 SEQ ID NO: 7.

Mutante C121T de beta-1,2-N-acetilglucosaminiltransferasa I (GnT1 Δ 103) humana truncada

AVIPILVIACDRSTVRRSLDKLLHYRPSAELFPIIVSQDCGHEETAQAIAASYGSAVTHIRQPDLSIAVPPDHRK
FQGYKIHARHYRWALGQVFRQFRPAAVVVEDDLEVAPDFFEYFRATYPLLKADPSLWCVSAWNDNGKEQMVDAS
RPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISRTMTFGRKGVSHGQFFDQH
LKFIKLNQQFVHFTQLDLSYLQREAYDRDRLARVYGAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAKALGV
MDDLKSGVPRAGYRGIVTFQFPGRRVHLAPPPTWEGYDPSWN

SEQ ID NO: 8.

Mutante C121A de beta-1,2-N-acetilglucosaminiltransferasa I (GnT1 Δ 103) humana truncada

AVIPILVIACDRSTVRRALDKLLHYRPSAELFPIIVSQDCGHEETAQAIAASYGSAVTHIRQPDLSIAVPPDHRK
FQGYKIHARHYRWALGQVFRQFRPAAVVVEDDLEVAPDFFEYFRATYPLLKADPSLWCVSAWNDNGKEQMVDAS
RPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISRTMTFGRKGVSHGQFFDQH
LKFIKLNQQFVHFTQLDLSYLQREAYDRDRLARVYGAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAKALGV
MDDLKSGVPRAGYRGIVTFQFPGRRVHLAPPPTWEGYDPSWN

10

SEQ ID NO: 9.

Mutante doble R120A C121H de beta-1,2-N-acetilglucosaminiltransferasa I (GnT1 Δ 103) humana truncada

AVIPILVIACDRSTVRAHLDKLLHYRPSAELFPIIVSQDCGHEETAQAIAASYGSAVTHIRQPDLSIAVPPDHRK
FQGYKIHARHYRWALGQVFRQFRPAAVVVEDDLEVAPDFFEYFRATYPLLKADPSLWCVSAWNDNGKEQMVDAS
RPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISRTMTFGRKGVSHGQFFDQH
LKFIKLNQQFVHFTQLDLSYLQREAYDRDRLARVYGAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAKALGV
MDDLKSGVPRAGYRGIVTFQFPGRRVHLAPPPTWEGYDPSWN

SEQ ID NO: 10.

15 beta-1,2-N-acetilglucosaminiltransferasa I (MBP-GnT1 Δ 103) humana truncada marcada con proteína de unión a maltosa

MKIEEGKLVWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSG
 LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
 NLQEPYFTWPLIAADGGYAFKYENGYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAM
 TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
 GAVALKSYEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNMNN
 NNNNNNLGIEGRISEFGSAVIPILVIACDRSTVRRCLDKLLHYRPSAELFPIIVSQDCGHEETAQAIAASYGSAVT
 HIRQPDLSIAVPPDHRKFQGYKIAHYRWALGQVFRQFRFPAAVVEDDLEVAPDFFEYFRATYPLLKADPSL
 WCVSAWNDNGKEQMVDASRPPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISR
 TMTFGRKGVSHGQFFDQHLKFIKLNQQFVHFTQLDLSYLOREAYDRDFLARVYGAPQLQVEKVRTNDRKELGEVR
 VQYTGDRDSFKAFKALGVMDLKGSGVPRAGYRGIVTFQFRGRRVHLAPPPTWEGYDPSWN

SEQ ID NO: 11.

Mutante C121S de beta-1,2-N-acetilglucosaminiltransferasa I (MBP-GnT1 Δ 103) humana truncada marcada con proteína de unión a maltosa

MKIEEGKLVWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSG
 LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
 NLQEPYFTWPLIAADGGYAFKYENGYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAM
 TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
 GAVALKSYEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNMNN
 NNNNNNLGIEGRISEFGSAVIPILVIACDRSTVRRSLDKLLHYRPSAELFPIIVSQDCGHEETAQAIAASYGSAVT
 HIRQPDLSIAVPPDHRKFQGYKIAHYRWALGQVFRQFRFPAAVVEDDLEVAPDFFEYFRATYPLLKADPSL
 WCVSAWNDNGKEQMVDASRPPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISR
 TMTFGRKGVSHGQFFDQHLKFIKLNQQFVHFTQLDLSYLOREAYDRDFLARVYGAPQLQVEKVRTNDRKELGEVR
 VQYTGDRDSFKAFKALGVMDLKGSGVPRAGYRGIVTFQFPGRRVHLAPPPTWEGYDPSWN

5

SEQ ID NO: 12.

beta-1,4-galactosiltransferasa (GalT1) humana

MRLREPLLSGAAMPGLASLQACRLLVAVCALHLGVTLVYYLAGRDLSRLPQLVGVSTPLQGGSNSAAAIGQSSGE
 LRTGGARPPPPLGASSQPRPGDSSPVVDSGPGPASNLTSVPVPHTTALSPLACPEESPLLVGPMLEFNMPVDL
 ELVAKQNPVNMGGRYAPRDCVSPHKVAIIIPFRNRQEHLYWLYLHPVLQRQQLDYGIYVINQAGDTIFNRAK
 LLNVGFQEALKDYDYTCFVFSVDLI PMNDHNAYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSQQLFTIN
 GFPNNYWGWGGEDDDI FNRLVFRGMSISRPNNAVVGRCRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDGLNSLTY
 QVLDVQRYPLYTQITVDIGTPS

SEQ ID NO: 13.

10 beta-1,4-galactosiltransferasa (GalT1) de ratón

MRFREQFLGGSAAMPGATLQACRLLVAVCALHLGVTLVYYLSGRDLSRLPQLVGVSSSTLQGGTNGAAASKQPPG
EQRPRGARPPPPLGVSPKPRPGLDSSPGAASGPGLKSNLSSLPVPTTGLLSLPACPEESPLLVGPMMLIDFNIAV
DLELLAKKNPEIKTGGRYSPKDCVSPHKVAIIIPFRNRQEHLYWLYLHPILQRQQLDYGIYVINQAGDTMFNR
AKLLNIGFQEALKDYDYNCFVFSVDVLIIPMDNRNAYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSQKQFLA
INGFPNNYWGWWGGEDDDIFNRLVHKGMSISRPNAVVGRCRMIRHSRDKKNEPNPQRFDRIAHTKETMRFDGLNSL
TYKVLVDQRYPLYTQITVDIGTPR

SEQ ID NO: 14.

beta-1,4-galactosiltransferasa (GalT1) bovina

MKFREPLLGGSAAMPGASLQACRLLVAVCALHLGVTLVYYLAGRDLRRLPQLVGVHPPPLQGSSHGAAAIGQPSG
ELRLRGVAPPPPLQNSSKPRSRAPSNLDAYSHPGPGPGPGSNLTSAPVPSTTTTSLTACPEESPLLVGPMMLIEFN
IPVDLKLIEQQNPVKVLGGRYTPMDCISPHKVAIIILFRNRQEHLYWLYLHPMVQRQQLDYGIYVINQAGESM
FNRAKLLNVGFKEALKDYDYNCFVFSVDVLIIPMDHNTYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSQKQ
FLSINGFPNNYWGWWGGEDDDIYNRLAFRGMSVSRPNAVIGKCRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDDL
NSLTVMVLEVQRYPLYTKITVDIGTPS

5 SEQ ID NO: 15.

beta-1,4-galactosiltransferasa (GalT1) bovina

MKFREPLLGGSAAMPGASLQACRLLVAVCALHLGVTLVYYLAGRDLRRLPQLVGVHPPPLQGSSHGAAAIGQPSG
ELRLRGVAPPPPLQNSSKPRSRAPSNLDAYSHPGPGPGPGSNLTSAPVPSTTTTSLTACPEESPLLVGPMMLIEFN
IPVDLKLIEQQNPVKVLGGRYTPMDCISPHKVAIIILFRNRQEHLYWLYLHPMVQRQQLDYGIYVINQAGESM
FNRAKLLNVGFKEALKDYDYNCFVFSVDVLIIPMDHNTYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSQKQ
FLSINGFPNNYWGWWGGEDDDIYNRLAFRGMSVSRPNAVIGKCRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDDL
NSLTVMVLEVQRYPLYTKITVDIGTPS

SEQ ID NO: 16.

beta-1,4-galactosiltransferasa (GalT1, CKI) de pollo

MKEPALPGTSLQACRLLVAFALHLSATLLYYLAGSSLTPPRSPEPPPPRRPPANLSLPPSRPPPPPAARPRPG
PVSAQPRNLPSAPSGLCPDPSPLLVGPLRVEFSQPVNLEEVASTNPEVREGGRFAPKDCALQKVAIIIPFRNR
EEHLKYWLYYMHPIQRQQLDYGVYVINQDGEDEFNRAKLLNVGFTEALKEYDYDCVFSVDVLIIPMDNRNTYKC
YSQPRHLSVSMDFKGFRLPYNQYFGGVSALSKEQFTKINGFPNNYWGWWGGEDDDIYNRLVFKGMGISRPDAVIGK
CRMIRHSRDRKNEPNPERFDRIAHTRTMSDDLNSLSYEVLRDRFPLYTRITVDIGAPGS

10

SEQ ID NO: 17.

beta-1,4-galactosiltransferasa (GalT1 Δ 129) bovina truncada

SLTACPEESPLLVGPMMLIEFNIPVDLKLIEQQNPVKVLGGRYTPMDCISPHKVAIIILFRNRQEHLYWLYLHP
MVQRQQLDYGIYVINQAGESMFNRAKLLNVGFKEALKDYDYNCFVFSVDVLIIPMDHNTYRCFSQPRHISVAMDK
FGFSLPYVQYFGGVSALSQKQFLSINGFPNNYWGWWGGEDDDIYNRLAFRGMSVSRPNAVIGKCRMIRHSRDKKNE
PNPQRFDRIAHTKETMLSDDLNSLTVMVLEVQRYPLYTKITVDIGTPS

SEQ ID NO: 18.

15 Mutante C342T de beta-1,4-galactosiltransferasa (GalT1 Δ 129) bovina truncada

SLTACPEESPLLVGPMLEIFNIPVDLKLIEQQNPKVKLGGRYTPMDCISPHKVAIIILFRNRQEHLYWLYYLHP
MVQRQQLDYGIIYVINQAGESMFNRAKLLNVGFKEALKDYDYNCFVFSVDVLI PMNDHNTYRCFSQPRHISVAMDK
FGFSLPYVQYFGGVSALSQQFLSINGFPNNYWGWWGGEDDDIYNRLAFRGMSVSRPNAVIGKTRMIRHSRDKKNE
PNPQRFDRIAHTKETMLSDGLNSLTVMLEVQRYPLYTKITVDIGTPS

SEQ ID NO: 19.

beta-1,4-galactosiltransferasa (MBP-GalT1 Δ129) bovina truncada marcada con proteína de unión a maltosa

MKIEEGKLVIWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDII FWAHDRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAFNKGEMTAM
TINGPWAWSNIDTSKVNIGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSNNNN
NNNNNNLGIEGRISEFSLTACPEESPLLVGPMLEIFNIPVDLKLIEQQNPKVKLGGRYTPMDCISPHKVAIIILF
RNRQEHLYWLYYLHPMVQRQQLDYGIIYVINQAGESMFNRAKLLNVGFKEALKDYDYNCFVFSVDVLI PMNDHNT
YRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSQQFLSINGFPNNYWGWWGGEDDDIYNRLAFRGMSVSRPNAV
IGKCRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDGLNSLTVMLEVQRYPLYTKITVDIGTPS

5 SEQ ID NO: 20.

Mutante C342T de beta-1,4-galactosiltransferasa (MBP-GalT1 Δ129) bovina truncada marcada con proteína de unión a maltosa

MKIEEGKLVIWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDII FWAHDRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAFNKGEMTAM
TINGPWAWSNIDTSKVNIGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSNNNN
NNNNNNLGIEGRISEFSLTACPEESPLLVGPMLEIFNIPVDLKLIEQQNPKVKLGGRYTPMDCISPHKVAIIILF
RNRQEHLYWLYYLHPMVQRQQLDYGIIYVINQAGESMFNRAKLLNVGFKEALKDYDYNCFVFSVDVLI PMNDHNT
YRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSQQFLSINGFPNNYWGWWGGEDDDIYNRLAFRGMSVSRPNAV
IGKTRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDGLNSLTVMLEVQRYPLYTKITVDIGTPS

SEQ ID NO: 21.

10 beta-1,4-galactosiltransferasa 2 (CKII) de pollo

MTRLLLGVTLERICKAVLLLCLLHFVIMILYFDVYAQHLDFFSRFNARNTSRVHPFSNSSRPNSTAPSYGPRGA
EPPSPSAKPNTRSVTEKPLQPCQEMPSGLVGRLLIEFSSEPMSEMERVQRENPDVSLGGKYTPPDCLPRQKVAILI
PFRHREHHLKYWLHYLHPILRRQKVAYGIYIINQYGEDTFNRAKLLNVGFLEALKDDEEYDCFI FSDVDLIPMDD
RNLYRCYEQPRHFAVGMDKFGFRLPYAGYFGGVSGLSKSQFLKINGFPNEYWGWWGGEDDDIYNRLISLNGMKVSRP
DIRMGRYRMIKHERDKHNEPNPQRFRTKIQNTKMTMKRDGISSLOQYRLVEVSRQPMYTNITVEIGRPPPLARG

SEQ ID NO: 22.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) humana

MGLLVFVRNLLLALCLFLVLGFLYYSAWKLHLLQWEEDSNSVVLSDSAGQTLGSEYDRLGFLNLD SKLPAELA
TKYANFSEGACKPGYASALMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTP
ALDSLRCRRCIIIVGNGGVLANKSLGSRIDDYDIVVRLNSAPVKGFEKDVGSKTTLRITYPEGAMQRPEQYERDSL
FVLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNGLMGRGNIP
TLGSVAVTMALHGCDEVAVAGFGYDMSTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLSSGI

SEQ ID NO: 23.

beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) de cerdo

MGLLVFVRNLLLALCLFLVLGFLYYSAWKLHLLQWEDSNSVVLSDSAGQTLGSEYDRLGFLNLD SKLPAELAT
KYANFSEGACKPGYASAVMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPA
LDLSLSCRRCIIIVGNGGVLANKSLGSRIDDYDIVVRLNSAPVKGFEKDVGSKTTLRITYPEGAMQRPEQYERDSL
VLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNGLMGRGNIPT
LGSVAVTMALHGCDEVAVAGFGYDMSTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLTSGI

5 SEQ ID NO: 24.

beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) de vaca

MGPLIFVRNLLLALCLFLVLGFLYYSAWKLHLLQWEDSNSVVLSDSAGQTLGSEYDRLGFLNLD SKLPAELAT
KYENFSEGACKPGYASALMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPA
LDLSLSCRRCIIIVGNGGVLANKSLGSRIDDYDIVVRLNSAPVKGFEKDVGSKTTLRITYPEGAMQRPEQYERDSL
VLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNGLMGRGNIPT
LGSVAVTMALHGCDEVAVAGFGYDMSTPNAPLHYEAVRMAAIRESWTHNIQREKEFLRKLVKARVITDLTSGI

SEQ ID NO: 25.

beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) de pollo

MGLLVFMRNLLLALCLFLVLGFLYYSAWKLHLLRWEDSNSVVLSDSVGHTIGSEYDRLGFLNLD SKLPAELAT
KYANFSEGACKPGYASAVMTAIFPRFSKPAPMFLDDSFRRWARIRDFVPPFGIKGQDNLIKAILSATKDYRLTPA
LDLSLSCRRCIIIVGNGGVLANKSLGLKIDDYDVVRLNSAPVKGFEKDVGGKTTLRITYPEGAIQKMEQYEKDSL
VLGFKWQDFKWLKYIVYKEKVSASDGFWKS VATRVPREPHEIRILNPYFIQEAAFSFIGLPFNGLMGRGNIPT
LGSVAITMALHNCDEVAVAGLGYDMSSPNAPLHYENIKMSAIKESWTHNIQREKEFLRKLVKARVITDLTSGI

10

SEQ ID NO: 26.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) de rata

MGLLVFVRNLLLALCLFLVLGFLYYSAWKLHLLQWEDSNSLILSLDSAGQTLGTEYDRLGFLNLD SKLPAELAT
KYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPA
LDLSLHCRRCIIIVGNGGVLANKSLGSRIDDYDIVIRLNSAPVKGFEKDVGSKTTLRITYPEGAMQRPEQYERDSL
VLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNGLMGRGNIPT
LGSVAVTMALDGCDEVAVAGFGYDMNTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLSSGI

SEQ ID NO: 27.

15 ST3Gal III, SIAT3/SIAT6, EC 2.4.99.6

Mutante D311H de Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) de rata

MGLLVFVRNLLLALCLFLVLGFLYSAWKLHLLQWEDSNLILSLDSAGQTLGTEYDRLGFLCLKLDSKLP AELAT
 KYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPA
 LDSLHCRRCIIIGNGGVLANKSLGSRIDDYDIVIRLNSAPVKGFEKDVGSKTTLRITYPEGAMQORPEQYERDSL
 FVLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNNGLMGRGNIPT
 LGSVAVTMALHGCDEVAVAGFGYDMNTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLSSGI

SEQ ID NO: 28.

ST3Gal III, SIAT3/SIAT6, EC 2.4.99.6

Mutante V162I de Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) de rata

MGLLVFVRNLLLALCLFLVLGFLYSAWKLHLLQWEDSNLILSLDSAGQTLGTEYDRLGFLCLKLDSKLP AELAT
 KYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPA
 LDSLHCRRCIIIGNGGVLANKSLGSRIDDYDIVIRLNSAPVKGFEKDVGSKTTLRITYPEGAMQORPEQYERDSL
 FVLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNNGLMGRGNIPT
 LGSVAVTMALDGCDEVAVAGFGYDMNTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLSSGI

5

SEQ ID NO: 29.

Mutante V162I D311H de Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3) de rata

MGLLVFVRNLLLALCLFLVLGFLYSAWKLHLLQWEDSNLILSLDSAGQTLGTEYDRLGFLCLKLDSKLP AELAT
 KYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPA
 LDSLHCRRCIIIGNGGVLANKSLGSRIDDYDIVIRLNSAPVKGFEKDVGSKTTLRITYPEGAMQORPEQYERDSL
 FVLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNNGLMGRGNIPT
 LGSVAVTMALHGCDEVAVAGFGYDMNTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLSSGI

SEQ ID NO: 30.

10 Mutante V162I D311H de Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (ST3Gal3 $\Delta 72$) de rata truncada

LATKYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIKGQDNLIKAILSVTKEYRL
 TPALDSLHCRRCIIIGNGGVLANKSLGSRIDDYDIVIRLNSAPVKGFEKDVGSKTTLRITYPEGAMQORPEQYERD
 SLFVLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNNGLMGRGN
 IPTLGSVAVTMALHGCDEVAVAGFGYDMNTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLSS
 GI

SEQ ID NO: 31.

Mutante V162I D311H de Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (MBP-ST3Gal3 $\Delta 72$) de rata truncada
 marcada con proteína de unión a maltosa

MKIEEGKLVIIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFWAHDRFGGYAQSG
 LLAEITPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
 NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFVLVDLIKXKHMNADTDYSIAEAFNKGETAM
 TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
 GAVALKSYYEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKDAQTSSNNNNN
 NNNNNNLGIEGRISEFGSLATKYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIK
 GQDNLIKAILSVTKEYRLTPALDSLHCRRCIIIGNGGVLANKSLGSRIDDYDIVIRLNSAPVKGFEKDVGSKTTL
 RITYPEGAMQORPEQYERDSLFLGFKWQDFKWLKYIVYKERSASDGFWKS VATRVPKEPPEIRILNPYFIQEA
 AFTLIGLPFNNGLMGRGNIPTLGSVAVTMALHGCDEVAVAGFGYDMNTPNAPLHYETVRMAAIKESWTHNIQRE
 KEFLRKLVKARVITDLSSGI

15

SEQ ID NO: 32.

Mutante V162I D311H de Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa III (MBP-SBD-ST3Gal3 Δ 72) de rata truncada marcada con dominio de unión a almidón-proteína de unión a maltosa

MKIEEGKLVWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFFWAHDRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAFNKGETAM
TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYYEELAKDPRIAATMENAQKEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSSNNNN
NNNNNNLIEGRISIEFGSIVATGGTTTTATPTGSGSVTSTSKTTATASKTSTSTSSCTTPTAVAVTFDLTATT
TYGENIYLVGSISQLGDWETSDGIALSADKYTSSDPLWVYVTVTLPAGESEYKFIRESDDSVESDPNREYTV
PQACGTSTATVTDTRWGSATKYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRRWARIREFVPPFGIK
GQDNLIKAILSVTKYRLTPALDSLHCRCIIIGNGGVLANKSLGSRIDYDIVIRLNSAPVKGEKDVGSKTTL
RITYPEGAMQRPEQYERDSLFLVLAGFKWQDFKWLKYIVYKERSASDGFWSVATRVKPEPPEIRILNPYFIQEA
AFTLIGLFPNGLMGRGNIPTLGSAVMTALHGCDEVAVAGFGYDMNTPNAPLHYETVRMAAIKESWTHNIQRE
KEFLRLKLVKARVITDLSSGI

5 SEQ ID NO: 33.

UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2) humana

MRRRSRMLLCFAFLWVLGIAYMYSGGGSALAGGAGGAGRKEDWNEIDPIKKDLHHSNGEKAQSMETLPPGK
VRWPDFNQEAAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQCQRKQWRVDLPATSVVITFHNEARS
ALLRTVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRREGLMRSRVRGADAAQKVLTFLLDS
HCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGDFDNLVFKWDYMTPEQRRSRQGNPVAP
IKTPMIAGGLFVMDKFYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEIIPCSRVGHVFRKQHPYTFPGSGT
VFARNTRRAAEVWMDYKFNFYAAVPSARNVPYGNIQSRLELRKKLSCKPFKWLNVPELVPELVRPDHQAIFGAL
QQGTNCLDTLGHFADGVVGVYECHNAGGNQEWALTKEKSVKHMDCLTIVVDRAVPSLIKLQGCRENDSRQKWEQI
EGNSKLRHVGSNLCDSRTAKSGGLSVEVCGPALSQQWKFTLNLQQ

SEQ ID NO: 34.

UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2) de *Drosophila*

MRRNIKLIVFVSIWFMVYVYFQSSTEKVENRALRLREVATAMQQYQDDSSAAAASARQWAPAGGGAGPGAA
AGAAGSGADDPGGNVILIGSVKDFERNAVHGLKLNGLIVALEETSQGLSGGTGGPGGRLPVAPSGRGTEVEYFNEA
GYIRAGALRNGEDPYIRNRFNQEASDALPSNRDIPDTRNPMCRTKKYREDLPETSVIITFHNEARSTLLRTIVSV
LNRSPEHLIREIVLVDDYSNHPEDGLELAKIDKVRVIRNDKREGLVRSRVKGADAAVSSVLTFLDSHVECNEML
EPLLERVREDPTRVVCVIDVISMDNFQYIGASADLKGDFDNLVFKWEYLSPSERAMRHNDPTTAIRTPMIAGG
LFVIDKAYFNKLGKYDMKMDVWGGENLEISFRVWQCGGSLEIIPCSRVGHVFRKRHPYTFPGSGNVFARNTRRA
AEVWMDYKQHYNAVPLAKNIPFGNIDRLALKEKLHCKPFKWLNVPELVPELVRPDHQAIFGAL
GHLIDGTVGIFPCHNTGGNQEWAFTRGEIKHDDLCLTLVTFARGSQVVLKACDDSENQRWIMREGGLVRHYKIN
VCLDSRDQSQQGVSAQHCSALGTQRWSFGKYA

10

SEQ ID NO: 35.

UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2) de ratón

MRRRSRMLLCFALLWVLGIAYMYSGGGSALAAGGGGAGRKGDWNDIDS I KKKDLHHSRGDEKAQGVETLPPGKV
RWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQCQRKQWRVLDLPATSVVITFHNEARSA
LLRTTVSVLKRSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRREGLMRSRVRGADAAQAKVLTFLDSH
CECNERWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGFDWNLVFKWDYMTPEQRRSRQGNPVAPI
KTPMIAGGLFVMDKLYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEIIPCSRVGHVFRKQHPYTFPGGSGTV
FARNTRRAAEVWMDYKHFYAAVPSARNVPYGNIQSRLELRKKLGCKPFKWYLDNVYPELRVPDHDQDIAFGALQ
QGTNCLDTLGHFADGVVGIYECHNAGGNQEWALTKEKSVKHMDLCLTVVDRSPGSLIRLQGCRENDSRQKWEQIE
GNSKLRHVGSNLCLDSRTAKSGGLSVEVCGPALSQQWKFSNLQQ

SEQ ID NO: 36.

UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ 51) humana truncada

KKKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQC
QRKQWRVLDLPATSVVITFHNEARSALLRTTVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRR
EGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGFDW
NLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKLYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEI
IPCSRVGHVFRKQHPYTFPGGSGTVFARNTRRAAEVWMDYKHFYAAVPSARNVPYGNIQSRLELRKKLSCKPF
KWYLENVYPELRVPDHDQDIAFGALQQGTNCLDTLGHFADGVVGVYECHNAGGNQEWALTKEKSVKHMDLCLTVVD
RAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCLDSRTAKSGGLSVEVCGPALSQQWKFTLNQQ

5 SEQ ID NO: 37.

UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ 1-51 Δ 445-571) humana truncada

KKKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQC
QRKQWRVLDLPATSVVITFHNEARSALLRTTVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRR
EGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGFDW
NLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKLYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEI

IPCSRVGHVFRKQHPYTFPGGSGTVFARNTRRAAEVWMDYKHFYAAVPSARNVPYGNIQSRLELRKKLSCKPF
KWYLENVYPELRVPDHDQD

SEQ ID NO: 38.

10 Forma alterna de UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ 51) humana truncada

MSKKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHD
QCQRKQWRVLDLPATSVVITFHNEARSALLRTTVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRND
RREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGF
DWNLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKLYFEELGKYDMMMDVWGGENLEISFRVWQCGGSL
EIPCSRVGHVFRKQHPYTFPGGSGTVFARNTRRAAEVWMDYKHFYAAVPSARNVPYGNIQSRLELRKKLSCK
PFKWYLENVYPELRVPDHDQDIAFGALQQGTNCLDTLGHFADGVVGVYECHNAGGNQEWALTKEKSVKHMDLCLTV
VDRAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCLDSRTAKSGGLSVEVCGPALSQQWKFTLNQQ

SEQ ID NO: 39.

Forma alterna de UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ 1-51 Δ 445-571) humana truncada

MSKKKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHD
 QCQRKQWRVDLPATSVVITFHNEARSALLRTTVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRND
 RREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGF
 DWNLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWGGENLEISFRVWQCGGSL
 EIIPCSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVPYGNIQSRLELRKKLSCK
 PFKWYLENVYPELRVPDHDQD

SEQ ID NO: 40.

UDP-N-acetilgalactosaminiltransferasa 2 (MBP-GalNAc-T2 Δ 51) humana truncada marcada con proteína de unión a maltosa

MKIEEGKLVIIWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSG
 LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
 NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNAGAKAGLTFLVDLIKKNHMNADTDYSIAEAAFNKGETAM
 TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
 GAVALKSYYEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSSNNNN
 NNNNNNLGIEGRISEFGSKKKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVE
 SDKLRMDRAIPDTRHDQCQRKQWRVDLPATSVVITFHNEARSALLRTTVSVLKKSPPHLIKEIILVDDYSNDPED
 GALLGKIEKVRVLRNDRREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINM
 DNFQYVGASADLKGGFDWNLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWGG
 ENLEISFRVWQCGGSL EIIPCSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVPY
 GNIQSRLELRKKLSCKPFKWYLENVYPELRVPDHDQDIAFGALQQGTNCLDTLGHFADGVVGVYECHNAGGNQEWA
 LTKEKSVKHMDLCLTVVDRAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCCLDSRTAKSGGLSVEVCGPA
 LSQQWKFTLNLQQ

5

SEQ ID NO: 41.

UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1) humana

MASKSWNLFTFLCGSAIGFLLCSQLFSILLGEKVDTPQNVLHNDPHARHSDDNGQNHLEGQMNFNADSSQHKDE
 NTDIAENLYQKVRILCWVMTGPQNLEKKAKHVKATWAQRCNKVLFMSSEENKDFPAVGLKTKEGRDQLYWKTIKA
 FQYVHEHYLEDADWFLKADDDTYVILDNLRWLLSKYDPEEPIYFGRRFKPYVKQGYMSGGAGYVLSKEALKRFVD
 AFKTDKCTHSSSIDLALGRCMEIMNVEAGDSRDTIGKETTFHPFVPEHHLIKGYLPRTFWYWNYNYPFVEGPGC
 CSDLAVSFHYVDSTTMYELEYLRYHLRPYGYLYRYQPTLPERILKEISQANKNEDTKVKLGNP

SEQ ID NO: 42.

10 UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1) de *Drosophila*

MTANSLGRSILNEGRSNKRSFVSLIVGLIVGFCLAEFVYSTPERSEFMPYDGHHRGVDNDAHSHDMMMSGP
 EQDVGGHEHVHENSTIAERLYSEVRVLCWIMTNPSNHQKKARHVKRTWKGKCNKLIFMSSAKDDELDAVALPVGE
 GRNNLWGKTKYAYKYYEHHINDADWFLKADDDTYTIVENMRYMLYPSPETPVYFGCKFKPYVKQGYMSGGAGY
 VLSREAVRRFVVEALPNPKLCKSDNSGAEDVEIGKCLQNVNVLGDSRDSNGRGRFFPFVPEHHLIPSHTDKKFW
 YWQYIFYKTDEGLDCCSDNAISFHYVSPNQMYVLDYLIYHLRPYGIINTPDALPNKLAUGELMPEIKEQATESTS
 DGVSKRSAETKTQ

SEQ ID NO: 43.

UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1) de ratón

MASKSWLNFLVFLCGSAIGFFLCSQLLSILLREEAAIQPNMLHNDPHARHSDDNHSHLKGQMFNADSSQHKDE
NIDVAENLYQVKILCWVMTSPQNLEKKAKHV KATWAQRCNKVLFMSSEENQDFPTVGLKTKEGREQLYWKTIKA
FQYVHDHYLEDADWFMKADDDTYVIVDNLRWLLSKYDPEQPIYFGRRFKPYVKQGYMSGGAGYVLSKEALRRFVN
AFKTEKCTHSSSIEDLALGRCMEIINVEAGDSRDTIGKETFHPFVPEHHLIKGYLPKTFWYWNYNYPPIEGPGC
CSDIAVSFHYVDGTTMYELEYLVYRLRPYGCYLYRYQPALPENILKEINQVNRKEDTKIKLGNP

SEQ ID NO: 44.

UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1) de rata

MASKSWLNFLTFLCGSAIGFFLCSQLLNILLQEADVQPNMLHNDPHARHSDDSGHNHLKGQMDFNADSSQHKDE
NTDVAENLYQVKVLCWVMTSPQNLEKKAKHV KATWAQRCNKVLFMSSEENKDFPTVGLKTKEGREQLYWKTIKA
FQYVHDHYLEDADWFMKADDDTYVILDNLRWLLSKYNPEQPIYFGRRFKPYVKQGYMSGGAGYVLSKEALRRFVD
AFKTEKCTHSSSIEDLALGRCMEI IKVEAGDSRDP TGKETFHPFVPEHHLIKGYLPKTFWYWNYNYPPEVEGPGC
CSDIAVSFHYVDSTTMYELEYLVYHLRPYGYLYRYQPALPENILKEINQVNRKEDTKIKLGNP

5 SEQ ID NO: 45.

Mutante A383T de UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1) de *Drosophila*

MTANSLLGRSILNEGRSNKRSFVSLIVGLIVGFCLAE LFVYSTPERSEFMPYDGHRHGDVNDAAHSHDMMMSGP
EQDVGGHEHVHENSTIAERLYSEVRVLCWIMTNPSNHQKKARHV KRTWGKRCNKLI FMSSAKDDELDAVALPVGE
GRNNLWGKTKEAYKYIYEHHINDADWFLKADDDTYTIVENMRMYLPYSPETPVYFGCKFKPYVKQGYMSGGAGY
VLSREAVRRFVVEALPNPKLCKSDNSGAEDVEIGKCLQNVNVL AGDSRDSNGRGRFFPFVPEHHLIPSHTDKKFW
YWQYIFYKTDEGLDCCSDNAISFHYVSPNQMYVLDYLIYHLRPYGIINTPDALPNKLAVGELMPEIKEQATESTS
DGVSKRSTETKTQ

SEQ ID NO: 46.

10 Mutante A383T de UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal- T1 Δ31) de *Drosophila* truncada

GFCLAE LFVYSTPERSEFMPYDGHRHGDVNDAAHSHDMMMSGPEQDVGGHEHVHENSTIAERLYSEVRVLCWIM
TNPSNHQKKARHV KRTWGKRCNKLI FMSSAKDDELDAVALPVGEGRNNLWGKTKEAYKYIYEHHINDADWFLKAD
DDTYTIVENMRMYLPYSPETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFVVEALPNPKLCKSDNSGAEDV
EIGKCLQNVNVL AGDSRDSNGRGRFFPFVPEHHLIPSHTDKKFWYQYIFYKTDEGLDCCSDNAISFHYVSPNQ
YVLDYLIYHLRPYGIINTPDALPNKLAVGELMPEIKEQATESTSDGVSKRSTETKTQ

SEQ ID NO: 47.

15 Forma alterna de mutante A383T de UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal- T1 Δ31) de *Drosophila* truncada

MGFCLAE LFVYSTPERSEFMPYDGHRHGDVNDAAHSHDMMMSGPEQDVGGHEHVHENSTIAERLYSEVRVLCWI
MTNPSNHQKKARHV KRTWGKRCNKLI FMSSAKDDELDAVALPVGEGRNNLWGKTKEAYKYIYEHHINDADWFLKA
DDDTYTIVENMRMYLPYSPETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFVVEALPNPKLCKSDNSGAED
VEIGKCLQNVNVL AGDSRDSNGRGRFFPFVPEHHLIPSHTDKKFWYQYIFYKTDEGLDCCSDNAISFHYVSPNQ
MYVLDYLIYHLRPYGIINTPDALPNKLAVGELMPEIKEQATESTSDGVSKRSTETKTQ

SEQ ID NO: 48.

UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal- T1 Δ50) de *Drosophila* truncada

PYDGHHRHGDVNDAAHSHDMMMSGPEQDVGGHEHVHENSTIAERLYSEVRVLCWIMTNPSNHQKKARHVKRTWGK
RCNKLIFMSSAKDDELDAVALPVGEGRNNLWGKTKEAYKYIYEHHINDADWFLKADDDTYTIVENMRYMLYPYSP
ETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFVVEALPNPKLCKSDNSGAEDVEIGKCLQNVNVLGDSRDS
NGRGRFFPFVPEHHLIPSHTDKCFWYQYIFYKTDEGLDCCSDNAISFHYVSPNQMYVLDYLIYHLRPGYGIINTP
DALPNKLAVGELMPEIKEQATESTSDGVSKRSAETKTQ

SEQ ID NO: 49.

Mutante A383T de UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1 Δ50) de *Drosophila* truncada

PYDGHHRHGDVNDAAHSHDMMMSGPEQDVGGHEHVHENSTIAERLYSEVRVLCWIMTNPSNHQKKARHVKRTWGK
RCNKLIFMSSAKDDELDAVALPVGEGRNNLWGKTKEAYKYIYEHHINDADWFLKADDDTYTIVENMRYMLYPYSP
ETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFVVEALPNPKLCKSDNSGAEDVEIGKCLQNVNVLGDSRDS

5

NGRGRFFPFVPEHHLIPSHTDKCFWYQYIFYKTDEGLDCCSDNAISFHYVSPNQMYVLDYLIYHLRPGYGIINTP
DALPNKLAVGELMPEIKEQATESTSDGVSKRSTETKTQ

SEQ ID NO: 50.

Forma alterna de mutante A383T de UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1 Δ50) de *Drosophila* truncada

MSPYDGHHRHGDVNDAAHSHDMMMSGPEQDVGGHEHVHENSTIAERLYSEVRVLCWIMTNPSNHQKKARHVKRTW
GKRCNKLIFMSSAKDDELDAVALPVGEGRNNLWGKTKEAYKYIYEHHINDADWFLKADDDTYTIVENMRYMLYPY
SPETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFVVEALPNPKLCKSDNSGAEDVEIGKCLQNVNVLGDSR
DSNGRGRFFPFVPEHHLIPSHTDKCFWYQYIFYKTDEGLDCCSDNAISFHYVSPNQMYVLDYLIYHLRPGYGIIN
TPDALPNKLAVGELMPEIKEQATESTSDGVSKRSTETKTQ

10

SEQ ID NO: 51.

Mutante A383T de UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (MBP-Core-1-Gal-T1 Δ50) de *Drosophila* truncada marcada con proteína de unión a maltosa

MKIEEGKLVIIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFVLVDLIKNKHMNADTDYSIAEAFNKGETAM
TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYYEELAKDPRIAATMENAQKEIMPNIPOMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSNNNN
NNNNNNLIGIEGRISEFGSPYDGHHRHGDVNDAAHSHDMMMSGPEQDVGGHEHVHENSTIAERLYSEVRVLCWIMT
NPSNHQKKARHVKRTWGKRCNKLIFMSSAKDDELDAVALPVGEGRNNLWGKTKEAYKYIYEHHINDADWFLKADD
DTYTIVENMRYMLYPYSPETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFVVEALPNPKLCKSDNSGAEDVE
IGKCLQNVNVLGDSRDSNGRGRFFPFVPEHHLIPSHTDKCFWYQYIFYKTDEGLDCCSDNAISFHYVSPNQMY
VLDYLIYHLRPGYGIINTPDALPNKLAVGELMPEIKEQATESTSDGVSKRSTETKTQ

15

SEQ ID NO: 52.

Core-1-Gal-T1 chaperona 1/COSMC humana

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERMELSKSFRVYCIILV
KPKDVS LWAAVKETWTKHCDKAEPFSSSENVKVFESINMDTNDMWLMMRKAYKYAFDKYRDQYNWFFLARPTTFAI
IENLKYFLLKKDPSQPFYLGHITIKSGDLEYVGMEGGIVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAV
CLKYAGVFAENAEDADGKDVFNTKSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMYGVYRLRAFG
HIFNDALVFLPPNGSDND

SEQ ID NO: 53.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (ST3Gal-1) humana

MVTLRKRTLKVLTFVLVLFIFLTSFFLNYSHTMVATTWFPKQMVLELSENLRLIKHRPCTCTHCIGQRKLSAWFD
ERFNQTMQPLLTQAONALLEDDTYRWLRLQREKKPNNLNDTIKELFRVVPGNVDPMLEKRSVGCRRCAVVGNSGN
LRESSYGPEIDSHDFVLRMNKAPTAGFEADVGTKTTHHLVYPESFRELGDNVSMILVFPKTTIDLEWVVSATTTGT

ISHTYIPVPAKIRVKQDKILIYHPAFIKYVFDNWLQGHGRYPSTGILSVIFSMHVCDEVLDLYGFGADSKGNWHHY
WENNPSAGAFRKTGVHDADFESNVTATLASINKIRIFKGR

SEQ ID NO: 54.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (ST3Gal1) porcina

MAPMRKKSTLKLTLVLFIFLTSFFLNYSHTVVTAWFPKQMVIELSENFKKLMKYPYRPTCTRCIEEQRVSA
WFDERFNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNLNDTIRELFQVVPGNVDPLLEKRLVSCRRCAVVGNS
SGNLKESYYPQIDSHDFVLRMNKAPTEGFEADVGSKTTHHFVYPESFRELAQEVSMILVFPKTTIDLEWVISATT
TGTISHTYVPVPAKIKVKEKILIYHPAFIKYVFDNWLQGHGRYPSTGILSVIFSLHICDEVLDLYGFGADSKGNW
HHYWENNPSAGAFRKTGVHDGDFESNVTILASINKIRIFKGR

SEQ ID NO: 55.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (ST3Ga1-1) de ratón

MRRKTLKYLTFLLFIFLTSFVLNYSNTGVPSAWFPKQMLLELSENFRFKIKSQPCTCRHCISQDKVSYWFDQRF
NKTMQPLLTVHNALEEDTYRWLRLQREKPNLSDTVKELFRLVPGNVDPMLNKRLVGCRRCAVVGNSGNLKD
SSYGPEIDSHDFVLRMNKAPTGVFEADVGSRTTHHLVYPESFRELGENVMVLVFPKTTDLQWVISATTTGTITH
TYVPVPPKIKVQEKILIYHPAFIKYVFDNWLQGHGRYPSTGILSVIFSIHICDEVLDLYGFGADSKGNWHHYWEN
NPSAGAFRKTGVHDGDFEYNITTTLAANKIRIFKGR

SEQ ID NO: 56.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (ST3Gal-1) de rata

MVMNRKRTLKYLTFLLFIFLTSFVLNYSNSGVPSAWFPKQMVLEFSENFRKFIKSQPCTCRHCISQGVSYWFD
QRFNKTMQPLLTVHNALEEDTYRWLRLQREKPNLSDTVKELFRLVPGNVDPMLNKRLVGCRRCAVVGNSGN
LKDSSYGPEIDSHDFVLRMNRAPTGVFEADVGSRTTHHLVYPESFRELGENVMVLVFPKITDLQWVISATTTGT
ITHTYVPVPPKIKVQEKILIYHPAFIKYVFDNWLQGHGRYPSTGILSVIFSIHICDEVLDLYGFGADSKGNWHHY
WENNPSAGAFRKTGVHDGDFEYNVTTTTLAANKIRIFKGR

SEQ ID NO: 57.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (ST3Gal-1) de chimpancé

MVTLRKRTLKVLTFVLVLFIFLTSFFLNYSHTMVATTWFPKQMVLELSENKRLIKHRPCTCTHCIGQRKLSAWFD
ERFNQTVQPLLTQAQNALLEDITYRWLRLQREKKPNNLNDTIKELFRVVPGNVDPMLEKRSVGCRCRAVVGNSGN
LRESSYGPEDSHDFVLRMNKAPTAGEADVGKTTTHHLVYPESFRELGDNVSMILVPFKTIDLEWVVSATTGT
ISHTYTPVLVKIRVKQDKILIYHPAFIKYVFDNLQGHGRYPSTGILSVIFSMHVCDEVDLYGFGADSKGNWHHY
WENNPSAGAFRKTGVHDADFESNVTATLAANKIRIFKGR

SEQ ID NO: 58.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (ST3Gal-1 Δ45) de cerdo truncada

ELSENFKKLMKYPYRPTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLNDTI
RELQVVPGNVDPLLEKRLVSCRCRAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPTEGFEADVGSKTTHHFVYP

ESFRELAQEVSMILVPFKTTDLEWVISATTTGTISHTYVPVPAKIKVKEKILIYHPAFIKYVFDRLQGHGRYP
STGILSVIFSLHICDEVDLYGFGADSKGNWHHYWENNPSAGAFRKTGVHDGDFESNVTILASINKIRIFKGR

5

SEQ ID NO: 59.

Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (ST3Gal-1 Δ56) de cerdo truncada

YPYRPTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLNDTIRELQVVPGNV
DPLLEKRLVSCRCRAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPTEGFEADVGSKTTHHFVYPESFRELAQEV
MILVPFKTTDLEWVISATTTGTISHTYVPVPAKIKVKEKILIYHPAFIKYVFDRLQGHGRYPSTGILSVIFSL
HICDEVDLYGFGADSKGNWHHYWENNPSAGAFRKTGVHDGDFESNVTILASINKIRIFKGR

SEQ ID NO: 60.

- 10 Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (MBP-ST3Gal-1 Δ45) de cerdo truncada marcada con proteína de unión a maltosa

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVHEHPDKLEEKFPQVAATGDGPDIIIFWAHDRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAFNKGETAM
TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYYEELAKDPRIAATMENAQGEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSNNNN
NNNNNLGIEGRISEFGSELSNFKKLMKYPYRPTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYK
WWLRLQREKQPNNLNDTIRELQVVPGNVDPLLEKRLVSCRCRAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPT
EGFEADVGSKTTHHFVYPESFRELAQEVSMILVPFKTTDLEWVISATTTGTISHTYVPVPAKIKVKEKILIYHP
AFIKYVFDRLQGHGRYPSTGILSVIFSLHICDEVDLYGFGADSKGNWHHYWENNPSAGAFRKTGVHDGDFESN
VTILASINKIRIFKGR

SEQ ID NO: 61.

- 15 Gal beta-1,3-GalNAc alfa-2,3-sialiltransferasa 1 (MBP-SBD-ST3Gal-1 Δ45) de cerdo truncada marcada con dominio de unión a almidón-proteína de unión a maltosa

MKIEEGKLVWINGDKGYNGLAEVGKKFEKDTGIIKVTVEHPDKLEEKFPQVAATGDGPDIIIFWAHDRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAFNKGETAM
TINGPWAWSNIDTSKVNYGVTVLPTTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQGEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNN
NNNNNNLGIEGRISEFGSIVATGGTTTTATPTGSGSVTSTSKTTATASKTSTSTSTSTCTTPTAVAVTFDLTATT
TYGENIYLVGSISQLGDWETSDGIALSADKYTSSDPLWYVTVTLPAGESEFYKFIRIESDDSVEWESDPNREYTV
PQACGTSTATVTDTWGRSELSNFKKLMKYPYRPTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYK
WWLRLQREKQPNLNDTIRELFQVVPGNVDPLLEKRLVSCRRCAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPT
EGFEADVGSKTTTHHFVYVPEFRELAQEVSMILVPFKTTDLEWVISATTTGRISHTYVVPVPAKIKVKKEKILIIYHP
AFIKYVFDRLWLGHRYPSTGILSVIFSLHICDEVLDYGFAGDSKGNWHHYWENNPSAGAFRKTGVHDGDFESNV
TTILASINKIRIFKGR

SEQ ID NO: 62.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) humana

MRSCLWRCRHLSQGVQWSLLLAFLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKPKSQAPTRARRTTI
YAEPVPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGR
TEAQSWKSQDTKTKTGNGGQTRKLTASRTVSEKHQGAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIP
PKEKKPQATPPAPFQSPTTQRNQLKAAFPKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKFLPN
LTLFLDSRHFNQSEWDRLEHFAPPPFGFMELNYSLVQKVTRFPVPPVQQQLLLASLPAGSLRCITCAVVGNGGIILN
NSHMQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLILGNRGFKNVPLGKDVRYLHFLEGTDRD
YEWLEALLMNQTVMSKNLFWFRHRPQEAFFREALHMDRYLLHLPDFLRVMKNRFLRSKTLTGAHWRIYRPTTGALL
LLTALQLCDQVSAYGFITEGHERFSDHYDTSWKRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

5 SEQ ID NO: 63.

N-acetilgalactosamina-alfa2,6-sialiltransferasa I (ST6GalNAcI) humana, de pcDNA3.1(+)-ST6GalNAcI-N1C1#1

MRSCLWRCRHLSQGVQWSLLLAFLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKPKSQAPTRARRTTI
YAEPVPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGR
TEAQSWKSQDTKTKTGNGGQTRKLTASRTVSEKHQGAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIP
PKEKKPQATPPAPFQSPTTQRNQLKAAFPKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKFLPN
LTLFLDSRHFNQSEWDRLEHFAPPPFGFMELNYSLVQKVTRFPVPPVQQQLLLASLPAGSLRCITCAVVGNGGIILN
NSHMQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLILGNRGFKNVPLGKDVRYLHFLEGTDRD
YEWLEALLMNQTVMSKNLFWFRHRPQEAFFREALHMDRYLLHLPDFLRVMKNRFLRSKTLTGAHWRIYRPTTGALL
LLTALQLCDQVSAYGFITEGHERFSDHYDTSWKRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

SEQ ID NO: 64.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de ratón

MTRYCRGLSQRQAFLLLTVLALLFILLFVVKDPRAKDSRRQFILNNDSSAQEILQKAEPQGPIMTLSPRVHNKEA
TSVSSKDLKKQEREAVQGEQAEQKEKRKLETIRPAPENPQSKAEPAAKTPVSEHLDKVPRTPGALSTRKTPMATG
AVPAKKKVQATKSPASSPHPTTRRRQRLKASEFKSEPRWDFEEYSLDMSSLQTNCASVKIKASKSPWLQNI
LPNITLFLDSGRFTQSEWNRLEHFAPPFGFMELNQSLVQKVVTFRPPVRQQQLLLASLPTGYSKCITCAVVGNGG
ILNDSRVGREIDSHDYVFRLSGAVIKGYEQDVGTRTSFYGFTAFSLTQSILILGRRGFQHVPLGKDVRYLHFLEG
TRNYEWLEAMFLNQTLAKTHLSWFRHRPQEAFRNALDLDRYLLHHPDFLRYMKNRFLRSKTLDTAXWRIYRPTTG
ALLLLTALHLCDKVSAYGFITEGHERFSDHYDTSWKRLIFYINHDFRLERMVWKRLHDEGI IWLYQRPQSDKAK
N

SEQ ID NO: 65.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de pollo

MGFLIRRLPKDSRIFRWLLILTVFSFIITSFSALFGMEKSIQRQLKIYQSIHMLQVDTQDQQGSNYSANGRISK
VGLERDIAWLELNTAVSTPSGEGKEEQKTKVPKPAKVEEAKKVTVPFPEVMGITNTTASTASVVERTKEKTTA

RPVPGVGEADGKRRTTIALPSMKEDKEKATVKPSFGMKVAHANSTSKDKPKAEEPASVKAIRPVTQAATVTEKKK
LRAADFKEPQWDFDDEYILDSSSPVSTCSESVRAKAKSDWLRLDLPNITLFDKSYFNVSEWDRLEHFAPPY
GFMELNYSLVEEVSRLPPNPHQQQLLANSSSNVSTCISCAVVGNGGILNNSGMGQEQIDSHDYVFRVSGAVIKGY
EKDVGTKTSFYGFTAYSLVSSLQNLGHKGFKKIPQKGHIRYIHFLAVRDYEWLKALLLDKDIRKGLNYYGRRP
RERFDEDFTMNKYLVAHPDFLRYLKNRFLKSKNLQKPYWRLYRPTTGALLLLTALHLCDRVSAVGYTEGHQKYS
DHYYDKWKRLVFYVNHDFNLEKQVWKRLHDENIMKLYQRS

SEQ ID NO: 66.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de rata

MENRCRGLSQGQTFLLLTGLMLLFILPSVVKPESTRVSRQFIEDNESSLQGVPOKPAPOGPIVTLTPTVHNKKT
SVRTKWVELQKQDRATARGERGEVEKKLQAIRLAPENPKGAPEVKTPASKHLDKLPRATGALSTRKTPMATG
AAPAKKKVQPTPTPASFPHTTQRRQRLKASDFKSEPRWDFEEYSLDGGSLQTLPWFLKITVSHSPWVQNI
PNITLFLDSGRFNQSEWYRLEHFTPPFGFMELNQSLVQKVVSFRPPVPQQQLLLASLPTRNLTCTITCAVVGNGGI
LNNSRMGQEQIDSHDYVFRLSGTVIKGYEQDVGTRTSFYGFTAFSLGQSILNLGSSRFSACAFLGKDVRYLHFLEG
TRDYEWLEAMFLNRTMANTKLYWFRHRPQEAFFREALDLDRYFLVHPDFLRYMKNRFLRSKTLDTAHWRLYRPTTG
ALLLLTALHLCDKVSAYGFITQGHERFSDHYDTSWKRLIFYINHDFALERTVWKRLHDEGI IQLYQRP

SEQ ID NO: 67.

10 alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de chimpancé

MRSCLWRCRHLSSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSREPAPENNALNTQTQPKAHTAGDRGKEANQAP
PEEQDKVPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGRTQAQSWKSQDTKTTQGNNGGQTRKLMASRTVSEK
HQGKAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPPTQRNQLKAANFKS
EPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFPLNLTFLDSRHPNQSEWDRLEHFAPPFGFMELNYS
LVQKVVTFRPPVPQQQLLLASLPAGSLRCITCAVVGNGGILNNSHMGQEQIDSHDYVFRLSGVLIKGYEQDAVDRT
SFYGFTAFSLTQSLILGNRGFKNVPLGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFFREAL
HMDRYLLHHPDFLRYMKNRFLRSKTLDTGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSW
KRLIFYTNHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

SEQ ID NO: 68.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de vaca

MATCPACGRLRLSLRPGSRKNEHQONIKERSPELLQONATSQAPTPTRRRATTHVESVQGTRTRDTHPKATTLTADQ
 RRRVQTSKARAEPPGRVPTPIGKAAPQTQASKDTRADTLPPMAGGGGVASSRTEAPSLNSQNPRMTKSGDRKAR
 PTGPRAVPTKLKRDSPSPATQRSQKLKATNFKSEPQWDFEEYSLEVGGLOTGTRTRDTHPKATTLTADQRRRVQT
 SKARAEPPGRVPTPIGKAAPQTQASKDTRADTLPPMAGGGGVASSRTEAPSLNSQNPRMTKSGDRKDRAQPTRS
 SAPLQSPATQRSQKLKATNFKSEPQWDFEYSLEVGGLOTTCPSVKIKASKSPWLRTLFLPNLTLFLDSGHFN
 QSEWDRLEHFAPPFGFMELNFSLVQKVVARFPPVPPQQQLLASLPAGSSRCISCAVVGNGGILNNSHVGPIDSH
 DYVFRLSGAIKGYEHDVGTRTSFYGFTAFSLTQSLILGSRGFPHAPLGQDVRYLHFLEGTDRDYEWLEALLNR

TLTSRNLSWFRRRPQEAFOEALQLDRYLLLHPDLLRYMKNRFLRSKTLNTHAWRIYRPTTGALLLLTALQLCDQV
 SAYGFITEGHERFSDDHYDKSWKRTIFYTNHDFKLERALWKRLHDEGIIRLYQRPITSKPTI

SEQ ID NO: 69.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de perro

MKIRLCRLSHLDQTTQWFLLSAVLTFFVFIPLPSFVKEPNTKPFRRHQHIRNEERSPESLHEVTGQAPTAGDRTMAP
 AVASAEKSVWAPQALATAYAAQEDTRAEALEATSQPTSMRGQERKDTMPDTLPPRAQDKGVASDRTGLPSVMS
 QDTRTIKGRDQKEEPTATRKVTPGPQSKARTTRTPVPESQAKALTPGAGPMGRTRKGATTAAAPHRDTARAT
 PPSTSSQGRTRRSPLRAANFKSEPRWDFEEQYSFDTEALQPTCPDSVKVKASKSPWLQNLFLANLTLFLDSSR
 FSQREWDRLEHFAPPFGFMELNQSLVQKGVTRFPPVPPQQQLLRASLPAGSSQCITCAVVGNGGILNDSRVGQID
 GHHDVFRLSGAVTKGYEQDVGTRTSFYGFTAFSITQSLALGGRGFPHVPLGKDVRYLHFLEGTDRDYEWLEALL
 NQTLVKSSLSWFRRRPQEAFRDTLQLDRYLLLHPDFMRYMKNRFLRSKTLNNIHWRIYRPTTGALLLLTALHLCD
 QVSAYGFITEGHERFSDDHYDKSWKRTIFYTNHDFPLERTLWKRLHDEGIIRLYQRPISKPKMVQTRPQAPPDS
 PGDTGTKEENPHCRRLDLAIQRHPHFRALFDLSTPVLSSGLSFTQELWDSLSQHKAPYGWQGLSHQAINSTLSL
 LNGSESSQLFATSRKPLSSCIRCAVVGNGGILNGSRQGLNIDAHDYVFRQTLKTERGGRVPARRGPPGRMSPEVR
 RCGDGLDGSDDRQHSYFPDGNVSVQGHAAKIKVQPPLTTPALGLPPPSLSCLPDDLRYIFIPSDIRDYVM
 LRSAILGVPVPEPDKGDRSPSHSPATSASYIGPVRGIPSYIAPYVPRFLKEPPFPQPPTVLVGGKPHVPTVLA
 ATFYTPFFRYAQPWPYRPHRKRPSLSGARSSRGLVWVEPNLIWDNPNYKGPGRSLDLELGLYPVSAYGFITSNYW
 KPSDHYFERVKKPLIFYVNHDLSEASLWRDLHKAGILQLYQR

5

SEQ ID NO: 70.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de pez cebra

MDYFLWPSVLFLLLAGIVAAFLVKNKPAHHQSSVVVEEEKKADAGVSSQEQHCEHWDESCQLKTVEKRLQTEDK
 TLEKAKDLNRDDLEEDALVQPTAREEEPGIRSALREYAFSPDAESKPLKYMAGMLRQCQLEKMMTKEELEEEQSK
 TIISSSILSPLGNTNLFKTQKHLNVERKVNITPIPVLYKKNFTKLPVWDFEDVYLRDSNARKPTCPKSLHNTED
 PEFKESVLPDIQLWLKYGLNMSEWNRLAHFNNPFGFMEYNYNEIKRAVDLIPKPRSSILLVPKSGDKGCIRCA
 VVGAGGILNNSKMGREIDSHDYVFRVNGAVTKGYEEDVGNRTSVYVHTAFSLYATILTLKKYGFHNIPQDEGIKY
 VMIPEGLRDFEWLQGLLQGAANGSFKGVRPLNFFNGHFNESEFNVLHPDFLRYIRNRFMPSKQMQGNYWAMYRP
 TNGAFALFLAIHTCDMVNAYGFITEDHHKYSNYEYKFKKTSVIFYINHGYGLEIKTWKKLHDSGIIRLFQRH

SEQ ID NO: 71.

