

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 995 336**

51 Int. Cl.:

C12N 9/10 (2006.01)

C07K 16/28 (2006.01)

C07K 16/32 (2006.01)

C12P 21/00 (2006.01)

A61K 47/68 (2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.04.2016 E 17208517 (7)**

97 Fecha y número de publicación de la concesión europea: **30.10.2024 EP 3354726**

54 Título: **Proceso para la modificación de una glucoproteína utilizando una glucosiltransferasa que es o que deriva de una beta-(1,4)-n-acetilgalactosaminiltransferasa**

30 Prioridad:

23.04.2015 EP 15164864

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

10.02.2025

73 Titular/es:

**SYNAFFIX B.V. (100.00%)
Kloosterstraat 9
5349 AB Oss, NL**

72 Inventor/es:

**VAN BERKEL, SANDER SEBASTIAAN;
VAN GEEL, REMON;
WIJDEVEN, MARIA ANTONIA y
VAN DELFT, FLORIS LOUIS**

74 Agente/Representante:

ELZABURU, S.L.P

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 995 336 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCION

Proceso para la modificación de una glucoproteína utilizando una glucosiltransferasa que es o que deriva de una beta-(1,4)-n-acetilgalactosaminiltransferasa

Campo técnico de la invención

- 5 La presente invención se refiere a un proceso para la modificación enzimática de una glucoproteína. Más en particular, la invención se refiere a un proceso para la modificación de una glucoproteína con un nucleótido derivado de azúcar, utilizando una enzima glucosiltransferasa que es o deriva de una β -(1,4)-N- acetilgalactosaminiltransferasa. La invención también se refiere a una glucoproteína que se puede obtener mediante el proceso y a un bioconjugado que se puede obtener conjugando la glucoproteína con un conjugado enlazador.

10 Antecedentes de la invención

- 15 Las glucosiltransferasas constituyen una superfamilia de enzimas que participan en la síntesis de hidratos de carbono complejos presentes en las glucoproteínas y los glucolípidos. El papel fundamental de una glucosiltransferasa es transferir el resto glucosilo de un derivado de nucleótido a un aceptor de azúcar específico. Las β -1,4-galactosiltransferasas (β 4GalT) (EC 2.4.1.38) constituyen una de las subfamilias de la superfamilia de las glucosiltransferasas, que comprende al menos siete miembros Gal-T1 a Gal-T7, que catalizan la transferencia de galactosa (Gal) desde UDP- Gal a diferentes aceptores de azúcar. Un motivo común resultante de una galactosa transferasa en un residuo de GlcNAc terminal es la secuencia de lactosamina Gal β 4GlcNAc-R (LacNAc o LN), que posteriormente se modifica de diversas maneras mediante las adiciones de otros azúcares y grupos sulfato. La estructura de azúcar más común e importante de los glucoconjugados de membrana es la poli-N-acetil-lactosamina (poli-LN), que, ligada a proteínas (o lípidos), desempeña un papel importante en la comunicación, adhesión y señalización celular y son moléculas clave en la regulación de las respuestas inmunes.

- 20 Otro motivo terminal común encontrado en glucoconjugados de vertebrados e invertebrados es la secuencia GalNAc β 4GlcNAc-R (LacdiNAc o LDN). El motivo LDN está presente en las hormonas glucoproteicas de la hipófisis de mamíferos, donde los residuos GalNAc terminales están sulfatados en la posición 4-O y funcionan como marcadores de reconocimiento para el aclaramiento por el receptor Man/S4GnM de la célula endotelial. Sin embargo, las glucoproteínas de mamífero que no son de la pituitaria también contienen determinantes LDN. Además, la LDN y las modificaciones de las secuencias de LDN son determinantes antigénicos comunes en muchos nematodos y trematodos parásitos. La biosíntesis de LDN implica la transferencia de GalNAc a un GlcNAc terminal, un proceso ejecutado por transferasas GalNAc altamente específicas. Por ejemplo, Miller et al. en J. Biol. Chem. 2008, 283, p. 1985 publicaron que se cree que dos β -1,4-N-acetilgalactosaminiltransferasas estrechamente relacionadas, β 4GalNAc-T3 y β 4GalNAc-T4, explican la adición específica para proteínas de GalNAc unida en β 1,4 a oligosacáridos unidos a Asn en una serie de glucoproteínas, incluida la glucoproteína luteinizante (LH) y la anhidrasa carbónica 6 (CA6).

- 25 Se han identificado β -(1,4)-acetilgalactosaminiltransferasas (β -(1,4)-GalNAcT) en una variedad de organismos, incluidos los seres humanos, *Caenorhabditis elegans* (Kawar et al, J. Biol. Chem. 2002, 277, 34924), *Drosophila melanogaster* (Hoskins et al. Science 2007, 316, 1625) y *Trichoplusia ni* (Vadaie et al., J. Biol. Chem. 2004, 279, 33501).

- 30 Finalmente, además de las GalT y GalNAcT implicadas en la modificación de la N-glucoproteína, una clase no relacionada de enzimas llamada UDP-N-acetilgalactosamina:polipéptido N- acetilgalactosaminiltransferasas (también denominadas ppGalNAcT) es la responsable de la biosíntesis de enlaces de tipo mucina (GalNAc- α -1-O-Ser/Thr). Estas enzimas transfieren GalNAc del donante de azúcar UDP- GalNAc a residuos de serina y treonina, formando un enlace alfa anomérico típico en las O-glucoproteínas. A pesar de la aparente simplicidad de la función catalítica de las ppGalNAcT, se estima en función del análisis in silico que existen 24 genes humanos únicos solo de ppGalNAcT. Debido a que la glucosilación ligada a O tiene lugar por etapas, la adición de GalNAc a serina o treonina representa la primera etapa comprometida en la biosíntesis de mucina. A pesar de esta aparente simplicidad, múltiples miembros de la familia de las ppGalNAcT parecen ser necesarios para glucosilar completamente sus sustratos proteicos.

- 35 Se ha demostrado que las galactosiltransferasas son capaces de transferir, además del sustrato UDP-Gal natural, una gama de derivados de galactosa no naturales a un sustrato GlcNAc aceptor. Por ejemplo, Elling et al. en ChemBioChem 2001, 2, 884, mostraron que las proteínas que contienen GlcNAc terminal pueden ser biotiniladas mediante la transferencia de una galactosa modificada en 6 a partir de un azúcar UDP bajo la acción de una gama de galactosiltransferasas. Del mismo modo, fue demostrado por Pannecoucke et al. en Tetrahedron Lett. 2008, 49, 2294, que la 6-azido-6-desoxigalactosa se puede transferir (en cierta medida) del azúcar UDP correspondiente a un sustrato GlcNAc de molécula pequeña tras someterse a β 1,4- galactosiltransferasa bovina. El uso de glucosiltransferasas para derivados de galactosa modificados también se mencionó anteriormente en US 2008/0108557 (WO 2006/035057, Novo Nordisk A / S), donde se afirma que una amplia gama de derivados de galactosa modificados en C-6 (p.ej. tiol, azida, O-propargilo, aldehído) se puede transferir a un sustrato GlcNAc bajo la acción de β 1,4-galactosiltransferasa (bovina o humana), utilizando 2-10 equivalentes de azúcares UDP. Sin embargo, los datos proporcionados para sustentar tales afirmaciones se refieren únicamente a la variante 6-O-propargilo y 6-aldehído de galactosa. Una gama

de derivados de GalNAc con un controlador químico en C2 también se indican como sustratos para glucosiltransferasas pero no se proporcionaron ejemplos.

En particular, la mutación del residuo Tyr-289 en Leu-289 en β 4Gal-T1 bovina, según lo informado por Ramakrishnan et al. J. Biol. Chem. 2002, 23, 20833, crea un bolsillo catalítico de la enzima que puede facilitar que una molécula UDP-Gal lleve un controlador químico en C2, como 2-ceto-Gal. Mediante un procedimiento de dos etapas que implica primero la transferencia del resto de galactosa no natural seguida de la ligación de oxima en el controlador C-2, esta enzima mutante, β 4GalT (Y289L), se ha utilizado para la detección in vitro de residuos de O-GlcNAc en proteínas o de la presencia de un resto de GlcNAc terminal en los glucanos de la superficie celular de tejidos tumorales normales y malignos.

Por ejemplo Khidekel et al, J. Am. Chem. Soc. 2003, 125, 16162, revela la instalación quimioselectiva de una funcionalidad de cetona innatural para proteínas modificadas por O-GlcNAc con β 4GalT (Y289L). El resto de cetona sirve como un marcador único para "marcar" proteínas glucosiladas por O-GlcNAc con biotina utilizando ligación con oxima. Una vez biotinilados, los glucoconjugados se pueden detectar fácilmente por quimioluminiscencia usando estreptavidina conjugada con peroxidasa de rábano picante (HRP).

Por ejemplo WO 2007/095506, WO 2008/029281 (ambas de Invitrogen Corporation), WO 2014/065661 (SynAffix B.V.) y Clark et al. J. Am. Chem. Soc. 2008, 130, 11576, describen un enfoque similar, utilizando β 4GalT (Y289L) y variantes acidoacetil de galactosamina, con éxito similar.

Por ejemplo, US 8697061 (Glykos), describe un enfoque similar, utilizando β 4GalT (Y289L) y azúcares modificados en 2, con éxito similar.

β 4GalT mutante (Y289L) también se ha aplicado más recientemente de forma preparatoria para el radiomarcaje selectivo según el sitio de anticuerpos en los glucanos de cadena pesada, según describen Zeglis et al. en Bioconj. Chem. 2013, 24, 1057. En particular, la incorporación de monosacáridos de N-acetilgalactosamina modificados con azida (GalNAz) en los glucanos del anticuerpo permitió el marcado controlado con ^{89}Zr después de la introducción por química click del quelante apropiado.

Ramakrishnan et al. en Biochemistry 2004, 43, 12513, describen que el doble mutante β 4GalT (Y289L, M344H) pierde el 98% de su actividad dependiente de Mn^{2+} , pero muestra un 25-30% de actividad en presencia de Mg^{2+} , incluyendo una capacidad para transferir sustratos de galactosa modificados en C-2. El doble mutante β 4GalT (Y289L, M344H) resultó útil para ensayos de galactosilación in vitro, ya que se sabe que el requisito típico de 5-10 mM Mn^{2+} tiene efectos citotóxicos potenciales para las células.

Mercer et al., Bioconjugate Chem. 2013, 24, 144, describen que una enzima doble mutante Y289L- M344H- β 4Gal-T1 transfiere GalNAc y azúcares análogos al receptor GlcNAc en presencia de Mg^{2+} .

Los intentos de emplear una β -(1,4)-N-acetilgalactosaminiltransferasa de tipo salvaje, denominada aquí también β -(1,4)-GalNAcT, para la transferencia de GalNAc modificada en C-2 han tenido poco éxito hasta la fecha.

Bertozzi et al. en ACS Chem. Biol. 2009, 4, 1068, aplicaron la técnica del reportero químico bioortogonal para la obtención in vivo de imágenes moleculares de O-glucanos de tipo mucina en *C. elegans*. Los gusanos se trataron con la variante azida-azúcar de N-acetilgalactosamina (GalNAz) permitiendo la incorporación in vivo de este azúcar no natural. Aunque se observó la incorporación metabólica de GalNAz en glucoproteínas, la digestión tanto con condroitinasa ABC como con péptido N-glucosidasa F (PNGasa F) de lisado de *C. elegans*, seguida de ligación de Staudinger usando un marcador de fosfina y el sondeo posterior de las glucoproteínas mediante transferencia Western blot utilizando un anticuerpo α -Flag, indicó que la mayoría de los residuos GalNAz en las glucoproteínas estaban situados en otros tipos de glucanos distintos de los N-glucanos. Además, no se observó una unión detectable de glucoproteínas marcadas con azida a la lectina concanavalina A (ConA) específica de N-glucano, lo que concuerda con la hipótesis de que la gran mayoría de los glucanos marcados están unidos por O y no están unidos por N. En función de estas observaciones, se puede concluir que GalNAz no se incorpora metabólicamente a proteínas N-GlcNAcizadas en este organismo.

Una conclusión similar fue extraída más recientemente por Burnham-Marusch et al. en Plos One 2012, 7, e49020, donde también se observó la falta de reducción de la señal en el tratamiento con PNGasa, lo que indicaba que no había una incorporación aparente de GalNAz en las N-glucoproteínas. Burnham-Marusch et al. describen un estudio que utiliza la reacción de cicloadición de azida-alquino catalizada por Cu(I) de una sonda de alquino terminal con una glucoproteína marcada con azido para detectar glucoproteínas marcadas metabólicamente. Los resultados indicaron que la mayoría del marcador GalNAz se incorpora en las clases de glucanos que no son sensibles a la PNGasa F, por lo tanto no son N-glucoproteínas.

La alta especificidad de sustrato de una β -(1,4)-GalNAcT para UDP-GalNAc se hace evidente por el escaso reconocimiento de UDP-GlcNAc, UDP-Glc y UDP-Gal, para las cuales solo el 0,7%, 0,2% y 1% de actividad transferasa permanece, respectivamente, igual que describieron Kowar et al, J. Biol. Chem. 2002, 277, 34924.

En función de lo anterior, no es sorprendente que no se haya descrito un método in vitro para la modificación de glucoproteínas por medio de GalNAc-transferasa de un derivado de GalNAc no natural tal como un derivado 2-ceto o 2-azidoacetilo.

5 Al mismo tiempo, se ha descrito en Qasba et al., J. Mol. Biol. 2007, 365, 570 que la sustitución del residuo del sitio activo Ile o Leu en GalNAcT de invertebrados -correspondiente al residuo Tyr-289 en enzimas ortólogas β 4Gal-T1 humanas- por un residuo Tyr convierte a la enzima en una β -(1,4)galactosiltransferasa al reducir su actividad de N-acetilgalactosaminiltransferasa, casi 1000 veces, a la vez que aumenta su actividad de galactosiltransferasa 80 veces.

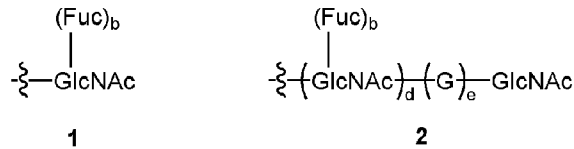
Taron et al, Carbohydr. Res. 2012, 362, 62, describe la incorporación metabólica en vivo de GalNAz en anclajes GPI.

Resumen de la invención

10 La presente invención se refiere a un proceso para la modificación de una glucoproteína, el proceso comprende poner en contacto una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal, con un nucleótido derivado de azúcar Su(A)-Nuc en presencia de una glucosiltransferasa, donde:

15 (i) la glucosiltransferasa es una β -(1,4)-N-acetilgalactosaminiltransferasa que comprende una secuencia que tiene al menos un 95% de identidad de secuencia con una cualquiera de SEQ. Id NO: 3, SEQ. ID NO: 4, SEQ. ID NO: 7 y SEQ. ID NO: 8;

(ii) el glucano que comprende un resto de GlcNAc terminal es según la fórmula (1) o (2):



donde:

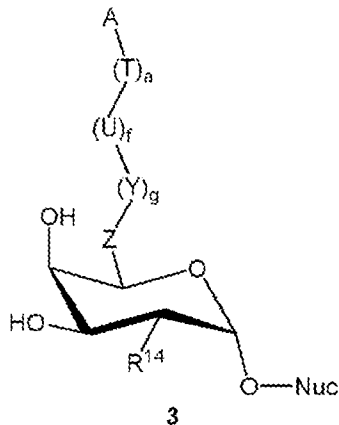
b es 0 o 1;

20 d es 0 o 1;

e es 0 o 1; y

G es un monosacárido, o un oligosacárido lineal o ramificado que comprende de 2 a 20 restos de azúcar; y

(iii) el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3):



25 donde:

a es 0 o 1;

f es 0 o 1;

g es 0 o 1;

Nuc es un nucleótido;

30

U es $[C(R^1)_2]_n$ o $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_q[C(R^1)_2]_q$, donde n es un número entero en el intervalo de 1 a 24; o es un número entero en el intervalo de 0 a 12; p y q son independientemente 0, 1 o 2; y R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I, OH y un grupo alquilo opcionalmente sustituido $C_1 - C_{24}$;

5 T es un grupo (hetero)arileno $C_3 - C_{12}$, donde el grupo (hetero)arileno está opcionalmente sustituido; A se selecciona del grupo que consiste en:

(a) $-N_3$

(b) $-C(O)R^3$

donde R^3 es un grupo alquilo opcionalmente sustituido $C_1 - C_{24}$;

10 (c) grupo (hetero)cicloalquinilo o un resto $-(CH_2)_iC\equiv C-R^4$ donde i es 0 - 10 y R^4 es hidrógeno o un grupo alquilo opcionalmente sustituido $C_1 - C_{24}$;

(d) $-SH$

(e) $-SC(O)R^8$

donde R^8 es un grupo alquilo $C_1 - C_{24}$ o grupo fenilo opcionalmente sustituido;

(f) $-SC(V)OR^8$

15 donde V es O o S, y R^8 es un grupo alquilo $C_1 - C_{24}$ o grupo fenilo opcionalmente sustituido;

(g) $-X$

donde X se selecciona del grupo que consiste en F, Cl, Br y I;

(h) $-OS(O)_2R^5$

20 donde R^5 se selecciona del grupo que consiste en grupos alquilo $C_1 - C_{24}$, grupos arilo $C_6 - C_{24}$, grupos alquilarilo $C_7 - C_{24}$ y grupos arilalquilo $C_7 - C_{24}$, donde los grupos alquilo, grupos arilo, grupos alquilarilo y grupos arilalquilo están opcionalmente sustituidos;

(i) R^{12}

donde R^{12} se selecciona del grupo que consiste en grupos alquenilo $C_2 - C_{24}$, grupos cicloalquenilo $C_3 - C_5$ y grupos alcadienilo $C_4 - C_8$ terminales opcionalmente sustituidos; y

25 (j) R^{13}

donde R^{13} es un grupo alenilo $C_3 - C_{24}$ terminal opcionalmente sustituido; y

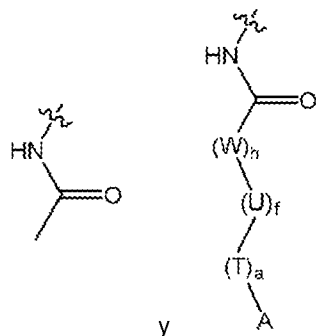
(k) $N(R^{17})_2$

donde R^{17} se selecciona independientemente del grupo que consiste en H y grupos alquilo $C_1 - C_{12}$;

Z es CH_2 , CF_2 o $C(O)$; o Z es $CHOH$ con la condición de que, cuando Z es $CHOH$, g es 0, f es 1 y a es 0 o 1;

30 Y se selecciona del grupo que consiste en O, S, $N(R^{15})$, $N(R^{15})C(O)$, $N(R^{15})C(O)N(R^{15})$, $N(R^{15})C(O)O$, $OC(O)N(R^{15})S(O)_2N(R^{15})$ y $N(R^{15})C(O)N(R^{15})S(O)_2O$, donde R^{15} se selecciona independientemente del grupo que consiste en H, grupos alquilo $C_1 - C_{12}$ y $(U)_f - (T)_a - A$ donde f, a, U, T y A son como se ha definido anteriormente; y

R^{14} se selecciona del grupo que consiste en:



en donde:

a, f, T, A y U son como se define anteriormente;

h es 0 o 1; y

W se selecciona del grupo que consiste en O, S, NR¹⁵, NHS(O)₂O y NHS(O)₂NR¹⁵,

5 en donde R¹⁵ es como se define anteriormente.

La invención se refiere además a una glucoproteína que se puede obtener mediante el proceso de acuerdo con la invención.

Descripción de las figuras

10 En la Figura 1, se muestran varios ejemplos de una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal, que puede modificarse mediante el proceso de acuerdo con la invención.

En la Figura 2, se muestra una forma de realización del proceso para la modificación de una glucoproteína, donde la glucoproteína es un anticuerpo. En esta forma de realización, un derivado de azúcar Su(A)-Nuc está unido al resto de GlcNAc terminal de un glucano de anticuerpo bajo la acción de una glucosiltransferasa, donde la glucosiltransferasa es, o deriva de, una β -(1,4)-N-acetilgalactosaminiltransferasa, para formar un anticuerpo modificado.

15 La Figura 3 muestra diferentes glucoformas de glucanos de anticuerpos G0, G1, G2, G0F, G1F y G2F.

La Figura 4 muestra un proceso para proporcionar una glucoproteína que comprende un glucano de acuerdo con la fórmula (10) mediante el tratamiento de una mezcla de glucoformas G0, G1, G2, G0F, G1F y G2F con sialidasa y galactosidasa, y un proceso para proporcionar una glucoproteína que comprende un glucano de acuerdo con la fórmula (1) por tratamiento de una mezcla de glucoformas G0, G1, G2, G0F, G1F y G2F con una endoglucosidasa. La incubación de las glucoproteínas que comprenden un glucano de acuerdo con la fórmula (10) o (1) con un derivado de UDP-GalNAc modificado con azida, p. ej. 6-azidoGalNAc, conduce a una glucoproteína modificada con azida (33) o (32), respectivamente.

20 La Figura 5 muestra un conjunto representativo de grupos funcionales (A) en la glucoproteína modificada de acuerdo con la invención, que tras la reacción con un grupo reactivo Q¹ lleva a la conexión del grupo CG y una glucoproteína funcionalizada.

Descripción detallada de la invención

Definiciones

30 El verbo "comprender" según se usa en esta descripción y en las reivindicaciones, y sus conjugaciones, se usa en su sentido no limitativo para significar que se incluyen los artículos que siguen a la palabra, pero los artículos no mencionados específicamente no se excluyen.

Además, la referencia a un elemento mediante el artículo indefinido "un" o "una" no excluye la posibilidad de que más de uno de los elementos esté presente, a menos que el contexto requiera claramente que haya uno y solo uno de los elementos. El artículo indefinido "un" o "una" por lo general significa "al menos uno/a".

35 Los grupos alquilo no sustituidos tienen la fórmula general C_nH_{2n+1} y pueden ser lineales o ramificados. Los grupos alquilo no sustituidos también pueden contener un resto cíclico y, por lo tanto, tienen la fórmula general concomitante C_nH_{2n-1}. Opcionalmente, los grupos alquilo están sustituidos por uno o más sustituyentes especificados adicionalmente en este documento. Los ejemplos de grupos alquilo incluyen metilo, etilo, propilo, 2-propilo, t-butilo, 1-hexilo, 1-dodecilo, etc.

40 Un grupo arilo comprende de seis a doce átomos de carbono y puede incluir estructuras monocíclicas y bicíclicas. Opcionalmente, el grupo arilo puede estar sustituido por uno o más sustituyentes adicionalmente especificados en este documento. Algunos ejemplos de grupos arilo son fenilo y naftilo.

45 Los grupos arilalquilo y los grupos alquilarilo comprenden al menos siete átomos de carbono y pueden incluir estructuras monocíclicas y bicíclicas. Opcionalmente, los grupos arilalquilo y alquilarilo pueden estar sustituidos por uno o más sustituyentes especificados adicionalmente en este documento. Un grupo arilalquilo es, por ejemplo, bencilo. Un grupo alquilarilo es, por ejemplo, 4-t-butilfenilo.

50 Los grupos heteroarilo comprenden al menos dos átomos de carbono (es decir, al menos C₂) y uno o más heteroátomos N, O, P o S. Un grupo heteroarilo puede tener una estructura monocíclica o bicíclica. Opcionalmente, el grupo heteroarilo puede estar sustituido por uno o más sustituyentes adicionalmente especificados en este documento. Los ejemplos de grupos heteroarilo adecuados incluyen piridinilo, quinolinilo, pirimidinilo, pirazinilo, pirazolilo, imidazolilo, tiazolilo, pirrolilo, furanilo, triazolilo, benzofuranilo, indolilo, purinilo, benzoxazolilo, tienilo, fosfolilo y

oxazolilo.

Los grupos heteroarilalquilo y los grupos alquilheteroarilo comprenden al menos tres átomos de carbono (es decir, al menos C₃) y puede incluir estructuras monocíclicas y bicíclicas. Opcionalmente, los grupos heteroarilo pueden estar sustituidos por uno o más sustituyentes adicionalmente especificados en este documento.

- 5 Cuando un grupo arilo se denota como un grupo (hetero)arilo, la notación pretende incluir un grupo arilo y un grupo heteroarilo. De forma similar, se pretende que un grupo alquilo (hetero)arilo incluya un grupo alquilarilo y un grupo alquilheteroarilo, y (hetero)arilalquilo incluye un grupo arilalquilo y un grupo heteroarilalquilo. Un grupo (hetero)arilo C₂ - C₂₄ debe interpretarse como que incluye un grupo heteroarilo C₂ - C₂₄ y un grupo arilo C₆ - C₂₄. Del mismo modo, se pretende que un grupo alquil(hetero)arilo C₃ - C₂₄ incluya un grupo alquilarilo C₇ - C₂₄ y un grupo alquilheteroarilo C₃ - C₂₄, y se pretende que un grupo (hetero)alrilalquilo C₃ - C₂₄ incluya un grupo arilaquilo C₇ - C₂₄ y un grupo heteroarilalquilo C₃ - C₂₄.

- 15 A menos que se indique lo contrario, grupos alquilo, grupos alquenilo, alquenos, alquinos, grupos (hetero)arilo, grupos (hetero)arilalquilo, grupos alquilo(hetero)arilo, grupos alquilenilo, grupos alquenilenilo, grupos cicloalquilenilo, grupos (hetero)arilenilo, grupos alquilo(hetero)arilenilo, grupos(hetero)arilalquilenilo, grupos alquenilo, grupos alquinilo, grupos cicloalquilo, grupos alcoxi, grupos alqueniloxi, grupos (hetero)ariloxi, grupos alquiniloxi y grupos cicloalquiloxi pueden estar sustituidos con uno o más sustituyentes seleccionados independientemente del grupo que consiste en grupos alquilo C₁ - C₁₂, grupos alquenilo C₂ - C₁₂, grupos alquinilo C₂ - C₁₂, grupos cicloalquilo C₃ - C₁₂, grupos cicloalquenilo C₅ - C₁₂, grupos cicloalquinilo C₈ - C₁₂, grupos alcoxi C₁ - C₁₂, grupos alqueniloxi C₂ - C₁₂, grupos alquiniloxi C₂ - C₁₂, grupos cicloalquiloxi C₃ - C₁₂, halógenos, grupos amino, grupos oxo y sililo, donde los grupos sililo pueden estar representados por la fórmula (R²)₃Si-, donde R₂ se selecciona independientemente del grupo que consiste en grupos alquilo C₁ - C₁₂, grupos alquenilo C₂ - C₁₂, grupos alquinilo C₂ - C₁₂, grupos cicloalquilo C₃ - C₁₂, grupos alcoxi C₁ - C₁₂, grupos alqueniloxi C₂ - C₁₂, grupos alquiniloxi C₂ - C₁₂ y grupos cicloalquiloxi C₃ - C₁₂, en donde los grupos alquilo, grupos alquenilo, grupos alquinilo, grupos cicloalquilo, grupos alcoxi, grupos alqueniloxi, grupos alquiniloxi y grupos cicloalquiloxi están opcionalmente sustituidos, los grupos alquilo, los grupos alcoxi, los grupos cicloalquilo y los grupos cicloalcoxi están opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo que consiste en O, N y S.

- 30 Un grupo alquinilo comprende un enlace triple carbono-carbono. Un grupo alquinilo no sustituido que comprende un enlace triple tiene la fórmula general C_nH_{2n-3}. Un alquinilo terminal es un grupo alquinilo en el que el enlace triple está ubicado en una posición terminal de una cadena de carbono. Opcionalmente, el grupo alquinilo está sustituido con uno o más sustituyentes adicionalmente especificados en este documento, y / o interrumpido por heteroátomos seleccionados del grupo de oxígeno, nitrógeno y azufre. Los ejemplos de grupos alquinilo incluyen etinilo, propinilo, butinilo, octinilo, etc.

- 35 Un grupo cicloalquinilo es un grupo alquinilo cíclico. Un grupo cicloalquinilo no sustituido que comprende un enlace triple tiene la fórmula general C_nH_{2n-5}. Opcionalmente, un grupo cicloalquinilo está sustituido con uno o más sustituyentes adicionalmente especificados en este documento. Un ejemplo de un grupo cicloalquinilo es ciclooctinilo.

- Un grupo heterocicloalquinilo es un grupo cicloalquinilo interrumpido por heteroátomos seleccionados del grupo de oxígeno, nitrógeno y azufre. Opcionalmente, un grupo heterocicloalquinilo está sustituido con uno o más sustituyentes adicionalmente especificados en este documento. Un ejemplo de un grupo heterocicloalquinilo es azaciclooctinilo.

- 40 Un grupo (hetero)arilo comprende un grupo arilo y un grupo heteroarilo. Un grupo alquilo(hetero)arilo comprende un grupo alquilarilo y un grupo alquilheteroarilo. Un grupo (hetero)arilalquilo comprende un grupo arilalquilo y un grupo heteroarilalquilo. Un grupo (hetero)alquinilo comprende un grupo alquinilo y un grupo heteroalquinilo. Un grupo (hetero)cicloalquinilo comprende un grupo cicloalquinilo y un grupo heterocicloalquinilo.

Un compuesto (hetero)cicloalquino se define aquí como un compuesto que comprende un grupo (hetero)cicloalquinilo.

- 45 Varios de los compuestos descritos en esta descripción y en las reivindicaciones se pueden describir como compuestos (hetero)cicloalquino fusionados, es decir, compuestos (hetero)cicloalquino en los que una segunda estructura de anillo se fusiona, es decir, se anilla, al grupo (hetero)cicloalquinilo. Por ejemplo, en un compuesto (hetero)ciclooctileno fusionado, un cicloalquilo (por ejemplo, un ciclopropilo) o un areno(por ejemplo, benceno) se puede anillar al grupo (hetero)ciclooctinilo. El enlace triple del grupo (hetero)ciclooctinilo en un compuesto fusionado (hetero)ciclooctileno puede estar situado en cualquiera de las tres ubicaciones posibles, es decir, en la posición 2, 3 o 4 del resto de deciclooctileno (numeración según la "Nomenclatura de Química Orgánica de la IUPAC", Regla A31.2). La descripción de cualquier compuesto (hetero)ciclooctileno fusionado en esta descripción y en las reivindicaciones pretende incluir los tres regioisómeros individuales del resto de ciclooctileno.

- 55 El término general "azúcar" se usa en la presente memoria para indicar un monosacárido, por ejemplo glucosa (Glc), galactosa (Gal), manosa (Man) y fucosa (Fuc). El término "derivado de azúcar" se usa en este documento para indicar un derivado de un azúcar monosacárido, es decir, un azúcar monosacárido que comprende sustituyentes y/o grupos funcionales. Los ejemplos de un derivado de azúcar incluyen aminoazúcares y ácidos de azúcares, como glucosamina (GlcNH₂), galactosamina (GalNH₂) N-acetilglucosamina (GlcNAc), N-acetilgalactosamina (GalNAc), ácido siálico (Sia) que también se conoce como ácido N-acetilneuramínico (NeuNAc), y ácido N-acetilmurámico (MurNAc), ácido

glucurónico (GlcA) y ácido idurónico (IdoA).

El término "nucleótido" se usa en este documento en su significado científico normal. El término "nucleótido" se refiere a una molécula que está compuesta por una nucleobase, un azúcar de cinco carbonos (ribosa o 2-desoxirribosa) y uno, dos o tres grupos fosfato. Sin el grupo fosfato, la nucleobase y el azúcar componen un nucleósido. Así, un nucleótido también puede denominarse nucleósido monofosfato, un nucleósido difosfato o un nucleósido trifosfato. La nucleobase puede ser adenina, guanina, citosina, uracilo o timina. Los ejemplos de un nucleótido incluyen difosfato de uridina (UDP), difosfato de guanosina (GDP), difosfato de timidina (TDP), difosfato de citidina (CDP) y monofosfato de citidina (CMP).

El término "proteína" se usa en este documento en su significado científico normal. Aquí, los polipéptidos que comprenden aproximadamente 10 o más aminoácidos se consideran proteínas. Una proteína puede comprender aminoácidos naturales, pero también no naturales.

El término "glucoproteína" se usa en este documento en su significado científico normal y se refiere a una proteína que comprende una o más cadenas de monosacáridos u oligosacáridos ("glucanos") unidas de manera covalente a la proteína. Un glucano se puede unir a un grupo hidroxilo en la proteína (glucano unido a un átomo de oxígeno), por ejemplo al grupo hidroxilo de serina, treonina, tirosina, hidroxilisina o hidroxiprolina, o a una función de nitrógeno en la proteína (N-glucoproteína), por ejemplo asparagina o arginina, o a un carbono en la proteína (C-glucoproteína), por ejemplo triptófano. Una glucoproteína puede comprender más de un glucano, puede comprender una combinación de uno o más monosacáridos y uno o más oligosacáridos glucanos, y puede comprender una combinación de glucanos unidos a N, unidos a O y unidos a C. Se estima que más del 50% de todas las proteínas tienen alguna forma de glucosilación y, por lo tanto, se pueden considerar glucoproteínas. Los ejemplos de glucoproteínas incluyen PSMA (antígeno prostático específico de membrana), CAL (lipasa de candida antártica), gp41, gp120, EPO (eritropoyetina), proteína anticongelante y anticuerpos.

El término "glucano" se usa en este documento en su significado científico normal y se refiere a una cadena de monosacáridos u oligosacáridos que está unida a una proteína. El término glucano se refiere a la parte de carbohidrato de una glucoproteína. El glucano se une a una proteína a través del carbono C-1 de un azúcar, que puede estar sin sustitución adicional (monosacárido) o puede estar sustituido adicionalmente en uno o más de sus grupos hidroxilo (oligosacárido). Un glucano de origen natural comprende típicamente de 1 a aproximadamente 10 restos de sacáridos. Sin embargo, cuando una cadena de sacáridos más larga está unida a una proteína, dicha cadena de sacáridos también se considera un glucano en este caso.

Un glucano de una glucoproteína puede ser un monosacárido. Típicamente, un glucano monosacárido de una glucoproteína consiste en una sola N-acetilglucosamina (GlcNAc), glucosa (Glc), manosa (Man) o fucosa (Fuc) unidas de manera covalente a la proteína.

Un glucano también puede ser un oligosacárido. Una cadena de oligosacáridos de una glucoproteína puede ser lineal o ramificada. En un oligosacárido, el azúcar que está directamente unido a la proteína se llama azúcar central. En un oligosacárido, un azúcar que no está directamente unido a la proteína y está unido a al menos otros dos azúcares se llama azúcar interno. En un oligosacárido, un azúcar que no está directamente unido a la proteína sino a otro azúcar, es decir, que no lleva más sustituyentes de azúcar en uno o más de sus otros grupos hidroxilo, se llama azúcar terminal. Para evitar dudas, pueden existir múltiples azúcares terminales en un oligosacárido de una glucoproteína, pero solo un azúcar central.

Un glucano puede ser un glucano unido a O, un glucano unido a N o un glucano unido a C. En un glucano unido a O, un glucano monosacárido u oligosacárido está unido a un átomo de O en un aminoácido de la proteína, típicamente a través de un grupo hidroxilo de serina (Ser) o treonina (Thr). En un glucano unido a N, un glucano monosacárido u oligosacárido se une a la proteína a través de un átomo de N en un aminoácido de la proteína, típicamente a través de un nitrógeno amídico en la cadena lateral de asparagina (Asn) o arginina (Arg). En un glucano unido a C, un glucano monosacárido u oligosacárido está unido a un átomo de C en un aminoácido de la proteína, típicamente a un átomo de C de triptófano (Trp).

El extremo de un oligosacárido que está directamente unido a la proteína se denomina el extremo reductor de un glucano. El otro extremo del oligosacárido se denomina el extremo no reductor de un glucano.

Para los glucanos unidos a O, existe una gran diversidad de cadenas. Los glucanos unidos a O de origen natural normalmente presentan un resto de α -O-GalNAc unido a serina o treonina, además sustituido con otra GalNAc, galactosa, GlcNAc, ácido siálico y/o fucosa. El aminoácido hidroxilado que lleva la sustitución de glucano puede formar parte de cualquier secuencia de aminoácidos de la proteína.

Para los glucanos unidos a N, existe una gran diversidad de cadenas. Los glucanos unidos a N de origen natural presentan típicamente un resto de β -N-GlcNAc unido a asparagina, a su vez sustituido en su 4-OH con β -GlcNAc, a su vez sustituido adicionalmente en su 4-OH con β -Man, a su vez sustituido en su 3-OH y 6-OH con α -Man, lo que lleva al pentasacárido de glucano $\text{Man}_3\text{GlcNAc}_2$. El resto de GlcNAc central puede estar sustituido adicionalmente en su 6-OH por α -Fuc. El pentasacárido $\text{Man}_3\text{GlcNAc}_2$ es el andamiaje común de oligosacáridos de casi todas las glucoproteínas unidas a N y puede llevar una amplia variedad de otros sustituyentes, que incluyen pero no se limitan

a Man, GlcNAc, Gal y ácido siálico. La asparagina que se sustituye con el glucano en su cadena lateral típicamente forma parte de la secuencia Asn-X-Ser/Thr, donde X es cualquier aminoácido excepto prolina y Ser/Thr es serina o treonina.

El término "anticuerpo" se utiliza en este documento con su significado científico normal. Un anticuerpo es una proteína generada por el sistema inmunitario que es capaz de reconocer y unirse a un antígeno específico. Un anticuerpo es un ejemplo de una glucoproteína. El término anticuerpo en la presente memoria se usa en su sentido más amplio e incluye específicamente anticuerpos monoclonales, anticuerpos policlonales, dímeros, multímeros, anticuerpos multiespecíficos (por ejemplo, anticuerpos biespecíficos), fragmentos de anticuerpos y anticuerpos de cadena doble y simple. El término "anticuerpo" en este documento también pretende incluir anticuerpos humanos, anticuerpos humanizados, anticuerpos quiméricos y anticuerpos que se unen específicamente al antígeno del cáncer. El término "anticuerpo" pretende incluir anticuerpos completos, pero también fragmentos de un anticuerpo, por ejemplo un fragmento Fab de anticuerpo, F(ab')₂, un fragmento Fv o fragmento Fc de un anticuerpo escindido, un fragmento scFv-Fc, un minicuerpo, un diacuerpo o un scFv. Además, el término incluye anticuerpos genéticamente modificados y derivados de un anticuerpo. Los anticuerpos, fragmentos de anticuerpos y anticuerpos genéticamente modificados se pueden obtener mediante métodos que son conocidos en la técnica. Los anticuerpos comercializados adecuados incluyen, entre otros, abciximab, rituximab, basiliximab, palivizumab, infliximab, trastuzumab, alemtuzumab, adalimumab, tositumomab-l131, cetuximab, ibrituximab tiuxetan, omalizumab, bevacizumab, natalizumab, ranibizumab, panitumumab, eculizumab, certolizumab pegol, golimumab, canakinumab, catumaxomab, ustekinumab, tocilizumab, ofatumumab, denosumab, belimumab, ipilimumab y brentuximab.

