

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 983 352**

51 Int. Cl.:

C07K 14/245 (2006.01)
A61K 38/48 (2006.01)
A61P 35/00 (2006.01)
C12N 9/52 (2006.01)
C07K 16/12 (2006.01)
C07K 16/40 (2006.01)
G01N 33/574 (2006.01)
G01N 33/68 (2006.01)
G01N 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **09.11.2017** **PCT/AU2017/051230**
87 Fecha y número de publicación internacional: **17.05.2018** **WO18085888**
96 Fecha de presentación y número de la solicitud europea: **09.11.2017** **E 17870370 (8)**
97 Fecha y número de publicación de la concesión europea: **27.03.2024** **EP 3538544**

54 Título: **Mutante de la subunidad B de la citotoxina subtilasa**

30 Prioridad:

09.11.2016 AU 2016904572

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
22.10.2024

73 Titular/es:

GRIFFITH UNIVERSITY (50.0%)
170 Kessels Road
Nathan, Queensland 4111, AU y
THE UNIVERSITY OF ADELAIDE (50.0%)

72 Inventor/es:

JENNINGS, MICHAEL, PAUL;
DAY, CHRISTOPHER;
PATON, ADRIENNE, WEBSTER y
PATON, JAMES, CLELAND

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 983 352 T3

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

DESCRIPCIÓN

Mutante de la subunidad B de la citotoxina subtilasa

5 Campo técnico de la invención

Esta invención se refiere a proteínas de toxinas bacterianas. Más particularmente, esta invención se refiere a una proteína mutante de la subunidad B de la citotoxina subtilasa que tiene capacidad para unirse al ácido *N*-glicolilneuramínico ligado a α 2-6 al tiempo que conserva la capacidad de unirse al ácido *N*-glicolilneuramínico ligado a α 2-3.

10

Antecedentes de la invención

Las toxinas AB5 ejercen sus efectos en un proceso de dos pasos: (i) unión de la subunidad B pentamérica a receptores glicanos específicos en la superficie de la célula objetivo; (ii) internalización de la toxina AB5, seguida de la inhibición o corrupción de funciones esenciales del hospedero mediada por la subunidad A ¹. Las subunidades B de las toxinas AB5 reconocen los receptores glicanos de la superficie celular, dirigiendo la internalización y el tráfico intracelular de la holotoxina. La especificidad de estas interacciones proteína-glicano es crítica para la patogénesis, ya que determina la susceptibilidad del hospedero y el tropismo tisular. Además, las interacciones pentavalentes entre las subunidades B de la toxina AB5 y sus glicanos afines dan lugar a uniones de muy alta afinidad, lo que las convierte en potentes ligandos para la detección de glicanos, siendo un ejemplo notable el uso de la subunidad B de la toxina del cólera para la detección del gangliósido GM1 en secciones histopatológicas ² y para el etiquetado de balsas lipídicas en membranas ³.

15

20

25

30

35

40

45

En 2004, Paton *et al.* describieron el descubrimiento y la caracterización biológica inicial de una nueva subfamilia de toxinas bacterianas AB5 cuyo prototipo se denominó citotoxina subtilasa (SubAB) ⁴. En el caso de SubAB, se descubrió que la subunidad A (SubA) es una serina proteasa de la familia de las subtilasas con una especificidad excelente para la chaperona esencial del retículo endoplásmico BiP/GRP78 ⁵. Los estudios estructurales revelaron que, a diferencia de la mayoría de las subtilasas, la SubA poseía una hendidura del sitio activo inusualmente profunda, lo que explicaba su excelente especificidad de sustrato ⁵. La SubA ha demostrado ser una poderosa herramienta para examinar el papel de BiP en diversos procesos celulares y también tiene potencial como un agente terapéutico contra el cáncer ^{6,7}. Significativamente, el análisis de matrices de glicanos ha demostrado que la subunidad B de la toxina (SubB) tiene un alto grado de especificidad de unión para los glicanos que terminan con ácido *N*-glicolilneuramínico ligado a α 2-3 (Neu5Gc), un ácido siálico que los humanos no pueden sintetizar ⁸. De todos los glicanos de la matriz, la mejor unión se produjo con Neu5Gc α 2-3Gal β 1-4GlcNAc β -. La unión de la toxina etiquetada a la matriz se redujo 20 veces si se cambiaba el Neu5Gc por Neu5Ac; más de 30 veces si se cambiaba el enlace Neu5Gc de α 2-3 a α 2-6; y 100 veces si se eliminaba el ácido siálico. El patrón general de unión a las estructuras representadas en la matriz indicó que la SubB tiene una gran afinidad por el Neu5Gc terminal ligado a α 2-3 con poca discriminación por el penúltimo resto. La estructura cristalina del complejo SubB-Neu5Gc reveló la base de esta especificidad. El hidroxilo adicional en el grupo metilo del resto *N*-acetilo que distingue al Neu5Gc del Neu5Ac interactúa con el Tyr78^{OH} de la SubB y establece uniones de hidrógeno con la cadena principal del Met10 ⁸. Estas interacciones clave no podrían producirse con Neu5Ac, explicando así la marcada preferencia por Neu5Gc. Guiados por los datos estructurales, se mutagenizaron residuos clave en el bolsillo de unión previsto, y esto anuló el reconocimiento del glicano, la unión celular y la toxicidad. Los aminoácidos S12 e Y78 de la SubB forman uniones estabilizadoras cruciales con el Neu5Gc⁸. Una mutación S12A abolió completamente la unión al glicano, mientras que una mutación Y78F que impide las interacciones con el grupo^{C11} OH que distingue el Neu5Gc del Neu5Ac redujo la unión al glicano en un 90 % y abolió la preferencia de la proteína SubB mutante por Neu5Gc frente a Neu5Ac8.

50

55

60

La forma más prominente de glicosilación aberrante en los cánceres humanos es la expresión de glicanos terminados por Neu5Gc. El Neu5Gc no se expresa en niveles significativos en las células humanas sanas normales ⁹⁻¹², ya que los humanos no pueden sintetizar Neu5Gc debido a una mutación inactivadora en el gen CMAH ¹³. No obstante, las investigaciones sugieren que la presentación de Neu5Gc en pacientes con cáncer puede explicarse por la absorción de Neu5Gc a través de la ingesta dietética de carne roja y productos lácteos, que son las fuentes más ricas en Neu5Gc ¹⁴. La presencia de Neu5Gc tiene importancia pronóstica, ya que su expresión se correlaciona frecuentemente con la invasividad, la metástasis y el grado tumoral ¹⁰. La visualización preferente de los glicanos Neu5Gc en las células cancerosas puede explicarse, al menos en parte, por el entorno tumoral hipóxico, que induce notablemente la expresión del transportador de ácido siálico sialina, lo que da lugar a una mayor visualización de Neu5Gc y otros ácidos siálicos en la superficie celular ¹⁵. Debido a que los sialil-conjugados regulan la adhesión y promueven la movilidad celular, estas alteraciones en la sialilación superficial pueden influir en la colonización y el potencial metastásico de las células tumorales ¹⁶. Se han observado niveles elevados de ácidos siálicos anormales, tal como el Neu5Gc, en el cáncer de mama, ovario, próstata, colon y pulmón ^{11,12}. Es importante destacar que la incorporación de Neu5Gc en las células cancerosas es más prominente en las glicoproteínas solubles que se encuentran tanto en el espacio extracelular como en el interior de la célula, y el Neu5Gc es el ácido siálico dominante en las glicoproteínas secretadas por las células cancerosas a los tejidos circundantes ⁹. También se sabe que la expresión de Neu5Gc en el cáncer impulsa la producción de xeno-autoanticuerpos contra Neu5Gc ^{17,18}. Estos anticuerpos anti-Neu5Gc se están investigando para determinar su potencial para nuevos diagnósticos, pronósticos y terapias en carcinomas humanos ¹⁷.

65

Byres *et al.* (2008) *Nature* 456: 648-53 demostraron que la SubB tiene una fuerte preferencia por los glicanos que terminan en el ácido siálico *N*-glicolilneuramínico (Neu5Gc) e identificaron residuos específicos implicados en esta especificidad,

cuya mutagénesis anulaba el reconocimiento del glicano.

Breve descripción de la invención

5 La presente invención se refiere a una proteína mutante de la subunidad B de la citotoxina subtilasa (SubB) que puede unirse a glicanos que tienen ácido *N*-glicolilneuramínico ligado a α 2-3y a glicanos que tienen ácido *N*-glicolilneuramínico ligado a α 2-6. Por lo tanto, la proteína SubB mutante tiene una capacidad no descrita anteriormente para unirse a glicanos que tienen ácido *N*-glicolilneuramínico ligado a α 2-6, mientras que no pierde la capacidad de unirse a glicanos que tienen ácido *N*-glicolilneuramínico ligado a α 2-3. Esto proporciona una proteína SubB mutante que puede utilizarse para detectar y tener como objetivo un espectro más amplio de glicanos que contienen ácido *N*-glicolilneuramínico de lo que era posible anteriormente.

15 Un aspecto de la invención proporciona una proteína aislada que comprende una secuencia de aminoácidos de SubB (SEQ ID NO:2) o un fragmento de la misma que comprende la secuencia de aminoácidos TTSTE (SEQ ID NO:3), que comprende una delección de uno o más de los residuos de aminoácidos de la secuencia de aminoácidos TTSTE (SEQ ID NO:3), y en donde la proteína aislada o el fragmento de la misma es capaz de unirse a ácido *N*-glicolilneuramínico ligado a α 2-3y ácido *N*-glicolilneuramínico ligado a α 2-6.

20 En una realización, la proteína aislada o fragmento de la misma comprende una delección de al menos uno de los residuos subrayados de TTSTE (SEQ ID NO:3) y opcionalmente una delección del residuo subrayado de TTSTE (SEQ ID NO:3).

25 En otra realización particular, la proteína aislada o fragmento de la misma comprende una delección de los residuos subrayados de TTSTE (SEQ ID NO:3). De forma adecuada, la proteína aislada o el fragmento de esta realización puede unir glicanos Neu5Ac, tales como Neu5Ac- α 2-6-lac y Neu5Ac- α 2-3-lac.

En una realización, la proteína aislada o fragmento de la misma comprende la secuencia de aminoácidos establecida en SEQ ID NO: 1 o una secuencia de aminoácidos que sea idéntica en un 98 % en toda la longitud completa de SEQ ID NO: 1.

30 Otro aspecto de la invención proporciona un complejo molecular aislado que comprende la proteína aislada o el fragmento del primer aspecto y un glicano que comprende ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6.

35 Otro aspecto más de la invención proporciona una composición que comprende la proteína aislada o el fragmento del primer aspecto.

En una realización la composición es una composición farmacéutica.

40 En otra realización, la composición es una composición diagnóstica.

45 Todavía otro aspecto de la invención proporciona un método para detectar el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6, en donde dicho método incluye el paso de combinar la proteína aislada o el fragmento del primer aspecto con una muestra para formar así un complejo detectable que comprenda la proteína aislada o el fragmento del primer aspecto y el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6.

En algunas realizaciones, el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6 pueden ser expresados por una célula tumoral o células sanguíneas felinas.

50 En otra realización, el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6 puede ser un contaminante en una muestra o preparación que comprenda fármacos glicosilados recombinantes, anticuerpos y otras biomoléculas terapéuticas para administración humana.

55 Otro aspecto de la invención proporciona un método de aislar un glicano o una célula que expresa el glicano, en donde el glicano comprende ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6, en donde dicho método incluye los pasos de: combinar la proteína aislada o el fragmento divulgado en la presente con una muestra para formar así un complejo que comprenda la proteína aislada o el fragmento y el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6; y aislar la proteína o la célula.

60 En algunas realizaciones, la célula es una célula tumoral o una célula sanguínea felina.

En otra realización, el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6 puede ser un contaminante en una preparación o formulación que comprenda fármacos glicosilados recombinantes, anticuerpos y otras biomoléculas terapéuticas para administración humana.

65 Otro aspecto de la invención proporciona una proteína aislada o un fragmento de acuerdo con la presente invención para

su uso en el tratamiento del cáncer en un sujeto para así dirigirse selectivamente a una célula cancerosa que exprese un ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6.

De forma adecuada, la proteína aislada o el fragmento del primer aspecto se acopla a un agente citotóxico.

Otro aspecto de la invención proporciona un ácido nucleico aislado que codifica la proteína aislada o el fragmento del primer aspecto.

Otro aspecto adicional de la invención proporciona un constructo genético que comprende el ácido nucleico aislado de la invención.

Otro aspecto adicional de la invención proporciona una célula hospedera que comprende el constructo genético del aspecto mencionado.

Aspectos relacionados de la invención proporcionan kits que comprenden la proteína aislada o fragmento de la misma, ácido nucleico aislado, composición, y/o constructo genético. Pueden utilizarse, por ejemplo, en la detección del ácido *N*-glicolilneuramínico ligado a α 2-3 y/o del ácido *N*-glicolilneuramínico ligado a α 2-6 o en el direccionamiento terapéutico de las células tumorales que expresan el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6.

A lo largo de esta memoria descriptiva, a menos que se indique lo contrario, "comprenden", "comprende" y "que comprende" se utilizan de forma inclusiva y no exclusiva, de modo que un número entero o grupo de números enteros establecido puede incluir uno o más números enteros o grupos de números enteros no establecidos.

Por "*consistir esencialmente en*" se entiende en este contexto que la proteína aislada o el fragmento tiene uno, dos o no más de tres residuos de aminoácidos además de la secuencia de aminoácidos recitada. Los residuos de aminoácidos adicionales pueden producirse en los extremos N- y/o C-terminales de la secuencia de aminoácidos recitada, aunque sin limitación al respecto.

También se entenderá que los artículos indefinidos "un" y "una" no deben leerse como singulares o como si de otro modo excluyeran más de uno o más de un único sujeto al que se refiera el artículo indefinido. Por ejemplo, "una" proteína incluye una proteína, una o más proteínas o una pluralidad de proteínas.

Breve descripción de las figuras

Figura 1. Representación de la superficie de la SubB en complejo con (A) Neu5Gc α 2-3Gal β 1-3GlcNAc (determinado a partir de una estructura cristalina de rayos X (Byres et al., 2008)) y (B) Neu5Gc α 2-6Gal β 1-3Glc (modelado con la estructura cristalina de rayos X). Los trisacáridos se muestran como una barra verde o cian con residuos rojos y azules que representan el oxígeno y el nitrógeno, respectivamente.

Figura 2. Representación de la superficie de los mutantes de tipo silvestre y SubB modelados con Neu5Gc α 2-6Gal β 1-3Glc (mostrado como una barra cian). Los residuos SubB mutados se muestran como barras grises, y los residuos rojos y azules representan el oxígeno y el nitrógeno, respectivamente.

Figura 3. ELISA de SubB modificada frente a suero humano y bovino etiquetado con FITC. SubB (A) y SubB Δ S106/ Δ T107 (B) recubiertas sobre placas ELISA fueron capaces de capturar proteínas séricas humanas y bovinas etiquetadas con FITC. Las barras de error muestran +1SD de la media de ensayos duplicados.

Figura 4. Ensayo de superposición de lectinas. Unión de SubB Δ S106/ Δ T107 a diluciones seriadas de AGP humana o bovina aplicadas como manchas sobre nitrocelulosa (se indican las cantidades totales de proteína por mancha).

Figura 5. Iones de oxonio NeuAc y NeuGc procedentes de digestiones tríplicas de glicoproteínas alfa-1 ácidas humanas y bovinas. Región de baja masa de los espectros MS/MS de (A) el ion glucopéptido a un *m/z* de 1183,163+ correspondiente al péptido TFMLAASWN[Hex2HexNAc2NeuAc1NeuGc1+Man3GlcNAc2]GTK de la glucoproteína alfa-1 ácida bovina y (B) el ion glucopéptido a un *m/z* de 1122,284+ correspondiente al péptido QDQCIYN[Hex3HexNAc3NeuAc2+Man3GlcNAc2]TTYLVNQR de la glucoproteína alfa-1 ácida humana que muestra abundantes iones oxonio, incluyendo los 274,1 y 292,1 específicos de NeuAc, y los 290,1 y 308,1 específicos de NeuGc. (C) Intensidad de los iones de oxonio específicos de NeuAc y NeuGc como una proporción de la intensidad total de iones de todos los espectros MS/MS procedentes del análisis LC-MS/MS de digestiones tríplicas de glicoproteínas alfa-1 ácidas humanas y bovinas. (D) Gel de proteínas de la AGP humana y bovina utilizada en MS, ELISA, Biacore y transferencia en mancha.

Figura 6. Imágenes en bruto de SubB de tipo silvestre (WT) y mutante de delección SubB Δ S106/ Δ T107 (2M) de las matrices Z-biotech. Glicanos Neu5Gc de la Región A (Parte superior y 3 manchas en la parte inferior derecha de cada submatriz). Glicanos Neu5Ac de la Región B (Parte inferior de cada submatriz). Manchas de control de la Región C (Lado derecho de cada submatriz).

Figura 7. Interacción SubB WT con la matriz de glicanos Neu5Ac/Gc Z-Biotech. Los glicanos Neu5Gc se muestran en azul, los Neu5Ac se muestran en rojo.

Figura 8. Interacción del mutante de delección SubB Δ S106/ Δ T107 (Sub2M) con la matriz de glicanos Neu5Ac/Gc Z-Biotech. Los glicanos Neu5Gc se muestran en azul, los Neu5Ac se muestran en rojo.

Figura 9. Glicanos en la matriz Z-biotech.

Figura 10. SubB WT con AGP bovina mezclada con un 1 % de suero humano normal.

Figura 11. SubB 2M (mutante de delección Δ S106/ Δ T107) con AGP bovina mezclada con un 1 % de suero humano normal.

Descripción detallada

La presente invención se refiere a una proteína mutante modificada de la subunidad B de la citotoxina subtilasa (SubB) o a un fragmento de la misma que tiene uno o más residuos de aminoácidos delecionados de la secuencia de aminoácidos TTSTE que pueden unirse a glicanos que terminan en Neu5Gc ligados a α 2-3 o α 2-6. La proteína aislada o un fragmento de la misma pueden utilizarse en métodos de detección de glicanos que comprenden ácido N-glicolilneuramínico ligado a α 2-3 y/o ácido N-glicolilneuramínico ligado a α 2-6, tales como los expresados por ciertas células tumorales y también expresados por las células sanguíneas felinas tipo A. Estos glicanos también pueden ser contaminantes en las preparaciones de fármacos y otras biomoléculas. Por lo tanto, otros aspectos de la invención pueden estar relacionados con la purificación, eliminación o agotamiento de glicanos que comprenden ácido N-glicolilneuramínico ligado a α 2-3 y/o ácido N-glicolilneuramínico ligado a α 2-6, o células que expresan estos glicanos. Otro aspecto de la invención se refiere a los usos terapéuticos de la proteína aislada para el suministro dirigido de agentes anticancerígenos a determinadas células tumorales.

A los efectos de la presente invención, por “aislado” se entiende un material que ha sido extraído de su estado natural o sometido a cualquier otra manipulación humana. El material aislado puede estar parcial, sustancial o esencialmente libre o agotado de los componentes que normalmente lo acompañan en su estado natural. El material aislado puede estar en forma nativa, sintética química o recombinante. En algunas realizaciones, el material aislado puede estar en forma enriquecida, parcialmente purificada o purificada.

Tal como se utiliza en la presente, una “muestra” puede ser cualquier fracción, pieza, porción o parte que sea representativa de una entidad mayor. La muestra puede ser de un producto farmacéutico, un fármaco, un anticuerpo u otra formulación o preparación terapéutica, o una muestra biológica, tal como una obtenida de un ser humano, un animal u otra fuente biológica. En algunas realizaciones, una muestra biológica puede ser una muestra de célula o de tejido, tal como una biopsia, un frotis, una sección de tejido o un sedimento celular, o una muestra de fluido, tal como orina, suero, plasma, líquido cefalorraquídeo o saliva, aunque sin limitaciones.

En general, tal como se utiliza en la presente, una “composición” comprende una proteína aislada o un fragmento de la misma, un ácido nucleico, un constructo genético, un anticuerpo u otra molécula junto con uno o más componentes, tales como agua u otros disolventes, sales, agentes amortiguadores y/o estabilizadores, aunque sin limitación de los mismos. En una realización particular, una composición “diagnóstica” puede comprender uno o más componentes moleculares que faciliten la detección de proteínas que comprendan ácido N-glicolilneuramínico ligado a α 2-3 y/o ácido N-glicolilneuramínico ligado a α 2-6. Dichos componentes pueden incluir sustratos enzimáticos, anticuerpos secundarios, reactivos de color, etiquetas y catalizadores (por ejemplo, “reactivos de detección”), como se describirá con más detalle a continuación. En otras realizaciones particulares, una composición “farmacéutica” puede comprender uno o más componentes moleculares que faciliten la administración terapéutica de la proteína aislada o del fragmento de la misma divulgado en la presente (por ejemplo, un portador, diluyente o excipiente), como se describirá con más detalle en lo sucesivo.

Por “proteína” se entiende un polímero de aminoácidos. Los aminoácidos pueden ser aminoácidos naturales o no naturales, D- o L-aminoácidos como son bien conocidos en la técnica. El término “proteína” incluye y abarca “péptido”, que se utiliza típicamente para describir una proteína que no tiene más de cincuenta (50) aminoácidos, y “polipéptido”, que se utiliza típicamente para describir una proteína que tiene más de cincuenta (50) aminoácidos.