10

2 alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de pez cebra

MFLLRIFLVTTFIASLPLFIFVTFYNGTSTLQIKWRQYYYWLNQGGVAGSLLTIVEPEPTNSSALTDILANTTTM
RLDPFETLSTPLPIMDKHKFTSLPHWKFDLLYRLDPHFKEPSECATSLQNSSNPTFKKKFIPNIQLFLQSDHLNMS
EWNRLYHFNNPFGYMGMLNYTAIKAAVETIHKPASSQLQVHPGVKDGCIKCAVVGASGILNGSKLGKEIDSSDYV
FRMNAAIIAGHEEDVGKRTSVYVHTAHSIIQSLMIHKKRGFKQIPTDKDIKYVLIPEGPRDYNFLESMLKNRKIP
SGAYRGRTPRKYYSGHFNNESSYIILHPDPLRYVRNRLRAKQLKTKRWWVVRPTNGAFTLLLAMHTCDIVRVYGF
CTADYRKYPSSYYDTKHTKLVFYGNHDYRLEMKTWKKFHDDKLIWMYLGKSN

SEQ ID NO: 72.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de *Xenopus*

MRPRWLRAVGITIGALTVCSLFYGHYYSSLVSVARSPQGGVGYDLPSKGLEWIPAVPHGRERNRELAETLRVQRP
SKVPGTCPTSLQLRVKNDSEYFKTIFNFEVPLLLWNSHLTENNWDTLKRPVPYGWKDLPREVASALKLLNDSAN
KAMFERQGPQKCIRCAVVGNGGILNGSKKGKEIDGHDYVFRNLNGAVIKGFEDVGTGKTSFYGFTVNTMKNLSISY
QEHGFTETPKGKDLHYIFIPSDLRDYVMLRSGILGVKVPESGYDEGDKPSEYFGPKPSPKKFKMLHPDFLLYTRDR
FLKSDILNTEYASLYMPSTGGMLMLLTALHSCDQVSAYGFITPDYNRYSDHYERQKVELEVNDANHMLMEMQLWG
RLHDRGIIKLYKR

5 SEQ ID NO: 73.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de *Takifugu*

SSGSLESMKLSDDQIKVDSRTTATPAYVKSSGPRQHTEHREAREQEIQNVSRSTPRSRGANQTRIEVPVKKPQQPT
QETRITDPPFIDGTMYSEEIPQTTCPDGIKTRRTNTEFGNFLNNVPVLQWARHATHEEYGRLSQHRGAHGWWG
VDYNTLVDAISVLNSSANWQMLDDWKDRSNNSECIRCAVVGNGGILKDSKRGAEIDSHHYVFRNGAVIKGFED
VGSRTTHYTFSATLNMNSIRAYAGVGFKAPPVSKETRYIFLPENDRDYLLVKAATHTLVERGPQRNQKPAFFG
KDVSAEKVKMYHPDFVRYLRNRLRSATLTKYKDIYRSTGAVMLLVALHTCDKVSAYGFMTDYMKYSDHYD
KKQKSVVFFKNHDLRMEMALWQRLHQAGLIQLYMRQ

SEQ ID NO: 74.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1) de erizo de mar

MRLWRKLTSCKTRSRFVLLAFWYTMIALSVSLLYSLRGTHHSQDDSGTTSRRSRFYNSLFPFSRNDVADYATLTQ
HDVADDIVAPPHVGATGEIGPEKNVAIDDNKQIQHLIRNSGGIAGQSGNLIGDMLNDEEQRTFEVVEEBVAEQE
EIEVVNEFRDLNGKRPSGDDPRAPVTSTIDVTSIHFNATVKPTPPPEVQGYNVQTKDKLRMRCSCALVSSSG
HLVNTSAGAEIDSYPCLVRMNSAPVRGYEVDVGRRTTIRIMGHVNLKVLNASNELQDEILINSTTRAEKIIPWL
YNVKVNQATDMYYKSARNLSSLYPHVEFYLLTPDKMKIAESLFQTETGLTRQETRTWLSTGWMNMLYAVDVCDKV
DIFGLVPENYCLKNPNSSVLHYHYESDGLKECDYTTISEERLTSGHKFITEKAVFARWATKFNMVIPPTSLEPHG
RCPY

10

SEQ ID NO: 75.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (ST6GalNAc-1 Δ35) humana truncada

KEPQTKPSRHQRTENIKERSLQSLAKPKSQAPTRARRTTIYAEPVPENNALNTQTQPKAHTTGDRGKEANQAPPE
EQDKVPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGGGQTRKLTASRTVSEKHQ
GKAATTAKTLIPKSQHRMLAPTGAIVSTRQKGVTTAVIPPEKKPQATPPPAPFQSPTTQRNQLKAANFKSEP
RWDFEKYSFEIGGLQTTCPDSVKIKASKSLWLQKFLPNLTFLDSRHFNQSEWDRLEHFAPPFGFMELNYSLV
QKVVTFRFPVPQQLLLASLPAGSLRCITCAVVGNGGILNNSHMQEIDSHDYVFRLSGALIKGYEQDVGTSTF
YGFTAFSLTQSLILGNRGFKNVPLGKDVRYLHFLEGRDYEWEALLMNQTVMSKNLFWFRHRPQEAFFREALHM

DRYLLLLHPDFLRYMKNRFLRSKTLGDAHWRIRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSWKR
LIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

SEQ ID NO: 76.

alfa-N-acetilgalactosaminida alfa-2,6-sialiltransferasa I (MBP-ST6GalNAc-1 Δ 35) humana truncada marcada con proteína de unión a maltosa

MKIEEGKLVIWINGDKGYNGLAIEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFWAHDRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKOLLNPPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAFNKGETAM
TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYYEELAKDPRIATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSSNNNN
NNNNNNLGIIEGRISEFGSKEPQTKPSRHQRTENIKERSLQSLAKPKSQAPTRARRTTIYAEPVPENNALNTQTQP
KAHTTGDRGKEANQAPPEEQDKVPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGN
GGQTRKLTASRTVSEKHQGAATTAKTLIPKSHRMLAPTGAVSTRTRQKGVTTAVIPPEKKPQATPPAPFQS
PTTQRNQLKAANFKSEPRWDFEEKYSFBIIGGLQTTCPDSVKIKASKSLWLQKFLPNLTLFLDSRHFNQSEWDR
LEHFAPPFGFMELNYSLVQKVTRFPVPVQQQLLASLPAGSLRCITCAVVGNGGILNNSHMGQEIDSHDYVFRLL
SGALIKGYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVPLGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKN
LFWFRHRPQEAFFREALHMDRYLLLLHPDFLRYMKNRFLRSKTLGDAHWRIRPTTGALLLLTALQLCDQVSAYGFI
TEGHERFSDHYDTSWKRLLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

5

SEQ ID NO: 77.

N-acetilgalactosamina-alfa2,6-sialiltransferasa I (ST6GalNAc-1 Δ 231) de pollo truncada

KTEPQWDFDDEYILDSSSPVSTCSSESVRAKASDWRDLFLPNITLFDKSYFNVSEWDRLEHFAPPYGFMEIN
YSLVEEVMSRLPPNPHQQLLLANSSSNVSTCISCAVVGNGGILNNSGMGQEIDSHDYVFRVSGAVIKGYEKDVGT
KTSFYGFTAYSLSVSSLQNLGHKGFKKIPQKGHIRYIHFLEAVRDYEWLKALLLDKDIRKGFLNYYGRRPRERFDE
DFTMNKYLVAHPDFLRYLKNRFLKSKNLQKPYWRLYRPTTGALLLLTALHLCDRVSAYGYITEGHQKYS DHYYDK
EWKRLVIFYVNHDFNLEKQVWKRLHDENIMKLYQRS

SEQ ID NO: 78.

10 beta-galactosida alfa-2,6-sialiltransferasa (ST6Gal I) humana

MIHTNLKKKFSCCVLVFLLFAVICVWKEKKKGSYYDSFKLQTKFQVLKSLGKLAMGSDSQSVSSSTQDPHRGR
QTLGSLRGLAKAKPEASFQVWNKDSSSKNLI PRLQKIWKNYLSMNKYKVSYKGPGPQGIKFSAEALRCHLRDHVNV
SMVEVTDFFNTSEWEGYLPKESIRTKAGPWGRCAVVSSAGSLKSSQLGREIDDHDAVLRFNAGAPTANFQQDVGT
KTTIRLMNSQLVTTTEKRFLKDSLYNEGILIVWDPSVYHSDIPKWYQNPDPYNNFYKTYRKLHPNQPFYILKPQM
PWELWDILQEISPEEIQPNPPSSGMLGIIIMMTLCDQVDIYEFPLPSKRKTDVCYYYQKFFDSACTMGAYHPLLYE
KNLVKHLNQGTDEDIYLLGKATLPGFRTIHC

SEQ ID NO: 79.

beta-galactosida alfa-2,6-sialiltransferasa (ST6Gal I) de ratón

MIHTNLKRKFSCFVLVFLFAIICVWKKGS DYEALTQAKVFQMPKSQEKVAVGPAPQAVFSNSKQDPKEGVQIL
SYPRVTAKVKPQPSLQVWDKDYSTYSLNPRLLKIWRNYLNMNMYKVSYKGPGPGRVFSVEGLRCHLRDHVNVSMI
EATDSPFNTTEWEGYLPKETFRTKAGPCTKCAVVSSAGSLKNSQLGREIDNHDAVLRFNAGPTDNFQQDVGTKTT
IRLVNSQLVTTTEKRFLKDSLYTEGILILWDPSVYHADIPQWYQKPDYNNFFETYKSYRRLHPSQPFYILKPQMPWE
LWDIIQEISPDLIQPNPPSSGMLGIIIMMTLCDQVDIYEFPLPSKRKTDVCYHQQKFFDSACTMGAYHPLLF EKNM
VKHLNEGTDEDIYLF GKATLSGFRNNRC

SEQ ID NO: 80.

beta-galactosida alfa-2,6-sialiltransferasa (ST6Gal I) de rata

MIHTNLKKKFSLFILVFLFAVICVWKKGSDYEALTLQAKEFQMPKSQEKVAMGSASQVVFNSKQDPKEDIPIIL
SYHRVTAKVKPQPSFQVWDKDYSTYKLNPRLLKIWRNYLNMNKYKVSYKGGPGVKFSVEALRCHLRDHVNVSMI
EATDFPFNTTEWEGYLPKENFRTKVGPWQRCVAVSSAGSLKNSQLGREIDNHDAVLRFNAGPTDNFQQDVGSKTT
IRLMNSQLVTTEKRFLKDSLYTEGILIVWDPSVYHADIPKWKYQKPDYNFFETYKSYRRLNPSQPFYILKPQMPWE
LWDIIQEISADLIQPNPPSSGMLGIIIMMTLCDQVDIYEFLLPSKRKTDVCYHQQKFDSACTMGAYDPLLFEKNM
VKHLNEGTDEDIYLF GKATLSGFRNIRC

SEQ ID NO: 81.

5 beta-galactosida alfa-2,6-sialiltransferasa (ST6Gal I) de vaca

MTRTSLKKKVFSCCVLIFLLFAIICVWKEKKKGNYYEFLKLQNKYQVLQGLEKLAVSSSSQPVSSSSTHNPQRN
IQALGGPKAKLKATFQVWDKSSSKNLAPRLQTIKKNYLNMNKYKVITYKGGPGVKFSAEALLCHLRDHVNI SMI
EASDFPFNTSDWEGYLPQEDIRSKAGPWGRCAVSSAGSLKSSRLGREIDDHDAVLRFNAGPTVKFQQDVGTKTT
IRLVNSQLVTTEAGFLKDSLYNEGILIVWDPSVYHSDIPKWRNPDYSFFNNFKSYRKLHPDQPFYILKPQMPWE
LWDIIQEISSELIQPNPPSSGMLGIAIMMSLCDQVDIYEFLLPSKRKTDVCYYYQRYFDSACTMGAYHPLLFEKNM
VKYLNLTGTDEDIYLLGKATLPGFRTIRCGA

SEQ ID NO: 82.

beta-galactosida alfa-2,6-sialiltransferasa (ST6Gal I) de pollo

MVHINVLKKFMCVLVILIALTVCLWKETRGSYYVPLKNDGTQVHRALDKWNLLKSQGLLHEAAGEMGQMPKALP
NNQNKVKGITSGAVEKSRKAAEHVKVWDKSSSRNLIPRLQKVRKNYLSMNKYNVITYNGKMNAAKLSPEQLLCRL
RDRVNVTMIRGSDGPFNSSEWQHYPDKSLNETVGRLGRCAVSSAGSLKSSHLGPEIDSHDAVLRFNAGPVKGF
QEDVGQKTTIRLVNSQLVTVEEQFLKDALYNTGILIVWDPAPYHAEIHEWYRKPDKYKFFEAYKSYRIRHPEQPF
YIILNPKMQWLQDILQENSLEHIQPNPPSSGMLGIVIMMTLCDEVDVYEFLLPSKRQTDICHYYQKFHDHACTMGA
YHPLLFEKNLVKHLNQGTDEDIYTHGKVTLPGFRNVHC

10 SEQ ID NO: 83.

alfa-N-acetilneuraminida alfa-2,8-sialiltransferasa (ST8Sia I) humana

MSPCGRARRQTSRGAMAVLAWKFPRTRLPMGASALCVVVLWCWLYIFPVYRLPNEKEIVQGVLLQGTAWRRNQTA
RAFRKQMEDCCDPAHLFAMTKMNSPMGKSMWYDGEFLYSFTIDNSTYSLFPQATPFQLPLKKCAVVGNGGILKKS
GCGRQIDEANFVMRCNLPLSSEYTKDVGSKSQLVTANPSIIRQRFQNLWSRKTFFVDMKIYNHSYIYMPAFSM

KTGTEPSLRVYYTSLSDVGANQTVLFANPNFLRSIGKFWKSRGIHAKRLSTGLFLVSAALGLCEEVAIYGFWPFSV
NMHEQPI SHHYDNLVLPFSGFHAMPPEEFLQLWYLHKIGALRMQLDPCEDTSLQPTS

SEQ ID NO: 84.

15 alfa-N-acetilneuraminida alfa-2,8-sialiltransferasa (ST8Sia I) de rata

MSPGGGPLHTSRGAMAMLARKFPRTRLPGVGSALCVVVLWCWLYVFPVYRLPNEKEIVQGVLAQRTAWRRNQTSAR
LFRKQMEDCCNPAHLFAMTKVNSPMGKSLWYDGEFLYSLTIDTSTYSLFPQATPFQLPLKKCAVVGNGGILKMSG
CGRQIDEANFVMRCNLPLSSEYTRDVGSKTQLVTANPSIIRQRFENLLWSRKKFVDMKIYNHSYIYMPAFSMK
TGTEPSLRVYYTSLDAGANQTVLFANPNFLRNIGKFWKGRGIHAKRLSTGLFLVSAALGLCEEVSIYGFWPFSVN
MQGEPI SHHYDNLVLPFSGFHAMPPEEFLQLWYLHKMGALRMQLDPCENPPSPQPTS

SEQ ID NO: 85.

alfa-N-acetilneuraminida alfa-2,8-sialiltransferasa (ST8Sia I) de vaca

MSPCGRRRHTSRGAMAVLAWKFPRTRLPVGASALCVVVLWCWLYVFPVYRLPDEKEIVQGVLQQGTAWRRNRATA
GIFRKQMEDCCDPAHLFAMTKMNAPMGKSLWYDGEFLYSFTIDNSTYSLFPQATPFQLPLKKCAVVGNGGILKKS
GCGRQIDEADFVMRCNLPPLSSEYTKOVGSKSHLVNTANPSIIRQRFQNLWSRKTFFVDMKVYNHSYIYMPAFSM
KTGTEPSLRVYYTLSDVGANQTVLFANPNFLRSIGKFWKSRGIHAKRLSTGLFLVSAALGLCEEVAIYGFWPFVS
NMHEQPISSHYYDNVLPFSGFHAMPEEFQLWYLHKIGALRMQLDPCEDNSLQPTS

SEQ ID NO: 86.

ST8Sia II, PSA-T, SIAT8B, EC 2.4.99.-

5 alfa-2,8-sialiltransferasa 8B (ST8Sia II) humana

MLQQFRSWMLAALTLLVVFLIFADISEIEEEEIGNSGGRGTIRSAVNSLHSSKNRAEVVINGSSSPAVVDRSNESI
KHNIQPASSKWRHNQTLRLRIRKQILKFLDAEKDISVLKGTLPKPGDIHYIFDRDSTMNVSONLYELLPRTSPLK
NKHFGTCAIVGNSGVLLNSGCGQEIDAHSFVIRCNLAPVQYARDVGLKTDLVMTNPSVIQRAFEDLVNATWREK
LLQRLHSLNGSILWIPAFMARGGKERVEWVNELILKHVNVRTAYPSLRLLHAVRGYWLTKNVHIKRPTTGLLMY
TLATRFCKQIYLYGFWPFPLDQONQNPVKYHYDSLKYGYTSQASPTMPLEFKALKSLHEQGALKLTVGQCDGAT

SEQ ID NO: 87.

alfa-2,8-sialiltransferasa 8B (ST8Sia II) de *Xenopus*

MNPDFRSWVLATVTLVVFLILADISELEEEEIGSTGGSGTIRSAAGNVHSSKNRAELILNNSSTSTAGRSNESI
KYRTKPVPIWRHNQTLALRIRKQILHFLDAERDISVLKGTLPKPGDIHYVFDSDSTMNISKNLYELLPRTSPLK
NKHFKTCAIVGNSGILLNSGCGKEIDSHDFVIRCNLAPVEEYAKDVGTNTLVMTNPSVVQRAFEDLVNDTWKDK
FLQRLKSLNESILWIPAFMAKGGEERVEVNDLIKHINVTAYPSLRLLHAVRGYWLTKNVHIKRPTTGILMY
TLATRFCNRIYLYGFWPFPRDLHQNPVKYHYDSLKYGYTSQAGPHAMPLEFKALKNLHLQGALKLVNVECEAAT

10 SEQ ID NO: 88.

alfa-2,8-sialiltransferasa 8B (ST8Sia II) de rata

MLQQFRSWMLAALTLLVVFLIFADISEIEEEEIGNSGGRGTIRSAVNSLHSSKNRAEVVINGSSLPVADRSNESL
KHSIQPASSKWRHNQTLRLRIRKQILKFLDAEKDISVLKGTLPKPGDIHYIFDRDSTMNVSONLYELLPRTSPLK
NKHFTCAIVGNSGVLLNSGCGQEIDTHSFVIRCNLAPVQYARDVGLKTDLVMTNPSVIQRAFEDLVNATWREK

LLQRLHGLNGSILWIPAFMARGGKERVEWVNALILKHVNVRTAYPSLRLLHAVRGYWLTKNVHIKRPTTGLLMY
TLATRFCNQIYLYGFWPFPLDQONQNPVKYHYDSLKYGYTSQASPTMPLEFKALKSLHEQGALKLTVGQCDGAT

SEQ ID NO: 89.

15 alfa-2,8-sialiltransferasa 8B (ST8Sia II) de pollo

MPLPCRGTWTLALLTLVGFLIFADISEIEGEGSSGGRTIRSAVNSLHSSKNRAEVVINDSSSPAVVDRSNESI
KHNIKPASSKWRHNQTLRLRIRKQILKFLDAEKDISVLKGTLPKPGDVIHYVFDSDSTMNVSONLYELLPRTSPLK
GKQFPSCAIVGNSGVLLNSGCGPEIDTHSFVIRCNLAPVQEYSQDVGTNTLVMTNPSVIQRAFEDLMNETWREK
LLQRLHSLNGSILWIPAFMAKGGERVEWVNELILKHVINVTAYPSLRLLHAVRGYWLTKNVHIKRPTTGLLMY
TLATRFCNRIHLYGFWPFPLDQONQNPVKYHYDSLKYGYTSQASPTMPLEFKALKTLHQGALKLTVGECEGDT

SEQ ID NO: 90.

alfa-2,8-sialiltransferasa 8C (ST8Sia III) humana

MRNCKMARVASVLGLVMLSVALLLILSLISYVSLKKENIFTTPKYASPGAPRMYMFHAGFRSQFALKFLDPSFVPI
TNSLTQELQEKPSKWKFNRTAFLHQRQEILQHVDVIKNFSLTKNSVRIGQLMHYDYSSHKYVFSISNNFRSLLPD
VSPIMNKHYNICAVVGNSGILTFIQCGREIDKSDVFRCNFAPSEAFQRDVGRKTNLTTFNPSILEKYNNLLTI
QDRNNFFLSLKKLDGAILWIPAFFHSTSATVTRTLVDFFVEHRGQLKVQLAWPGNIMQHVNRWKNKHLSPKRLS
TGILMYTLASAICEEIHLYGFWPFGFDPNTREDLPYHYDCKGKTKFTTKWQESHQLPAEFQLLYRMHGEGLTKLT
LSHCA

SEQ ID NO: 91.

alfa-2,8-sialiltransferasa 8C (ST8Sia III) de rata

MRNCKMARVASILGLVMLSVALLLILSLISYVSLKKENIFTTPKYASPGAPRMYMFHAGFRSQFALKFLEPSFVPI
TNFLTHELQEKPSKWTFNRTAFLHQRQEILQHVDVIKNFSLTKNSVRIGQLIHYDYSSHKYVFSISNNFRSLLPD
VSPILNKRYNICAVVGNSGILTGSQCGQEIDKSDVFRCNFAPTEAFHKDVGKKTNLTTFNPSILEKYNNLLTI
QDRNNFFLSLKKLDGAILWIPAFFHSTSATVTRTLVDFFVEHRGQLKVQLAWPGNIMQHVNRWKNKHLSPKRLS
TGFLMYTLASAICEEIHLYGFWPFGFDPNTREDLPYHYDCKGKTKFTTKWQESHQLPAEFQLLYRMHGEGLTKLT
LSHCA

5 SEQ ID NO: 92.

alfa-2,8-sialiltransferasa 8C (ST8Sia III) de pollo

MVRVASVLGLVMLSIALLLILSLISYVSLKKDNIFGAPRAAGPGGPRMYMFHAGFRSQFALKFLDPSFVPIITNSLG
QELQEKPSKWFNRTAFAQQRQEILQHVDVIKNFSLTKNSVRIGQLMHYDYSSHKYVFSISNNFRSLLPDVSPIL
NKHYNVCAVVGNSGILTGSQCGQEIDKSDVFRCNFAPTEAFQKDVGRKTNLTTFNPSILEKYNNLLTIQDRNN
FFLSLKKLDGAILWIPAFFHSTSATVTRTLVDFFVEHRGQLKVQLAWPGNIMQHVNRWKNKHLSPKRLSTGILM
YTLASAICEEIHLYGFWPFGFDPNTREDLPYHYDCKGKTKFTTKWQESHQLPAEFQLLYRMHGEGLAKLTLSRCA

SEQ ID NO: 93.

alfa-2,8-sialiltransferasa 8C (ST8Sia III) de ratón

MRNCKMARVASVLGLVMLSVALLLILSLISYVSLKKENIFTTPKYASPGAPRMYMFHAGFRSQFALKFLDQSFVPI
TNSLTHELQEKPSKWTFNRTAFLHQRQEILQHVDVIKNFSLTKSSVRIGQLMHYDYSSHKYVFSISNNFRSLLPD

10

VSPIMNKRYNVCAVVGNSGILTGSQCGQEIDKSDVSRCNFAPTEAFHKDVGRKTNLTTFNPSILEKYNNLLTI
QDRNNFFLSLKKLDGAILWIPAFFHSTSATVTRTLVDFFVEHRGQLKVQLAWPGNIMQHVNRWKNKHLSPKRLS
TGILMYTLASAICEEIHLYGFWPFGFDPNTREDLPYHYDCKGKTKFTTKWQESHQLPAEFQLLYRMHGEGLTKLT
LSHCA

SEQ ID NO: 94.

alfa-2,8-sialiltransferasa 8D (ST8Sia IV) humana

MRSIRKRWITICTISLLLIFYKTKETARTEEHQETQLIGDGELSLSRSLVNSSDKIIRKAGSSIFQHNVEGWKINS
SLVLEIRKNILRFLDAERDVSVVKSFSKPGDVIHYVLDRRRTNLNISHDLHSLLEPVSPMKNNRRFKTCAVVGNSGI
LLDSECGKEIDSHNFVIRCNLAPVVEFAADVGTSDFITMNPVSVQRAFGGFRNESDREKFVHRLSMLNDSVLWI
PAFMVKGGEKHVEWVNALILKNLKVRTAYPSLRLIHAVRGYWLTKVPKIRPSTGLLMTLATRFCDIEHLYGF
WPFPPKDLNGKAVKYHYDDLKYRYFSNASPHRMPLEFKTLNVLHNRGALKLTGKCVKQ

15

SEQ ID NO: 95.

alfa-2,8-sialiltransferasa 8D (ST8Sia IV) de rata

MRSIRKRWITICTISLLLI FYKTK EIARTEEHQETQLIGDGELCLSRSLVNNSDKITRKAGSTIFQHSVQGWKINS
SLVLEIRKNILRFLDAERDVSVVKSSFKPGDVIHYVLD RRRRTLNIS HDLHSLLEPVSPMKNRRFKTCAVVGNSGI
LLDSGCGKEIDSHNFVIRC NLAPVVEFAADVGT KSDFITMNP SVVQRAFGGFRNESDREK FVHRLSMLNDSVLWI
PAFMVKGGEKHVEWVNALILKNKLKVRTAYPSLR LIHAVRGYWL TNKVPIKRPITGLLMYTLATRFCD EIHLYGF
WPF PKDLNGKAVKYHYDD LKYRYFSNASPHRMPLEFKTLNVLHNRGALKLTT

SEQ ID NO: 96.

alfa-2,8-sialiltransferasa 8D (ST8Sia IV) de pollo

MRSVRKRWTVCTISLLLI FYKTK EIARTEERQEAPLAGDGELSLSRSMINSSDKIIRKGGSAIFQHSVEGWKINS
TLVLEIRKSILRFLDAERDVSVVKSSFKPGDVIHYVLD RRRRTLNIS QDLHSLLEPVSPMKNRRFKTCAVVGNSGI
LLDSGCGKEIDTHDFVIRC NLAPVVEFAADVGNK SDFITMNP SVVQRAFGGFRNESDREK FGHRLSMLNDSVLWI
PAFMVKGGEKHLEWVNALILKNKLKVRTAYPSLR LIHAVRGYWL TNKVHIKRPSTGLLMYTLATRFCD EIHLYGF
WPF PKDLHGKPKVKYHYDD LKYRYFSNASPHRMPLEFKTLV LHNRGALKLTTGKCVKQ

5

SEQ ID NO: 97.

alfa-2,8-sialiltransferasa 8D (ST8Sia IV) de vaca

MRSIRKRWITICTISLLLI FYKTK EMARTEEHQETQLIGDGELSLSRSLVNSSDKIIRKAGSSIFQHSVEGGKINS
SLVLEIRKNILRFLDAERDVSVVKSSFKPGDVIHYVLD RRRRTLNIS QDLHSLLEPVSPMKNRRFKTCAVVGNSGI
LLDSECGKEIDSHNFVIRC NLAPVVEFAADVGT KSDFITMNP SVVQRAFGGFRNESDREK FVHRLSMLNDSVLWI
PAFMVKGGEKHVEWVNALILKNKLKVRTAYPSLR LIHAVRGYWL TNKVPIKRPSTGLLMYTLATRFCD EIHLYGF
WPF PKDLNGKAVKYHYDD LKYRYFSNASPHRMPLEFKTLNVLHNRGALKLTTGKCIKQ

SEQ ID NO: 98.