Identidad/similitud

En el contexto de la invención, una proteína o un fragmento de proteína se representa mediante una secuencia de aminoácidos.

Debe entenderse que cada proteína o fragmento de proteína o péptido o péptido derivado o polipéptido derivado como se identifica en el presente documento por un número de identidad de secuencia dado (SEQ ID N°) no está limitado a esta secuencia específica como se describe. La "identidad de secuencia" se define aquí como una relación entre dos o más secuencias de aminoácidos (polipéptidos o proteínas), según se determina comparando las secuencias. En la técnica, "identidad" también significa el grado de similitud de secuencia entre las secuencias de aminoácidos, según el caso, tal y como se determina por la coincidencia entre cadenas de tales secuencias. A menos que se indique lo contrario en este documento, la identidad o similitud con un SEQ ID N° dado significa identidad o similitud en función de la longitud completa de dicha secuencia (es decir, en toda su longitud o como un todo).

Cualquier enzima abarcada por la invención que tenga menos de 100% de identidad de secuencia con la secuencia específicamente indicada definida por su SEQ ID N° preferiblemente tiene una actividad enzimática que es al menos 10%, 20%, 30%, 40%, 50%, 60 %, 70% o preferiblemente al menos 80% o 90% o al menos 100% de la actividad enzimática de la enzima que tiene una identidad del 100% respecto de dicha secuencia definida por la SEQ ID N°.

La "similitud" entre dos secuencias de aminoácidos se determina comparando la secuencia de aminoácidos y sus sustitutos de aminoácidos conservados de un polipéptido con la secuencia de un segundo polipéptido. La "identidad" y la "similitud" se pueden calcular fácilmente mediante métodos conocidos, incluyendo, entre otros, los que se describen en (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heine, G., Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; y Carillo, H., and Lipman, D., SIAM J. Applied Math., 48:1073 (1988).

Los métodos preferidos para determinar la identidad están diseñados para proporcionar la mayor coincidencia entre dos o más secuencias sometidas a evaluación. Los métodos para determinar la identidad y la similitud están codificados en programas de ordenador disponibles públicamente. Los métodos preferidos del programa de ordenador para determinar la identidad y la similitud entre dos secuencias incluyen, por ejemplo, el paquete de programas GcG (Devereux, J., et al., Nucleic Acids Research 12 (1): 387 (1984)), BestFit, BIASTP, BLASTN y FASTA (Altschul, S. F. et al., J. Mol. Biol. 215: 403 - 410 (1990)) El programa BLAST X está puesto a disposición del público por el NCBI y otras fuentes (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990)). El conocido algoritmo de Smith-Waterman también se puede usar para determinar la identidad.

Los parámetros preferidos para la comparación de secuencias polipeptídicas incluyen los siguientes: Algoritmo: Needleman and Wunsch, J. Mol. Biol. 48: 443-453 (1970); Matriz de comparación: BLOSSUM62 de Hentikoff y Hentikoff, Proc. Natl. Acad. Sci. EE. UU. 89: 10915 - 10919 (1992)); Penalización por hueco: 12; y Penalización por Longitud de Hueco: 4. Un programa útil con estos parámetros está disponible públicamente como el programa "Ogap" de Genetics Computer Group, ubicado en Madison, WI. Los parámetros anteriormente mencionados son los parámetros predeterminados para las comparaciones de aminoácidos (junto con la ausencia de penalizaciones para los huecos finales).

Opcionalmente, para determinar el grado de similitud de aminoácidos, la persona experta también puede tener en cuenta las denominadas sustituciones de aminoácidos "conservadoras", como será evidente para el experto en la materia. Las sustituciones conservadoras de aminoácidos se refieren a la intercambiabilidad de residuos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales hidroxilas alifáticas es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Los grupos conservadores preferidos de sustitución de aminoácidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina y asparagina-glutamina. Las variantes de sustitución de la secuencia de aminoácidos descrita en este documento son aquellas en las que se ha eliminado al menos un residuo en las secuencias descritas y se ha insertado un residuo diferente en su lugar. Preferiblemente, el cambio del aminoácido es conservador. Las sustituciones conservadoras preferidas para cada uno de los aminoácidos de origen natural son las siguientes: Ala a Ser; Arg a Lys; Asn a Gln o a His; Asp a Glu; Cys a Ser o Ala; Gln a Asn; Glu a Asp; Gly a Pro; His a Asn o Gln; Ile a Leu o Val; Leu a Ile o Val; Lys a Arg; Gln o Glu; Met a Leu o He; Phe to Met, Leu o Tyr; Ser a Thr; Thr to Ser; Trp a Tyr o His; Tyr a Trp o Phe; y, Val a Ile o Leu.

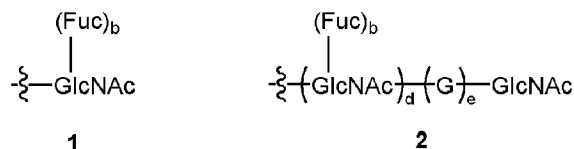
Proceso para la modificación de una glucoproteína

La presente invención se refiere a un proceso para la modificación de una glucoproteína, bajo la acción de una glucosiltransferasa, donde la glucosiltransferasa es una o, se deriva de una, β -(1,4)-N-acetilgalactosaminiltransferasa para obtener una glucoproteína modificada. Preferiblemente el proceso es un proceso in vitro.

En particular, la invención se refiere a un proceso para la modificación de una glucoproteína, proceso que comprende poner en contacto una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal con un nucleótido derivado de azúcar Su(A)-Nuc en presencia de una glucosiltransferasa, donde:

(i) la glucosiltransferasa es una β -(1,4)-N-acetilgalactosaminiltransferasa que comprende una secuencia que tiene al menos un 95% de identidad de secuencia con cualquiera de la SEQ. ID NO: 3, SEQ. ID NO: 4, SEQ. ID NO: 7 y SEQ. ID NO: 8;

(ii) el glucano que comprende un resto de GlcNAc terminal es según la fórmula (1) o (2):



donde:

b es 0 o 1;

d es 0 o 1;

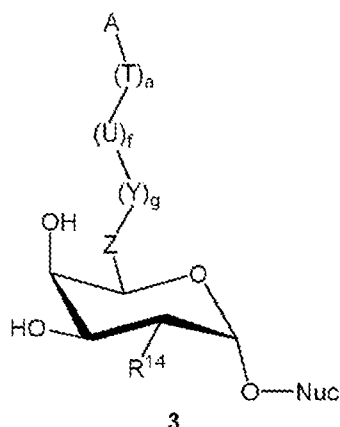
e es 0 o 1;

G es un monosacárido, o un oligosacárido lineal o ramificado que comprende de 2 a 20 restos de azúcar;

y

(iii) el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3):

donde:



a es 0 o 1;

f es 0 o 1;

g es 0 o 1;

5 Nuc es un nucleótido;

U es $[C(R^1)_2]_n$ o $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$, donde n es un número entero en el intervalo de 1 a 24; o es un número entero en el intervalo de 0 a 12; p y q son independientemente 0, 1 o 2; y R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I, OH y un grupo alquilo opcionalmente sustituido $C_1 - C_{24}$;

10 T es un grupo (hetero)arileno $C_3 - C_{12}$, donde el grupo (hetero)arileno está opcionalmente sustituido; A se selecciona del grupo que consiste en:

(a) $-N_3$

(b) $-C(O)R^3$

donde R^3 es un grupo alquilo opcionalmente sustituido $C_1 - C_{24}$;

15 (c) grupo (hetero)cicloalquinilo o un resto $-(CH_2)_iC\equiv C-R^4$ donde i es 0 - 10 y R^4 es hidrógeno o un grupo alquilo opcionalmente sustituido $C_1 - C_{24}$;

(d) $-SH$

(e) $-SC(O)R^8$

donde R^8 es un grupo alquilo $C_1 - C_{24}$ o grupo fenilo opcionalmente sustituido;

(f) $-SC(V)OR^8$

20 donde V es O o S, y R^8 es un grupo alquilo $C_1 - C_{24}$ o grupo fenilo opcionalmente sustituido;

(g) $-X$

donde X se selecciona del grupo que consiste en F, Cl, Br y I;

(h) $-OS(O)_2R^5$

25 donde R^5 se selecciona del grupo que consiste en grupos alquilo $C_1 - C_{24}$, grupos arilo $C_6 - C_{24}$, grupos alquilarilo $C_7 - C_{24}$ y grupos arilalquilo $C_7 - C_{24}$, donde los grupos alquilo, grupos arilo, grupos alquilarilo y grupos arilalquilo están opcionalmente sustituidos;

(i) R^{12}

donde R^{12} se selecciona del grupo que consiste en grupos alquenilo $C_2 - C_{24}$, grupos cicloalquenilo $C_3 - C_5$ y grupos alcadienilo $C_4 - C_8$ terminales opcionalmente sustituidos; y

30 (j) R^{13}

donde R^{13} es un grupo alenilo $C_3 - C_{24}$ terminal opcionalmente sustituido; y

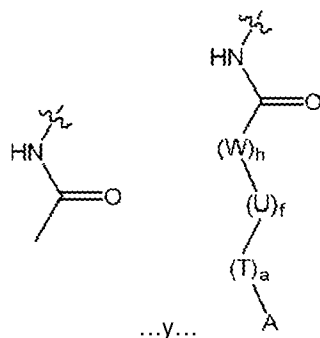
(k) $N(R^{17})_2$

donde R^{17} se selecciona independientemente del grupo que consiste en H y grupos alquilo $C_1 - C_{12}$;

Z es CH_2 , CF_2 o C(O); o Z es CHOH con la condición de que, cuando Z es CHOH, g es 0, f es 1 y a es 0 o 1;

5 Y se selecciona del grupo que consiste en O, S, $N(R^{15})$, $(R^{15})C(O)$, $N(R^{15})C(O)N(R^{15})$, $N(R^{15})C(O)O$, $OC(O)N(R^{15})S(O)_2N(R^{15})$ y $N(R^{15})C(O)N(R^{15})S(O)_2O$, donde R^{15} se selecciona independientemente del grupo que consiste en H, grupos alquilo $C_1 - C_{12}$ y $(U)_f - (T)_a - A$ donde f, a, U, T y A son como se ha definido anteriormente; y

R^{14} se selecciona del grupo que consiste en:



donde:

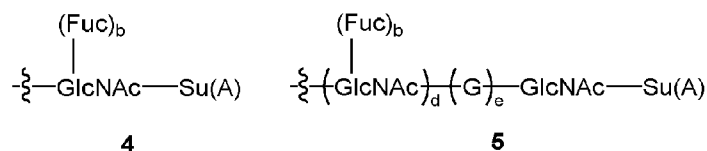
a, f, T, A y U son como se ha definido anteriormente;

10 h es 0 o 1; y

W se selecciona del grupo que consiste en O, S, NR^{15} , $NHS(O)_2O$ y $NHS(O)_2NR^{15}$, en donde R^{15} es como se ha definido anteriormente.

15 En una forma de realización, A en $Su(A)$ -Nuc es según la fórmula (3) y se selecciona del grupo que consiste en las opciones (a) a (j) tal como se ha definido anteriormente. En otra forma de realización, A en $Su(A)$ -Nuc es según la fórmula (3) y se selecciona del grupo que consiste en las opciones (a) a (d) y (g) a (k) como se ha definido anteriormente, más preferiblemente del grupo que consiste en las opciones (a) a (d) y (g) a (j).

Como se ha descrito anteriormente, el proceso de acuerdo con la invención para la modificación de una glucoproteína proporciona una glucoproteína modificada. Una glucoproteína modificada se define aquí como una glucoproteína que comprende un glucano de acuerdo con la fórmula (4) o (5):

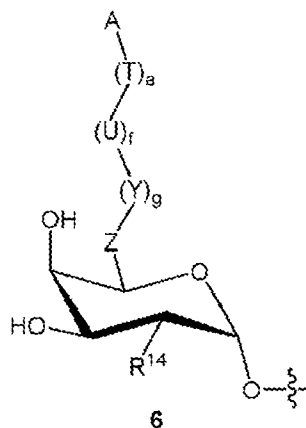


20

donde:

b, d, e y G son como se ha definido anteriormente; y

$Su(A)$ es un derivado de azúcar de acuerdo con la fórmula (6):



donde:

R¹⁴, Z, Y, U, T, A, g, f y a son como se ha definido anteriormente.

En el glucano de la glucoproteína modificada según la fórmula (4) y (5), el C1 de Su(A) derivado de azúcar está conectado al C4 del resto de GlcNAc a través de un enlace β-1,4-O-glucosídico.

- 5 El proceso para la modificación de una glucoproteína puede comprender además la etapa de proporcionar una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal. Por lo tanto, también se describe un proceso para la modificación de una glucoproteína que comprende las etapas de:

10 (1) proporcionar una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal, donde el glucano que comprende un resto de GlcNAc terminal es según la fórmula (1) o (2) como se ha definido anteriormente; y

(2) poner en contacto dicha glucoproteína con un nucleótido derivado de azúcar Su(A)-Nuc, en presencia, más particularmente bajo la acción de, una glucosiltransferasa, donde la glucosiltransferasa, es o se deriva de una β-(1,4)-N-acetilgalactosaminiltransferasa y donde Su(A)-Nuc es según la fórmula (3) como se ha definido anteriormente.

- 15 La glucoproteína que comprende un glucano que comprende un resto GlcNAc terminal, el nucleótido derivado de azúcar Su(A)-Nuc y la glucoproteína modificada, y las formas de realización preferidas de los mismos, se describen con más detalle a continuación.

La glucosiltransferasa que es o deriva de una β-(1,4)-N-acetilgalactosaminiltransferasa se describe con más detalle a continuación.

- 20 Preferiblemente, la β-(1,4)-N-acetilgalactosaminiltransferasa es o deriva de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 46, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 71, 72 y 73. Esto es particularmente preferido cuando R¹⁴ es -NHC(O)CH₃, -NHC(O)-(W)_h-(U)_f-(T)_a-A o -N₃.

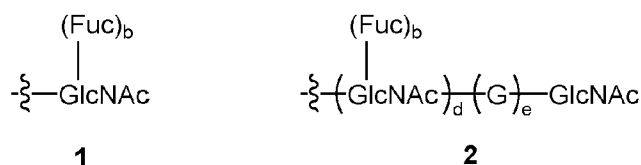
- 25 Alternativamente, la β-(1,4)-N-acetilgalactosaminiltransferasa es o deriva del grupo que consiste en SEQ ID N°: 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 48, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70 y 74. Esto es particularmente preferido cuando R¹⁴ es -OH.

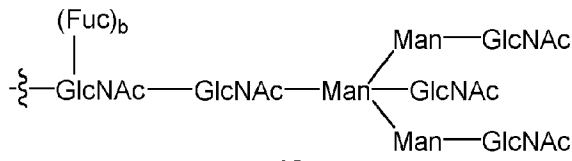
- 30 Alternativamente, la β-(1,4)-N-acetilgalactosaminiltransferasa tiene al menos 50% de identidad con una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 46, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 71, 72 y 73. En este contexto, se prefiere además que la β-(1,4)-N-acetilgalactosaminiltransferasa tenga al menos 55% de identidad de secuencia, preferiblemente al menos 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o preferiblemente 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 46, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 71, 72 y 73. Estas enzimas son particularmente preferidas cuando R¹⁴ es -NHC(O)CH₃, -NHC(O)-(W)_h-(U)_f-(T)_a-A o -N₃.

- 35 En otro aspecto preferido del proceso descrito en este documento, la β-(1,4)-N-acetilgalactosaminiltransferasa tiene al menos 50% de identidad con una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 48, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70 y 74. En esta forma de realización, se prefiere además que la β-(1,4)-N-acetilgalactosaminiltransferasa tenga al menos 55% de identidad de secuencia, preferiblemente al menos 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o preferiblemente 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 48, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70 y 74. Estas formas de realización son particularmente preferidas cuando R¹⁴ es -OH.

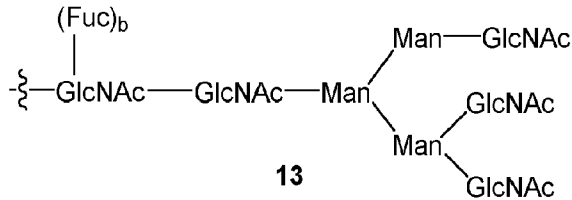
Glucoproteína

- 45 La glucoproteína que se modifica en el proceso de acuerdo con la invención comprende un glucano, comprendiendo dicho glucano un resto de GlcNAc terminal, es decir, un resto de GlcNAc que está presente en el extremo no reductor del glucano. Dicho glucano comprende uno o más restos de sacáridos, y puede ser lineal o ramificado. El glucano que comprende un resto de GlcNAc terminal está de acuerdo con la fórmula (1) o (2):

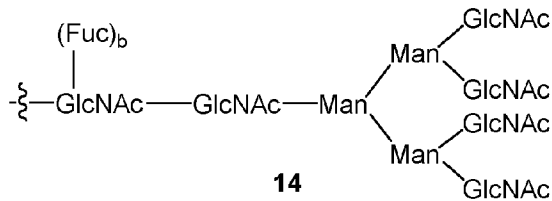




12



13



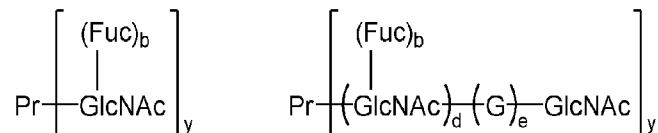
14

donde b es 0 o 1.

- 5 En una forma de realización preferida del proceso según la invención, el glucano que comprende un resto de GlcNAc terminal es un glucano según la fórmula (1), (9), (10), (11), (12), (13) o (14), más preferiblemente un glucano unido a N según la fórmula (1), (9), (10), (11), (12), (13) o (14). En una forma de realización preferida adicional, el glucano que comprende un resto de GlcNAc terminal es un glucano según la fórmula (1), (9), (10) o (11), más preferiblemente un glucano unido a N según la fórmula (1), (9), (10) o (11). Más preferiblemente, el glucano que comprende un resto de GlcNAc terminal es un glucano según la fórmula (1) o (10), más preferiblemente un glucano unido a N según la fórmula (1).
- 10

La glucoproteína que comprende un glucano que comprende un resto GlcNAc terminal es preferiblemente de acuerdo con la fórmula (7), (8) u (8b):

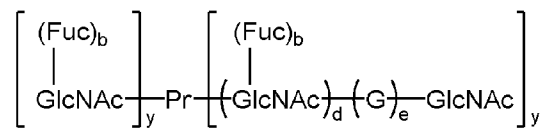
donde:



15

7

8



8b

b, d, e y G, y sus formas de realización preferidas, son como se ha definido anteriormente;

y es independientemente un número entero en el intervalo de 1 a 24; y

Pr es una proteína.

- 20 La glucoproteína que se modifica en el proceso según la invención comprende uno o más glucanos que comprenden un resto de GlcNAc terminal ("y" es 1 a 24). Preferiblemente, "y" es un número entero en el intervalo de 1 a 12, más preferiblemente un número entero en el intervalo de 1 a 10. Más preferiblemente, "y" es 1, 2, 3, 4, 5, 6, 7 u 8, y aún más preferiblemente "y" es 1, 2, 3, 4, 5 o 6. Incluso más preferiblemente, "y" es 1, 2, 3 o 4. Cuando la glucoproteína que se va a modificar comprende más de un glucano ("y" es 2 o más), los glucanos pueden diferir el uno del otro.
- 25 Como se ha descrito anteriormente, la glucoproteína puede comprender además uno o más glucanos que no tienen un resto de GlcNAc terminal.

- 5 Cuando la glucoproteína que se modifica en el proceso según la invención es según la fórmula (7), (8) u (8b), también se prefiere que el glucano que comprende un resto de GlcNAc terminal sea un glucano, preferiblemente un glucano unido a N, según la fórmula (1), (9), (10), (11), (12), (13) o (14) como se ha descrito anteriormente, más preferiblemente un glucano, preferiblemente un glucano unido a N según la fórmula (1), (9), (10) o (11) y aún más preferiblemente de acuerdo con la fórmula (1) o (10). Más preferiblemente, el glucano que comprende un resto de GlcNAc terminal es un glucano unido a N según la fórmula (1).
- 10 En una forma de realización preferida del proceso de acuerdo con la invención, la glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal es un anticuerpo, más preferiblemente un anticuerpo de acuerdo con la fórmula (7), (8) u (8b), donde la proteína (Pr) es un anticuerpo (Ab). Además, cuando la glucoproteína que se modifica es un anticuerpo y el anticuerpo comprende más de un glucano ("y" es 2 o más), los glucanos pueden diferir entre sí. El anticuerpo puede comprender adicionalmente uno o más glucanos que no tienen un resto de GlcNAc terminal. Además, cuando la glucoproteína que se modifica es un anticuerpo, se prefiere que el glucano que comprende un resto de GlcNAc terminal sea un glucano de acuerdo con la fórmula (1), (9), (10), (11), (12), (13) o (14) como se ha definido anteriormente, más preferiblemente de acuerdo con la fórmula (1), (9), (10) o (11), incluso más preferiblemente de acuerdo con la fórmula (1) o (10). En esta forma de realización, se prefiere además que el glucano que comprende un resto de GlcNAc terminal sea un glucano unido a N según la fórmula (1), (9), (10), (11), (12), (13) o (14), más preferiblemente un glucano unido a N según la fórmula (1), (9), (10) o (11), y de la manera más preferible un glucano unido a N según la fórmula (1) o (10).
- 15 Cuando la glucoproteína que se modifica es un anticuerpo, se prefiere que "y" sea 1, 2, 3, 4, 5, 6, 7 u 8, más preferiblemente que "y" sea 1, 2, 4, 6 u 8, incluso más preferiblemente que "y" sea 1, 2 o 4, y más preferiblemente que "y" sea 1 o 2.
- 20 Como se ha definido anteriormente, dicho anticuerpo puede ser un anticuerpo completo, pero también un fragmento de anticuerpo. Cuando el anticuerpo es un anticuerpo completo, dicho anticuerpo comprende preferiblemente uno o más, más preferiblemente uno, de glucano GlcNAc no reductor terminal en cada cadena pesada. De este modo, dicho anticuerpo completo preferiblemente comprende 2 o más, preferiblemente 2, 4, 6 u 8 de dichos glucanos, más preferiblemente 2 o 4, y de la manera más preferible 2 glucanos. En otras palabras, cuando dicho anticuerpo es un anticuerpo completo, "y" es preferiblemente 2, 4, 6 u 8, más preferiblemente "y" es 2 o 4, y más preferiblemente "y" es 2. Cuando el anticuerpo es un fragmento de anticuerpo, se prefiere que "y" sea 1, 2, 3 o 4, y más preferiblemente que "y" sea 1 o 2.
- 25 En una forma de realización preferida, dicho anticuerpo es un anticuerpo monoclonal (mAb). Preferiblemente, dicho anticuerpo se selecciona del grupo que consiste en anticuerpos IgA, IgD, IgE, IgG e IgM. Más preferiblemente, dicho anticuerpo es un anticuerpo IgG1, IgG2, IgG3 o IgG4, y de la manera más preferible dicho anticuerpo es un anticuerpo IgG1.
- 30 En el proceso según la invención, se puede usar una mezcla de glucoproteína que comprende glucanos fucosilados así como no fucosilados como la glucoproteína de partida. Dicha mezcla puede comprender, por ejemplo, glucoproteínas que comprenden uno o más glucanos fucosilados (b es 1) (1) y/o (2) y/o uno o más glucanos no fucosilados (b es 0) (1) y/o (2). La eliminación de fucosa de un glucano fucosilado antes del proceso de acuerdo con la invención, por lo tanto, no es necesaria, sino opcional.
- 35 Una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal también se denomina en la presente "proteína GlcNAc terminal no reductora", y un glucano que comprende un resto de GlcNAc terminal también se denomina en la presente "glucano GlcNAc no reductor terminal". Debe observarse que la expresión "proteína GlcNAc terminal no reductora" incluye una proteína de fórmula (7) en la que b es 1, y que la expresión "glucano GlcNAc terminal no reductor" incluye un glucano de fórmula (1) en la que b es 1.
- 40 La proteína GlcNAc no reductora terminal puede comprender uno o más glucanos de GlcNAc no reductores lineales y/o uno o más ramificados. Un glucano se une a la proteína por el C1 del resto de azúcar central de glucano, y dicho resto de azúcar central preferiblemente es un resto de GlcNAc central. En consecuencia, cuando el glucano GlcNAc no reductor terminal unido a la proteína es un glucano según la fórmula (2), se prefiere que d sea 1. Más preferiblemente, cuando el glucano es según la fórmula (2), d es 1 y e es 1.
- 45 En una forma de realización preferida, el C1 del resto de azúcar central del glucano GlcNAc no reductor terminal se une a la proteína a través de un enlace N-glucosídico a un átomo de nitrógeno en un residuo de aminoácido en dicha proteína, más preferiblemente al átomo de nitrógeno de la cadena lateral de un aminoácido asparagina (Asn) o arginina (Arg). Sin embargo, el C1 del resto de azúcar central del glucano GlcNAc no reductor también se puede unir a la proteína a través de un enlace O-glucosídico a un átomo de oxígeno en un residuo de aminoácido en dicha proteína, más preferiblemente a un átomo de oxígeno en la cadena lateral de un aminoácido serina (Ser) o treonina (Thr).
- 50 En esta forma de realización, se prefiere que el resto de azúcar central de dicho glucano sea un resto de GlcNAc o un resto de GalNAc, preferiblemente un resto de GlcNAc. El C1 del resto del azúcar central del glucano GlcNAc no reductor también se puede unir a la proteína a través de un enlace C-glucosídico a un átomo de carbono en la proteína, por ejemplo a triptófano (Trp). Como se ha descrito anteriormente, una glucoproteína puede comprender más de un glucano, y puede comprender una combinación de glucanos unidos a N, unidos a O y/o unidos a C.
- 55

El glucano GlcNAc no reductor terminal puede estar presente en un sitio de glucosilación nativo de una proteína, pero también se puede introducir en un sitio diferente de una proteína.

Cuando la glucoproteína es un anticuerpo, se prefiere que el glucano que comprende un resto de GlcNAc terminal se una al sitio de N-glucosilación conservado en el fragmento Fc en asparagina en la región 290-305, típicamente en N297.

En la Figura 1 se muestran varios ejemplos de una proteína GlcNAc terminal no reductora que puede modificarse en el proceso según la invención. La Figura 1 (A) muestra una glucoproteína que comprende un único resto de GlcNAc, opcionalmente fucosilado. Este glucano de GlcNAc puede estar unido, por ejemplo, a la proteína a través de un enlace N-glucosídico o un enlace O-glucosídico. La glucoproteína de la Figura 1(A) puede obtenerse, por ejemplo, mediante expresión regular seguida de recorte con una endoglucosidasa o una combinación de endoglucosidasas. La Figura 1(B) muestra una glucoproteína que comprende un oligosacárido glucano ramificado en el que una de las ramificaciones comprende un resto de GlcNAc terminal (este glucano también se denomina GnM). El resto de GlcNAc central se puede fucosilar opcionalmente. La glucoproteína de la Figura 1(B) puede obtenerse, por ejemplo, mediante la expresión de una glucoproteína en un sistema de mamífero en presencia de swainsonina o mediante la expresión en un organismo hospedador manipulado genéticamente, por ejemplo Lec1 CHO o Pichia. La Figura 1(C) muestra un anticuerpo que comprende un único resto de GlcNAc opcionalmente fucosilado. Este glucógeno de GlcNAc se une preferiblemente al anticuerpo a través de un enlace N-glucosídico. La glucoproteína de la Figura 1(C) puede obtenerse, por ejemplo, mediante expresión regular seguida de recorte con una endoglucosidasa o una combinación de endoglucosidasas. La Figura 1(D) muestra un anticuerpo que comprende un oligosacárido glucano ramificado, en el que el resto de GlcNAc central está opcionalmente fucosilado y en el que todas las ramificaciones comprenden un resto de GlcNAc terminal. La glucoproteína de la Figura 1(D) puede obtenerse, por ejemplo, recortando la mezcla regular de glucoformas de anticuerpos (G0, G1, G2, G0F, G1F y G2F) bajo la acción combinada de sialidasa y galactosidasa.

En la Figura 2 se muestra una forma de realización del proceso para la modificación de una glucoproteína, donde la glucoproteína es un anticuerpo. En esta forma de realización, un derivado de azúcar Su(A) se transfiere de Su(A)-Nuc a un resto de GlcNAc terminal de un glucano anticuerpo, bajo la acción de una glucosiltransferasa que es o deriva de una β -(1,4)-N-acetilgalactosaminiltransferasa, para formar un anticuerpo modificado.

Como se ha descrito anteriormente, el proceso para la modificación de una glucoproteína puede comprender además la etapa de proporcionar una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal, y por lo tanto también se describe un proceso para la modificación de una glucoproteína que comprende las etapas de:

(1) proporcionar una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal, donde el glucano que comprende un resto de GlcNAc terminal es según la fórmula (1) o (2) como se ha definido anteriormente; y

(2) poner en contacto dicha glucoproteína con un nucleótido derivado de azúcar Su(A)-Nuc, en presencia de, más en particular bajo la acción de, una glucosiltransferasa, donde la glucosiltransferasa, es o se deriva de una β -(1,4)-N-acetilgalactosaminiltransferasa, y donde Su(A)-Nuc es según la fórmula (3) como se ha definido anteriormente.

Cuando, por ejemplo, la glucoproteína que se modifica en el proceso de acuerdo con la invención comprende un glucano según la fórmula (1), en la etapa (1) del proceso, la glucoproteína que se va a modificar se puede proporcionar mediante un proceso que comprende la etapa de eliminar una glucoproteína que comprende un oligosacárido glucano mediante la acción de una enzima adecuada, preferiblemente una endoglucosidasa.

En un gran número de glucanos, un segundo residuo de GlcNAc se une al residuo de GlcNAc que está directamente unido a la glucoproteína, como también se ve en la Figura 1(B) y (C). Un glucano en el que un segundo resto de GlcNAc está unido al resto de GlcNAc que está directamente unido a la glucoproteína se puede recortar para obtener una glucoproteína que comprende un glucano según la fórmula (1). El recorte se produce entre dichos dos restos de GlcNAc.

Una "enzima adecuada" se define como una enzima para la cual el glucano que se va a recortar es un sustrato. El tipo preferido de enzima que se utilizará en la etapa (1) depende del glucano o glucanos específicos que se recorte(n). Preferiblemente, la enzima en la etapa (1) de esta forma de realización particular del proceso se selecciona de entre el grupo de las endoglucosidasas.

Las endoglucosidasas son capaces de escindir enlaces glucosídicos internos en estructuras de glucanos, lo que proporciona un beneficio para la remodelación y los esfuerzos sintéticos. Por ejemplo, pueden emplearse endoglucosidasas para la homogeneización fácil de poblaciones heterogéneas de glucanos, cuando se escinden en sitios predecibles dentro de las regiones conservadas de glucanos. Una de las clases más importantes de endoglucosidasas a este respecto comprende las endo- β -N-acetilglucosaminidasas (EC 3.2.1.96, comúnmente conocidas como Endos y ENGAsas), una clase de enzimas hidrolíticas que eliminan los N-glucanos de las glucoproteínas al hidrolizar el enlace β -1,4-glucosídico en el núcleo de N,N'-diacetilquitobiosa (descritas por Wong et

- al. Chem. Rev. 2011, 111, 4259), dejando un residuo de GlcNAc ligado a N de un solo núcleo. Las endo- β -N-acetilglucosaminidasas se encuentran ampliamente distribuidas en la naturaleza con variantes quimioenzimáticas comunes incluyendo Endo D, que es específica para pauci-manosa; Endo A y Endo H, que son específicas para alta manosa; subtipos de Endo F, que van desde alta manosa hasta complejo biantenarico; y Endo M, que puede escindir la mayoría de las estructuras de N-glucanos (alta manosa / de tipo complejo / de tipo híbrido), excepto glucanos fucosilados, y la actividad hidrolítica para los oligosacáridos de tipo alta manosa es significativamente mayor que la de los oligosacáridos de tipo complejo y de tipo híbrido. Estas ENGAsas muestran especificidad para la estructura de N-glucano distal y no para la proteína que la tiene, lo que las hace útiles para escindir la mayoría de los glucanos unidos a N de glucoproteínas en condiciones nativas.
- 10 Las endoglucosidasas F1, F2 y F3 son las más adecuadas para la desglucosilación de proteínas nativas. Las especificidades de unión de endo F1, F2 y F3 sugieren una estrategia general para la desglucosilación de proteínas que puede eliminar todas las clases de oligosacáridos unidos a N sin desnaturalizar la proteína. Las estructuras biantenarias y triantenarias pueden eliminarse inmediatamente mediante las endoglucosidasas F2 y F3, respectivamente. La oligomanosa y las estructuras híbridas pueden ser eliminadas por Endo F1.
- 15 La Endo F3 es singular en que su escisión es sensible al estado del enlace peptídico del oligosacárido, así como al estado de fucosilación del núcleo. La endoglucosidasa F3 escinde los oligosacáridos complejos biantenarios y triantenarios unidos a asparagina. Escindirán estructuras biantenarias y triantenarias no fucosiladas a un ritmo lento, pero solo si están unidas a péptidos. Las estructuras biantenarias fucosiladas centrales son sustratos eficaces para Endo F3, con una actividad hasta 400 veces mayor. No hay actividad en las moléculas de oligomanosa e híbridas.
- 20 Véase, por ejemplo, Tarentino et al. Glycobiology 1995, 5, 599.
- La Endo S es una endoglucosidasa secretada por *Streptococcus pyogenes*, y también pertenece a la familia de glucósido hidrolasas 18, según se describe por Collin et al. (EMBO J. 2001, 20, 3046)) A diferencia de las ENGAsas mencionadas anteriormente, la Endo S tiene una especificidad más definida y es específica para escindir solo el N-glucano conservado en el dominio Fc de las IgG humanas (no se ha identificado hasta la fecha ningún otro sustrato), lo que sugiere que una interacción proteína-proteína entre la enzima y la IgG proporciona esta especificidad.
- 25 La Endo S49, también conocida como Endo S2, se describe en WO 2013/037824 (Genovis AB). La Endo S49 se aísla de *Streptococcus pyogenes* NZ131 y es una homóloga de Endo S. La Endo S49 tiene una actividad de endoglucosidasa específica en la IgG nativa y escinde una mayor variedad de glucanos Fc que la Endo S.
- 30 En una forma de realización preferida, la enzima en la etapa (1) de esta forma de realización es una endo- β -N-acetilglucosaminidasa. En una forma de realización preferida adicional, la endo- β -N-acetilglucosaminidasa se selecciona del grupo que consiste en Endo S, Endo S49, Endo F1, Endo F2, Endo F3, Endo H, Endo M y Endo A, o una combinación de las mismas.
- 35 Cuando el glucano que se va a recortar es una estructura diantenaria de tipo complejo, la endo- β -N-acetilglucosaminidasa se selecciona preferiblemente del grupo que consiste en Endo S, Endo S49, Endo F1, Endo F2 y Endo F3, o una combinación de las mismas.
- 40 Cuando la glucoproteína es un anticuerpo y el oligosacárido que se va a recortar es una estructura diantenaria de tipo complejo (es decir, según la Figura 1 (C)), y está presente en el sitio de N-glucosilación conservado de la IgG en N297, la endo- β -N-acetilglucosaminidasa se selecciona preferiblemente del grupo que consiste en Endo S, Endo S49, Endo F1, Endo F2 y Endo F3, o una combinación de las mismas, más preferiblemente del grupo que consiste en Endo S y Endo S49, o una combinación de las mismas.
- 45 Cuando la glucoproteína es un anticuerpo y el glucano que se va a recortar es una estructura diantenaria de tipo complejo, y no está presente en el sitio de N-glucosilación conservado de la IgG en N297, la endo- β -N-acetilglucosaminidasa se selecciona preferiblemente del grupo que consiste en Endo F1, Endo F2 y Endo F3, o una combinación de las mismas.
- 50 Cuando el glucano que se va a recortar es una alta manosa, la endo- β -N-acetilglucosaminidasa se selecciona preferiblemente del grupo que consiste en Endo H, Endo M, Endo A y Endo F1.
- Por lo tanto, cuando la glucoproteína que se modifica en el proceso según la invención comprende un glucano según la fórmula (1), en la etapa (1) del proceso la glucoproteína que se modifica se proporciona preferiblemente mediante un proceso que comprende la etapa de recortar un glucano de una glucoproteína que comprende un oligosacárido glucano por la acción de una endo- β -N-acetilglucosaminidasa, con el fin de proporcionar una glucoproteína que comprende un glucano según la fórmula (1).
- 55 En una forma de realización preferida adicional, la endo- β -N-acetilglucosaminidasa se selecciona del grupo que consiste en Endo S, Endo S 49, Endo F1, Endo F2, Endo F3, Endo H, Endo M y Endo A, y cualquier combinación de las mismas. Más preferiblemente, la endo- β -N-acetilglucosaminidasa se selecciona del grupo que consiste en Endo S, Endo S 49, Endo H, Endo F1, Endo F2 y Endo F3, y cualquier combinación de las mismas. Incluso más preferiblemente, la endo- β -N-acetilglucosaminidasa es Endo S o Endo S49. Más preferiblemente, la endo- β -N-

acetilglucosaminidasa es una combinación de Endo H y Endo S.

El proceso para proporcionar una glucoproteína que comprende un glucano según la fórmula (1) mediante el tratamiento de una mezcla de glucoformas G0, G1, G2, G0F, G1F y G2F con una endoglucosidasa se muestra en la Figura 4. La Figura 4 muestra que el tratamiento de una glucoproteína, en este caso un anticuerpo, que comprende de N-glucoproteína híbrida en presencia de swainsonina, como por ejemplo se describe en Kanda et al., Glycobiology 2006, 17, 104 y, si fuera necesario, seguido de tratamiento con sialidasa / galactosidasa. Un enfoque alternativo incluye la modificación genética de un organismo hospedador. Por ejemplo, Lec1 CHO es una línea celular CHO con inactivación de un gen, en concreto que carece del gen para la expresión de Mns-II. Como consecuencia, la biosíntesis del N-glucano se detiene inevitablemente en la fase GnM₅ del glucano, que puede aislarse en forma pura del sobrenadante. Un enfoque más extenso implica la modificación genética de organismos hospedadores que normalmente no están programados para producir N-glucanos híbridos o complejos, como células de levadura o de insecto. Sin embargo, se ha demostrado ampliamente que estas células hospedadoras no mamíferas (por ejemplo, Glycoswitch™) también se pueden emplear para la expresión selectiva de una única glucoforma de una N-glucoproteína particular, incluidos los glucanos de tipo GnM₅ y de tipo M₅.