Tal como se utiliza generalmente en la presente, un “glicano” es una glicoproteína, un glicolípido u otra macromolécula que contiene carbohidratos, e incluye moléculas que pueden denominarse peptidoglicanos, glicoproteínas, glicopéptidos, glicolipoproteínas y similares. Un glicano particular comprende el ácido N-glicolilneuramínico (Neu5Gc). El glicano puede comprender ácido N-glicolilneuramínico ligado a α 2-3 o ácido N-glicolilneuramínico ligado a α 2-6. Convenientemente, el ácido N-glicolilneuramínico ligado a α 2-3y los ácidos N-glicolilneuramínicos ligados a α 2-6 son ácidos siálicos terminales. A modo de ejemplo, el ácido N-glicolilneuramínico ligado a α 2-3 es un ácido siálico terminal, tal como en Neu5Gc α 2-3Gal β 1-4GlcNAc β -, y el ácido N-glicolilneuramínico ligado a α 2-6 es un ácido siálico terminal en Neu5Gc α 2-6Gal β 1-3Glc-

Como se entenderá a partir de lo anterior, un aspecto preferido de la invención proporciona una proteína aislada que comprende una secuencia de aminoácidos de la proteína SubB (SEQ ID NO:2) o un fragmento de la misma que comprende la secuencia de aminoácidos TTSTE (SEQ ID NO:3), que tiene uno o más residuos de aminoácidos delecionados de la secuencia de aminoácidos TTSTE, en donde la proteína aislada o el fragmento es capaz de unir un glicano que comprende ácido N-glicolilneuramínico ligado a α 2-3 y un glicano que comprende ácido N-glicolilneuramínico ligado a α 2-6. Un aspecto relacionado de la invención proporciona un complejo molecular aislado que comprende la proteína aislada o el fragmento del primer aspecto y un glicano que comprende ácido N-glicolilneuramínico ligado a α 2-3 y/o ácido N-glicolilneuramínico ligado a α 2-6.

En este contexto, por “capaz de unirse a ácido N-glicolilneuramínico ligado a α 2-3 y ácido N-glicolilneuramínico ligado a

α 2-6" se entiende que la proteína aislada o el fragmento une glicanos de ácido N-glicolilneuramínico ligado a α 2-6 con una afinidad sustancialmente mayor que la de una proteína SubB de tipo silvestre, mientras que también une glicanos de ácido N-glicolilneuramínico ligado a α 2-3 con una afinidad comparable a la de una proteína SubB de tipo silvestre. En una realización particular, la proteína aislada o el fragmento se une a glicanos de ácido N-glicolilneuramínico ligado a α 2-6 con una afinidad de aproximadamente 5-15 nM, aproximadamente 7-12 nM o aproximadamente 8-10 nM. En una realización particular, la proteína aislada o el fragmento se une a glicanos de ácido N-glicolilneuramínico ligado a α 2-3 con una afinidad de aproximadamente 8-20 nM, 10-18 nM o aproximadamente 14-16 nM.

La secuencia de aminoácidos TTSTE está normalmente presente en una proteína SubB de tipo silvestre. La proteína SubB de tipo silvestre puede comprender una secuencia de aminoácidos establecida en SEQ ID NO:2:

```

1  MTIKRFFYCA GIRGCLSLNP ANAEWTDAR DQNTGQVIT QFNTGQIDNK FYFCIECKQS
61  AGESISACSE ENSEWQASR STLNQALYF ITTCQFVAIY YFQVWTYFP EVFALTSNAL
121 VELSPCTTST SCFGPRRKR S

```

Los residuos subrayados son una región N-terminal que está ausente en la forma madura de la SubB. Por lo tanto, la numeración utilizada en la presente comienza en el residuo de glutamato 24 (es decir, Glu24 = residuo 1). Utilizando esta numeración, los residuos **en negrita** TTSTE (SEQ ID NO:3) corresponden al "bucle T104-E108". Se propone que el azúcar terciario de la estructura α 2-6 se pliega de nuevo sobre la superficie de la proteína SubB, haciendo estrecho contacto con un bucle que comprende los residuos T104-E108 de SubB. Este bucle está estabilizado por un enlace disulfuro entre C103 y C109. El impedimento estérico resultante distorsiona el acoplamiento del Neu5Gc terminal en el bolsillo de unión, lo que explica la unión significativamente más deficiente de las estructuras Neu5Gc enlazadas con α 2-6 observada en el análisis original de la matriz de glicanos de la SubB⁸. La modificación de uno o más residuos en el bucle potencia la unión de una proteína SubB mutante a la estructura α 2-6 al tiempo que permite la unión a las estructuras α 2-3. En general, las deleciones de aminoácidos que reducen o disminuyen la "altura" del bucle pueden ser ventajosas para mejorar la unión de las estructuras de Neu5Gc ligado α 2-6. En este contexto, la "altura" puede ser una función de la distancia que un grupo R del aminoácido proyecta o se extiende desde la cadena principal del péptido en el espacio tridimensional (por ejemplo, la valina tiene mayor altura que la leucina). Por lo tanto, en una realización se prefiere la deleción de uno o más de los residuos TTSTE (SEQ ID NO:3) del bucle (denominado en la presente "mutante de deleción"). Preferiblemente, están subrayados en TTSTE (SEQ ID NO:3). Con base en la numeración de la secuencia de aminoácidos SubB madura en SEQ ID NO:2, se delecionan los residuos S106 y/o T107. En una realización particularmente preferida, se delecionan los residuos S106 y T107.

Por lo tanto, una realización particular de una proteína SubB mutante comprende la secuencia de aminoácidos de SEQ ID NO:1:

```

1  ENTQARDEN FEQVITQFR TGQIDNKPYF CIECKQAGS SIEACSKKNS SYNGASFTEL
61  INQALYFYTT QGFVAIYKFP GWWTYFPVX ALTSNALVGL STQTECEGFG GRKINS

```

En otra realización, la proteína aislada o el fragmento comprenden una deleción de los residuos de aminoácidos subrayados en la secuencia de aminoácidos TTSTE (SEQ ID NO:3).

Con base en la numeración de la secuencia de aminoácidos SubB madura en SEQ ID NO:2, se delecionan los residuos T107 y E108. Como resultará evidente a partir de los datos mostrados en la Tabla 1, la proteína aislada "mutante de deleción" que carece de T107 y E108 se une a glicanos de ácido N-glicolilneuramínico ligado a α 2-6 con una afinidad sustancialmente mayor que la de una proteína SubB de tipo silvestre, al tiempo que también se une a glicanos de ácido N-glicolilneuramínico ligado a α 2-3. Sin embargo, a diferencia de un mutante de deleción S106 y T107 (tal como el de la SEQ ID NO:1), el mutante de deleción T107 y E108 puede unirse ampliamente a glicanos, incluidos glicanos Neu5Ac, tales como Neu5Ac- α 2-6-lac y Neu5Ac- α 2-3-lac (por ejemplo, véase la Tabla 1). También se observa que la SubB WT no se une de forma detectable a Neu5Ac- α 2-6-lac. Por lo tanto, la proteína mutante de deleción T107 y E108 puede ser una proteína útil para unir o detectar glicanos Neu5Gc, tales como los glicanos de ácido N-glicolilneuramínico ligado a α 2-6 y los glicanos de ácido N-glicolilneuramínico ligado a α 2-3, y también los glicanos Neu5Ac, tales como Neu5Ac- α 2-6-lac y Neu5Ac- α 2-3-lac.

También se divulgan en la presente variantes, fragmentos y derivados de la proteína aislada. Convenientemente, las variantes, fragmentos y derivados de la proteína aislada conservan la capacidad de unirse a glicanos que terminan en Neu5Gc ligado a α 2-3 y ligado a α 2-6. En realizaciones particulares, se trata de al menos el 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de la capacidad de la proteína aislada de una proteína aislada divulgada en la presente para unirse a glicanos que terminan en Neu5Gc ligados a α 2-3 y ligados a α 2-6.

Como se utiliza en la presente, una “*variante*” peptídica tiene al menos un 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia con una secuencia de aminoácidos de una proteína aislada divulgada en la presente. La “*variante*” peptídica divulgada en la presente puede tener uno o más aminoácidos delecionados o sustituidos por aminoácidos diferentes. Es bien sabido en la técnica que algunos aminoácidos pueden sustituirse o delecionarse sin cambiar la actividad biológica del péptido (sustituciones conservativas).

Los términos utilizados en general en la presente para describir las relaciones de secuencia entre las proteínas y los ácidos nucleicos respectivos incluyen “*ventana de comparación*”, “*identidad de secuencia*”, “*porcentaje de identidad de secuencia*” e “*identidad sustancial*”. Debido a que los ácidos nucleicos/proteínas respectivos pueden comprender cada uno (1) sólo una o más porciones de una secuencia completa de ácido nucleico/proteína que comparten los ácidos nucleicos/proteínas, y (2) una o más porciones que son divergentes entre los ácidos nucleicos/proteínas, las comparaciones de secuencias se realizan típicamente comparando secuencias sobre una “*ventana de comparación*” para identificar y comparar regiones locales de similitud de secuencias. Una “*ventana de comparación*” se refiere a un segmento conceptual de típicamente 6, 9 o 12 residuos contiguos que se compara con una secuencia de referencia. La ventana de comparación puede comprender adiciones o delecciones (es decir, huecos) de aproximadamente un 20 % o menos en comparación con la secuencia de referencia para un alineamiento óptimo de las secuencias respectivas. El alineamiento óptimo de las secuencias para alinear una ventana de comparación puede realizarse mediante implementaciones informáticas de algoritmos (programa Geneworks de Intelligenetics; GAP, BESTFIT, FASTA y TFASTA en el paquete de software Wisconsin Genetics Software Package Liberación 7,0, Genetics Computer Group, 575 Science Drive Madison, WI, EE. UU.), o por inspección y el mejor alineamiento (es decir, el que tenga como resultado el mayor porcentaje de homología sobre la ventana de comparación) generado por cualquiera de los diversos métodos seleccionados.

La similitud y la identidad de secuencias se definen comúnmente con referencia al algoritmo GAP (Wisconsin Package, Accelrys, San Diego EE. UU.). GAP utiliza el algoritmo de Needleman y Wunsch para alinear dos secuencias completas de forma que se maximice el número de coincidencias y se minimice el número de huecos. En general, se utilizan los parámetros por defecto, con una penalización por creación de hueco = 12 y una penalización por ampliación de hueco = 4.

También se hace referencia a la familia de algoritmos BLAST, que utiliza el método de Altschul et al. (1990) J. Mol. Biol. 215: 405-410), el algoritmo psi-Blast (Nucl. Acids Res. (1997) 25 3389-3402), FASTA (que utiliza el método de Pearson y Lipman (1988) PNAS USA 85: 2444-2448), el algoritmo Smith-Waterman (Smith y Waterman (1981) J. Mol Biol. 147: 195-197), o el programa TBLASTN, de Altschul et al. (1990) *supra*, empleando generalmente parámetros por defecto.

Se encontrará una discusión detallada sobre el análisis de secuencias en la Unidad 19,3 de CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Eds. Ausubel et al. (John Wiley & Sons Inc NY, 1995-2015).

La comparación de secuencias puede realizarse sobre la longitud completa de la secuencia pertinente descrita en la presente.

El término “*identidad de secuencia*” se utiliza en la presente en su sentido más amplio para incluir el número de coincidencias exactas de nucleótidos o aminoácidos teniendo en cuenta un alineamiento adecuado mediante un algoritmo estándar, teniendo en cuenta hasta qué punto las secuencias son idénticas a lo largo de una ventana de comparación. Por lo tanto, un “*porcentaje de identidad de secuencia*” se calcula comparando dos secuencias óptimamente alineadas durante la ventana de comparación, determinando el número de posiciones en las que aparece la base de ácido nucleico idéntica (por ejemplo, A, T, C, G, I) en ambas secuencias. para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación (es decir, el tamaño de la ventana) y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia. Por ejemplo, por “*identidad de secuencia*” puede entenderse el “*porcentaje de coincidencia*” calculado por el programa informático DNASIS (Versión 2,5 para Windows; disponible sw Hitachi Software Engineering Co., Ltd., South San Francisco, California, EE. UU.).

La invención también proporciona fragmentos del péptido aislado divulgado en la presente. En algunas realizaciones, los fragmentos pueden comprender, consistir esencialmente o estar formados por 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115 aminoácidos contiguos de una proteína aislada divulgada en la presente.

También se divulgan en la presente los derivados del péptido aislado.

Tal como se utilizan en la presente, las proteínas o péptidos “*derivados*” han sido alterados, por ejemplo, por conjugación o formación de complejos con otras moléculas químicas, por modificación postraducciona (por ejemplo, fosforilación, ubiquitinación, glucosilación), modificación química (por ejemplo, reticulación, acetilación, biotilación, oxidación o reducción y similares), conjugación con etiquetas (por ejemplo, fluoróforos, enzimas, isótopos radiactivos) y/o inclusión de secuencias de aminoácidos adicionales, según se entienda en la técnica.

Con respecto a esto, se remite al experto al Capítulo 15 de CURRENT PROTOCOLS IN PROTEIN SCIENCE, Eds. Coligan et al. (John Wiley & Sons NY 1995-2015) para una metodología más amplia relativa a la modificación química de

las proteínas.

Las secuencias de aminoácidos adicionales pueden incluir secuencias de aminoácidos de compañeros de fusión que crean una proteína de fusión. A modo de ejemplo, las secuencias de aminoácidos del compañero de fusión pueden ayudar en la detección y/o purificación de la proteína de fusión aislada. Entre los ejemplos no limitativos se incluyen las proteínas de fusión de unión a metales (por ejemplo, polihistidina), la proteína de unión a la maltosa (MBP), la proteína A, la glutatión S-transferasa (GST), secuencias de proteínas fluorescentes (por ejemplo, GFP), etiquetas epitópicas, tales como las etiquetas myc, FLAG y hemaglutinina.

Los péptidos aislados de la presente invención pueden producirse por cualquier medio conocido en la técnica, incluidos, entre otros, la síntesis química y la tecnología del ADN recombinante.

La síntesis química incluye la síntesis en fase sólida y en fase de disolución. Tales métodos son bien conocidos en la técnica, aunque se hace referencia a ejemplos de técnicas de síntesis química tal como se proporcionan en el Capítulo 9 de SYNTHETIC VACCINES, Ed. Nicholson (Blackwell Scientific Publications) y Capítulo 15 de CURRENT PROTOCOLS IN PROTEIN SCIENCE, Eds. Coligan et al., (John Wiley & Sons, Inc. NY EE.UU. 1995-2014). A este respecto, también se hace referencia a la Publicación Internacional WO 99/02550 y la Publicación Internacional WO 97/45444.

Las proteínas recombinantes pueden ser convenientemente preparadas por un experto en la materia utilizando protocolos estándar como los descritos, por ejemplo, en Sambrook et al., MOLECULAR CLONING. A Laboratory Manual (Cold Spring Harbor Press, 1989), en particular las Secciones 16 y 17; CURRENT PROTOCOLS IN MOLECULAR BIOLOGY Eds. Ausubel et al., (John Wiley & Sons, Inc. NY, EE. UU. 1995-2014), en particular los capítulos 10 y 16; y CURRENT PROTOCOLS IN PROTEIN SCIENCE Eds. Coligan et al., (John Wiley & Sons, Inc. NY EE.UU. 1995-2014), en particular los Capítulos 1, 5 y 6.

Un aspecto de la invención proporciona un método para detectar el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6, en donde dicho método incluye el paso de combinar la proteína aislada o el fragmento divulgado en la presente con una muestra para formar así un complejo detectable que comprenda la proteína aislada o el fragmento del primer aspecto y el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6.

El ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6 pueden ser componentes de glicanos expresados por células tumorales, y ciertas células sanguíneas, tales como las células sanguíneas felinas. En realizaciones particulares, los glicanos que comprenden ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6 pueden ser expresados por carcinomas humanos, detectándose una expresión elevada en el cáncer de mama, ovario, próstata, colon y pulmón. En otras realizaciones particulares, el ácido *N*-glicolilneuramínico define el grupo sanguíneo "A" de los felinos, mientras que el ácido *N*-acetilneuramínico define el grupo sanguíneo "B" de los felinos.

En consecuencia, la proteína aislada o el fragmento pueden utilizarse para detectar la presencia de células tumorales que expresen ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6 en una muestra de un paciente, tal como una biopsia, una muestra de fluido, un frotis o similares. En otra realización, la proteína aislada o el fragmento pueden utilizarse para la tipificación sanguínea felina mediante la detección de células sanguíneas que expresen glicanos de ácido *N*-glicolilneuramínico, tales como los que comprenden ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6, en una muestra de sangre felina.

En otras realizaciones particulares, los glicanos que comprenden ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6 pueden estar presentes en una preparación o formulación que comprenda fármacos, anticuerpos u otras biomoléculas terapéuticas para administración humana.

Los fármacos glicosilados recombinantes, los anticuerpos y otras biomoléculas terapéuticas para administración humana se producen a menudo en líneas celulares de mamíferos no humanos que pueden sintetizar y/o incorporar metabólicamente el ácido siálico no humano ácido *N*-glicolilneuramínico (Neu5Gc). Algunos humanos tienen altos niveles de anticuerpos anti-Neu5Gc circulantes. A modo de ejemplo, el mAb antiEGFR Cetuximab, clínicamente eficaz, puede tener Neu5Gc unido covalentemente. Los anticuerpos anti-Neu5Gc de humanos normales interactúan con el Cetuximab de forma Neu5Gc-específica y generan complejos inmunes *in vitro*. Estos anticuerpos pueden potenciar la depuración del cetuximab (u otro anticuerpo terapéutico) *in vivo*. Por lo tanto, la contaminación por Neu5Gc de fármacos, anticuerpos y otras biomoléculas terapéuticas puede afectar negativamente a la vida media, la eficacia y las reacciones inmunológicas de los pacientes a los que se administran dichos fármacos. Aunque sea posible evitar la contaminación por Neu5Gc utilizando células y medios deficientes en Neu5Gc, puede que ésta no sea una disolución óptima en todos los casos. Por lo tanto, ciertas realizaciones de la invención proporcionan la detección del ácido *N*-glicolilneuramínico ligado a α 2-3 y/o del ácido *N*-glicolilneuramínico ligado a α 2-6 en fármacos, anticuerpos u otras biomoléculas terapéuticas para la administración humana. Otras realizaciones de la invención proporcionan el aislamiento, agotamiento o eliminación de contaminantes que contienen ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6 de fármacos, anticuerpos u otras biomoléculas terapéuticas para la administración humana.

En consecuencia, puede ser ventajoso acoplar, unir, fijar o ligar de otro modo la proteína aislada (por ejemplo, de una proteína aislada divulgada en la presente como la que comprende la secuencia de aminoácidos de SEQ ID NO:1), o un fragmento de la misma, a un agente que facilite la detección del ácido *N*-glicolilneuramínico ligado a α 2-3 y/o del ácido *N*-glicolilneuramínico ligado a α 2-6. En una forma preferida, la proteína aislada o el fragmento se acopla covalentemente a una etiqueta.

En otras realizaciones, puede utilizarse un agente de unión secundario etiquetado, tal como un anticuerpo o un fragmento de anticuerpo, para detectar la proteína aislada cuando se une a glicanos que comprenden ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6.

De acuerdo con cualquiera de las realizaciones anteriores, una etiqueta puede seleccionarse de entre un grupo que incluya un cromógeno, un catalizador, biotina, avidina, digoxigenina, una enzima, un fluoróforo, una molécula quimioluminiscente o un radioisótopo, aunque sin limitación al respecto.

El fluoróforo puede ser, por ejemplo, isotiocianato de fluoresceína (FITC), tintes Alexa, isotiocianato de tetrametilrhodamina (TRITL), alofocianina (APC), rojo Texas, FAM, ROX, Cy5, Cy3, o R-ficoeritrina (RPE) aunque sin limitación de los mismos.

La enzima puede ser peroxidasa de rábano picante (HRP), fosfatasa alcalina (AP), β -galactosidasa o glucosa oxidasa, aunque sin limitación de las mismas. Entre los sustratos apropiados se incluyen la diaminobanzidina (DAB), el rojo permanente, el ácido 3-etilbenzotiazolina sulfónico (ABTS), el 5-bromo-4-cloro-3-indolil fosfato (BCIP), el azul de nitro tetrazolio (NBT), la 3,3',5,5'-tetrametil bencidina (TNB) y el 4-cloro-1-naftol (4-CN), aunque sin limitación de los mismos. Un ejemplo no limitativo de sustrato quimioluminiscente es el Luminol™, que se oxida en presencia de HRP y peróxido de hidrógeno para formar un producto de estado excitado (3-aminofalato).

Las etiquetas radioisotópicas pueden incluir ^{125}I , ^{131}I , ^{51}Cr y ^{99}Tc , aunque sin limitación de las mismas.

En el caso de una etiqueta visual directa, puede utilizarse una partícula coloidal metálica o no metálica, una partícula de tinte, un polímero orgánico, una partícula de látex, un liposoma, una minicélula u otra vesícula que contenga una sustancia productora de señales y similares.

La proteína aislada etiquetada o el fragmento pueden utilizarse en sistemas de detección, tales como la histoquímica, la citometría de flujo, la microscopía de fluorescencia y los ensayos ELISA, la imagenología corporal (por ejemplo, las tomografías por emisión de positrones) y la medicina nuclear, aunque sin limitación de los mismos.

En un aspecto adicional, la invención proporciona un método de aislar un glicano o una célula que expresa el glicano, en donde el glicano comprende ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6, en donde dicho método incluye los pasos de: combinar la proteína aislada o el fragmento divulgado en la presente con una muestra para formar así un complejo que comprenda la proteína aislada o el fragmento y el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6; y aislar el glicano o la célula.

En este contexto, el término "aislar" se refiere preferentemente a purificar, enriquecer o agotar o eliminar el glicano que comprende el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6, o las células que lo expresan.

En algunas realizaciones, la proteína aislada (por ejemplo, que comprende la secuencia de aminoácidos de SEQ ID NO: 1), o un fragmento del mismo, se acopla a una etiqueta como la descrita anteriormente, que facilita la detección del glicano que comprende el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6, o las células que lo expresan. Un ejemplo no limitativo incluye una etiqueta fluorescente (como la descrita anteriormente) que facilita la clasificación por citometría de flujo de células tumorales o células sanguíneas felinas.

En otra realización, la proteína aislada divulgada en la presente, o un fragmento de la misma, puede acoplarse, unirse, fijarse o enlazarse de otro modo a un sustrato que facilite el aislamiento, enriquecimiento, purificación, agotamiento o eliminación de un glicano que comprenda ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6, o células que lo expresen.