10 alfa-1,2-fucosiltransferasa, FUT1, humana

MWLRSHRQLCLAFLLVCVLSVIFFLHIHQDSFPHGLGLSILCPDRRLVTPPVAIFCLPGTAMGPNASSSCPQHPA
SLSGTWTVPNGRFGNQMGQYATLLALAQ LNGRRAFILPAMHAALAPVFRITLPVLAPEVDSRTPWRELQLHDMW

SEFYADLRDPFLKLSGFPCSWTFFHHLREQIRREFTLHDHLREEAQSVLGQLRLGRTGDRPRTFVG VHVRRGDYL
QVMPQRWKGVVGD SAYLRQAMDWFRARHEAPV FVVT SNGMEWCKENIDTSQGDVTFAGDGQEATPWKDFALLTQC
NHTIMTIGTGF GFWAAYLAGD TVYLANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPLWTLAKP

SEQ ID NO: 99.

alfa-1,2-fucosiltransferasa, FUT2, humana

MLVVQMPFSFPMAHFILFVFTVSTIFHVQQLAKIQAMWELPVQIPVLASTSKALGPSQLRGMWTINAIGRLGNQ
MGEYATLYALAKMNGRPAFIPAQMHS TLAPIFRITLPVLHSATASRIPWQNYHLNDWMEEEYRHIPGEYVRFTGY
PCSWTFYHHLRQEILQEFTLHDHVREEAQKFLRGLQVNGSRPGTFVG VHVRRGDYVHVMPKVWKGVVADRRYLQQ
ALDWFRARYSSLIFVVT SNGMAWCRENIDTSHGDVVFAGD GIEGSPAKDFALLTQC NHTIMTIGT FGIWAAYLTG
GDTIYLANYTLPD SPFLKIFKPEAAFLPEWTGIAADLSPLLKH

15

SEQ ID NO: 100.

alfa-1,2-fucosiltransferasa, FUT1, de cerdo

MWVPSRRHLCLTFLLVCLAAIFFLVNYQDLFYSGLDLLALCPDHNVSPPVAIFCLAGTPVHPNASDSCPKHPA
SLSGTWTIYPDGRFGNQMGQYATLLALAQLNGRQAFIQPAMHAVLAPVFRITLPVLAPEVDRHAPWRELELHDWM
SEDYAHLKEPWLKLTGFPCSWTFFHHLREQIRSEFTLHDHLRQEAQGVLSQFRLPRTGDRPSTFVGHVRRGDYL
RVMPKRWKGVVGDGAYLQQAMDWFRARYEAPVFVVTSSNGMEWCRKNIDTSRGDVI FAGDGREAAPARDFALLVQC
NHTIMTIGTGFWAAYLAGGDTIYLANFTLPTSSFLKIFKPEAAFLPEWVGINADLSPLQMLAGP

SEQ ID NO: 101.

alfa-1,2-fucosiltransferasa, FUT2, de cerdo

MLSMQASFFFPPTGPFILFVFTASTIFHLQQRMVKIQPTWELQMVTVTTESPPSPQLKGMWTINAIGRLGNQMGE
YATLYALARMNGRPAFIPPEMHSTLAPIFRITLPVLHASTARRIPWQNYHLNDWMEERYRHIPGEYVRLTGYPSC
WTFYHHLRTEILREFTLHNHVREEAQDFLRGLRVNGSRPSTYVGHVRRGDYVHVMPNVWKGVVADRRYLEQALD
WFRARYRSPVFVSSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTVMTIGTFGIWAAYLAGGET
IYLANYTLPDSPFLKLFKPEAAFLPEWIGIEADLSPLLKX

5 SEQ ID NO: 102.

alfa-1,2-fucosiltransferasa, FUT1, de rata

MWTPSRKQLCLAFLSVCVLSAGSFFHNLNGNFFQNALTFSVLCPDYHLLKSPVAMVCLPYPSNASSGSPSCPEQ
SLLSGTWTITPGGRFGNQMGQYATLLALAQLNGRRAFIQPEMHTTLPVFRISLPVLDPEVDSLTPWQHLVLHDW
MSEESHLEDPFLKLSGFPSCSWTFFHHLREQIRREFTLHDHLREDAQRLLSGLRIGPAGIRPRTYVGHVRRGDY
LEVMPNWRWKGVDRAYLQKAMDWFRARHKDPIFVVTSSNGMRWCLENIDTSHGDVVFAGNGQEGTPGKDFALLTQ
CNHTIMTIGTGFWAAYLAGGDTVYLANFTLPDSEFLKIFRPKAAFLPEWVGINADLSPLQAQFDPWETDSLFR
A

SEQ ID NO: 103.

alfa-1,2-fucosiltransferasa, FUT2, de rata

MASAQVPFSFPLAHFLIFVFTSTIIHLQQRIVKLQPLSEKELPMTTQMSSGNTESPEMRRDSEQHNGELRGMF
TINSIGRLGNQMGEYATLFALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEERYRH

10

IPGHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLRGLRVNGSQPSTFVGHVRRGDYVHVMPNVWK
GVVADRGYLEKALDMFRARYSSPVFVVTSSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIG
TFGIWAAYLAGGDTIYLANYTLPDSPFLKVFKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYV
AGRAF

SEQ ID NO: 104.

fucosiltransferasa, FUT3, humana

MDPLGAAKPQWPWRRCLAALLFQLLVAVCFSSYLRSRDDATGSPRAPS GSSRQDTTPTRPTLLILLWTWPFHIP
VALSRCSEMVPGTADCHITADRKVYPQADTVIVHHWDIMSNPKSRLPPSPRPQQRWIWFNLEPPNQCQHLEALD
RYFNLTMSYRSDSDIFTYPGWLEPWSGQPAHPPLNLSAKTELVAWAVSNWKPD SARVYYQSLQAHLKVDVYGRS
HKPLPKGTMMETLSRYKFYLAFENSLHPDYITEKLWRNALEAWVPVVLGSPRSNYERFLPPDAFIHVDDFQSPK
DLARYLQELDKDHARYLSYFRWRETLP RSFSWALDFCKACWKLQQESRYQTVRSIAAWFT

15 SEQ ID NO: 105.

fucosiltransferasa, FUT6/FCT3A, humana

MDPLGPAKPQSWRCCLTTLFQLLMAVCFFSYLRVSQDDPTVYPNGSRFPDSTGTPAHSIPLILLWTWPFNKPI
ALPRCSEMPGTADCNITADRKVYPQADAVIVHHREVMYNPSAQLPRSPRRQGQRWIWFSMESPSHCWQLKAMDG
YFNLTMSYRSDSDIFTPYGWLEPWSGQPAHPPLNLSAKTELVAWAVSNWGPNSARVRYQSLQAHLKVDVYGRSH
KPLPQGTMMETLSRYKFYLAFENSLHPDYITEKLWRNALEAWAVPVVLGPPSRSNYERFLPPDAFIHVDDFQSPKD
LARYLQELDKDHARYLSYFRWRETLRPRSFSWALAFCKACWKLQESRYQTRGIAAWFT

SEQ ID NO: 106.

fucosiltransferasa, FUT5, humana

MDPLGPAKPQWLWRRCLAGLLFQLLVAVCFFSYLRVSRRDATGSPRPGLMAVEPVTGAPNGSRCQDSMATPAHPT
LLILLWTWPFNTFPVALPRCSEMPGAADCNITADSSVYPQADAVIVHHWDIMYNPSANLPPTRPQGQRWIWFSM
ESPSNCRHLEALDGYFNLTMSYRSDSDIFTPYGWLEPWSGQPAHPPLNLSAKTELVAWAVSNWKPD SARVRYQSL
LQAHLKVDVYGRSHKPLPKGTMMETLSRYKFYLAFENSLHPDYITEKLWRNALEAWAVPVVLGPPSRSNYERFLPP
DAFIHVDDFQSPKDLARYLQELDKDHARYLSYFRWRETLRPRSFSWALAFCKACWKLQESRYQTVRSIAAWFT

5 SEQ ID NO: 107.

fucosiltransferasa, FUT7, humana

MNNAGHGPTRRRLRGLGVLGVALLAALWLLWLLGSAPRGTPAPQPTITILVHWPFTDQPPPELPSDTCTRYGIAR
CHLSANRSLASADAVVFHRELQTRRSHLPLAQRPRGQPVWVWASMESPSHTHGLSHLRGIFNWVLSYRRSDIF
VPYGRLEPHWGPSPPLPAKSRVAWVVSINFQERQLRARLYRQLAPHLRVDVFGGRANGRPLCASCLVPTVAQYRFY
LSFENSQHRDYITEKFWRNALVAGTVPVVLGPPRATYEA FVPADAFVHVDDFGSARELA AFLTGMNESRYQRFFA
WRDRLRVRLFTDWRERFCAICDRYPHLPRSQVYEDLEGWFQA

SEQ ID NO: 108.

fucosiltransferasa, FUT9, humana

MTSTSKGILRPFLIVCIILGCFMACLLIYIKPTNSWIFSPMESASSVLKMKNFFSTKTDYFNETTILVWVWPFQ
TFDLTSCQAMFNIQGCHLTTRSLYNKSHAVLIHHRDISWDLTNLPQARPPFQKWIWMNLESPHTHPQKSGIEH

10

LFNLTLTYRRSDIQVPYGFLLTVSTNPFVFEVPSKEKLVCVVVSNNWPEHARVKYYNELSKSIEIHTYGQAFGEY
VNDKNLIPTISACKFYLSFENSIHKDYITEKLYNAFLAGSVPVVLGPPSRRENYENYIPADSFHVEDYNSPSELAK
YLKEVDKNNKLYLSYFNWRKDFTVNLPRFWESHACLADHVKRHQEYKSVGNLEKWFVN

SEQ ID NO: 109.

fucosiltransferasa, FUT11, humana

MAAGPIRVVLVLLGVLSVCAASGHGSVAEREAGGEAEWAEPWDGAVFRPPAALGAVGVTRSSGTTPRGREEAGDL
PVLLWVSPGLFPHFPDGSERIECARGACVASRNRRLRDSRTRALLFYGTDFRASAAPLPRLAHQSWALLHEESP
LNNFLLSHGPGIRLFNLSTFSRHSYPLSLQWLPGTAYLRRPVPPMERAERRRRGYAPLLYLQSHCDVPADRD
RYVRELMRHIPVDSYGKCLQNRELPTARLQDTATATTEDPELLAFLSRYKFHLALENAICNDYMEKLRWPMHLG
AVPVYRGSPSVRDWMPNNHVSILIDD FESPQKLA EIDFLDKNDEEYMKYLAYKQPGGITNQFLDLSLKHREWG
NDPLLPNYLN GFECFVCDYELARLDAEKAAASPGDSPVFEPHIAQPSHMDCPVPTPGFGNVEEIPENDSWKEMW
LQDYWQGLDQGEALTAMIHNNETEQT KFDYDLHEIFMKRQHL

15 SEQ ID NO: 110.

fucosiltransferasa, FUT3, de chimpancé

MDPLGAAKPQWPWRRCLAALLFQLLVAVCFFSYLRVSRDDATGSPRPGLMAVEPVTGAPSGSSRQDTTPTRPTLL
ILLWTWPFHIPVALSRCSEMVPGAADCHITADRKVYPQADAVIVHHWDIMYNPKSRLPPSPRQQRWIWFNLEP
PPNCQHLEALDRYFNLMTSYRSDSDIFTPYGWLEPWSGQPAHPPLNLSAKTELVAWAVSNWKLDSARVRYYSLSQ
AHLKVDVYGRSHKPLPKGTMMETLSRYKFYLAFENSLHPDYITEKLWRNALEAWAVPVVLGPPSRSNYERFLPPDA
FIHVDDFQSPKDLARYLQELDKDHARYLSYFRWRETLPFRSFSWALDFCKACWKLQESRYQTMRSIAAWFT

SEQ ID NO: 111.

fucosiltransferasa, FUT4, de chimpancé

MGAPWGSPTAAAGRRRWRRGRGLPWTVCVLAAGLTCTALITYACWGQLPPLPWASPTSPRPVGVLLWWEFAG
RHSAPRPPDCRLRFNISGCRLLTDRASYGEAQAVLFHHRDLVKGPPDWPPWPGVQAHTAEVDLRLDYEEAAA
AAEALATSSPRPPGQRWVWMNFESPSPGLRSLASNLFNWTLSTRADSDVFVPYGYLYPRSHPGDPPSGLAPPL
SRKQGLVAWVSHWDERQARVRYHQLSQHVTVDVFRGGGQPVPEIGLLHTVARYKFYLAFENSLHPDYITEK
LWRNALLAGAVPVVLGPDNRANYERFVPRGAFIHVDDFPSASSLASYLFLDRNPAVYRRYFHWRRSYAVHITSFW
DEPWCRVCQAVQRAGDRPKSIRNLASWFER

5 SEQ ID NO: 112.

fucosiltransferasa, FUT5, de chimpancé

MDPLGPAKPQWLWRRCLAGLLFQLLVAVCFFSYLRVSRDDATGSPRPGLMAVEPVTGAPNGSRCQDSMATPAHPT
LLILLWTWPFNTPVALPRCSEMVPGAADCNITADSNVYPQADAVIVHHWDIMYNPSANLPPPTRPQQRWIWFSM
ESPSNCRHLEALDGYFNLMTSYRSDSDIFTPYGWLQPWGQPVHPPLNLSAKTELVAWAVSNWGPNSARVRYYS
LQAHKVDVYGRSHKPLPQGTMMETLSRYKFYLAFENSLHPDYITEKLWRNALEAWAVPVVLGPPSRSNYERFLPP
DAFIHVDDFQSPKDLARYLQELDKDHARYLSYFRWRETLPFRSFSWALDFCKACWKLQESRYQTVRSIAAWFT

SEQ ID NO: 113.

fucosiltransferasa, FUT6, de chimpancé

MDPLGPAKPQSWRCCLTTLFQLLVAVCFFSYLRVSRDDPTVYPNGSHFPDSTGTPAHSIPLILLWTWPFNKPI
ALPRCSEMVPGTADCNITADRKVYPQADAVIVHHREVMYNPSAQLPRSPRRQQRWIWFSMESPSNCRHLEALDG
YFNLMTSYRSDSDIFTPYGWLQPWGQPVHPPLNLSAKTELVAWAVSNWGPNSARVRYYSLSQAHKVDVYGRSH
KPLPQGTMMETLSRYKFYLAFENSLHPDYITEKLWRNALEAWAVPVVLGPPSRSNYERFLPPDAFIHVDDFQSPKD
LARYLQELDKDHARYLSYFRWRETLPFRFFSWALAFCKACWKLQESRYQTRRSIAAWFT

10

SEQ ID NO: 114.

fucosiltransferasa, fut7, de perro

MGPWGKPGKTASPVSHVGYSLSRRLRAVGVLGAALLSAIWLLWLLGSPPTGGPAPPHTLIILIWHWPFSDRPPE
LPSNTCISFGVAHCHLSTNHSLLASADAVVFHRELQTQARLPLAKRPRGQPVWVSMESPSHTHGLGRLGGIF
NWLVSRRSDIFVYPYQLEPHQGPAPPLPAKKRVAWVVSNFQERQRRVQVYRQLAPHLQVDVFRANRQPLCA
NCLLPYTAQYLFYLSFENSQHODYITEKFWRNALAAGTVPVVLGPPRATYEAFAFADAFVHVDDFSTAQELASFL
ANMNESHYRRYFAWRDRLRVRLFQDWRERFCAICTHYPHLPRGQVYQDLQGWFOA

SEQ ID NO: 115.

15 Fucosiltransferasa, FUT4, de ratón

MAPARQELQHESRCRPSRTVDAWRAAVATRGRHMETPGYRRRTRCGGWGLPRSVSSLAAVGLLCTALTTFICWGO
LPPLPWASAPAPQRLVGVLLWWEPPFRGRGGYPKSPDCSLRFNISGCRLLTDRAAYGAEQAVLFHHRDLVKELHDW
PPPWGARERTDKALVLRVFDQEGAVTLTGKALETVGSRPPGQRWVWMNFESPSHTPGLRGLAKDLFNWTLSTYRT
DSDVFPYGFYLSRSDPTEQPSGLGPQLARKRGLVAWVVSNNWNEHQARVRYHQLSRHVSVDVFGRTGPGRPVPA
IGLLHTVARYKFYLAFENSRHVDYITEKLWRNAFLAGAVPVVLGPDRANYERFVPRGAFIHVDDFPNAASLAAYL
LFLDRNAVYRRYFRWRRSFAVHITSFWDEQWCRTCQAVQTSQDQPKSIHNLADWFQR

SEQ ID NO: 116.

fucosiltransferasa, FUT7, de ratón

MPTPCPPACLSTPGTHRLLPFPDWKAPSWESRKEATCNSSSPGPWAEPTVQGYHPTRRRLRAWGGLAGGATFMVW
FFWLWGSAPGSAPVPQSTLTILIHWPFTNRPELPGDTC TRYGMASCRLSANRSLASADAVVFHRELQTRQS
LLPLDQRPBGQPVWVWASMESPSNTHGLHRFRGIFNVVLSYRRSDIFVPYGRLEPLSGPTSPLPAKSMAAWVIS
NFQERQORAKLYRQLAPHLQVDVFGGRASGRPLCANCLLPTLARYRFYLAFENSQHRDYITEKFWRNALAGAVPV
ALGPPRATYRAFVPPDAFVHVDDFSSARELAVFLVSMNESRYRGFFAWRDRLRVRLGDRWRERFCTICARYPYLP
RSQVYEDLESWFQA

5 SEQ ID NO: 117.

fucosiltransferasa, FUT9, de ratón

MTSTSKGILRPFLIVCIILGCFMACLLIYIKPTNSWVFS PMESASSVLKMNFFSTKTDYFNETTILVWVWPFQ
TFDLTSCQAMFNIQGCHLTDRSLYNKSHAVLIHHRDISWDLTNLPQARPPFQKWIWMNLESPHTPQKSGIEH
LFNLTLTYRRSDIQVPYGFLLTVSTNPFVFEVPSKEKLVCVVSNWNPEHARVKYYNELSKSIEIHTYGQAFGEY
VNDKNLIPTISTCKFYLSFENS IHKDYITEKLYNAFLAGSVPVVLGSPRENYENYIPADSF IHVEDFNPSSELA
YLKEVDKNNKLYLSYFNWRKDFTVNLPRFWESHACLACDHVKRHQYKSVGNLEKWFVN

SEQ ID NO: 118.

fucosiltransferasa, FUT4, de rata

MAPAGRKLQHESRCRPSRPVDAWRAAATTRGRCMGTPGARRTARRGGWGLPRTSSGLAAAGLLCTALTACLWGO
LPPLPWASAPAPQRPVSVLLWWEPPFGRGGHSGPPDCSLRFNISGCRLLTDRAAYGAEQAVLFHHRDLVKGPDPW
PPPWGAQERTDEALELRVFDQEGAVMLAREALETTGSRPPGQRWVWMNFESPSHTPGLRGLAKDLFNWTLSTYRT
DSDIFVPYGFYLPYRSHPAEQPSGLGPPLARKRGLVAWVVS HWNERQARVRYHQLRRHVSVDVFGRAGPGQPVPA
VGLLHTVARYKFYLAFENSQHVDYNTKELWRNAFLAGAVPVLLGPDRANYEGFVPRGSFIHVDDFPSAASLAAYL
LFLDRNAVYRRYFHWRRSYAVHITSFWDEPWCQTCRAVQTSQDQPKSIHNLADWFQR

10

SEQ ID NO: 119.

fucosiltransferasa, FUT9, de rata

MTSTSKGILRPFLIVCIILGCFVACLLIYIKPTNSWVFS PMESASSVLKMNFFSRKTDYFNETTILVWVWPFQ
TFDLTSCQAMFNIQGCHLTDRSLYNKSHAVLIHHRDISWDLTNLPQARPPFQKWIWMNLESPHTPQKSGIEH
LFNLTLTYRRSDIQVPYGFLLTVSTSPFVFEVPSKEKLVCVVSNWNPEHARVKYYNELSKSIEIHTYGQAFGEY
VNDKNLIPTISTCKFYLSFENS IHKDYITEKLYNAFLAGSVPVVLGSPRENYENYIPADSF IHVEDFNPSSELA
YLKEVDKNNKLYLSYFNWRKDFTVNLPRFWESHACLACDHVKRHQYKSVGNLEKWFVN

SEQ ID NO: 120.

15 fucosiltransferasa, FUT8, humana, isoforma A

MRPWTGSRWIMLILFAWGTLIFYIGGHLVRDNDHPDHSSRELSKILAKLERLKQONEDLRRMAESLRIPEGPID
 QGPAIGRVRLVEEQLVKAKEQIENYKKQTRNGLGKDHEILRRRIENGAKELWFFLQSELKKLKNLEGNELQRHAD
 EFLDLGHHERSIMTDLYYLSQTDGAGDWREKEAKDLTELVRRTITYLQNPDKCSKAKKLVCNINKGCGYGCQLH
 HVVYCFMIAYGTQRTLILESQNWRYATGGWETVFRPVSETCTDRSGISTGHWSGEVKDKNVQVVELPIVDSLHPR
 PPYLPLAVPEDLADRLVRVHGDPVWVWSQFVKYLIRPQPWLEKEIEEATKKLGFKHPVIGVHVRRTDKVGTEAA
 FHPIEYVMVVEEHFQLLARRMQVDKRVYLATDDPSLLKEAKTKYPNYEFISDNSISWSAGLHNRYTENSRLRGV
 ILDIHFLSQADFLVCTFSSQVCRVAYEIMQTLHPDASANFHSLLDDIYFGGQNAHNQIAIYAHQPRTADEIPMEP
 GDIIGVAGNHWDGYSKGVNRKLGRTGLYPSYKVREKIETVKYPTYPEAEK

SEQ ID NO: 121.

Fucosiltransferasa, FUT8, humana, isoforma C

MHRQIWHLHWTLVRYPNYEFISDNSISWSAGLHNRYTENSRLRGVILDIHFLSQADFLVCTFSSQVCRVAYEIMQTL
 LHPDASANFHSLLDDIYFGGQNAHNQIAIYAHQPRTADEIPMEPGDIIGVAGNHWDGYSKGVNRKLGRTGLYPSY
 KVREKIETVKYPTYPEAEK

5 SEQ ID NO: 122.

Fucosiltransferasa, FUT8, de cerdo

MRPWTGSRWIMLILFAWGTLIFYIGGHLVRDNDHSDHSSRELSKILAKLERLKQONEDLRRMAESLRIPEGPID
 QGPASGRVRALVEEQFMKAKEQIENYKKQTKNGPGKDHEILRRRIENGAKELWFFLQSELKKLKNLEGNELQRHAD
 EFLSDLGHHERSIMTDLYYLSQTDGAGDWREKEAKDLTELVRRTITYLQNPDKCSKAKKLVCNINKGCGYGCQLH
 HVVYCFMIAYGTQRTLALESHNWRYATGGWETVFRPVSETCTDRSGSSTGHWSGEVKDKNVQVVELPIVDSVHPR

 PPYLPLAVPEDLADRLVRVHGDPVWVWSQFVKYLIRPQPWLEKEIEEATKKLGFKHPVIGVHVRRTDKVGAEAA
 FHPIEYTVHVEEDFQLLARRMQVDKRVYLATDDPALKEAKTKYPSYEFISDNSISWSAGLHNRYTENSRLRGV
 ILDIHFLSQADFLVCTFSSQVCRVAYEIMQALHPDASANFRSLDDIYFGGPNAHNQIAIYPHQPRTEGEIPMEP
 GDIIGVAGNHWDGYPKGVNRKLGRTGLYPSYKVREKIETVKYPTYPEADK

SEQ ID NO: 123.

10 Fucosiltransferasa, fut8, de perro

MRPWTGSRWIMLILFAWGTLIFYIGGHLVRDNDHPDHSSRELSKILAKLERLKQONEDLRRMAESLRIPEGPID
 QAPASGRVRALVEEQLLKAKEQIENYKKQTRNGLGKDHEILRRRIENGAKELWFFLQSELKKLKNLEGNVLQRHAD
 EFLSDLGHHERSIMTDLYYLSQTDGAGDWREKEAKDLTELVRRTITYLQNPDKCSKAKKLVCNINKGCGYGCQLH
 HVVYCFMIAYGTQRTLILESQNWRYATGGWETVFRPVNETCTDRSGTSTGHWSGEVKDKNVQVVELPIVDSLHPR
 PPYLPLAVPEDLADRLVRVHGDPVWVWSQFVKYLIRPQPWLEKEIEEATKKLGFNIPVIGVHVRRTDKVGTEAA
 FHPIEYVMVVEEHFQLLARRMQVDKRRVYLATDDPSLLKEAKTKYPTYPEFISDNSISWSAGLHNRYTENSRLRGV
 ILDIHFLSQADFLVCTFSSQVCRVAYEIMQTLHPDASANFHSLLDDIYFGGQNAHNQIAIYPHQPRTADEIPMEP
 GDIIGVAGNHWDGYSKGVNRKLGRTGLYPSYKVREKIETVKYPTYPEAEK

SEQ ID NO: 124.

Fucosiltransferasa, FUT8, de ratón

MRAWTGSRWIMLILFAWGTLFFYIGGHLVRDNDHPDHSTRELSKILAKLERLKQQNEDLRRMAESLRIPGPID
QGTATGRVRVLEEQLVKAKEQIENYKKQARNGLGKDHEILRRRIENGAKELWFFLQSELKKLKHLEGNELQRHAD
EILLDLGHHERSIMTDLYYLSQTDGAGDWREKEAKDLTELVRRTITYLQNPDCSKARKLVCNINKGCGYGCQLH
HVVCYCFMIAYGTQRTLILESQNWRYATGGWETVFRPVSETCTDRSGLSTGHSWGEVNDKNIQVVELPIVDSLHPR
PPYLPPLAVPEDLADRLRLRVHGDPAVWVVSQFVKYLIRPQPWLEKEIEEATKKLGFKHPVIGVHVRRTDKVGTEAA
FHPIEEYMVHVEQHFQQLLARRMQVDKRVYLATDDPTLLKEANTKYSNYEFISDNSISWSAGLHNRYTENSLRGV
ILDIHFLSQADFLVCTFSSQVCRVAYEIMQTLHPDASANFHSLLDDIYFGGQNAHNQIAVYPHKPRTEEEIPMEP
GDIIGVAGNHWGDYSGKINRKLKGTGLYPSYKVREKIETVKYPTYPEAEK

SEQ ID NO: 125.

Fucosiltransferasa, fut8, de pollo

MRPWTGSRWIMLILFAWGTLFFYIGGHLVRDSEHPDHSSRELSKILAKLERLKQQNEDLRRMAESLRIPDGPID
QGPAAGKVHALEEQLLKAKEQIENYKKQTDGDLGKDHEILRRRIENGAKELWFFLQSELKKLKNLEGSSELQRRID
EFLSDLGHQERSIMTDLYYLSQTDGAGDWREKEAKDLTLVQRRITYLQNPDCSKAKKLVCNINKGCGYGCQLH
HVVCYCFMIAYGTQRTLILESQNWRYATGGWETVFRPVSETCTDRSGTTGHSWGETNDKDVQVVELPIVDSLHPR
PPYLPPLAVPEDLADRLIRVHGDPAVWVVSQFVKYLIRPQPWLEKEIEEATRKLGFHPVIGVHVRRTDKVGTEAA
FHPIEEYMVHVEERFELLARRMHVDKRVYLATDDPSLLQEAQSKYPNYEFISDNSISWSAGLHNRYTENSLRGV
ILDIHFLSQADFLVCTFSSQVCRVPYEIMQTLHPDASAYFHSLLDDIYFGGQNAHNQIAVYAHHPRTADEIPMEP
GDIIGVAGNHWGDYSGKINRKLKGTGLYPSYKVKEKIETVKYPTYPEAEK

5 SEQ ID NO: 126.

UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1) humana

MRKFAYCKVVLATSLIWVLLDMFLLLYFSECNKCEKKERGLPAGDVLEPVQKPHEGPGEMGKPVVIPKEDQEKM
KEMFKINQFNLMASEMIALNRSPLDVRLEGCKTKVYPDNLPPTTSVVIVFHNEAWSTLLRTVHVSVINRSRPHMIEE
IVLVDDASERDFLKRPLESYVKKLKVPHVIRMEQSRGLIRARLKGAAVSKGQVITFLDAHCECTVGWLEPLLAR
IKHDRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKLNFRWYPVPQREMDRRKGDRTLPVRTPTMAGGLFSIDR
DYFQEIGTYDAGMDIWGGENLEISFRIWQC GG TLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMD
EFKNFFYIISPGVTKVDYGDISSRVLRLHKLQCKPFSWYLENIYPDSQIPRHYFSLGEIRNVETNQCLDNMARKE
NEKVGIFNCHGMGGNQVFSYTANKEIRTDLCLDVSKLNGPVTMLKCHHLKGNQLWEYDPVKLTLQHVNSNQCLD
KATEEDSQVPSIRDCNGSRSQQWLLRNVTLP E I F

SEQ ID NO: 127.

UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1) de vaca

MRKFAYCKVVLATSLIWVLLDMFLLLYFSECNKCEKKERGLPAGDVLEPVQKPHEGPGEMGKPVVIPKEDQEKM
KEMFKINQFNLMASEMIALNRSPLDVRLEGCKTKVYPDNLPPTTSVVIVFHNEAWSTLLRTVHVSVINRSRPHMLEE
IVLVDDASERDFLKRPLESYVKKLKVPHVIRMEQSRGLIRARLKGAAVSKGQVITFLDAHCECTVGWLEPLLAR
IKHDRKTVCPIIDVISDDTFEYMAGSDMTYGGFNWKLNFRWYPVPQREMDRRKGDRTLPVRTPTMAGGLFSIDR
DYFQEIGTYDAGMDIWGGENLEISFRIWQC GG TLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMD
EFKNFFYIISPGVTKVDYGDISSRLGLRHLQCRPFSWYLENIYPDSQIPRHYFSLGEIRNVETNQCLDNMARKE
NEKVGIFNCHGMGGNQVFSYTANKEIRTDLCLDVSKLNGPVTMLKCHHLKGNQLWEYDPVKLTLQHVNSNQCLD
KATDEDSQVPSIRDCSGSRSQQWLLRNVTLP E I F

10

SEQ ID NO: 128.

UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1) de *Drosophila*

MLPRFRSFYGLKLIIFILVALCFILYSKVQONGSPPEPPVAPLVRAAALRGHGRERFEAYSDSENEIARPATQSPY
EQIIQLDLQKQKVGLEQGVAVHLSGAAKERGDEIYKKIALNEELSEQLTYNRSVGDHRNPLCAKQRFDSDSLPT
ASVVIIFNEPYSVLLRTVHSTLSTCNEKALKEIILVDDGSDNVELGAKLDYYVTRIPSGKVITILRLKNRLGLI
RARLAGARIATGDVLIIFLDAHCEGNIGWCEPLLQRIKESRTSVLVPIIDVIDANDFQYSTNGYKSFQVGGFQWNG
HFDWINLPEREKQQRRECKQEREICPAYSPTMAGGLFAIDRRYFWEVGSYDEQMDGWGGENLEMSFRIWQCGGT
IETIPCSRVGHIIFRDFHPYKFPNDRDTHGINTARMALVWMEYINIFFLNRPDLKFHADIGDVTHRVMLRKKLRRC
KSFEWYLNKIYPEKFVPTKDVQGWGKVHAVNSNICLDDLLQNNKPYNAGLYPCGKVLQKSQLFSFTNTNVLRNE
LSCATVQHSESPPYRVVMVPCMEDEFNEQWRYEHQHIHSNTGMCLDHQGLKSLDDAQVAPCDPHSESQRWTIE
H

SEQ ID NO: 129.

UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1) de ratón

MRKFAYCKVVLATSLVWVLLDMFLLLYFSECNKCEEKQERGLPAGDVLELVQKPHEGPGEMGKPVVIPKEDQEKM
KEMFKINQFNLMASEMIALNRSPLDVRLEGCKTKVYPDNLPSTSVVIVFHNEAWSTLLRTVHVSINRSRPHMIEE
IVLVDDASERDFLKRPLESYVKKLVPHVIRMEQSRGLIRARLKGAASRGQVITFLDAHCECTAGWLEPLLAR
IKHDRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKNFRWYFVPQREMDRRKGDRTLPVRTPTMAGGLFSIDR
DYFQEIPTYDAGMDIWGGENLEISFRIWQCGGTLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMD
EFKNFFYIISPGVTKVDYGDISSRLGLRRKLQCKPFSWYLENIYPDSQIPRHYFSLGEIRNVETNQCLDNMARKE

NEKVGIFNCHGMGGNQVFSYTANKEIRTDLCLDVSKLNGPVTMLKCHHLKGNQLWEYDPVKLTQLQHVNSNQCLD
KATEEDSQVPSIRDCTGSRSSQQWLLRNVTLP EIF

SEQ ID NO: 130.

UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1) de cerdo

MRKFAYCKVVLATSLIWVLLDMFLLLYFSECNKCEKKERGLPAGDVLEPVQKPHEGPGEMGKPVVIPKEDQDKM
KEMFKINQFNLMASEMIALNRSPLDVRLEGCKTKVYPDNLPSTSVVIVFHNEAWSTLLRTVHVSINRSRPHMLEE
IVLVDDASERDFLKRPLESYVKKLVPHVIRMEQSRGLIRARLKGAASRGQVITFLDAHCECTAGWLEPLLAR
IKHDRKTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKNFRWYFVPQREMDRRKGDRTLPVRTPTMAGGLFSIDR
DYFQEIPTYDAGMDIWGGENLEISFRIWQCGGTLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMD
EFKTFYIISPGVTKVDYGDISSRLGLRHLKQCRPFSWYLENIYPDSQIPRHYSSLGEIRNVETNQCLDNMARKE
NEKVGIFNCHGMGGNQVFSYTANKEIRTDLCLDVSKLNGPVTMLKCHHLKGNQLWEYDPVKLTQLQHVNSNQCLD
KATEEDSQVPSIRDCSGSRSSQQWLLRNVTLP EIF

SEQ ID NO: 131.

UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1) de *Xenopus*

MKKFAYCKVVLATSLVWVFLDMFLLLYFSECNKCDKKEKGLPGGDGPDINQKPHEGPGEMGKPVVIPKEEQEKM
KEMFKINQFNLMASEMIALNRSPLDVRLEGCKTKVYPDSLPSTSVVIVFHNEAWSTLLRTVHVSINRSRPHLLQE
IILVDDASERDFLKRPLETYVKKLTVPHVLRMEQSRGLIRARLKGAASRGQVITFLDAHCECTAGWLEPLLAR
IKHDRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKNFRWYFVPQREMDRRRGDRTLPVRTPTMAGGLFSIDR
DYFQEIPTYDAGMDIWGGENLEISFRIWQCGGTLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMD
EFKNFFYIISPGVTKVDYGDISTRVGLRHLKQCKPFSWYLENVYPDSQIPRHYSSLGEVSNIIIKLCFINVILLI
HI

SEQ ID NO: 132.

UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1) de rata

MRKFAYCKVVLATSLVWVLLDMFLLLYFSECNKCEEKKERGLPAGDVLELVQKPHEGPGEMGKPVVIPKEDQEKM
KEMFKINQFNLMASEMIAFNRSLPDVRLEGCKTKVYPDSLPTTSVVIVFHNEAWSTLLRTVHVSINRSPRHMIEE
IVLVDDASERDFLKRPLESYVKKLKVPVHVIRMEQSRGLIRARLKGAAVSKGQVITFLDAHCECTVGWLEPLLAR
IKHDRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKLNFRWYPVPQREMDRRKGDRTL PVRTPTMAGGLFSIDR
DYFQEIGTYDAGMDIWGGENLEISFRIWQCGGTLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMD
EFKNFFYIISPQVTKVDYGDISSRVGLRHKLQCKPFSWYLENIYPDSQIPRHYFSLGEIRNVETNQCLDNMARKE
NEKVGIFNCHGMGGNQVFSYTANKEIRTDCLDVLKNGPVTMLKCHHLKGNQLWEYDPVKLTLOHVNSNQCLD
KATEEDSQVPSIRDCTGSRSSQWLLRNVTLPETIF

SEQ ID NO: 133.

UDP-N-acetilgalactosaminiltransferasa 3 (GalNAc-T3) humana

MAHLKRLVKLHIKRHYHKFWKLGAIVFFFIIVLVLMQREVSQYSKEESRMERNMKNKNKMLDLMLEAVNNIKD
AMPKMQIGAPVRQONIDAGERPCLOQYYTAAELKPVLDPRPQDSNAPGASGKAFKTTNLSVEEQKEKERGEAKHCF
NAFASDRISLHRDLGPDTRPPECIEQKFKRCPLPTTSVIVFHNEAWSTLLRTVHVSILYSSPAILLKEIILVDD
ASVDEYLHDKLDEYVKQFSIVKIVRQERKGLITARLLGATVATAETLTFLDAHCECFYGWLEPLLARIAENYTA

5

VVSPDIASIDLNTFEFNKPSPYGSNNHNRGNFDWSLSFGWESLPDHEKQRRKDETYPIKTPTFAGGLFSISKEYFE
YIGSYDEEMEIWGGENIEMSFVRWQCGGLEIMPCSVVGHVFRSKSPHSFPKGTQVIARNQVRLAEVWMDYKEI
FYRNTDAKIVKQKAFGLSKRFEIKHRLRCKNFTWYLNNIYPEVYVPDLNPVISGYIKSVGQPLCLDVGENNO
GGKPLIMYTCHGLGNGQYFEYSAQHEIRHNIQKELCLHAAQGLVQLKACTYKGHKT VVTGEQIWEIQDQLLYNP
FLKMCLSANGEHPSLVSCNPSDPLQKWILSQND

SEQ ID NO: 134.

UDP-N-acetilgalactosaminiltransferasa 3 (GalNAc-T3) de *Drosophila*

MGLRFQQLKKLWLLYLFLFFAFFMFASINLYVASIQGGDAEMRHPKPPPKRRSLWPHKNIVAHYIGKGDIFGN
MTADDYNINLFPINGEGADGRPVVPPRDRFRMRFRNLNSFNLLASDRIPLNRTLKDYRTPECRDKKYASGLP
STSVIIVFHNEAWSVLLRTITSVINRSPRHLLKEIILVDDASDRSYLKRQLESYVKVLAVPTRIFRMKKRSGLVP
ARLLGAENARGDVLTFDAHCECSRQWLEPLLRIKESRKVVICPVIDIISDDNFSYTKTFENHWGAFNWQLSFR
WFSSDRKRQTAGNSSKDSTDPATPGMAGGLFAIDRKYFYEMGSYDSNMVRVWGGENVEMSFRIWQCGGRVEISPC
SHVGHVFRSSTPYTFPGMSEVLTDNLARAATVWMDWQYFIMLYTSGTLGAKDKVNVTERVALRERLQCKPFS
WYLENIWPEHFFPAPDRFFGKIIWLDGETECAQAYSKHMKNLPGRLSREWKRAFEEIDSKAEELMALIDLERDK
CLRPLKEDVPRSSLSAVTVGDCTSHAQSMDFVITPKGQIMTNDNVCLTYRQQLGVIKMLKNNRATTSSNVMQAQ
CASDSSQLWTYDMDTQQISHRDTKLCLTLKAATNSRLQKVEKVLSMECDFKDITQKWGFIPLPWRM

10 SEQ ID NO: 135.

UDP-N-acetilgalactosaminiltransferasa 3 (GalNAc-T3) de ratón

MAHLKRLVKLHIKRHYHRKFWKLGAIVFFFLVVLILMQREVSVQYSKEESKMERNLKNKNKMLDFMLEAVNNIKD
AMPKMQIGAPIKENIDVRERPCLQGYTTAAELKPVFDRPPQDSNAPGASGKPFKITHLSPEEQKEKERGETKHCF
NAFASDRISLHRDLGPDTRPPECIEQKFKRCPLPTTSVIVFHNEAWSTLLRTVHVSPLYSSPAILLKEIILVDD
ASVDDYLHEKLEEYIKQFSIVKIVRQQRKGLITARLLGAAVATAETLTFLDAHCECFYGWLEPLLARIAENYTA
VVSPDIASIDLNTFEFNKPSPYGNNHNRGNFWDLSLFGWESLPDHEKQRRKDETYPIKTPTFAGGLFSISKKYFE
HIGSYDEEMEIWGGENIEMSFVRWQCGGQLEIMPCSVVGHVFRSKSPHTFPKGTQVIARNQVRLAEVWMDEYKEI
FYRRNTDAKIVKQKSFGLSKRFEIKKRLQCKNFTWYLNITYPEAYVPDLNPVISGYIKSVGQPLCLDVGENNQ
GGKPLILYTCHGLGGNQYFEYSAQREIRHNIQKELCLHATQGVVQLKACVYKGHRTIAPGEQIWEIRKDQLLYNP
LFKMCLSSNGEHPNLVPCDATDLLQKWIFSQND

SEQ ID NO: 136.

GalT al, 4

alfa 1,4-galactosiltransferasa, A4GALT humana

MSKPPDLLLRLLRGAPRQVCTLFIIIGFKFTFFVSIMYWHVVGEPKEKGQLYNLP AEIPCPTLTPTPTPSHGPT
PGNIFFLETSDRTNPNFLFMCVESAARTHPESHVLVLMKGLPGGNASLPRHLGISLLSCFPNVQMLPLDLRELF
RDTPLADWYAAVQGRWEPYLLPVLSASRIALMWKFGGIYLDTFIVLKNLRNLTNVLGTQSRVVLNGAFLAFER
RHEFMALCMRDFVDHYNGWIWGHQGPQLLTRVFKKWCIRS LAESRACRGVTTLPEAFYPIPWQDWKKYFEDIN
PEELPRLLSATYAVHVWNKKSQGTREATS RALLAQLHARYCPTTHEAMKMYL

5

SEQ ID NO: 137.

Transferasa A de grupo sanguíneo ABO humano

MAEVLRTLAGKPKCHALRPMILFLIMLVLVLFYGYVLSPRSLMPGSLERGFCAVREPDHLQRVSLPRMVYPQPK
VLTPCRKDVLVVTPWLAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVAFLKLFLETAEKHFVGHVHY
VFTDQPAAPRVTLGTGRQLSVLEV RAYKRWQDVSMRRMEMISDFCERRFLSEVDYLCVDVDMEFRDHVGEIL
TPLFGTLHPGFYGS SREAFYERRPQS QAYIPKDEGDFYLLGGFFGGSVQEVQRLTRACHQAMMVDQANGIEAVW
HDESHLNKYLLRHKPTKVLSP EYLWDQQLLGWPAVLRKL RFTAVPKNHQAVRNP

SEQ ID NO: 138.

10 beta-1,3-galactosiltransferasa 1, B3GALT1 humana

MASKVSCLYVLTVVCWASALWYLSITRPTSSYTGSKPFSHLTVARKNFTFGNIRTRPINPHSFEFLINEPNKCEK
NIPFLVILISTTHKEFDARQAIRETWGDENNFKGIKIATLFLLGKNADPVLNQMVEQESQIFHDIIVEDFIDSYH
NLTLKTLGMGRWVATFCSKAKYVMKTDSDIFVNMDNLIYKLLKPSTKPRRRYFTGYVINGGPIRDVRSKWYMPRD
LYPDSNYPPFCSGTGYIFSADVAELIYKTS LHTRLLHLEDVYVGLCLRKLGIHPFQNSGFNHWKMAYSLCRYRRV
ITVHQISPEEMHRIWNDMSSKKHLRC

SEQ ID NO: 139.

beta-1,3-galactosiltransferasa 1, B3Galt1 de ratón

MASKVSCLYVLTVVCWASALWYLSITRPTSSYTGSKPFSHLTVARKNFTFGNIRTRPINPHSFEFLINEPNKCEK
NIPFLVILISTTHKEFDARQAIRETWGDENNFKGIKIATLFLLGKNADPVLNQMVEQESQIFHDIIVEDFIDSYH
NLTLKTLGMGRWVATFCSKAKYVMKTDSDIFVNMDNLIYKLLKPSTKPRRRYFTGYVINGGPIRDVRSKWYMPRD
LYPDSNYPPFCSGTGYIFSADVAELIYKTS LHTRLLHLEDVYVGLCLRKLGIHPFQNSGFNHWKMAYSLCRYRRV
ITVHQISPEEMHRIWNDMSSKKHLRC

15 SEQ ID NO: 140.

beta-1,3-N-acetilglucosaminiltransferasa 1, B3GNT1 humana

MSVGRRRIKLLGILMMANVFIYFIMEVSKSSSQEKNKGGEVVIIPKEKFWKISTPPEAYWNREQEKLNRQYNPILS
MLTNQTGEAGRLSNISHLNYCEPDLRVTSVVTGFFNNLPDRFKDFLLYLRCRNYSLIDQPKCAKKPFLLLAICS
LTPHFARRQAIRESWGQESNAGNQTIVRVFLLGQTPPEDNHPDLSDMLKFESEKHQDILMWNYRDTFFNLSLKEV
LFLRWVSTSCPDTEFVFKGDDDFVNTHHILNLYNSLSKTKAKDLFIGDVIHNAGPHRDKKLYYIPEVYVSGLY
PPYAGGGGFLYSGHLALRLYHITDQVHLYPIDDVYTMCLQKLGVLPEKHKGFRFTDIEEKNKNICSIVDLMLV
HSRKPQEMIDIWSQLQSAHLKC

SEQ ID NO: 141.

beta-1,3-N-acetilglucosaminiltransferasa 1, B3gnt1 de ratón

MSVGRRRVKLLGILMMANVFIYLIVEVSKNSSQDKNGKGGVVIIPKEKFWKPPSTPRAYWNREQEKLNRWYNPILN
RVANQTGELATSPNTSHLSYCEPDSTVMTAVTDFNNLPDRFKDFLLYLRCRNYSLIDQPKCAKKPFLLLAICS
LIPHFARRQAIRESWGRETNGVGNQTVRVFLLGKTPPEDNHPDLSDMLKFESDKHQDILMWNYRDTFFNLSLKEV
LFLRWVSTSCPDAEFVFKGDDDFVNTHHILNLYNSLSKSKAKDLFIGDVIHNAGPHRDKKLYYIPEVFYTGIVY
PPYAGGGGFLYSGPLALRLYSATSRVHLYPIDDVYTMCLQKLGVLPEKHKGFRFTDIEEKNKNICSIVDLMLV
HSRKPQEMIDIWSQLQSPNLKC

SEQ ID NO: 142.

N-acetilglucosaminiltransferasa 2 (GnT-II) humana

MRFRIYKRKVLILTLVVAACGFVLWSSNGRQKNEALAPPLDAEPARGAGGRGGDHPSVAVGIRRVSNVSAASL
VPAVPQPEADNLTLYRSLVYQLNFDQTLRNVDKAGTWAPRELVLVVQVHNRPYLRLLDLSLRKAQGIDNVLVI
FSHDFWSTEINQLIAGVNFVCPVLQVFFPFSIQLYPNEFPFGSDPRDCPRDLPKNAALKGCINAEPDSFGHYREA
KFSQTKHHWWKHLHFVWERVKILRDYAGLILFLEEDHYLAPDFYHVFKMWKLKQECPECDVLSLGTYSASRSF
YGMADKVDVKTWKSTEHNMGALALTRDAYQKLIECTDTFCTYDDYNWDWTLQYLTVSCLPKFWKVLVPQIPRIFHA
GDCGMHHKKTCTPSTQSAQIESLLNNKQYMPETLTISEKFTVVAISPFRKNGGWGDIRDHELCKSYRRLQ

SEQ ID NO: 143.

N-acetilglucosaminiltransferasa 2 (GnT-II) de cerdo

MRFRIYKRKVLILTFVVAACGFVLWSSNGRQKNEALAPPLDAEPVRGAGARAGDHPAISVGIRRGSNDSAAPL
VAAAPQPEVDNLTLYRSLVYQLNFDQTLRNVDKVSWSVPRELVLVVQVHNRAEYLRLLDLSLRKAQIDNVLVI
FSHDFWSTEINQLIAGVDFCPVLQVFFPFSIQLYPNEFPFGTDPDCPRDLEKNAALKMGCINAEPDSFGHYREA
KFSQTKHHWWKHLHFVWERVKVLRDYAGLILFLEEDHYVAPDFYHVFKMWNLKQECPECDVLSLGTYYTTRVSF
RDVADKVDVKTWKSTEHNMGALALTRDAYQKLIECTDTFCTYDDYNWDWTLQYLTVSCLPKFWKVLVPQVPRIFHA
GDCGMHHKKTCTPSTQSAQIESLLNSNKQYMPETLTISEKLTAAISPFRKNGGWGDIRDHELCKSYRRLQ

SEQ ID NO: 144.

N-acetilglucosaminiltransferasa 2 (GnT-II) de ratón

MRFRIYKRKVLILTLVVAACGFVLWSSNGRQKSDALGPPLDAEPVRGAGHLAVSVGIRRVSNESAAPLVPAVP
RPEVDNLTLYRSLVYQLNFDQMLRNVDGDTWSGELVLVVQVHNRPYLRLLIDSLRKAQGIQEVLVIFSHDF
WSAEINSLISRVDVFCPVLQVFFPFSIQLYPNEFPFGSDPRDCPRDLKKNAAALKGCINAEPDSFGHYREAKFSQT
KHHWWKHLHFVWERVKVLQDYTGILFLEEDHYLAPDFYHVFKMWKLKQECPCGCDVLSLGTYYTTRSFYGIAD
KVDVKTWKSTEHNMGALALTRDAYQKLIECTDTFCTYDDYNWDWTLQYLTACLPKIWKVLVPQAPRIFHAGDCGM
HHKKTCTPSTQSAQIESLLNSNKQYLPETLVIGEKFPMMAISPFRKNGGWGDIRDHELCKSYRRLQ

SEQ ID NO: 145.

N-acetilglucosaminiltransferasa 2 (GnT-II) de rata

MRFRYYKRVLIITLVVAACGFVLWSSNGRQRKNDALAPPLLDSEPLRGAGHFAASVGIRRVSNDSAAPLVPAVP
RPEVDNLTLYRSLVYQLNFDQMLRNVDKDGTSWPGELVLVQVHNRPEYLRLIDSLRKAQGIREVLVIFSHDF
WSAEINSLISSVDFCPVLQVFFPFSIQLYPSEFPGSDPRDCPRDLKKNAALKLGCINAEPDSFGHYREAKFSQT
KHHWWKLFHVWERVKVLQDYTGILFLFLEEDHYLAPDFYHVFKMWKWKQCECPGCDVLSLGTYYTIRSFYGIAD
KVDVKTWKSTEHNMGALALTRDAYQKLIECTDTFTCTYDDYNWDWTLQYLTACLPLKVKVLVPQAPRI FHAGDCGM
HHKKTCPSTQSAQIESLLNNKQYLPFETLVIGEKFPMMAISPPRKNGGWGDIRDHELCKSYRRLQ

SEQ ID NO: 146.

N-acetilglucosaminiltransferasa 3 (GnT-III) humana

MKMRRYKFLMFCMAGLCLISFLHFFKTL SYVTFFPRELASLSPNLVSSFFWNNAPVTPQASPEPGGPDLLRTPLY
SHSPLLQPLPPSKAAEELHRVDLVLPEDTTEYFVRTKAGGVCFKPGTKMLERPPPGRPEEKPEGANGSSARRPFR

5

YLLSARERTGGRGARRKWVECVCLPGWHGPSCGVPTVVQYSNLPTKERLVPREVPRRVINAINVNEHFDLLDVR
HELGDVDAFVVCESNFTAYGEPRPLKFREMLTNGTFEYIRHKVLYVFLDHFPFGGRQDGIADYLRFTLTQDG
VSRLRLRPDDVFI IDDADEIPARDGVLFLKLYDGWTEPFAFHMRSYLYGFFWKQPGTLEVVSCTMDMLQAVYG
LDGIRLRRRQYYTMPNFRQYENRTGHILVQWSLGSPLHFAGWHCSWCFTPEGIYFKLVSAQNGDFPRWGDYEDKR
DLNYIRGLIRTGGWFDGTQQEYPPADPSEHMYAPKYLLKNYDRFHYLLDNPHYQEPRSTAAGGWRHRGPEGRPPAR
GKLDEAEV

SEQ ID NO: 147.

N-acetilglucosaminiltransferasa 3 (GnT-III) de ratón

MKMRRYKFLMFCMAGLCLISFLHFFKTL SYVTFFPRELASLSPNLISSFFWNNAPVTPQASPEPGDPLLRTP
SHSPLLQPLSPSKATEELHRVDFVLPEDTTEYFVRTKAGGVCFKPGTRMLEKPSPGRTEEKTEVSEGSARGPAR
RPMRHVLSRERLGRGTRRKWVECVCLPGWHGPSCGVPTVVQYSNLPTKERLVPREVPRRVINAININHEFDLL
DVRFHELGDVDAFVVCDSNFTAYGEPRPLKFREMLTNGTFEYIRHKVLYVFLDHFPFGGRQDGIADYLRFTLT
TQDGVSRRLRLRPDDVFI IDDADEIPARDGVLFLKLYDGWTEPFAFHMRSYLYGFFWKQPGTLEVVSCTMDMLQ
AVYGLDGIRLRRRQYYTMPNFRQYENRTGHILVQWSLGSPLHFAGWHCSWCFTPEGIYFKLVSAQNGDFPRWGDY
EDKRDLYIRSLIRTGGWFDGTQQEYPPADPSEHMYAPKYLLKNYDQFRYLLENPHYREPSTVEGGRQNGSDGR
SSAVRGKLDTAEG

10

SEQ ID NO: 148.

N-acetilglucosaminiltransferasa 3 (GnT-III) de rata

MKMRRYKFLMFCMAGLCLISFLHFFKTL SYVTFFPRELASLSPNLISSFFWNNAPVTPQASPEPGDPLLRTP
SHSPLLQPLSPSKATEELHRVDFVLPEDTTEYFVRTKAGGVCFKPGTRMLEKPSPGRTEEKTKVAEGSSVRGP
RPMRHVLSARERLGRGTRRKWVECVCLPGWHGPSCGVPTVVQYSNLPTKERLVPREVPRRVINAININHEFDLL
DVRFHELGDVDAFVVCESNFTAYGEPRPLKFREMLTNGTFEYIRHKVLYVFLDHFPFGGRQDGIADYLRFTLT
TQDGVSRRLRLRPDDVFI IDDADEIPARDGVLFLKLYDGWTEPFAFHMRSYLYGFFWKQPGTLEVVSCTIDMLQ
AVYGLDGIRLRRRQYYTMPNFRQYENRTGHILVQWSLGSPLHFAGWHCSWCFTPEGIYFKLVSAQNGDFPRWGDY
EDKRDLYIRSLIRTGGWFDGTQQEYPPADPSEHMYAPKYLLKNYDQFRYLLENPHYREPSTVEGGRNQS
SSAVRGKLDTTEG

SEQ ID NO: 149.

15 N-acetilglucosaminiltransferasa 4 (GnT-IV) de ratón

MLKFYQMKYIFQILDKMRCLRKSTVSFLGVLVVFLLFMNLYIEDSYVLEGDQKQLIRETSTHQLNSERYVHTFKD
LSNFSGTINVITYRYLAATPLQKRKYLITIGLSSVKKKGNLYLDTIKSIFEQSSYEELKEISVVVHLADFNSSWRD
AMVQDITQKFAHHIAGRLMVIHAPPEEYYPVLDGLKRNYPEDRVRFRSKQNVDAFLLNFCANTSDDYYVMLED
DVRCSRNLTAIKKVIASLEGTYWVTFLEFSKLGYGKLYHSHDLPRLAHFLLMFYQEMPCDWLLTHFRGLLAQKN
VIRFKPSLFQHMGGYSSYKGTENKLDKDDFEEESFDIPDNPPASFYTNMNVFENYEASKAYSSVDEYFWGKSPSM
GDTFVIVFENPITIKKIKVNTGTEDRQNDILQHGALDVGEKLIFFSKQIRQCDTYLRLGEFKNGYFEMSDVNQKIP
FDIHCMRICVTKTQKEWLIIRSISIWTS

SEQ ID NO: 150.

N-acetilglucosaminiltransferasa (GnT-IV) de cerdo

MFKFHGVKHIFEILDKMRCLRKFTVSFLGVLVIFLLFMNLYIEDSYVLEGDQKQLVRETSIHQLNPERYVHTFKY
LSNFSGTINVITYRYLAVMPFQKRFLITIGLSSVRRKKGNLYLETIKSIFEQSSYEELKEISVVVHLADFNSSWRE
VMVQDITQKFAHHIAGRLMVIHAPPEEYYPVLNGLKRNYPEDRVRFRSKQNVDAFLLNFCANISDDYYVMLED
DVRCSKNFLTAIKKVITSLQGTWVTFLEFSKLGYGKLYHSHDLPRLAHFLLMFYQEMPCDWLLTHFRGLLAQKN
VIRFKPSLFQHMGGYSSYKGTENKLDKDDFEEELIDLDPNPPASLYTNMSVFENYDASKAYSSVDGYFWGKPPST
GDVFFVVFFENPVIIKKIKVSTGTEDRQNDILHGHGALDVGGYIRYFKQNRQCITYIRLGEFKNGNFVSDVNQKIP
FDVHCMRIHVTKTQKEWLIIRSISIWTS

5 SEQ ID NO: 151.

N-acetilglucosaminiltransferasa 4a (GnT-IVa) humana

MRLRNGTVATALAFITSFLTLSWYTTWQNGKEKLIAYQREFLALKERLRVIAEHRISQRSSELNTIVQQFKRVGAE
TNGSKDALNKPFSNDTLKLLKELTSKKSQVPSIYYHLPHLLKNEGSLQPAVQIGNRTGVSIVMGIPTVKREVKS
YLIELTLHSLIDNLYPEEKLDKVIIVFVIGETDIDYVHGUVANLEKEFSKEISSGLVEVISPPESYYPDLTNLKETF
GDSKERVRWRWKQNLDYCFLLMYAQEKGIYYIQLEDDIIVKQNYFNTIKNFALQLSSEEWMIFFSGLGFIGKMF
QAPDLTLIVBFIFMFYKEKPIDWLLDHILWVKVCNPEKDAKHCDRQKANLRIRFRPSLFQHVGLHSSLSGKIQKL
TDKDYMKPLLLKIHVNPPAEVSTSLKVYQGHTEKTYMGEDFFWAITPIAGDYILFKFDKPVNVESYLFHSGNQE
HPGDILLNTTVEVLFPKSEGLEISKETKDKRLEDGYFRIGKFENGVAEGMVDPSLNPISAFRLSVIQNSAVWAIL
NEIHIKKATN

SEQ ID NO: 152.

N-acetilglucosaminiltransferasa 4a (GnT-IVa) de ratón

MRLRNGTVATALVFVTSFLTLSWYTTWQNGKEKLIAYQREFLALKERLRVAEHRISQRSSELNTIVQQFRRAGAE
TNGSKTALSTISDNTIKLLKELTSKKSRLVPSIYYHLPHLLQNERSLQPAVQIGSGRTGVSIVMGIPTVKREVKS
YLVELTLHSLIDNLYPEEKLDKVIIVFVIGETDLDYVHVSVANLEKEFSREISSGLLEIISPESYYPDLTNLKETF
GDSKERVRWRWKQNLDYCFLLMYAQEKGIYYIQLEDDIIVKQNYFNTIKNFALQLSSEEWMIFFSGLGFIGKMF
QAPDLALVVEFILMFYKEKPIDWLLDHILWVKVCNPEKDAKHCDRQKANLRIRFRPSLFQHVGLHSSLSGKIQKL
TDKDYMKPLLLKVHVNPPEAEVSTSLKVYQGHTEKTYMGEDFFWAITPTAGDYILFKFDKPVNVESYLFHSGNQE
HPGDILLNTTVDVLPKSDSLEISKETKDKRLEDGYFRIGKFEGYVAEGIVDPGLNPISAFRLSVIQNSAVWAIL
NEIHIKKVTS

10

SEQ ID NO: 153.

N-acetilglucosaminiltransferasa 4b (GnT-IVb) humana

MRLRNGTFLTLTLFCLCAFLSLSWYAALSGQKGDVVDVYQREFLALRDLHAAEQESLKRSKELNLVLDEIKRAV
SERQALRDGDGNRTWGRLTEDPRLKPWNGSHRHVLHLPVTFHHLPHLLAKESSLQPAVRVGQGRGTGVSVVMGIPS
VRREVHSYLTDTLHSLISELSPQEKEDSVIVVLIETDSQYTSAVTENIKALFPTEIHSGLLEVISPSPHFYPDF
SRLRESFGDPKERVWRWTKQNLDYCFLLMYAQSKGIYYVQLEDDIVAKPNYLSTMKNFALQQPSEDWMILEFSQL
GFIGKMFKSLDLSLIVEFILMFYRDKPIDWLLDHILWVKVCNPEKDAKHCDRQKANLRIRFKPSLFQHVGTSSSL
AGKIQKLKDKDFGKQALRKEHVNPPAEVSTSLKTYQHFTLEKAYLREDDFFWAFTPAAGDFIRFRFFQPLRLERFF

FRSGNIEHPEDKLFNTSVEVLPFDNPQSDKEALQEGRTATLRYPRSPDGYLQIGSFYKGVAEGEVDPAFGPLEAL
RLSIQTDSPVWVILSEIFLKKAD

SEQ ID NO: 154.