Cuando, por ejemplo, la glucoproteína que se modifica en el proceso según la invención comprende un glucano según la fórmula (9), la glucoproteína que comprende dicho glucano, también denominada "GnM₅", se puede proporcionar de varias maneras. En esta forma de realización, se prefiere que la glucoproteína sea proporcionada por una expresión de N-glucoproteína híbrida en presencia de swainsonina, como por ejemplo se describe en Kanda et al., Glycobiology 2006, 17, 104 y, si fuera necesario, seguido de tratamiento con sialidasa / galactosidasa. Un enfoque alternativo incluye la modificación genética de un organismo hospedador. Por ejemplo, Lec1 CHO es una línea celular CHO con inactivación de un gen, en concreto que carece del gen para la expresión de Mns-II. Como consecuencia, la biosíntesis del N-glucano se detiene inevitablemente en la fase GnM₅ del glucano, que puede aislarse en forma pura del sobrenadante. Un enfoque más extenso implica la modificación genética de organismos hospedadores que normalmente no están programados para producir N-glucanos híbridos o complejos, como células de levadura o de insecto. Sin embargo, se ha demostrado ampliamente que estas células hospedadoras no mamíferas (por ejemplo, Glycoswitch™) también se pueden emplear para la expresión selectiva de una única glucoforma de una N-glucoproteína particular, incluidos los glucanos de tipo GnM₅ y de tipo M₅.

Por lo tanto, cuando la glucoproteína que se modifica en el proceso según la invención comprende un glucano según la fórmula (9), en la etapa (1) del proceso la glucoproteína que comprende un glucano opcionalmente fucosilado de fórmula (9) se proporciona preferiblemente mediante un proceso que comprende la expresión de la glucoproteína en un organismo hospedador, en presencia de swainsonina. Preferiblemente, dicho organismo hospedador es una línea celular de mamífero, por ejemplo HEK293 o NS0 o una línea celular CHO. Las glucoproteínas resultantes se pueden obtener como una mezcla de proteínas que comprende un glucano de fórmula (9) (también conocido como GnM₅), un glucano denominado GalGnM₅, un glucano sialilado denominado SiaGalGnM₅ y/o una mezcla de los mismos. El resto de ácido siálico y/o galactosa no reductor, si está presente, puede eliminarse procesando la glucoproteína con sialidasa (eliminación del resto de ácido siálico) y/o β -galactosidasa (eliminación del resto de galactosa), mediante lo cual se obtiene una glucoproteína que comprende un glucano de fórmula (9). Preferiblemente, el tratamiento con sialidasa y β -galactosidasa se produce en una única etapa en (1b). En esta forma de realización, se prefiere adicionalmente que en la etapa (1) del proceso la glucoproteína que se modifica se proporcione mediante un proceso que comprenda las etapas de:

(1a) expresión de una glucoproteína en un organismo hospedador en presencia de swainsonina; y

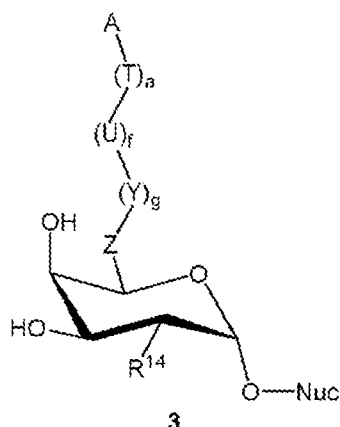
(1b) tratamiento de la glucoproteína obtenida con sialidasa y/o β -galactosidasa con el fin de obtener una glucoproteína que comprenda un glucano de fórmula (9).

Cuando la glucoproteína que se modifica en el proceso según la invención comprende un glucano según la fórmula (10), en la etapa (1) del proceso, la glucoproteína que se modifica puede proporcionarse, por ejemplo, mediante un proceso que comprende un tratamiento de una mezcla de glucoformas G0, G1, G2, G0F, G1F y G2F de la glucoproteína con sialidasa y galactosidasa. En la Figura 3 se muestran las glucoformas G0, G1, G2, G0F, G1F y G2F de un anticuerpo que comprende un glucano biantenarico.

La Figura 4 muestra un proceso para proporcionar una glucoproteína, en este caso un anticuerpo, que comprende un glucano según la fórmula (10) mediante el tratamiento de una mezcla de glucoformas G0, G1, G2, G0F, G1F y G2F con sialidasa y galactosidasa, seguido de transferencia del resto de azúcar de un nucleótido derivado de azúcar Su(A)-UDP en donde A es un grupo azida, por ejemplo 6-azida-GalNAc-UDP, bajo la acción de una glucosiltransferasa que es, o deriva de, una β -(1,4)-GalNAcT, que proporciona un anticuerpo modificado según la fórmula (33).

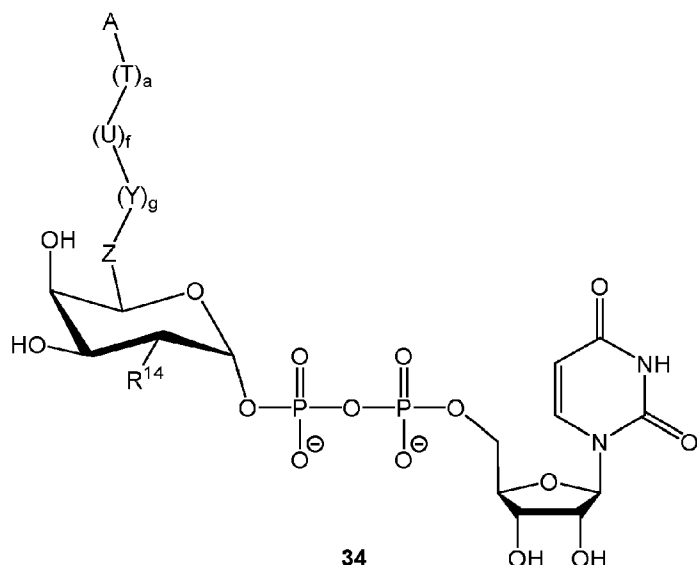
Nucleótido derivado de azúcar Su(A)-Nuc

En el proceso para la modificación de una glucoproteína según la invención, una glucoproteína que comprende un glucano según la fórmula (1) o (2) se pone en contacto, bajo la acción de la glucosiltransferasa, con un nucleótido derivado de azúcar Su(A)-Nuc. El nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3):



donde Nuc, R¹⁴, a, f, g, U, T, A, Z e Y son como se ha definido anteriormente.

Nuc se define aquí como un nucleótido. Nuc se selecciona preferiblemente del grupo que consiste en un nucleósido monofosfato y un nucleósido difosfato, más preferiblemente del grupo que consiste en uridina difosfato (UDP), guanosina difosfato (GDP), timidina difosfato (TDP), citidina difosfato (CDP) y citidina monofosfato (CMP), más preferiblemente del grupo que consiste en uridina difosfato (UDP), guanosina difosfato (GDP) y citidina difosfato (CDP). Más preferiblemente, Nuc es uridina difosfato (UDP). Por lo tanto, en una forma de realización preferida del proceso según la invención, Su(A)-Nuc (3) es Su(A)-UDP (34):



donde R¹⁴, a, f, g, U, T, A, Z e Y son como se ha definido anteriormente.

En una forma de realización, A es un grupo azida -N₃.

En otra forma de realización, A es un grupo ceto -C(O)R³, donde R³ es un grupo alquilo C₁ - C₂₄ opcionalmente sustituido, preferiblemente un grupo alquilo C₁ - C₁₂ opcionalmente sustituido, y más preferiblemente un grupo alquilo C₁ - C₆ opcionalmente sustituido. Aún más preferiblemente, R³ es metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo o t-butilo y, más preferiblemente, R³ es metilo.

En otra forma de realización, A es un grupo alquínilo. En otras palabras, A es un resto funcional que comprende un enlace C≡C, preferiblemente un grupo (hetero)cicloalquínilo o un resto de -(CH₂)_iC≡C-R⁴. En una forma de realización, el grupo alquínilo es un grupo (hetero)cicloalquínilo, preferiblemente un grupo (hetero)ciclooctínilo. En una forma de realización preferida, el grupo alquínilo es -(CH₂)_iC≡C-R⁴, en donde i es 0 - 10 y R⁴ es hidrógeno o un grupo alquilo C₁ - C₂₄ opcionalmente sustituido, preferiblemente hidrógeno o un grupo alquilo C₁ - C₁₂ opcionalmente sustituido, y más preferiblemente hidrógeno o un grupo alquilo C₁ - C₆ opcionalmente sustituido. Aún más preferiblemente, R⁴ es hidrógeno, metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo o t-butilo y más preferiblemente R⁴ es hidrógeno o metilo. Preferiblemente, i es 0, 1, 2, 3, 4, 5 o 6, más preferiblemente i es 0, 1, 2, 3 o 4, incluso más preferiblemente i es 0, 1 o 2, aún más preferiblemente i es 0 o 1 y más preferiblemente es 1. Más preferiblemente, R⁴ es hidrógeno, metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo o t-butilo, e i es 0, 1 o 2. Aún más preferiblemente, R⁴ es hidrógeno o metilo, e i es 0, 1 o 2. En esta forma de realización, se prefiere adicionalmente que el grupo alquínilo sea un grupo alquínilo

terminal, por ejemplo que R^4 sea más preferiblemente hidrógeno. En una forma de realización particularmente preferida, el grupo alquinilo es $-\text{CH}_2-\text{C}\equiv\text{CH}$ o $-\text{C}\equiv\text{CH}$, de la forma más preferible $-\text{CH}_2-\text{C}\equiv\text{CH}$.

En otra forma de realización, A es un grupo tiol $-\text{SH}$.

- 5 En otra forma de realización, A es un precursor de un grupo tiol $-\text{SC}(\text{O})\text{R}^8$, donde R^8 es un grupo alquilo o grupo fenilo $\text{C}_1 - \text{C}_{24}$ opcionalmente sustituido. Preferiblemente, R^8 es un grupo alquilo o grupo fenilo $\text{C}_1 - \text{C}_{12}$ opcionalmente sustituido, más preferiblemente R^8 es un grupo alquilo o grupo fenilo $\text{C}_1 - \text{C}_6$ opcionalmente sustituido, e incluso más preferiblemente R^8 es metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo, t-butilo o fenilo. Aún más preferiblemente R^8 es metilo o fenilo, y más preferiblemente, R^8 es metilo. En el proceso según la invención para la modificación de una glucoproteína, se puede usar un nucleótido derivado de azúcar en el que A es un precursor de un grupo tiol. Durante el proceso, el precursor de tiol se convierte en un grupo tiol.

- 10 En otra forma de realización, A es $-\text{SC}(\text{V})\text{OR}^8$, donde V es O o S, y R^8 es un grupo alquilo o grupo fenilo $\text{C}_1 - \text{C}_{24}$ opcionalmente sustituido. En una forma de realización preferida, A es $-\text{SC}(\text{O})\text{OR}^8$. En otra forma de realización preferida, A es $-\text{SC}(\text{S})\text{OR}^8$. Cuando A es $-\text{SC}(\text{O})\text{OR}^8$ y cuando A es $-\text{SC}(\text{S})\text{OR}^8$, R^8 es preferiblemente un grupo alquilo o grupo fenilo $\text{C}_1 - \text{C}_{12}$ opcionalmente sustituido, más preferiblemente R^8 es un grupo alquilo o grupo fenilo $\text{C}_1 - \text{C}_6$ opcionalmente sustituido, e incluso más preferiblemente R^8 es metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo, t-butilo o fenilo. Aún más preferiblemente R^8 es metilo o fenilo, y más preferiblemente, R^8 es metilo.

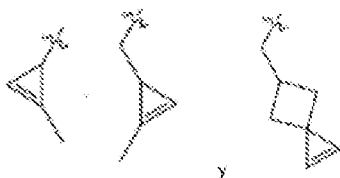
- 15 En otra forma de realización, A es un halógeno X. X se selecciona del grupo que consiste en F, Cl, Br y I, preferiblemente del grupo que consiste en Cl, Br y I, más preferiblemente del grupo que consiste en Cl y Br. Más preferiblemente, X es Cl.

- 20 En otra forma de realización, A es un grupo sulfonilo $-\text{OS}(\text{O})_2\text{R}^5$, donde R^5 se selecciona del grupo que consiste en grupos alquilo $\text{C}_1 - \text{C}_{24}$, grupos arilo $\text{C}_6 - \text{C}_{24}$, grupos alquilarilo $\text{C}_7 - \text{C}_{24}$ y grupos arilalquilo $\text{C}_7 - \text{C}_{24}$, donde los grupos alquilo, grupos arilo, grupos alquilarilo y grupos arilalquilo están opcionalmente sustituidos. Preferiblemente, R^5 es un grupo alquilo $\text{C}_1 - \text{C}_{12}$, grupo arilo $\text{C}_6 - \text{C}_{12}$, grupo alquilarilo $\text{C}_7 - \text{C}_{12}$ o un grupo arilalquilo $\text{C}_7 - \text{C}_{12}$. Más preferiblemente R^5 se selecciona del grupo que consiste en $-\text{CH}_3$, $-\text{C}_2\text{H}_5$, un grupo alquilo C_3 lineal o ramificado, un grupo alquilo C_4 lineal o ramificado, un grupo arilo $\text{C}_6 - \text{C}_{10}$ y un grupo alquilarilo C_7 . Aún más preferiblemente, R^5 es un grupo metilo, un grupo etilo, un grupo fenilo o un grupo p-tolilo. Más preferiblemente, el grupo sulfonilo es un grupo mesilato, $-\text{OS}(\text{O})_2\text{CH}_3$, un grupo bencenosulfonato ($-\text{OS}(\text{O})_2(\text{C}_6\text{H}_5)$) o un grupo tosilato ($-\text{OS}(\text{O})_2\text{C}_6\text{H}_4\text{CH}_3$).

- 25 En otra forma de realización, A es R^{12} , donde R^{12} se selecciona del grupo que consiste en grupos alqueno $\text{C}_2 - \text{C}_{24}$, grupos cicloalqueno $\text{C}_3 - \text{C}_5$ y grupos alcadieno $\text{C}_4 - \text{C}_8$ terminales opcionalmente sustituidos.

- 30 El término "grupo alqueno terminal" en este documento se refiere a un grupo alqueno en el que el enlace doble carbono-carbono está situado en un extremo del grupo alqueno. Cuando R^{12} es un grupo alqueno $\text{C}_2 - \text{C}_{24}$ terminal opcionalmente sustituido, el grupo alqueno $\text{C}_2 - \text{C}_{24}$ terminal termina preferiblemente con un resto $\text{C}=\text{CH}_2$, más preferiblemente un resto $\text{C}(\text{H})=\text{CH}_2$. Preferiblemente R^{12} es un grupo alqueno $\text{C}_2 - \text{C}_{12}$ terminal opcionalmente sustituido, y más preferiblemente un grupo alqueno $\text{C}_2 - \text{C}_6$ terminal opcionalmente sustituido. Más preferiblemente, el grupo alqueno terminal es un grupo alqueno lineal, preferiblemente un grupo alqueno lineal no sustituido. Aún más preferiblemente, R^{12} se selecciona del grupo que consiste en $-\text{C}(\text{H})=\text{CH}_2$, $-\text{CH}_2-\text{C}(\text{H})=\text{CH}_2$, $-\text{CH}_2-\text{CH}_2-\text{C}(\text{H})=\text{CH}_2$, $-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{C}(\text{H})=\text{CH}_2$ y $-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{C}(\text{H})=\text{CH}_2$. Sin embargo, aún más preferiblemente R^{12} se selecciona del grupo que consiste en $-\text{C}(\text{H})=\text{CH}_2$, $-\text{CH}_2-\text{C}(\text{H})=\text{CH}_2$ y $-\text{CH}_2-\text{CH}_2-\text{C}(\text{H})=\text{CH}_2$. Sin embargo, aún más preferiblemente R^{12} es $-\text{C}(\text{H})=\text{CH}_2$ o $-\text{CH}_2-\text{C}(\text{H})=\text{CH}_2$, y más preferiblemente, R^{12} es $-\text{C}(\text{H})=\text{CH}_2$.

- 40 Cuando R^{12} es un grupo cicloalqueno $\text{C}_3 - \text{C}_5$, opcionalmente sustituido, R^{12} preferiblemente comprende un grupo ciclopropeno. Más preferiblemente, el grupo cicloalqueno $\text{C}_3 - \text{C}_5$ (opcionalmente sustituido) se selecciona del grupo que consiste en:



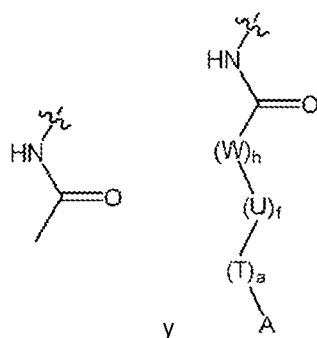
- 45 Cuando R^{12} es un grupo alcadieno $\text{C}_4 - \text{C}_8$ opcionalmente sustituido, se prefiere que el grupo alcadieno $\text{C}_4 - \text{C}_8$ termine con un resto $\text{C}=\text{CH}_2$, más preferiblemente un resto $\text{C}=\text{C}(\text{H})-\text{C}(\text{H})=\text{CH}_2$. Preferiblemente, el grupo alcadieno $\text{C}_4 - \text{C}_8$ se selecciona del grupo que consiste en $\text{C}(\text{H})=\text{C}(\text{H})-\text{C}(\text{H})=\text{CH}_2$, $\text{CH}_2-\text{C}=\text{C}(\text{H})-\text{C}(\text{H})=\text{CH}_2$ y $\text{CH}_2-\text{CH}_2-\text{C}=\text{C}(\text{H})-\text{C}(\text{H})=\text{CH}_2$, más preferiblemente de $\text{C}(\text{H})=\text{C}(\text{H})-\text{C}(\text{H})=\text{CH}_2$ y $\text{CH}_2-\text{C}=\text{C}(\text{H})-\text{C}(\text{H})=\text{CH}_2$. Cuando R^{12} es un grupo alcadieno $\text{C}_4 - \text{C}_8$ opcionalmente sustituido, más preferiblemente R^{12} es $\text{C}(\text{H})=\text{C}(\text{H})-\text{C}(\text{H})=\text{CH}_2$.

En otra forma de realización, A es R^{13} , donde R^{13} es un grupo alenilo $C_3 - C_{24}$ terminal opcionalmente sustituido. El término "grupo alenilo terminal" en el presente documento se refiere a un grupo alenilo en el que el resto $C=C=C$ está situado en un extremo del grupo alenilo. El grupo alenilo $C_3 - C_{24}$ terminal termina preferiblemente con el resto $-C(H)=C=CH_2$. Preferiblemente R^{13} es un grupo alenilo $C_3 - C_{12}$ terminal opcionalmente sustituido, y más preferiblemente un grupo alenilo $C_3 - C_6$ terminal opcionalmente sustituido. Más preferiblemente, el grupo alenilo terminal es un grupo alenilo lineal, preferiblemente un grupo alenilo lineal no sustituido. Aún más preferiblemente R^{13} se selecciona del grupo que consiste en $-C(H)=C=CH_2$, $-CH_2-C(H)=C=CH_2$, $-CH_2-CH_2-C(H)=C=CH_2$ y $-CH_2-CH_2-CH_2-C(H)=C=CH_2$. Sin embargo, aún más preferiblemente, R^{13} se selecciona del grupo que consiste en $-C(H)=C=CH_2$ y $-CH_2-C(H)=C=CH_2$. Más preferiblemente, R^{13} es $-C(H)=C=CH_2$. Cuando A es R^{13} , se prefiere particularmente que en Su(A)-Nuc (3), tanto U como T estén ausentes, es decir, se prefiere particularmente que a sea 0 y f sea 0.

En otra forma de realización, A es $N(R^{17})_2$, donde R^{17} se selecciona independientemente del grupo que consiste en H, grupos alquilo $C_1 - C_{12}$. Los grupos alquilo preferidos en el contexto de R^{17} son grupos alquilo $C_1 - C_6$, más preferiblemente grupos alquilo $C_1 - C_4$. Preferiblemente, al menos uno de R^{17} es H y A es NHR^{17} , más preferiblemente los dos R^{17} son H y A es NH_2 . Cuando A es $N(R^{17})_2$, se prefiere que en Su(A)-Nuc (3) Y esté ausente, es decir, que g sea 0, más preferiblemente que tanto U como T también estén ausentes, es decir, se prefiere particularmente que g sea 0, a sea 0 y f sea 0.

En una forma de realización preferida del proceso según la invención, A en Su(A)-Nuc (3), y formas de realización preferidas de (3) como se describe con más detalle a continuación, se selecciona del grupo que consiste en $-N_3$, $-C(O)R^3$, $-SH$, $-(CH_2)C\equiv CR^4$ y R^{12} , en donde i, R^3 , R^4 , R^{12} y las formas de realización preferidas de los mismos son como se ha definido anteriormente. Más preferiblemente, A se selecciona del grupo que consiste en $-N_3$, $-C(O)CH_3$, $-SH$, $-CH=CH_2$ y $-CH_2C=CH$. Más preferiblemente, A es N_3 .

En el presente documento se divulga el nucleótido derivado de azúcar Su(A)-Nuc (3) y las formas de realización preferidas del mismo, tales como, por ejemplo (34), R^{14} se seleccionan del grupo que consiste en:

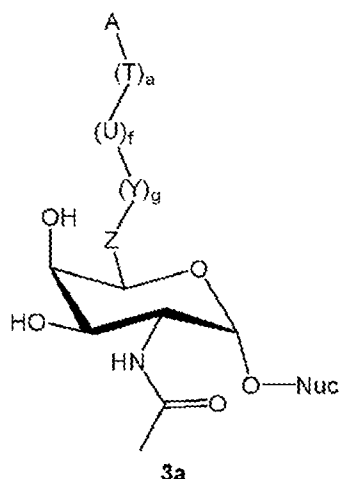


donde W, h, a, f, T, A y U son como se ha definido anteriormente.

Más preferiblemente, R^{14} es $-NHAc$.

Cuando R^{14} es $-NHC(O)-(W)h-(U)f-(T)a-A$, W, h, a, f, T, A y U son como se ha definido anteriormente. Las formas de realización preferidas de W, h, a, f, T y U se describen con más detalle a continuación. Las formas de realización preferidas de A son como se ha descrito con más detalle anteriormente.

En una realización preferida del proceso según la invención, R^{14} en Su(A)-Nuc (3) es $-NHC(O)CH_3$. En esta realización, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3a):



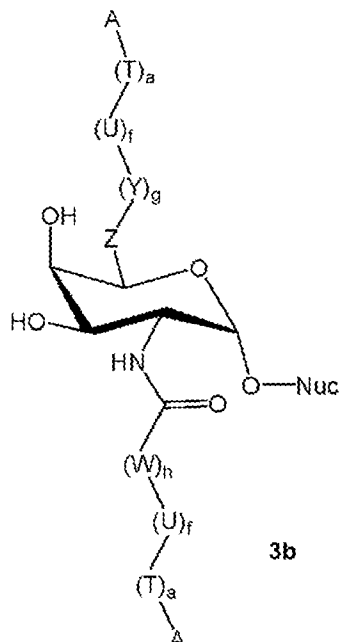
donde Nuc, Z, Y, U, T, A, g, f y a son como se ha definido anteriormente.

También cuando el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3a) o formas de realización preferidas de la misma, se prefiere que Nuc sea UDP.

- 5 Además, en Su(A)-Nuc (3a), A también se selecciona preferiblemente del grupo que consiste en $-N_3$, $-C(O)R^3$, $-SH$, $-(CH_2)_iC\equiv CR^4$ y R^{12} , donde i, R^3 , R^4 , R^{12} , y las formas de realización preferidas de los mismos, son como se ha definido anteriormente. Más preferiblemente, A se selecciona del grupo que consiste en $-N_3$, $-C(O)CH_3$, $-SH$, $-CH=CH_2$ y $-CH_2C\equiv CH$. Más preferiblemente A es N_3 .

- 10 En una forma de realización preferida particular, en Su(A)-Nuc (3a), Nuc es UDP y A se selecciona del grupo que consiste en $-N_3$, $-C(O)R^3$, $-SH$, $-(CH_2)_iC=CR^4$ y R^{12} , donde i, R^3 , R^4 , R^{12} , y las formas de realización preferidas de los mismos, son como se ha definido anteriormente. Incluso más preferiblemente, Nuc es UDP y A se selecciona del grupo que consiste en $-N_3$, $-C(O)CH_3$, $-SH$, $-CH=CH_2$ y $-CH_2C\equiv CH$. Más preferiblemente Nuc es UDP y A es N_3 .

En otra forma de realización preferida del proceso según la invención, R^{14} es $-NHC(O)-(W)_h-(U)_f-(T)_a-A$. En esta forma de realización, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3b):



donde Nuc, Z, Y, U, T, A, W, h, g, f y a son como se ha definido anteriormente.

En el nucleótido derivado de azúcar Su(A)-Nuc (3b), A, T, U, a y f se seleccionan independientemente. En otras palabras, A, T, U, a y f en el sustituyente en C2 de (3b) pueden diferir de A, T, U, a y f en el sustituyente en C6 de (3b).

Además, cuando Su(A)-Nuc es según la fórmula (3b), o sus formas de realización preferidas, se prefiere que Nuc sea UDP.

Además, en Su(A)-Nuc (3b) A también se selecciona preferiblemente del grupo que consiste en $-N_3$, $-C(O)R^3$, $-SH$, $-(CH_2)_iC\equiv CR^4$ y R^{12} , en donde i , R^3 , R^4 , R^{12} y las formas de realización preferidas de los mismos son como se ha definido anteriormente. Más preferiblemente, A se selecciona del grupo que consiste en $-N_3$, $-C(O)CH_3$, $-SH$, $-CH=CH_2$ y $-CH_2C\equiv CH$. Más preferiblemente, A es N_3 .

5 En una forma de realización preferida particular, en Su(A)-Nuc (3b), Nuc es UDP y A se selecciona del grupo que consiste en $-N_3$, $-C(O)R^3$, $-SH$, $-(CH_2)_iC\equiv CR^4$ y R^{12} , donde i , R^3 , R^4 , R^{12} y las formas de realización preferidas de los mismos, son como se ha definido anteriormente. Incluso más preferiblemente, Nuc es UDP y A se selecciona del grupo que consiste en $-N_3$, $-C(O)CH_3$, $-SH$, $-CH=CH_2$ y $-CH_2C\equiv CH$. Más preferiblemente, Nuc es UDP y A es N_3 .

10 En Su(A)-Nuc (3), y sus formas de realización preferidas, tales como, por ejemplo, (34), (3a) o (3b), T es un grupo (hetero)arileno $C_3 - C_{12}$, donde el grupo (hetero)arileno está opcionalmente sustituido. En una forma de realización preferida, T está ausente (a es 0). En otra forma de realización preferida, T está presente (a es 1). Cuando a es 1, el grupo (hetero)arileno T en (3) se sustituye por A, donde A es como se ha definido anteriormente.

15 El grupo (hetero)arileno T está opcionalmente sustituido adicionalmente con uno o más sustituyentes R^2 , donde R^2 se selecciona independientemente del grupo que consiste en halógeno ($-F$, $-Cl$, $-Br$, $-I$, preferiblemente $-F$, $-Cl$, $-Br$), $-CN$, $-NO_2$, $-C(O)R^9$, $-C(O)OR^9$, $-C(O)N(R^{10})_2$, grupos alquilo $C_1 - C_{12}$, grupos alquenilo $C_2 - C_{12}$, grupos alquinilo $C_2 - C_{12}$, grupos cicloalquilo $C_3 - C_{12}$, grupos cicloalquenilo $C_5 - C_{12}$, grupos cicloalquinilo $C_8 - C_{12}$, grupos alcoxi $C_1 - C_{12}$, grupos alquenilo $C_2 - C_{12}$, grupos alquinoxilo $C_2 - C_{12}$, grupos cicloalquinoxilo $C_3 - C_{12}$, grupos amino (preferiblemente $-N(R^{10})_2$), grupos oxo y grupos $-Si(R^7)_3$, donde los grupos alquilo, grupos alquenilo, grupos alquinilo, grupos cicloalquilo, grupos cicloalquenilo, grupos cicloalquinilo, grupos alcoxi, grupos alquenilo, grupos alquinoxilo y grupos cicloalquinoxilo están opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo que consiste en O, N y S, y donde R^7 se selecciona independientemente del grupo que consiste en grupos alquilo $C_1 - C_{12}$, grupos alquenilo $C_2 - C_{12}$, grupos alquinilo $C_2 - C_{12}$, grupos cicloalquilo $C_3 - C_{12}$, grupos alcoxi $C_1 - C_{12}$, grupos alquenilo $C_2 - C_{12}$, $C_2 - C_{12}$ grupos alquinoxilo y grupos cicloalquinoxilo $C_3 - C_{12}$ en los que los grupos alquilo, grupos alquenilo, grupos alquinilo, grupos cicloalquilo, grupos alcoxi, grupos alquenilo, grupos alquinoxilo y grupos cicloalquinoxilo están opcionalmente sustituidos, donde R^9 es un grupo alquilo $C_1 - C_{12}$, y en donde R^{10} se selecciona independientemente de hidrógeno y un grupo alquilo $C_1 - C_{12}$. Preferiblemente, R^9 es un grupo alquilo $C_1 - C_6$, incluso más preferiblemente un grupo alquilo $C_1 - C_4$, y de la manera más preferible un grupo metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo o t-butilo. Preferiblemente, R^{10} es un hidrógeno o un grupo alquilo $C_1 - C_6$, más preferiblemente hidrógeno o un grupo alquilo $C_1 - C_4$, y más preferiblemente R^{10} es hidrógeno, un grupo metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo o t-butilo.

30 Cuando R^2 es un grupo $-Si(R^7)_3$, preferiblemente R^7 es, independientemente, un grupo alquilo $C_1 - C_{12}$, de forma más preferible independientemente un grupo alquilo $C_1 - C_6$, incluso de forma más preferible independientemente un grupo alquilo $C_1 - C_4$, y más preferiblemente R^7 es, independientemente, un grupo metilo, etilo, n-propilo, i-propilo, n-butilo, s-butilo o t-butilo.

35 Preferiblemente, R^2 , cuando está presente, se selecciona independientemente del grupo que consiste en $-F$, $-Cl$, $-Br$, $-I$, $-CN$, $-NO_2$, $-C(O)R^9$, $-C(O)OR^9$, $-C(O)N(R^{10})_2$, grupos alquilo $C_1 - C_{12}$, grupos alcoxi $C_1 - C_{12}$, grupos amino ($-N(R^{10})_2$), grupos oxo y grupos $-Si(R^7)_3$, donde R^7 , R^9 , R^{10} y formas de realización preferidas de R^7 , R^9 , R^{10} son como se ha definido anteriormente.

40 Más preferiblemente, R^2 , cuando está presente, se selecciona independientemente del grupo que consiste en $-F$, $-Cl$, $-Br$, $-CN$, $-NO_2$, $-C(O)R^9$, $-C(O)OR^9$, $-C(O)N(R^{10})_2$, grupos alquilo $C_1 - C_6$, grupos alcoxi $C_1 - C_6$, grupos amino, grupos oxo y grupos $-Si(R^7)_3$, en donde R^7 , R^9 , R^{10} y formas de realización preferidas de R^7 , R^9 , R^{10} son como se ha definido anteriormente.

Aún más preferiblemente, R^2 , cuando está presente, se selecciona independientemente del grupo que consiste en $-F$, $-Cl$, $-Br$, $-CN$, $-NO_2$, $-C(O)R^9$, $-C(O)OR^9$, $-C(O)N(R^{10})_2$, grupos alquilo $C_1 - C_4$ y grupos alcoxi $C_1 - C_4$, donde R^9 y R^{10} y formas de realización preferidas de R^9 y R^{10} son como se ha definido anteriormente.

45 Sin embargo, aún más preferiblemente, R^2 , cuando está presente, se selecciona independientemente del grupo que consiste en $-F$, $-Cl$, $-Br$, $-CN$, $-NO_2$, metilo, metoxi, etilo, etoxi, n-propilo, n-propoxi, i-propilo, i-propoxi, n-butilo, n-butoxi, s-butilo, s-butoxi, t-butilo y t-butoxi. Más preferiblemente, R^2 , cuando está presente, se selecciona independientemente del grupo que consiste en $-F$, $-Cl$, $-Br$, $-CN$, $-NO_2$, metilo y metoxi.

50 En una forma de realización preferida, el grupo (hetero)arileno en (3) no está sustituido. En otra forma de realización preferida, el grupo (hetero)arileno en (3) comprende uno o más sustituyentes R^2 , donde R^2 y formas de realización preferidas de R^2 se han definido antes.

55 El término "grupo(hetero)arileno" en este documento se refiere a grupos arileno así como a grupos heteroarileno. El término "grupo (hetero)arileno" en este documento se refiere a grupos (hetero)arileno monocíclicos así como a grupos (hetero)arileno bicíclicos. El grupo (hetero)arileno en Su(A)-Nuc (3) puede ser cualquier grupo arileno o cualquier grupo heteroarileno.

En una forma de realización preferida del proceso según la invención, el grupo (hetero)arileno T en (3) se selecciona del grupo que consiste en grupos fenileno, grupos naftileno, grupos antracileno, grupos pirrolileno, grupos pirroliomileno, grupos furanileno, grupos tiofenileno (es decir, grupos tiofuranileno), grupos pirazolileno, grupos imidazolileno, grupos pirimidinomileno, grupos imidazoliomileno, grupos isoxazolileno, grupos oxazolileno, grupos oxazoliomileno, grupos isotiazolileno, grupos tiazolileno, grupos 1,2,3-triazolileno, grupos 1,3,4-triazolileno, grupos diazolileno, grupos 1-oxa-2,3-diazolileno, grupos 1-oxa-2,4-diazolileno, grupos 1-oxa-2,5-diazolileno, grupos 1-oxa-3,4-diazolileno, grupos 1-tia-2,3-diazolileno, grupos 1-tia-2,4-diazolileno, 1-grupos tia-2,5-diazolileno, grupos 1-tia-3,4-diazolileno, grupos tetrazolileno, grupos piridinileno, grupos piridazinileno, grupos pirimidinileno, grupos pirazinileno, grupos piradizinileno, grupos piridiniomileno, grupos pirimidiniomileno, grupos benzofuranileno, grupos benzotiofenileno, grupos bencimidazolileno, grupos indazolileno, grupos benzotriazolileno, grupos pirrolo[2,3-b]piridinileno, grupos pirrolo[2,3-c]piridinileno, grupos pirrolo[3,2-c]piridinileno, grupos pirrolo[3,2-b]piridinileno, grupos imidazo[4,5-b]piridinileno, grupos imidazo [4,5-c] piridinileno, grupos pirazolo [4,3-d] piridinileno, grupos pirazolo[4,3-c]piridinileno, grupos pirazolo[3,4-c] piridinileno, grupos pirazolo[3,4-b]piridinileno, grupos isoindolileno, grupos indazolileno, grupos purinileno, grupos indolinileno, grupos imidazo[1,2-a]piridinileno, grupos imidazo[1,5-a]piridinileno, grupos pirazolo[1,5-a]piridinileno, grupos pirrolo[1,2-b]piridazinileno, grupos imidazo[1,2-c]pirimidinileno, grupos quinolinileno, grupos isoquinolinileno, grupos cinolinileno, grupos quinazolinileno, grupos quinoxalinileno, grupos ftalazinileno, grupos 1,6-naftiridinileno, grupos 1,7-naftiridinileno, grupos 1,8-naftiridinileno, grupos 1,5-naftiridinileno, grupos 2,6-naftiridinileno, grupos 2,7-naftiridinileno, grupos pirido[3,2-d]pirimidinileno, grupos pirido[4,3-d]pirimidinileno, grupos pirido[3,4-d]pirimidinileno, grupos pirido[2,3-d]pirimidinileno, grupos pirido[2,3-b]pirazinileno, grupos pirido [3,4-b]pirazinileno, grupos pirimido[5,4-d]pirimidinileno, grupos pirazino[2,3-b]pirazinileno y grupos pirimido[4,5-d] pirimidinileno, todos los grupos opcionalmente sustituidos con uno o más sustituyentes R^2 , en donde R^2 y formas de realización preferidas de R^2 son como se ha definido anteriormente.

En una forma de realización preferida adicional, el grupo (hetero)arileno T se selecciona del grupo que consiste en grupos fenileno, grupos piridinomileno, grupos piridinileno, grupos pirimidinileno, grupos pirimidiniomileno, grupos pirazinileno, grupos piridizinileno, grupos pirrolileno, grupos pirroliomileno, grupos furanileno, grupos tiofenileno (es decir, grupos tiofuranileno), grupos diazolileno, grupos quinolinileno, grupos imidazolileno, grupos pirimidiniomileno, grupos imidazolio, grupos oxazolileno y grupos oxazoliomileno, todos los grupos opcionalmente sustituidos con uno o más sustituyentes R^2 , donde R^2 y formas de realización preferidas de R^2 son como se ha definido anteriormente.

Incluso más preferiblemente, el grupo (hetero)arileno T se selecciona del grupo que consiste en grupos fenileno, grupos piridinileno, grupos piridiniomileno, grupos pirimidinileno, grupos pirimidiniomileno, grupos imidazolileno, grupos pirimidiniomileno, grupos imidazoliomileno, grupos pirrolileno, grupos furanileno y grupos tiofenileno, todos los grupos opcionalmente sustituidos con uno o más sustituyentes R^2 , donde R^2 y formas de realización preferidas de R^2 son como se ha definido anteriormente.

De la forma más preferible, el grupo (hetero)arilo T se selecciona del grupo que consiste en grupos fenileno, grupos imidazolileno, grupos imidazoliomileno, grupos pirimidiniomileno, grupos piridinileno y grupos piridiniomileno, todos los grupos opcionalmente sustituidos con uno o más sustituyentes R^2 , donde R^2 y formas de realización preferidas de R^2 son como se ha definido anteriormente.

En $Su(A)-Nuc$ (3), y sus formas de realización preferidas, tales como, por ejemplo, (34), (3a) o (3b), U puede estar presente (f es 1) o ausente (f es 0). Cuando está presente, U es $[C(R^1)_2]_n$, donde n es un número entero en el intervalo de 1 a 24; o U es $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_q-[C(R^1)_2]_q$, donde o es un número entero en el intervalo de 0 a 12, p y q son independientemente 0, 1 o 2, y R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I, OH y un grupo alquilo $C_1 - C_{24}$ opcionalmente sustituido. Cuando U es $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_q-[C(R^1)_2]_q$, se prefiere que al menos uno de p, o y q no sea 0.

En una forma de realización preferida, U está ausente, es decir, f es 0.

En otra forma de realización preferida, U está presente, es decir, f es 1.

Cuando U es $[C(R^1)_2]_n$, n es un número entero en el intervalo de 1 a 24, preferiblemente un número entero en el intervalo de 1 a 12. Más preferiblemente n es 1, 2, 3, 4, 5, 6, 7 u 8, incluso más preferiblemente n es 1, 2, 3, 4, 5 o 6, aún más preferiblemente n es 1, 2, 3 o 4, aún más preferiblemente n es 1, 2 o 3, y más preferiblemente, n es 1 o 2.

R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I y un grupo alquilo $C_1 - C_{24}$ opcionalmente sustituido, preferiblemente del grupo que consiste en H, F, Cl, Br, I y un grupo alquilo $C_1 - C_{12}$ opcionalmente sustituido, y más preferiblemente del grupo que consiste en H, F, Cl, Br, I y un grupo alquilo $C_1 - C_6$ opcionalmente sustituido. Aún más preferiblemente, R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I, un grupo metilo, un grupo etilo, un grupo n-propilo, un grupo i-propilo, un grupo n-butilo, un grupo s-butilo o un grupo t-butilo. Aún más preferiblemente, R^1 se selecciona independientemente del grupo que consiste en H, F, Cl y metilo y, de la manera más preferible, R^1 se selecciona independientemente del grupo que consiste en H y F.