La proteína aislada divulgada en la presente, o un fragmento de la misma, puede acoplarse, unirse, fijarse o ligarse de otro modo a un sustrato que puede ser una perla, una matriz, un polímero reticulado, un gel, una partícula, una superficie u otro sustrato sólido o semisólido. En realizaciones particulares, el sustrato puede ser o comprender sefarosa, agarosa, proteína A, proteína G, una perla magnética, una partícula paramagnética o la superficie de un chip sensor (por ejemplo, para BIAcore o resonancia de plasmón superficial). De manera adecuada, una muestra comprende una mezcla de moléculas que pueden comprender, o que se sospecha que comprenden, ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6, o células que lo expresan.

En ciertas realizaciones, la proteína aislada divulgada en la presente o un fragmento de la misma, acoplada, unida, fijada

o ligada de otro modo a un sustrato puede ser adecuada para la cromatografía (por ejemplo, cromatografía de afinidad), agotamiento por perlas magnéticas u otras técnicas que faciliten el aislamiento, enriquecimiento, purificación, agotamiento o eliminación de un glicano que comprenda ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6, o células que lo expresen.

En una realización particular, el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6 pueden estar presentes como contaminantes en una muestra, por lo que el complejo formado entre la proteína aislada y el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6 elimina los contaminantes de la muestra. Un ejemplo particular es una preparación o formulación que comprende fármacos glicosilados recombinantes, anticuerpos y otras biomoléculas terapéuticas para administración humana, como se ha descrito anteriormente.

La proteína aislada divulgada en la presente, tal como la que comprende la secuencia de aminoácidos de SEQ ID NO:1, o un fragmento puede ser adecuada para el suministro dirigido de compuestos anticancerígenos a células tumorales que expresan glicanos que comprenden ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6. La capacidad de unirse tanto al ácido *N*-glicolilneuramínico ligado a α 2-3 como al ácido *N*-glicolilneuramínico ligado a α 2-6 significa que la presente invención proporciona un sistema de suministro dirigido mucho más eficaz que el que podría proporcionarse utilizando una proteína SubB de tipo silvestre.

Por consiguiente, otro aspecto de la invención proporciona una proteína aislada o un fragmento de la misma de acuerdo con la invención para su uso en el tratamiento del cáncer en un sujeto para así dirigirse selectivamente a una célula cancerosa que exprese un ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6.

Como se ha descrito anteriormente, algunas células tumorales expresan glicanos que comprenden ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6, mientras que las células normales no suelen expresar estos azúcares. En realizaciones particulares, los glicanos que comprenden ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6 pueden ser expresados por carcinomas humanos, detectándose una expresión elevada en el cáncer de mama, ovario, próstata, colon y pulmón, aunque sin limitación al respecto.

En una realización, la proteína aislada o el fragmento pueden acoplarse, unirse, fijarse o enlazarse de otro modo a un agente citotóxico que facilite la unión a células tumorales que expresen ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6, así como su destrucción o desactivación.

El agente citotóxico puede ser un radionúclido, un fármaco quimioterapéutico, un mutágeno, una toxina, un inhibidor de la mitosis u otro agente antiproliferativo, un agente proapoptótico, un agente intercalante del ADN, o cualquier otro agente que ayude o provoque la muerte o desactivación de las células tumorales.

Ejemplos no limitativos de radionúclidos incluyen ^{211}At , ^{212}Bi , ^{213}Bi , ^{125}I , ^{111}In , ^{90}Yt , ^{193}Pt , ^{177}Lu , ^{134}Eu y ^{67}Ga , aunque sin limitación de los mismos.

Los fármacos quimioterapéuticos, mutágenos, toxinas, inhibidores de la mitosis, agentes proapoptóticos y agentes intercalantes del ADN pueden incluir la doxorrubicina, la *N*-acetil-y-calicheamicina, los maytansinoides, los taxoides, las auristatinas y las duocarmicinas, aunque sin limitación de los mismos. Los fármacos quimioterapéuticos, mutágenos, toxinas, inhibidores de la mitosis, agentes proapoptóticos y agentes intercalantes del ADN pueden acoplarse a la proteína aislada o al fragmento mediante un enlazador escindible o no escindible para formar un conjugado escindible. Típicamente, el conjugado escindible es internalizado por la célula tumoral, en donde el enlazador escindible se escinde para liberar el fármaco en la célula. En el caso de los enlazadores no solubles, éstos pueden ser preferibles cuando sea esencial que el fármaco esté totalmente localizado en la célula tumoral objetivo y no haya "fugas" del fármaco de la célula tumoral objetivo a células, tejidos o fluidos adyacentes. En algunas realizaciones, los fármacos quimioterapéuticos, mutágenos, toxinas, inhibidores de la mitosis, proapoptóticos y agentes intercalantes del ADN pueden estar en forma de profármaco que se activa al internalizarse en el interior de una célula tumoral objetivo.

En las realizaciones relativas a los usos terapéuticos, la proteína aislada (por ejemplo, que comprende la secuencia de aminoácidos de SEQ ID NO:1), o un fragmento de la misma, puede administrarse como una composición farmacéutica.

De manera adecuada, la composición farmacéutica comprende un portador, diluyente o excipiente farmacéuticamente aceptable.

Por "*portador, diluyente o excipiente farmacéuticamente aceptable*" se entiende un material de relleno sólido o líquido, diluyente o sustancia encapsulante que pueda utilizarse con seguridad en la administración sistémica. Dependiendo de la vía de administración particular, se pueden usar una variedad de portadores bien conocidos en la técnica. Estos portadores pueden seleccionarse entre un grupo que incluya azúcares, almidones, celulosa y sus derivados, malta, gelatina, talco, sulfato cálcico, liposomas y otros portadores a base de lípidos, aceites vegetales, aceites sintéticos, polioles, ácido alginico, disoluciones amortiguadoras de fosfatos, emulsionantes, disolución salina isotónica y sales como sales de ácidos minerales, incluidos clorhidratos, bromuros y sulfatos, ácidos orgánicos, tales como acetatos, propionatos y malonatos y agua libre de pirógenos.

Una referencia útil que describe portadores, diluyentes y excipientes farmacéuticamente aceptables es Remington's Pharmaceutical Sciences (Mack Publishing Co. N.J. EE.UU., 1991).

- 5 Puede emplearse cualquier vía de administración segura para proporcionar a un paciente la composición de la invención. Por ejemplo, pueden emplearse por vía oral, rectal, parenteral, sublingual, bucal, intravenosa, intraarticular, intramuscular, intradérmica, subcutánea, inhalatoria, intraocular, intraperitoneal, intracerebroventricular, transdérmica y similares. La inyección intramuscular y subcutánea es apropiada, por ejemplo, para la administración de composiciones inmunoterapéuticas, vacunas proteínicas y vacunas de ácidos nucleicos.
- 10 Las formas de dosificación incluyen comprimidos, dispersiones, suspensiones, inyecciones, disoluciones, jarabes, troqueles, cápsulas, supositorios, aerosoles, parches transdérmicos y similares. Estas formas de dosificación también pueden incluir la inyección o el implante de dispositivos de liberación controlada diseñados específicamente para este fin u otras formas de implantes modificados para actuar adicionalmente de esta manera. La liberación controlada del agente terapéutico puede efectuarse mediante el recubrimiento del mismo, por ejemplo, con polímeros hidrófobos, incluyendo
- 15 resinas acrílicas, ceras, alcoholes alifáticos superiores, ácidos poliláctico y poliglicólico, y ciertos derivados de la celulosa, tal como la hidroxipropilmetilcelulosa. Además, la liberación controlada puede efectuarse utilizando otras matrices poliméricas, liposomas y/o microesferas.
- 20 Las composiciones de la presente invención adecuadas para la administración oral o parenteral pueden presentarse como unidades separadas, tales como cápsulas, sobres o comprimidos que contengan cada uno una cantidad predeterminada de uno o más agentes terapéuticos de la invención, como un polvo o gránulos o como una disolución o una suspensión en un líquido acuoso, un líquido no acuoso, una emulsión de aceite en agua o una emulsión líquida de agua en aceite. Dichas composiciones pueden prepararse por cualquiera de los métodos de farmacia, pero todos los métodos incluyen el
- 25 paso de poner en asociación uno o varios de los agentes descritos anteriormente con el portador que constituye uno o varios ingredientes necesarios. En general, las composiciones se preparan mezclando uniforme e íntimamente los agentes de la invención con portadores líquidos o sólidos finamente divididos, o ambos, y luego, si es necesario, dando forma al producto en la presentación deseada.
- 30 Las composiciones anteriores pueden administrarse de una manera compatible con la formulación de la dosis y en la cantidad que sea farmacéuticamente eficaz. La dosis administrada a un paciente, en el contexto de la presente invención, debe ser suficiente para efectuar una respuesta beneficiosa en un paciente durante un periodo de tiempo apropiado. La cantidad de agente(s) a administrar puede depender del sujeto a tratar, incluyendo la edad, el sexo, el peso y el estado general de salud del mismo, factores que dependerán del criterio del profesional.
- 35 También se describe en la presente un anticuerpo o fragmento de anticuerpo que se une a la proteína aislada o a un fragmento de la misma. De manera adecuada, el anticuerpo o fragmento de anticuerpo no se une a una proteína SubB de tipo silvestre (como la que comprende la secuencia de aminoácidos de la SEQ ID NO:2), o se une con una afinidad al menos 5 o 10 veces menor en comparación con la afinidad con la que se une a la proteína aislada o fragmento divulgado en la presente (como la que comprende la secuencia de aminoácidos de la SEQ ID NO:1). De forma adecuada, el
- 40 anticuerpo o fragmento de anticuerpo se une a un epítipo de SEQ ID NO:1 que comprende uno o más residuos de aminoácidos modificados de la secuencia de aminoácidos TTSTE.
- 45 Tal como se utiliza en la presente, un "*anticuerpo*" es o comprende una inmunoglobulina. El término "*inmunoglobulina*" incluye cualquier producto proteico de unión a antígeno de un complejo génico de inmunoglobulina de mamífero, incluidos los isotipos de inmunoglobulina IgA, IgD, IgM, IgG e IgE y sus fragmentos de unión a antígeno. Se incluyen en el término "*inmunoglobulina*" las inmunoglobulinas que son quiméricas o humanizadas o que comprenden de otro modo residuos de aminoácidos, secuencias y/o glicosilación alterados o variantes, ya sean naturales o producidos por intervención humana (por ejemplo, mediante tecnología de ADN recombinante).
- 50 Los fragmentos de anticuerpos incluyen fragmentos Fab y Fab'2, diacuerpos y fragmentos de anticuerpos de cadena única (por ejemplo, scVs), aunque sin limitación a los mismos. Típicamente, un anticuerpo comprende regiones variables respectivas de cadena ligera y cadena pesada que comprenden cada una secuencias de aminoácidos CDR 1, 2 y 3. El anticuerpo o fragmento de anticuerpo puede comprender al menos una parte de una secuencia de aminoácidos CDR1, 2 y/o 3. Un fragmento de anticuerpo preferido comprende al menos una región variable CDR de la cadena ligera completa
- 55 y/o al menos una región variable CDR de la cadena pesada completa.
- 60 Los anticuerpos y los fragmentos de anticuerpos pueden ser policlonales o, preferiblemente, monoclonales. Los anticuerpos monoclonales pueden producirse utilizando el método estándar como, por ejemplo, el descrito en un artículo de Köhler & Milstein, 1975, Nature 256, 495, o mediante modificaciones más recientes del mismo como por ejemplo las descritas en Capítulo 2 de Coligan et al., CURRENT PROTOCOLS IN IMMUNOLOGY, inmortalizando células del bazo u otras células productoras de anticuerpos derivadas de una especie de producción a la que se haya inoculado la proteína aislada (por ejemplo, que comprenda la secuencia de aminoácidos de SEQ ID NO: 1) o un fragmento de la misma. También se entenderá que los anticuerpos pueden producirse como anticuerpos sintéticos recombinantes o fragmentos de anticuerpos expresando un ácido nucleico que codifique el anticuerpo o el fragmento de anticuerpo en una célula
- 65 hospedera apropiada. El anticuerpo sintético recombinante o el fragmento de anticuerpo de cadena pesada y ligera pueden coexpresarse a partir de diferentes vectores de expresión en la misma célula hospedera o expresarse como un

anticuerpo de cadena única en una célula hospedera. Se proporcionan ejemplos no limitativos de técnicas de expresión y selección de anticuerpos recombinantes en Capítulo 17 de Coligan et al., CURRENT PROTOCOLS IN IMMUNOLOGY *supra* y Zuberbuhler et al., 2009, Protein Engineering, Design & Selection 22 169.

5 Los anticuerpos y los fragmentos de anticuerpos pueden modificarse para que sean administrables a una especie habiendo sido producidos en, u originarios de, otra especie sin provocar una respuesta inmunitaria deletérea al anticuerpo "extraño". En el contexto de los humanos, se trata de la "humanización" del anticuerpo producido en, u originario de, otra especie. Tales métodos son bien conocidos en la técnica y generalmente implican el "injerto" recombinante de regiones determinantes de complementariedad (CDR) de anticuerpos no humanos en un andamiaje o cadena principal de anticuerpos humanos.

El anticuerpo o fragmento de anticuerpo puede estar etiquetado. Las etiquetas pueden ser como las descritas anteriormente.

15 La invención también proporciona un ácido nucleico aislado que codifica la proteína aislada divulgada en la presente, o un fragmento de la misma. En una realización, el ácido nucleico aislado codifica la proteína aislada que comprende la secuencia de aminoácidos establecida en SEQ ID NO:1, o un fragmento de la misma.

20 El término "ácido nucleico", tal y como se utiliza en la presente, designa el ADN y el ARN monocatenarios o bicatenarios. El ADN incluye el ADN genómico y el ADNc. El ARN incluye ARNm, ARN, ARNi, ARNip, ARNc y ARN autocatalítico. Los ácidos nucleicos también pueden ser híbridos de ADN-ARN. Un ácido nucleico comprende una secuencia de nucleótidos que típicamente incluye nucleótidos que comprenden una base A, G, C, T o U. Sin embargo, las secuencias de nucleótidos pueden incluir otras bases, tales como la inosina, la metilcitosina, la metilinosina, la metiladenosina y/o la tiouridina, aunque sin limitación alguna.

25 También se contemplan variantes del ácido nucleico aislado.

En una realización, una variante de ácido nucleico puede tener al menos un 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia de nucleótidos con una secuencia de nucleótidos que codifica SEQ ID NO:1.

En otra realización, una variante de ácido nucleico puede hibridarse con una secuencia de nucleótidos que codifica la SEQ ID NO:1 en condiciones de alta rigurosidad.

35 Las condiciones de alta rigurosidad son bien conocidas en la técnica, tales como las descritas en los capítulos 2,9 y 2,10 de Ausubel et al. CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (John Wiley & Sons NY, 1995-2015) y en particular en las páginas 2,9.1 a 2,9.20.

40 Generalmente, la rigurosidad depende de la concentración de uno o más factores durante la hibridación y/o el lavado. Dichos factores pueden incluir la fuerza iónica, el tipo y/o la concentración de detergente, la temperatura, el tiempo, el tipo y/o la concentración de desnaturalizante, como es bien sabido en la técnica.

Los ejemplos específicos y no limitativos de condiciones de alta rigurosidad se incluyen:

45 (i) de al menos aproximadamente 31 % v/v a al menos aproximadamente 50 % v/v de formamida y de al menos aproximadamente 0,01 M a al menos aproximadamente 0,15 M de sal para la hibridación a 42 °C, y de al menos aproximadamente 0,01 M a al menos aproximadamente 0,15 M de sal para el lavado a 42 °C;
(ii) BSA al 1 %, 1 mM EDTA, 0,5 MNaHPO₄ (pH 7,2), SDS al 7 % para la hibridación a 65 °C, y (a) 0,1 x SSC, SDS al 0,1 %; o (b) BSA al 0,5 %, 1mM EDTA, 40 mMNaHPO₄ (pH 7,2), SDS al 1 % para el lavado a una temperatura superior a 65 °C durante aproximadamente una hora; y
50 (iii) 0,2 x SSC, SDS al 0,1 % para el lavado a 68 °C o más durante aproximadamente 20 minutos.

En general, el lavado se realiza a $T_m = 69,3 + 0,41 (G + C) \% - 12$ °C. En general, la T_m de un ADN dúplex disminuye aproximadamente 1 °C con cada aumento del 1 % en el número de bases con falta de concordancia.

55 En un aspecto, el ácido nucleico aislado se encuentra en un constructo genético que comprende el ácido nucleico aislado enlazado o conectado operativamente a otro u otros componentes genéticos. Un constructo genético puede ser adecuado para el suministro terapéutico del ácido nucleico aislado o para la producción de proteínas recombinantes en una célula hospedera.

60 En términos generales, el constructo genético tiene la forma de, o comprende componentes genéticos de, un plásmido, un bacteriófago, un cósmido, un cromosoma artificial de levadura o bacteriano, tal y como se conocen bien en la técnica. Los constructos genéticos pueden ser adecuados para el mantenimiento y la propagación del ácido nucleico aislado en bacterias u otras células hospedadoras, para su manipulación mediante la tecnología de ADN recombinante y/o la expresión de ácido nucleico o de una proteína codificada de la invención.

65 A efectos de la expresión en células hospedadoras, el constructo genético es un constructo de expresión. De forma

adecuada, el constructo de expresión comprende el ácido nucleico de la invención ligado operativamente a una o más secuencias adicionales en un vector de expresión. Un “vector de expresión” puede ser un vector extracromosómico autorreplicante, tal como un plásmido, o un vector que se integra en un genoma hospedero.

5 Por “ligado operativamente” se entiende que dicha secuencia o secuencias nucleotídicas adicionales están posicionadas con respecto al ácido nucleico de la invención preferentemente para iniciar, regular o controlar de otro modo la transcripción.

10 Las secuencias de nucleótidos reguladoras serán generalmente apropiadas para la célula hospedera utilizada para la expresión. Se conocen en la técnica numerosos tipos de vectores de expresión apropiados y secuencias reguladoras adecuadas para una gran variedad de células hospedadoras.

15 Típicamente, dichas una o más secuencias nucleotídicas reguladoras pueden incluir, entre otras, secuencias promotoras, secuencias líder o señal, sitios de unión ribosómica, secuencias de poliadenilación, secuencias de inicio y terminación transcripcional, secuencias de inicio y terminación traslacional, y secuencias potenciadoras o activadoras.

La invención contempla los promotores constitutivos, reprimibles o inducibles como son conocidos en la técnica.

20 El constructo de expresión también puede incluir una secuencia de nucleótidos adicional que codifique un compañero de fusión (normalmente proporcionado por el vector de expresión) de modo que la proteína recombinante se exprese como una proteína de fusión, tal como se ha descrito anteriormente.

25 El constructo de expresión también puede incluir una secuencia de nucleótidos adicional que codifique un marcador de selección como amp^R, neo^R o kan^R, aunque sin limitación de los mismos.

30 En realizaciones particulares relacionadas con el suministro de ácidos nucleicos aislados a una herida o a un sujeto, el constructo de expresión puede estar en forma de ADN plasmídico, que comprende adecuadamente un promotor operable en una célula animal (por ejemplo, un CMV, un promotor de SV40 o α A-cristalina). En otras realizaciones, el ácido nucleico puede estar en forma de un constructo viral, tal como un vector viral adenoviral, vaccinia, lentiviral o adenoasociado.

En otro aspecto, la invención proporciona una célula hospedera transformada con una molécula de ácido nucleico o un constructo genético descrito en la presente.

35 Las células hospedadoras adecuadas para la expresión pueden ser procariotas o eucariotas. Por ejemplo, las células hospedadoras adecuadas pueden incluir, entre otras, células de mamífero (por ejemplo, células HeLa, Cos, NIH-3T3, HEK293T, Jurkat), células de levadura (por ejemplo, *Saccharomyces cerevisiae*), células de insecto (por ejemplo, Sf9, *Trichoplusia ni*) utilizadas con o sin un sistema de expresión de baculovirus, células vegetales (por ejemplo, *Chlamydomonas reinhardtii*, *Phaeodactylum tricornutum*) o células bacterianas, tal como *E. coli*. La introducción de constructos genéticos en células hospedadoras (ya sean procariotas o eucariotas) es bien conocida en la técnica, como, por ejemplo, se describe en CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Eds. Ausubel et al., (John Wiley & Sons, Inc. 1995-2015), en particular los Capítulos 9 y 16.

45 Los aspectos relacionados de la invención proporcionan kits que comprenden la proteína aislada o fragmento de la misma, ácido nucleico aislado, y/o constructo genético, tal como para la expresión de la proteína aislada o fragmento, su uso en la detección del ácido *N*-glicolilneuramínico ligado a α 2-3 y/o del ácido *N*-glicolilneuramínico ligado a α 2-6 o en el objetivo terapéutico de células tumorales que expresen el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6, aunque sin limitación de los mismos.

50 A modo de ejemplo, los kits para detectar el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6 pueden comprender la proteína aislada o el fragmento, que puede estar etiquetado o no etiquetado, opcionalmente un agente de unión secundario etiquetado como un anticuerpo que se une a la proteína aislada o al fragmento, opcionalmente uno o más sustratos para enzimas, tales como AP o HRP, e instrucciones de uso.

55 En otro ejemplo, los kits para la expresión de la proteína o fragmento aislado pueden comprender un constructo genético que codifique la proteína o fragmento aislado, células hospedadoras adecuadas para la transfección y expresión de la proteína aislada e instrucciones de uso.

Para que la invención pueda comprenderse fácilmente y llevarse a la práctica, se hace referencia a los siguientes Ejemplos no limitativos.

60 Ejemplos

Introducción

65 Debido a su conocida implicación en el cáncer y a su nivel normalmente bajo en los tejidos humanos no cancerosos, la detección de una gran cantidad de Neu5Gc en el suero y en los tejidos se consideraría anormal y sería indicativa de la

presencia de un tumor. Esto plantea la posibilidad de explotar la especificidad de la SubB para Neu5Gc con el fin de desarrollar una prueba de cribado diagnóstico de alto rendimiento para una serie de cánceres. Sin embargo, la escasa afinidad por el Neu5Gc ligado a α 2-6 podría repercutir en la sensibilidad de una prueba de este tipo. En el presente estudio, hemos examinado la interacción entre la SubB y los glicanos que terminan en Neu5Gc ligado a α 2-3, o α 2-6, con vistas a diseñar un mutante SubB con capacidad para reconocer ambos tipos de estructuras con alta afinidad.