N-acetilglucosaminiltransferasa 4b (GnT-IVb) de ratón

MRLRNGTFLTLTLFCLCAFLSLSWYAALSGQKGDVVDIYQREFLALRDLHAAEQESLKRSKELNLVLEEIKRAV
SERQALRDGEGRNRTWGRLTEDPRLKPWNVSHRHVLHLPVTFHHLPHLLAKESSLQPAVRVGQGRGTGVSVVMGIPS
VRREVHSYLTDTLHSLISELSPQEKEDSVIVVLIETDPQYTSAVTENIKALFPTEIHSGLLEVISPSPHFYPDF
SRLRESFGDPKERVWRWTKQNLDYCFLLMYAQSKGIYYVQLEDDIVAKPNYLSTMKNFALQQPSEDWMILEFSQL
GFIGKMFKSLDLSLIVEFILMFYRDKPIDWLLDHILWVKVCNPEKDAKHCDRQKANLRIRFKPSLFQHVGTSSSL
AGKIQKLKDKDFGKHARKEHVNPPAEVSTSLKTYQHFTLEKAYLREDDFFWAFTPAAGDFIRFRFFQPLRLERFF
FRSGNIEHPEDKLFNTSVEVLPFDNPQSEKEALQEGRSATLRYPRSPDGYLQIGSFYKGVAEGEVDPAFGPLEAL
RLSIQTDSPVWVILSEIFLKKAD

5

SEQ ID NO: 155.

N-acetilglucosaminiltransferasa 5 (GnT-V) humana

MALFTPWKLSSQKLGFFLVTFGFIWGMMLLHFTTIQRTQPESSSMLREQILDLSKRYIKALAEENRNVDGPYAG
VMTAYDLKKTAVLLDNILQRIGKLESKVDNLVVGNGTGTNSTNSTTAVPSLVALEKINVADIINGAQEKCVLPPM
DGYPHCEGKIKWMKDMWRSDPCYADYGVGDSTCSFFIYLSEVENWCPHLPWRAKNPYEADHNSLAETIRTDNFIL
YSMMKKHEEFWRMLRIRRMADAWIQAISLAEKQNLKRRKRVLHVLGLLTKESGFKIAETAFSGGPLGELVQ
WSDLITSLYLLGHDIRISASLAEKEIMKKVGNRSGCPTVGDRIVELIYIDIVGLAQFKKTLGPSWVHYQCMRL
VLDSFGTEPEFNHANYAQSKGHKTPWGKWNLPQQFYTMFPTPDNSFLGFVVEQHLNSSDIHHINEIKRONQSL
VYGKVDSEFWKNKKIYLDIIHTYMEVHATVYGSSTKNIPSYVKNHGILSGRDLQFLLRETKLFVGLGFPYEGPAPL
EAIANGCAFLNPKFNPPKSSKNTDFFIGKPTLRELTSHQHPYAEVFIGRPHVWTVDLNNQEEVEDAVKAILNQKIE
PYMPYEFTCEGMLQRINAFIEKQDFCHGQVMWPPLSALQVKLAEPGQSCQVCQESQLICEPSFFQHLNKKDKML
KYKVTCQSSELAKDILVPSFDPKNKHCVFQGDLLLFSAGAHPRHQRVCPCRDFIKGQVALCKDCL

10 SEQ ID NO: 156.

N-acetilglucosaminiltransferasa 5 (GnT-V) de ratón

MAFFSPWKLSSQKLGFFLVTFGFIWGMMLLHFTIQQRTQPESSSMLREQILDLSKRYIKALAEENRDVVDGPYAG
 VMTAYDLKKTTLAVLLDNILQRIQKLESKVDNLVNGTGANSTNSTTAVPSLVLEKINVADIINGVQEKCVLPMPD
 GYPHCEGKI KWMKDMWRSDPCYADYGVDTSCSFFIYLSEVENWCPRLPWRAKNPYEADHNSLAEIRTDNFNILY
 GMMKKHEEFWRMLRIRRMADAWIQAISLAEKQNLKRRKKILVHLGLLTKESGFKIAETAFSGGPLGELVQW
 SDLITSLYLLGHDIRISASLAELKEIMKKVGNRSGCPTVGDRIVELIYIDIVGLAQFKKTLGPSVWHYQCMLRV
 LDSFGTEPEFNHASYAQSKGHKTPWGKWNLPQQFYTMFPHTPDNSFLGFVVEQHLNSSDIHHINEIKRQNQSLV
 YGKVDSFWKNKKIYLDIIHTYMEVHATVYGSSTKNIPSYVKNHGILSGRDLQFLLRETKLFVGLGFPPYEGPAPLE
 AIANGCAFLNPKFNPPKSSKNTDFFIGKPTLRELTSQHPYAEVFIGRPHVWTVDLNNREEVEDAVKAILNQKIEP
 YMPYEFTCEGMLQRINAFIEKQDFCHGQVMWPPLSALQVKLAEPGQSCQVCQESQLICEPSFFQHLNKEKDLLK
 YKVTCSSELYKDILVPSFYPKSKHCVFQGDLLLFSCAGAHPTHQRICPCRDFIKGQVALCKDCL

SEQ ID NO: 157.

N-acetilglucosaminiltransferasa 5 (GnT-V) de rata

MAFFSPWKLSSQKLGFFLVTFGFIWGMMLLHFTIQQRTQPESSSMLREQILDLSKRYIKALAEENRNVDGPYAG
 VMTAYDLKKTTLAVLLDNILQRIQKLESKVDNLVNGTGANSTNSTTAVPSLVLEKINVADIINGVQEKCVLPMPD
 GYPHCEGKI KWMKDMWRSDPCYADYGVDTSCSFFIYLSEVENWCPRLPWRAKNPYEADHNSLAEIRTDNFNILY
 GMMKKHEEFWRMLRIRRMADAWIQAISLAEKQNLKRRKKILVHLGLLTKESGFKIAETAFSGGPLGELVQW
 SDLITSLYLLGHDIRISASLAELKEIMKKVGNRSGCPTVGDRIVELIYIDIVGLAQFKKTLGPSVWHYQCMLRV
 LDSFGTEPEFNHASYAQSKGHKTPWGKWNLPQQFYTMFPHTPDNSFLGFVVEQHLNSSDIHHINEIKRQNQSLV
 YGKVDSFWKNKKIYLDIIHTYMEVHATVYGSSTKNIPSYVKNHGILSGRDLQFLLRETKLFVGLGFPPYEGPAPLE
 AIANGCAFLNPKFNPPKSSKNTDFFIGKPTLRELTSQHPYAEVFIGRPHVWTVDLNNREEVEDAVKAILNQKIEP
 YMPYEFTCEGMLQRINAFIEKQDFCHGQVMWPPLSALQVKLAEPGQSCQVCQESQLICEPSFFQHLNKEKDLLK
 YKVICQSSELYKDILVPSFYPKSKHCVFQGDLLLFSCAGAHPTHQRICPCRDFIKGQVALCKDCL

5 SEQ ID NO: 158.

N-acetilglucosaminiltransferasa 6 (GnT-VI) de pollo

MRCSPKRSLTAVIAASFLLLLLLLLLLHGRSQDQPQEVQFRDLPSDAVLKILKQGSLLHILQDTDNLCAHNI SYHL
 LAGSPLPHKKFLAVGLSSVRRPRGYLPDTLQSLFKQSSEELQEMVVVHLADADPIWNAQVAADISHRFAHHI
 LLGRLVLIHTPHEFYPTLEGLKRNNDPEERVKFRSKQNVDAFLFTFAANLSSYYLMIEDDVWSAKSFFTAIRK
 AVASQEGSNWATLEFSKLG YIGKLYRSSDLPRLARFLLLFYQEMPCDWLLTHFRLLLTQKDVIRFKPSL FQHMG
 YSSFQGTVNRLEDDEFQADAMDLPDNPAAFLTMMVVFENYEPSKAYSTARGYFWGKNPAVGSI FSVFHQPARV
 TRVRVQTGSSERPGDFLHAGVLELGRGRRADGRDCSVYTTVTGTFEKG NLEWRGLEKGMNPVECVRIRVTQSQSE
 WLIIQSIGIWTAGT

SEQ ID NO: 159.

beta-1,4-N-acetilglucosaminiltransferasa IV, GNTIVH humana

MFKFHQMKHIFEILDKMRCLRKRSTVSFLGVLVIFLLFMNLYIEDSYVLEGDKQLIRETSTHQLNSERYVHTFKD
 LSNFSGAINVTYRYLAATPLQRKRYLTIGLSSVKRKKGNLYLETIKSIFEQSSYEELKEISVVVHLADFNSWRD
 AMVQDITQKFAHHIIAGRLMVIHAP EYYPI LDGLKRNNDPEDRVKFRSKQNVDAFLNFCANTS DYYVMLED
 DVRCSKNFLTAKKVIASLEGTYWVTFEFSKLG YIGKLYHSHDLPRLAHFLLMFYQEMPCDWLLTHFRGLLAQKN
 VIRFKPSL FQHMGYYSSYKGTENKLDKDDFEEESFDIPDNPPASLYTNMNVFENYEASKAYSSVDEYFWGKPPST
 GDVFVIVFENPII IKKIKVNTGTEDRQNDILHGHGALDVGENVMP SKQRRQCSSYLRLGEFKNGNFEMSGV NQKIP
 FDIHCMRIYVTKTQKEWLIIRSISIWTS

10

SEQ ID NO: 160.

beta-1,4-N-acetilglucosaminiltransferasa IV, Gnt-IV (GNTIVH) de cerdo

MFKFHQVKHIFEILDKMRLRKRFTVSFLGVLVIFLLFMNLYIEDSYVLEGDKQLVRETSIHQLNPERYVHTFKY
LSNFSGTINVTRYRLAVMPFORKRFLTIGLSSVRRKKNYLLLETIKSIFEQSSYEELKEISVVVHLADFNSSWRE
VMVQDITQKFAHHI IAGRLMVIHAPEEYYPVLNGLKRNNDPEDRVRFRSKQNVDAFLNFCANISDYVVMLED
DVRCSKNFLTAKKVITSLQGTYWVTLFESKLGYGKLYHSHDLPRLAHFLLMFYQEMPCDWLLTHFRGLLAQKN
VIRFKPSLFQHMGGYSSYKGTENKLDKDDFEELIDLDPNPPASLYTNMSVFENYDASKAYSSVDGYFWGKPPST

GDVFFVVFENPVIKKIKVSTGTEDRQNDILHGHGALDVGGYIRYFKQNRQCITYIRLGEFKNGNFEVSDVNQKIP
FDVHCMRIHVTKTQKEWLIIRSISIWTS

SEQ ID NO: 161.

5 Dolichil-fosfato beta-glucosiltransferasa, ALG5 humana

MAPLLLQLAVLGAALAAAALVLISIVAFTTATKMPALHRHEEEKFFLNAKGQKETLPSIWDSPKQLSVVPSYN
EEKRLPVMDEALSYLEKRQKRDPATYEVIVDDGSKDQTSKVAFKYCQKYGSDKVRVITLVKNRGKGGAIRMG
IFSSRGEKILMADADGATKFPDVEKLEKGLNDLQWPQNMAIACGSRAHLEKESIAQRSYFRTLLMYGFHFLVWF
LCVKGIRDTCGFKLFTREASRTFSSHLVERWAFDVELLYIAQFFKIPAEIAVNWTEIEGSKLVFPFWSWLQMG
KDLLFIRLRYLTGAWRLEQTRKMN

SEQ ID NO: 162.

Polipéptido de dolichil-fosfato manosiltransferasa 1, Dpm1 humana

MASLEVSRSPPRRSRRELEVRSPRQNKYSVLLPTYNERENLPLIVWLLVKSFSESGINYEIIIDDGSPDGTDRVA
EQLEKIYGSDRILLRPREKKLGLGTAYIHGMKHATGNYYIIMDADLSHHPKFIPEFIRKQKEGNFDIVSGTRYKG
NGGVYGDWLKRKIIISRGANFLTQILLRPGASDLTGFSRLYRKEVLEKLIBKCVSKGYVFQMEMIVRARQLNYTIG
EVPISFVDRVYGESKLGNEIVSFLKGLLTLFATT

10 SEQ ID NO: 163.

alfa-1,6-manosiltransferasa, ALG12 humana

MAGKGSSGRPLLLGLLVAVATVHLVICPYTKVEESFNLQATHDLLYHWQDLEQYDHLFPGVVPRTFLGPVVIA
VFSSPAVYVLSLLEMSKFYSQLIVRGVLGLGVIFGLWTLQKEVRRHFGAMVATMFCWVTAMQFHLMFYCTRTPN
VLALPVLVLLAALAWRHEWARFIWLSAFAIIVFRVELCLFLGLLLLLLALGNRKVSVVRLRHAVPAGILCLGLTV
AVDSYFWRQLTWPEGKVLWYNTVLNKSNNWGTSPLLWYFYSALPRGLGCSLLFIPLGLVDRRTHAPTVALGFMMA
LYSLLPHKELRFIYAFPMNLNITAARGCSYLLNNYKKSPLYKAGSLLVIGHLVVNAAYSATALYVSHFNYPGGVA
MQRLHQLVPPQTDVLLHIDVAAAQTVSRFLQVNSAWRYDKREDVQPGTGMLAYTHILMEAPGLLALYRDTHRV
LASVVGTTGVSLNLTQLPPFNVHLQTKLVLLERLPRPS

SEQ ID NO: 164.

alfa-1,3-manosiltransferasa, ALG3 humana

MAAGLRKRGRSGSAAQAEGLCQWLQRAWQERRLLREPRYTLLVAACLCLAEVGITFWVIHRVAYTEIDWKAYM
AEVEGVINGTYDYTLQDGTGPLVPAGFVYIFMGLYYATSRGTDIRMAQNIFAVLYLATLLLVFLIYHQTCCKVP
PFVFFFMCCASYRVHSIFVLRFLFNDPVMVLLFLSINLLLAQRWGWGCCFFSLAVSVKMNVLFLAPGLLFLLLLTQ
FGFRGALPKLGICAGLQVVLGLPFLLENPSGYLSRSFDLGRQFLFHWTNVNWRFLPEALFLHRAFLHALLTAHLTL
LLLALCRWHRTGESILSLLRDPSKRKVPPQPLTPNQIVSTLFTSNFIGICFSRSLHYQFYVWYFHTLPYLLWAM
PARWLTHLLRLLVLGLIELSWNTYPSTSCSSAALHICHAVILLQLWLGPQFPKSTQHSKKAH

15

SEQ ID NO: 165.

beta-1,4-manosiltransferasa, HMT1/ALG1 humana

MAASCLVLLALCLLLPLLLGGWKRWRGRRAARHVAVVLGDVGRSPRMQYHALSLAMHGFSVTLLGFCNSKPHD
ELLQNNRIQIVGLTELQSLAVGPRVFQYGVKVVQLAMYLLWKLMWREPGAYIFLQNPPLPSIAVCWFVGCCLCGS

KLVIDWHNYGYSIMGLVHGPNHPLVLLAKWYEKFFGRLSHLNLCVTNAMREDLADNWHIRAVTVYDKPASFFKET
PLDLQHRLEFMKLGSMHSPFRARSEPEDPVTERSATERDAGSGLVTRLRERPALLVSSTSWTEDEDFSIILAALE
KFEQLTLDGHNPLSLVCVITGKGFLREYYSRLIHQKHQHIQVCTPWLEAEDYPLLLGSADLGVCCLHTSSSGLDL
PMKVDMFGCCLPVCVAVNFKCLHELKHEENGLVFEDSEELAAQLQMLFSNFPDPAGKLNQFRKNLRESQQLRWD
ESWVQTVLPLVMDT

5 SEQ ID NO: 166.

alfa-1,3-manosiltransferasa, ALG2 humana

MAEEQGRERDSVPKPSVLFHLPDLGVGGAERLVDAALALQARGCSVKIWTAHYDPGHCFAESRELTPVRCAGDWL
PRGLGWGGRGAAVCAYVMVFLALYVFLADEEFDVVCDQVSACIPVFLARRRKKILFYCHFDDLLLTKRDSF
LKRLYRAPIDWIEEYTTGMADCILVNSQFTAAVFKETFKSLSHIDPDVLYPSLNVTSDSVVPEKLDLVPKGGK
FLLLSINRYERKKNLTALAEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVEHYQELKKMVQQSDLGQYVTF
RSFSDKQKISLLHSCCTVLYTPSNEHFGIVPLEAMYMQCPVIAVNSGGPLESIDHSVTGFLCEPDVHFSEATEK
FIREPSLKATMGLAGRARVKEKFSPEAFTEQLYRYVTCLLV

SEQ ID NO: 167.

péptido-O-fucosiltransferasa 1, POFUT1 humana

MGAAAWARPLSVSFLLLLLPLPGMPAGSWDPAGYLLYCPCMGRFGNQADHFLGSLAFKLLNRTLAVPPWIEYQH
HKPPTNLHVSQYQYFKLEPLQAYHRVISLEDFMEKLAPTHWPPEKRVAYCFEVAQRSPDKKTCPMKEGNPFGP
FWDQFHVSVFNKSELFTGISFSASYREQWSQRFSPKEHPVLALPGAPAQFPVLEEHRPLQKYMVWSDVMVKTGEAQ
IHAIHLVRPYVGIHLRIGSDWKNACAMLDGTAGSHFMAAPQCVGYSRSTAAPLTMTMCLPDLKEIQRAVKLWVRS
LDAQSVYVATDSESYVPELQQLFKGKVKVSLKPEVAQVDLYILGQADHFIGNCVSSFTAFVKRERDLQGRPSSF
FGMDRPPKLRDEF

10

SEQ ID NO: 168.

péptido-O-fucosiltransferasa 1, pofut1 de vaca

MGAAAWAPSLPPRVSLLLLLPLPGLPVGSWDPAGYLLYCPCMGRFGNQADHFLGSLAFKLLNRTLAVPPWIE
YQHHKPPFTNVHVSQYQYFKLEPLQAYHRVISLEDFMEKLAPTHWPPEKRVAYCFEVAQRSPDKKTCPMKEGNP
FGPFWQFHVSVFNKSELFAGISFSASYKDQWIQRFSPEEHPVLALPGAPAQFPVLEEHRPLQKYMVWSDVMVRTG
EAQIHAIHLIRPYVGIHLRIGSDWKNACAMLDGTAGAHFMAAPQCVGYSRHTTAPLTMTMCLPDLKEIRRALKLW
VTALNAQSVYIATDSESYLPEIQQLFKGKVKVSLKPEVAQIDLYILGQADHFIGNCVSSFTAFVKRERDLHGRP
SSFFGMDRPPQLRDEF

SEQ ID NO: 169.

15 péptido-O-fucosiltransferasa 1, pofut1 de ratón

MGAAAWAPPHLLLRASFLLLLLLLPLRGRSAGSWDLAGYLLYCPCMGRFGNQADHFLGSLAFAKLLNRTLAVPPW
IEYQHKKPPFTNLHVSQYKFLEPLQAYHRVVSLEDFMENLAPSHWPPEKRVAYCFEVAQSPDKKTCPMKEG
NPFPGFWDQFHVSNKSELFTGISFSASYKEQWTQRFPAKEHPVLALPGAPAQFPVLEEHRPLQKYMVWSDVMR
TGEALISAHLVRPYVGIHLRIGSDWKNACAMLDGTAGSHFMA SPQCVGYSRSTATPLTMTMCLPDLKEIQRAVT
LWVRALNARSVVIATDSESYVSEIQQLFKDKVRVVS LKPEVAQIDLYILGQADHFIGNCVSSFTAFVKRERDLHG
RQSSFFGMDRPSQLRDEF

SEQ ID NO: 170.

péptido-O-fucosiltransferasa 1, POFUT1 de cerdo

SSPLPPLVSLLLLLLLPLPGLPAVSWDPAGYLLYCPCMGRFGNQADHFLGSLAFAKLLNRTLAVPPWIEYQHKKPP
FTNLHVSQYKFLEPLQVYHRVISLEDFMEKLAPTHWPPERRVAYCFEVAQSPDKKTCPMKEGNPFPGFWDQ
FHVSNKSELFTAGISFSASYKDQWIRFSPKEHPVLALPGAPAQFPVLEEHRPLQKYMVWSDVMVRTGEAQIRAH
LIRPYVGIHLRIGSDWKNACAMLDGTAGAHFMA SPQCVGYSRHTAAPLTMTMCLPDLKEIRRALKLWVTALSAQ
SVVIATDSESYLPEIQQLFKGKVKVVS LKPEVAQIDLYILGQADHFIGNCVSSFTAFVKRERDLQGRSSFFGMD
RPP

5 SEQ ID NO: 171.

beta-1,3-N-acetilglucosaminiltransferasa humana, maniac fringe

MQCRLPRGLAGALLTLLCMGLLCLRYHLNLSQPRVQGTPELSQPNPGPPKLQLHDVFI AVKTTRAFHRLRLLELLL
DTWVSRTREQTFVFTDSPDKGLQERLGSHLVVTNCSAEHSHPALSCMAAEFDTF LASGLRWFCHVDDDNVYNPR
ALLQLLRAFPLARDVYVGRPSLNRPIHASEPQPHNRTRLVQFWFATGGAGFCINRKLALKMAPWASGSRFMDTSA
LIRLPDDCTMGYIIECKLGGRLQPSPLFHSLETQLLRLTAQLPEQVTL SYGVFEGKLNVIKLQGPFSPEEDPSR
FRSLHCLLYPDTPWCPQLGAR

SEQ ID NO: 172.

Beta-1,2-N-acetilglucosaminiltransferasa, MGNT1 humana

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQALRRFCQTGAVLFLLVTIVIVNIKLILDTRRAISEANEDPEPE
QDYDEALGRLEPPRRRSGSGPRRVL DVEVYSSRSKVYVAVDGTTVLEDEAREQGRGIHVIVLNQATGHVMAKR VFD
TYSPEHEAMVLFNLNVAPGRVLICTVKDEGSFHLKDTAKALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKH
SKSPALSSWGD PVLLKTDVPLSSAEAECHWADTELNRRRRRFC SKVEGYGVCCKDPTPIEFSPDPLDNKVL
NVPVAVIAGNRPNLYRMLRSLLSAQGVSPQMITVTFIDGYEPM DVVALFGLRGIQHTPISIKNARVSQHYKAS
LTATFNLFPFAKFAVLEEDLDIAVDFFSFLSQS IHLLEEDDSLYCISAWNDQGYEHTAEDPALLYRVETMPGLG
WVLRRLSYKEELEPKWPTPEKLWDWDMWMRMPEQRRGRECIIPDVSRSYHFGIVGLNMNGYFHEAYFKKHKFNTV
PGVQLRNVD SLKKEAYEVEVHRLLSAEVLDH SKNPCEDSFLPDTEGHTYVAFIRMEKODDFTTWTQLAKCLHIW
DL DVRGNHRGLWRLFRKKNHFLVVGVPASPSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

10

SEQ ID NO: 173.

O-unido N-acetilglucosamina transferasa, OGT1 humana

MASSVGNVADSTEPTKRMLSFQGLAELAHREYQAGDFEAAERHCMQLWRQEPDNTGVLLLLSSIHFCRRRLDRSA
 HFSTLAIKQNPLLAAYSNLGNVYKERGQLQEAIEHYRHALRLKPDFIDGYINLAAALVAAGDMEGAVQAYVSAL
 QYNPDLYCVRSDLGNNLLKALGRLEEAKACYLKAIETQPNFAVAWSNLGCVFNAQGEIWLAIHHFEKAVTLDPNFL
 DAYINLGNVLKEARIFDRAVAAYLRALSLSPNHAVVHGNLACVYVEQGLIDLDAIDTYRRAIELQPHFPDAYCNLA
 NALKEKGSVAEAEEDCYNTALRLCPTHADSLNNLANIKREQNIEEAVRLYRKALEVFPEFAAAHSNLASVLQQQG
 KLQEALMHYKEAIRISPTFADAYSNMGNLTKEMQDVQGLQCYTRAIQINPAFADAHSNLASIHKDSGNIPEAIA
 SYRTALKLKPDFPDAYCNLAHCLQIVCDWTDYDERMCKLVSIVADQLEKNRLPSVHPHSMPLYPLSHGFRKAIAE
 RHGNLCCLKINVLHKPPYEHKDLKLSGRLRVGYVSSDFGNHPTSHLMQSIIPGMHNPKDFEVFCYALSPDDGTN
 FRVKVMAEANHFIIDLSQIPCNKGAADRIHQDGIHILVNMNGYTKGARNELFALRPAPIQAMWLGYPGTSGALFMD

YIITDQETSPAEEVAEQYSEKLAYMPHTFFIGDHANMFPHLKKKAVIDFKSNGHIYDNRIVLNGIDLKAFDLSLPD
 VKIVKMKCPDGGDNADSSNTALNMPVIMNTIAEAVIEMINRGQIQITINGFSISNGLATTQINNKAATGEEVPR
 TIIIVTTRSQYGLPEDAIVYCNFNQLYKIDPSTLQMWANILKRVPNVSVLWLLRFPVAVGEPNIQQYAQNMGLPQNR
 IFSPVAPKEEHVRRGQLADVCLDTPLCNHGTGMDVLWAGTPMVTMPGETLASRVAASQLTCLGCELELIAKNRQE
 YEDIAVKLGTDLEYLKKVRGKVKQRISSPLFNTKQYTMELERLYLQWHEHYAAGNKPDMIKPVEVTESA

SEQ ID NO: 174.

O-unido N-acetilglucosamina transferasa, Ogt de ratón

MASSVGNVADSTEPTKRMLSFQGLAELAHREYQAGDFEAAERHCMQLWRQEPDNTGVLLLLSSIHFCRRRLDRSA
 HFSTLAIKQNPLLAAYSNLGNVYKERGQLQEAIEHYRHALRLKPDFIDGYINLAAALVAAGDMEGAVQAYVSAL
 QYNPDLYCVRSDLGNNLLKALGRLEEAKACYLKAIETQPNFAVAWSNLGCVFNAQGEIWLAIHHFEKAVTLDPNFL
 DAYINLGNVLKEARIFDRAVAAYLRALSLSPNHAVVHGNLACVYVEQGLIDLDAIDTYRRAIELQPHFPDAYCNLA
 NALKEKGSVAEAEEDCYNTALRLCPTHADSLNNLANIKREQNIEEAVRLYRKALEVFPEFAAAHSNLASVLQQQG
 KLQEALMHYKEAIRISPTFADAYSNMGNLTKEMQDVQGLQCYTRAIQINPAFADAHSNLASIHKDSGNIPEAIA
 SYRTALKLKPDFPDAYCNLAHCLQIVCDWTDYDERMCKLVSIVAEQLEKNRLPSVHPHSMPLYPLSHGFRKAIAE
 RHGNLCCLKINVLHKPPYEHKDLKLSGRLRVGYVSSDFGNHPTSHLMQSIIPGMHNPKDFEVFCYALSPDDGTN
 FRVKVMAEANHFIIDLSQIPCNKGAADRIHQDGIHILVNMNGYTKGARNELFALRPAPIQAMWLGYPGTSGALFMD
 YIITDQETSPAEEVAEQYSEKLAYMPHTFFIGDHANMFPHLKKKAVIDFKSNGHIYDNRIVLNGIDLKAFDLSLPD
 VKIVKMKCPDGGDNPDSSNTALNMPVIMNTIAEAVIEMINRGQIQITINGFSISNGLATTQINNKAATGEEVPR
 TIIIVTTRSQYGLPEDAIVYCNFNQLYKIDPSTLQMWANILKRVPNVSVLWLLRFPVAVGEPNIQQYAQNMGLPQNR
 IFSPVAPKEEHVRRGQLADVCLDTPLCNHGTGMDVLWAGTPMVTMPGETLASRVAASQLTCLGCELELIAKSRQE
 YEDIAVKLGTDLEYLKKIRGKVKQRISSPLFNTKQYTMELERLYLQWHEHYAAGNKPDMIKPVEVTESA

SEQ ID NO: 175.

O-unido N-acetilglucosamina transferasa, Ogt de rata

MASSVGNVADSTGLAELAHREYQAGDFEAAERHCMQLWRQEPDNTGVLLLLSS IHFQCRRLDRSAHFSTLAIKQN
 PLLAEAYSNLGNVYKERGQLQEAI EHYRHALRLKPDFIDGYINLAAALVAAGDMGAVQAYVSALQYNPDLYCVR
 SDLGNNLLKALGRLEEAKACYLKAIETQPNFAVAWSNLGCVFNAQGEIWLAIHHFEKAVTLDPNFLDAYINLGNVL
 KEARIFDRAVAAYLRALSLSPNHAVVHGNLACVYYEQGLIDLAI DTYRRAIELQPHFPDAYCNLANALKEKGSVA
 EAEDCYNTALRLCPTHADSLNNLANIKREQGNIEEAVRLYRKALEVFPEFAAAHSNLASVLQQQKGLQEALMHYK
 EAIRISPTFADAYSNMGNLTKEMQDVQGALQCYTRAIQINPAFADAHSNLAS I HKDSGNIPEAIAS YRTALKLKP
 DFPDAYCNLAHCLQIVCDWTDYDERMKKLVSIVAEQLEKNRLPSVHPHSMPLYPLSHGFRKAI AERHGNLCLDKI
 NVLHKPPYEHKDLKLS DGRLRVG YVSSDFGNHPTSHLMQSIPGMHNPDKFEVFCYALSPDDGTNFRVKVMAEAN
 HFIDLSQIPCNKGAADRIHQDGIHILVMNGYTKGARNELFALRPAPIQAMWLGYPGTSGALFMDYIITDQETSP
 AEVAEQYSEKLAYMPHTFFIGDHANMFPHLKKKAVIDFSNGHIYDNRIVLNGIDLKAFDLSLDPVKIVKMKCPD
 GGDNADTTNTALNMPVIMPNTIAEAVIEMINRGQIQITINGFSISNGLATTQINNKAATGEEVPRTIIVTTRSQY
 GLPEDAIVYC NFNQLYKIDPSTLQMGANILKRV PNSVLWLLRFPVAVGEPNIQQYAQNMGLPQNR IIFSPVAPKEE
 HVRRGQLADVCLDTPLCNGHTTGMDVLWAGTPMVTMPGETLASRVAASQLTCLGCL ELIAKSRQEYEDIAVKLGT
 DLEYLKKIRGKVWKQRIS SPLFNTKQYTMELERLYLQMW EHYAAGNKPDHMIKPV EVTESA

SEQ ID NO: 176.