Cuando U es $[C(R^1)_2]_n$ y n es 1 o 2, los ejemplos preferidos del resto de $-[C(R^1)_2]_n-$ en Su(A)-Nuc incluyen $-(CH_2)-$, $-(CF_2)-$, $-(CCl_2)-$, $-(CBr_2)-$, $-(CMe_2)-$, $-(CH_2CH_2)-$, $-(CH_2CF_2)-$, $-(CH_2CCl_2)-$, $-(CH_2CBr_2)-$, $-(CH_2CMe_2)-$, $-(CF_2CF_2)-$, $-(CCl_2CCl_2)-$, $-(CBr_2CBr_2)-$ y $-(CMe_2CMe_2)-$, más preferiblemente $-(CH_2)-$, $-(CF_2)-$, $-(CH_2CH_2)-$, $-(CH_2CF_2)-$ y $-(CF_2CF_2)-$.

- 5 Cuando U es $[C(R^1)_2]_n$ y n es 3 o más, los ejemplos preferidos del resto de $-[C(R^1)_2]_n-$ en Su(A)-Nuc incluyen $-(C_nH_{2n})-$, $-(C_nF_{2n})-$, $-(C_nCl_{2n})-$, $-(C_nBr_{2n})-$, $-(C_{(n-1)}H_{2(n-1)}CF_2)-$, $-(C_{(n-1)}H_{2(n-1)}CCl_2)-$, $-(C_{(n-1)}H_{2(n-1)}CBr_2)-$ y $-(C_{(n-1)}H_{2(n-1)}CMe_2)-$, por ejemplo $-(C_3H_6)-$, $-(C_3F_6)-$, $-(C_3Cl_6)-$, $-(C_3Br_6)-$, $-(CH_2CH_2CF_2)-$, $-(CH_2CH_2CCl_2)-$, $-(CH_2CH_2CBr_2)-$ y $-(C_4H_8)-$. Los ejemplos más preferidos incluyen $-(C_nH_{2n})-$, $-(C_nF_{2n})-$, p. ej. $-(C_3H_6)-$, $-(C_4H_8)-$, $-(C_3F_6)-$ y $-(C_4F_8)-$.

- 10 Cuando U es $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$, "o" es un número entero en el intervalo de 0 a 12 y "p" y "q" son independientemente 0, 1 o 2. Preferiblemente, "o" es un número entero en el intervalo de 1 a 10, más preferiblemente "o" es 1, 2, 3, 4, 5, 6, 7 u 8, incluso más preferiblemente "o" es 1, 2, 3, 4, 5 o 6, aún más preferiblemente "o" es 1, 2, 3 o 4, aún más preferiblemente "o" es 1, 2 o 3, aún más preferiblemente, "o" es 1 o 2 y más preferiblemente "o" es 1. En otra forma de realización preferida, "o" es 0. Se prefiere particularmente que "o" sea 0, 1 o 2. Cuando "o" es 0, se prefiere adicionalmente que cuando p es 0, q sea 1 o 2, y cuando q es 0, p sea 1 o 2.

- 15 Cuando U es $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$ y "o" y/o "p" y/o "q" son 1 o más, R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I y un grupo alquilo $C_1 - C_{24}$ opcionalmente sustituido, preferiblemente del grupo que consiste en H, F, Cl, Br, I y un grupo alquilo $C_1 - C_{12}$ opcionalmente sustituido, y más preferiblemente del grupo que consiste en H, F, Cl, Br, I y un grupo alquilo $C_1 - C_6$ opcionalmente sustituido. Aún más preferiblemente, R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I, un grupo metilo, un grupo etilo, un grupo n-propilo, un grupo i-propilo, un grupo n-butilo, un grupo s-butilo o un grupo t-butilo. Aún más preferiblemente, R^1 se selecciona independientemente del grupo que consiste en H, F, Cl y metilo. Más preferiblemente, R^1 es H.

- 25 Cuando U es $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$, los ejemplos preferidos del resto de $-[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q-$ en Su(A)-Nuc incluyen CH_2-O- , $-(CH_2)_2-O-$, $-O-CH_2-$, $-O-(CH_2)_2-$, $-CH_2-O-(CH_2CH_2O)_o-$, $-(CH_2)_2-O-(CH_2CH_2O)_o-$, $-O-(CH_2CH_2O)_o-$, $-O-(CH_2CH_2O)_o-CH_2-$, $-O-(CH_2CH_2O)_o-(CH_2)_2-$, $-CH_2-O-(CH_2CH_2O)_o-CH_2-$, $-CH_2-O-(CH_2CH_2O)_o-(CH_2)_2-$, $-(CH_2)_2-O-(CH_2CH_2O)_o-CH_2-$ y $-(CH_2)_2-O-(CH_2CH_2O)_o-(CH_2)_2-$, donde o es 1, 2, 3, 4, 5 o 6, preferiblemente o es 1, 2, 3 o 4, más preferiblemente o es 1 o 2 y más preferiblemente o es 1.

- 30 En el nucleótido derivado de azúcar Su(A)-Nuc (3), y las formas de realización preferidas del mismo, tales como, por ejemplo, (34), (3a) o (3b), se prefiere que a y f no sean los dos 0. En otra forma de realización preferida, a es 0 y f es 1 o a es 1 y f es 0. En estas formas de realización, g puede ser 0 o 1.

- 35 En una forma de realización preferida del proceso según la invención, a es 0, f es 1 y U es $[C(R^1)_2]_n$. En esta forma de realización, se prefiere adicionalmente que a sea 0, f sea 1 y n esté en el intervalo de 1 a 12, más preferiblemente que a sea 0, f sea 1 y n sea 1, 2, 3, 4, 5, 6, 7 u 8, incluso más preferiblemente que a sea 0, f sea 1 y n sea 1, 2, 3, 4, 5 o 6, pero aún más preferiblemente que a sea 0, f sea 1 y n sea 1, 2, 3 o 4, aún más preferiblemente que a sea 0, f sea 1 y n sea 1 o 2, y de la manera más preferible que a sea 0, f sea 1 y n sea 1. Los ejemplos preferidos de $[C(R^1)_2]_n$ se han descrito con más detalle anteriormente.

- 40 En otra forma de realización preferida del proceso según la invención, a es 0, f es 1 y U es $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$. Más preferiblemente, p, o y q no son todos 0, por ejemplo, o es un número entero en el intervalo de 1 a 12 y/o p es 1 o 2 y/o q es 1 o 2. En esta forma de realización, se prefiere además que a sea 0, f sea 1 y que o esté en el intervalo de 1 a 12, más preferiblemente que a sea 0, f sea 1 y que o esté en el intervalo de 1 a 10, incluso más preferiblemente que a sea 0, f sea 1 y que o sea 1, 2, 3, 4, 5, 6, 7 u 8, aún más preferiblemente que a sea 0, f sea 1 y que o sea 1, 2, 3, 4, 5 o 6, y aún más preferiblemente que a sea 0, f sea 1 y que o sea 1, 2, 3 o 4, aún más preferiblemente que a sea 0, f sea 1 y que o sea 1 o 2, y de la forma más preferible que a sea 0, f sea 1 y que o sea 1. También en esta forma de realización, p y q son independientemente 0, 1 o 2. Los ejemplos preferidos de $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$ son como se ha descrito antes con mayor detalle.

- 50 En otra forma de realización preferida, a es 1, f es 1 y U es $[C(R^1)_2]_n$. En esta forma de realización, se prefiere adicionalmente que n esté en el intervalo de 1 a 12, más preferiblemente que n sea 1, 2, 3, 4, 5, 6, 7 u 8, incluso más preferiblemente que n sea 1, 2, 3, 4, 5 o 6, aún más preferiblemente que n sea 1, 2, 3 o 4, aún más preferiblemente que n sea 1 o 2, y más preferiblemente que n sea 1. Los ejemplos preferidos de $[C(R^1)_2]_n$ son como se ha descrito antes con mayor detalle.

- 55 En otra forma de realización preferida, a es 1, f es 1 y U es $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$, o es un número entero en el intervalo de 1 a 12 y p y q son independientemente 0, 1 o 2. En esta forma de realización, se prefiere además que o esté en el intervalo de 1 a 10, más preferiblemente que o sea 1, 2, 3, 4, 5, 6, 7 u 8, incluso más preferiblemente que o sea 1, 2, 3, 4, 5 o 6, aún más preferiblemente que o sea 1, 2, 3 o 4, aún más preferiblemente que o sea 1 o 2, y más preferiblemente que o sea 1. También en esta forma de realización, p y q son independientemente 0, 1 o 2. Los ejemplos preferidos de $[C(R^1)_2]_p-O-[C(R^1)_2C(R^1)_2O]_o-[C(R^1)_2]_q$ son como se ha descrito antes con mayor detalle.

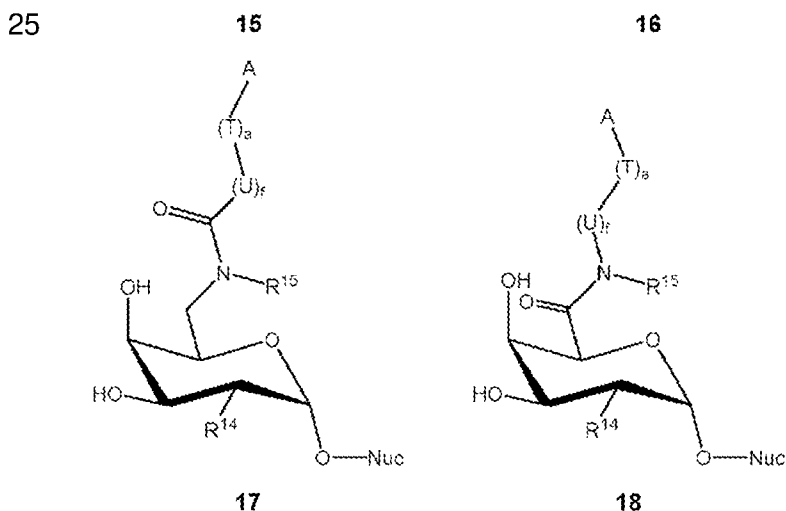
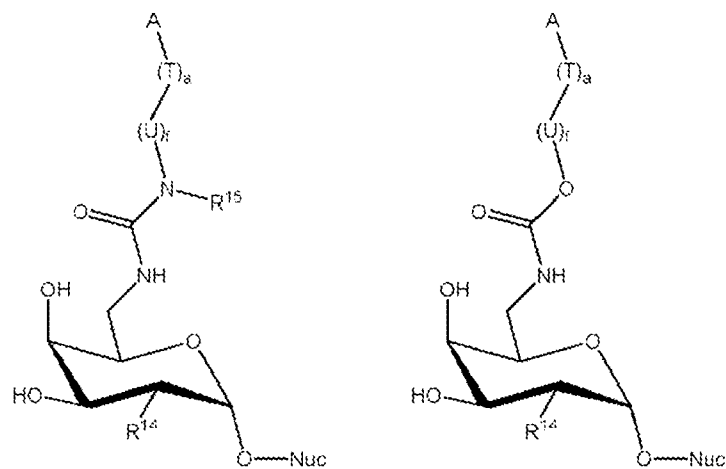
Como se ha definido anteriormente, Z en Su(A)-Nuc (3), y sus formas de realización preferidas, tales como, por ejemplo, (34), (3a) o (3b), es CH₂, CF₂ o C(O); o Z es CHOH con la condición de que g sea 0, f sea 1 y a sea 0 o 1. En una forma de realización preferida, Z se selecciona del grupo que consiste en CH₂, CF₂ y C(O). En otra forma de realización preferida, Z es CHOH con la condición de que g sea 0, f sea 1 y a sea 0 o 1.

- 5 En Su(A)-Nuc (3), y sus formas de realización preferidas, tales como, por ejemplo, (34), (3a) o (3b), Y puede estar ausente (g es 0) o presente (g es 1). Cuando Y está presente, Y se selecciona del grupo que consiste en O, S, N(R¹⁵), N(R¹⁵)C(O), N(R¹⁵)C(O)N(R¹⁵), N(R¹⁵)C(O)O, OC(O)N(R¹⁵)S(O)₂N(R¹⁵) y N(R¹⁵)C(O)N(R¹⁵)S(O)₂O, donde R¹⁵ se selecciona independientemente del grupo que consiste en H, grupos alquilo C₁ - C₁₂ grupos alquilo y (U)_f-(T)_a-A donde f, a, U, T y A son como se ha definido anteriormente. Preferiblemente, Y se selecciona del grupo que consiste en O, S, N(R¹⁵), NHC(O), NHC(O)N(R¹⁵), NHC(O)O, OC(O) NHS(O)₂ NH y NHC(O) NHS(O)₂O, donde R¹⁵ se selecciona independientemente del grupo que consiste en H, grupos alquilo C₁ - C₁₂ y (U)_f-(T)_a-A donde f, a, U, T y A son como se ha definido anteriormente. En estas formas de realización, se prefiere adicionalmente que R¹⁵ se seleccione independientemente del grupo que consiste en H, grupos alquilo C₁ - C₆ y (U)_f-(T)_a-A en donde f, a, U, T y A son como se ha definido anteriormente. Más preferiblemente, R¹⁵ se selecciona independientemente del grupo que consiste en H, metilo, etilo, i-propilo, n-propilo y (U)_f-(T)_a-A donde f, a, U, T y A son como se ha definido anteriormente. Más preferiblemente R¹⁵ se selecciona del grupo que consiste en H y metilo.

- 20 En una forma de realización preferida, Z es CH₂ y g es 1. En esta forma de realización, se prefiere adicionalmente que Y se seleccione del grupo que consiste en O, S, N(R¹⁵), N(R¹⁵)C(O), N(R¹⁵)C(O)N(R¹⁵) y N(R¹⁵)C(O)O, más preferiblemente del grupo que consiste en O, S, N(R¹⁵), NHC(O), NHC(O)N(R¹⁵) y NHC(O)O, donde R¹⁵ y formas de realización preferidas de R¹⁵ son como se ha definido anteriormente.

En otra forma de realización preferida Z es C(O) g es 1. En esta forma de realización, se prefiere adicionalmente que Y sea N(R¹⁵), donde R¹⁵ y formas de realización preferidas de R¹⁵ son como se ha definido anteriormente.

Por lo tanto, en una forma de realización preferida del proceso según la invención, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (15), (16), (17) o (18):



donde Nuc, a, f, R¹⁴, R¹⁵, A, U y T son como se ha definido antes.

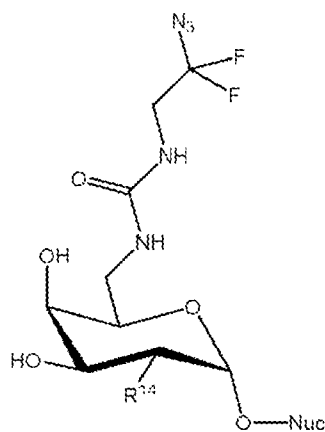
En una forma de realización preferida de (15), (16), (17) y (18), R^{14} es -OH. En otra forma de realización preferida, R^{14} es -N₃. En otra forma de realización preferida, R^{14} es -NHC(O)CH₃. En otra forma de realización preferida R^{14} es -NHC(O)-(W)_h-(U)_f-(T)_a-A, donde W, U, T, A, h, f y a son como se ha definido anteriormente. En estas formas de realización, se prefiere adicionalmente que Nuc sea UDP.

- 5 Las formas de realización preferidas para U, T, a y f en (15), (16), (17) y (18) son como se ha descrito anteriormente. Las formas de realización preferidas para A como se ha definido anteriormente también son válidas para (15), (16), (17) y (18).

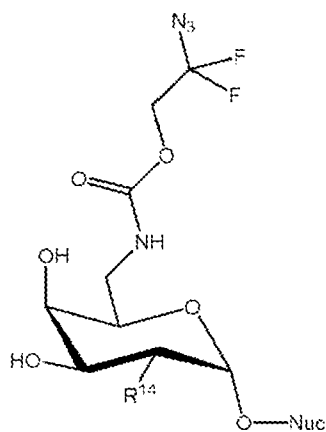
En una forma de realización particularmente preferida de (15), (16), (17) y (18), a es 0, f es 1 y U es -CH₂CF₂. En esta forma de realización, se prefiere adicionalmente que A sea N₃.

- 10 En otra forma de realización particularmente preferida de (15), (16), (17) y (18), a es 1 y T es preferiblemente un grupo fenilo opcionalmente sustituido. Como se ha descrito anteriormente, el grupo fenilo está opcionalmente sustituido con R², y preferiblemente R² se selecciona del grupo que consiste en H, F, Cl y Br, más preferiblemente del grupo que consiste en H, F y Cl y, de la manera más preferible, del grupo que consiste en H y F. En esta forma de realización, f es 0 o 1, y cuando f es 1, U es preferiblemente -CH₂-. En estas formas de realización, se prefiere adicionalmente que A sea N₃. Preferiblemente, A está presente en la posición para del grupo fenilo opcionalmente sustituido.

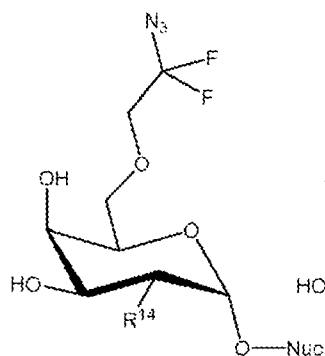
En una forma de realización preferida del proceso según la invención, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (19), (20), (21), (22), (23), (24), (25), (26), (65) o (66), preferiblemente según la fórmula (19), (20), (21), (22), (23), (24), (25) o (26):



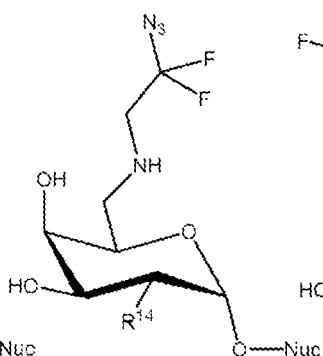
19



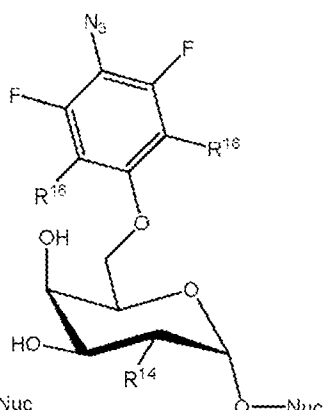
20



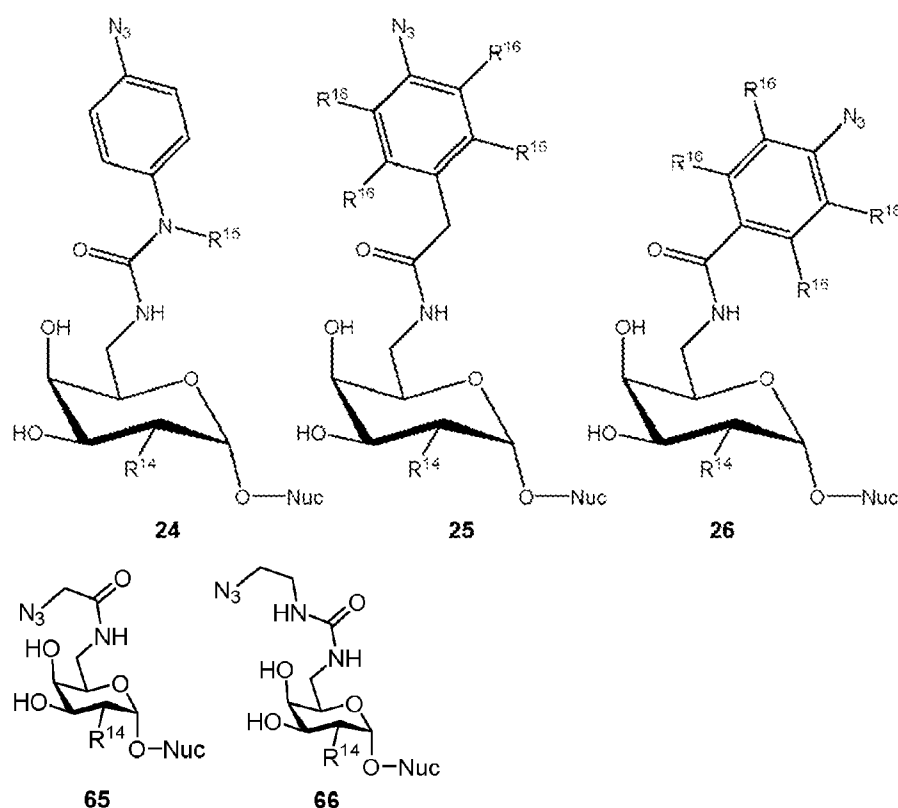
21



22



23

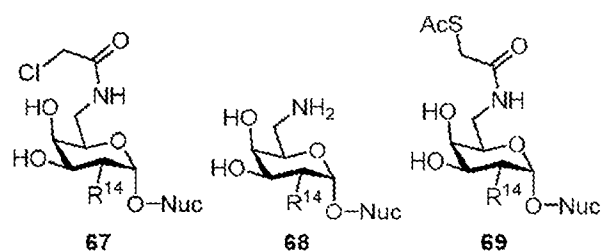


donde:

R^{14} y R^{15} son como se ha definido anteriormente; y

- 5 R^{16} se selecciona independientemente del grupo que consiste en H y F.

En una forma de realización preferida del proceso según la invención, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (67), (68) o (69):

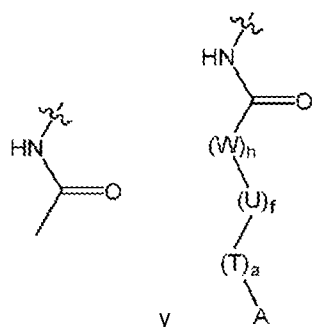


donde R^{14} es como se ha definido anteriormente.

- 10 En una forma de realización preferida adicional, R^{15} se selecciona del grupo que consiste en H y grupos alquilo C_1 - C_{12} , preferiblemente del grupo que consiste en H y grupos alquilo C_1 - C_6 , más preferiblemente del grupo que consiste en H, metilo, etilo, i-propilo, n-propilo, n-butilo, s-butilo y t-butilo y, de la manera más preferible, del grupo que consiste en H y metilo. En otra forma de realización preferida adicional, R^{15} es (U)_f-(T)_a-A, donde f, a, U, T y A, y sus formas de realización preferidas son como se ha definido anteriormente. Cuando R^{15} es (U)_f-(T)_a-A, se prefiere que el grupo (U)_f-(T)_a de R^{15} corresponda a un grupo de (U)_f-(T)_a-A originario del resto de Z-(Y)_g-(U)_f-(T)_a-A en Su(A)-Nuc (3). Por ejemplo, cuando R^{15} en (24) es (U)_f-(T)_a-A se prefiere que R^{15} sea -(C₆H₄(N₃)), con N₃ en la posición para de fenilo. En estas formas de realización, se prefiere además que Nuc sea UDP.
- 15

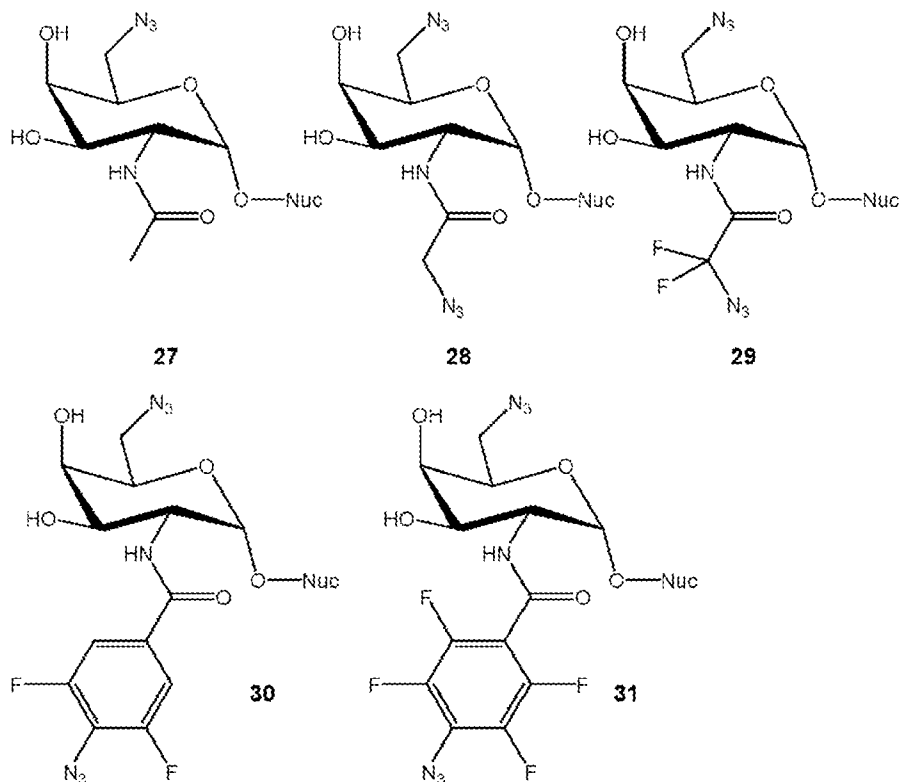
- 20 En una realización preferida de (19), (20), (21), (22), (23), (24), (25), (26), (65), (66), (67), (68) y (69) y sus formas de realización preferidas como se ha descrito anteriormente, R^{14} es -NHC(O)CH₃. En otra realización preferida, R^{14} es -NHC(O)-(W)_h-(U)_f-(T)_a-A, donde W, U, T, A, h, f y a son como se ha definido anteriormente. También en estas formas de realización se prefiere además que Nuc sea UDP.

En el proceso según la invención, R^{14} se selecciona del grupo que consiste en:



donde a, f, h, T, A, U y W, y sus formas de realización preferidas, son como se ha descrito anteriormente.

- 5 En una realización preferida del proceso según la invención, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (27), (28), (29), (30) o (31):



y

donde Nuc es como se ha definido anteriormente.

- 10 También cuando Su(A)-Nuc es según la fórmula (27), (28), (29), (30) o (31), se prefiere que Nuc sea UDP.

Enzima

- 15 El proceso según la invención comprende la etapa de poner en contacto una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal con un nucleótido derivado de azúcar Su(A)-Nuc en presencia, más particularmente bajo la acción de una glucosiltransferasa, donde la glucosiltransferasa es una β -(1,4)-N-acetilgalactosaminiltransferasa que comprende una secuencia que tiene al menos un 95% de identidad de secuencia con cualquiera de SEQ. ID NO: 3, SEQ. ID NO: 4, SEQ. ID NO: 7 y SEQ. ID NO: 8, para proporcionar una glucoproteína modificada. Una β -(1,4)-N-acetilgalactosaminiltransferasa también se denomina en la presente memoria enzima β -(1,4)-GalNAcT, o β -(1,4)-GalNAcT o GalNAcT.

- 20 Las β -(1,4)-N-acetilgalactosaminiltransferasas (β -(1,4)-GalNAcT) son conocidas en la técnica. Típicamente, una β -(1,4)-GalNAcT es una enzima que cataliza la transferencia de N-acetilgalactosamina (GalNAc) de uridina difosfato-GalNAc (UDP-GalNAc, también denominada GalNAc-UDP) a un resto de GlcNAc terminal de un glucoproteína

glucano, donde el C1 del resto de GalNAc está unido al C4 del resto de GlcNAc mediante un enlace β -1,4-O-glucosídico. Como se ha descrito con más detalle anteriormente, el resto de GlcNAc de un glucano según la fórmula (1), en la que b es 1, es decir, el resto de GlcNAc de un glucano que consiste en una GlcNAc fucosilada, también se considera un resto de GlcNAc terminal.

En el proceso según la invención, la glucosiltransferasa cataliza la transferencia del derivado de azúcar Su(A) de un nucleótido derivado de azúcar Su(A)-Nuc a un resto GlcNAc terminal de una glucoproteína glucano para proporcionar una glucoproteína modificada, donde Su(A) es según la fórmula (6), Su(A)-Nuc es según la fórmula (3), el glucano que comprende un resto GlcNAc terminal es según la fórmula (1) o (2), y la glucoproteína modificada es según la fórmula (4) o (5), como se ha descrito anteriormente. En este proceso, el C1 del resto de Su(A) se une al C4 del resto de GlcNAc a través de un enlace β -1,4-O-glucosídico.

La enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención es o se deriva de una enzima β -(1,4)-GalNAcT de invertebrado, es decir, es o se deriva de una β -(1,4)-GalNAcT procedente de especies de animales invertebrados. La enzima β -(1,4)-GalNAcT es, o se deriva de, una enzima β -(1,4)-GalNAcT procedente de *Ascaris suum* o *Trichoplusia ni*. En una realización preferida adicional, la enzima β -(1,4)-GalNAcT es, o se deriva de, una enzima β -(1,4)-GalNAcT procedente de *Ascaris suum*. En otra realización preferida adicional, la enzima β -(1,4)-GalNAcT es, o se deriva de, una enzima β -(1,4)-GalNAcT procedente de *Trichoplusia ni*. *Ascaris suum* también se denomina en esta memoria As, *Trichoplusia ni* se denomina Tn y *Drosophila melanogaster* se denomina Dm.

Preferiblemente, la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención tiene al menos 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia respecto de una secuencia seleccionada del grupo que consiste en SEQ ID N°: 3 y 4. En otras palabras, preferiblemente la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención tiene al menos un 95%, 96%, 97 %, 98%, 99% o preferiblemente un 100% de identidad de secuencia respecto de una secuencia seleccionada del grupo que consiste en SEQ ID N°: 3 y SEQ ID N°: 4 y, de la manera más preferible, a la SEQ ID N°: 4.

La enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención es o se deriva de cualquiera de las enzimas β -(1,4)-GalNAcT naturales o de tipo salvaje seleccionadas del grupo que consiste en β -(1,4)-GalNAcT de *Ascaris suum*, denominada en este documento AsGalNAcT (SEQ ID N°: 3) y β -(1,4)-GalNAcT de *Trichoplusia ni*, denominada en este documento TnGalNAcT (SEQ ID N°: 4).

En una forma de realización particularmente preferida, la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención es o deriva de β -(1,4)-GalNAcT de *Trichoplusia ni*, denominada en este documento TnGalNAcT (SEQ ID N°: 4).

En otra forma de realización preferida, la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención es una enzima β -(1,4)-GalNAcT que es o deriva de una enzima β -(1,4)-GalNAcT que procede de una especie de invertebrado del filo Nematodo, de la clase Secementea, del orden Ascaridida, de la familia Ascarididae, del género *Ascaris*. Dicha especie de invertebrado es de *Ascaris Sum*. Preferiblemente, la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención tiene al menos 95%, 96%, 97%, 98%, 99% o preferiblemente 100% de identidad de secuencia respecto de una secuencia del grupo que consiste en la SEQ ID N°: 3.

En otra forma de realización preferida, la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención es una enzima β -(1,4)-GalNAcT que es o deriva de una enzima β -(1,4)-GalNAcT que procede de una especie de invertebrado del filo Anthropoda, de la clase Insecta, del orden Lepidópteros, de la familia Noctuidae, del género *Trichoplusia*. Dicha especie de invertebrado es de *Trichoplusia Ni*. *Trichoplusia Ni* también se conoce a veces como *Phytometra brassicae*, *Plusia innata* u oruga de la col. Preferiblemente, la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención tiene al menos 95%, 96%, 97%, 98%, 99% o preferiblemente 100% de identidad de secuencia respecto de una secuencia del grupo que consiste en la SEQ ID N°: 4.

"Que deriva de" una enzima β -(1,4)-GalNAcT se debe entender aquí como una enzima β -(1,4)-GalNAcT que tiene una secuencia de aminoácidos que ha sido alterada a partir de una enzima β -(1,4)-GalNAcT mediante la sustitución, inserción, eliminación o adición de uno o más, preferiblemente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20 o más aminoácidos, respectivamente. Una enzima β -(1,4)-GalNAcT que deriva de una enzima β -(1,4)-GalNAcT también se denomina en el presente documento enzima β -(1,4)-GalNAcT derivada o una enzima β -(1,4)-GalNAcT modificada o una enzima β -(1,4)-GalNAcT mutante o una mutante de β -(1,4)-GalNAcT.

Las enzimas derivadas son conocidas en la técnica e incluyen enzimas que han experimentado una modificación convencional y estándar de la secuencia de aminoácidos, tal como la eliminación del dominio transmembrana, la inclusión de un marcador, tal como un marcador de solubilidad y/o de purificación como se menciona aquí. Tales procedimientos que conducen a una enzima que tiene una secuencia de aminoácidos modificada son ampliamente conocidos en la técnica, y están cubiertos por el proceso según la presente invención.

- En una forma de realización, la enzima derivada, es decir, que tiene menos del 100% de identidad de secuencia con las enzimas β -(1,4)-GalNAcT de origen natural mencionadas aquí, preferiblemente tiene una actividad enzimática que es al menos 10%, 20%, 30%, 40%, 50%, 60%, 70% o preferiblemente al menos 80% o 90% o al menos 100% de la actividad enzimática de la enzima β -(1,4)-GalNAcT de origen natural. En este caso, la actividad se mide convenientemente como la eficacia en la incorporación de un residuo de GalNAc (modificado) al residuo de GlcNAc terminal de una glucoproteína.
- La enzima no es una galactosiltransferasa. En una forma de realización, la enzima no es una enzima categorizada como E.C. 2.4.1.38 o como E.C. 2.4.1.133, preferiblemente no como E.C. 2.4.1.22, como E.C. 2.4.1.38, como E.C. 2.4.1.90 o como E.C. 2.4.1.133.
- En una forma de realización, la enzima es una enzima categorizada como E.C. 2.4.1.41, como E.C. 2.4.1.92, como E.C. 2.4.1.174 o como E.C. 2.4.1.244, preferiblemente como E.C. 2.4.1.92 o como E.C. 2.4.1.244.
- Preferiblemente, dicha enzima β -(1,4)-GalNAcT derivada se modifica añadiendo aminoácidos N- o C- terminales adicionales o restos químicos o eliminando aminoácidos N- o C-terminales para aumentar la estabilidad, la solubilidad, la actividad y/o la facilidad de purificación.
- Preferiblemente, la enzima β -(1,4)-GalNAcT se modifica eliminando el dominio citoplasmático N-terminal y el dominio transmembrana, lo que en este documento se denomina enzima truncada. En la técnica se sabe que la eliminación de estos dominios da como resultado una enzima que muestra una mayor solubilidad en soluciones acuosas.
- Por ejemplo, CeGalNAcT (30-383) debe entenderse aquí como una enzima β -(1,4)-GalNAcT de *Caenorhabditis elegans* truncada que consiste en la secuencia de aminoácidos representada por los aminoácidos en la posición 30-383 de la SEQ ID N°: 2. De manera similar, AsGalNAcT (30-383) debe entenderse aquí como una enzima β -(1,4)-GalNAcT de *Ascaris Sum* truncada que consiste en la secuencia de aminoácidos representada por los aminoácidos en la posición 30-383 de la SEQ ID N°: 3, TnGalNAcT (33-421) debe entenderse aquí como una enzima β -(1,4)-GalNAcT de *Trichoplusia Ni* truncada que consiste en la secuencia de aminoácidos representada por los aminoácidos en la posición 33-421 de la SEQ ID N°: 4, y DmGalNAcT (47-403) debe entenderse aquí como una enzima β -(1,4)-GalNAcT de *Drosophila melanogaster* truncada que consiste en la secuencia de aminoácidos representada por los aminoácidos en la posición 47-403 de la SEQ ID N°: 5.
- Preferiblemente, la enzima β -(1,4)-GalNAcT utilizada en el proceso de la invención tiene al menos 95%, 96%, 97%, 98%, 99% o preferiblemente al menos 100% de identidad de secuencia con cualquiera de las secuencias de SEQ ID N°: 7 o SEQ ID N°: 8, e incluso más preferiblemente la SEQ ID N°: 8.
- Una enzima β -(1,4)-GalNAcT en la que se ha sustituido, añadido o eliminado uno o más aminoácidos se denomina en la presente memoria enzima β -(1,4)-GalNAcT derivada. Preferiblemente, la enzima β -(1,4)-GalNAcT se modifica eliminando el dominio citoplasmático N-terminal y el dominio transmembrana, y sustituyendo uno o más aminoácidos. Una sustitución de uno o más aminoácidos también se denomina aquí mutación. Una enzima que comprende uno o más aminoácidos sustituidos también se denomina enzima mutante.
- Debe observarse que la numeración de la posición de los aminoácidos se basa en la numeración de la posición de aminoácidos en la enzima β -(1,4)-GalNAcT de tipo salvaje. Cuando una enzima β -(1,4)-GalNAcT es, por ejemplo, una enzima truncada, el número utilizado aquí para indicar, por ejemplo, la posición de una sustitución de aminoácido corresponde a la numeración de la posición del aminoácido en la correspondiente enzima de tipo salvaje β -(1,4)-GalNAcT.
- Por ejemplo, en CeGalNAcT de tipo salvaje (SEQ ID N°: 2) está presente una isoleucina (Ile, I) en la posición del aminoácido 257. En CeGalNAcT (I257L), el aminoácido isoleucina en la posición 257 se sustituye por un aminoácido leucina (Leu, L). Como se ha descrito anteriormente, CeGalNAcT (30-383) debe entenderse aquí como una enzima CeGalNAcT truncada que consiste en la secuencia de aminoácidos representada por los aminoácidos en la posición 30-383 de la SEQ ID N°: 2, mientras que la propia CeGalNAcT (30-383) está representada por la SEQ ID N°: 6. En CeGalNAcT (30-383; I257L), el número "257" en I257L indica que es el aminoácido I en la posición 257 en la CeGalNAcT de tipo salvaje correspondiente (es decir, el número 257 de la SEQ ID N°: 2) el que está sustituido con un aminoácido L. El aminoácido isoleucina en la posición 257 de la SEQ ID N°: 2 está representado por el aminoácido isoleucina en la posición 228 de la SEQ ID N°: 6.
- En el proceso según la invención, cuando la glucosiltransferasa deriva de enzima β -(1,4)-GalNAcT de *Trichoplusia Ni* o enzima β -(1,4)-GalNAcT de *Trichoplusia Ni* truncada, se prefiere que la enzima comprenda además una o más mutaciones. Las mutaciones preferidas incluyen la sustitución del triptófano (Trp, también denominado W) en la posición 336 por fenilalanina (Phe, también denominada F), histidina (His, también denominada H) o valina (Val, también denominada V). Por consiguiente, cuando la glucosiltransferasa deriva de TnGalNAcT o TnGalNAcT (33-421), se prefiere que la enzima comprenda una mutación W336F, W336H o W336V. Las mutaciones preferidas de TnGalNAcT o TnGalNAcT (33-421) también incluyen la sustitución del ácido glutámico (Glu, también denominado E) en la posición 339 por alanina (Ala, también denominada A), glicina (Gly, también denominada G), ácido aspártico (Asp, también conocido como D) o serina (Ser, también conocido como S). Por consiguiente, cuando la