Resultados

Mutación guiada por la estructura del sitio de unión del glicano de la SubB

Para comprender la base molecular de la preferencia por las estructuras ligadas a α 2-3, hemos comparado la interacción entre SubB y Neu5Gc α 2-3Gal β 1-3GlcNAc (determinada por cristalografía de rayos X) frente a Neu5Gc α 2-6Gal β 1-3Glc (Figura 1). Mientras que los azúcares subterminales del primer glicano se extienden libremente hacia el disolvente, tal y como se informó anteriormente⁸, el azúcar terciario de la estructura α 2-6 se pliega de nuevo sobre la superficie SubB, haciendo estrecho contacto con un bucle que comprende los residuos de SubB T104-E108. Este bucle está estabilizado por un enlace disulfuro entre C103 y C109. El impedimento estérico resultante distorsiona el acoplamiento del Neu5Gc terminal en el bolsillo de unión, lo que explica la unión significativamente más deficiente de las estructuras Neu5Gc enlazadas con α 2-6 observada en el análisis original de la matriz de glicanos.

Dado que los ácidos siálicos ligados a α 2-6 son marcadores comunes del cáncer de colon^{19,20} y están ligados al pronóstico en una serie de cánceres²¹, utilizamos la ingeniería molecular para mejorar la unión de las estructuras de Neu5Gc ligado a α 2-6 a SubB mediante el diseño de una serie de mutantes de sustitución y/o delección para reducir la altura del bucle T104-E108. Hemos modelizado las interacciones entre estos mutantes SubB y Neu5Gc α 2-6Gal β 1-3Glc y predecimos que habrían mejorado el reconocimiento de Neu5Gc α 2-6-ligado sin impactar significativamente en el enlace de Neu5Gc ligado a α 2-3, como se muestra en la Figura 2. A continuación, construimos genes subB recombinantes y expresamos y purificamos las distintas proteínas como proteínas de fusión marcadas con His6 terminal C a partir de *E. coli* recombinante (véanse los Métodos). Se purificaron con éxito proteínas SubB con sustituciones de aminoácidos simples o dobles (T107A y S106A/T107A), un mutante de delección doble (Δ S106/ Δ T107) y un mutante triple (Δ S106/ Δ T107/E108D).

Resonancia de plasmón superficial de mutantes SubB modificados por ingeniería

A continuación, la SubB purificada y los distintos derivados mutantes se inmovilizaron en chips Biacore y se probaron sus afinidades de unión a una serie de estructuras con terminación Neu5Ac o Neu5Gc (ácido siálico libre, ácido siálico- α 2-3-lactosa y ácido siálico- α 2-6-lactosa), así como a la glicoproteína α 1 ácida (AGP) humana y bovina, mediante resonancia de plasmón superficial (SPR) (Tabla 1). Los glicanos AGP humanos contienen Neu5Ac^{22,23} y los glicanos AGP bovinos contienen tanto Neu5Ac como Neu5Gc²³. El análisis glicoproteómico por EM (Figura 5) se realizó para confirmar la distribución de Neu5Ac y Neu5Gc en las AGP humana y bovina utilizadas en el estudio SPR. Se descubrió que la SubB de tipo silvestre tenía una gran afinidad por la Neu5Gc-lactosa ligada a α 2-3 y la Neu5Gc libre, tal y como se predijo a partir del resultado de la matriz de glicanos, observándose afinidades de unión nanomolares. No se observó ninguna unión para la Neu5Gc-lactosa ligada a α 2-6 (probada hasta una concentración máxima de 25 μ M) y se observó una afinidad de 2,2 μ M para Neu5Ac ligado a α 2-3, una disminución de más de 300 veces en la unión en comparación con la estructura Neu5Gc equivalente. La SubB de tipo silvestre también tenía una afinidad de unión 13 veces menor por la AGP humana en comparación con la bovina. La mutación en SubB_{T107A} no tuvo un efecto significativo en la unión a ninguna de las estructuras probadas en comparación con la proteína de tipo silvestre. SubB_{S106A/T107A} presentó una mejor unión a las estructuras ligadas a α 2-6, pero esta mejora se observó tanto para Neu5Ac como para Neu5Gc. Las afinidades de rango nanomolar observadas para todos los azúcares de enlace probados revelan que la SubB_{S106A/T107A} es una lectina que reconoce bien el ácido siálico. El mutante SubB _{Δ S106/ Δ T107/E108D} presentaba un reconocimiento mejorado del Neu5Gc ligado a α 2-6 sin cambiar la unión a las estructuras de Neu5Ac ligado a α 2-6. Sin embargo, la diferencia de afinidad entre el Neu5Ac ligado a α 2-3 y el Neu5Gc ligado a α 2-3 se redujo a 50 veces en comparación con las 300 veces observadas para el tipo silvestre. El mutante SubB _{Δ S106/ Δ T107} mejoró significativamente la discriminación Neu5Gc frente a Neu5Ac en comparación con la proteína de tipo silvestre, y tuvo la capacidad de unir Neu5Gc ligado a α 2-3 y Neu5Gc ligado a α 2-6 con afinidades de unión que no fueron significativamente diferentes entre las dos estructuras (15,3 nM frente a 8,5 nM, respectivamente; $P = 0,12$). Por lo tanto, SubB _{Δ S106/ Δ T107} mostró la combinación óptima de una mayor discriminación Neu5Gc frente a Neu5Ac y la capacidad de reconocer estructuras Neu5Gc tanto ligado a α 2-3 como ligado a α 2-6. El mutante de delección SubB _{Δ T107/ Δ E108} se unió a glicanos de ácido N-glicolilneuramínico ligado a α 2-6 con una afinidad sustancialmente mayor que la proteína SubB de tipo silvestre, mientras que también se unió a glicanos de ácido N-glicolilneuramínico ligado a α 2-3. Sin embargo, a diferencia de SubB _{Δ S106/ Δ T107}, SubB _{Δ T107/ Δ E108} puede unir ampliamente glicanos Neu5Ac, tal como Neu5Ac- α 2-6-lac, que no están unidos de forma detectable ni por la SubB de tipo silvestre ni por la SubB _{Δ S106/ Δ T107}.

El anticuerpo anti-Neu5Gc producido en pollo se utilizó como control y mostró menos selectividad y menor afinidad por los glicanos que contienen Neu5Gc que cualquiera de las proteínas SubB probadas.

Análisis de matrices de glicanos de SubB de tipo silvestre, SubB_{S106A/T107A} y SubB _{Δ S106/ Δ T107}.

Para valorar si la mutación SubB _{Δ S106/ Δ T107} preferida y específica para Neu5Gc introducía especificidad para

estructuras no sialiladas, no cubiertas por el análisis SPR, se realizaron análisis de matrices de glicanos en los mutantes SubB de tipo silvestre, SubBAS106/ Δ T107 y SubBS106A/T107A (Tabla 4). La SubB de tipo silvestre sólo mostró una unión significativa a cuatro de las 402 estructuras del conjunto de glicanos; Neu5Gc α 2-3 Gal, Neu5Gc α 2-3 Gal β 1-4GlcNAc y dos estructuras terminadas en Neu5Ac α 2-3Gal β 1-4GlcNAc. Esto concuerda con el análisis de matriz de glicanos de SubB8 publicado anteriormente (www.functionalglycomics.org/glycomics/HServlet?operation=view&sideMenu=no&psId=primscreen_1579#).

SubBAS106/ Δ T107 sólo habían mostrado una unión significativa a cuatro estructuras de la matriz. Éstas se limitaron a estructuras con terminación en Neu5Gc α 2-3Gal o Neu5Gc α 2-6 Gal. SubBS106A/T107A se unió a 18 glicanos en total en la matriz, incluyendo estructuras que contenían Neu5Gc y Neu5Ac. También reconoció estructuras sulfatadas, incluidos glicosaminoglicanos (heparina y condroitina-6-sulfato) y estructuras de lactosamina sulfatada (Tabla 4). SubBS106A/T107A también reconoció un rango con carga negativa de monosacáridos (NeuSAc, Neu5Gc, 9-NAc-NeuSAc, 3-O-Su-GlcNAc) en la matriz.

ELISA de SubB modificado por ingeniería frente a proteínas/suero humano y bovino

Para valorar la capacidad de los mutantes diseñados para detectar la presencia de Neu5Gc en muestras biológicas se realizaron ensayos ELISA. Utilizando placas recubiertas con una serie de diluciones de SubB, se probaron proteínas séricas etiquetadas de origen humano y bovino. Se identificó una mejora del doble en el reconocimiento diferencial de las proteínas séricas bovinas que contienen Neu5Gc con SubBAS106/ Δ T107. (Fig. 3).

Detección de AGP humana frente a bovina

Para verificar de forma independiente la capacidad de discriminar entre AGP humana y bovina (sólo la bovina muestra niveles significativos de glicanos Neu5Gc-terminales), se aplicaron como manchas glicoproteínas diluidas en serie sobre filtros de nitrocelulosa y, tras lavarlos y bloquearlos, se superpusieron los filtros con SubBAS106/ Δ T107 biotinilado purificado. A continuación, la lectina unida se detectó en los filtros lavados utilizando estreptavidina-AP (Figura 4). La unión de SubBAS106/ Δ T107 a la AGP bovina fue detectable hasta aproximadamente 200 ng/mancha, mientras que la unión significativa a la AGP humana no fue detectable ni siquiera en la cantidad máxima probada (12,5 μ g/mancha). Este poder discriminatorio es coherente con los datos de SPR anteriores.

Matrices de glicanos adicionales

Las distintas estructuras de glicanos Neu5Ac y Neu5Gc analizadas en las matrices de glicanos de Z-biotech se muestran en la Figura 9 y un ejemplo de matriz en la Figura 6. La Tabla 3 proporciona el código que enlaza los glicanos de la Figura 9 con los datos de la matriz de las Figuras 7 y 8. Los datos de la matriz resumidos en la Figura. 7 muestran que la unión a estructuras Neu5Gc es preferida por la SubB de tipo silvestre, pero hay 4/40 glicanos Neu5Ac que se unen con más de 5000 unidades de fluorescencia por encima del fondo y 14/41 estructuras Neu5Gc que tienen una unión por debajo de 5000. Todas las estructuras Neu5Ac registran alguna unión por encima del fondo. Como también es evidente en la Figura 8, la unión a estructuras Neu5Gc es preferida por la SubB 2M. Ningún glicano Neu5Ac se une con más de 5000 unidades de fluorescencia por encima del fondo y sólo 5/41 estructuras Neu5Gc tienen una unión por debajo de 5000. Sólo 7/14 glicanos Neu5Ac tienen alguna unión por encima del fondo. Estos resultados muestran una mejora definitiva con respecto a los resultados obtenidos con la SubB WT en términos de especificidad para el Neu5Gc y de reconocimiento mejorado de los diferentes enlaces y presentaciones de los glicanos que contienen Neu5Gc.

Desarrollo ulterior del cribado en chip de proteínas que contienen Neu5Gc en suero utilizando SubB 2M.

En referencia a las Figuras 10 y 11, todos los ensayos se realizaron en un fondo de suero humano normal al 1 % obtenido de Sigma-Aldrich. La SubB de tipo silvestre no pudo analizarse ya que todas las uniones observadas procedían del suero presente, con valores que caían por debajo del control de sólo suero cuando se añadía una proteína que contenía Neu5Gc. Esto indica que la SubB WT tenía preferencia por la proteína Neu5Gc sobre el suero pero se unía al suero a niveles elevados en ausencia de la proteína (Figura 10). En la FIG. 11, la SubB 2M funcionó mucho mejor con respuestas por encima del fondo sérico a partir de concentraciones entre 31,25 nM y 62,5 nM. Esto es a una concentración de proteína de -2 μ g/mL.

Discusión

El Neu5Gc es un importante marcador diagnóstico y pronóstico en los carcinomas humanos, habiéndose detectado una elevada expresión de Neu5Gc en el cáncer de mama, ovario, próstata, colon y pulmón^{11,12}. La SubB de tipo silvestre tenía una especificidad sin precedentes para los glicanos con terminación en Neu5Gc, pero se unía de manera deficiente al Neu5Gc ligado a α 2-6 y seguía reconociendo las estructuras de Neu5Ac ligado a α 2-3, aunque débilmente⁸. Para mejorar el reconocimiento de la SubB para el Neu5Gc ligado a α 2-6 y hacerla más específica para el Neu5Gc, diseñamos la SubB utilizando modificaciones asistidas por la estructura, centrándonos específicamente en el bucle T104-E108.

La manipulación de este bucle tuvo dos resultados específicos mediante la modificación de los mismos dos aminoácidos. En primer lugar, la sustitución por alanina de S106 y T107 (S106A/T107A) condujo a una pérdida de especificidad por Neu5Gc, produciendo una lectina capaz de unirse a todos los glicanos sialilados terminales probados,

independientemente del enlace (α 2-3 y α 2-6) o del tipo de ácido siálico (NeuSAc o Neu5Gc). La segunda fue que la delección de los mismos dos aminoácidos (Δ S106/ Δ T107) produjo una lectina con una especificidad excelente para Neu5Gc independientemente del enlace (α 2-3 y α 2-6). El mutante SubB $_{\Delta$ S106/ Δ T107 mejoró significativamente en el reconocimiento de estructuras que contienen Neu5Gc en comparación con la SubB de tipo silvestre. La SubB $_{\Delta$ S106/ Δ T107 tampoco presentaba diferencias en su capacidad para unirse a estructuras Neu5Gc ligado α 2-3 o Neu5Gc ligado α 2-6, lo que la convertía en una mejora significativa con respecto a la proteína de tipo silvestre. Otras modificaciones de la proteína SubB fuera de los aminoácidos S106 y T107 no produjeron ninguna mejora significativa de la especificidad. La proteína mutante SubB $_{\Delta$ S106/ Δ T107/E108D, que es la proteína SubB $_{\Delta$ S106/ Δ T107 con una mutación E108D también añadida, era menos capaz de distinguir el Neu5Gc ligado a α 2-3 del Neu5Ac ligado a α 2-3 que la SubB $_{\Delta$ S106/ Δ T107 y tenía una unión más fuerte a la glicoproteína α 1-ácido humana que el mutante SubB $_{\Delta$ S106/ Δ T107 (24 veces más proteína unida por SubB $_{\Delta$ S106/ Δ T107/E108D que SubB $_{\Delta$ S106/ Δ T107). Por el contrario, el mutante de delección SubB $_{\Delta$ T107/ Δ E108 no sólo se unió a glicanos de ácido N-glicolilneuramínico ligado a α 2-6 y glicanos de ácido N-glicolilneuramínico ligado a α 2-3, sino también glicanos Neu5Ac, tales como Neu5Ac- α 2-6-lac y Neu5Ac- α 2-3-lac, que no están unidos de forma detectable por SubB $_{\Delta$ S106/ Δ T107.

Estos mutantes SubB mejorados ofrecen una nueva herramienta para la prueba de muestras biológicas, en particular suero y otros fluidos de individuos con cáncer o con sospechas de tener cáncer.

Métodos

Modelado estructural de SubB.

La estructura tridimensional de los mutantes SubB se modeló utilizando Phyre2²⁴. Neu5GC α 2-6Gal β 1-3Glc se adquirió de PDB ID: 4EN8²⁵ y se modeló en las estructuras mutantes SubB y SubB manualmente utilizando Coot²⁶.

Construcción y expresión de mutantes SubB.

Las mutaciones se introdujeron en la secuencia codificante de la *subB* (cerca del extremo 3') mediante PCR directa de alta fidelidad utilizando el cebador sentido pETSubBF y los cebadores antisentido específicos de las mutaciones respectivos enumerados en la Tabla 2. Los productos de la PCR se clonaron en los sitios *Bam*HI y *Xho*I de pET-23(+) (Novagen) y se transformaron en *E. coli* BL21(DE3). Los derivados SubB se expresaron y purificaron como proteínas de fusión etiquetadas con His₆ mediante cromatografía de afinidad Ni-NTA, tal y como se ha descrito previamente⁴. Las proteínas tenían una pureza superior al 95 %, según se juzgó mediante SDS-PAGE y tinción con azul de Coomassie.

Resonancia de plasmón superficial de SubB y mutantes de SubB modificados por ingeniería.

La resonancia de plasmón superficial (SPR) se llevó a cabo utilizando el sistema Biacore T100 (GE), tal y como se ha descrito anteriormente²⁷. Brevemente, SubB, mutantes SubB e IgY anti-Neu5Gc (SiaMab; anteriormente Sialix/GC-Free Inc., San Diego, CA, EE. UU.) se inmovilizaron en la celda de flujo 2-4 de un chip sensor CMS serie S (GE) utilizando el kit de captura NHS y la celda de flujo 1 se utilizó como una inmovilización en blanco. Los monosacáridos, disacáridos, oligosacáridos y la glicoproteína α 1-ácido de origen humano y bovino (Sigma-Aldrich; véase la Tabla 1) se pasaron a 0,01-100 μ M en los experimentos iniciales de determinación del rango. Las concentraciones se ajustaron y todos los datos se analizaron mediante cinética de ciclo único utilizando el software Biacore T100 Evaluation.

Análisis espectroscópico de masas de la glicoproteína α 1-ácido.

La AGP del plasma humano (Sigma-Aldrich G9885) y del plasma bovino (Sigma-Aldrich G3643) (1 mg en cloruro de guanidinio 6M, Tris-HCl 50 mM pH 8) se redujo y alquiló con ditiotretol 10 mM y acrilamida 25 mM, respectivamente. A continuación, se precipitó la proteína añadiendo 4 volúmenes de metanol:acetona 1:1, incubando a -20 °C durante 16 h y centrifugando después (18,000 rcf, 10 min) para recoger el precipitado. La proteína precipitada se resuspendió en 50 μ l de Tris-HCl 50 mM pH8 y se digirió (37 °C, 16 h) con 1 μ g de tripsina (Trypsin Gold, Promega). A continuación, los péptidos digeridos se desalaron con C18 ZipTips (Millipore).

Análisis ELISA de SubB y el mutante SubB $_{\Delta$ S106/ Δ T107 modificado por ingeniería.

Los pozos de las placas NUNC Maxisorp negras de 96 pozos se recubrieron con proteína SubB o SubB $_{\Delta$ S106/ Δ T107 diluida dos veces en serie en un amortiguador de recubrimiento de bicarbonato/carbonato 100 mM (pH 9,6) comenzando con 1,25 μ g de proteína durante una noche a 4 °C. Los pozos se lavaron 3 veces con disolución salina tamponada con fosfato, Tween-20 al 0,05 % (PBS-T) antes de añadir disolución amortiguadora (BSA al 3 %) durante 1 hora a temperatura ambiente. Las proteínas del suero humano normal y del suero bovino se marcaron con fluorescencia combinando suero puro con 100 μ M de tinte FITC (Peirce) e incubando en hielo durante 1 hora. El exceso de tinte se eliminó utilizando una columna de centrifugación de exclusión por tamaño de 1 kDa. Se añadieron 100 μ l de suero humano normal o bovino marcado con FITC a los pozos recubiertos con SubB o SubB $_{\Delta$ S106/ Δ T107 y se incubaron los pozos durante 1 hora a temperatura ambiente. Los pozos se lavaron 3 veces con PBS-T. Se añadieron 100 μ l de PBS a cada pozo antes de medir la fluorescencia a 485/535 nm. Los valores de las unidades de fluorescencia se muestran como la media de duplicados \pm SD, restando las unidades de fluorescencia medias obtenidas para los pozos que contienen todos los reactivos excepto las proteínas SubB. Cualquier valor negativo se consideró 0.

Experimentos de superposición de SubB.

La SubB_{ΔS106/ΔT107} purificada se etiquetó con biotina utilizando el kit de biotilación EZ-Link[®] Sulfo-NHS (Thermo Scientific) de acuerdo con las instrucciones del fabricante. Glicoproteína α-1 ácida humana y bovina purificada (No. de cat. Sigma: G9885 y G3643) se disolvieron en agua a 5 mg/ml y se aplicaron como manchas volúmenes de 5 µl de diluciones dobles seriadas en filtros de nitrocelulosa y se secaron al aire a 37 °C durante la noche. A continuación, los filtros se bloquearon con leche desnatada al 5 % en disolución salina amortiguada con Tris con Tween 20 al 0,05 % (TTBS) durante 2 h. Después de lavarlos tres veces en TTBS, los filtros se recubrieron con 1 µg/ml de biotina-SubB_{ΔS106/ΔT107} en TTBS y se incubaron durante toda la noche a 4 °C. A continuación, los filtros se lavaron tres veces en TTBS y la biotina-SubB_{ΔS106/ΔT107} unida se detectó utilizando el conjugado de estreptavidina-fosfatasa alcalina (Roche). Los filtros se revelaron utilizando un sistema cromogénico de sustrato de tetrazolio nitroazul/X-fosfato (Roche).

Análisis de matrices de glicanos de SubB y de mutantes de SubB modificados por ingeniería.

Para los datos mostrados en la Tabla 3, los portaobjetos de las matrices de glicanos se imprimieron en sustratos activados con SuperEpoxy 3 (Arrayit) utilizando una impresora de contacto Arrayit Spotbot Extreme como se describió previamente²⁸. Para cada submatriz se precomplejaron 2 µg de proteínas SubB con anticuerpo anti-etiqueta His (Señalización celular) y anticuerpos secundarios y terciarios Alexa555 (conejo anti-ratón; cabra anti-conejo) en una relación de 2: 1:0,5:0,25 en un volumen final de 500 µl. Este complejo proteico de anticuerpos de 500 µl se añadió a un marco genético de 65 µl (Thermo Scientific) sin un cubreobjetos. El lavado y el análisis se realizaron como se ha descrito previamente²⁷.