Chaperona molecular tipo HSP-70 de *E.coli* dnaK, inducible por calor

MGKIIGIDLGTNSCVAIMDGTTPRVLENAEGDRTPSIIAYTQDGETLVGQPAKRQAVTNPQNTLFAIKRLIGR
 RFQDEEVQRDVSIMPFKIIAADNGDAWVEVKGQKMAPQISA EVLKMKKTAEDYLGE PVTEAVITVPAYFNDAQ
 RQATKDAGRIAGLEVKRIINEPTAAALAYGLDKGTGNRTIAVYDLGGGTFDISIIEIDEVDGEKTFEVLATNGDT
 HLGGEDFDSRLINYLVEEFKDKQIDLRNDPLAMQRLKEAAEKAKIELSSAQQTVDVNLPHYITADATGPKHMNIKV
 TRAKLES LVEDLVNRSIEPLKVALQDAGLSVSDIDDVILVGGQTRMPMVQKKVAEFFGKEPRKDVNPDEAVAIGA
 AVQGGVLTGDVKDVL LLDVTPLSLGIETMGGVMTTLIAKNTTIPTKHSQVFSTAEDNQSAVTIHLVQGERKRAAD
 NKS LGQFNLDGINPAPRGMPQIEVTFDIDADGILHVS AKDKNSGKEQKITIKASSGLNEDEIQKMVRDAEANA EA
 DRKFEELVQTRNQGDHLLHSTRKQVEEAGDKLPADDKTAIESALTALETALKGEDKAAIEAKMQELAQVSQKLME
 IAQQQHAQQQTAGADASANNAKDDDVVDAEFEEVKDKK

5 SEQ ID NO: 177.

Proteasa de *E.coli* EG10157 clpB ClpB, dependiente de ATP

MRLDRLTNKFQALALADAQSLALGHDNQFIEPLHLMSALLNQEGGSVSPLLTSAGINAGQLRTDINQALNRLPQVE
 GTGGDVQPSQDLVRVLNLCDKLAQKRGDNFISSELFVLAAL ESRGTLADILKAAGATTANITQAI EQMRGGESVN
 DQGAEDQRQALKKYTIDLTERAEQGLDPVIGRDEEIRRTIQVLQRRTKNNPVLIGEPGVGKTAIVEGLAQRI IN
 GEVPEGLKGRRLV LALDMGALVAGAKYRGEFEERLKGVLNDLAKQEGNVILFIDELHTMVGAGKADGAMDAGNMLK
 PALARGELHCVGATTLD EYRQYIEKDAALERRFQKV FVAEPSVEDTIAILRGLKERYELHHHVQITDPAIVAAAT
 LSHRYIADRQLPDKAIDLIDEAASSIRMQIDSKPEELDRLDRRIIQLKLEQQALMKESDEASKKRLDMLNEELSD
 KERQYSELEEEWKAEKASLSGTQTIKAELEQAKIAIEQARRVGD LARMSELQYKGKIP ELEGKLEAATQLEGKTMR
 LLRNKVTDAEIAEVLARWTGIPVSRMMESEREKLLRMEQELHHRVIGQNEAVDAVSNAIRRSRAGLADPNRPIGS
 FLFLGPTGVGKTELCKALANFMFDSDEAMVRIDMSEFM EKHSVSRLVGAPPGYVG YEEGGYLTEAVRRRPYSVIL
 LDEVEKAHPDVFNILLQVLDDGRLTDGQGR TVDFRNTVVIMTSNLGSDLIQERFGELDYAHMKELVLGVVSHNFR
 PEFINRIDEVVVFHPLGEQHIA SIAQIQLKRLYKRL EERGYEIHISDEALKLLSENGYDPVYGARPLKRAIQQQI
 ENPLAQQILSGELVPGKVRIRLEV NEDRIVAVQ

SEQ ID NO: 178.

GroEL, *E.coli*

AAKDVKFGNDARVKMLRGVNVVLADAVKVTGLGPKGRNVVLDKSFSGAPTITKDGVSVAAREIELEDKFENMGAQMVKE
VASKANDAAGDGTATVLAQAIITEGLKAVAAGMNPMDLKRIGDKAVTAAVEELKALSVPCSDSKAIAQVGTIS
ANSDETVGKLI AEAMDKVGKEGVITVEDGTGLQDELDVVEGMQFDRGYLSPYFINKPETGAVELES PFILLADKK
ISNIREMLPVLEAVAKAGKPLLI I AEDVEGEALATLVVNTMRGIVKVA AVKAPGFGDRRKAMLQDIATLTGGTVI
SEEIGMELEKATLEDLGQAKRVVINKDTTTIIDGVGEEAAIQGRVAQIRQQIEEATSDYDREKLQERVAKLAGGV
AVIKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVALIRVASKLADLRGQNEQNVGKVALRAMEAPL
RQIVLNCGEEPSVVANTVKGGDGNYGYNAAATEEYGNMIDMGILDPTKVTR SALQYAASVAGL MITTECMVTDL PK
NDAADLGAAGGMGGM

SEQ ID NO: 179.

Cpn10, *Oleispira antarctica*

MKIRPLHDRVVRKEEETATAGGIILPGAAAEPNQGVVISVGTGRILDNGSVQALAVNEG DVVVF GKYSQONT
IDIDGEELLIILNESDIYGVLEA

5 SEQ ID NO: 180.

Cpn 60, *Oleispira antarctica*

MAAKDVLFGDSARAKMLVGVN ILADAVRVTGLGPKGRNVVIEKSFSGAPIITKDGVSVAAREIELKDKFENMGAQMVK
EVASQANDQAGDGTATVLAQAI I SEGLKSVAAGMNPMDLKRIGDKATAAVVA AIKEQAQPCLDTKAIAQVGTI
SANADETVGR LIAEAMEKVGKEGVITVEEGKGLEDELDVVEGMQFDRGYLSPYFINNQEKMTVEMENPLIILLVDK
KIDNLQELLPILENVAKSGRPLLIVAEDVEGQALATLVVNNLRGTFKVA AVKAPGFGDRRKAMLQDLAILTGGQV
ISEELGMSLETADPSSLGTSKVV IDKENTVI VDGAGTEASVNTVRDQIRAEIESSTSDYDIEKLQERVAKLAGG
VAVIKVGAGSEMEMKEKKDRVDDALHATRAAVEEGVVAGGGVALIRALSSVTTVGDNEDQNVGIALALRAMEAPI
RQIAGNAGAE GSVVVDKVKSGTGSFGFNASTGEYGDMIAMGILDPAKVTRSSLQAAAS IAGL MITTEAMVADAPV
EEGAGGMPDMGGMGGMGMPGMM

SEQ ID NO: 181.

Proteína PDI, PDIA1_RAT de rata

MLSRALLCLALAWAARVGADALEEEDNVVLVKKSNFAEALAAHNYLLVEFYAPWCGHCKALAPEYAKAAAKLKAE
GSEIRLAKVDATEESDLAQQYGVRYPTIKFFKNGDTASPKYTAGREADDIVNWLKKRTGPAATTLSDTAAAES
LVDSSEVTVIGFFKDAGSDSAKQFLAAEA VDDIPFGITSNSDVFSKYQLDKDGVVLFKKFDEGRNNFEGETKE
KL LDFIKHNQLPLVIEFTEQTAPKIFGGEIKTHILLFLPKSVSDYDGKLSNFKKAABEGFKGKILFIFIDSHTDN
QRILEFFGLKKEBCPAVRLITLBEEMTKYKPESDELTAEKITQFCHHFLEGGIKPHLMSQELPEDWDKQPVKVLV
GKNFEEVAFDEKKNVFEFYAPWCGHCKQLAPIWDKLGETYKDHENIVIAKMDSTANEVEAVKVHSFPTLKFFPA
SADRTVIDYNGERTLDGFKKFLESGGQDGAGDNDLDLLEBALEPDMEEEDDDQKAVKDEL

SEQ ID NO: 182.

Ero1p, *Sacchromyces*

MRLRTAIATLCLTAFTSATSNN SYIATDQTQNAFNDTHFCVKDRNDHVS PSCNVTFNELNAINENIRDDL SALLK
SDFFKYFRLDLYKQCSFWDANDGLCLNRACSV DVVEDWDTLPEYWQPEILGSFNNDTMKEADDSDECKFLDQLC
QTSK KPVDIEDTINYCDVND FNGKNAVLIDL TANPERFTGYGGKQAGQI WSTIYQDNCFTIGETGESLAKDAFYR
LVSGFHASIGTHLSKEYLNTKTGKWEPNLDLFMARIGNFPDRVTNMYFNYAVVAKALWKIQPYLPEFSFCDLVNK
EIKNKMDNVISQLDTKIFNEDLVFANDLSLTLKDEFRSRFKNVTKIMDCVQCDCRLWGKIQTGTGYATALKILFE
INDADEFTKQHIVGKLT KYELIALLQT FGRLSESIESVNMFEKMYGKRLNGSENRLSSFFQNNFFN ILKEAGKS I
RYTIENINSTKEGKKKTNNSQSHVFDLKM PKAEI VPRPSNGTVNKWKKAWNTEVNNVLEAFRFIYRSYLDLPRN
IWELSLMKVYKFWNKFIGVADYVSEETREPISYKLDIQ

SEQ ID NO: 183.

Pdi1p, *Saccharomyces*, YIL005W

MKMNLKRLVVTFFSCITFLLKFTIAAAEPPEGGFPEPLNPTNFKEELSKGLHIIDFYSPYCPHCKHLAPVWMETWE
EFKEESKTLNITFSQVNCIESADLCGDENIEYFPEIRLYNPSGYIKSFTETPRTKESLIAFARRESMDPNMLDTD

LDSAKSESQYLEGFDFLELIAGKATRPHLVSWPTKDMKNSDDSLFKNCDKCHEFQRTWKIISRQLAVDDINTG
HVNCESNPTICEELGFGDLVKITNHRADREP KVALVLPNKTSNNLFDYPNGYSAKSDGYVDFARRTFTNSKFPNI
TEGELEKKANRDIDFLQERGRVTNNDIHLVFSYDPETVVIDFDILEYLIIEPLSKIPNIYLHQIDKNLINLSRNL
FGRMYEKINYDASQTQKVFNKEYFTMNTVTQLPTFFMFKDGDPISYVFPGYSTTEMRNIDAIMDWVKYSNPLVT
EVDSSNLKKLISFQTKSYSDLAIQILISSTDHKKHIKGSNKLIKNNLLASWEYEHIRMENNFEEINERRARKADGIK
KIKEKKAPANKITVDKMREEIPMDQKKLLGYLDISKEKNFRKYGITGEYKIGDVIIIDKSNNYYNKNDFGNS
LTSNPPQLLREAFVSLNIPSKALYSSKLKGRLLNSPFHNVL SFLDI IHGNGMPGYLIVIVLFIAILKGPSIYRRY
KVRKHYRAKRNAVGILGNMEKKKNQD

5 SEQ ID NO: 184.

Proteína DsbB, *E.coli*

MLRFLNQCSQGRGAWLLMAFTALALELTALWFQHVMLLKPCVLCIYERCALFGVLGAALIGAIAPKTPLRYVAMV
IWLYSAFRGVQLTYEHTMLQLYPSPFATCDFMVRFPPEWLPLDKWVPQVFVSGDCAERQWDFLGLEMPQWLLGIF
IAYLIVAVLVVISQPFKAKKRDLFGR

SEQ ID NO: 185.

Proteína DsbA, *E.coli*

MKKIWLALAGLVLAFSASAAQYEDGKQYTTLEKPVAGAPQVLEFFSFFCPHCYQFEEVLHISDNVKKKLPEGVKM
TKYHVNFMGGDLGKDLTQAWAVAMALGVEDKVTVPLEFQVQKTQTI RSASDIRDVF INAGIKGEEYDAWNSFVV
KSLVAQQEKAADVQLRGVPAMFVNGKYQLNPQGMDSNMDDVFVQYQYADTVKYLSEKK

10

SEQ ID NO: 186.

Proteína DsbC, *E.coli*

MKKGFMLFTLLAAFSGFAQADDAAIQQTAKMGIKSSDIQAPVAGMKTVLTNSGVLYITDDGKHIIQGPMYDVS
GTAPVNVNTKMLLKQLNALEKEMIVYKAPQEKHVITVFTDITCGYCHKLHEQMADYNALGITVRYLAFPRQGLDS
DAEKEMKAIWCAKDKNAFDDVMAGKSVAPASCDVDIADHYALGVQLGVSGTPAVVLSNGTLVPGYQPPKEMKEF
LDEHQMTSGK

SEQ ID NO: 187.

15 Proteína DsbD, *E.coli*

MAQRIFTLILLCLSTSVFAGLFDAPGRSQFVPADQAFAPDFQONQHDLNLTWQIKDGYLYRKQIRITPEHAKIA
DVQLPQGVWHEDEFYKGSEIYRDRLTLPVTINQASAGATLTVTYQGCADAGFCYPPETKTVPLEVVANNAAPQP
VSPVQQEQPTAQLPFSALWALLIGIGIAFTPCVLPMPYPLISGIVLGGKQRLSTARALLLTFIYVQGMALTYTALG
LVVAAAGLQFQAALQHPYVLIGLAIVFTLLAMSMFGLFTLQLPSSLQTRLTLMNSRQQGSPGGVFVMGAIAGLI
CSPCTTAPLSAILLYIAQSGNMWLGSGTLYLYALGMGLPLMLITVFGNRLLPKSGPWMEQVKTAFGFVILALPVF
LLERVIGDVWGLRLWSALGVAFWGAFITSLQAKRGWMRIVQIILLAAALVSVRPLQDWAFGATHTAQTQTHLNF
TQIKTVDELNQALVEAKGKPVMLDLYADWCVACKEFEKYTFSDPQVQKALADTVLLQANVTANDAQDVALLKHLN
VLGLPTILFFDQGQGEHPQARVTGFMDAETFS AHLRDRQP

SEQ ID NO: 188.

Proteína DsbG, *E.coli*

MTVIGYAFYSTFALTEKDKMLKILLALLPAIAFAEELPAPVKAIEKQGITI IKTFDAPGGMKGYLKGYQDMG
VTIYLTDPDGKHAISGYMYNEKGENLSNTLIEKEIYAPAGREMWQRMEQSHWLLDGKKDAPVIVYVFADPFCPYCK
QFWQQARPWVDSGKVLRTLLVGVIKPESPATAAAILASKDPAKTWQQYEASGGKLLNVPANVSTEQMKVLSDN
EKLMDLGGANVTTPAIYYMSKENTLQQAVGLPDQKTLNIIMGNK

SEQ ID NO: 189.

5 UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1 Δ 40) humana truncada

GLPAGDVLEPVQKPHEGPGEMGKPVVIPKEDQEKMKEMFKINQFNLMASEMIALNRSPLDVRLEGCKTKVYPDNL
PTTSVVIVFHNEAWSTLLRTVHNSVINRSPRHMIEEIVLDDASERDFLKRPLESYVKKLKVPVHVIRMEQRSGLI
RARLKGAAVSKGQVITFLDAHCECTVGWLEPLLARIKHDRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKLNF
RWYPVPQREMDRRKGDRTLPVRTPTMAGGLFSIDRDYFQEIGTYDAGMDIWGGENLEISFRIWQCGGTLEIVTCS
HVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMDDEFKNFFYIISPVGTVKVDYGDISSRVGLRHKLQCKPFSWYL
ENIYPDSQIPRHYFSLGEIRNVETNQCLDNMARKENKVGIFNCHGMGGNQVFSYTANKEIRTDLCLDVS KLNG
PVTMLKCHHLKGNQLWEYDPVKLTLQHVNSNQCLDKATEEDSQVPSIRDCNGSRSQQWLLRNVTLP EIF

SEQ ID NO: 190.

Forma alterna de UDP-N-acetilgalactosaminiltransferasa 1 (GalNAc-T1 Δ 40) humana truncada

MGLPAGDVLEPVQKPHEGPGEMGKPVVIPKEDQEKMKEMFKINQFNLMASEMIALNRSPLDVRLEGCKTKVYPDNL
LPTTSVVIVFHNEAWSTLLRTVHNSVINRSPRHMIEEIVLDDASERDFLKRPLESYVKKLKVPVHVIRMEQRSGL
IRARLKGAAVSKGQVITFLDAHCECTVGWLEPLLARIKHDRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKLNF
FRWYPVPQREMDRRKGDRTLPVRTPTMAGGLFSIDRDYFQEIGTYDAGMDIWGGENLEISFRIWQCGGTLEIVTCS
SHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMDDEFKNFFYIISPVGTVKVDYGDISSRVGLRHKLQCKPFSWY
LENIYPDSQIPRHYFSLGEIRNVETNQCLDNMARKENKVGIFNCHGMGGNQVFSYTANKEIRTDLCLDVS KLNF
GPVTMLKCHHLKGNQLWEYDPVKLTLQHVNSNQCLDKATEEDSQVPSIRDCNGSRSQQWLLRNVTLP EIF

10 SEQ ID NO: 191.

UDP-N-acetilgalactosaminiltransferasa 1 (MBP-GalNAc-T1 Δ 40) humana truncada marcada con proteína de unión a maltosa

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFWAHDFRFGGYAQSG
LLAEITPDKAFQDKLYPFTWDAVRVNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAFNKGGETAM
TINGPWAWSNIDTSKVNIGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSIEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSNNNN
NNNNNNLGIEGRISEFGSHMGLPAGDVLEPVQKPHEGPGEMGKPVVIPKEDQEKMKEMFKINQFNLMASEMIALN
RSLPDVRLEGCKTKVYPDNLPTTSVVIVFHNEAWSTLLRTVHNSVINRSPRHMIEEIVLDDASERDFLKRPLESY
VKKLKVPVHVIRMEQRSGLIRARLKGAAVSKGQVITFLDAHCECTVGWLEPLLARIKHDRRTVVCPIIDVISDDT
FEYMAGSDMTYGGFNWKLNF RWYPVPQREMDRRKGDRTLPVRTPTMAGGLFSIDRDYFQEIGTYDAGMDIWGGEN
LEISFRIWQCGGTLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMDDEFKNFFYIISPVGTVKVDYGD

ISSRVGLRHKLQCKPFSWYLENIYPDSQIPRHYFSLGEIRNVETNQCLDNMARKENKVGIFNCHGMGGNQVFSY
TANKEIRTDLCLDVS KLNGPVTMLKCHHLKGNQLWEYDPVKLTLQHVNSNQCLDKATEEDSQVPSIRDCNGSRS
QQWLLRNVTLP EIF

SEQ ID NO: 192.

UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ53) humana truncada

KDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQCQR
KQWRVDLPATSVVITFHNEARSALLRTVVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRREG
LMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGFDWNL
VFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEIIP
CSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVPYGNIQSRLELRKKLSCKPFW
YLENVYPELRVPDHDQDIAFGALQQGTNCLDTLGHFADGVVGVYECHNAGGNQEWALTKEKSVKHMCLCLTVVDRA
PGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCLDSRTAKSGGLSVEVCGPALSQQWKFTLNLQQ

SEQ ID NO: 193.

5 UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ1-53 Δ445-571) humana truncada

KDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQCQR
KQWRVDLPATSVVITFHNEARSALLRTVVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRREG
LMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGFDWNL
VFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEIIP
CSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVPYGNIQSRLELRKKLSCKPFW
YLENVYPELRVPDHDQD

SEQ ID NO: 194.

Forma alterna de UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ53) humana truncada

MSKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQC
QRKQWRVDLPATSVVITFHNEARSALLRTVVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRR
EGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGFDW
NLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEI
IPCSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVPYGNIQSRLELRKKLSCKPFW
KWYLENVYPELRVPDHDQDIAFGALQQGTNCLDTLGHFADGVVGVYECHNAGGNQEWALTKEKSVKHMCLCLTVVD
RAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCLDSRTAKSGGLSVEVCGPALSQQWKFTLNLQQ

10 SEQ ID NO: 195.

Forma alterna de UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ1-53 Δ445-571) humana truncada

MSKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVESDKLRMDRAIPDTRHDQC
QRKQWRVDLPATSVVITFHNEARSALLRTVVSVLKKSPPHLIKEIILVDDYSNDPEDGALLGKIEKVRVLRNDRR
EGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINMDNFQYVGASADLKGGFDW

NLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWGGENLEISFRVWQCGGSLEI
IPCSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVPYGNIQSRLELRKKLSCKPFW
KWYLENVYPELRVPDHDQD

SEQ ID NO: 196.

15 UDP-N-acetilgalactosaminiltransferasa 2 (MBP-GalNAc-T2 Δ1-53 Δ445-571) humana truncada marcada con proteína de unión a maltosa

MKIEEGKLVIIWINGDKGYNGLAIEVGGKFEKDTGIIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSG
 LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
 NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAM
 TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
 GAVALKSYEEELAKDPRIAATMENAQGEIMPNIPOMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNN
 NNNNNNLGIEGRISEFGSHMSKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQV
 ESDKLRMDRAIPDTRHDQCQRKQWRVDLPATSVVITFHNEARSALLRTTVSVLKKSPPHLIKEIILVDDYSNDPE
 DGALLGKIEKVRVLRNDRREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVIN
 MDNFQYVGASADLKGGFOWNLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWG
 GENLEISFRVWQCGGSLEIIPCSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVP
 YGNIQSRLELRKKLSCKPFKYLENVPELRVPDHQDIAFGALQQGTNCLDTLGHFADGVGVVYECHNAGGNQEW
 ALTKEKSVKHMCLCLTVVDRAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCCLDSRTAKSGGLSVEVCGP
 ALSQQWKFTLNLQQ

SEQ ID NO: 197.

UDP-N-acetilgalactosaminiltransferasa 2 (MBP-GalNAc-T2 Δ 53) humana truncada marcada con proteína de unión a maltosa

MKIEEGKLVIIWINGDKGYNGLAIEVGGKFEKDTGIIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSG
 LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
 NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAM
 TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
 GAVALKSYEEELAKDPRIAATMENAQGEIMPNIPOMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNN
 NNNNNNLGIEGRISEFGSHMSKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQV
 ESDKLRMDRAIPDTRHDQCQRKQWRVDLPATSVVITFHNEARSALLRTTVSVLKKSPPHLIKEIILVDDYSNDPE
 DGALLGKIEKVRVLRNDRREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVIN
 MDNFQYVGASADLKGGFOWNLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWG
 GENLEISFRVWQCGGSLEIIPCSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVP
 YGNIQSRLELRKKLSCKPFKYLENVPELRVPDHQDIAFGALQQGTNCLDTLGHFADGVGVVYECHNAGGNQEW
 ALTKEKSVKHMCLCLTVVDRAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCCLDSRTAKSGGLSVEVCGP
 ALSQQWKFTLNLQQ

5

SEQ ID NO: 198.

UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (Core-1-Gal-T1 Δ 31) de *Drosophila*

GFCLAEFLVYSTPERSEFMPYDGHHRHGDVNDAAHSHDMMMSGPEQDVGGHEHVHENSTIAERLYSEVRVLCWIM
 TNPSNHQKARHVKRTWKGKRCNKLIFMSSAKODELDAVALPVGEGRNNLWGKTKEAYKYIEHHINDADWFLKAD
 DDTYITIVENMRYMLYPSPETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFVVEALPNPKLCKSDNSGAEDV
 EIGKCLQNVNVLGDSRDSNGRGRFFPFVPEHHLIPSHTDKKFWYQYIFYKTDEGLDCCSDNAISFHYVSPNQM
 YVLDYLIYHLRPYGIINTPDALPNKLAVGELMPEIKEQATESTSDGVSKRSAETKTQ

SEQ ID NO: 199.

10 Mutante A383T de UDP-galactosa:N-acetilgalactosamina-alfa-R beta-1,3-galactosiltransferasa (MBP-Core-1-Gal-T1 Δ 31) de *Drosophila* truncada marcada con proteína de unión a maltosa

MKIEEGKLVIIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFWAHDFRGGYAQSG
LLAEITPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMF
NLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKKNKHMNADTDYSIAEAAFNKGETAM
TINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL
GAVALKSYEEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNN
NNNNNNLGIEGRISEFGSHMGFCLAE LFVYSTPERSEFMPYDGHHRGVDNDAHSHDMMEMSGPEQDVGGHEHVN
ENSTIAERLYSEVRVLCWIMTNPSNHQKKARHVKRTWKGRCNKLI FMSSAKDDELDAVALPVGEGRNNLWGKTKE
AYKYIYEHHINDADWFLKADDDTYTIVENMRYMLYPSPETPVYFGCKFKPYVKQGYMSGGAGYVLSREAVRRFV
VEALPNPKLCKSDNSGAEDVEIGKCLQNVNVLGDSRDSNGRGRFFFPFVPEHHLIPSHTDKKFWYQYIFYKTDE
GLDCCSDNAISFHYVSPNQMYVLDYLIYHLRPGYIINTPDALPNKLAVGELMPEIKEQATESTSDGVSKRSTETK
TQ

SEQ ID NO: 200.

UDP-N-acetilgalactosaminiltransferasa 2 (GalNAc-T2 Δ 1-51 Δ 445-571) humana truncada marcada con proteína de unión a maltosa

mkieegklviwingdkgynglaevgkkfekdtgikvtvehpdkleekfpqvaatgdgpdiiifwahdrfggyaqsg
llaeitpdkafqdklypftwdavryngkliaypiavealsliynkdllpnppktweeipaldkelkakgksalmf
nlqepyftwpliaadggayafkyengkydikdvgvndagakagltflvdlikknkhnadtdysiaeaafnkgetam
tingpwawsnidtskvnygtvlptfkgqpskpfvgvlsaginaaspnkelakeflenylltdegleavnkdkpl
gavalksyeeelakdpriaatmenaqkgeimpnipoqmsafwyavrtavinaasgrqtvdealkdaqtnssssnnnn
nnnnnnlgiegrisefgskkkDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDPYARNKFNQVE
SDKLRMDRAIPDTRHDQCQRKQWRVDLPATSVVITFHNEARSALLRTVSVLKKSPPHLIKEIILVDDYSNDPED
GALLGKIEKVRVLRNDRREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAEDRTRVVSPIIDVINM
DNFQYVGASADLKGFDWNLVFKWDYMTPEQRRSRQGNPVAPIKTPMIAGGLFVMDKIFYFEELGKYDMMMDVWGG
ENLEISFRVWQCGGSLEIIPCSRVGHVFRKQHPYTFPGGSGTVFARNTRRAAEVWMDEYKNFYAAVPSARNVPY
GNIQSRLELRKKLSCKPFKWYLENVYPELRVPDHDQ

REIVINDICACIONES

1. Método de producción de una glicosiltransferasa eucariota activa soluble en un microorganismo procariota, en el que el microorganismo procariota tiene un entorno oxidante,
comprendiendo el método las etapas de
 - a) expresar un ácido nucleico que codifica para la glicosiltransferasa eucariota en el microorganismo procariota; y
 - b) hacer crecer el microorganismo en condiciones que permiten la expresión de la glicosiltransferasa eucariota activa soluble dentro de un compartimiento celular del microorganismo procariota,en el que el microorganismo procariota tiene actividad reductasa ausente o reducida que resulta de una mutación en un ácido nucleico de reductasa endógeno, y se hace crecer a una temperatura inferior a una temperatura de crecimiento óptima.
2. Método de la reivindicación 1, en el que la glicosiltransferasa eucariota es un miembro seleccionado del grupo que consiste en una N-acetilglucosaminiltransferasa I (GnT o GNT) eucariota, una N-acetilgalactosaminiltransferasa (GalNAcT) eucariota, una galactosiltransferasa (GalT) eucariota y una sialiltransferasa eucariota.
3. Método de la reivindicación 2, en el que la galactosiltransferasa (GalT) eucariota es una β -1,4-galactosiltransferasa (GalT1) eucariota o una core 1 galactosiltransferasa (core 1 GalT1) eucariota; la sialiltransferasa eucariota es una α (2,3)sialiltransferasa (ST3Gal3) eucariota, una α -N-acetilgalactosaminida α -2,6-sialiltransferasa I (ST β GalNAcT1) eucariota, o una Gal β 1,3GalNAc α 2,3-sialiltransferasa (ST3GalI) eucariota.
4. Método de la reivindicación 1, en el que el microorganismo procariota es una *E. coli*.
5. Método de la reivindicación 4, en el que la *E. coli* tiene una mutación en un gen *txrB* y un gen *gor* y se hace crecer a una temperatura entre 12-30°C.
6. Método de la reivindicación 1, que comprende adicionalmente la etapa de aislar la glicosiltransferasa eucariota activa, soluble.
7. Método de la reivindicación 1, en el que la glicosiltransferasa eucariota activa soluble se produce en una escala de microgramos, miligramos o gramos.
8. Método de la reivindicación 1, en el que la glicosiltransferasa eucariota activa soluble comprende una etiqueta de purificación.
9. Método de la reivindicación 1, en el que el microorganismo procariota comprende una proteína disulfuro isomerasa (PDI) heteróloga.
10. Método de la reivindicación 1, en el que el microorganismo procariota comprende una proteína chaperona heteróloga.