- glucosiltransferasa deriva de TnGalNAcT o TnGalNAcT (33-421), se prefiere que la enzima comprenda una mutación E339A, E339G, E339D o E339S. Más preferiblemente, cuando la glucosiltransferasa deriva de TnGalNAcT o TnGalNAcT (33-421), tanto la posición 336 como la posición 339 están mutadas como se ha descrito anteriormente. En consecuencia, cuando la glucosiltransferasa deriva de TnGalNAcT o TnGalNAcT (33-421) se prefiere que la enzima comprenda una mutación W336F, W336H o W336V y una mutación E339A, E339G, E339D o E339S.
- 5 Las mutaciones preferidas de TnGalNAcT o TnGalNAcT (33-421) también incluyen la sustitución de la isoleucina (Ile, también denominada I) en la posición 311 por una tirosina (Tyr, también denominada Y). Por consiguiente, cuando la glucosiltransferasa deriva de TnGalNAcT o TnGalNAcT (33-421), se prefiere que la enzima comprenda una mutación I311Y.
- 10 Cuando la glucosiltransferasa deriva de TnGalNAcT o TnGalNAcT (33-421) y comprende una mutación I311Y, la enzima puede comprender además una mutación en la posición 336 como se ha descrito anteriormente y/o una mutación en la posición 339 como se ha descrito anteriormente. En consecuencia, cuando la glucosiltransferasa deriva de TnGalNAcT o TnGalNAcT (33-421) que comprende una mutación I311Y, la enzima puede comprender además una mutación W336F, W336H o W336V, y/o una mutación E339A, E339G, E339D o E339S.
- 15 En una forma de realización preferida del proceso según la invención, la glucosiltransferasa que es, o deriva de, una enzima β -(1,4)-GalNAcT es una enzima β -(1,4)-GalNAcT de Trichoplusia Ni seleccionada del grupo que consiste en TnGalNAcT (33-421; W336F) (SEQ ID N°: 25), TnGalNAcT (33-421; W336H) (SEQ ID N°: 26), TnGalNAcT (33 - 421; W336V) (SEQ ID N°: 27), TnGalNAcT (33 - 421; E339A) (SEQ ID N°: 28), TnGalNAcT (33 - 421; E339G) (SEQ ID N°: 29), TnGalNAcT (33 - 421; E339D) (SEQ ID N°: 30) y TnGalNAcT (33 - 421; E339S) (SEQ ID N°: 31).
- 20 En otra forma de realización preferida del proceso según la invención, la glucosiltransferasa que es, o deriva de, una enzima β -(1,4)-GalNAcT es una enzima β -(1,4)-GalNAcT de Trichoplusia Ni seleccionada del grupo que consiste en TnGalNAcT (33-421; W336H, E339A) (SEQ ID N°: 32), TnGalNAcT (33-421; W336H, E339D) (SEQ ID N°: 33 y TnGalNAcT (33-421; W336H, E339S) (SEQ ID N°: 34).
- 25 En otra forma de realización preferida del proceso según la invención, la glucosiltransferasa que es, o deriva de, una enzima β -(1,4)-GalNAcT es una enzima β -(1,4)-GalNAcT de Trichoplusia Ni TnGalNAcT (33 - 421; I311Y) (SEQ ID N°: 35).
- 30 En otra forma de realización preferida del proceso según la invención, la glucosiltransferasa que es, o deriva de, una enzima β -(1,4)-GalNAcT es una enzima β -(1,4)-GalNAcT de Trichoplusia Ni seleccionada del grupo que consiste en TnGalNAcT (33-421; I311Y, W336F) (SEQ ID N°: 36), TnGalNAcT (33-421; I311Y, W336H) (SEQ ID N°: 37), TnGalNAcT (33-421; I311Y, W336V) (SEQ ID N°: 38), TnGalNAcT (33-421; I311Y, E339A) (SEQ ID N°: 39), TnGalNAcT (33-421; I311Y, E339G) (SEQ ID NO: 40), TnGalNAcT (33-421; I311Y, E339D) (SEQ ID N°: 41) y TnGalNAcT (33-421; I311Y, E339S) (SEQ ID N°: 42).
- 35 En otra forma de realización preferida del proceso según la invención, la glucosiltransferasa que es, o deriva de, una enzima β -(1,4)-GalNAcT es una enzima β -(1,4)-GalNAcT de Trichoplusia Ni seleccionada del grupo que consiste en TnGalNAcT (33-421; I311Y, W336H, E339A) (SEQ ID N°: 43), TnGalNAcT (33-421; I311Y, W336H, E339D) (SEC. ID NO: 44) y TnGalNAcT (33-421; I311Y, W336H, E339S) (SEQ ID N°: 45).
- 40 En el proceso según la invención, cuando la glucosiltransferasa deriva de la enzima β -(1,4)-GalNAcT de Ascaris Sum o de la enzima β -(1,4)-GalNAcT de Ascaris Sum truncada, se prefiere que la enzima comprenda además una o más mutaciones. Las mutaciones preferidas incluyen la sustitución de triptófano (Trp, también denominado W) en la posición 282 por histidina (His, también denominada H), y/o la sustitución del ácido glutámico (Glu, también denominado E) en la posición 285 por ácido aspártico (Asp, también denominado D), y/o la sustitución de isoleucina (Ile, también denominada I) en la posición 257 por tirosina (Tyr, también denominada Y). En consecuencia, cuando la glucosiltransferasa deriva de AsGalNAcT o AsGalNAcT (30-383), se prefiere que la enzima comprenda una mutación W282H, una mutación E285D y/o una mutación I257Y.
- 45 En una forma de realización preferida del proceso según la invención, la glucosiltransferasa que es o deriva de una enzima β -(1,4)-GalNAcT es una β -(1,4)-GalNAcT de Ascaris Sum seleccionada del grupo que consiste en AsGalNAcT (30-383; W282H) (SEQ ID N°: 46) y AsGalNAcT (30-383; E285D) (SEQ ID N°: 47).
- En otra forma de realización preferida del proceso según la invención, la glucosiltransferasa que es o deriva de una enzima β -(1,4)-GalNAcT es β -(1,4)-GalNAcT de Ascaris Sum AsGalNAcT (30-383; I257Y) (SEQ ID N°: 48).
- 50 En otra forma de realización preferida del proceso según la invención, la glucosiltransferasa que es o deriva de una enzima β -(1,4)-GalNAcT es β -(1,4)-GalNAcT de Ascaris Sum seleccionada del grupo que consiste en AsGalNAcT (30-383; I257Y, W282H) y AsGalNAcT (30-383; I257Y, E285D).
- 55 En una forma de realización preferida del proceso según la invención, la glucosiltransferasa que es o deriva de una enzima β -(1,4)-GalNAcT como se define aquí comprende una secuencia que codifica un marcador para facilitar la purificación. Preferiblemente, dicho marcador se selecciona de, pero no se limita a, el grupo consistente en un

marcador FLAG, marcador poli(His), marcador HA, marcador Myc, marcador SUMO, marcador GST, marcador MBP o marcador CBP, más preferiblemente dicho marcador es un marcador 6xHis. Otros marcadores preferidos para incorporarse a la enzima son marcadores de solubilidad, tales como un marcador AFV, un marcador SlyD, un marcador Tsf, un marcador SUMO, un marcador Bla, un marcador MBP y un marcador GST. En una forma de realización preferida adicional, dicho marcador o marcadores están unidos covalentemente a la enzima β -(1,4)-GalNAcT en el extremo C de la enzima. En otra forma de realización preferida adicional, dicho marcador está unido covalentemente a la enzima β -(1,4)-GalNAcT en el extremo N- terminal de la enzima.

En una forma de realización preferida del proceso según la invención, cuando la enzima β -(1,4)-GalNAcT es, o deriva de, β -(1,4)-GalNAcT de *Trichoplusia Ni*, la enzima β -(1,4)-GalNAcT marcada con His es, o deriva de, His-TnGalNAcT (33-421) (SEQ ID N°: 49).

En otra forma de realización preferida del proceso según la invención, cuando la enzima β -(1,4)-GalNAcT es, o deriva de, β -(1,4)-GalNAcT de *Trichoplusia Ni*, la enzima β -(1,4)-GalNAcT marcada con His es o deriva de His-TnGalNAcT (33-421; W336F) (SEQ ID N°: 50), His -TnGalNAcT (33-421; W336H) (SEQ ID N°: 51), His- TnGalNAcT (33-421; W336V) (SEQ ID N°: 52), His-TnGalNAcT (33-421; 339A) (SEQ ID N°: 53), His-TnGalNAcT (33-421; E339G) (SEQ ID N°: 54), His-TnGalNAcT (33-421; E339D) (SEQ ID N°: 55), His-TnGalNAcT (33-421; E339S) (SEQ ID NO: 56), His-TnGalNAcT (33-421; W336H, E339A) (SEQ ID N°: 57), His-TnGalNAcT (33-421; W336H, E339D) (SEQ ID N°: 58) o His-TnGalNAcT (33 - 421; W336H, E339S) (SEQ ID N°: 59).

En otra forma de realización preferida del proceso según la invención, cuando la enzima β -(1,4)-GalNAcT es, o deriva de, β -(1,4)-GalNAcT de *Trichoplusia Ni*, la enzima β -(1,4)-GalNAcT marcada con His es o deriva de His-TnGalNAcT (33-421; I311Y) (SEQ ID N°: 60).

En otra forma de realización preferida del proceso según la invención, cuando la enzima β -(1,4)-GalNAcT es, o deriva de, β -(1,4)-GalNAcT de *Trichoplusia Ni*, la enzima β -(1,4)-GalNAcT marcada con His es o deriva de His-TnGalNAcT (33-421; I311Y, W336F) (SEQ ID N°: 61), His-TnGalNAcT (33-421; I311Y, W336H) (SEQ ID N°: 62), His-TnGalNAcT (33-421; I311Y, W336V) (SEQ ID N°: 63), His-TnGalNAcT (33-421; I311Y, E339A) (SEQ ID N°: 64), His-TnGalNAcT (33-421; I311Y, E339G) (SEQ ID N°: 65), His-TnGalNAcT (33-421; I311Y, E339D) (SEQ ID N°: 66), His-TnGalNAcT (33-421; I311Y, E339S) (SEQ ID N°: 67), His-TnGalNAcT (33-421; I311Y, W336H, E339A) (SEQ ID N°: 68), His-TnGalNAcT (33-421 ; I311Y, W336H, E339D) (SEQ ID N°: 69) o His-TnGalNAcT (33-421; I311Y, W336H, E339S) (SEQ ID N°: 70).

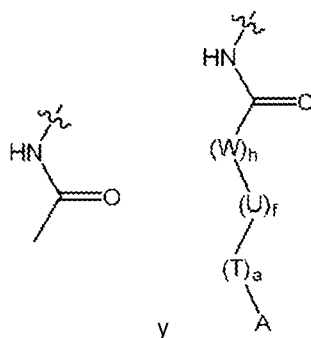
En otra forma de realización preferida del proceso según la invención, cuando la enzima β -(1,4)-GalNAcT es, o deriva de, β -(1,4)-GalNAcT de *Ascaris Sum*, la enzima β -(1,4)-GalNAcT marcada con His es o deriva de His-AsGalNAcT (30-383) (SEQ ID N°: 71).

En otra forma de realización preferida del proceso según la invención, cuando la enzima β -(1,4)-GalNAcT es, o deriva de, β -(1,4)-GalNAcT de *Ascaris Sum*, la enzima β -(1,4)-GalNAcT marcada con His es, o deriva de, His-AsGalNAcT (30-383; W282H) (SEQ ID N°: 72), His -AsGalNAcT (30-383; E285D) (SEQ ID N°: 73) o His- AsGalNAcT (30-383; I257Y) (SEQ ID N°: 74).

En este documento, el término "derivado de" comprende, por ejemplo, enzimas truncadas, enzimas mutantes y enzimas que comprenden un marcador para facilitar la purificación, y estas modificaciones se han descrito con más detalle anteriormente. El término "derivado de" también comprende enzimas que comprenden una combinación de las modificaciones descritas con más detalle anteriormente.

La β -(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso según la invención tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en SEQ ID N° : 3, SEQ ID N°: 4 y SEQ ID N° : 8.

En el proceso según la invención, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3), o formas de realización preferidas de la misma, como se ha descrito con más detalle anteriormente. R¹⁴ se selecciona del grupo que consiste en:



donde W, h, a, f, T, A y U, y las realizaciones preferidas de los mismos, son como se ha definido anteriormente.

Glucosiltransferasas preferidas cuando R^{14} es $-NHC(O)CH_3$

En una realización preferida del proceso según la invención, R^{14} es $-NHC(O)CH_3$. En esta forma de realización, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3a), como se ha definido antes.

5 Cuando Su(A)-Nuc es según la fórmula (3a), o formas de realización preferidas de (3a) como se ha descrito anteriormente, en una forma de realización preferida del proceso, la glucosiltransferasa que es, o se deriva de, β -(1,4)-GalNAcT es, o se deriva de, β -(1,4)-GalNAcT de Ascaris Sum (AsGalNAcT) o β -(1,4)-GalNAcT de Trichoplusia Ni (TnGalNAcT). β -(1,4)-GalNAcTs que son, o se derivan de, (AsGalNAcT) o (TnGalNAcT) se describen con más detalle anteriormente.

10 Cuando R^{14} en el nucleótido derivado de azúcar Su(A)-Nuc es $-NHC(O)CH_3$, se prefiere particularmente que la β -(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso sea, o derive de, una secuencia seleccionada del grupo consistente en la SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 7 y SEQ ID N°: 8. Más preferiblemente, cuando R^{14} es $-NHC(O)CH_3$, la β -(1,4)-N-acetilgalactosaminiltransferasa usada en el proceso es, o se deriva de, una secuencia seleccionada del grupo que consiste en SEQ ID NO: 7 y SEQ ID NO: 8. Lo más preferiblemente la β -(1,4)-N-acetilgalactosaminiltransferasa usada en el proceso es, o se deriva de la SEQ ID NO: 8.

15 En otra forma de realización particularmente preferida, cuando R^{14} en el nucleótido derivado de azúcar Su(A)-Nuc es $-NHC(O)CH_3$, la β -(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 7 y SEQ ID N°: 8. Más preferiblemente, cuando R^{14} es $-NHC(O)CH_3$, la β -(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 7 y SEQ ID N°: 8. De la manera más preferible, la β -(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de la SEQ ID N°: 8.

25 En otra forma de realización particularmente preferida del proceso según la invención donde R^{14} es $-NHC(O)CH_3$, la glucosiltransferasa es, o deriva de, β -(1,4)-GalNAcT de Trichoplusia Ni (TnGalNAcT).

En una forma de realización preferida adicional del proceso donde R^{14} es $-NHC(O)CH_3$, la TnGalNAcT es o deriva de la SEQ ID N°: 4 o SEQ ID N°: 8.

30 En otra forma de realización preferida adicional del proceso en el que R^{14} es $-NHC(O)CH_3$, la TnGalNAcT utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de la SEQ ID N°: 4 o SEQ ID N°: 8.

En otra forma de realización particularmente preferida del proceso según la invención donde R^{14} es $-NHC(O)CH_3$, la glucosiltransferasa es o deriva de β -(1,4)-GalNAcT de Ascaris Sum (AsGalNAcT).

35 En esta forma de realización del proceso donde R^{14} es $-NHC(O)CH_3$, se prefiere además que AsGalNAcT sea o derive de la SEQ ID N°: 3 o SEQ ID N°: 7.

En otra forma de realización preferida adicional del proceso donde R^{14} es $-NHC(O)CH_3$, la AsGalNAcT utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de la SEQ ID N°: 3 o SEQ ID N°: 7.

40 En las formas de realización preferidas descritas aquí del proceso según la invención donde R^{14} es $-NHC(O)CH_3$ y la glucosiltransferasa como se ha descrito anteriormente, se prefiere adicionalmente que el nucleótido derivado de azúcar Su(A)-Nuc sea según la fórmula (15), (16), (17) o (18) como se ha definido anteriormente, donde R^{14} es $-NHC(O)CH_3$; o según la fórmula (19), (20), (21), (22), (23), (24), (25), (26), (65), (66), (67), (68) o (69) como se ha definido anteriormente, donde R^{14} es $-NHC(O)CH_3$; o según la fórmula (27) como se ha definido arriba. En estas formas de realización particularmente preferidas, se prefiere adicionalmente que Nuc sea UDP.

Glucosiltransferasas preferidas cuando R^{14} es $-NHC(O)-(W)_h-(U)_r-(T)_a-A$

En otra realización preferida del proceso según la invención, R^{14} es $-NHC(O)-(W)_h-(U)_r-(T)_a-A$. En esta realización, el nucleótido derivado de azúcar Su(A)-Nuc es según la fórmula (3b), como se ha definido antes.

50 Cuando Su(A)-Nuc es según la fórmula (3b), o realizaciones preferidas de (3b) como se ha descrito anteriormente, en una realización preferida del proceso, la glucosiltransferasa que es o deriva de una β -(1,4)-GalNAcT es o deriva de una β -(1,4)-GalNAcT de tipo salvaje, (AsGalNAcT) de Ascaris Sum o β -(1,4)-GalNAcT de Trichoplusia Ni (TnGalNAcT). Las β -(1,4)-GalNAcT que derivan de (AsGalNAcT) o (TnGalNAcT) se han descrito con más detalle anteriormente.

Cuando R¹⁴ en el nucleótido derivado de azúcar Su(A)-Nuc es -NHC(O)-(W)_h-(U)_f-(T)_a-A, se prefiere particularmente que la β-(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso sea, o derive de, una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 7 y SEQ ID N°: 8, más preferiblemente del grupo que consiste en la SEQ ID N°: 7 y SEQ ID N°: 8. De la manera más preferible, la β-(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso es o deriva de la SEQ ID N°: 8.

En otra realización particularmente preferida, cuando R¹⁴ en el nucleótido derivado de azúcar Su(A)-Nuc es -NHC(O)-(W)_h-(U)_f-(T)_a-A, la β-(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso tiene al menos un 50% de identidad de secuencia, preferiblemente al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 7 y SEQ ID N°: 8. Más preferiblemente, cuando R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A, la β-(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 7 y SEQ ID N°: 8. Lo más preferiblemente, la β-(1,4)-N-acetilgalactosaminiltransferasa utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de la SEQ ID N°: 8.

En otra realización particularmente preferida del proceso de la invención donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A, la glucosiltransferasa es, o deriva de, β-(1,4)-GalNAcT de Trichoplusia Ni (TnGalNAcT).

En una realización preferida adicional del proceso donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A, la TnGalNAcT es o deriva de la SEQ ID N°: 4 o SEQ ID N°: 8.

En otra realización preferida adicional del proceso donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A, la TnGalNAcT utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de la SEQ ID N°: 4 o SEQ ID N°: 8.

En otra realización particularmente preferida del proceso según la invención donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A, la glucosiltransferasa es, o deriva de, β-(1,4)-GalNAcT de Ascaris Sum (AsGalNAcT).

En esta realización del proceso donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A, se prefiere adicionalmente que AsGalNAcT sea o derive de la SEQ ID N°: 3 o SEQ ID N°: 7.

En otra realización preferida adicional del proceso donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A, la AsGalNAcT utilizada en el proceso tiene al menos un 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia respecto de la SEQ ID N°: 3 o SEQ ID N°: 7.

En las realizaciones preferidas descritas aquí del proceso según la invención donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A y la glucosiltransferasa como se ha descrito anteriormente, se prefiere adicionalmente que el nucleótido derivado de azúcar Su(A)-Nuc sea según la fórmula (15), (16), (17) o (18) como se ha definido anteriormente, donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A; o según la fórmula (19) (20), (21), (22), (23), (24), (25), (26), (65), (66), (67), (68) o (69) como se ha definido anteriormente, donde R¹⁴ es -NHC(O)-(W)_h-(U)_f-(T)_a-A; o según la fórmula (28), (29), (30) o (31) como se ha definido anteriormente. En estas realizaciones particularmente preferidas, se prefiere además que Nuc sea UDP.

Enzimas

En este documento también se divulga la enzima como se define en la presente memoria, es decir, glucosiltransferasas que son, o derivan de, β-(1,4)-GalNAcT, en particular las glucosiltransferasas que derivan de β-(1,4)-GalNAcT, es decir, enzimas derivadas de β-(1,4)-GalNAcT. En un aspecto, la enzima deriva de una especie de invertebrado. En un aspecto, las enzimas según este aspecto están en forma aislada. Las enzimas y los aspectos preferidos de las mismas se han definido antes adicionalmente en el contexto del proceso según la invención, que se aplica de igual modo a la propia enzima, según el presente aspecto.

En un aspecto, la enzima según este aspecto de la divulgación deriva de una β-(1,4)-N-acetilgalactosaminiltransferasa, preferiblemente deriva de una β-(1,4)-N-acetilgalactosaminiltransferasa que tiene una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 2-23 y SEQ ID N°: 25-74, más preferiblemente del grupo que consiste en la SEQ ID N°: 10-14 y SEQ ID N°: 25-74, más preferiblemente del grupo que consiste en la SEQ ID N°: 10-13, SEQ ID N°: 25-48, SEQ ID N°: 50-70 y SEQ ID N°: 72-74. Las enzimas según este aspecto típicamente están en forma aislada.

En un aspecto, la enzima según este aspecto de la divulgación se refiere a β-(1,4)-N-acetilgalactosaminiltransferasas que tienen al menos un 40% de identidad de secuencia, preferiblemente al menos un 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 2-23 y SEQ ID N°: 25-74. Las enzimas según este aspecto típicamente están en forma aislada. Preferiblemente, la divulgación se refiere a β-(1,4)-N-acetilgalactosaminiltransferasas que tienen al menos un 40% de identidad de secuencia, preferiblemente

al menos un 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 10-14 y SEQ ID N°: 25-74, de la manera más preferible del grupo que consiste en la SEQ ID N°: 10-13, SEQ ID N°: 25-48, SEQ ID N°: 50-70 y SEQ ID N°: 72-74.

5 En un aspecto preferido, la enzima según este aspecto de la divulgación se refiere a β -(1,4)-N-acetilgalactosaminiltransferasas que tienen al menos un 40% de identidad de secuencia, preferiblemente al menos un 45%, 50%, 55%, 60%, 65 %, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 6-14, es decir, del grupo que consiste en la SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N° : 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13 y SEQ ID N°: 14, preferiblemente del grupo que consiste en la SEQ ID N°: 10-13, es decir, del grupo que consiste en la SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12 y SEQ ID N°: 13.

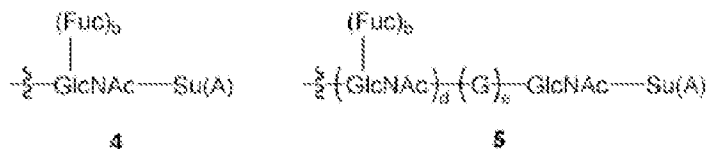
15 En otro aspecto preferido, la enzima según este aspecto de la divulgación se refiere a β -(1,4)-N-acetilgalactosaminiltransferasa que tiene al menos un 40% de identidad de secuencia, preferiblemente al menos un 45%, 50%, 55%, 60%, 65 %, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 25-45 y SEQ ID N°: 50-70, es decir, del grupo que consiste en la SEQ ID N°: 25, SEQ ID N°: 26, SEQ ID N°: 27, SEQ ID N°: 28, SEQ ID N°: 29, SEQ ID N°: 30, SEQ ID N°: 31, SEQ ID N°: 32, SEQ ID N°: 33, SEQ ID N°: 34, SEQ ID N°: 35, SEQ ID N° : 36, SEQ ID N°: 37, SEQ ID N°: 38, SEQ ID N°: 39, SEQ ID N°: 40, SEQ ID N°: 41, SEQ ID N°: 42, SEQ ID N°: 43, SEQ ID N°: 44 , SEQ ID N°: 45, SEQ ID N°: 50, SEQ ID N°: 51, SEQ ID N°: 52, SEQ ID N°: 53, SEQ ID N°: 54, SEQ ID N°: 55, SEQ ID N°: 56, SEQ ID N°: 57, SEQ ID N°: 58, SEQ ID N°: 59, SEQ ID N°: 60, SEQ ID N°: 61, SEQ ID N°: 62, SEQ ID N°: 63, SEQ ID N°: 64, SEQ ID N° : 65, SEQ ID N°: 66, SEQ ID N°: 67, SEQ ID N°: 68, SEQ ID N°: 69 y SEQ ID N°: 70.

25 En otro aspecto preferido, la enzima según este aspecto de la divulgación se refiere a β -(1,4)-N-acetilgalactosaminiltransferasa que tiene al menos un 40% de identidad de secuencia, preferiblemente al menos un 45%, 50%, 55%, 60%, 65 %, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o preferiblemente un 100% de identidad de secuencia, respecto de una secuencia seleccionada del grupo que consiste en la SEQ ID N°: 46-48 y SEQ ID N°: 72-74, es decir, del grupo que consiste en la SEQ ID N°: 46, SEQ ID N°: 47, SEQ ID N°: 48, SEQ ID N°: 72, SEQ ID N°: 73 y SEQ ID N°: 74.

30 La descripción también se refiere al uso de la enzima según la descripción, como se ha definido anteriormente, en un proceso para la modificación de una glucoproteína, preferiblemente en el proceso según la divulgación. El proceso comprende poner en contacto una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal con un nucleótido derivado de azúcar en presencia de la enzima. En un aspecto preferido, la glucoproteína es como se define en la presente memoria, de la manera más preferible una glucoproteína que comprende un gulfano según la fórmula (1) o (2) como se define adicionalmente en este documento. En un aspecto preferido, el nucleótido derivado de azúcar es un nucleótido derivado de azúcar Su(A)-Nuc según la fórmula (3) como también se define en este documento.

Glucoproteína modificada

40 La presente invención se refiere además a una glucoproteína modificada que se puede obtener mediante el proceso según la invención para la modificación de una glucoproteína. Más en particular, la invención se refiere a una glucoproteína que comprende un glucano según la fórmula (4) o (5):



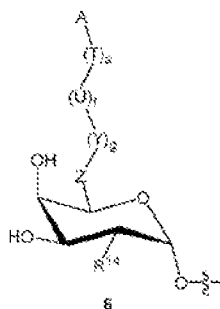
donde:

b es 0 o 1;

45 d es 0 o 1;

e es 0 o 1;

G es un monosacárido, o un oligosacárido lineal o ramificado que comprende de 2 a 20 restos de azúcar; y Su(A) es según la fórmula (6):



donde R^{14} , Z, Y, U, T, A, g, f y a son como se ha definido anteriormente para (3).

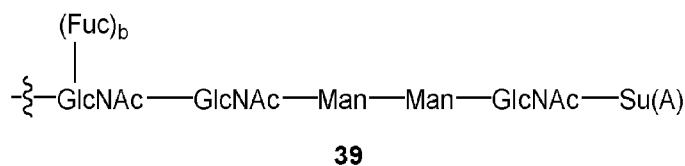
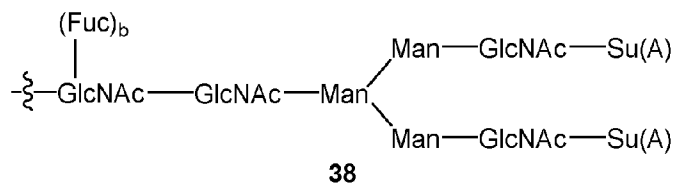
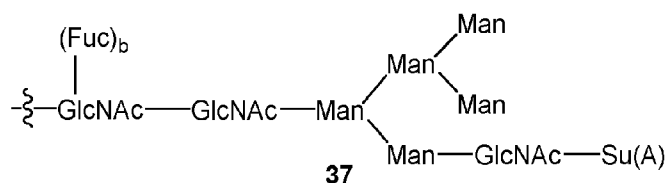
- 5 Las formas de realización preferidas de R^{14} , Z, Y, U, T, A, g, f y a en (6) son como se ha descrito anteriormente con más detalle para (3) y formas de realización preferidas de (3) como, por ejemplo, (3a) o (3b).

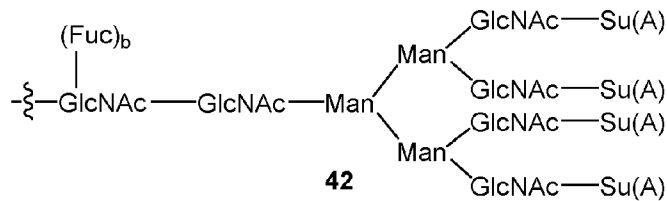
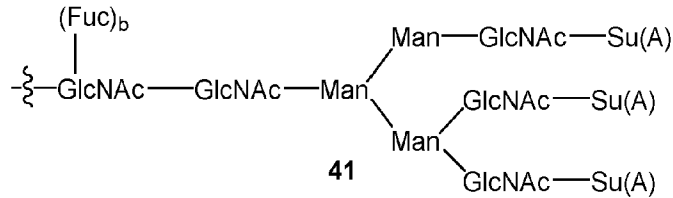
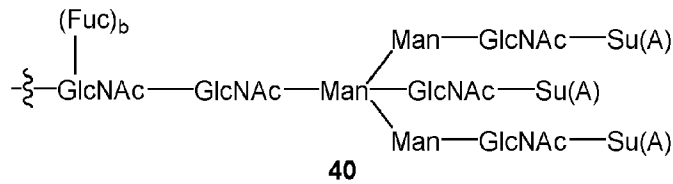
En la glucoproteína modificada según la invención, el C₁ del resto de Su(A) está unido al C₄ del resto de GlcNAc a través de un enlace β -1,4-O-glucosídico.

- 10 La glucoproteína modificada según la invención puede comprender más de un glucano según la fórmula (4) o (5). Cuando este es el caso, los dos o más glucanos pueden diferir entre sí. La glucoproteína también puede comprender uno o más glucanos adicionales que no comprendan un resto de Su(A).

En una forma de realización preferida, la glucoproteína modificada comprende un glucano según la fórmula (4), donde b es 0. En otra forma de realización preferida, la glucoproteína modificada comprende un glucano según la fórmula (4), donde b es 1.

- 15 En otra forma de realización preferida, la glucoproteína modificada comprende un glucano según la fórmula (5), donde b es 0. En otra forma de realización preferida, la glucoproteína modificada comprende un glucano según la fórmula (5), donde b es 1. En un glucano según la fórmula (5), G representa un monosacárido, o un oligosacárido lineal o ramificado que comprende de 1 a 20, preferiblemente de 1 a 12, más preferiblemente de 1 a 10, incluso más preferiblemente 1, 2, 3, 4, 5, 6, 7 u 8, y de la manera más preferible 1, 2, 3, 4, 5 o 6 restos de azúcar. En el glucano (5), se prefiere que cuando d sea 0, e sea 1 y cuando e sea 0, d sea 1. Más preferiblemente, en el glucano (5) d es 1, e incluso más preferiblemente d es 1 y e es 1. Los restos de azúcar que pueden estar presentes en un glucano son conocidos por un experto en la materia e incluyen, por ejemplo, glucosa (Glc), galactosa (Gal), manosa (Man), fucosa (Fuc), N-acetilglucosamina (GlcNAc), N-acetilgalactosamina (GalNAc), ácido N-acetilneuramínico (NeuNAc) o ácido siálico y xilosa (Xyl). Cuando el glucano es según la fórmula (5), se prefiere adicionalmente que el glucano sea según la fórmula (37), (38), (39), (40), (41) o (42):



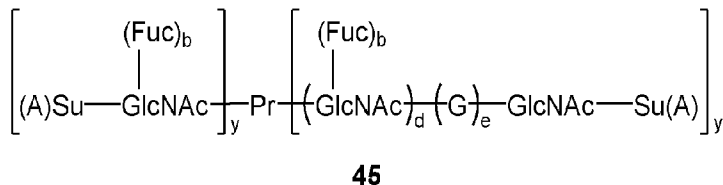
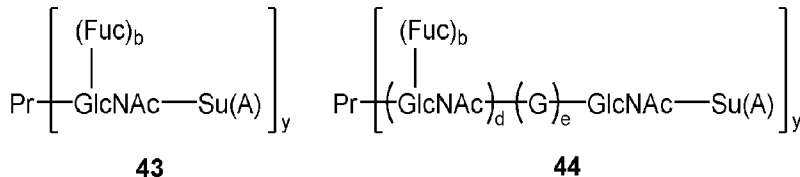


donde b es 0 o 1; y

5 Su (A) es según la fórmula (6) como se ha definido antes.

En una forma de realización preferida, la glucoproteína modificada según la invención comprende un glucano según la fórmula (4), (37), (38), (39), (40), (41) o (42), más preferiblemente un glucano unido a N según la fórmula (4), (37), (38), (39), (40), (41) o (42). En una forma de realización preferida adicional, la glucoproteína modificada según la invención comprende un glucano según la fórmula (4), (37), (38) o (39), más preferiblemente un glucano unido a N según la fórmula (4), (37), (38) o (39). Más preferiblemente, la glucoproteína modificada según la invención comprende un glucano según la fórmula (4) o (38), más preferiblemente un glucano unido a N según la fórmula (4) o (38).

La glucoproteína modificada según la invención es preferiblemente según la fórmula (43), (44) o (45):



15 donde:

b, d, e y G, y sus formas de realización preferidas, son como se ha definido anteriormente;

Su(A) es según la fórmula (6) como se ha definido anteriormente; y es independientemente un número entero en el intervalo de 1 a 24; y Pr es una proteína.

La glucoproteína modificada puede comprender uno o más glucanos (4) o (5) (y es 1 a 24). Preferiblemente, "y" es un número entero en el intervalo de 1 a 12, más preferiblemente un número entero en el intervalo de 1 a 10. Más preferiblemente, y es 1, 2, 3, 4, 5, 6, 7 u 8, y aún más preferiblemente y es 1, 2, 3, 4, 5 o 6. Incluso más preferiblemente, y es 1, 2, 3 o 4. Cuando y es 2 o más, los glucanos pueden diferir entre sí. La glucoproteína modificada también puede comprender una combinación de uno o más glucanos (4) y uno o más glucanos (5). Como se ha descrito anteriormente, la glucoproteína puede comprender además uno o más glucanos que no tienen un resto de Su(A).

25 Cuando la glucoproteína modificada según la invención es según la fórmula (43), (44) o (45), también se prefiere que la glucoproteína comprenda un glucano según la fórmula (4), (37), (38), (39), (40), (41) o (42) como se ha descrito anteriormente, más preferiblemente un glucano, preferiblemente un glucano unido a N según la fórmula (4), (37), (38)

o (39) y aún más preferiblemente según la fórmula (4) o (38). Más preferiblemente, el glucano que comprende un resto GlcNAc terminal es un glucano unido a N según la fórmula (4) o (38).

En una forma de realización preferida del proceso según la invención, la glucoproteína que comprende un glucano que comprende un resto GlcNAc terminal es un anticuerpo, más preferiblemente un anticuerpo según la fórmula (43), (44) o (45), donde la proteína (Pr) es un anticuerpo (Ab), o más específicamente Pr es la parte de polipéptido de un anticuerpo. Además, cuando la glucoproteína que se modifica es un anticuerpo y el anticuerpo comprende más de un glucano (y es 2 o más), los glucanos pueden diferir entre sí. El anticuerpo puede comprender además uno o más glucanos que no tienen un resto de Su(A). Además, cuando la glucoproteína modificada es un anticuerpo, se prefiere que el anticuerpo modificado comprenda un glucano según la fórmula (4), (37), (38), (39), (40), (41) o (42) como se ha definido anteriormente, más preferiblemente según la fórmula (4), (37), (38) o (39), incluso más preferiblemente según la fórmula (4) o (38). En esta forma de realización, se prefiere además que el anticuerpo comprenda un glucano unido a N según la fórmula (4), (37), (38), (39), (40),

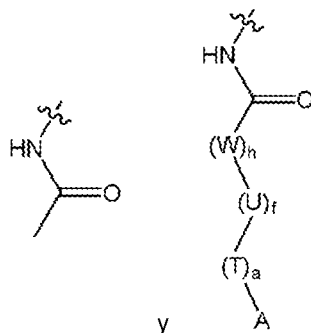
(41) o (42), más preferiblemente un glucano unido a N según la fórmula (4), (37), (38) o (39), y de la manera más preferible un glucano unido a N según la fórmula (4) o (38).

Cuando la glucoproteína modificada es un anticuerpo, se prefiere que y sea 1, 2, 3, 4, 5, 6, 7 u 8, más preferiblemente que y sea 1, 2, 4, 6 u 8, incluso más preferiblemente que y sea 1, 2 o 4, y más preferiblemente que y sea 1 o 2.

Como se ha definido anteriormente, dicho anticuerpo puede ser un anticuerpo completo, pero también un fragmento de anticuerpo. Cuando el anticuerpo es un anticuerpo completo, dicho anticuerpo comprende preferiblemente uno o más glucanos, más preferiblemente un glucano, en cada cadena pesada. Dicho anticuerpo completo, por lo tanto, comprende preferiblemente 2 o más, preferiblemente 2, 4, 6 u 8 de dichos glucanos, más preferiblemente 2 o 4, y de la forma más preferible 2 glucanos. En otras palabras, cuando dicho anticuerpo es un anticuerpo completo, "y" es preferiblemente 2, 4, 6 u 8, más preferiblemente "y" es 2 o 4, y más preferiblemente "y" es 2. Cuando el anticuerpo es un fragmento de anticuerpo, se prefiere que y sea 1, 2, 3 o 4, y más preferiblemente que y es 1 o 2.

En una forma de realización preferida, dicho anticuerpo es un anticuerpo monoclonal (mAb). Preferiblemente, dicho anticuerpo se selecciona del grupo que consiste en anticuerpos IgA, IgD, IgE, IgG e IgM. Más preferiblemente, dicho anticuerpo es un anticuerpo IgG1, IgG2, IgG3 o IgG4, y de la manera más preferible dicho anticuerpo es un anticuerpo IgG1.

En la glucoproteína modificada según la invención, R^{14} en Su(A) según la fórmula (6) es -NHAc se selecciona del grupo que consiste en:



donde W, h, a, f, T, A aynd U son como se ha definido anteriormente.

Lo más preferiblemente R^{14} en Su(A) según la fórmula (6) es -NHAc.

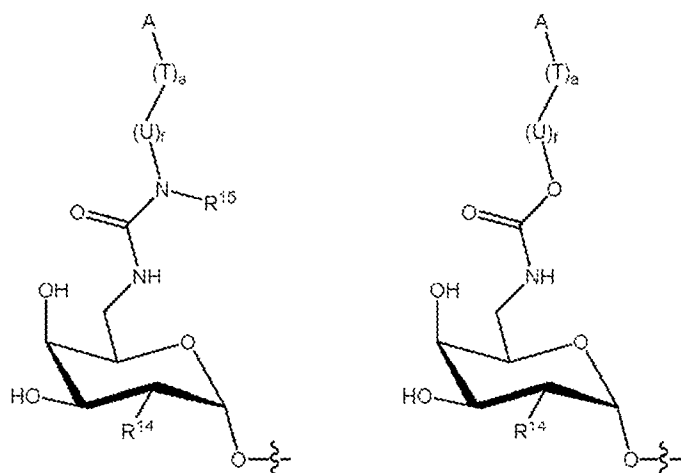
En una realización preferida adicional, la glucoproteína modificada según la invención comprende un glucano, más preferiblemente un glucano unido a N, según la fórmula (4) o (5), donde R^{14} en Su(A) (6) es -NHC(O)CH₃. En esta realización, se prefiere que la glucoproteína modificada sea según la fórmula (43), (44) o (45).