Matrices de glicanos Neu5Ac/Neu5Gc

Los datos mostrados en las Figuras 6-9, las matrices de glicanos Neu5Ac/Neu5Gc se obtuvieron de Z-biotech (<http://www.zbiotech.com/neu5gc-xenoantigen-microarray.html>). Las matrices se preformaron según las instrucciones del fabricante con un total de 2 µg de proteína aplicada a cada una de las áreas de la matriz. La detección se realizó con IgG anti-His de ratón (relación molar 1: 1 con proteína), IgG de conejo anti-ratón Alexa 555 (0,5 de cantidad molar de IgG de ratón) e IgG de cabra anti-conejo Alexa 555 (0,5 de cantidad molar de IgG de conejo). Las proteínas se incubaron durante 1 hora y se lavaron 3 veces en 1xPBS. Los portaobjetos se escanearon en un Innoscan 1100AL utilizando láseres de 488, 532 y 647. Las matrices se analizaron con el software Mapix. Todos los datos se tomaron del canal láser 532 y en el análisis se utilizó la fluorescencia con sustracción de fondo.

Referencias

1. Beddoe, T., Paton, A.W., Le Nours, J., Rossjohn, J. y Paton, J.C. Structure, biological functions and applications of the AB5 toxins. *Trends Biochem Sci* 35, 411-8 (2010).
2. Petr, T. et al. Histochemical detection of GM1 ganglioside using cholera toxin-B subunit. Evaluation of critical factors optimal for in situ detection with special emphasis to acetone pre-extraction. *Eur J Histochem* 54, e23 (2010).
3. Kenworthy, A.K., Petranova, N. y Edidin, M. High-resolution FRET microscopy of cholera toxin B-subunit and GPI-anchored proteins in cell plasma membranes. *Mol Biol Cell* 11, 1645-55 (2000).
4. Paton, A.W., Srimanote, P., Talbot, U.M., Wang, H. y Paton, J.C. A new family of potent AB(5) cytotoxins produced by Shiga toxigenic *Escherichia coli*. *J Exp Med* 200, 35-46 (2004).
5. Paton, A.W. et al. AB5 subtilase cytotoxin inactivates the endoplasmic reticulum chaperone BiP. *Nature* 443, 548-52 (2006).
6. Backer, J.M. et al. Chaperone-targeting cytotoxin and endoplasmic reticulum stress-inducing drug synergize to kill cancer cells. *Neoplasia* 11, 1165-73 (2009).
7. Martin, S. et al. Targeting GRP78 to enhance melanoma cell death. *Pigment Cell Melanoma Res* 23, 675-82 (2010).
8. Byres, E. et al. Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. *Nature* 456, 648-52 (2008).
9. Inoue, S., Sato, C. y Kitajima, K. Extensive enrichment of N-glycolylneuraminic acid in extracellular sialoglycoproteins abundantly synthesized and secreted by human cancer cells. *Glycobiology* 20, 752-62 (2010).
10. Malykh, Y.N., Schauer, R. y Shaw, L. N-Glycolylneuraminic acid in human tumours. *Biochimie* 83, 623-34 (2001).
11. Marquina, G. et al. Gangliosides expressed in human breast cancer. *Cancer Res* 56, 5165-71 (1996).
12. Samraj, A.N., Laubli, H., Varki, N. y Varki, A. Involvement of a non-human sialic Acid in human cancer. *Front Oncol* 4, 33 (2014).
13. Varki, N.M. y Varki, A. Diversity in cell surface sialic acid presentations: implications for biology and disease. *Lab Invest* 87, 851-7 (2007).
14. Lofling, J.C., Paton, A.W., Varki, N.M., Paton, J.C. y Varki, A. A dietary non-human sialic acid may facilitate hemolytic-uremic syndrome. *Kidney Int* 76, 140-4 (2009).
15. Yin, J. et al. Hypoxic culture induces expression of sialin, a sialic acid transporter, and cancer-associated gangliosides containing non-human sialic acid on human cancer cells. *Cancer Res* 66, 2937-45 (2006).
16. Dennis, J.W., Laferte, S., Yagel, S. y Breitman, M.L. Asparagine-linked oligosaccharides associated with metastatic cancer. *Cancer Cells* 1, 87-92 (1989).
17. Padler-Karavani, V. et al. Human xeno-autoantibodies against a non-human sialic acid serve as novel serum

biomarkers and immunotherapeutics in cancer. *Cancer Res* 71, 3352-63 (2011).

18. Pham, T. et al. Evidence for a novel human-specific xeno-auto-antibody response against vascular endothelium. *Blood* 114, 5225-35 (2009).

19. Murayama, T. et al. Colon carcinoma glycoproteins carrying alpha 2,6-linked sialic acid reactive with Sambucus nigra agglutinin are not constitutively expressed in normal human colon mucosa and are distinct from sialyl-Tn antigen. *Int J Cancer* 70, 575-81 (1997).

20. Sata, T., Roth, J., Zuber, C., Stamm, B. y Heitz, P.U. Expression of alpha 2,6-linked sialic acid residues in neoplastic but not in normal human colonic mucosa. A lectin-gold cytochemical study with Sambucus nigra and Maackia amurensis lectins. *Am J Pathol* 139, 1435-48 (1991).

21. Hedlund, M., Ng, E., Varki, A. y Varki, N.M. alpha 2-6-Linked sialic acids on N-glycans modulate carcinoma differentiation in vivo. *Cancer Res* 68, 388-94 (2008).

22. Imre, T. et al. Glycosylation site analysis of human alpha-1-acid glycoprotein (AGP) by capillary liquid chromatography-electrospray mass spectrometry. *J Mass Spectrom* 40, 1472-83 (2005).

23. Nakano, M., Kakehi, K., Tsai, M.H. y Lee, Y.C. Detailed structural features of glycan chains derived from alpha1-acid glycoproteins of several different animals: the presence of hypersialylated, O-acetylated sialic acids but not disialyl residues. *Glycobiology* 14, 431-41 (2004).

24. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. y Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10, 845-58 (2015).

25. Yamashita, S. et al. Carbohydrate recognition mechanism of HA70 from Clostridium botulinum deduced from X-ray structures in complexes with sialylated oligosaccharides. *FEBS Lett* 586, 2404-10 (2012).

26. Emsley, P., Lohkamp, B., Scott, W.G. y Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486-501 (2010).

27. Shewell, L.K. et al. The cholesterol-dependent cytolysins pneumolysin and streptolysin O require binding to red blood cell glycans for hemolytic activity. *Proc Natl Acad Sci USA* 111, E5312-20 (2014).

Tabla 1. Análisis de resonancia de plasmón superficial de proteínas de unión a Neu5Gc

Variante/anticuerpo SubB	AGP α 1 humana	AGP α 1 bovina	Neu5 Ac- α 2-3-lac	Neu5 Gc- α 2-3-lac	Neu5 Ac- α 2-6-lac	Neu5 Gc- α 2-6-lac	Neu5 Ac libre	Neu5 Gc libre	Ma nS	malt osa	Lact osa	GT2	Condro itina 6 sulfato
Anticuerpo anti-NeuSGc (IgY IgY)	<i>n.t.</i>	<i>n.t.</i>	249 \pm 46 μ M	2,34 \pm 0,85 μ M	<i>n.t.</i>	<i>n.t.</i>	NCDI	35,7 \pm 4,2 μ M	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>
SubB de tipo silvestre	2,12 \pm 0,56 μ M (Rmáx= 125)	155,8 \pm 22 nM (Rmáx= 525)	2,24 \pm 0,93 μ M	6,62 \pm 2,17 nM	NCDI	NCDI	NCDI	18,1 \pm 5,9 nM	NC DI	NCD I	NCD I	NCDI	NCDI
S106A/T107 A	723 \pm 129 nM (Rmáx= 142)	164 \pm 10 nM (Rmáx= 499)	489 \pm 171 nM	1,52 \pm 0,50 nM	348 \pm 52 nM	8,05 \pm 0,14 nM	3,27 \pm 0,29 μ M	6,61 \pm 1,6 nM	NC DI	NCD I	NCD I	8,97 \pm 2,2 μ M	33,0 \pm 7,6 μ M
T107A	<i>n.t.</i>	<i>n.t.</i>	4,18 \pm 1,6 μ M	15,2 \pm 0,02 nM	NCDI	208 \pm 123 nM	NCDI	16,8 \pm 0,99 nM	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>
ΔS106/ΔT107	1,65 \pm 0,42 μ M (Rmáx= 7)	115 \pm 37 nM (Rmáx= 299)	NCDI	15,3 \pm 5,8 nM	NCDI	8,53 \pm 0,15 nM	NCDI	17,8 \pm 4,0 nM	NC DI	NCD I	NCD I	NCDI	NCDI
ΔS106/ΔT107 E108D	2,82 \pm 0,15 μ M (Rmáx= 165)	32,5 \pm 2,6 nM (Rmáx= 276)	371 \pm 64 nM	7,39 \pm 0,72 nM	NCDI	3,45 \pm 0,87 nM	NCDI	45,1 \pm 1,2 nM	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>
AT107/AE108	308 \pm 24 nM (Rmáx= 542)**	98,8 \pm 43 nM (Rmáx= 865)**	9,65 \pm 0,70 nM	4,32 \pm 0,65 nM	4,94 \pm 0,35 nM	3,71 \pm 0,41 nM	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>
ΔS106/ΔT107 AE108*	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>	<i>n.t.</i>

*Insoluble en proteína ** Realizado en otra ocasión. Capturó 3 veces más proteína que la proteína anterior.

5

Leyenda de la Tabla 1: Las afinidades de unión de la SubB de tipo silvestre, de varios derivados mutantes y de un anticuerpo IgY anti-Neu5Gc, a tri- y monosacáridos purificados y/o a la glicoproteína $\alpha 1$ ácida (AGP) humana o bovina, se determinaron mediante SPR, tal y como se describe en los Materiales y Métodos. NCDI indica que no se observó ninguna interacción dependiente de la concentración con concentraciones de hasta 100 μ M; ND: No realizado; R_{máx}: la cantidad total de unidades de respuesta (RU) del analito unido a la proteína (cuanto mayor sea el número, mayor será la unión del glicano/glicoproteína a la SubB inmovilizada).

Tabla 2. Oligonucleótidos

Cebador	Secuencia 5'-3'
pETSubBF	TTGTAAGGATCCGGAGGTGCATATGACG (SEQ ID NO:4)
pETSubB _{T107A} R	GATTATCTCGAGTGAGTTCTTTTCTGTGTCAGGACCAAAACATTCTGCCGATG TGGTGCAGGTTG (SEQ ID NO:5)
pETSubB _{S106A/T107A} R	GATTATCTCGAGTGAGTTCTTTTCTGTGTCAGGACCAAAACATTCTGCCGCT GTGGTGCAGGTTG (SEQ ID NO:6)
pETSubB _{ΔS106/ΔT107R}	GATTATCTCGAGTGAGTTCTTTTCTGTGTCAGGACCAAAACATTCTGTGGTGC AGGTTGATAACCC (SEQ ID NO:7)
pETSubB _{ΔS106/ΔT107/E108D} R	GATTATCTCGAGTGAGTTCTTTTCTGTGTCAGGACCAAAACAGTCTGTGGTG CAGGTTGATAACCC (SEQ ID NO:8)

10

Tabla 3. Códigos de glicanos para las Figuras 7-9

ID de Glicano Gc	Glicanos Neu5Gc	ID de Glicano Ac	Glicanos Neu5Ac
GC-1	N002G	AC-1	N002
GC-2	N003G	AC-2	N003
GC-3	N005G	AC-3	N005
GC-4	N012G	AC-4	N012
GC-5	N013G	AC-5	N013
GC-6	N015G	AC-6	N015
GC-7	N022G	AC-7	N022
GC-8	N023G	AC-8	N023
GC-9	N025G	AC-9	N025
GC-10	N032G	AC-10	N032
GC-11	N033G	AC-11	N033
GC-12	N042G	AC-12	N042
GC-13	N043G	AC-13	N043
GC-14	N045G	AC-14	N045
GC-15	N052G	AC-15	N052
GC-16	N053G	AC-16	N053
GC-17	N055G	AC-17	N055
GC-18	N112G	AC-18	N112
GC-19	N113G	AC-19	N113
GC-20	N115G	AC-20	N115
GC-21	N 122G	AC-21	N 122
GC-22	N123G	AC-22	N123
GC-23	N125G	AC-23	N125

ES 2 983 352 T3

ID de Glicano Gc	Glicanos Neu5Gc	ID de Glicano Ac	Glicanos Neu5Ac
GC-24	N133G	AC-24	N133
GC-25	N134G	AC-25	N134
GC-26	N135G	AC-26	N135
GC-27	N 144G	AC-27	N144
GC-28	N145G		
GC-29	N155G	AC-29	N155
GC-30	N212G	AC-30	N212
GC-31	N213G	AC-31	N213
GC-32	N215G	AC-32	N215
GC-33	N222G	AC-33	N222
GC-34	N223G	AC-34	N223
GC-35	N225G	AC-35	N225
GC-36	N233G	AC-36	N233
GC-37	N235G		
GC-3 8	N245G		
GC-39	N255G	AC-39	N255
GC-40	N003G1		
GC-41	N003G2		

Tabla 4

Número	Estructura	SubB WT	SubBAS196/31167	SubBS466/47107A
MONOSACÁRIDOS				
1	Fucose-p3	-0,003	0,119	0,250
2	Glucose-p3	-0,533	0,117	0,058
3	Galactose-p3	0,000	0,167	0,109
4	GalNAc-p3	-0,006	0,253	0,168
5	GalNAc-p3	-0,010	0,531	0,445
6	GalNAc-p3	0,004	0,154	0,249
7	Glucose-p3	0,002	0,183	0,006
8	Galactose-p3	-0,001	0,156	0,767
10	Glucose-p3	0,001	0,128	0,748
14	Glucose-p3	0,006	0,153	0,423
15	6- α -D-GlcNAc-6S, 6- α -D-GlcNAc-6S	-0,011	0,103	0,119
16	Mannose-p3	0,016	0,172	0,122
18	Mannose-p3	-0,003	0,372	0,300
19	Mannose-p3	0,021	0,086	0,056
20	Mannose-p3	0,002	0,205	0,237
22	Glucose-p3	-0,001	0,178	0,211
17	3-O-Sul-Galactose-p3	-0,039	0,151	0,140
18	3-O-Sul-Galactose-p3	-0,002	0,123	0,027
43	6-O-Sul-GlcNAc-p3	-0,272	0,195	0,118

Valor promedio del pliegue por suscripción del fondo (fondo
prom + 3x error estándar de la media)

44	GlcA-ep3	-0,022	0,120	-0,308
45	GlcA-ep3	-0,009	0,306	0,100
46	6-EGlcA-GlcA-ep4	0,012	0,196	0,113
47	6-EGlcA-Mann-ep3	0,000	0,223	0,040
48	Neu5Ac-ep3	0,009	0,219	1,192
49	Neu5Ac-ep3	0,002	0,137	1,848
51	Neu5Glc-ep3	0,004	0,163	2,205
54	9-NAc-Neu5Ac-ep3	-0,020	0,108	1,293
55	3-O-Su-GlcNAc-ep3	-0,003	0,131	1,403
SubB WT SubB5106A7107 SubB5106A7107A				
Galactose terminal				
75	GalNAc-GalB-ep3	-0,009	0,149	0,627
76	GalNAc-GalB-ep3	0,014	0,259	0,352
77	GalNAc-GalNAc-ep3	-0,004	0,127	0,466
78	GalNAc-GalNAc-ep3	0,008	0,174	0,335
80	GalNAc-GlcNAc-ep3	0,009	0,129	0,119
81	GalNAc-GlcNAc-ep3	-0,004	0,131	0,213
83	GalNAc-GlcB-ep4	0,008	0,217	0,109
84	GalNAc-GalB-ep3	-0,024	0,139	0,056
85	GalNAc-GlcNAc-ep3	0,043	0,071	0,052
87	GalNAc-GalB-ep3	0,229	0,512	0,279
88	GalNAc-GalNAc-ep3	0,002	0,188	0,173
89	GalNAc-GalNAc-ep3	-0,003	0,134	0,240
91	GalNAc-GlcB-ep4	-0,004	0,131	0,158
94	GalNAc-GalB-ep3	-0,005	0,134	0,154
97	GalNAc-GlcNAc-ep3	-0,004	0,143	0,095
100	GalNAc-GalB-ep4	0,011	0,288	0,496
145	GalNAc-6-O-Su-GlcNAc-ep3	-0,016	0,310	0,674

146	Galβ1-4(6-O-Su)Glcβ-sp2	0.085	0.187	0.144
147	Galβ1-4(6-O-Su)GlcNAcβ-sp3	0.003	0.162	0.380
150	3-O-Su-Galβ1-3GalNAcα-sp3	0.003	0.146	0.372
151	6-O-Su-Galβ1-3GalNAcα-sp3	0.009	0.118	0.387
152	3-O-Su-Galβ1-4Glcβ-sp2	0.008	0.360	0.086
153	6-O-Su-Galβ1-4Glcβ-sp2	0.000	0.130	0.279
155	3-O-Su-Galβ1-3GlcNAcβ-sp3	-0.014	0.219	0.838
157	3-O-Su-Galβ1-4GlcNAcβ-sp3	-0.003	0.206	0.330
159	4-O-Su-Galβ1-4GlcNAcβ-sp3	0.007	0.193	0.786
161	6-O-Su-Galβ1-3GlcNAcβ-sp3	0.004	0.129	0.146
163	6-O-Su-Galβ1-4GlcNAcβ-sp3	-0.020	0.161	0.646
176	3-O-Su-Galβ1-4(6-O-Su)Glcβ-sp2	-0.007	0.110	0.102
177	3-O-Su-Galβ1-4(6-O-Su)GlcNAcβ-sp3	-0.427	0.188	0.868
178	6-O-Su-Galβ1-4(6-O-Su)Glcβ-sp2	-0.022	0.424	0.873
179	6-O-Su-Galβ1-3(6-O-Su)GlcNAcβ-sp2	0.884	0.166	0.076
180	6-O-Su-Galβ1-4(6-O-Su)GlcNAcβ-sp2	0.003	0.057	0.613
181	3,4-O-Su ₂ -Galβ1-4GlcNAcβ-sp3	-0.002	0.140	0.301
182	3,6-O-Su ₂ -Galβ1-4GlcNAcβ-sp2	0.018	0.218	0.950
183	4,6-O-Su ₂ -Galβ1-4GlcNAcβ-sp2	0.058	0.262	0.133
184	4,6-O-Su ₂ -Galβ1-4GlcNAcβ-sp3	-0.003	0.396	1.760
189	3,6-O-Su ₂ -Galβ1-4(6-O-Su)GlcNAcβ-sp2	-0.041	0.438	4.455
201	3,4-O-Su ₂ -Galβ1-4GlcNAcβ-sp3	-0.009	0.112	0.885
203	Galβ1-4(6-O-Su)GlcNAcβ-sp2	-0.015	0.360	3.957
220	Galα1-3Galβ1-4Glcβ-sp2	-0.002	0.216	0.242
222	Galα1-3Galβ1-4GlcNAcβ-sp3	0.014	0.132	0.357
224	Galα1-4Galβ1-4Glcβ-sp3	-0.062	0.077	0.130
225	Galα1-4Galβ1-4GlcNAcβ-sp2	0.020	0.218	0.936

228	Galβ1-2Galα1-4GlcNAcβ-sp4	-0,075	0,091	0,678
229	Galβ1-3Galβ1-4GlcNAcβ-sp4	0,002	0,210	0,230
231	Galβ1-4GlcNAcβ1-3GalNAcα-sp3	0,000	0,206	0,742
232	Galβ1-4GlcNAcβ1-6GalNAcα-sp3	0,008	0,168	0,640
254	Galβ1-3GlcNAcβ1-6GalNAcα-sp3	-0,007	0,285	0,832
362	Galβ1-3GalNAcβ1-3Gal-sp2	-0,045	0,169	0,102
264	Galβ1-4Galβ1-4GlcNAc-sp2	0,013	0,174	0,876
373	Galα1-4Galβ1-4GlcNAcβ1-5Galβ-sp3	0,012	0,184	0,395
375	Galα1-4GlcNAcβ1-3Galβ1-4GlcNAcβ-sp3	-0,009	0,649	0,260
376	Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ-sp4	0,098	0,231	0,381
377	Galβ1-3GlcNAcβ1-4Galβ1-3GlcNAcβ-sp2	0,001	0,184	0,069
378	Galβ1-3GlcNAcβ1-3Galβ1-4GlcNAcβ-sp3	-0,001	0,214	0,824
379	Galβ1-3GlcNAcβ1-3Galβ1-4GlcNAcβ-sp3	-0,014	0,251	0,434
380	Galβ1-3GlcNAcβ1-6Galβ1-4GlcNAcβ-sp2	0,028	0,258	0,683
381	Galβ1-3GlcNAcβ1-6Galβ1-4GlcNAcβ-sp2	0,140	0,258	0,613
382	Galβ1-3GalNAcβ1-4Galβ1-4Glcβ-sp3	0,387	0,108	0,438
383	Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ-sp2	-0,070	0,114	0,078
385	Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ-sp3	0,277	0,115	0,125
387	Galβ1-4GlcNAcβ1-6Galβ1-4GlcNAcβ-sp2	0,033	0,373	0,076
388	Galβ1-4GlcNAcβ1-6Galβ1-3GalNAcα-sp3	0,008	0,152	0,229
504	(A-GN-M)-3,6-M-GN-GMβ-sp4	-0,006	0,280	0,821
1A	Galβ1-3Glcβ-Ac	-0,056	0,107	0,042
1B	Galβ1-4Glcβ-Ac	-0,023	0,203	0,523
1C	Galβ1-4Gal	-0,022	0,199	0,458
1D	Galβ1-6Glcβ-Ac	-0,020	0,352	0,398
1E	Galβ1-3Galβ-Ac	-0,013	0,178	0,775
1F	Galβ1-3GalNAcβ1-4Galβ1-4Glc	0,081	0,395	-0,131