Figura 1

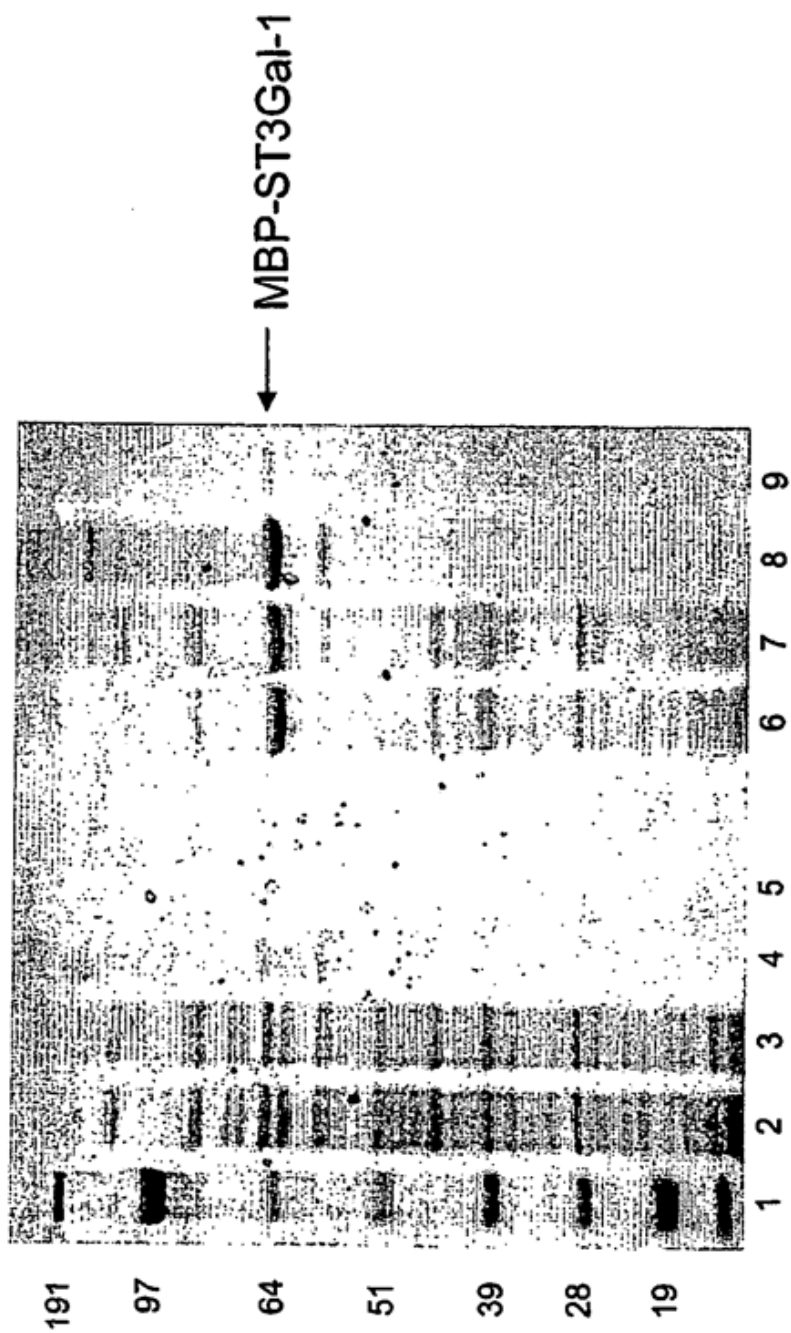


Figure 2a

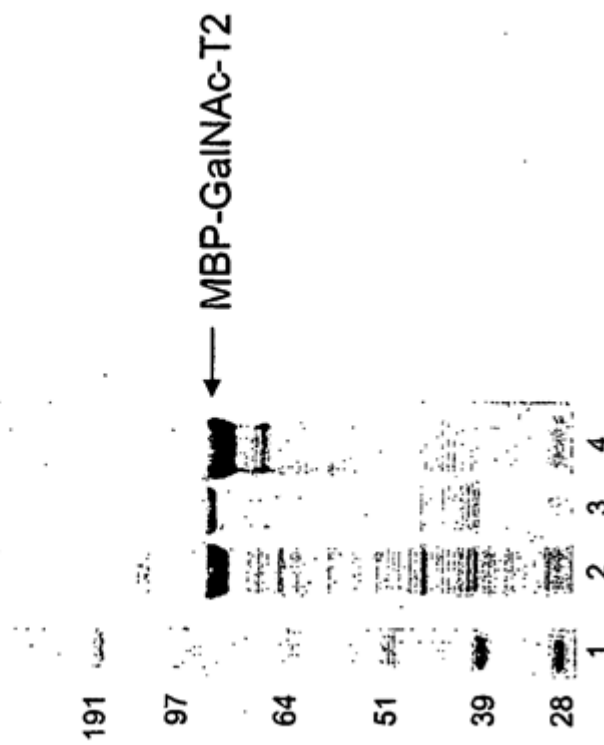


Figure 2b

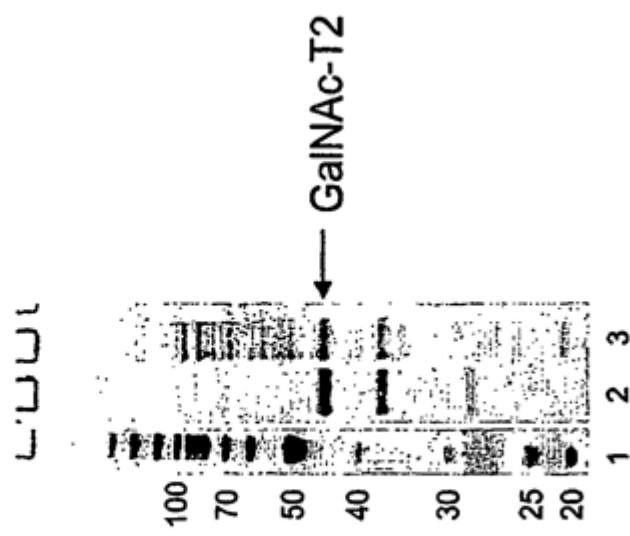


Figura 3

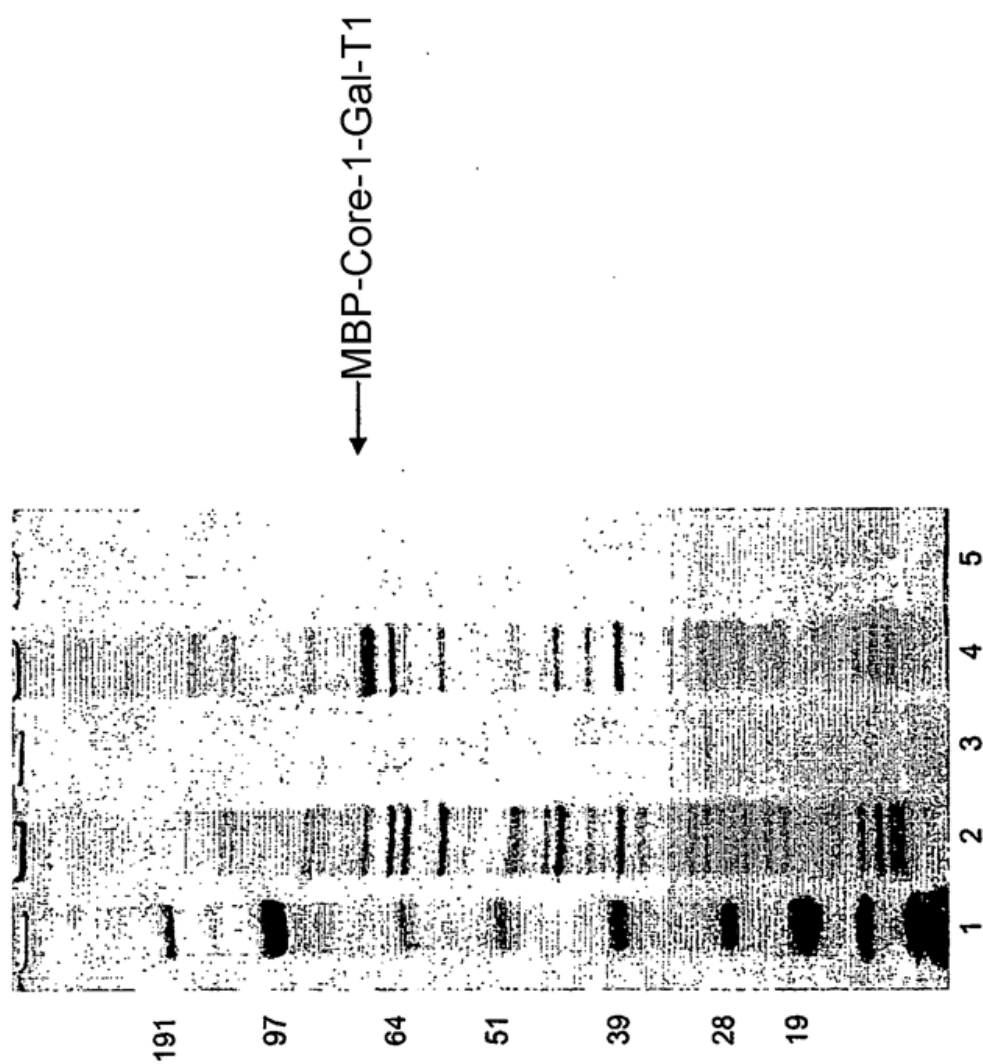


Figure 4

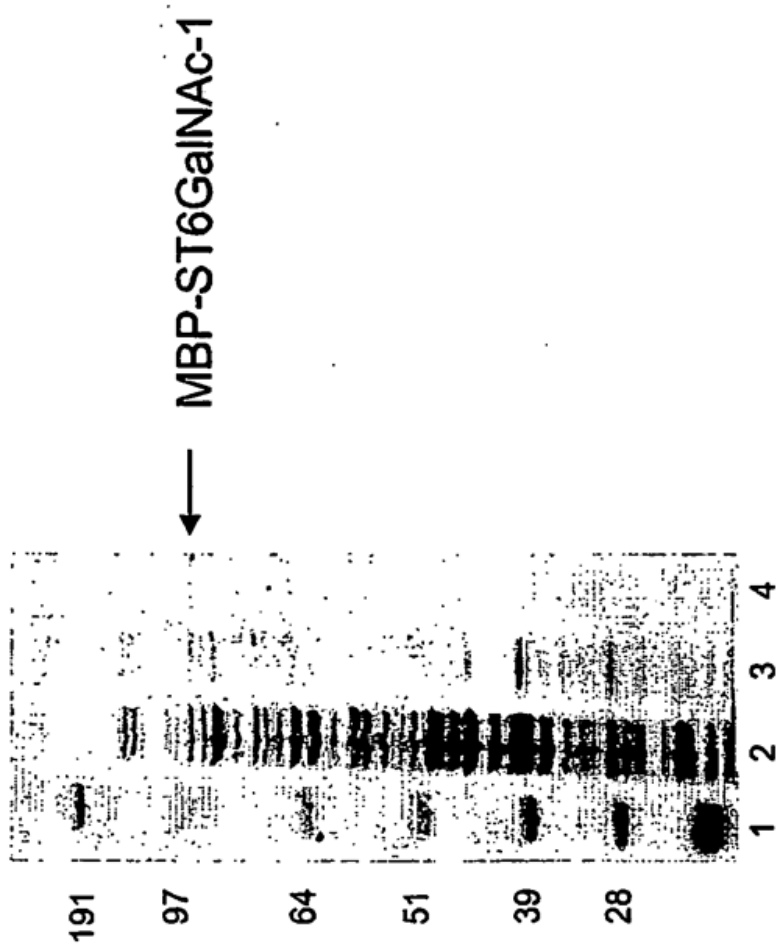


Figura 5

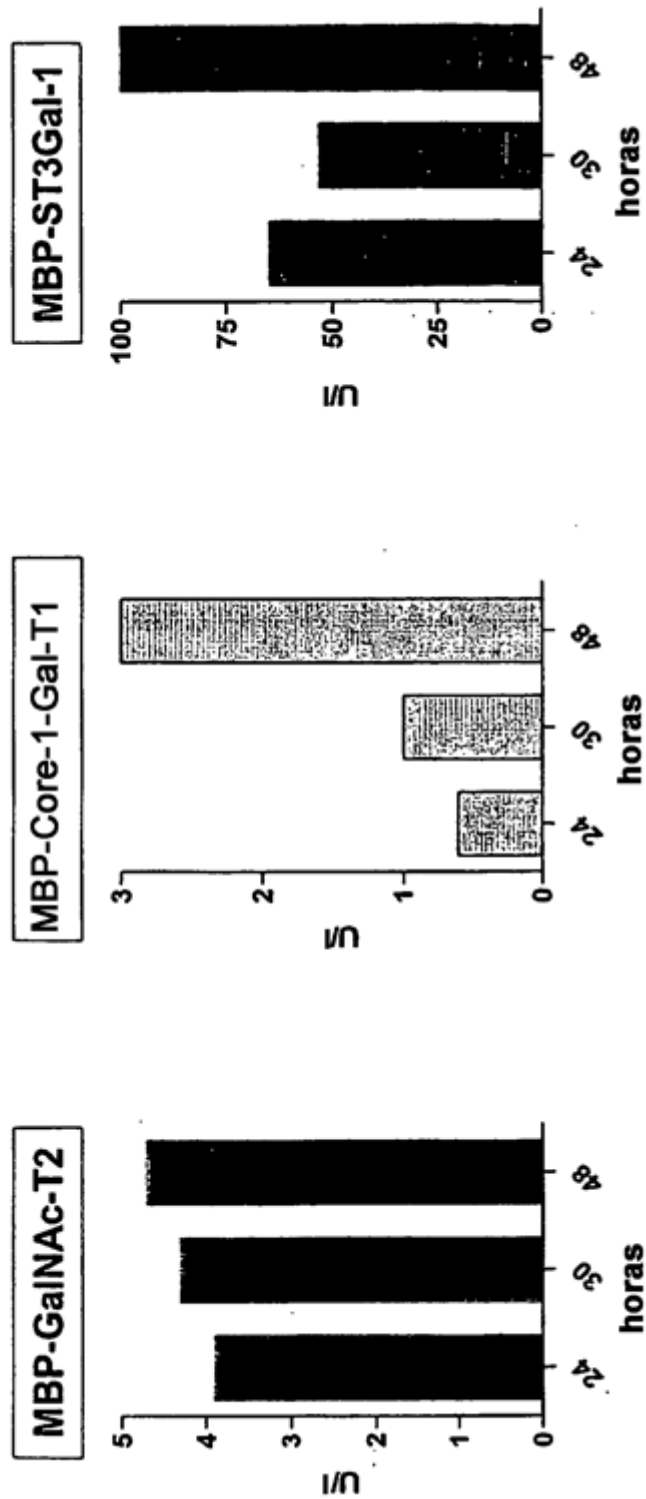


Figura 6

MALDI-TOF MS; Interferón alfa-2b

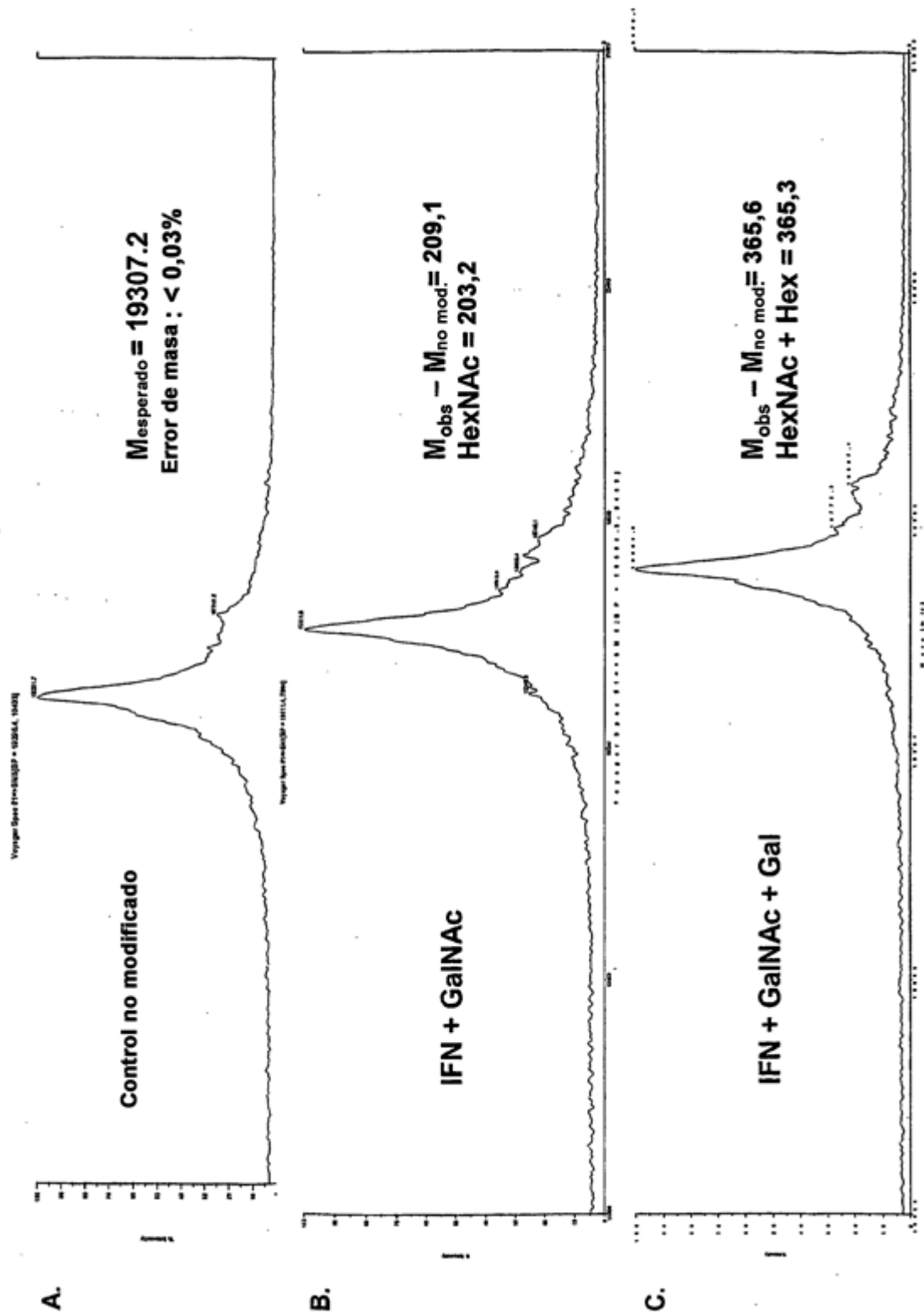


Figura 7

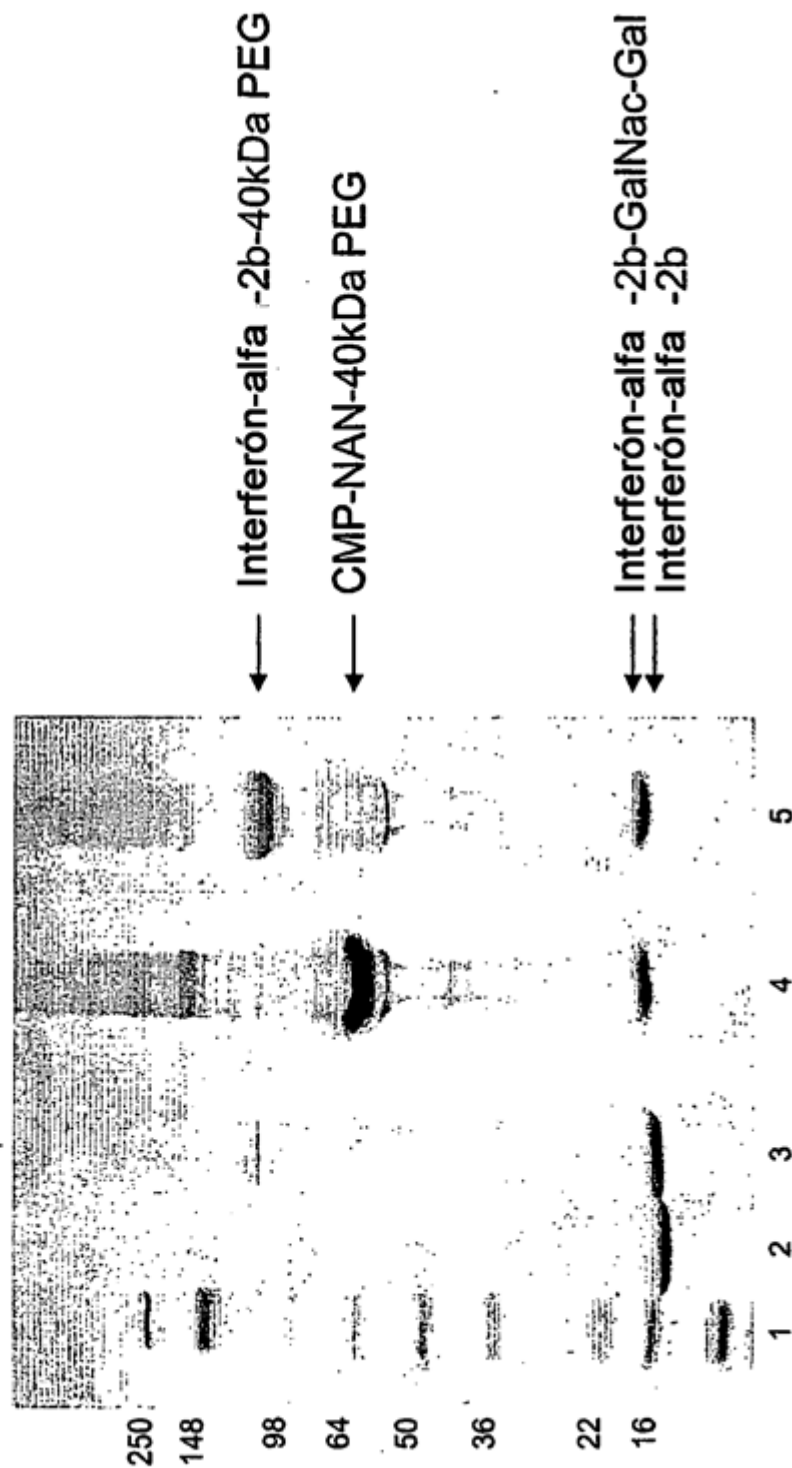


Figura 8

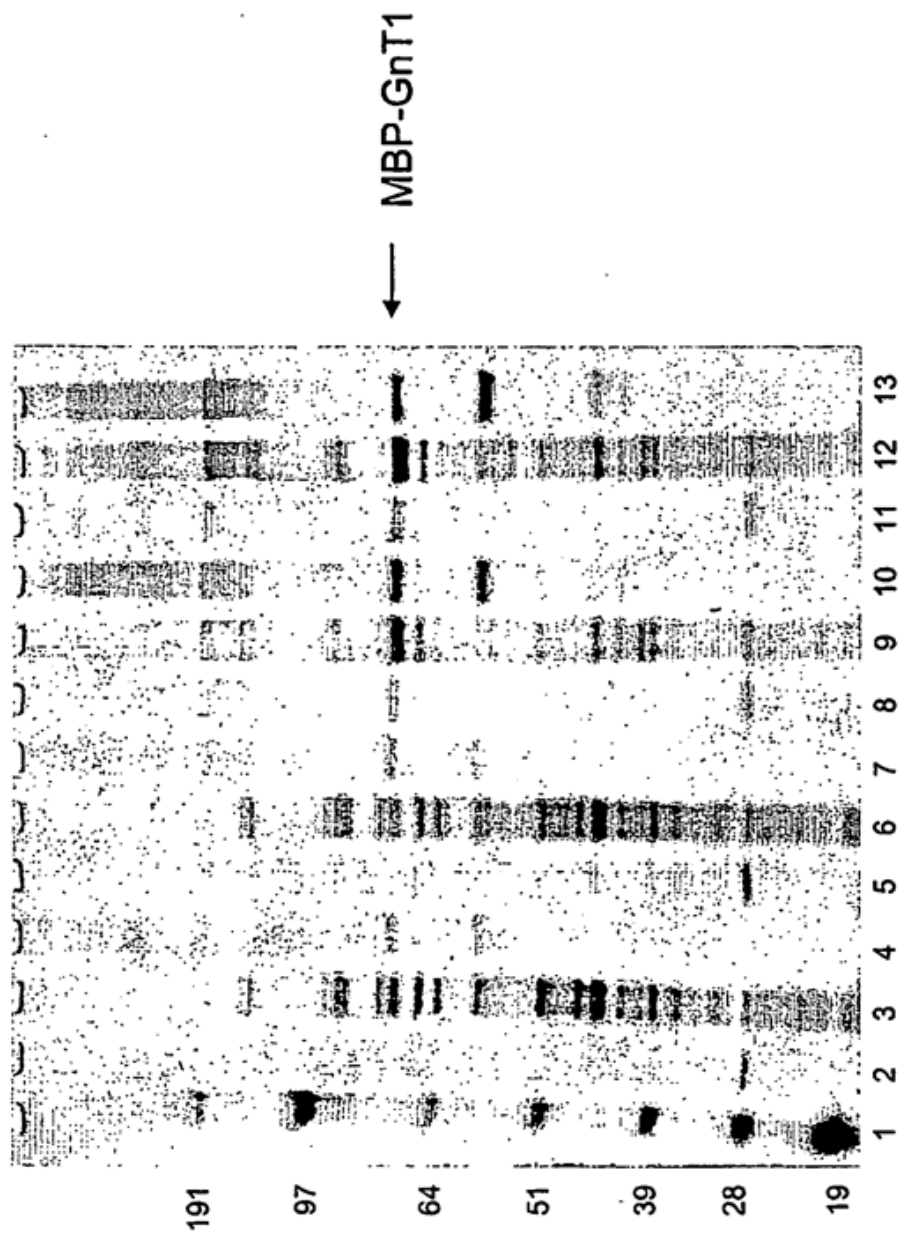


Figura 9

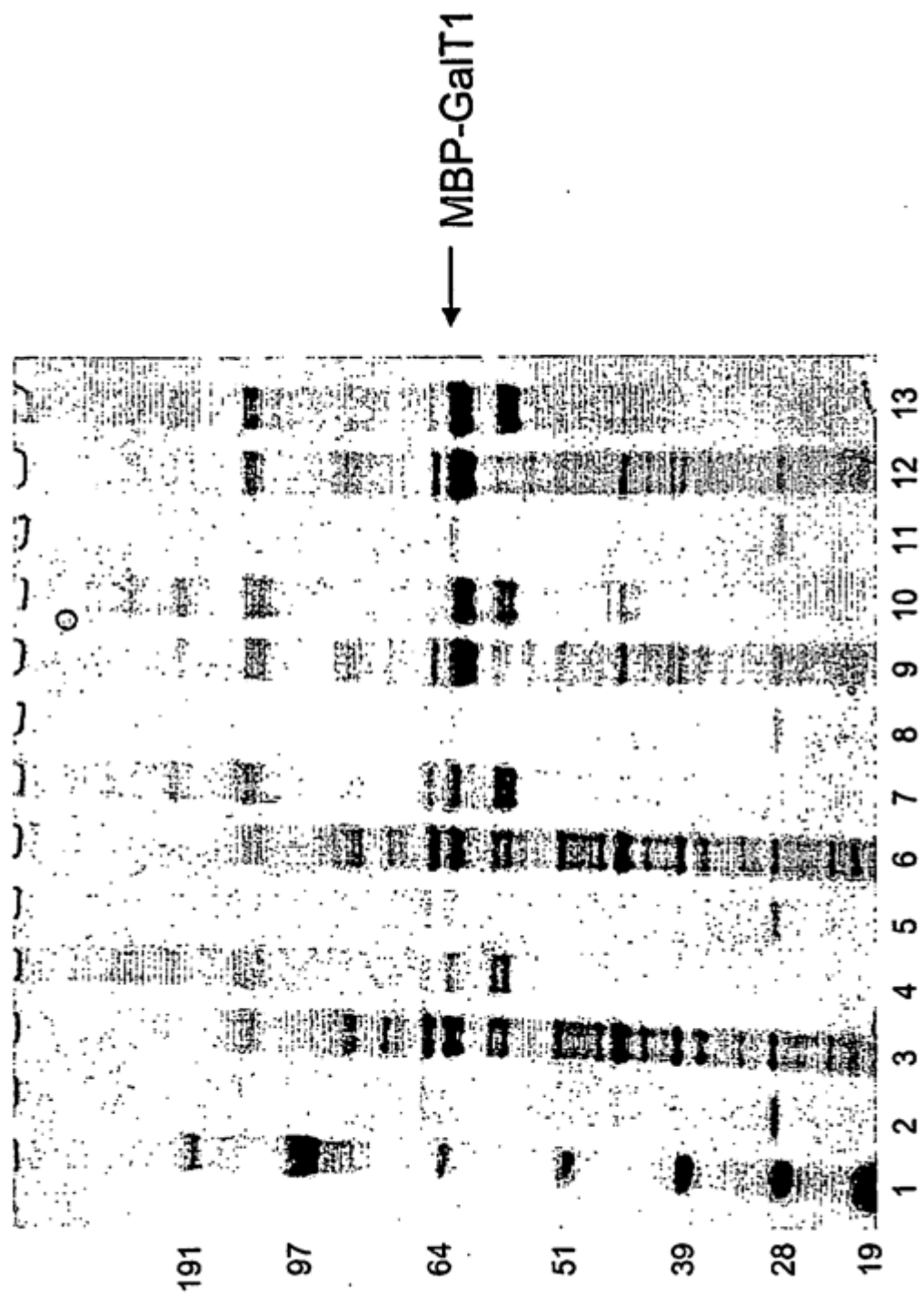


Figure 10

