

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 986 557**

51 Int. Cl.:

C12N 9/24 (2006.01)

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.05.2018** **PCT/EP2018/063833**

87 Fecha y número de publicación internacional: **29.11.2018** **WO18215657**

96 Fecha de presentación y número de la solicitud europea: **25.05.2018** **E 18728577 (0)**

97 Fecha y número de publicación de la concesión europea: **29.05.2024** **EP 3630968**

54 Título: **Enzimas para el análisis de glucanos**

30 Prioridad:

26.05.2017 GB 201708471

26.05.2017 GB 201708476

24.04.2018 GB 201806655

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.11.2024

73 Titular/es:

GENOVIS AB (100.0%)

Box 4

244 21 Kaevlinge, SE

72 Inventor/es:

LEO, FREDRIK;

LOOD-ALAYÓN, ROLF;

BJORK, STEPHAN;

MEJARE, MALIN y

OLSSON, FREDRIK

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 986 557 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

Enzimas para el análisis de glucanos

5 **Campo de la invención**

La presente invención se refiere a enzimas y combinaciones de las mismas útiles para el estudio de glucoproteínas, y a los métodos de uso correspondientes.

10 **Antecedentes de la invención**

La glucosilación de proteínas desempeña un papel fundamental en muchas funciones fisiológicas en humanos, incluida la señalización, el transporte, la protección contra la actividad proteolítica, la adherencia, la respuesta inflamatoria, la colonización microbiana, etc. La mayoría de las cadenas de glucano unidas a las proteínas, ya sean unidas por O o N, están decoradas con ácidos siálicos terminales. Al ser el glucano más externo en las glucoproteínas, su presencia/ausencia a menudo es crítica para el efecto aguas abajo, incluido, por ejemplo, el potencial inflamatorio en las inmunoglobulinas. Los ácidos siálicos en las proteínas son heterogéneos, tanto en términos de presencia/ausencia en una proteína dada, como de modificaciones estructurales individuales. También están generalmente cargados negativamente, lo que complica el análisis de espectrometría de masas. Esto dificulta el estudio de las glucoproteínas, además de reducir la capacidad de los fabricantes para confirmar que un lote de glucoproteínas funcionará en una materia homóloga. Para superar estos problemas, se han realizado intentos de diseñar genéticamente células CHO para reducir la complejidad de los glucanos, aunque esto puede afectar la función. También se han utilizado enfoques químicos, pero a menudo dañan las proteínas. Existe la necesidad de un enfoque alternativo para eliminar el ácido siálico de las glucoproteínas. Además, una vez que se eliminan los ácidos siálicos, existe la necesidad de más herramientas para estudiar las cadenas de glucano restantes, particularmente aquellas que están unidas a O.

Breve descripción de la invención

30 La presente invención proporciona:

Una composición que comprende una primera y una segunda sialidasa,

35 en donde la primera sialidasa es un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 2; y

en donde la segunda sialidasa es un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 5; y

40 en donde la composición además comprende opcionalmente una O-glucosidasa que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 9,

45 opcionalmente en donde dicha primera y/o segunda sialidasa incluye una metionina adicional en el extremo N y/o una etiqueta de histidina en el extremo C, cuya etiqueta puede unirse al extremo C mediante un enlazador.

La composición puede comprender adicionalmente una glucosidasa y/o una proteasa, que está opcionalmente presente en forma altamente purificada o aislada.

50 La presente invención también proporciona un método para modificar una glucoproteína que comprende poner en contacto una muestra que contiene la glucoproteína con una composición de acuerdo con la invención,

opcionalmente en donde se analizan los productos resultantes.

55 **Breve descripción de las figuras**

Figura 1: Expresión y purificación de sialidasas *Akkermansia*. Las cuatro sialidasas, representadas en la presente por Am0707 y Am1757, se expresaron bien y pudieron purificarse hasta una alta homogeneidad. Los valores (mg/mL) indican la concentración obtenida después de la purificación de His.

60 **Figura 2:** Actividad de sialidasa sobre diferentes enlaces de ácido siálico. Los dos productos de sialidasa (Am1757 y Mix (Am1757 y Am0707)) se incubaron con tres sustratos diferentes (que representan los 3 tipos indicados de enlace de ácido siálico) durante 30 minutos, después de lo cual se midió la cantidad de ácido siálico libre.