En otra realización preferida adicional, la glucoproteína modificada según la invención comprende un glucano, más preferiblemente un glucano unido a N, según la fórmula (4) o (5), donde R^{14} en Su(A) (6) es -NHC(O)-(W)_h-(U)_f-(T)_a-A donde W, h, U, f, T, a y A, y las realizaciones preferidas de los mismos, son como se ha descrito con más detalle anteriormente. En esta realización se prefiere que la glucoproteína modificada sea según la fórmula (43), (44) o (45).

En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, más preferiblemente un glucano unido a N, según la fórmula, (37), (38), (39), (40), (41) o (42), donde R^{14} en Su(A) (6) es -NHC(O)CH₃. En esta realización se prefiere que la glucoproteína modificada sea según la fórmula (43), (44) o (45).

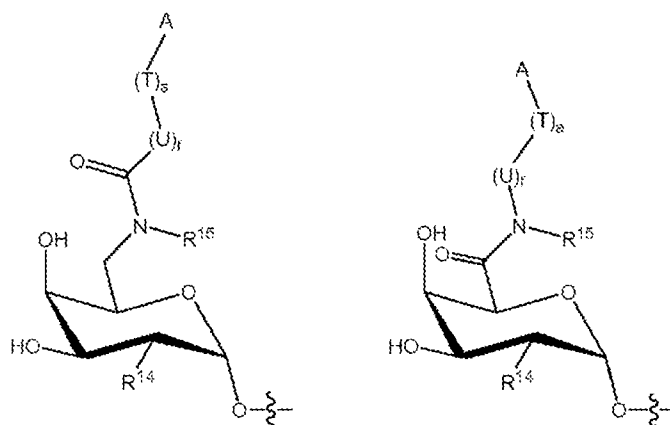
En otra forma de realización preferida adicional, la glucoproteína modificada según la invención comprende un glucano, más preferiblemente un glucano unido a N, según la fórmula (37), (38), (39), (40), (41) o (42), donde R^{14} en Su(A) (6) es $-NHC(O)-(W)_h-(U)_f-(T)_a-A$. En esta forma de realización, se prefiere que la glucoproteína modificada sea según la fórmula (43), (44) o (45).

- 5 En otra forma de realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (4) o (5), más preferiblemente un glucano, incluso preferiblemente un glucano unido a N, según la fórmula (37), (38), (39), (40), (41) o (42), donde Su(A) (6) es según la fórmula (46), (47), (48) o (49):



46

47



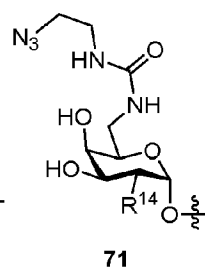
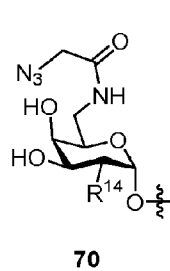
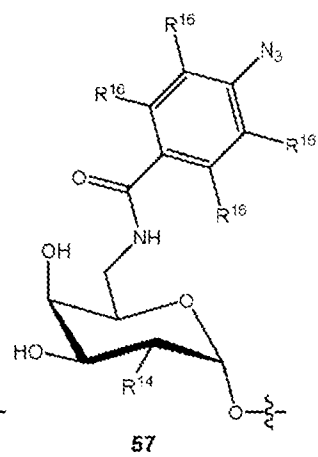
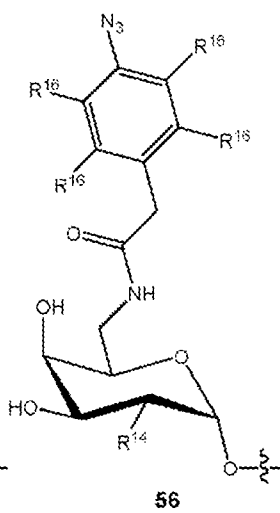
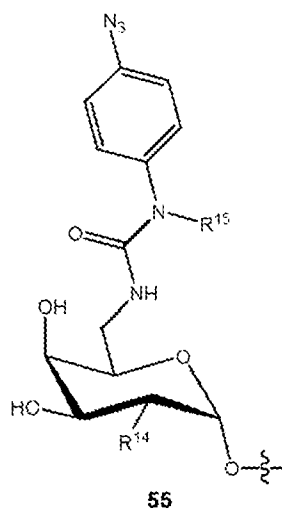
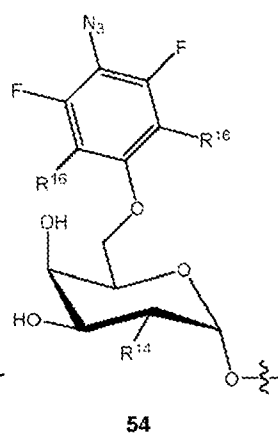
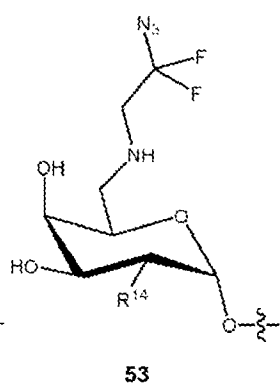
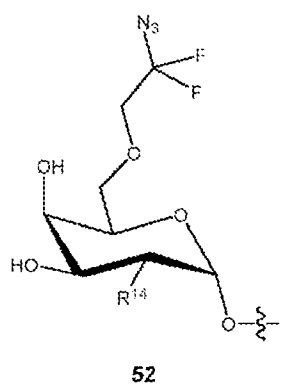
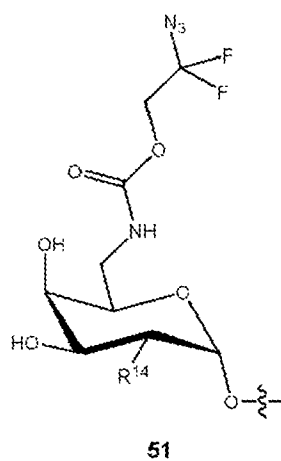
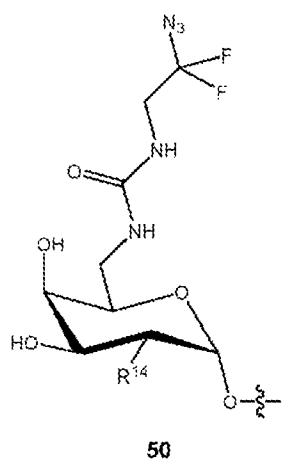
48

49

donde a, f, R^{14} , R^{15} , A, U y T, y las realizaciones preferidas de los mismos, son como se ha definido anteriormente para (15), (16), (17) y (18).

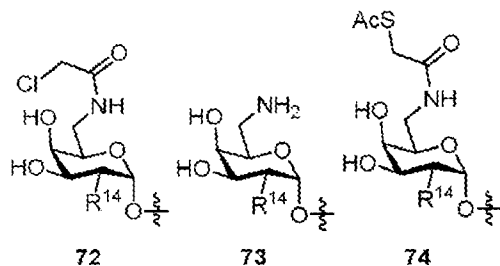
En estas realizaciones donde Su(A) (6) es según la fórmula (46), (47), (48) o (49), en una realización preferida adicional R^{14} es $-NHC(O)CH_3$. En otra realización preferida adicional, R^{14} es $-NHC(O)-(W)_h-(U)_f-(T)_a-A$ donde W, h, U, f, T, a y A, y las realizaciones preferidas de los mismos, son como se ha descrito con más detalle anteriormente. También en estas realizaciones se prefiere que la glucoproteína modificada sea según la fórmula (43), (44) o (45).

En otra forma de realización preferida adicional, la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (4) o (5), más preferiblemente un glucano, incluso más preferiblemente un glucano unido a N, según la fórmula (37), (38), (39), (40), (41) o (42), donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70) o (71), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57):



5 donde R^{14} , R^{15} y R^{16} , y las realizaciones preferidas de los mismos, son como se ha definido anteriormente para (19), (20), (21), (22), (23), (24), (25), (26), (65), (66), (67), (68) y (69).

En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (4) o (5), más preferiblemente un glucano, incluso más preferiblemente un glucano unido a N, según la fórmula (37), (38), (39), (40), (41) o (42), donde Su(A) (6) es según la fórmula (72), (73) o (74):

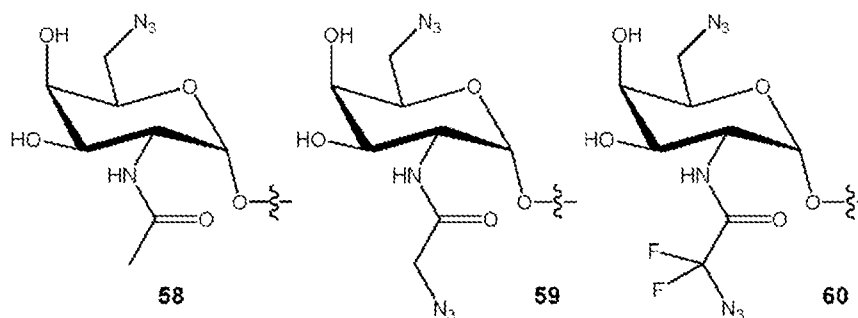


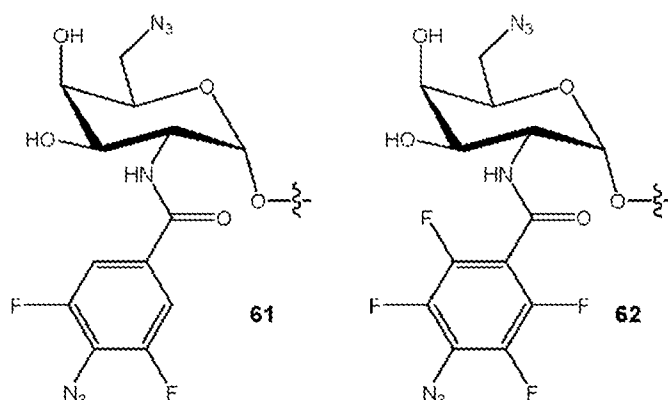
donde R¹⁴ y las realizaciones preferidas de los mismos, son como se ha definido anteriormente para (72), (73) y (74).

En una realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (37), donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70), (71), (72), (73) o (74), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (38), donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70), (71), (72), (73) o (74), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (39), donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70), (71), (72), (73) o (74), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (40), donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70), (71), (72), (73) o (74), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (41), donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70), (71), (72), (73) o (74), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (42), donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70), (71), (72), (73) o (74), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57).

En estas realizaciones donde Su(A) (6) es según la fórmula (50), (51), (52), (53), (54), (55), (56), (57), (70), (71), (72), (73) o (74), preferiblemente según la fórmula (50), (51), (52), (53), (54), (55), (56) o (57), en una realización preferida adicional R¹⁴ es -NHC(O)CH₃. En otra realización preferida adicional R¹⁴ es -NHC(O)-(W)_h-(U)_i-(T)_a-A donde W, h, U, f, T, a y A, y las realizaciones preferidas de los mismos, son como se ha descrito con más detalle anteriormente. También en estas realizaciones se prefiere que la glucoproteína modificada sea según la fórmula (43), (44) o (45).

En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (4) o (5), más preferiblemente un glucano, incluso más preferiblemente un glucano unido a N, según la fórmula (37), (38), (39), (40), (41) o (42), donde Su(A) (6) es según la fórmula (58), (59), (60), (61) o (62):





En una realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (38), donde Su(A) (6) es según la fórmula (58), (59), (60), (61) o (62). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (39), donde Su(A) (6) es según la fórmula (58), (59), (60), (61) o (62). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (40), donde Su(A) (6) es según la fórmula (58), (59), (60), (61) o (62). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (41), donde Su(A) (6) es según la fórmula (58), (59), (60), (61) o (62). En otra realización preferida adicional la glucoproteína modificada según la invención comprende un glucano, preferiblemente un glucano unido a N, según la fórmula (42), donde Su(A) (6) es según la fórmula (58), (59), (60), (61) o (62).

En estas realizaciones donde Su(A) (6) es según la fórmula (58), (59), (60), (61) o (62) se prefiere que la glucoproteína modificada sea según la fórmula (43), (44) o (45).

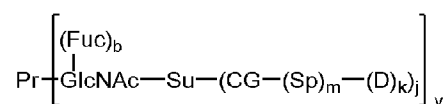
La invención también se refiere al uso de la glucoproteína modificada según la invención, como se ha definido anteriormente, en un proceso para preparar un bioconjugado, preferiblemente el bioconjugado según la invención. El proceso es preferiblemente para preparar un conjugado anticuerpo-fármaco (aDc). El proceso comprende poner en contacto la glucoproteína modificada con un conjugado enlazador. En una forma de realización preferida, el conjugado enlazador es como se define aquí. En una forma de realización preferida, el bioconjugado es un bioconjugado según la fórmula (75), (76) o (77), como se define más adelante en este documento.

Bioconjugado

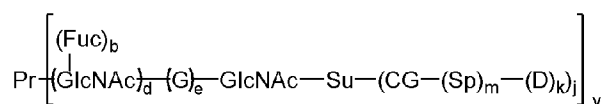
La presente invención se refiere además a un bioconjugado que puede obtenerse conjugando un conjugado enlazador a la glucoproteína modificada según la invención. Los conjugados enlazadores se conocen en la técnica como uno de los reactivos en una reacción de bioconjugación, donde una glucoproteína, tal como una glucoproteína modificada según la invención, es el otro reactivo. Un conjugador enlazador se define aquí como un compuesto en el que una molécula de actuación está conectada covalentemente a un grupo reactivo Q1 a través de un enlazador. El grupo reactivo Q1 es capaz de reaccionar con el grupo funcional A presente en la glucoproteína modificada según la invención. Un conjugado enlazador puede comprender más de un grupo reactivo Q1 y/o más de una molécula de actuación. Entre los conjugados enlazadores adecuados se incluyen aquellos descritos en WO 2014/065661 y WO 2016/053107.

Las reacciones de bioconjugación son conocidas en el campo de los conjugados de anticuerpos, tales como los conjugados anticuerpo-fármaco (ADC), donde se usan para preparar conjugados de un anticuerpo con una molécula de actuación, típicamente una citotoxina. En dicha reacción de bioconjugación, la glucoproteína modificada según la invención se acopla o se conjuga al conjugado enlazador en virtud de una reacción entre el grupo funcional A presente en la glucoproteína modificada y un grupo reactivo Q1 presente en el conjugado enlazador. El bioconjugado según la invención es preferiblemente un conjugado de anticuerpo, en el que un anticuerpo está conjugado a una molécula de actuación, de la forma más preferible como un conjugado anticuerpo-fármaco, en el que un anticuerpo está conjugado con un fármaco, típicamente una citotoxina.

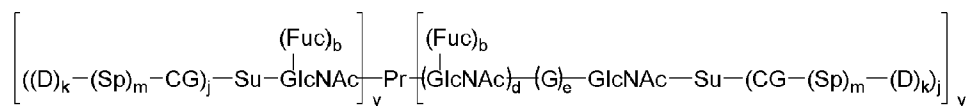
Más en particular, la invención se refiere a un bioconjugado según la fórmula (75), (76) o (77):



75



76



77

donde:

Pr es una proteína

5 y es independientemente como se ha definido anteriormente para (43);

b, d, e y G son independientemente como se ha definido anteriormente para (5);

CG es un grupo conector que conecta Su a Sp o D;

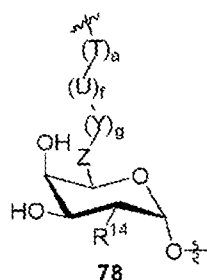
Sp es un espaciador;

D es una molécula de actuación;

10 j es independientemente 1, 2, 3, 4 o 5, preferiblemente j es 1;

k es independientemente un número entero en el intervalo de 1 a 10, preferiblemente k es 1, 2, 3 o 4, más preferiblemente k es 1; m es 0 o 1, preferiblemente m es 1.

Su es según la fórmula (78):



78

15 donde R14, Z, Y, U, T, g, f y a son como se ha definido anteriormente para (3), y Su está conectado a través de C1 a C4 del resto de GlcNAc a través de un enlace β -1,4-O-glucosídico y a cG a través de Z, Y, U o T.

Las formas de realización preferidas de Pr e "y" en (75), (76) o (77) son como se ha descrito anteriormente con más detalle para (43), (44) y (45). En una forma de realización preferida, la glucoproteína es un anticuerpo. El bioconjugado, en particular el anticuerpo, puede comprender más de un glucano funcionalizado ("y" es 2 o más), los glucanos pueden diferir entre sí. El anticuerpo puede comprender además uno o más glucanos que no tienen un resto Su-(CG-(Sp)- (D)_k)_j. Se prefiere adicionalmente que el glucano funcionalizado sea un glucano unido a N. Cuando el bioconjugado según la invención es un conjugado de anticuerpo, se prefiere que y sea 1, 2, 3, 4, 5, 6, 7 u 8, más preferiblemente y es 1, 2, 4, 6 u 8, incluso más preferiblemente que y sea 1, 2 o 4, y más preferiblemente que y sea 1 o 2.

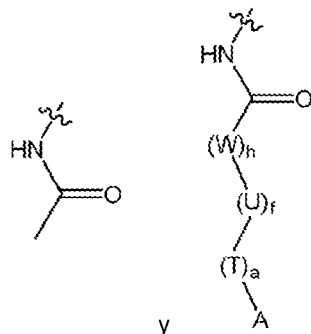
25 Como se ha definido anteriormente, dicho anticuerpo puede ser un anticuerpo completo, pero también un fragmento de anticuerpo. Cuando el anticuerpo es un anticuerpo completo, dicho anticuerpo comprende preferiblemente uno o más glucanos, más preferiblemente un glucano, en cada cadena pesada. De este modo, dicho anticuerpo completo comprende preferiblemente 2 o más, preferiblemente 2, 4, 6 u 8 de dichos glucanos funcionalizados, más preferiblemente 2 o 4, y de la manera más preferible 2 glucanos funcionalizados. En otras palabras, cuando dicho anticuerpo es un anticuerpo completo, y es preferiblemente 2, 4, 6 u 8, más preferiblemente y es 2 o 4, y más preferiblemente y es 2. Cuando el anticuerpo es un fragmento de anticuerpo, se prefiere que y sea 1, 2, 3 o 4, y más preferiblemente que y sea 1 o 2.

En una forma de realización preferida, dicho anticuerpo es un anticuerpo monoclonal (mAb). Preferiblemente, dicho anticuerpo se selecciona del grupo que consiste en anticuerpos IgA, IgD, IgE, IgG e IgM. Más preferiblemente, dicho anticuerpo es un anticuerpo IgG1, IgG2, IgG3 o IgG4, y de la manera más preferible dicho anticuerpo es un anticuerpo IgG1.

Las formas de realización preferidas de la cadena de glucano, en particular de b, d, e y G, en (75), (76) o (77) son como se ha descrito anteriormente en más detalle para (4) y (5) y sus formas de realización preferidas, tales como, por ejemplo, (37), (38), (39), (40), (41) o (42).

- 5 Las formas de realización preferidas de R^{14} , Z, Y, U, T, g, f y a en (78) son como se ha descrito anteriormente con más detalle para (3) y formas de realización preferidas de (3) como, por ejemplo, (3a) o (3b). Las formas de realización preferidas para Su según la fórmula (78) corresponden a Su(A) según cualquiera de (46) a (62) y (70) a (74), y sus formas de realización, como se ha descrito anteriormente con más detalle para (6), aunque con haciendo reaccionar A con Q1 para formar CG.

En el bioconjugado según la invención, R14 en Su según la fórmula (78) se selecciona del grupo que consiste en:



donde W, h, a, f, T, A y U son como se ha definido anteriormente.

Lo más preferiblemente R^{14} en Su según la fórmula (78) es -NHAc.

15 D es una molécula de actuación. Las moléculas de actuación se definen aquí como estructuras moleculares que poseen una propiedad deseada que se imparte en la biomolécula tras la conjugación. La molécula de actuación D se selecciona preferiblemente del grupo que consiste en sustancias activas, moléculas indicadoras, polímeros, superficies sólidas, hidrogeles, nanopartículas, micropartículas y biomoléculas. Más preferiblemente, la molécula de actuación D es una sustancia activa.

20 El término "sustancia activa" en el presente documento se refiere a una sustancia farmacológica y/o biológica, es decir, una sustancia que es biológicamente y/o farmacéuticamente activa, por ejemplo un fármaco, un profármaco, un agente de diagnóstico, una proteína, un péptido, un polipéptido, un marcador peptídico, un aminoácido, un glucano, un lípido, una vitamina, un esteroide, un nucleótido, un nucleósido, un polinucleótido, ARN o ADN. Los ejemplos de marcadores peptídicos incluyen péptidos que penetran en la célula, como la poliarginina o la lactoferrina humana. Un ejemplo de glucano es la oligomanosa. Un ejemplo de aminoácido es la lisina. Cuando la molécula de actuación es una sustancia activa, la sustancia activa se selecciona preferiblemente del grupo que consiste en fármacos y profármacos. Más preferiblemente, la sustancia activa se selecciona del grupo que consiste en compuestos farmacéuticamente activos, en particular compuestos de peso molecular bajo a medio (por ejemplo, de aproximadamente 200 a aproximadamente 2500 Da, preferiblemente de aproximadamente 300 a aproximadamente 1750 Da). En una forma de realización preferida adicional, la sustancia activa se selecciona del grupo que consiste en citotoxinas, agentes antivirales, agentes antibacterianos, péptidos y oligonucleótidos. Los ejemplos de citotoxinas incluyen colchicina, alcaloides de la vinca, antraciclinas, camptotecinas, doxorubicina, daunorubicina, taxanos, calicheamicinas, tubulisin, irinotecanos, un péptido inhibidor, amanitina, deBouganina, duocarmicinas, maitansinas, auristatinas o pirrolobenzodiazepinas (PBD).

35 El término "molécula indicadora" en este documento se refiere a una molécula cuya presencia se detecta fácilmente, por ejemplo, mediante un agente de diagnóstico, un colorante, un fluoróforo, un marcador isotópico radiactivo, un agente de contraste, un agente de formación de imágenes por resonancia magnética o un marcador de masa. El experto en la técnica conoce una amplia variedad de fluoróforos, también denominados sondas fluorescentes. Se describen varios fluoróforos con más detalle en, por ejemplo, G.T. Hermanson, "Bioconjugate Techniques", Elsevier, 3a ed. 2013, Capítulo 10: "Sondas fluorescentes", pp. 395 - 463. Los ejemplos de un fluoróforo incluyen todos los tipos de Alexa Fluor (por ejemplo, Alexa Fluor 555), colorantes de cianina (por ejemplo, Cy3 o Cy5) y derivados de colorantes de cianina, derivados de cumarina, fluoresceína y derivados de fluoresceína, rodamina y derivados de rodamina, derivados borodipirrometeno, derivados de pireno, derivados de naftalimida, derivados de fibobiliproteína (por ejemplo, alofocianina), cromomicina, quelatos de lantánidos y nanocristales o puntos cuánticos. Los ejemplos de una 40 marcador isotópico radiactivo incluyen ^{99m}Tc , ^{111}In , ^{114}mIn , ^{115}In , ^{18}F , ^{14}C , ^{64}Cu , ^{131}I , ^{125}I , ^{123}I , ^{212}Bi , ^{88}Y , ^{90}Y , ^{67}Cu , ^{186}Rh , ^{188}Rh , ^{66}Ga , ^{67}Ga y ^{10}B , que está opcionalmente conectado a través de un resto quelante tal como, por ejemplo, dTPA (anhídrido dietilentriaminopentaacético), DOTA (1,4,7,10-tetraazaciclododecano-N, N, N', N''-Ácido tetraacético), NOTA (1,4,7-triazaciclononano N, N', N''-triacético), TETA (1,4,8,11-tetraazaciclotetradecano-N, N', N'', N'''-ácido tetraacético), DTTA ($\wedge$ -(p-isotiocianatobencilo) -dietaniletanotriamina-Ní,N2,N3,N3-ácido tetraacético), deferoxamina o DFA (N'-[5-[[4-[[5-(acetilhidroxiamino) pentil]amino]-1,4-dioxobutol]hidroxiamino]pentil]-N-(5-aminopentilo)-N-hidroxitbutanodiamida) o HYNIC (hidrazinonicotinamida). Las técnicas de marcado isotópico son 45

conocidas por una persona experta en la técnica, y se describen con más detalle en, p. G.T. Hermanson, "Bioconjugate Techniques", Elsevier, 3a ed. 2013, Capítulo 12: "Isotopic Labelling Techniques", p. 507 - 534.

Los polímeros adecuados para usar como molécula de actuación D en el compuesto según la invención son conocidos por una persona experta en la técnica, y varios ejemplos se describen con más detalle en, por ejemplo, G.T. Hermanson, "Bioconjugate Techniques", Elsevier, 3a ed. 2013, Capítulo 18: "PEGylation and synthetic polymer modification", pp. 787 - 838. Cuando la molécula de actuación D es un polímero, la molécula de actuación D se selecciona preferiblemente de manera independiente respecto del grupo que consiste en un poli(etilenglicol) (PEG), un óxido de polietileno (PEG), un polipropilenglicol (PPG), un óxido de polipropileno (PPO), un polímero diaminoalcano 1,x (donde x es el número de átomos de carbono del alcano, y preferiblemente x es un número entero en el intervalo de 2 a 200, preferiblemente de 2 a 10), una (poli)etilenglicol diamina (por ejemplo, 1,8-diamino-3,6-dioxaoctano y equivalentes que comprenden cadenas de etilenglicol más largas), un polisacárido (por ejemplo, dextrano), un poli(aminoácido) (por ejemplo, una poli(L-lisina)) y un poli (alcohol vinílico), una poli(2-oxazolina)s (PAOx).

Las superficies sólidas adecuadas para uso como una molécula de actuación D son conocidas por una persona experta en la técnica. Una superficie sólida es, por ejemplo, una superficie funcional (por ejemplo, una superficie de un nanomaterial, un nanotubo de carbono, un fullereno o una cápside vírica), una superficie de metal (por ejemplo, una superficie de titanio, oro, plata, cobre, níquel, estaño, rodio o zinc), una superficie de aleación de metal (donde la aleación es, por ejemplo, de aluminio, bismuto, cromo, cobalto, cobre, galio, oro, indio, hierro, plomo, magnesio, mercurio, níquel, potasio, plutonio, rodio, escandio, plata, sodio, titanio, estaño, uranio, zinc y/o zirconio), una superficie de polímero (en la que el polímero es, por ejemplo, poliestireno, poli(cloruro de vinilo), polietileno, polipropileno, poli(dimetilsiloxano) o polimetilmetacrilato, poli(acrilamida), una superficie de vidrio, una superficie de silicona, una superficie de soporte de cromatografía (donde el soporte de cromatografía es, por ejemplo, un soporte de sílice, un soporte de agarosa, un soporte de celulosa o un soporte de alúmina), etc. Cuando la molécula de actuación D es una superficie sólida, se prefiere que D se seleccione independientemente del grupo que consiste en una superficie funcional o una superficie polimérica.

Los hidrogeles son conocidos por los expertos en la técnica. Los hidrogeles son redes hinchadas por el agua, formadas por enlaces cruzados entre los constituyentes poliméricos. Véase, por ejemplo, A. S. Hoffman, Adv. Drug Delivery Rev. 2012, 64, 18. Cuando la molécula de actuación es un hidrogel, se prefiere que el hidrogel esté compuesto de poli(etilenglicol) (PEG) como base polimérica.

Las micro y nanopartículas adecuadas para uso como una molécula de actuación D son conocidas por una persona experta en la técnica. Se describe una variedad de micro y nanopartículas adecuadas en, por ejemplo, G.T. Hermanson, "Bioconjugate Techniques", Elsevier, 3a ed. 2013, Capítulo 14: "microparticles and nanoparticles", p. 549 - 587. Las micro o nanopartículas pueden tener cualquier forma, por ejemplo esferas, barras, tubos, cubos, triángulos y conos. Preferiblemente, las micro o nanopartículas tienen forma esférica. La composición química de las micro y nanopartículas puede variar. Cuando la molécula de actuación D es una micro o nanopartícula, la micro o nanopartícula es, por ejemplo, una micro o nanopartícula polimérica, una micro o nanopartícula de sílice o una micro o nanopartícula de oro. Cuando la partícula es una micropartícula o nanopartícula polimérica, el polímero es preferiblemente poliestireno o un copolímero de estireno (por ejemplo, un copolímero de estireno y divinilbenceno, butadieno, acrilato y/o viniltolueno), polimetilmetacrilato (PMMA), poliviniltolueno, poli(hidroxietilmetacrilato) (pHEMA) o poli(etilenglicoldimetacrilato / 2-hidroxietilmetacrilato) [poli (EDGMA/HEMA)]. Opcionalmente, la superficie de las micro o nanopartículas se modifica, por ejemplo, con detergentes, mediante polimerización por injerto de polímeros secundarios o mediante unión covalente de otro polímero o de restos espaciadores, etc.

La molécula de actuación D también puede ser una biomolécula. Cuando la molécula de actuación D es una biomolécula, se prefiere que la biomolécula se seleccione del grupo que consiste en proteínas (incluyendo glucoproteínas y anticuerpos), polipéptidos, péptidos, glucanos, lípidos, ácidos nucleicos, oligonucleótidos, polisacáridos, oligosacáridos, enzimas, hormonas, aminoácidos y monosacáridos.

CG es un grupo conector. El término "grupo conector" en este documento se refiere al elemento estructural que conecta una parte de un compuesto y otra parte del mismo compuesto. Típicamente, se prepara un bioconjugado por reacción de un grupo reactivo Q1 presente en un conjugador enlazador con un grupo funcional A presente en la glucoproteína modificada según la invención. CG es el resto formado por la reacción del grupo reactivo Q1 con el resto funcional A. Como entenderá la persona experta en la técnica, la naturaleza del CG depende del tipo de reacción orgánica utilizada para establecer la conexión entre la glucoproteína modificada según la invención y el conjugado enlazador. En otras palabras, la naturaleza del CG depende de la naturaleza del grupo reactivo Q1 del conjugado enlazador y la naturaleza del grupo funcional A de la biomolécula. Dado que existe una gran cantidad de reacciones químicas diferentes disponibles para establecer la conexión entre la glucoproteína modificada y el conjugado enlazador, en consecuencia, hay un gran número de posibilidades para el CG. En la figura 5 se muestran varios ejemplos de combinaciones adecuadas de F1 y Q1 y del grupo conector Z3, que estará presente en un bioconjugado cuando un conjugado enlazador que comprenda Q1 esté conjugado con una biomolécula que comprenda un grupo funcional complementario F1.

- Cuando A es, por ejemplo, un grupo tiol, los grupos complementarios Q1 incluyen grupos N-maleimidilo y grupos alquenilo, y los grupos conectores CG correspondientes son como se muestra en la Figura 5. Cuando A es un grupo tiol, los grupos complementarios Q1 también incluyen grupos alenamida.
- 5 Cuando A es, por ejemplo, un grupo amino, los grupos complementarios Q1 incluyen grupos cetona, grupos éster activados y grupos azido, y los grupos conectores correspondientes CG son como se muestra en la Figura 5.
- Cuando A es, por ejemplo, un grupo cetona, los grupos complementarios Q1 incluyen grupos (O- alquil)hidroxilamino y grupos hidrazina, y los grupos conectores correspondientes CG son como se muestra en la Figura 5.
- Cuando A es, por ejemplo, un grupo alquínilo, los grupos complementarios Q1 incluyen grupos azido, y el grupo conector correspondiente CG es como se muestra en la Figura 5.
- 10 Cuando A es, por ejemplo, un grupo alqueno, los grupos complementarios Q1 incluyen grupos tioles, dienos o heterodienos que son reactivos en una cicloadición de Diels-Alder y grupos tetrazinilo, y el grupo conector correspondiente CG puede ser tioéteres, aductos de Diels-Alder (ciclohexenos o sus análogos) o dihidropiridazina, respectivamente.
- 15 Cuando A es, por ejemplo, un grupo azido, los grupos complementarios Q1 incluyen grupos alquínilo, y el grupo conector correspondiente CG es como se muestra en la Figura 5.
- Cuando A es, por ejemplo, un grupo ciclopropenilo, un grupo trans-cicloocteno o un grupo ciclooctileno, los grupos complementarios Q1 incluyen grupos tetrazinilo y el grupo conector Z3 correspondiente es como se muestra en la Figura 5. En estos casos particulares, Z3 es solo una estructura intermedia y expulsará N₂, generando así una dihidropiridazina (a partir de la reacción con alqueno) o piridazina (a partir de la reacción con alquino).
- 20 Cuando A es, por ejemplo, un halógeno (X), los grupos complementarios Q1 incluyen tioles y los grupos conectores correspondientes CG pueden ser un tioéter.
- Cuando A es, por ejemplo, -OS(O)R⁵, los grupos complementarios Q1 incluyen grupos hidroxilo y amina (primaria y secundaria), y los grupos conectores correspondientes CG pueden ser un o un grupo amina (secundaria o terciaria) o éter.
- 25 Cuando A es, por ejemplo, un grupo alenilo, los grupos complementarios Q1 incluyen tioles y los grupos conectores correspondientes CG pueden ser un tioéter, típicamente un tioéter sustituido con metilo.
- Cuando A es, por ejemplo, -SC(O)R⁸ o -SC(V)OR⁸, típicamente A reacciona primero a un tiol, y los grupos complementarios Q1 incluyen grupos N-maleimidilo, grupos alquenilo, grupos alenamida. Los grupos conectores correspondientes CG pueden ser como se muestra en la figura 5 siendo A tiol.
- 30 Las combinaciones adicionales adecuadas de A y Q1, y la naturaleza del grupo conector CG resultante son conocidas por una persona experta en la técnica, y están descritas por ejemplo en G.T. Hermanson, "Bioconjugate Techniques", Elsevier, 3a ed. 2013 (ISBN: 978-0-12-382239-0), en particular en el Capítulo 3, páginas 229 - 258. Una lista de grupos reactivos complementarios adecuados para procesos de bioconjugación se describe en la Tabla 3.1, páginas 230 - 232 del Capítulo 3 de G.T. Hermanson, "Bioconjugate Techniques", Elsevier, 3a ed. 2013 (ISBN: 978-0-12-382239-0).
- 35 Sp es un espaciador o un enlazador. Un espaciador se define aquí como un resto que conecta dos o más elementos de un compuesto. Por ejemplo, en un bioconjugado, una biomolécula y una molécula de actuación están conectadas covalentemente entre sí a través de un enlazador; en un conjugado enlazador un grupo reactivo Q1 está conectado de manera covalente a una molécula de actuación a través de un enlazador. Cualquier enlazador conocido en la técnica que sea adecuado para su uso en bioconjugados, en particular conjugados de anticuerpos, se puede usar
- 40 como Sp. Dichos restos espaciadores son conocidos por una persona experta en la técnica. Los ejemplos de restos espaciadores adecuados incluyen (poli)etilenglicol diaminas (por ejemplo, 1,8-diamino-3,6-dioxaoctano o equivalentes que comprenden cadenas de etilenglicol más largas), cadenas de polietilenglicol o cadenas de óxido de polietileno, cadenas de polipropilenglicol o cadenas de óxido de polipropileno y 1, x-diaminoalcanos donde x es el número de átomos de carbono en el alcano. Otra clase de restos espaciadores adecuados comprende restos espaciadores escindibles o enlazadores escindibles. Los enlazadores escindibles son ampliamente conocidos en la técnica. Por ejemplo Shabat et al., Soft Matter 2012, 6, 1073, describen enlazadores escindibles que comprenden restos autoinmolutivos que se liberan tras un desencadenante biológico, por ejemplo una división enzimática o una oxidación. Algunos ejemplos de enlazadores escindibles adecuados son enlazadores disulfuro que se escinden tras la reducción, enlazadores peptídicos que se escinden tras el reconocimiento específico por una proteasa, por ejemplo cathepsina,
- 45 plasmína o metaloproteasas, o enlazadores basados en glucósidos que se escinden tras el reconocimiento específico por una glucosidasa, por ejemplo glucoronidasa, o nitroaromáticos que se reducen en áreas hipoxicas pobres en oxígeno. Aquí, los restos espaciadores escindibles adecuados también incluyen restos espaciadores que comprenden una secuencia de aminoácidos específica escindible. Los ejemplos incluyen, por ejemplo, restos espaciadores que comprenden un resto Val-Ala (valina-alanina) o Val-Cit (valina-citrulina).
- 50

- En una forma de realización preferida, Sp se selecciona del grupo que consiste en grupos lineales o ramificados alquilenos C₁-C₂₀₀, grupos alquilenos C₂-C₂₀₀, grupos alquilenos C₂-C₂₀₀, grupos cicloalquilenos C₃-C₂₀₀, grupos cicloalquilenos C₅-C₂₀₀, grupos cicloalquilenos C₈-C₂₀₀, grupos alquilarilenos C₇-C₂₀₀, grupos arilalquilenos C₇-C₂₀₀, grupos arilalquilenos C₈-C₂₀₀ y grupos arilalquilenos C₉-C₂₀₀, donde los grupos alquilenos, grupos alquilenos, grupos alquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos alquilarilenos, grupos arilalquilenos, grupos arilalquilenos y grupos arilalquilenos están opcionalmente sustituidos y opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo de O, S y NR¹⁹, donde R¹⁹ se selecciona independientemente del grupo que consiste en hidrógeno, grupos alquilo C₁ - C₂₄, grupos alquilenos C₂ - C₂₄, grupos alquilenos C₂ - C₂₄ y grupos cicloalquilo C₃ - C₂₄, donde los grupos alquilo, grupos alquilenos, grupos alquilenos y grupos cicloalquilo están opcionalmente sustituidos. Cuando los grupos alquilenos, grupos alquilenos, grupos alquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos alquilarilenos, grupos arilalquilenos, grupos arilalquilenos y grupos arilalquilenos están interrumpidos por uno o más heteroátomos como se ha definido anteriormente, se prefiere que dichos grupos estén interrumpidos por uno o más átomos de O, y/o por uno o más grupos S-S.
- Más preferiblemente, Sp se selecciona del grupo que consiste en grupos lineales o ramificados alquilenos C₁-C₁₀₀, grupos alquilenos C₂-C₁₀₀, grupos alquilenos C₂-C₁₀₀, grupos cicloalquilenos C₃-C₁₀₀, grupos cicloalquilenos C₅-C₁₀₀, grupos cicloalquilenos C₈-C₁₀₀, grupos alquilarilenos C₇-C₁₀₀, grupos arilalquilenos C₇-C₁₀₀, grupos arilalquilenos C₈-C₁₀₀ y grupos arilalquilenos C₉-C₁₀₀, donde los grupos alquilenos, grupos alquilenos, grupos alquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos alquilarilenos, grupos arilalquilenos, grupos arilalquilenos y grupos arilalquilenos están opcionalmente sustituidos y opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo de O, S y NR¹⁹, donde R¹⁹ se selecciona independientemente del grupo que consiste en hidrógeno, grupos alquilo C₁ - C₂₄, grupos alquilenos C₂ - C₂₄, grupos alquilenos C₂ - C₂₄ y grupos cicloalquilo C₃ - C₂₄, donde los grupos alquilo, grupos alquilenos, grupos alquilenos y grupos cicloalquilo están opcionalmente sustituidos.
- Incluso más preferiblemente, Sp se selecciona del grupo que consiste en grupos lineales o ramificados alquilenos C₁-C₅₀, grupos alquilenos C₂-C₅₀, grupos alquilenos C₂-C₅₀, grupos cicloalquilenos C₃-C₅₀, grupos cicloalquilenos C₅-C₅₀, grupos cicloalquilenos C₈-C₅₀, grupos alquilarilenos C₇-C₅₀, grupos arilalquilenos C₇-C₅₀, grupos arilalquilenos C₈-C₅₀ y grupos arilalquilenos C₉-C₅₀, donde los grupos alquilenos, grupos alquilenos, grupos alquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos alquilarilenos, grupos arilalquilenos, grupos arilalquilenos y grupos arilalquilenos están opcionalmente sustituidos y opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo de O, S y NR¹⁹, donde R¹⁹ se selecciona independientemente del grupo que consiste en hidrógeno, grupos alquilo C₁ - C₂₄, grupos alquilenos C₂ - C₂₄, grupos alquilenos C₂ - C₂₄ y grupos cicloalquilo C₃ - C₂₄, donde los grupos alquilo, grupos alquilenos, grupos alquilenos y grupos cicloalquilo están opcionalmente sustituidos.
- Aún más preferiblemente, Sp se selecciona del grupo que consiste en grupos lineales o ramificados alquilenos C₁-C₂₀, grupos alquilenos C₂-C₂₀, grupos alquilenos C₂-C₂₀, grupos cicloalquilenos C₃-C₂₀, grupos cicloalquilenos C₅-C₂₀, grupos cicloalquilenos C₈-C₂₀, grupos alquilarilenos C₇-C₂₀, grupos arilalquilenos C₇-C₂₀, grupos arilalquilenos C₈-C₂₀ y grupos arilalquilenos C₉-C₂₀, donde los grupos alquilenos, grupos alquilenos, grupos alquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos alquilarilenos, grupos arilalquilenos, grupos arilalquilenos y grupos arilalquilenos están opcionalmente sustituidos y opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo de O, S y NR¹⁹, donde R¹⁹ se selecciona independientemente del grupo que consiste en hidrógeno, grupos alquilo C₁ - C₂₄, grupos alquilenos C₂ - C₂₄, grupos alquilenos C₂ - C₂₄ y grupos cicloalquilo C₃ - C₂₄, donde los grupos alquilo, grupos alquilenos, grupos alquilenos y grupos cicloalquilo están opcionalmente sustituidos.
- En estas formas de realización preferidas, también se prefiere que los grupos alquilenos, grupos alquilenos, grupos alquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos cicloalquilenos, grupos alquilarilenos, grupos arilalquilenos, grupos arilalquilenos y grupos arilalquilenos no estén sustituidos y estén opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo de O, S y NR¹⁹, preferiblemente O, donde R¹⁹ se selecciona independientemente del grupo que consiste en hidrógeno y grupos alquilo C₁ - C₄, preferiblemente hidrógeno o metilo.
- De la manera más preferible, Sp se selecciona del grupo que consiste en grupos alquilenos C₁-C₂₀ lineales o ramificados, estando los grupos alquilenos opcionalmente sustituidos y opcionalmente interrumpidos por uno o más heteroátomos seleccionados del grupo de O, S y NR¹⁹, donde R¹⁹ se selecciona independientemente del grupo que consiste en hidrógeno, grupos alquilo C₁ - C₂₄, grupos alquilenos C₂ - C₂₄, grupos alquilenos C₂ - C₂₄ y grupos cicloalquilo C₃ - C₂₄, donde los grupos alquilo, grupos alquilenos, grupos alquilenos y grupos cicloalquilo están opcionalmente sustituidos. En esta forma de realización, se prefiere adicionalmente que los grupos alquilenos no estén sustituidos y opcionalmente estén interrumpidos por uno o más heteroátomos seleccionados del grupo de O, S y NR¹⁹, preferiblemente O y/o S-S, donde R¹⁹ se selecciona independientemente del grupo que consiste en hidrógeno y grupos alquilo C₁ - C₄, preferiblemente hidrógeno o metilo.
- Los restos de Sp particularmente preferidos incluyen -(CH₂)_n-, -(CH₂CH₂)_n-, -(CH₂CH₂O)_n-, -(OCH₂CH₂)_n-, -(CH₂CH₂O)_nCH₂CH₂-, -CH₂CH₂(OCH₂CH₂)_n-, -(CH₂CH₂CH₂OV, -(OCH₂CH₂CH₂V, -(CH₂CH₂CH₂O)_nCH₂CH₂CH₂- y -CH₂CH₂CH₂(O-CH₂CH₂CH₂)_n-, donde n es un número entero en el intervalo de 1 a 50, preferiblemente en el intervalo

de 1 a 40, más preferiblemente en el intervalo de 1 a 30, incluso más preferiblemente en el intervalo de 1 a 20 y aún más preferiblemente en el intervalo de 1 a 15. Más preferiblemente n es 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10, más preferiblemente 1, 2, 3, 4, 5, 6, 7 u 8, incluso más preferiblemente 1, 2, 3, 4, 5 o 6, aún más preferiblemente 1, 2, 3 o 4.