1G	Gal(1-3GlcNAc(1-4Gal(1-4Glc	-0.031	0.114	0.091
1H	Gal(1-4GlcNAc(1-3Gal(1-4Glc	0.006	0.148	0.802
1I	Gal(1-4GlcNAc(1-4Gal(1-4GlcNAc(1-3Gal(1-4Glc	0.070	0.220	0.282
1J	Gal(1-4GlcNAc(1-4Gal(1-3GlcNAc(1-3Gal(1-4Glc	0.004	0.273	0.646
1K	Gal(1-4Gal(1-4Glc	-0.026	0.121	0.299
1L	GalNAc(1-6Gal	-0.002	0.662	0.098
1M	Gal(1-3GalNAc(1-3Gal	0.030	0.192	-0.725
1N	Gal(1-4Gal	-0.026	0.308	0.618
1O	Gal(1-3Gal(1-4GlcNAc	-0.003	0.217	0.103
1P	Gal(1-3Gal(1-4Glc	-0.010	0.291	0.844
1A	Gal(1-3Gal(1-4Gal(1-3Gal	0.089	0.159	0.609
1B	Gal(1-4Gal	0.321	0.104	0.136
1C	GalNAc(1-3Gal	0.000	0.326	0.899
1D	GalNAc(1-4Gal	0.045	0.125	0.667
1E	Gal(1-4Gal(1-4GlcNAc	0.017	0.524	0.023
1F	GalNAc(1-3Gal(1-4Glc	-0.004	0.324	0.386
1G	Gal(1-3GlcNAc(1-3Gal(1-4GlcNAc(1-4Gal(1-3Gal(1-4Glc	-0.001	0.113	0.750
1H	Gal(1-3GlcNAc(1-3Gal(1-4GlcNAc(1-3Gal(1-4Glc	-0.075	0.100	0.559
1I	Gal(1-3GalNAc(1-3Gal(1-4Glc	-0.016	0.142	0.757
1J	Gal(1-4GalNAc(1-3Gal	0.091	0.208	0.659
1K	Gal(1-4Glc	-0.035	0.201	0.201
1L	Gal(1-4Gal	0.011	0.382	0.493
1M	Gal(1-4Gal	0.000	0.093	0.185
1N	Gal(1-4Gal	Sub WT	Sub WT	Sub WT
N-acetylglucosamine terminal				
1O1	GalNAc(1-3GalNAc(1-3Gal	0.015	0.148	0.015
1O2	GalNAc(1-3Gal(1-3Gal	-0.012	0.103	0.396
1O3	GalNAc(1-3GalNAc(1-3Gal	-0.007	0.177	0.087

194	Ga ₂ N ₂ β1-5Gaβ-q ³	0,000	0,139	0,050
196	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,031	0,062	0,427
192	Ga ₂ N ₂ β1-4Ga ₂ O ₂ N ₂ β-q ³	-0,196	0,303	0,049
193	3Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,253	0,175	0,161
194	3Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	0,000	0,105	0,077
195	6Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,009	0,292	-0,009
196	3Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,028	0,148	0,810
197	3Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,021	0,210	0,116
198	4Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,004	0,203	0,123
199	4Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,300	0,347	0,834
200	4Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	0,000	0,108	0,109
202	6Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	0,332	0,125	0,073
204	4Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	0,262	0,115	0,503
208	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,005	0,168	0,560
209	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,018	0,127	0,044
IL	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	0,022	0,175	0,052
20	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	-0,009	0,142	0,070
20	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	0,094	0,112	0,049
20	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	0,000	0,125	0,145
20	Ga ₂ N ₂ β1-4Ga ₂ N ₂ β-q ³	Substrat	Substrat	Substrat
Furnace				
71	Furnace-2Ga ₂ N ₂ β-q ³	-0,075	0,272	0,066
71	Furnace-2Ga ₂ N ₂ β-q ³	-0,262	0,077	0,114
71	Furnace-2Ga ₂ N ₂ β-q ³	-0,014	0,128	0,104
215	Furnace-2Ga ₂ N ₂ β-q ³	-0,018	0,228	0,095
216	Furnace-2Ga ₂ N ₂ β-q ³	0,039	0,166	0,154
217	Furnace-2Ga ₂ N ₂ β-q ³	0,007	0,159	0,206
219	Furnace-2Ga ₂ N ₂ β-q ³	0,000	0,136	0,145

246	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}000\text{-}ap3$	0,000	0,150	0,067
249	$\text{Ca}001 \rightarrow \text{Pw}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,000	0,140	0,437
254	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,325	0,109	0,677
259	$\text{Pw}001 \rightarrow \text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}000\text{-}ap3$	0,006	0,260	0,433
287	$3\text{Ca}000\text{-}ap3 \rightarrow \text{Pw}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,029	0,125	0,507
288	$\text{Pw}001 \rightarrow 3\text{Ca}000\text{-}ap3 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,067	0,268	0,546
359	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,738	0,343	0,102
360	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,000	0,147	0,126
362	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,403	0,298	0,249
363	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,019	0,140	0,269
364	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,354	0,106	0,995
368	$\text{Pw}001 \rightarrow \text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,022	0,160	0,418
368	$\text{Pw}001 \rightarrow \text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,003	0,106	0,283
374	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,002	0,307	0,195
372	$\text{Pw}001 \rightarrow \text{Pw}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,085	0,917	0,190
382	$\text{Pw}001 \rightarrow \text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,166	0,145	0,125
479	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,012	0,344	0,329
480	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,120	0,093	0,107
483	$\text{Pw}001 \rightarrow \text{Pw}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	0,000	0,116	0,070
486	$\text{Pw}001 \rightarrow \text{Pw}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,013	0,330	0,159
497	$\text{Pw}001 \rightarrow \text{Pw}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,012	0,177	0,086
518	$3\text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}000\text{-}ap3$	0,003	0,096	0,153
519	$3\text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}000\text{-}ap3$	0,000	0,135	0,177
543	$3\text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}000\text{-}ap3$	-0,101	0,222	0,061
542	$3\text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}000\text{-}ap3$	0,000	0,133	0,115
543	$3\text{Ca}000\text{-}ap3 \rightarrow 3\text{Ca}000\text{-}ap3$	0,415	0,128	0,309
74	$\text{Pw}001 \rightarrow \text{Ca}001 \rightarrow 3\text{Ca}001 \rightarrow 4\text{Ca}000\text{-}ap3$	-0,225	0,173	0,152

7B	Galβ1-3(Fucal1-4)GlcNAcβ1-3Galβ1-4Glc	0,017	0,151	0,006
7C	Galβ1-4(Fucal1-2)GlcNAcβ1-3Galβ1-4Glc	0,001	0,255	0,305
7D	Fucal1-2Galβ1-3(Fucal1-4)GlcNAcβ1-3Galβ1-4Glc	0,002	0,160	-0,029
7E	Galβ1-3(Fucal1-4)GlcNAcβ1-3Galβ1-4(Fucal1-3)Glc	0,000	0,120	0,114
7F	Fucal1-2Gal	-0,217	0,117	-0,589
7G	Fucal1-2Galβ1-4Glc	0,409	0,154	0,452
7H	Galβ1-4(Fucal1-3)Glc	-0,009	0,109	0,356
7I	Galβ1-4(Fucal1-3)GlcNAc	-0,002	0,227	0,370
7J	Galβ1-3(Fucal1-4)GlcNAc	-0,010	0,199	0,991
7K	GalNAcα1-3(Fucal1-2)Gal	0,016	0,360	0,343
7L	Fucal1-2Galβ1-4(Fucal1-3)Glc	-0,012	0,165	0,502
7M	Galβ1-3(Fucal1-2)Gal	0,030	0,832	0,106
7N	Fucal1-2Galβ1-4(Fucal1-3)GlcNAc	0,012	0,170	0,151
7O	Fucal1-2Galβ1-3GlcNAc	-0,003	0,342	0,372
7P	Fucal1-2Galβ1-3(Fucal1-4)GlcNAc	0,008	0,223	0,347
8A	SO ₃ -Galβ1-4(Fucal1-4)GlcNAc	0,011	0,169	0,636
8B	SO ₃ -Galβ1-4(Fucal1-3)GlcNAc	0,024	0,192	0,660
8C	Galβ1-3GlcNAcβ1-2Galβ1-4(Fucal1-3)GlcNAcβ1-3Galβ1-4Glc	0,004	0,104	0,337
8D	Galβ1-4(Fucal1-3)GlcNAcβ1-4Galβ1-3GlcNAcβ1-3Galβ1-4Glc	0,005	0,266	0,990
8E	Galβ1-4(Fucal1-3)GlcNAcβ1-4(Fucal1-2)Galβ1-3GlcNAcβ1-3Galβ1-4Glc	0,004	0,309	0,522
8F	Galβ1-4(Fucal1-3)GlcNAcβ1-4(Fucal1-2)Galβ1-3(Fucal1-4)GlcNAcβ1-3Galβ1-4Glc	0,012	0,445	0,265
8G	Galβ1-4GlcNAcβ1-3Galβ1-4(Fucal1-3)Glc	0,016	0,183	-0,019
8H	Fucal1-2Galβ1-4(Fucal1-3)GlcNAcβ1-3Galβ1-4Glc	0,013	0,690	0,314
8I	Fucal1-3Galβ1-4GlcNAcβ1-3Galβ1-4(Fucal1-3)Glc	0,008	0,243	0,277
8J	Fucal1-2Galβ1-4(Fucal1-3)GlcNAcβ1-3(Fucal1-2)Galβ1-4Glc	0,011	0,133	0,674
8K	Galβ1-4(Fucal1-3)GlcNAcβ1-4Galβ1-4GlcNAcβ1-3Galβ1-4Glc	0,023	0,243	0,429
8L	Galβ1-4(Fucal1-2)GlcNAcβ1-4Galβ1-4(Fucal1-3)GlcNAcβ1-3Galβ1-4Glc	0,020	0,216	0,118

884	Pheal-2-Galp1-4-Pheal-3-CleNAcb1-6-Galp1-4-CleNAcb1-3-Xalp1-4-Glc	0,001	0,934	0,738
885	Gal11-3-CleNAcb1-4-Galp1-4-Pheal-3-CleNAcb1-6-Galp1-4-CleNAcb1-3-Galp1-4-Glc	0,022	0,858	0,750
886	Pheal-2-Galp1-3-CleNAcb1-4-Pheal-3-CleNAcb1-6-Galp1-4-CleNAcb1-3-Galp1-4-Glc	0,000	0,176	0,256
887	GalNAcb1-3-Pheal-2-Galp1-4-Glc	0,006	0,200	0,389
888	Gal11-3-Pheal-2-Galp1-4-Pheal-3-Glc	0,007	0,563	0,283
889	Gal11-4-CleNAcb1-6-Pheal-3-Galp1-3-CleNAcb1-3-Galp1-4-Glc	-0,008	0,161	0,186
890	Gal11-3-Pheal-2-Galp1-4-Glc	-0,004	0,259	0,459
891	GalNAcb1-3-Pheal-2-Galp1-4-Pheal-3-Glc	0,002	0,116	0,149
892	Gal11-4-Pheal-3-GleNAcb1-3-Gal	0,004	0,183	0,181
893	Pheal-2-Galp1-4-Pheal-3-CleNAcb1-3-Gal	0,018	0,342	0,370
894	Gal11-3-Pheal-4-GleNAcb1-3-Gal	0,018	0,193	0,453
895	Pheal-3-Galp1-3-Pheal-4-CleNAcb1-3-Gal	0,004	0,729	0,182
Sialilado				
896	Neu5Ac2-6Gal1-4-Gal	-0,023	0,304	0,891
897	Neu5Ac2-6Gal1-4-Gal	0,018	0,132	0,861
898	Neu5Ac2-6GalNAc-4-Gal	-0,030	0,174	0,987
899	Neu5Ac2-6GalNAc-4-Gal	-0,003	0,649	0,992
900	Neu5Ac2-6GalNAc-4-Gal	0,004	0,269	0,940
901	Neu5Ac2-6GalNAc-4-Gal	0,031	0,632	0,911
902	Neu5Ac2-6GalNAc-4-Gal	0,022	0,246	0,910
903	Neu5Ac2-6GalNAc-4-Gal	3,359	1,190	3,798
904	Gal11-3-Neu5Ac2-6GalNAc-4-Gal	0,009	0,454	0,809
905	Gal11-4-Neu5Ac2-6GalNAc-4-Gal	-0,008	0,150	0,559
906	Neu5Ac2-3-Galp1-3-GalNAc-4-Gal	0,002	0,231	1,008
907	Neu5Ac2-4-Galp1-4-Gal	0,005	0,151	0,631
908	Neu5Ac2-3-Galp1-4-Gal	-0,001	0,330	0,408
909	Neu5Ac2-6Gal1-4-Gal	-0,015	0,286	0,033

296	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	1,248	0,255	0,278
299	Na ₂ S ₂ O ₃ ·10H ₂ O-5Cl ₂ N ₂ O ₂ -ap ³	0,005	0,266	0,816
300	Na ₂ S ₂ O ₃ ·6H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,000	0,308	0,478
301	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	2,572	1,838	3,094
304	Na ₂ S ₂ O ₃ ·6H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,000	1,135	2,798
306	Na ₂ S ₂ O ₃ ·Na ₂ S ₂ O ₃ ·2H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	-0,005	0,179	0,261
311	Na ₂ S ₂ O ₃ ·10H ₂ O-4H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	-0,003	0,459	0,980
317	Na ₂ S ₂ O ₃ ·10H ₂ O-4H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,007	0,201	0,019
318	Na ₂ S ₂ O ₃ ·6H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,037	0,182	0,125
319	Na ₂ S ₂ O ₃ ·2H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,000	0,559	0,105
321	(Na ₂ S ₂ O ₃ ·8H ₂ O)-ap ³	0,017	0,198	0,399
323	Na ₂ S ₂ O ₃ ·2H ₂ O-4Cl ₂ N ₂ O ₂ -5Cl ₂ N ₂ O ₂ -ap ³	0,126	0,166	0,116
324	Na ₂ S ₂ O ₃ ·6H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	-0,295	0,160	0,071
331	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,316	1,091	1,703
421	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -4Cl ₂ N ₂ O ₂ -ap ²	-0,039	0,124	0,312
422	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5Cl ₂ N ₂ O ₂ -ap ³	-0,005	0,194	0,068
423	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5Cl ₂ N ₂ O ₂ -ap ³	0,843	0,150	0,603
426	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5Cl ₂ N ₂ O ₂ -ap ³	0,141	0,156	0,116
428	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,118	0,193	0,067
429	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,066	0,063	0,091
433	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,753	0,143	0,132
434	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,029	0,144	0,034
517	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,401	0,068	0,377
528	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,769	0,103	2,216
529	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,098	0,082	0,111
531	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	-0,021	0,114	0,072
533	Na ₂ S ₂ O ₃ ·10H ₂ O-4Cl ₂ N ₂ O ₂ -5H ₂ O-5H ₂ O-4Cl ₂ N ₂ O ₂ -ap ³	0,053	0,345	0,181

533	(Neu5Ac)2-3-(Neu5Ac)2-3-GalNAc β 1-4-Gal β 1-4-Glc \rightarrow q2	-0,001	0,163	0,381
534	Neu5Ac)2-3-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-GlcNAc \rightarrow q3	0,633	0,113	0,115
536	Neu5Ac)2-3-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-Glc \rightarrow q4	0,028	0,126	0,260
537	Neu5Ac)2-3-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-Glc \rightarrow q4	0,547	0,102	0,619
540	Lac1-4(Fuc)NAc1-3-Lac \rightarrow q4	0,062	0,350	0,580
10A	Neu5Ac)2-3-Gal β 1-4(Fuc)1-4-GlcNAc	-0,007	0,088	0,864
10B	Neu5Ac)2-3-Gal β 1-4(Fuc)1-4-GlcNAc	-0,002	0,294	0,791
10C	Neu5Ac)2-3-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-Glc	0,307	0,193	0,974
10D	Gal β 1-4(Fuc)1-3-GlcNAc β 1-4(GlcNAc)2-6-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-Glc	0,024	0,155	0,628
10E	Neu5Ac)2-3-Gal β 1-4(Fuc)NAc2-6-GlcNAc	0,118	0,122	0,957
10H	Neu5Ac)2-3-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4(Fuc)1-3-Glc	0,015	0,109	0,421
10I	Gal β 1-4-GlcNAc β 1-4(Fuc)NAc2-6-Gal β 1-4-GlcNAc β 1-4-Glc	0,422	0,096	0,632
10J	Neu5Ac)2-6-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-GlcNAc β 1-4-Glc	0,246	0,097	0,915
10K	Neu5Ac)2-3-Gal β 1-4-GlcNAc	-0,021	0,129	0,758
10L	Neu5Ac)2-4-Gal β 1-4-GlcNAc	0,384	0,186	0,753
10M	Neu5Ac)2-3-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-Glc	0,047	0,411	0,795
10N	Gal β 1-3(Neu5Ac)2-6-GlcNAc β 1-3-Gal β 1-4-Glc	-0,037	0,463	0,925
10O	Neu5Ac)2-4-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-Glc	0,000	0,135	0,384
10P	Neu5Ac)2-4-Gal β 1-4(Fuc)NAc2-6-GlcNAc β 1-3-Gal β 1-4-Glc	0,079	0,083	0,527
11A	Neu5Ac)2-3-Gal β 1-4-Glc	0,001	0,099	0,835
11B	Neu5Ac)2-4-Gal β 1-4-Glc	0,188	0,345	0,687
11C	(Neu5Ac)2-4(Neu5Ac)1-6(Glc)	0,007	0,132	0,599
11A	Neu5Ac)2-3-Gal β 1-4-GlcNAc β 1-3-Gal β 1-4-Glc	0,213	0,351	0,669
11K	9-NAc-Neu5Ac	-0,001	0,183	0,968
11O	Neu5Ac	0,035	0,283	0,650
11K	Neu5Ac)2-3-Gal β 1-4(Fuc)1-4-GlcNAc β 1-3-Gal	0,009	0,124	0,851
Manova		Subb WI	Subb25106/21107	Subb5106/47107A

		SubB WT	SubBΔ3106/ΔT107	SubBΔ106/ΔT107Δ
119	Mannitol-2Manβ-sp4	0.015	0.105	0.086
120	Mannitol-3Manβ-sp4	0.063	0.076	0.377
121	Mannitol-4Manβ-sp4	-0.006	0.110	0.112
122	Mannitol-6Manβ-sp4	0.426	0.133	0.065
123	Mannitol-4GlcNAcβ-sp4	0.035	0.233	0.213
124	Mannitol-2Manα-sp4	0.033	0.119	0.058
126	Mannitol-4S(Mannitol-6)Manβ-sp4	0.039	0.663	0.291
129	Mannitol-3S(Mannitol-3)Manβ-sp4	0.360	0.077	0.137
131	GlcNAcβ1-2Man	0.580	0.106	0.338
133	GlcNAcβ1-2Manβ1-6(GlcNAcβ1-2)Manβ1-3)Man	0.411	0.176	0.070
135	Mannitol-2Man	-0.002	0.133	0.959
137	Mannitol-3Man	-0.008	0.389	0.154
139	Mannitol-4Man	0.021	0.384	0.407
141	Mannitol-6Man	0.007	0.220	0.306
143	Mannitol-6(Mannitol-3)Man	0.034	0.059	0.121
145	Mannitol-4S(Mannitol-3)Manβ1-6S(Mannitol-3)Man	-0.001	0.207	0.019
N-acetylglucosamine terminal				
113	GlcNAcβ1-3GalNAcα-sp3	0.000	0.146	0.060
114	GlcNAcβ1-3Manβ-sp4	0.126	0.267	0.018
115	GlcNAcβ1-4GlcNAcα-sp4	-0.008	0.332	-0.007
117	GlcNAcβ1-4GlcNAcβ-sp4	0.004	0.159	0.015
118	GlcNAcβ1-6GalNAcα-sp3	-0.005	0.132	0.146
149	GlcNAcβ1-4GlcNAcβ-sp2	0.006	0.138	0.483
127	GlcNAcβ1-4-4HCOOCH ₃ CH ₂ CH ₂ -3-O-GlcNAcβ-sp4	-0.009	0.138	0.048
116	GlcNAcβ1-βHCOO(CH ₂ CH ₂ -3-O-GlcNAcβ1-αD-glucosaminyl-2-Asnase	-0.001	0.118	0.096
126	GlcNAcβ1-2Galβ1-3GalNAcα-sp3	-0.006	0.080	0.414
127	GlcNAcβ1-2Galβ1-3GalNAcα-sp3	-0.015	0.126	0.341

248	GlcNAc β 1-3Gal β 1-4Glc β -sp2	-0,007	0,097	0,084
249	GlcNAc β 1-3Gal β 1-4GlcNAc β -sp3	-0,030	0,133	0,278
251	GlcNAc β 1-4Glc β 1-4GlcNAc β -sp2	-0,002	0,071	0,210
252	GlcNAc β 1-4GlcNAc β 1-4GlcNAc β -sp4	-0,001	0,103	0,351
253	GlcNAc β 1-4Gal β 1-4GlcNAc β -sp2	-0,005	0,103	0,102
255	GlcNAc β 1-3(GlcNAc β 1-5GalNAc α -sp3)	-0,009	0,123	0,003
395	GlcNAc β 1-3(GlcNAc β 1-5Gal β 1-4GlcNAc β -sp3)	-0,012	0,225	0,746
402	(GlcNAc β 1-4) β -sp4	-0,015	0,067	0,076
503	(GlcNAc β 1-4) β -sp4	-0,003	0,124	0,175
505	(GlcNAc β 1-3,6-M-GlcNAc β -sp4)	-0,007	0,129	0,091
5A	GlcNAc β 1-4GlcNAc	0,007	0,189	0,150
5B	GlcNAc β 1-4GlcNAc β 1-4GlcNAc	0,001	0,230	0,511
5C	GlcNAc β 1-4GlcNAc β 1-4(GlcNAc β 1-4GlcNAc)	0,003	0,096	0,062
5D	GlcNAc β 1-4GlcNAc β 1-4GlcNAc β 1-4GlcNAc β 1-4GlcNAc β 1-4GlcNAc	-0,001	0,113	0,227
5E	Disacárido ramificado de la pared celular	-0,009	0,089	0,045
5F	GlcNAc β 1-4GlcNAc β 1-4GlcNAc β 1-4GlcNAc β 1-4GlcNAc	-0,010	0,071	0,051
180	6-O-Sp-GlcNAc	0,002	0,190	0,046
183	GlcNAc	0,014	0,184	0,591
Glucosa		SubB WT	SubBMS106/37107	SubBMS106/37107A
110	Glc α 1-4Glc β -sp3	-0,016	0,065	0,121
111	Glc β 1-4Glc β -sp4	-0,005	0,114	0,089
112	Glc β 1-6Glc β -sp4	-0,002	0,080	0,572
164	Glc β 1-3GlcNAc β -sp3	-0,010	0,136	0,564
165	Glc β 1-3Gal β -sp3	0,005	0,862	0,075
166	Glc β 1-4Gal β -sp3	0,007	0,156	0,181
240	(Glc α 1-4) β -sp4	-0,014	0,060	0,110
241	(Glc α 1-6) β -sp4	-0,007	0,144	0,507