EJEMPLOS

5 Ejemplo 1. Selección y diseño de GalNAc transferasas

Se seleccionaron cuatro secuencias específicas para la evaluación inicial, en particular el número de depósito Uniprot: Q9GUM2 (*C. elegans*; identificada aquí como SEQ ID N°: 2), U1MEV9 (*A. suum*; identificada aquí como SEQ ID N°: 3), Q6J4T9 (*T. ni*; identificada aquí como SEQ ID N°: 4) y Q7KN92 (*D. melanogaster*, identificada aquí como SEQ ID N°: 5).

10 Los polipéptidos se construyeron basándose en la delección del dominio citoplasmático predicho y el dominio transmembrana. Estos polipéptidos comprenden los anteriormente mencionados *C. elegans* (CeGalNAcT [30-383] identificada por la SEQ ID N°: 6), *A. suum* (AsGalNAcT [30-383] identificada por la SEQ ID N°: 7), *T. ni* (TnGalNAcT [33-421] identificada por la SEQ ID N°: 8) y *D. melanogaster* (DmGalNAcT [47-403] identificada por la SEQ ID N°: 9).

15 Además, se construyeron variantes polipeptídicas que contenían un marcador His en el extremo N- terminal para AsGalNAcT [30-383] (His-AsGalNAcT [30-383] identificada por la SEQ ID N°: 71) y TnGalNAcT [33-421] (His-TnGalNAcT). [33 - 421] identificada por la SEQ ID N°: 49).

Ejemplo 2. Diseño de mutantes de GalNAcT de *T. ni* y mutantes de GalNAcT de *A. suum*

20 Las mutantes de TnGalNAcT y AsGalNAcT se diseñaron basándose en la estructura cristalina de β -(1,4)-Gal-T1 bovina en complejo con UDP-N-acetil-galactosamina (entrada de PDB 1OQM) y la mutante de β -(1,4)-Gal-T1 (Y289L) descrita por Qasba et al. (*J. Biol. Chem.* 2002, 277: 20833 - 20839). Las mutantes de TnGalNAcT y AsGalNAcT se diseñaron basándose en una alineación de secuencia de TnGalNAcT y AsGalNAcT con β -(1,4)Gal-T1 bovina. Los residuos de aminoácidos correspondientes entre estas proteínas se muestran en la Tabla 1.

Tabla 1. Números de aminoácidos correspondientes en diferentes especies de GalNAcT/GalT

TnGalNAcT	AsGalNAcT	β -(1,4)-Gal-T1 bovina
I311	I257	Y289
W336	W282	W314
E339	E285	E317

25 Ejemplo 3. Mutagénesis dirigida de mutantes de His-TnGalNAcT (33-421)

Un vector pET15b que contiene la secuencia codificada optimizada que codifica los residuos 33-421 de TnGalNAcT (identificada por la SEQ ID N°: 8) entre los sitios NdeI-BamHI se obtuvo de Genscript, lo que dio como resultado His-TnGalNAcT (33-421) (identificada por la SEQ. ID NO: 49). Los genes mutantes de TnGalNAcT se amplificaron a partir de la construcción descrita anteriormente usando un conjunto de cebadores solapantes mediante una PCR de amplificación lineal. Los conjuntos de cebadores solapantes utilizados para cada mutante se muestran en la tabla 2. Para la construcción de His-TnGalNAcT (33-421; W336F) (identificada por la SEQ ID N°: 50), el fragmento de ADN se amplificó con un par de cebadores definidos en este documento como SEQ ID N°: 79 y SEQ ID N°: 80. Para la construcción de His-TnGalNAcT (33-421; W336H) (identificada por la SEQ ID N°: 51), el fragmento de DNA se amplificó con un par de cebadores definidos en la presente memoria como SEQ ID N°: 81 y SEQ ID N°: 82. Para la construcción de His-TnGalNAcT (33-421; W336V) (identificada por la SEQ ID N°: 52), el fragmento de ADN se amplificó con un par de cebadores definidos en la presente memoria como SEQ ID N°: 83 y SEQ ID N°: 84. Para la construcción de His-TnGalNAcT (33-421; E339A) (identificada por la SEQ ID N°: 53), el fragmento de DNA se amplificó con un par de cebadores definidos aquí como SEQ ID N°:

85 y SEQ ID N°: 86. Para la construcción de His-TnGalNAcT (33-421; E339G) (identificada por la SEQ ID N°: 54), el fragmento de DNA se amplificó con un par de cebadores definidos en la presente como SEQ ID N°: 87 y SEQ. ID NO: 88. Para la construcción de His-TnGalNAcT (33-421; E339D) (identificada por la SEQ ID N°: 55) el fragmento de ADN se amplificó con un par de cebadores definidos en este documento como SEQ ID N°: 89 y 5 SEQ ID N°: 90. Para la construcción de His-TnGalNAcT (33-421; I311Y) (identificada por la SEQ ID N°: 60) el fragmento de ADN se amplificó con un par de cebadores definidos aquí como SEQ ID N°: 91 y SEQ ID N°: 92. Después de la amplificación por PCR, la mezcla de reacción se trató con DpnI para digerir el molde de ADN seguido de la transformación en células competentes NEB 10-beta (obtenidas de New England Biolabs). Se aisló el ADN y las secuencias se confirmaron por análisis de secuencia para las mutantes His-TnGalNAcT (33-421; W336F) (identificada por la SEQ ID N°: 50), His-TnGalNAcT (33-421; W336V) (identificada por la SEQ ID N°: 52), His-TnGalNAcT (33-421; E339A) (identificada por la

SEQ ID N°: 53) e His-TnGalNAcT (33-421; I311Y) (identificada por la SEQ ID N°: 60).

Tabla 2. Identificación de secuencia de los cebadores utilizados. Los codones correspondientes al aminoácido mutado están en negrita.

SEQ ID N°	Nombre	Secuencia de nucleótidos
SEQ ID N°: 79	W336F,dir	C TCG AAT AAA TAT TGG GGT TTT GGC GGT GAA GAT GAC GAT ATG
SEQ ID N°: 80	W336F,inv	CAT ATC GTC ATC TTC ACC GCC AAA ACC CCA ATA TTT ATT CGA G
SEQ ID N°: 81	W336H,dir	CG AAT AAA TAT TGG GGT CAC GGC GGT GAA GAT GAC G
SEQ ID N°: 82	W336H,inv	C GTC ATC TTC ACC GCC GTG ACC CCA ATA TTT ATT CG
SEQ ID N°: 83	W336V,dir	CG AAT AAA TAT TGG GGT GTG GGC GGT GAA GAT GAC G
SEQ ID N°: 84	W336V,inv	C GTC ATC TTC ACC GCC CAC ACC CCA ATA TTT ATT CG
SEQ ID N°: 85	E339A,dir	G GGT TGG GGC GGT GCG GAT GAC GAT ATG AGC
SEQ ID N°: 86	E339A,inv	GCT CAT ATC GTC ATC CGC ACC GCC CCA ACC C
SEQ ID N°: 87	E339G,dir	G GGT TGG GGC GGT GGA GAT GAC GAT ATG AG
SEQ ID N°: 88	E339G,inv	CT CAT ATC GTC ATC TCC ACC GCC CCA ACC C
SEQ ID N°: 89	E339D,dir	G GGT TGG GGC GGT GAT GAT GAC GAT ATG AGC
SEQ ID N°: 90	E339D,inv	GCT CAT ATC GTC ATC ATC ACC GCC CCA ACC C
SEQ ID N°: 91	I311Y,dir	G CCG TAC GAA GAT TAT TTC GGC GGT GTC TCA G
SEQ ID N°: 92	I311Y,inv	C TGA GAC ACC GCC GAA ATA ATC TTC GTA CGG C

- 5 Ejemplo 4. Expresión y repliegamiento de His-TnGalNAcT (33-421), His-TnGalNAcT (33-421; W336F), His-TnGalNAcT (33-421; W336V) e His-TnGalNAcT (33-421; E339A) en *E. coli*

10 His-TnGalNAcT (33-421), His-TnGalNAcT (33-421; W336F), His-TnGalNAcT (33-421; W336V) e His-TnGalNAcT (33-421; E339A) se expresaron a partir de las correspondientes construcciones pET15b, que se obtienen como se describe en el Ejemplo 3. La expresión, el aislamiento de cuerpos de inclusión y el repliegamiento se realizaron según el procedimiento descrito por Qasba et al. (Prot. Expr. Pur. 2003, 30, 219-76229). Después del repliegamiento, la proteína insoluble se eliminó por centrifugación (10 minutos 8000 x g) seguida de filtración a través de un filtro con poros de 0,45 µm de diámetro. La proteína soluble se purificó y se concentró usando una columna HisTrap HP de 5 ml (GE Healthcare). La columna se lavó primero con tampón A (tampón Tris 20 mM, imidazol 20 mM, NaCl 500 mM, pH 7,5). La proteína retenida se eluyó con tampón B (Tris 20 mM, NaCl 500 mM, imidazol 250 mM, pH 7,5, 10 ml). Las fracciones se analizaron mediante SDS-PAGE en geles de poliacrilamida (12%), y las fracciones que contenían proteína de actuación purificada se combinaron y el tampón se intercambió por Tris 20 mM, pH 7,5 y NaCl 150 mM mediante diálisis realizada durante la noche a 4°C. La proteína purificada se concentró hasta al menos 2 mg/ml usando una membrana Ultracel-10 Amicon Ultra-0.5 (Millipore) y se almacenó a -80°C antes de volver a utilizarse.

Ejemplo 5. Expresión transitoria de GalNAcTs y mutantes en CHO

20 Las proteínas se expresaron transitoriamente en células CHO K1 de Evitria (Zurich, Suiza) en una escala de 20 mL. Se expresaron las siguientes variantes de GalNAcT: CeGalNAcT (30-383) (identificada por la SEQ ID N°: 6), AsGalNAcT (30-383) (identificada por la SEQ ID N°: 7), TnGalNAcT (33-421) (identificada por la SEQ ID N°: 8), DmGalNAcT (47-403) (identificada por la SEQ ID N°: 9) y TnGalNAcT (33-421; E339A) (identificada por la SEQ ID N°: 28). En un experimento de purificación típico, se dializó sobrenadante producido por CHO que contenía la GalNAcT expresada con tampón Tris 20 mM, pH 7,5. El sobrenadante (típicamente 25 ml) se filtró a través de un filtro con poros de 0,45 µm de diámetro y posteriormente se purificó en una columna de intercambio catiónico (columna HiTrap SP HP 5mL, GE Healthcare), que se equilibró con tampón Tris 20 mM, pH 7,5 antes del uso. La purificación se realizó en un sistema de cromatografía AKTA Prime equipado con un colector de fracciones externo. Las muestras se cargaron desde la bomba de sistema A. Las proteínas no unidas se eluyeron de la columna lavando la columna con 10

volúmenes de columna (CV) de tampón Tris 20 mM, pH 7,5. La proteína retenida se eluyó con tampón de elución (Tris 20 mM, 1 NaCl, pH 7,5, 10 ml). Las fracciones recogidas se analizaron mediante SDS-PAGE en geles de poliacrilamida (12%), y las fracciones que contenían la proteína de actuación se combinaron y concentraron usando filtración por centrifugación hasta un volumen de 0,5 mL. Excepto la TnGalNAcT (33-421; E339A), las proteínas se purificaron a continuación en una columna de cromatografía de exclusión por tamaño Superdex200 10/300 GL (GE Healthcare) usando un sistema AKTA purifier-10 (UNICORN v6.3) para obtener las fracciones monoméricas puras. Las fracciones se analizaron mediante SDS-PAGE y las fracciones que contenían la proteína monomérica se almacenaron a -80 °C antes de volver a ser utilizadas.

Protocolo general para el análisis espectral de masas de IgG

Antes del análisis espectral de masas, las IgG se trataron con DTT, que permite el análisis tanto de la cadena ligera como pesada, o se trataron con Fabricator (disponible comercialmente de Genovis, Lund, Suecia), lo que permite el análisis del fragmento Fc/2. Para el análisis tanto de la cadena ligera como pesada, se incubó una solución de 20 pg de IgG (modificada) durante 5 minutos a 37 °C con 100 mM de DTT en un volumen total de 4 µl. Si están presentes, las funcionalidades azida se reducen a aminas en estas condiciones. Para el análisis del fragmento Fc/2, se incubó una solución de 20 pg de IgG (modificada) durante 1 hora a 37 °C con Fabricator (1,25 U/µl) en solución salina tamponada con fosfato (PBS) pH 6.6 en un volumen total de 10 µl. Después de la reducción o de la digestión del fabricante, las muestras se lavaron tres veces con miliQ usando una membrana Amicon Ultra-0.5, Ultracel-10 (Millipore) dando como resultado un volumen de muestra final de aproximadamente 40 µl. A continuación, se analizaron las muestras por tiempo de vuelo con ionización por electrospray (ESI-TOF) en un JEOL AccuTOF. Los espectros desconvolucionados se obtuvieron usando el software Magtran.

Ejemplo 6. Preparación de trastuzumab recortado mediante tratamiento con endo S.

El recorte de glucanos de trastuzumab se realizó con endo S de *Streptococcus pyogenes* (disponible comercialmente en Genovis, Lund, Suecia). Por lo tanto, se incubó trastuzumab (10 mg/ml) con endo S (40 U / ml) en Tris 25 mM, pH 8,0 durante aproximadamente 16 horas a 37°C. La IgG desglucosilada se concentró y se lavó con 10 mM de MnCl₂ y 25 mM de Tris-HCl pH 8,0 usando una membrana Amicon Ultra-0.5, Ultracel-10 (Millipore). Después de la desconvolución de los picos, el espectro de masas mostró un pico de la cadena ligera y dos picos de la cadena pesada. Los dos picos de la cadena pesada pertenecían a un producto principal (49496 Da, 90% de la cadena pesada total), que resulta del núcleo de trastuzumab sustituido por GlcNAc (Fuc) y un producto menor (49351 Da, ± 10% de la cadena pesada total), que resulta del trastuzumab recortado.

Ejemplo 7. Transferencia de residuos de glucosa de 6-azido-Gal-UDP a trastuzumab recortado bajo la acción de β-(1,4)-Gal-T1 bovina

El trastuzumab recortado (10 mg/ml), obtenido mediante tratamiento con endo S de trastuzumab como se ha descrito anteriormente se incubó con 6-azido-Gal-UDP (1 mM, comercialmente disponible de GlycoHub) y 0,1 o 0,5 mg/ml de β-(1,4) Gal-T1 bovina (disponible comercialmente de Sigma Aldrich) en MnCl₂ 10 mM y Tris-HCl 25 mM, pH 8,0 a 30 °C durante la noche. El análisis espectral de masas de las muestras reducidas indicó que no hubo formación de producto para las dos concentraciones de β-(1,4)-Gal-T1 bovina (pico de cadena pesada principal de 49494 Da, 90% de la cadena pesada total, resultante del núcleo de trastuzumab sustituido por GlcNAc (Fuc)).

Ejemplo 8. Transferencia de residuos de glucosa de la 6-azido-N-acetilgalactosamina-UDP a trastuzumab recortado bajo la acción de β-(1,4)-Gal-T1 bovina (130-402; Y289L, C342T)

Se usó un mutante derivado de β-(1,4)-Gal-T1 bovina (identificado por la SEQ ID N°: 1) que contenía las mutaciones Y289L y C342T y que contenía solo el dominio catalítico (residuos de aminoácidos 130-402). Este mutante β-(1,4)-Gal-T1 bovino (130-402; Y289L, C342T) se describe por Qasba et al. (J. Biol. Chem. 2002, 277, 20833 - 20839) y se expresó, aisló y replegó a partir de cuerpos de inclusión según el procedimiento descrito por Qasba et al. (Prot. Expr. Pur. 2003, 30, 219-76229). El trastuzumab recortado (10 mg/ml), obtenido mediante tratamiento con endoS de trastuzumab como se ha descrito anteriormente, se incubó con 6-azido-GalNAc-UDP (2,5 mM, comercialmente disponible de GlycoHub) y 1 mg / ml β-(1,4) -Gal-T1 (130-402; Y289L, C342T) en MnCl₂ 10 mM y Tris-HCl 25 mM pH 7,5 a 37 °C durante la noche. El análisis espectral de masas de las muestras reducidas indicó que no hubo formación de producto (pico de cadena pesada principal de 49502 Da, 90% de la cadena pesada total, que resulta del núcleo de trastuzumab sustituido con GlcNAc (Fuc)).

Ejemplo 9. Transferencia de residuos de glucosa de la 6-azido-N-acetilgalactosamina-UDP a trastuzumab recortado bajo la acción de varias GalNAcT

Se sometió a prueba la incorporación de 6-azidoGalNAc para CeGalNAcT (30-383) (identificada por la SEQ ID N°: 6), AsGalNAcT (30-383) (identificada por la SEQ ID N°: 7), TnGalNAcT (33-421) (identificada por la SEQ. ID NO: 8) y DmGalNAcT (47-403) (identificada por la SEQ ID N°: 9), que se expresaron y purificaron como se describe en el ejemplo 5. Se incubó trastuzumab recortado (10 mg/ml), obtenido mediante tratamiento con endoS de trastuzumab como se ha descrito anteriormente, con 6-azido-GalNAc-UDP (1 mM, disponible comercialmente de GlycoHub) en MnCl₂ 10 mM y Tris-HCl 25 mM pH 7,5 y 0,2 o 0,5 mg/ml de una de las GalNAcT mencionadas anteriormente.

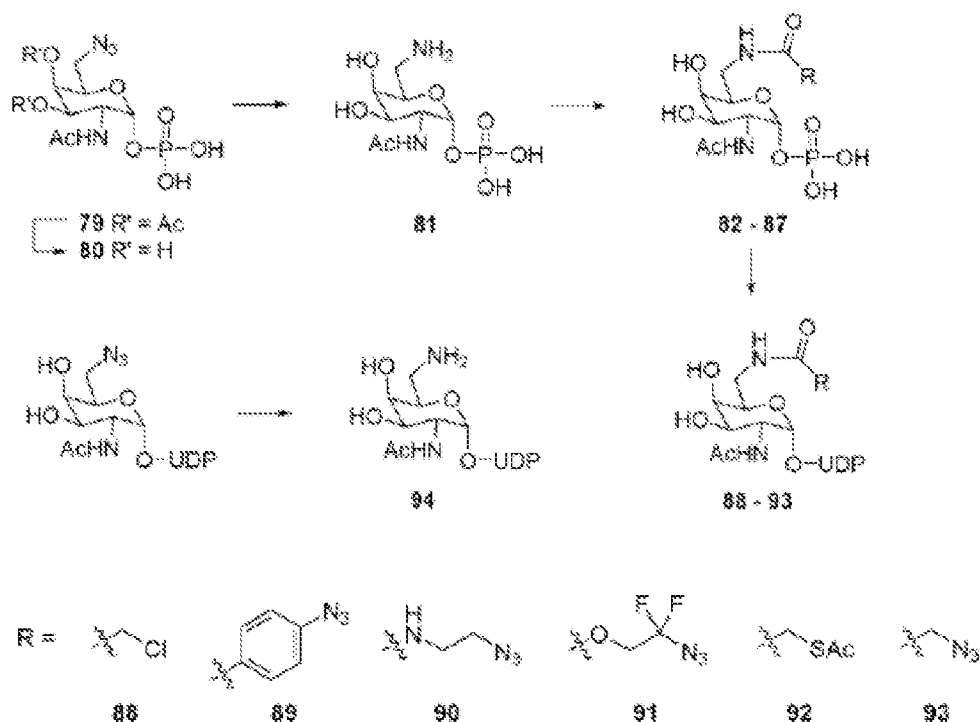
5

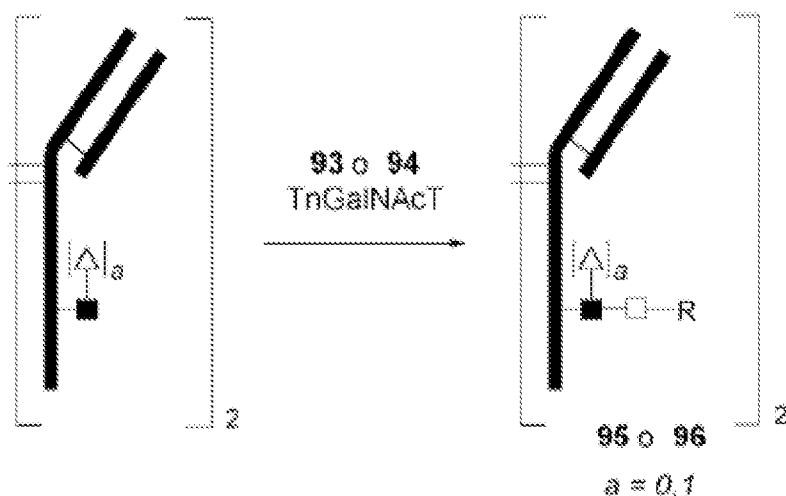
El análisis espectral de masas de las muestras recopiladas por Fabricator indicó que no se formaron productos para las dos concentraciones de CeGalNAcT (30-383) y DmGalNAcT (47-403) (pico principal Fc/2 de 24139 Da, 90% de la cadena pesada total, resultante del núcleo de trastuzumab sustituido con GlcNAc(Fuc)), mientras que tanto AsGalNAcT (30-383) como TnGalNAcT (33-421) mostraron una conversión parcial del núcleo de trastuzumab sustituido con GlcNAc(Fuc) (masa observada de 24139 Da) en el producto (masa observada de 24366 Da), que resulta de la transferencia de 6-azido-GalNAc al núcleo de trastuzumab sustituido con GlcNAc (Fuc). Las conversiones obtenidas se muestran en la Tabla 3.

Tabla 3. Conversiones (%) de trastuzumab sustituido con GlcNAc (Fuc) en trastuzumab sustituido con 6-Azido-GalNAc-GlcNAc (Fuc) por varias GalNAcT a diversas concentraciones enzimáticas.

	0,2 mg/ml de enzima	0,5 mg /ml de enzima
CeGalNAcT (30-383)	0	0
DmGalNAcT (47-403)	0	0
AsGalNAcT (30-383)	10	30
TnGalNAcT (33-421)	40	60

10





Esquema 1: Síntesis de los compuestos 88 - 94 y de las glucoproteínas modificadas 95 - 96 (Ejemplo 10 - 26)

Ejemplo 10. Síntesis de 6-azido-6-desoxi-GalNAc-1-monofosfato 80

El azúcar acetilado 79 puede prepararse según el procedimiento de Wang et al, Bioorg. Med. Chem. Lett., 2009, 19, 5433.

A una suspensión del azúcar acetilado 79 (4,9 g, 11,9 mmol) en MeOH (15 mL) se añadió NH₄OH en solución al 25% (60 mL). La reacción se dejó en agitación a temperatura ambiente y la conversión se analizó con LCMS. Después de 4 h, la mezcla se concentró bajo presión reducida y se almacenó a -20°C durante 2 días. El sólido se volvió a disolver en NH₄OH en solución al 25% (75 ml), se agitó a temperatura ambiente y, después de 3 h, la espectrometría de masas mostró una conversión completa. La concentración del disolvente proporcionó el producto bruto como un sólido amarillo. Se realizó una RMN cuantitativa, que mostró que el 60% en peso fue el producto 80, dando un rendimiento de 3,2 g (83%).

¹H-RMN (400 MHz, D₂O): δ 5,28 (dd, J=7,2, 3,2 Hz, 1H), 4,12 (dd, J=6,8; 6,4 Hz, 1H), 4,06 (ddd, J=10,8, 3,2, 2,0 Hz, 1H), 3,92-3,81 (m, 2H), 3,47 (dd, J=12,8, 7,2 Hz, 1H), 3,40 (dd, J=12,8 Hz, 6,4 Hz, 1H), 1,88 (s, 3H). LRMS (ESI-) calculado para C₈H₁₅N₄O₈P (M⁺H⁺) 325,06, resultado 325,30.

Ejemplo 11. Síntesis de 6-amino-6-desoxi-GalNAc-1-monofosfato 81

A una solución de azida 80 (5,9 mmol) en H₂O (30 ml) y MeOH (30 ml) se añadió Pd/C (400 mg) y se hizo burbujear H₂ a través de la mezcla de reacción durante 1 h. La conversión de la reacción se controló con TLC (MeOH 7: 3: MeCN). La mezcla de reacción se filtró sobre celite (diatomita), se enjuagó a fondo con MeOH y H₂O y se concentró in vacuo para lograr el producto bruto 81 con un rendimiento de 1,8 g (99%).

¹H-RMN (400 MHz, D₂O): δ 5,28 (dd, J=7,2; 3,6 Hz, 1H), 4,25 (dd, J=8,8; 4,0 Hz, 1H), 4,09-4,04 (m, 1H), 3,903,79 (m, 2H), 3,19-3,08 (m, 2H), 1,85 (s, 3H). LRMS (ESI-) calculado para C₈H₁₇N₂O₈P (M⁺H⁺) 299,06, resultado 229,29.

Ejemplo 12. Síntesis de 6-(2-cloroacetamido)-6-desoxi-GalNAc-1-monofosfato 82

El éster de succinimidilo de ácido cloroacético se preparó según el procedimiento de Hosztafi et al, Helv. Chim. Acta, 1996, 79, 133.

A una solución de azúcar 81 (12 mg, 0,040 mmol) en DMF seca (0,5 ml) en una atmósfera de nitrógeno se añadieron éster de succinimidilo de ácido cloroacético (9 mg, 0,044 mmol) y Et₃N (6,7 μ l, 0,048 mmol). La mezcla de reacción se dejó en agitación a temperatura ambiente y se concentró in vacuo para lograr el producto bruto 82.

¹H-RMN (400 MHz, D₂O): δ 5,42-5,32 (m, 1H), 4,13-4,02 (m, 4H), 3,92-3,81 (m, 2H), 3,53-3,46 (m, 1H), 3,33-3,26 (m, 1H), 1,94 (s, 3H). LRMS (ESI-) calculado para C₁₀H₁₈ClN₂O₉P (M⁺H⁺) 375,68 (100%), 377,03 (30%), resultado 3,75,08 (100%), 377,19 (25%).

Ejemplo 13. Síntesis de 6-(4-azidobenzamido)-6-desoxi-GalNAc-1-monofosfato 83

Se preparó éster de succinimidilo de ácido 4-azidobenzoico según Hartman et al, Chem. Comm., 2012, 48, 4755.

A una solución de azúcar 81 (38 mg, 0,127 mmol) en DMF seca (1,5 ml) en una atmósfera de nitrógeno se añadieron Et₃N (21 µl, 0,152 mmol) y éster de succinimidilo de ácido 4-azidobenzoico (36 mg, 0,139 mmol) y la mezcla de reacción se agitó a temperatura ambiente. Se añadieron más ácido de succinimidilo de ácido 4-azidobenzoico (36 mg, 0,139 mmol) y Et₃N (42 µl, 0,304 mmol) y la reacción se dejó en agitación durante 5 días a temperatura ambiente. La formación del producto se controló con TLC y MS. La mezcla de reacción se concentró después de 6 días para lograr el producto bruto 83.

LRMS(ESI-) calculada para C₁₅H₂₀N₅O₉P (M·H⁺) 444,09, resultado 444,20.

Ejemplo 14. Síntesis de 6-(N-2-azido-2,2-difluoroetilcarbamato)-6-desoxi-GalNAc-1-monofosfato 84

Se preparó 2-azido-2,2-difluoroetanol según el procedimiento descrito en WO2015/112016.

Se disolvió 2-azido-2,2-difluoroetanol (200 mg, 1,63 mmol) en DCM (10 ml) en atmósfera de nitrógeno, se añadieron cloroformiato de 4-nitrofenilo (295 mg, 1,46 mmol) y Et₃N (226 µl, 1,63 mmol) y la mezcla resultante se agitó a temperatura ambiente durante 1 h. Luego se disolvió azúcar 81 (122 mg, 0,41 mmol) en H₂O (2 ml), se añadieron Et₃N (113 µl, 0,81 mmol) y DMF (5 mL) y la solución resultante se añadió a la mezcla de reacción. La reacción se dejó en agitación a temperatura ambiente durante 16 h, cuando la TLC y la LCMS confirmaron el consumo total del azúcar 81. El disolvente se eliminó bajo presión reducida para proporcionar el producto bruto. La purificación se realizó con cromatografía de intercambio iónico (Q HTRAP, columnas de 3 x 5 ml y 1 x 15 ml). La primera unión en la columna se logró mediante carga con tampón A (NH₄HCO₃ 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente a B al 40% (NH₄HCO₃ 250 mM) para eluir el producto y la columna se lavó con B al 100% para eliminar los subproductos restantes. Las fracciones que contienen el producto se liofilizaron para proporcionar el producto deseado 84 (147 mg, 0,33 mmol, 80%).

¹H-RMN (400 MHz, D₂O): δ 5,17 (dd, J=6,4; 3,2 Hz, ¹H), 4,40-4,24 (m, 2H), 4,07-3,93 (m, 2H), 3,85-3,70 (m, 2H), 3,28-3,13 (m, 2H), 1,87 (s, 3H) LRMS (ESI-) se calculó para C₁₁H₁₈F₂N₅O₁₀P (M·H⁺) 448,07, resultado 448,14.

Ejemplo 15. Síntesis de 6-(N-1-(2-azidoetil)urea)-6-desoxi-GalNAc-1-monofosfato 85

Se preparó 2-azidoetilamina según el procedimiento descrito en Zhang et al., J. Am. Chem. Soc., 2015, 137, 6000. ¹H-NMR (400 MHz, CDCl₃): δ 3,40-3,33 (m, 2H), 2,91-2,81 (m, 2H).

Se disolvió carbonildiimidazol (377 mg, 2,32 mmol) en DMF seca (10 ml) y se agitó en una atmósfera de nitrógeno. Se disolvió 2-azidoetilamina (200 mg, 2,32 mmol) en DMF seca (5 ml) y se añadió gota a gota al CDI. La solución resultante se agitó durante 1 hora a temperatura ambiente, seguido de calentamiento a 60°C. El azúcar 81 se disolvió en H₂O (2 ml) y DMF (5 ml) y se añadió a la reacción. La suspensión resultante se agitó durante 16 h a 60°C. La formación del producto deseado se controló con LCMS. Después de agitar durante 16 h, se añadió H₂O (5 ml), seguido de la adición de 2-azidoetilamina recién activada en DCM (15 ml). La mezcla resultante se agitó de nuevo durante 16 h a 60°C y el disolvente se eliminó bajo presión reducida. El producto bruto se disolvió luego en MeOH (10 ml) y H₂O (15 ml) y se lavó con EtOAc (2 x 30 ml). La fase acuosa se concentró para lograr el producto bruto, que se purificó con cromatografía de intercambio iónico (Q-HITRAP, columnas de 3 x 5 ml y 1 x 15 ml). La primera unión en la columna se logró mediante la carga con tampón A (NH₄HCO₃ 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente de B al 40% (NH₄HCO₃ 250 mM) para eluir el producto y las columnas se enjuagaron con B al 100% para eliminar los subproductos restantes. Las fracciones que contenían el producto se liofilizaron para proporcionar el producto deseado 85 (147 mg, 0,33 mmol, 80%).

¹H-RMN (400 MHz, D₂O): δ 5,30 (br s, 1H), 4,13 - 3,93 (m, 2H), 3,91 - 3,76 (m, 2H), 3,35 - 3,10 (m, 5H), 1,92 (s, 3H). LRMS (ESI-) se calculó para C₁₁H₂₁N₆O₉P (M·H⁺) 411,10, resultado 411,24.

Ejemplo 16. Síntesis de 6-(N-(2-S-acetil)-mercaptoacetamido)-6-desoxi-GalNAc-1-monofosfato 86

El azúcar 81 (105 mg, 0,35 mmol) se disolvió en H₂O (1,7 ml) y se agitó a temperatura ambiente. En un vial separado se disolvió éster de pentafluorfenilo de ácido S- acetiltioglicólico (210 mg, 0,70 mmol) en DMF (1,7 ml) y se añadió a la reacción, junto con Et₃N (146 µl, 1,05 mmol). La mezcla resultante se agitó durante 16 h a temperatura ambiente cuando LCMS mostró el consumo total de azúcar 81. El disolvente se eliminó bajo presión reducida y el producto en bruto se purificó por cromatografía ultrarrápida (6:2:1-4:2:1 EtOAc:MeOH:H₂O) para lograr el producto 86 (95 mg, 0,23 mmol, 65%).

¹H-RMN (400 MHz, D₂O): δ 5,33 (dd, J=6,8, 3,6 Hz, 1H), 4,11-4,00 (m, 2H), 3,90-3,80 (m, 2H), 3,63-3,53 (m, 2H), 3,44 (dd, J=14,0, 5,2 Hz, 1H), 3,23 (dd, J=14,0, 8,0 Hz, 1 H), 2,31 (s, 3 H), 1,94 (s, 3 H). LRMS (ESI-) se calculó para C₁₂H₂₁N₂O₁₀PS (M·H⁺) 415,06, resultado 415,18.

Ejemplo 17. Síntesis de 6-(N-2-azidoacetamido)-6-desoxi-GalNAc-1-monofosfato 87

Se disolvió ácido azidoacético (101 mg, 1,0 mmol) en DMF (2 ml) y se añadieron EDC (192 mg, 1,0 mmol), NHS (115 mg, 1,0 mmol) y DMAP (4 mg, 0,03 mmol). Luego, el azúcar 81 (100 mg, 0,33 mmol) se disolvió en H₂O (3 mL), se añadió a la reacción y se agitó durante 16 h a temperatura ambiente. La formación del producto deseado se controló

con LCMS. Otra porción de ácido azidoacético se activó como se ha descrito anteriormente y se añadió a la reacción. Después de 4 h, la reacción se concentró in vacuo. La purificación se realizó con cromatografía de intercambio iónico (Q-HITRAP, columnas de 3 x 5 ml y 1 x 15 ml). La primera unión a la columna se logró mediante carga con tampón A (NH_4HCO_3 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente a tampón B al 40% (NH_4HCO_3 250 mM) para eluir el producto y la columna se lavó abundantemente con tampón B al 100% para eliminar los subproductos restantes. Las fracciones que contenían el producto se liofilizaron para proporcionar el producto deseado 87 (100 mg, 0,26 mmol, 79%).