396	(Glcot-4) β -sp4	-0,001	0,104	0,069
397	(Glcot-6) β -sp4	-0,016	0,130	0,094
492	(Glcot-6) β -sp4	0,029	0,102	0,451
502	(Glcot-6) β -sp4	-0,017	0,135	0,229
181	GlcA	-0,005	0,193	-0,072
187	6-O-(H ₂ PO ₃)Glc	0,000	0,244	0,026
190	Glcot-4Glcot-4	0,013	0,121	0,184
198	Glcot-4Glcot-4Glcot-4	-0,002	0,220	0,361
Subst W1 Subst W6 Subst W7 Subst W8 Subst W9				
Caragenano y glicaninoglicanos de bajo peso molecular				
12A	Neosaminetraosa-41, 3-di-O-sulfato (Na ⁺)	-0,001	0,087	0,156
12B	Neosaminetraosa-41-O-sulfato (Na ⁺)	0,063	0,217	0,216
12C	Neosaminetraosa-24, 41, 3, 5-tetra-O-sulfato (Na ⁺)	-0,007	0,098	0,498
12D	Neosaminetraosa-41, 3, 5-tetra-O-sulfato (Na ⁺)	-0,017	0,097	0,027
12E	Neosaminetraosa-41, 3, 5, 7-tetra-O-sulfato (Na ⁺)	-0,016	0,290	0,266
12F	Neosaminetraosa-41, 3, 5, 7, 9-penta-O-sulfato (Na ⁺)	0,001	0,283	0,532
12G	MUA-2S-GlcNAc-6S	-0,009	0,120	0,377
12H	MUA-GlcNAc-6S	-0,002	0,219	0,239
12I	MUA-2S-GlcNAc	0,072	0,101	0,014
12J	MUA-2S-GlcNAc-6S	-0,007	0,229	0,189
12K	MUA-GlcNAc-5S	0,017	0,146	0,031
12L	SUA-2S-GlcNAc	-0,009	0,272	0,154
12M	MUA-GlcNAc	-0,004	0,075	0,471
12N	MUA-GalNAc-4S (Delta D-4S)	-0,001	0,125	0,166
12O	MUA-GalNAc-5S (Delta D-5S)	-0,007	0,156	0,148
12P	MUA-GalNAc-4S-5S (Delta D-4S-5S)	0,004	0,192	0,109
13A	SUA-2S-GalNAc-4S (Delta D-4S-2S)	-0,002	0,599	0,064
13B	MUA-2S-GalNAc-4S (Delta D-4S-2S)	-0,001	0,167	0,226

13C	ΔUA-2S-GabAc-4S-6S (Delta D-ribs)	0,001	0,132	0,087
13D	ΔUA-2S-GabAc-6S (Delta D)-UAS	-0,007	0,115	0,087
13E	ΔUA-GlcNAc (Delta D-BA)	0,000	0,117	0,102
13M	ΔUA-→2S-GlcNAcS	-0,147	0,092	0,125
13N	ΔUA-→GlcNAcS	0,002	0,209	0,142
13O	ΔUA-→2S-GlcN	-0,051	0,098	0,198
13P	ΔUA-→GlcN	0,027	0,100	0,232
Caragenano y glucosaminoglicanos de alto peso molecular (GAGs)				
62S	(GlcAβ1-4GlcNAcβ1-3) _n -NH ₂ -ol	-0,012	0,143	0,418
11E	(GlcAβ1-3GlcNAcβ1-4) _n (n=4)	-0,011	0,143	0,048
11G	(GlcAβ1-3GlcNAcβ1-4) _n (n=6)	-0,009	0,151	0,178
11H	(GlcAβ1-3GlcNAcβ1-4) _n (n=16)	-0,003	0,160	0,494
11I	(GlcAβ1-3GlcNAcβ1-4) _n (n=12)	0,004	0,162	0,802
11J	(GlcAβ1-3GlcNAcβ1-4GlcNAcβ1-4) _n (n=200)	0,027	0,179	1,058
11K	(GlcAβ1-3GlcNAcβ1-4GlcNAcβ1-4) _n (n=300)	0,040	0,260	0,068
11L	(GlcAβ1-3GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4) _n (n=300)	0,159	0,135	0,070
11M	(GlcAβ1-3GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4) _n (n=300)	0,467	0,202	1,857
13Q	HA-4 (n=30)	-0,016	0,219	0,299
13R	HA-6 (n=30)	-0,009	0,455	0,446
13P	HA-8 (n=30)	-0,004	0,598	0,446
14A	HA-10 (n=30)	0,005	0,094	-0,004
14B	HA-12 (n=30)	0,290	0,101	0,005
14C	HA-14 (n=30)	-0,012	0,132	0,030
14D	HA-16 (n=30)	-0,002	0,133	0,048
14E	HA-2000 da 2,5mg/ml	0,005	0,149	0,058
14F	HA-10000 da 2,5mg/ml	0,007	0,264	0,060
14G	HA-10000 da 2,5 mg/ml	0,019	0,075	0,084

Sub WT SubB5196CWT07 SubB5196A/T107A

148	3A 22000 da 3,5 mg/ml	0,586	0,082	0,104
149	3A 10000 da 3,5 mg/ml	0,0005	0,087	0,105
149	Sulfato de heparina 5 mg/ml	0,01278	0,093	0,106
149	Glucano B1-3	0,00477	0,094	0,110
N-glicanos complexos				
627	(Sial2,6A,6B,AD), 3,6-A,4,6N,GNP-q4	Substancia 7107	Substancia 7107A	
19A	Galβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4Fucα1-6GlcNAc	0,02051	0,095	0,112
19B	Galβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	0,01818	0,098	0,113
19C	Manα6-2-4GlcNAcβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	0,00915	0,099	0,116
19D	Manα6-2-4GlcNAcβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	0,00948	0,101	2,135
19E	Galβ1-4GlcNAcβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	0,00765	0,103	1,151
19F	Manα6-2-4GlcNAcβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	0,00316	0,103	0,156
19F	Manα6-2-4GlcNAcβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	0,00284	0,103	0,172
19G	Galβ1-4GlcNAcβ1-4GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	0,00056	0,104	0,175
19H	GlcNAcβ1-2(GlcNAcβ1-2Manα1-6GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ1-4GlcNAc	-1E-05	0,105	0,175

Los valores de plegue superiores a 1 indican una unión significativamente

REIVINDICACIONES

1. Una proteína aislada que comprende una secuencia de aminoácidos de SEQ ID NO:2 o un fragmento de la misma que comprende la secuencia de aminoácidos TTSTE (SEQ ID NO:3), que comprende una delección de uno o más de los residuos de aminoácidos de la secuencia de aminoácidos TTSTE (SEQ ID NO:3), y en donde la proteína aislada o el fragmento de la misma es capaz de unirse a ácido *N*-glicolilneuramínico ligado a α 2-3y ácido *N*-glicolilneuramínico ligado a α 2-6.
2. La proteína aislada o fragmento de la reivindicación 1, que comprende una delección de al menos uno de los residuos subrayados de TTSTE (SEQ ID NO:3).
3. La proteína aislada o fragmento de la reivindicación 2, que comprende además una delección del residuo subrayado de TTSTE (SEQ ID NO:3).
4. La proteína aislada o fragmento de la reivindicación 2, que comprende una delección de los residuos subrayados de TTSTE (SEQ ID NO: 3).
5. La proteína aislada o fragmento de la reivindicación 1, que comprende la secuencia de aminoácidos establecida en SEQ ID NO: 1 o una secuencia de aminoácidos que sea idéntica en un 98 % en toda la longitud completa de SEQ ID NO: 1.
6. Un complejo molecular aislado que comprende la proteína aislada o el fragmento de cualquiera de las reivindicaciones 1-5 y un glicano que comprende ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6.
7. El complejo molecular aislado de la reivindicación 6, en donde el glicano que comprende el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6 es expresado por una célula tumoral o por células sanguíneas felinas.
8. Una composición que comprende la proteína aislada o el fragmento de cualquiera de las reivindicaciones 1-5.
9. Un método para detectar el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6, en donde dicho método incluye el paso de combinar la proteína aislada o el fragmento de cualquiera de las reivindicaciones 1-5 o la composición de la reivindicación 8 con una muestra para formar así un complejo detectable que comprenda la proteína aislada y el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6.
10. Un método para aislar un glicano o una célula que expresa el glicano, en donde el glicano comprende el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o un ácido *N*-glicolilneuramínico ligado a α 2-6, en donde dicho método incluye los pasos de: combinar la proteína o fragmento aislado de las reivindicaciones 1-5 o la composición de la reivindicación 8 con una muestra para formar así un complejo que comprenda la proteína o fragmento aislado y el glicano que comprende el ácido *N*-glicolilneuramínico ligado a α 2-3 y/o el ácido *N*-glicolilneuramínico ligado a α 2-6; y aislar la proteína o célula.
11. Una proteína aislada o fragmento de cualquiera de las reivindicaciones 1-5, o la composición de la reivindicación 8, para su uso en el tratamiento del cáncer en un sujeto para así dirigirse selectivamente a una célula cancerosa que exprese un ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6.
12. La proteína aislada o el fragmento de cualquiera de las reivindicaciones 1-5, o la composición de la reivindicación 8 para su uso en la detección de células tumorales que expresen ácido *N*-glicolilneuramínico ligado a α 2-3 y/o ácido *N*-glicolilneuramínico ligado a α 2-6.
13. Un ácido nucleico aislado que codifica la proteína aislada o el fragmento de cualquiera de las reivindicaciones 1-5.
14. Un constructo genético que comprende el ácido nucleico aislado de la reivindicación 13, opcionalmente en donde el constructo genético se encuentra en una célula hospedera.
15. Un kit que comprende la proteína aislada o el fragmento de cualquiera de las reivindicaciones 1-5, la composición de la reivindicación 8, el ácido nucleico aislado de la reivindicación 13 y/o el constructo genético de la reivindicación 14.

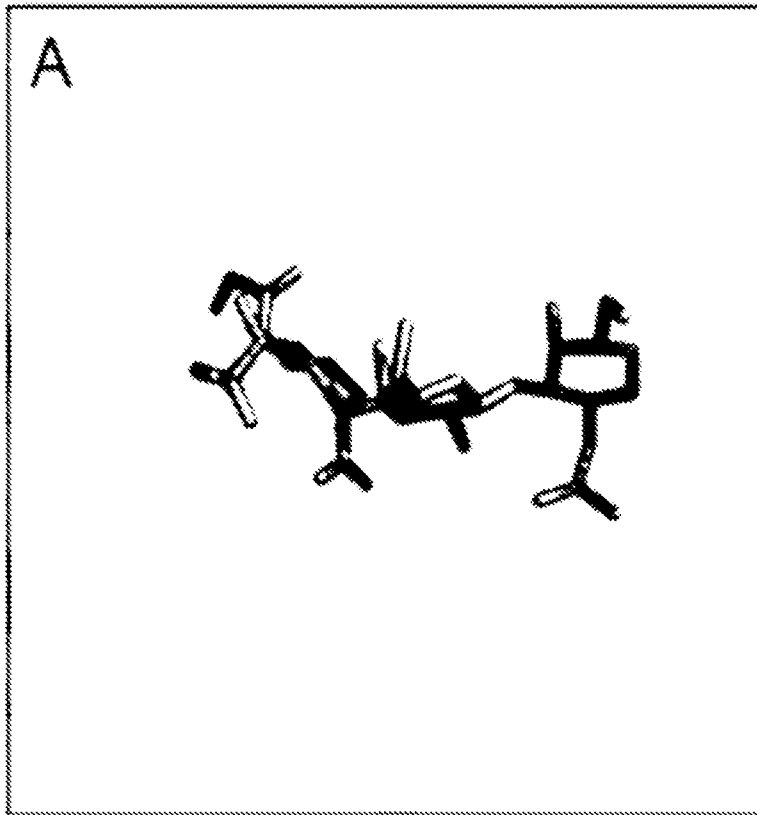


Figura 1A

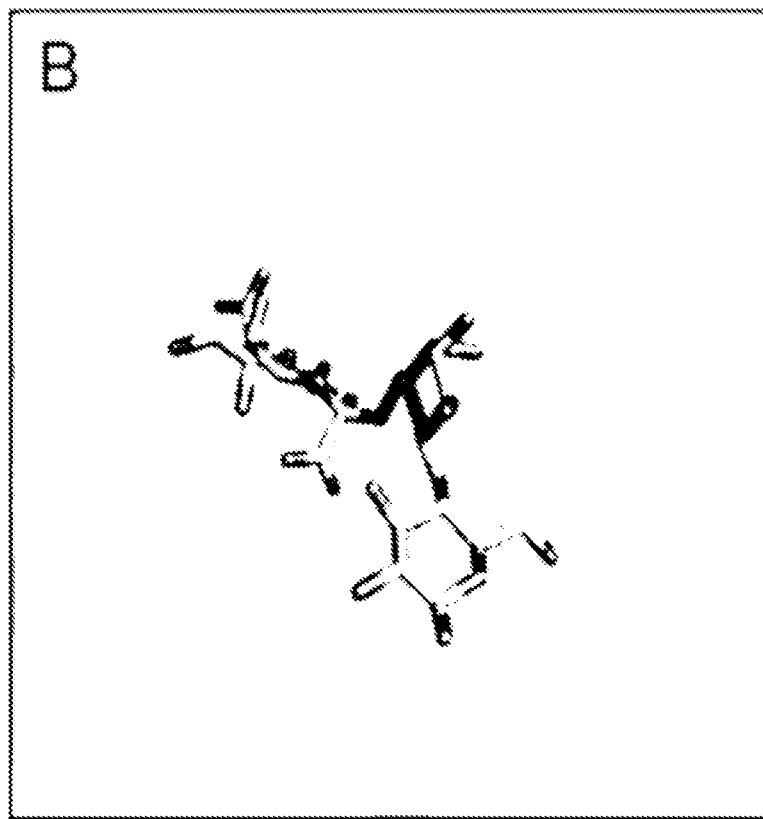
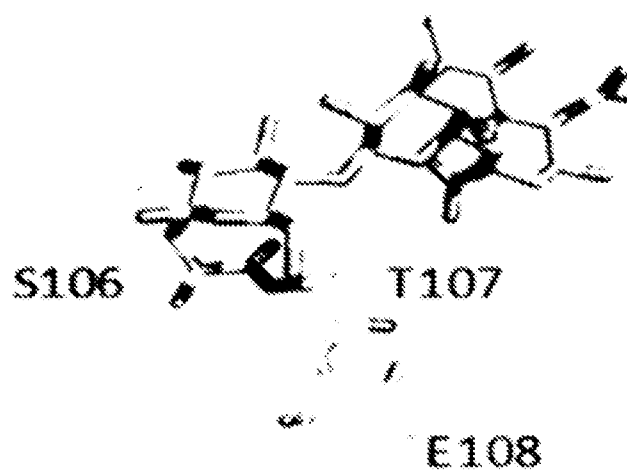


Figura 1B

Tipo silvestre



$\Delta S106/\Delta T107$

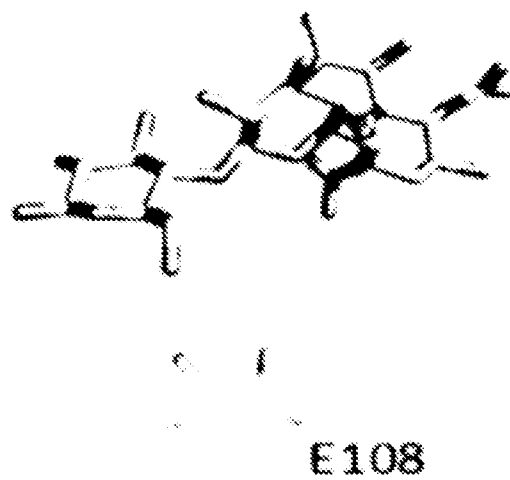


Figura 2

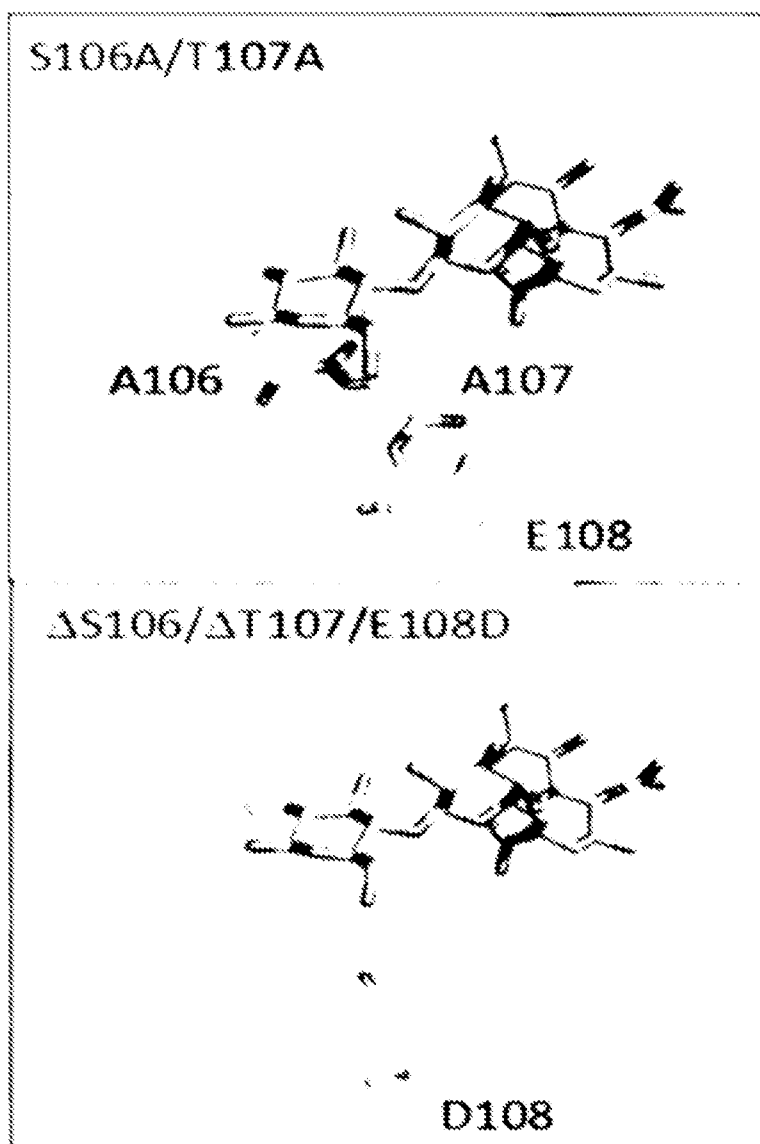


Figura 2 (continuación)

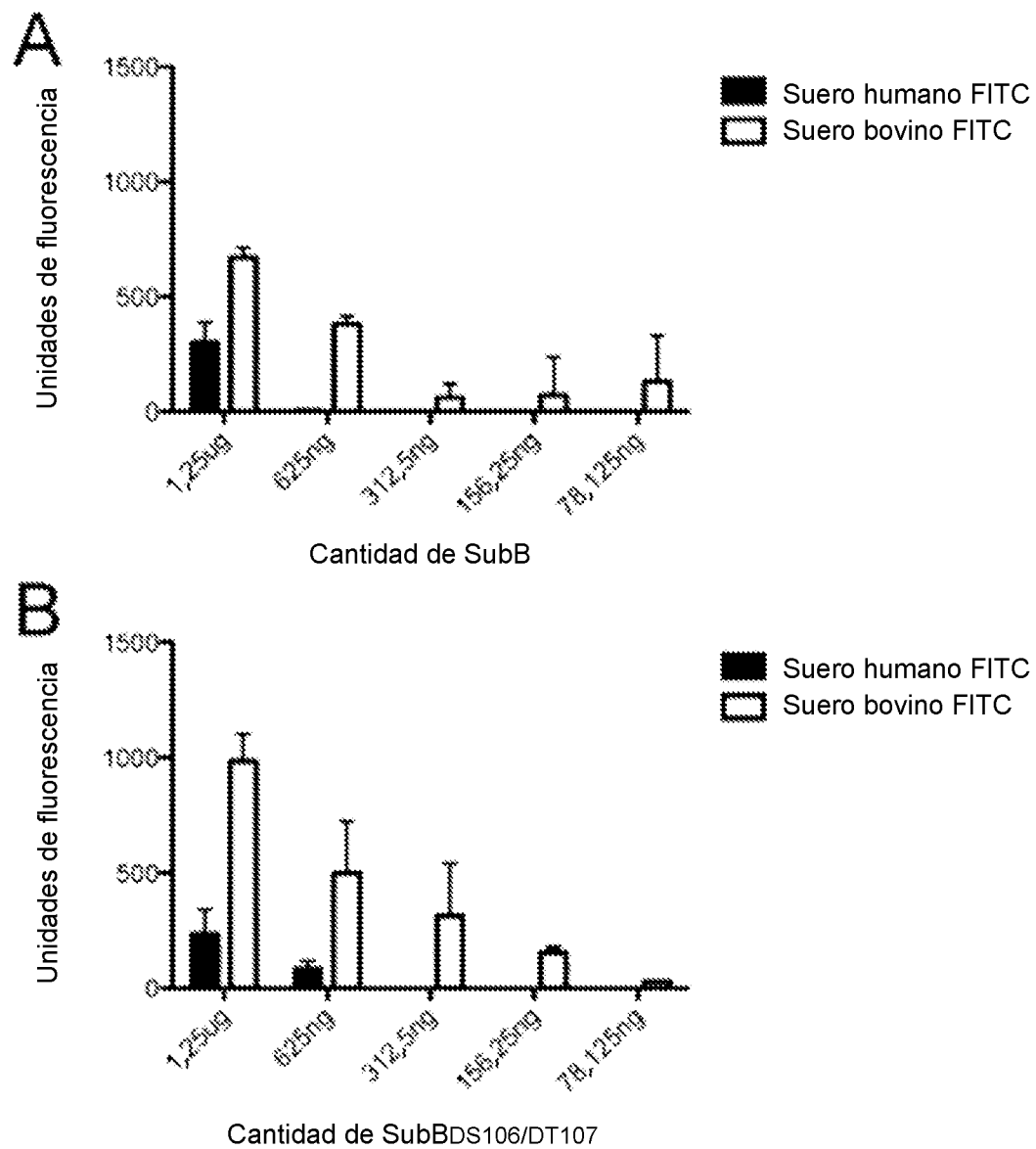


Figura 3

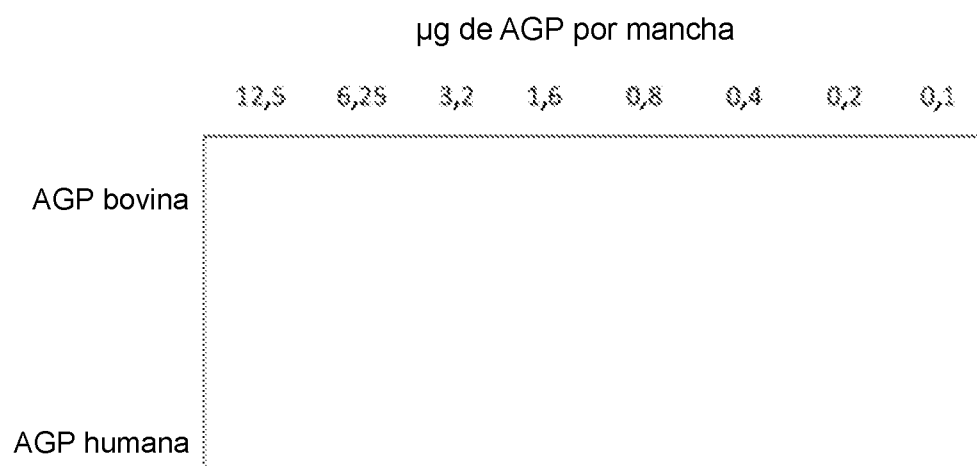


Figura 4

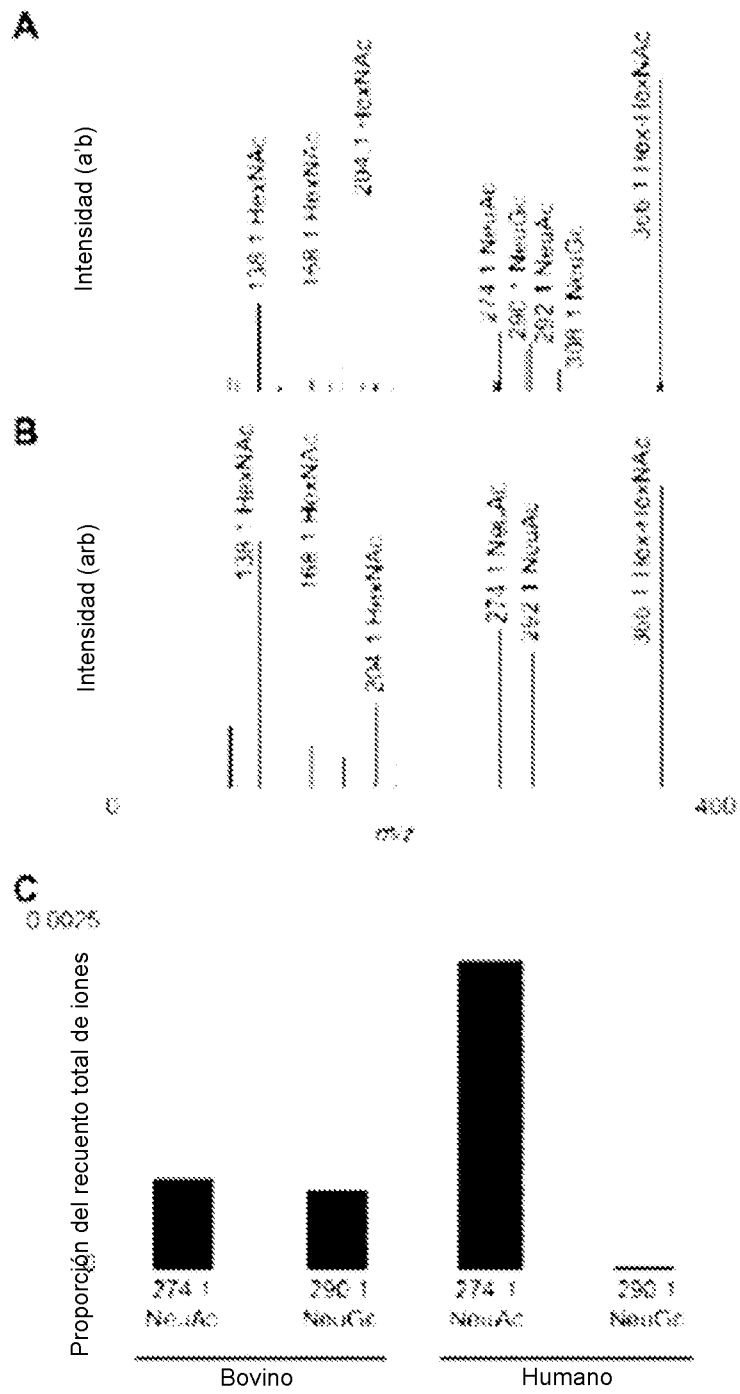


Figura 5A-C

D

glicoproteína $\alpha 1$ ácida
Bovina Humana

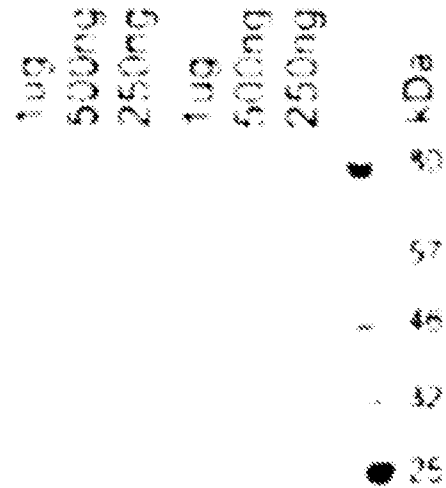


Figura 5D

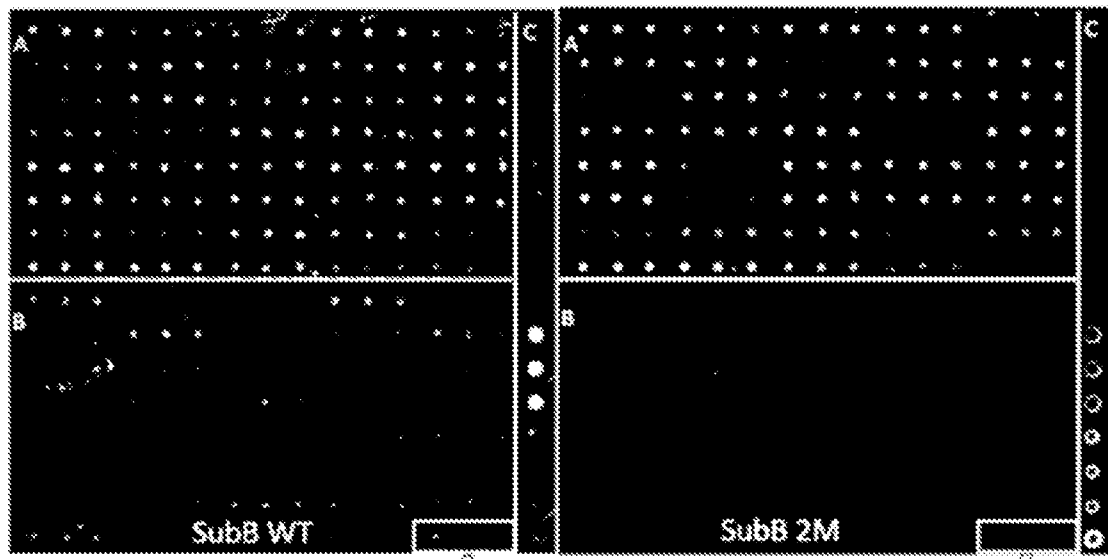


Figura 6

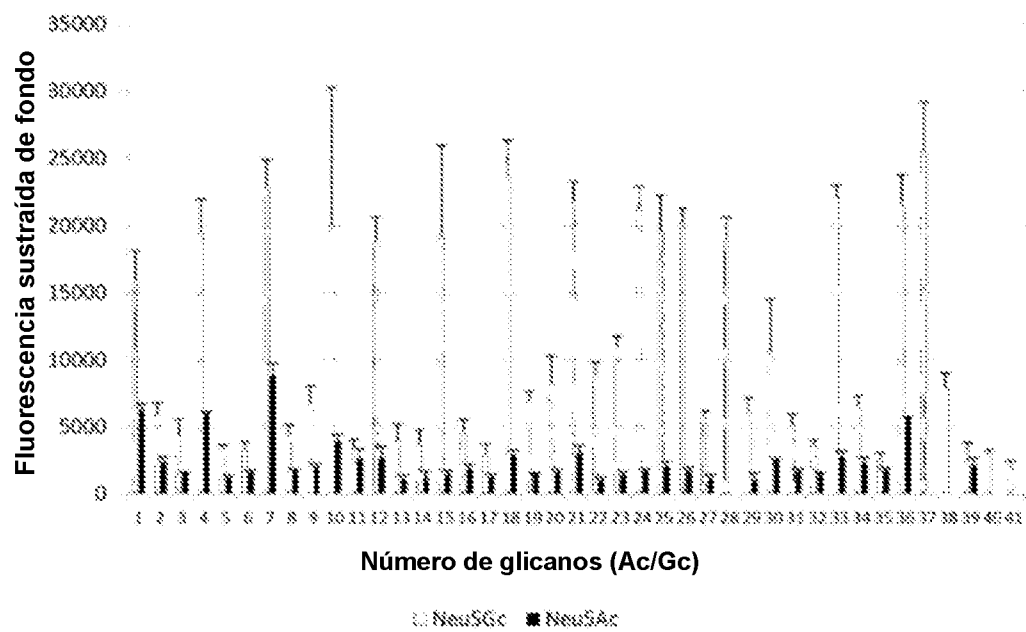


Figura 7

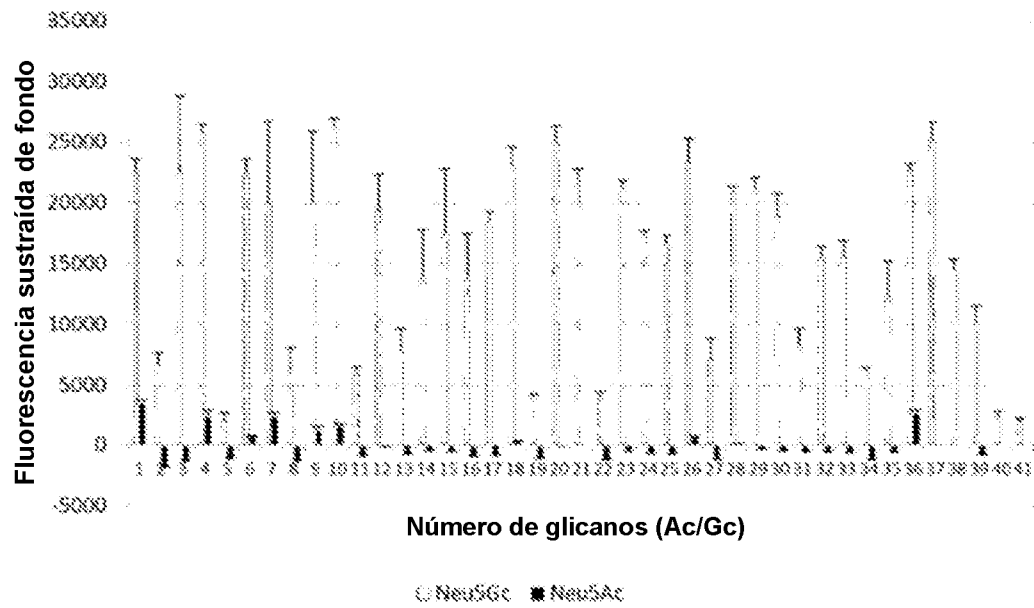


Figura 8

Glicanos Neu 5 Ac

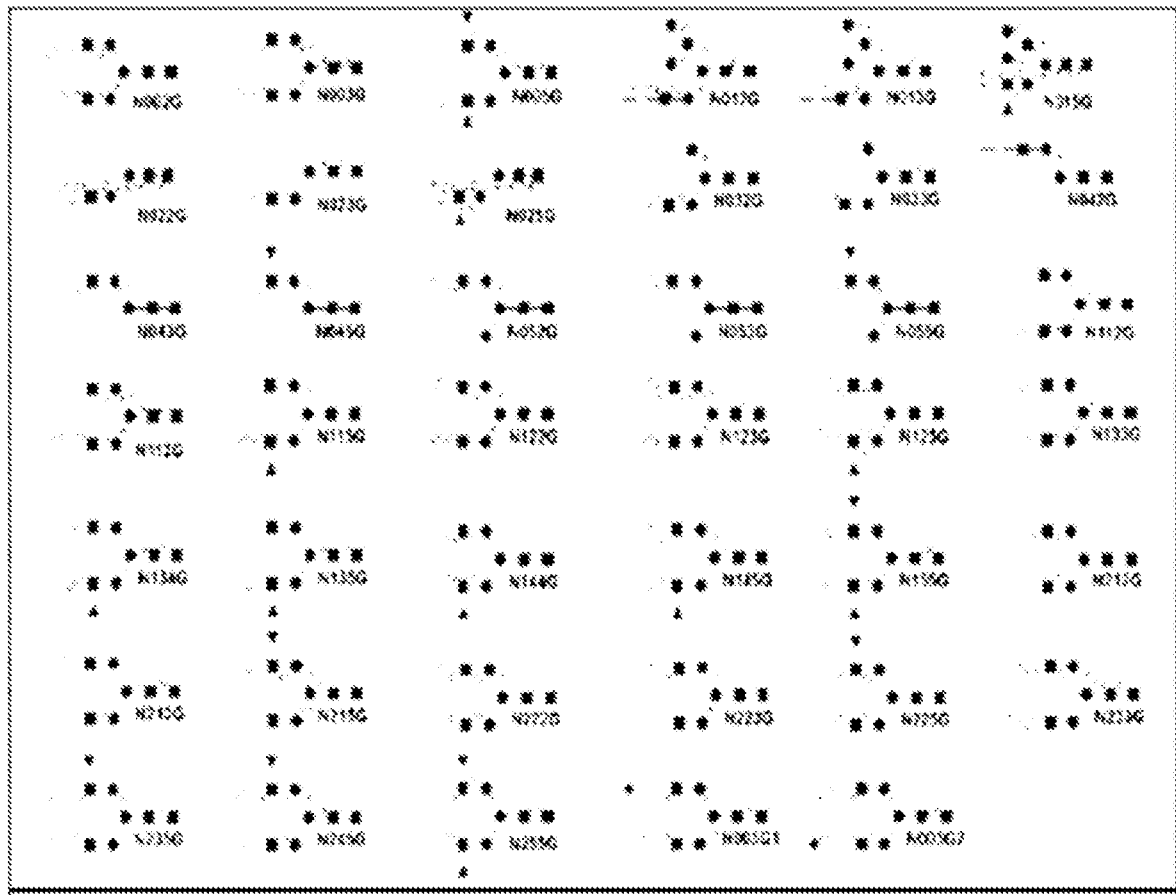
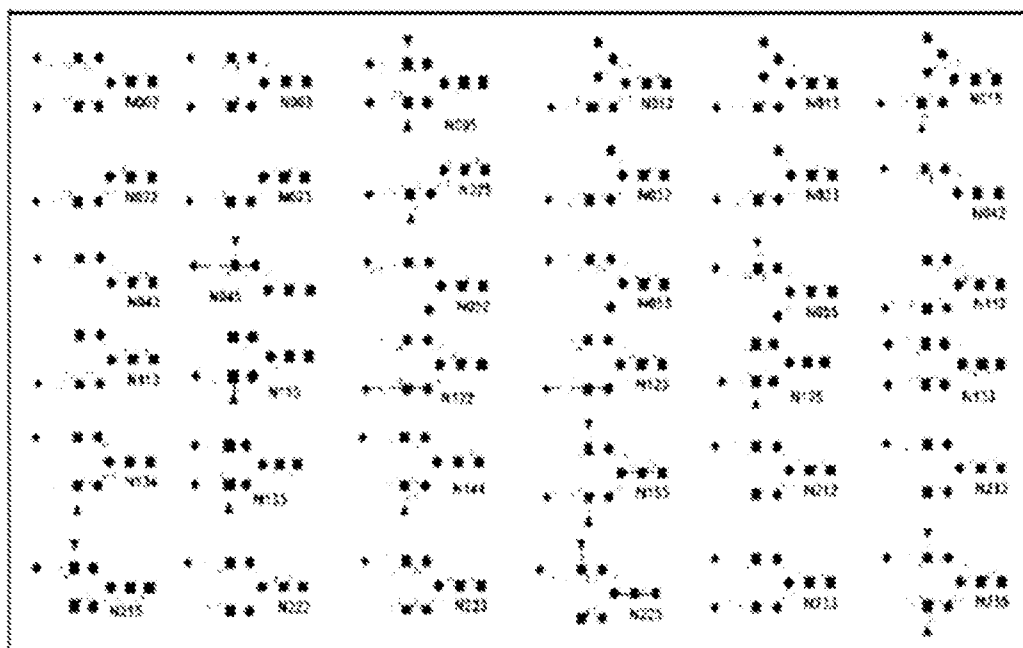


Figura 9

Glicanos Neu 5 Ac



Símbolos

- Man
- GcNAc
- Gal
- ▼ L-Fuc
- ◆ Neu5Ac
- ◇ Neu5Gc

Figura 9 (continuación)

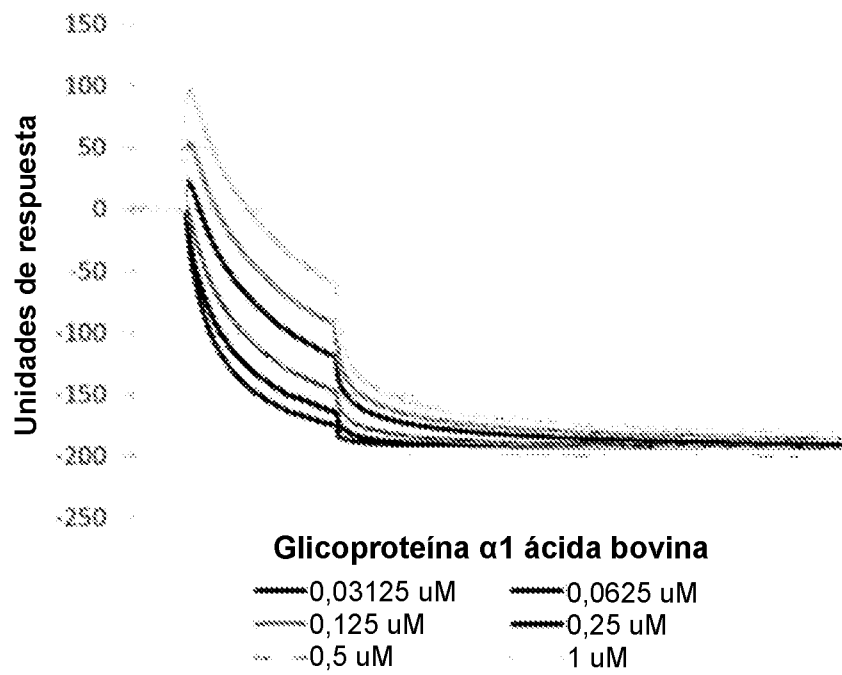


Figura 10

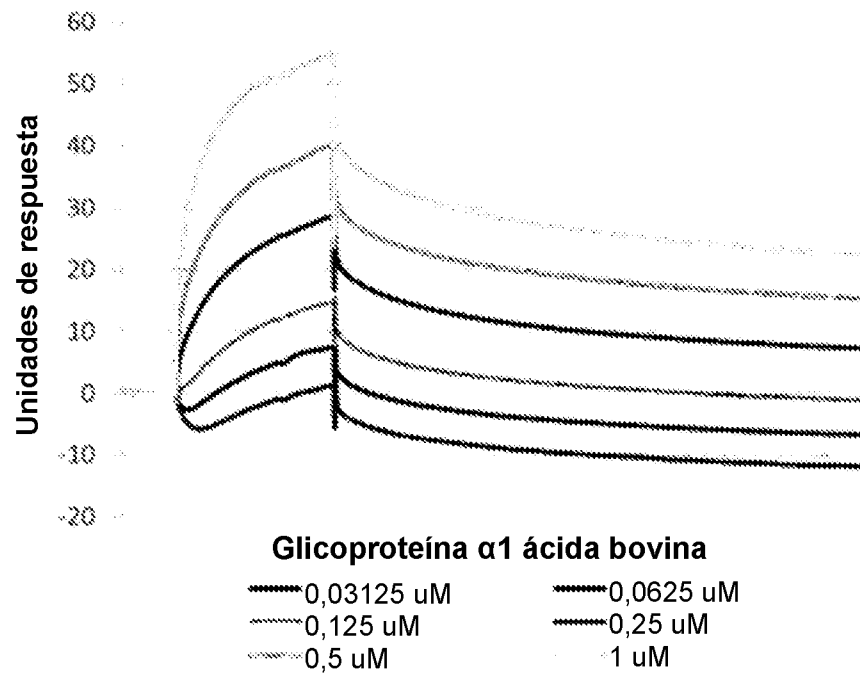


Figura 11