^1H -RMN (400 MHz, D_2O): δ 5,34 (br s, 1H), 4,13-4,03 (m, 2H), 3,93 (s, 2H), 3,92-3,81 (m, 2H), 3,48 (dd, $J=14,0$; 4,0 Hz, 1H), 3,29 (dd, $J=14,0$, 8,0 Hz, 1H), 1,95 (s, 3H). LRMS (ESI-) se calculó para $\text{C}_{10}\text{H}_{18}\text{N}_5\text{O}_9\text{P}$ (M^+H^-) 382,08, resultado 382,15.

Ejemplo 18. Síntesis de 6-(2-cloroacetamido)-6-desoxi-GalNAc-UDP 88

Se acopló monofosfato 82 a UMP según un procedimiento descrito por Baisch et al. *Bioorg. Med. Chem.*, 1997, 5, 383.

En resumen, se disolvió uridina-5'-monofosfato de tributilamonio (31 mg, 0,06 mmol) en DMF seca (0,5 ml) en una atmósfera de nitrógeno. Se añadió carbonildiimidazol (13 mg, 0,04 mmol) y la mezcla de reacción se agitó a temperatura ambiente durante 30 min. A continuación, se añadió MeOH seco (2,5 μl) y se agitó durante 15 minutos para eliminar el exceso de CDI. El MeOH sobrante se eliminó bajo alto vacío durante 15 minutos. Posteriormente, el monofosfato 82 (15 mg, 0,04 mmol) se disolvió en DMF seca (0,5 ml) y se añadió a la mezcla de reacción, seguido de N-metilimidazol, sal de HCl (25 mg, 0,16 mmol). La reacción se dejó en agitación a temperatura ambiente durante la noche antes de la concentración in vacuo. El consumo de la sustancia intermedia de monofosfato se controló por MS. La purificación se realizó con cromatografía de intercambio iónico (Q-HITRAP, columna de 1 x 5 ml). La primera unión a la columna se logró mediante carga con tampón A (NH_4HCO_3 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente a tampón B al 40% (NH_4HCO_3 250 mM) para eluir el producto y la columna se lavó abundantemente con tampón B al 100% para eliminar los subproductos restantes. Las fracciones que contenían el producto se liofilizaron para lograr el producto deseado 88 (1 mg, 1,46 pmol, 4%).

LRMS (ESI-) se calculó para $\text{C}_{19}\text{H}_{29}\text{ClN}_4\text{O}_{17}\text{P}_2$ (M^+H^-) 681,06 (100%), 683,06 (32%), resultado 681,13 (100%), 683,15 (40%).

Ejemplo 19. Síntesis de 6-(4-azidobenzamido)-6-desoxi-GalNAc-UDP 89

Se acopló monofosfato 83 a UMP según un procedimiento descrito por Baisch et al. *Bioorg. Med. Chem.*, 1997, 5, 383.

En resumen, se disolvió uridina-5'-monofosfato de tributilamonio (77 mg, 0,15 mmol) en DMF seca (1 ml) en una atmósfera de nitrógeno. Se añadió carbonildiimidazol (41 mg, 0,25 mmol) y la mezcla de reacción se agitó a temperatura ambiente durante 30 min. A continuación, se añadió MeOH seco (6,2 μl) y se agitó durante 15 minutos para eliminar el exceso de CDI. El MeOH sobrante se eliminó bajo alto vacío durante 15 minutos. Posteriormente, el monofosfato 83 (56 mg, 0,13 mmol) se disolvió en DMF seca (1 ml) y se añadió a la mezcla de reacción, seguido de N-metilimidazol, sal de HCl (79 mg, 0,51 mmol). La reacción se dejó en agitación a temperatura ambiente durante la noche antes de la concentración in vacuo. El consumo de la sustancia intermedia de monofosfato fue controlado por MS. La purificación se realizó con cromatografía de intercambio iónico (Q-HITRAP, columnas de 3 x 5 ml, columna de 1 x 15 ml). La primera unión a la columna se logró mediante carga con tampón A (NH_4HCO_3 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente de tampón B al 40% (NH_4HCO_3 250 mM) para eluir el producto y la columna se lavó con B al 100% de para eliminar los subproductos restantes. Las fracciones que contenían el producto se liofilizaron para proporcionar el producto deseado 89 (13 mg, 0,017 mmol, 14%).

LRMS (ESI-) se calculó para $\text{C}_{24}\text{H}_{31}\text{N}_7\text{O}_{17}\text{P}_2$ (M^+H^-) 750,12, resultado 750,33.

Ejemplo 20. Síntesis de 6-(N-2-azido-2,2-difluoroetilcarbamato)-6-desoxi-GalNAc-UDP 90

Se acopló monofosfato 84 a UMP según un procedimiento descrito por Baisch et al. *Bioorg. Med. Chem.*, 1997, 5, 383.

En resumen, se disolvió uridina-5'-monofosfato de tributilamonio (200 mg, 0,39 mmol) en DMF seca (3 ml) en una atmósfera de nitrógeno. Se añadió carbonildiimidazol (106 mg, 0,65 mmol) y la mezcla de reacción se agitó a temperatura ambiente durante 30 min. A continuación, se añadió MeOH seco (16 μl) y se agitó durante 15 minutos para eliminar el exceso de CDI. El MeOH sobrante se eliminó bajo alto vacío durante 15 minutos. Posteriormente, el monofosfato 84 (147 mg, 0,33 mmol) se suspendió en DMF seca (3 ml) y se añadió a la mezcla de reacción, seguido de N-metilimidazol, sal de HCl (204 mg, 1,31 mmol). El consumo de la sustancia intermedia de monofosfato se controló por MS. La reacción se dejó en agitación a temperatura ambiente durante 3 días. Otra porción de UMP se activó como se ha descrito anteriormente y se agregó a la reacción junto con 1 ml de H_2O . Después de agitar durante la noche, la reacción se completó y el disolvente se eliminó bajo presión reducida. La purificación se realizó con cromatografía de intercambio iónico (Q-HITRAP, columnas de 3 x 5 ml, columna de 1 x 15 ml). La primera unión a la columna se logró mediante carga con tampón A (NH_4HCO_3 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente de tampón B al 40% (NH_4HCO_3 250 mM) para eluir el producto y la columna se lavó abundantemente con tampón B al 100% para eliminar los subproductos restantes. Las fracciones que contenían el producto se

liofilizaron para proporcionar el producto deseado 90 (122 mg, 0,16 mmol, 49%).

LRMS (ESI-) se calculó para $C_{20}H_{29}F_2N_7O_{18}P_2$ ($M^{\cdot}H^+$) 754,09, resultado 754,16.

Ejemplo 21. Síntesis de 6-(N-1-(2-azidoetil)urea)-6-desoxi-GalNAc-UDP 91

Se acopló monofosfato 85 a UMP según un procedimiento descrito por Baisch et al. Bioorg. Med. Chem., 1997, 5, 383.

- 5 En resumen, se disolvió uridina-5'-monofosfato de tributilamonio (126 mg, 0,25 mmol) en DMF seca (2 ml) en una atmósfera de nitrógeno. Se añadió carbonildiimidazol (67 mg, 0,41 mmol) y la mezcla de reacción se agitó a temperatura ambiente durante 30 min. A continuación, se añadió MeOH seco (10 μ l) y se agitó durante 15 minutos para eliminar el exceso de CDI. El MeOH sobrante se eliminó bajo alto vacío durante 15 minutos. Posteriormente, el monofosfato 85 (85 mg, 0,21 mmol) se disolvió en DMF seca (2 ml) y se añadió a la mezcla de reacción, seguido de N-metilimidazol, sal de HCl (129 mg, 0,82 mmol). La reacción se dejó en agitación a temperatura ambiente durante 2 días antes de la concentración in vacuo. El consumo de la sustancia intermedia de monofosfato se controló por MS. La purificación se realizó con cromatografía de intercambio iónico (Q HITRAP, columnas de 3 x 5 ml, columna de 1 x 15 ml). La primera unión a la columna se logró mediante carga con tampón A (NH_4HCO_3 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente de tampón B al 40% (NH_4HCO_3 250 mM) para eluir el producto y la columna se lavó abundantemente con tampón B al 100% para eliminar los subproductos restantes. Las fracciones que contenían el producto se liofilizaron para lograr el producto deseado 91 (83 mg, 0,12 mmol, 56%).

LRMS (ESI-) se calculó para $C_{20}H_{32}N_8O_{17}P_2$ ($M^{\cdot}H^+$) 717,13, resultado 717,27.

Ejemplo 22. Síntesis de 6-(N-(2-S-acetil)-mercaptoacetamido)-6-desoxi-GalNAc-UDP 92

Se acopló monofosfato 86 a UMP según un procedimiento descrito por Baisch et al. Bioorg. Med. Chem., 1997, 5, 383.

- 20 En resumen, se disolvió uridina-5'-monofosfato de tributilamonio (139 mg, 0,27 mmol) en DMF seca (2 ml) en una atmósfera de nitrógeno. Se añadió carbonildiimidazol (74 mg, 0,46 mmol) y la mezcla de reacción se agitó a temperatura ambiente durante 30 min. A continuación, se añadió MeOH seco (11 μ l) y se agitó durante 15 minutos para eliminar el exceso de CDI. El MeOH sobrante se eliminó bajo alto vacío durante 15 minutos. Posteriormente, el monofosfato 86 (95 mg, 0,27 mmol) se disolvió en DMF seca (2 ml) y se añadió a la mezcla de reacción, seguido de N-metilimidazol, sal de HCl (142 mg, 0,91 mmol). La reacción se dejó en agitación a temperatura ambiente durante 3 días antes de la concentración in vacuo. El consumo de la sustancia intermedia de monofosfato se controló por MS. La purificación se realizó con cromatografía ultrarrápida (7:2:1-4:2:1 EtOAc:MeOH:H₂O) para lograr el producto 92 (97 mg, 0,13 mmol, 49%).

LRMS (ESI-) se calculó para $C_{21}H_{32}N_4O_{18}P_2S$ ($M^{\cdot}H^+$) 721,08, encontrado 721,39.

30 Ejemplo 23. Síntesis de 6-(2-azidoacetamido)-6-desoxi-GalNAc-UDP 93

Se acopló monofosfato 87 a UMP según un procedimiento descrito por Baisch et al. Bioorg. Med. Chem., 1997, 5, 383.

- 35 En resumen, se disolvió uridina-5'-monofosfato de tributilamonio (191 mg, 0,38 mmol) en DMF seca (3 ml) en una atmósfera de nitrógeno. Se añadió carbonildiimidazol (102 mg, 0,63 mmol) y la mezcla de reacción se agitó a temperatura ambiente durante 30 min. A continuación, se añadió MeOH seco (16 μ l) y se agitó durante 15 minutos para eliminar el exceso de CDI. El MeOH sobrante se eliminó bajo alto vacío durante 15 minutos. Posteriormente, el monofosfato 87 (120 mg, 0,31 mmol) se suspendió en DMF seca (3 ml) y se añadió a la mezcla de reacción, seguido de N-metilimidazol, sal de HCl (195 mg, 1,25 mmol). El consumo de la sustancia intermedia de monofosfato se controló por MS. La reacción se dejó en agitación a temperatura ambiente durante 16 h. Para disolver todos los componentes en la reacción, se añadió 1 ml de H₂O. Después de agitación durante 3 h, el disolvente se eliminó bajo presión reducida. La purificación se realizó con cromatografía de intercambio iónico (Q-HITRAP, columnas de 3 x 5 ml, columna de 1 x 15 ml). La primera unión a la columna se logró mediante carga con tampón A (NH_4HCO_3 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente de tampón B al 40% (NH_4HCO_3 250 mM) para eluir el producto y la columna se lavó abundantemente con tampón B al 100% para eliminar los subproductos restantes. Las fracciones que contenían el producto se liofilizaron para lograr el producto deseado 93 (10 mg, 0,015 mmol, 5%).

45 LRMS (ESI-) se calculó para $C_{19}H_{29}N_7O_{17}P_2$ ($M^{\cdot}H^+$) 688,10, resultado 688,10.

Ejemplo 24. Síntesis de 6-amino-6-desoxi-GalNAc-UDP 94

- 50 A una solución de 6-azido-GalNAc-UDP (25 mg, 0,04 mmol) en H₂O (0,5 ml) se añadieron DTT (6 mg, 0,04 mmol) y unas gotas de Et₃N. La reacción se agitó a temperatura ambiente durante 2 h, y se siguió con LCMS. Para acelerar la reacción, se añadió DTT adicional (12 mg, 0,08 mmol), y después de 1 h la reacción se completó y se concentró in vacuo. La purificación se realizó con cromatografía de intercambio iónico (Q-HITRAP, columna de 1 x 5 ml). La primera unión a la columna se logró mediante carga con tampón A (NH_4HCO_3 10 mM) y la columna se enjuagó con tampón A. A continuación, se realizó un gradiente de tampón B al 40% (NH_4HCO_3 250 mM) para eluir el producto y la columna se lavó abundantemente con tampón B al 100% para eliminar los 15 subproductos restantes. Las fracciones que

contenían el producto se liofilizaron para lograr el producto deseado 94 (12 mg, 0,019 mmol, 51%).

LRMS (ESI-) se calculó para $C_{17}H_{28}N_4O_{16}P_2$ ($M-H^+$) 605,09, resultado 605,11.

Ejemplo 25. Preparación de brentuximab-(6-amino-6-desoxi-GalNAc), 95

Bentuximab fue recortado de forma análoga al recorte de trastuzumab como se describe en el ejemplo 6.

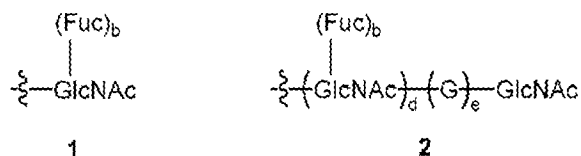
- 5 El bentuximab recortado (15 mg/mL) se incubó con 6-amino-6-desoxi-GalNAc-UDP 94 (5 mM) y TnGalNAcT (1,5 mg/ml) en $MnCl_2$ 10 mM y Tris-HCl 25 mM, pH 8,0 a 30 °C durante la noche. Una muestra de la mezcla de reacción (2 μ l) se incubó durante 1 hora a 37 °C con Fabricator (1,25 U / μ l) en solución salina 25 tamponada con fosfato (PBS) pH 6.6 en un volumen total de 10 μ l. El análisis de espectrometría de masas de esta muestra mostró una conversión completa al producto brentuximab-(6-amino-GalNAc) (24307 Da (70%) y 24435 (30%, variante de lisina C-terminal)).
- 10 Ejemplo 26. Preparación de brentuximab-(6-(2-azidoacetamido)-6-desoxi-GalNAc), 96
Se incubó bentuximab (15 mg/ml) con 6-(2-azidoacetamido)-6-desoxi-GalNAc 93 (5 mM) y TnGalNAcT (1,5 mg/ml) en $MnCl_2$ 10 mM y Tris-HCl 25 mM, pH 8,0 a 30 °C durante la noche. Una muestra de la mezcla de reacción (2 μ l) se incubó durante 1 hora a 37 °C con Fabricator (1,25 U / μ l) en solución salina tamponada con fosfato (PBS) pH 6,6 en un volumen total de 10 μ l. El análisis de espectrometría de masas de esta muestra mostró una conversión del 70% al producto brentuximab-(6-(2-azidoacetamido)-6-desoxi-GalNAc) (24391 Da 35 (70%) y 24518 (30%, variante de lisina C-terminal)).
- 15

REIVINDICACIONES

1. Proceso para la modificación de una glucoproteína, proceso que comprende poner en contacto una glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal, con un nucleótido de azúcar Su(A)-Nuc en presencia de una glucosiltransferasa, donde:

- 5 (i) la glucosiltransferasa es una β -(1,4)-N-acetilgalactosaminiltransferasa que comprende una secuencia que tiene al menos un 95% de identidad de secuencia con una cualquiera de las SEQ. ID NO: 3, SEQ. ID NO: 4, SEQ. ID NO: 7 y SEQ. ID NO: 8;

(ii) el glucano que comprende un resto de GlcNAc terminal es según la fórmula (1) o (2):



10 donde:

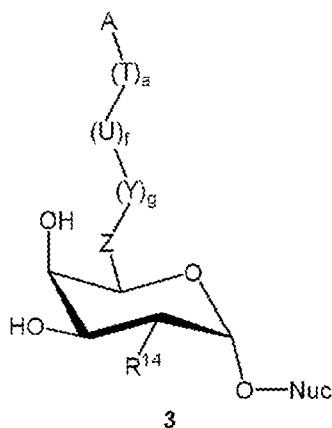
b es 0 o 1;

d es 0 o 1;

e es 0 o 1; y

G es un monosacárido, o un oligosacárido lineal o ramificado que comprende de 2 a 20 restos de azúcar; y

- 15 (iii) el nucleótido de azúcar Su(A)-Nuc es según la fórmula (3):



donde:

a es 0 o 1;

f es 0 o 1;

- 20 g es 0 o 1;

Nuc es un nucleótido;

U es $[\text{C}(\text{R}^1)_2]_n$ o $[\text{C}(\text{R}^1)_2]_p\text{--O--}[\text{C}(\text{R}^1)_2\text{C}(\text{R}^1)_2\text{O}]_o\text{--}[\text{C}(\text{R}^1)_2]_q$, donde "n" es un número entero en el intervalo de 1 a 24; "o" es un número entero en el intervalo de 0 a 12; "p" y "q" son independientemente 0, 1 o 2; y R^1 se selecciona independientemente del grupo que consiste en H, F, Cl, Br, I, OH y un grupo alquilo C_1 - C_{24} opcionalmente sustituido;

- 25 T es un grupo (hetero)arileno C_3 - C_{12} , en donde el grupo (hetero)arileno está opcionalmente sustituido;

A se selecciona del grupo que consiste en:

(a) -N_3

(b) -C(O)R^3

donde R^3 es un grupo alquilo C_1 - C_{24} opcionalmente sustituido;

(c) grupo (hetero)cicloalquinilo o un resto $-(CH_2)_iC\equiv C-R^4$

donde i es 0 - 10 y R^4 es hidrógeno o un grupo alquilo $C_1 - C_{24}$ opcionalmente sustituido;

(d) -SH

(e) $-SC(O)R^8$

5 donde R^8 es un grupo alquilo o grupo fenilo $C_1 - C_{24}$, opcionalmente sustituido;

(f) $-SC(V)OR^8$

doonde V es O o S, y R^8 es un grupo alquilo o grupo fenilo $C_1 - C_{24}$, opcionalmente sustituido,

(g) -X

donde X se selecciona del grupo que consiste en F, Cl, Br y I;

10 (h) $-OS(O)_2R^5$

donde R^5 se selecciona del grupo que consiste en grupos alquilo $C_1 - C_{24}$, grupos arilo $C_6 - C_{24}$, grupos alquilarilo $C_7 - C_{24}$ y grupos arilalquilo $C_7 - C_{24}$, donde los grupos alquilo, grupos arilo, grupos alquilarilo y grupos arilalquilo están opcionalmente sustituidos;

(i) R^{12}

15 donde R^{12} se selecciona del grupo que consiste en grupos alqueno $C_2 - C_{24}$, grupos cicloalqueno $C_3 - C_5$ y grupos alcadieno $C_4 - C_8$, opcionalmente sustituidos y terminales; y

(j) R^{13}

donde R^{13} es un grupo alenilo terminal $C_3 - C_{24}$ opcionalmente sustituido; y

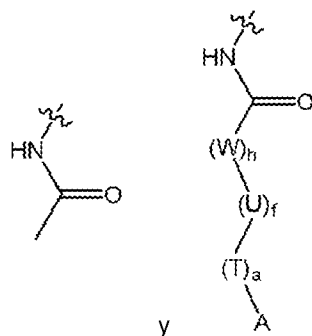
(k) $N(R^{17})_2$

20 donde R^{17} se selecciona independientemente del grupo que consiste en H y grupos alquilo $C_1 - C_{12}$;

"Z" es CH_2 , CF_2 o $C(O)$; o Z es $CHOH$ con la condición de que g sea 0, f sea 1 y a sea 0 o 1;

"Y" se selecciona del grupo que consiste en O, S, $N(R^{15})$, $N(R^{15})C(O)$, $N(R^{15})C(O)N(R^{15})$, $N(R^{15})C(O)O$, $OC(O)N(R^{15})S(O)_2N(R^{15})$ y $N(R^{15})C(O)N(R^{15})S(O)_2O$, donde R^{15} se selecciona independientemente del grupo que consiste en H, grupos alquilo $C_1 - C_{12}$ y $(U)_f(T)_aA$ donde f, a, U, T y A son como se ha definido anteriormente; y

25 R^{14} se selecciona del grupo que consiste en:



donde:

a, f, T, A y U son como se ha definido anteriormente;

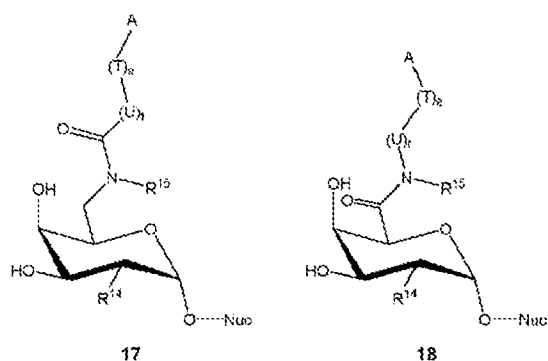
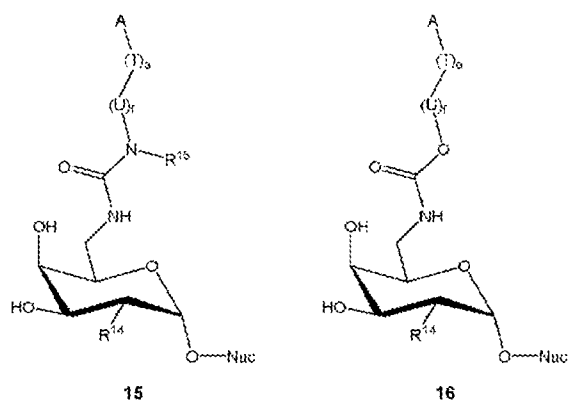
h es 0 o 1; y

30 W se selecciona del grupo que consiste en O, S, NR^{15} , $NHS(O)_2O$ y

$NHS(O)_2NR^{15}$, donde R^{15} es como se ha definido anteriormente.

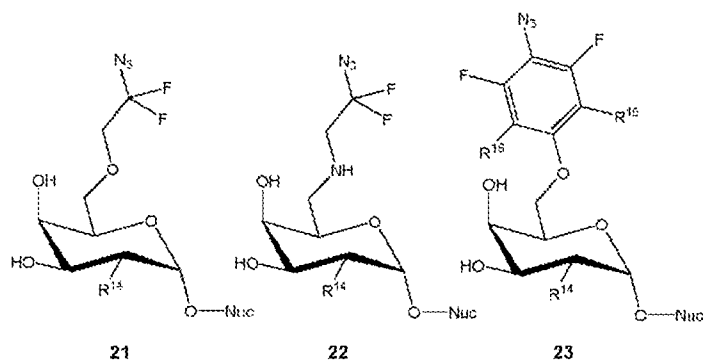
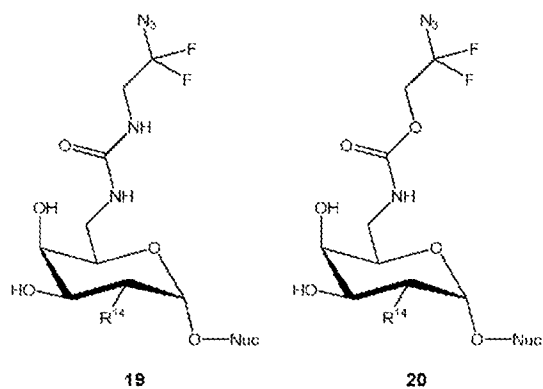
2. Proceso según la reivindicación 1, donde A se selecciona del grupo que consiste en $-N_3$, $-C(O)CH_3$, $-SH$, $-CH=CH_2$ y $-CH_2C\equiv CH$.

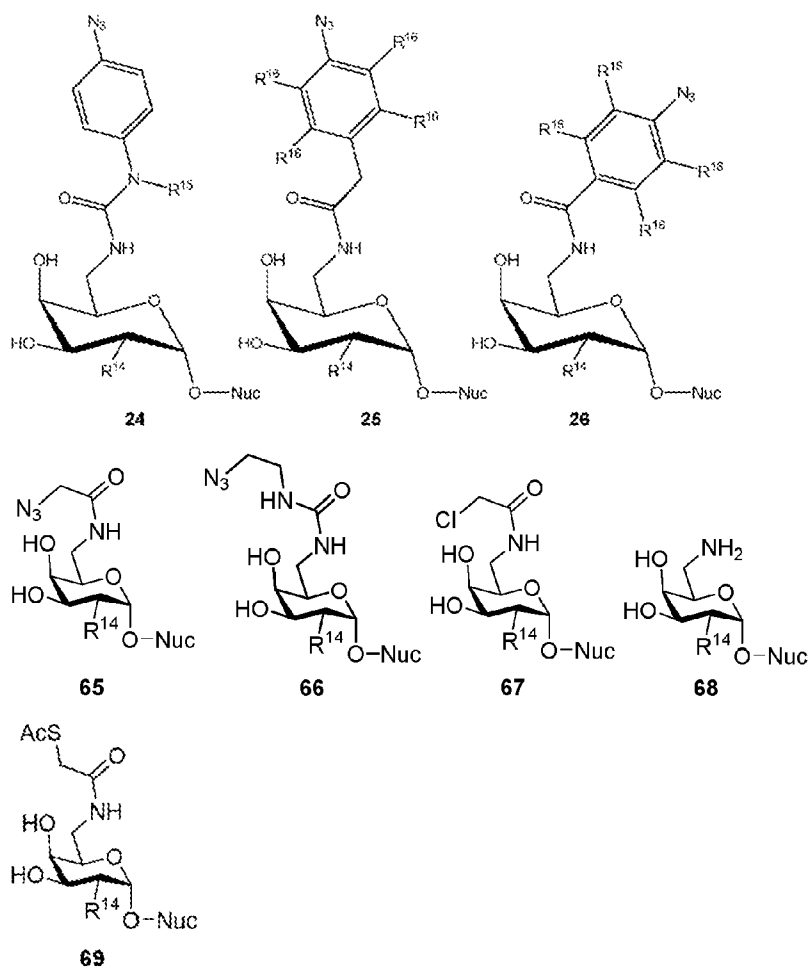
3. Proceso según la reivindicación 1 o 2, donde el nucleótido de azúcar Su(A)- Nuc es según la fórmula (15), (16), (17) o (18):



5 donde Nuc, a, f, R¹⁴, R¹⁵, A, U y T son como se ha definido en la reivindicación 1,

o donde nucleótido de azúcar Su(A)-Nuc es según la fórmula (19), (20), (21), (22), (23), (24), (25), (26), (65), (66), (67), (68) o (69):



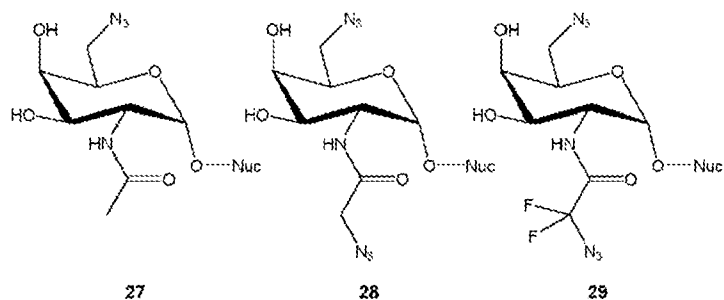


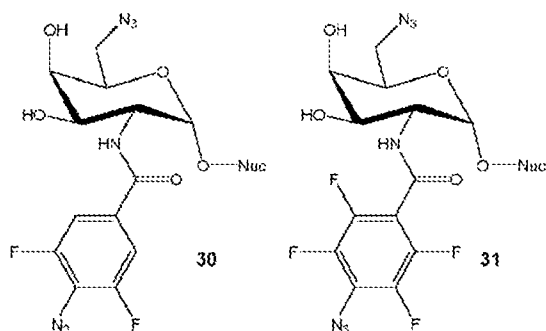
donde:

5 R^{14} y R^{15} son como se define en la reivindicación 1; y

R^{16} se selecciona independientemente del grupo que consiste en H y F.

4. Proceso según cualquiera de las reivindicaciones anteriores, donde el nucleótido de azúcar Su(A)- Nuc es según la fórmula (27), (28), (29), (30), o (31):

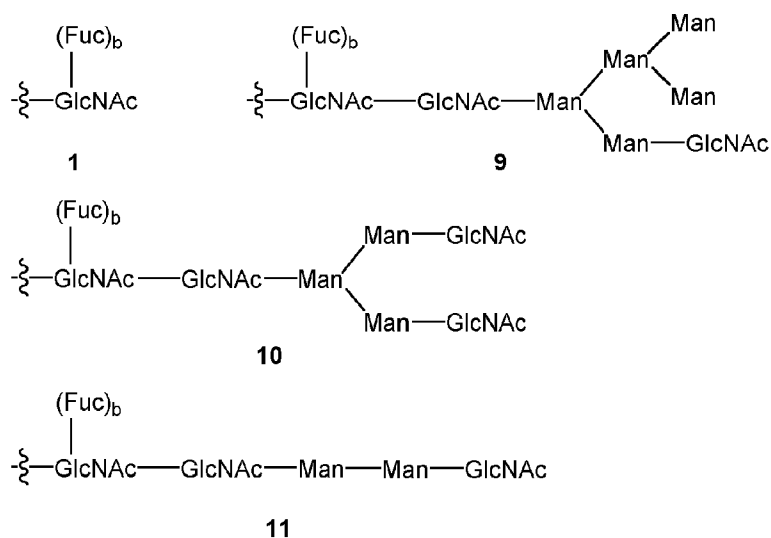




donde Nuc es como se define en la reivindicación 1.

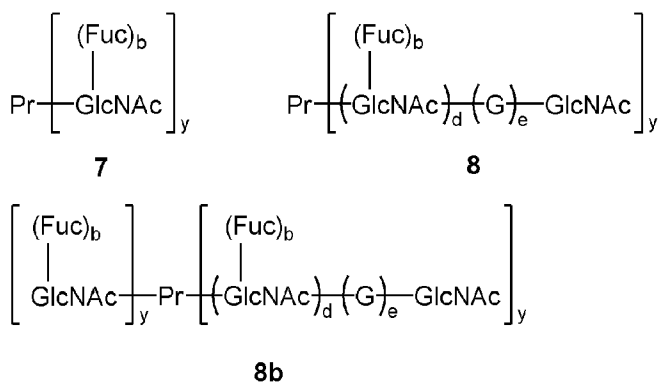
5. Proceso según cualquiera de las reivindicaciones precedentes, donde el nucleótido en el nucleótido de azúcar Su(A)-Nuc es UDP.

- 5 6. Proceso según cualquiera de las reivindicaciones precedentes, donde el glucano que comprende un resto de GlcNAc terminal es según la fórmula (1), (9), (10) u (11):



- 10 donde b es como se define en la reivindicación 1.

7. Proceso según cualquiera de las reivindicaciones precedentes, donde la glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal es según la fórmula (7), (8) o (8b):



- 15 donde:

b, d, e y G son como se define en la reivindicación 1;

"y" es independientemente un número entero en el intervalo de 1 a 24; y

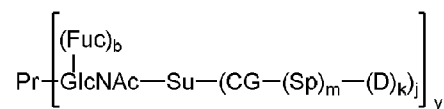
Pr es una proteína.

8. Proceso según cualquiera de las reivindicaciones precedentes, donde la glucoproteína que comprende un glucano que comprende un resto de GlcNAc terminal es un anticuerpo.

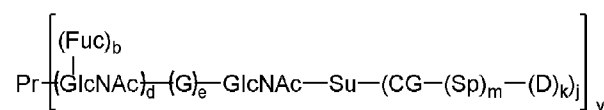
9. Glucoproteína que se puede obtener mediante el proceso según una cualquiera de las reivindicaciones 1 - 8.

10. Glucoproteína según la reivindicación 9, donde la glucoproteína es un anticuerpo.

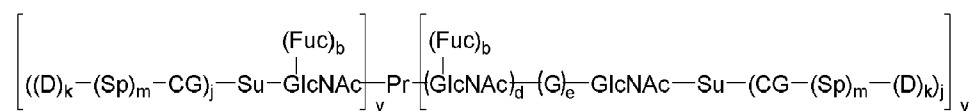
5 11. Bioconjugado según la fórmula (75), (76) o (77):



75



76



77

donde:

10 Pr es una proteína

"y" es independientemente un número entero en el intervalo de 1 a 24;

CG es un grupo conector que conecta Su a Sp o D;

Sp es un espaciador;

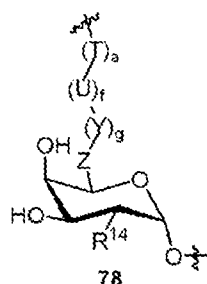
D es una molécula de actuación;

15 j es independientemente 1, 2, 3, 4 o 5;

k es independientemente un número entero en el intervalo de 1 a 10;

m es 0 o 1;

Su es según la fórmula (78):



78

20 y donde b, d, e, G, R¹⁴, Z, Y, U, T, g, f y a son como se define en la reivindicación 1, y Su está conectado a través de C1 a C4 del resto de GlcNAc a través de un enlace β-1,4-O-glucosídico y a CG a través de Z, Y, U o T.

12. Bioconjugado según la reivindicación 11, donde D es una citotoxina.

13. Bioconjugado según las reivindicaciones 11 o 12, que es un conjugado anticuerpo-fármaco.

Fig. 1

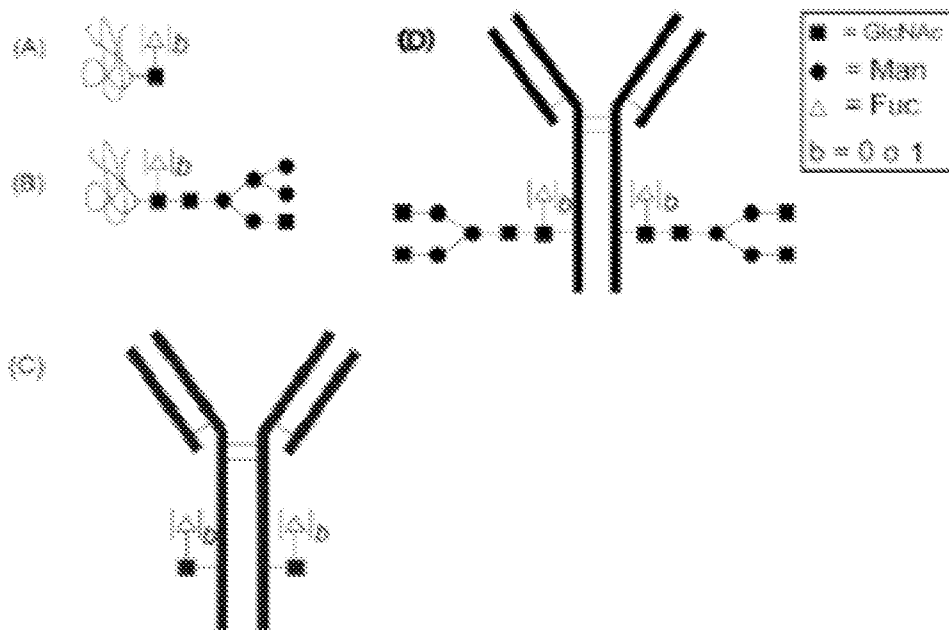


Fig. 2

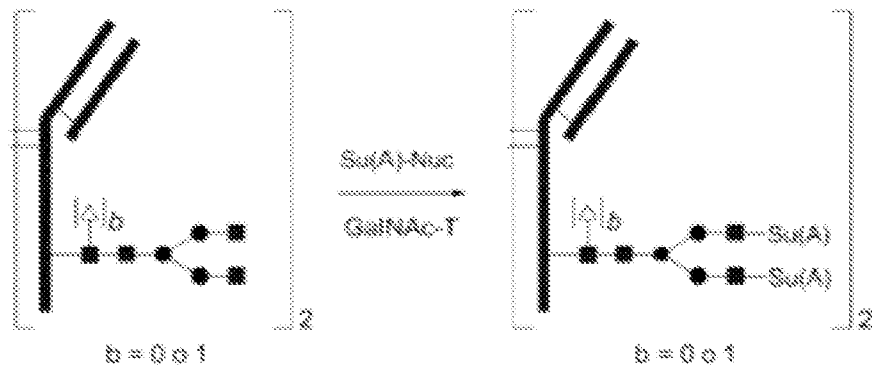


Fig. 3

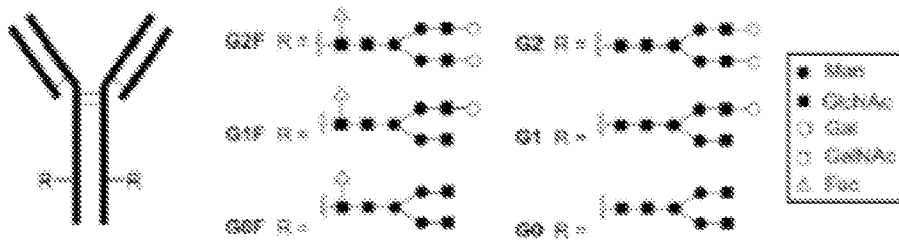


Fig. 4

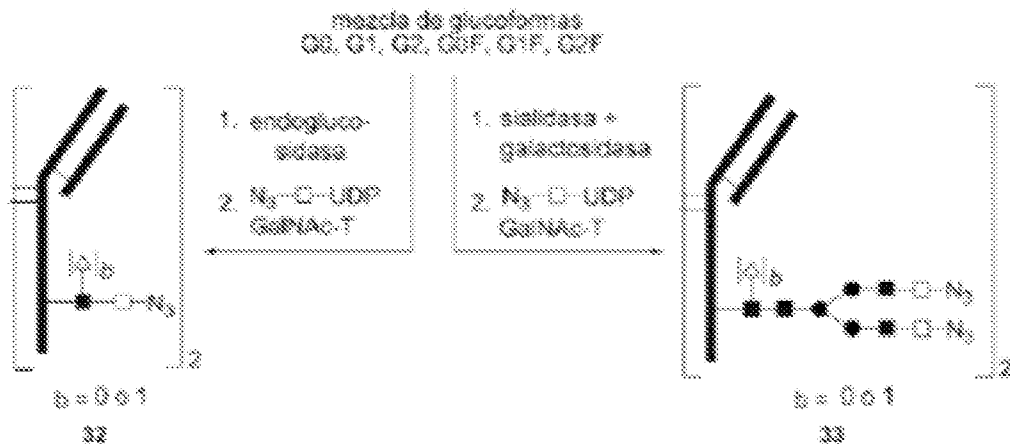


Fig. 5