65

Figura 3: Se determinaron las condiciones óptimas para Am0707 y Am1757. Las condiciones que afectan la actividad enzimática de las sialidasas, incluidos los iones, el NaCl y el pH, se investigaron utilizando un sustrato de sialillactosa 2-3, cuantificándose los ácidos siálicos libres después de 15 minutos de incubación.

Figura 4: El tratamiento combinado con sialidasa de las glucoproteínas es más eficiente que el uso de sialidasas individuales. Las sialidasas se incubaron con 0,5 µg de fetuina durante 60 minutos después de lo cual se separaron usando SDS-PAGE. Smix contiene tanto Am0707 como Am1757, mientras que se utilizó una sialidasa de *C. perfringens* como comparación de referencia. Todas las reacciones tuvieron lugar a 37 °C, en Tris-HCl 20 mM pH 6,8, excepto para el producto de referencia que se incubó de acuerdo con las instrucciones del fabricante.

Figura 5: GVS_Smix es superior a las sialidasas actuales disponibles comercialmente. Se probaron varias sialidasas comerciales existentes de New England Biolabs (primeras 3 columnas en cada conjunto) junto con la mezcla AM1757 +Am0707 (GVS_Smix) y cada una de las enzimas Am1757 y Am0707 individualmente. Cada una se incubó con sustratos de sialidasa específicos (que representan los 3 tipos indicados de enlace de ácido siálico) durante 30 min a 37 °C antes de analizarse. Las FU (unidades de fluorescencia) representan la cantidad de ácidos siálicos liberados.

Figura 6: GVS_Smix puede asialilar completamente la fetuina. Las técnicas de electroforesis en gel de poliacrilamida con dodecil sulfato de sodio [SDS-PAGE, por sus siglas en inglés] y de transferencia tipo Southern [SNA, por sus siglas en inglés] muestran que Smix (la mezcla AM1757+Am0707), así como dos productos de New England Biolabs (NEB A y NEB O) pueden asialilar completamente los enlaces de ácido siálico 2-3 y 2-6 en fetuina. Las dos 2-3 enzimas específicas Am1757 y New England Biolabs NEB S no pudieron asialilar completamente la fetuina.

Figura 7: El GVS_Smix (la mezcla AM1757 +Am0707; primer conjunto de barras) libera ácidos siálicos de proteínas nativas a un nivel similar o mejor que los tres productos comerciales probados (los últimos tres conjuntos de barras). También se muestra la enzima Am1757 sola (segundo conjunto de barras). Las proteínas se mezclaron con sialidasas en su respectivo amortiguador y se incubaron durante 15 minutos antes de la adición del amortiguador de desarrollo de ácido siálico. Todas las incubaciones tuvieron lugar a 37 °C en los tampones respectivos (para GVS-Smix esto fue Tris-HCl 20 mM pH 6,8).

Figura 8: Expresión recombinante de O-glucosidasas. (A) La O-glucosidasa de *S. oralis* y *B. bifidum* se expresó, se purificó por afinidad y se separó en SDS-PAGE. (B) Ensayo de estabilidad de O-glucosidasa de *S. oralis* almacenada a 4 °C durante 0-6 días.

Figura 9: Las O-glucosidasas pueden hidrolizar O-glucanos marcados con pNP de los tipos Núcleo 1, 2 y 3. Se incubaron diferentes concentraciones de O-glucosidasas de *S. oralis* (S.o) y *B. bifidum* (B.b) con (A) Núcleo 1, (B) Núcleo 2 y (C) Núcleo 3 O-glucanos, y se midió la liberación de pNP como cambio en la absorbancia a 405 nm.

Figura 10: La glucosidasa de *S. oralis* tiene mayor actividad en pH básico suplementado con MgCl₂. Las O-glucosidasas se incubaron a diferente pH (A), con diferentes iones (B) y diferentes concentraciones de MgCl₂ (C) para determinar el amortiguador óptimo.

Figura 11: La actividad de la O-glucosidasa es dependiente del tiempo y de la dosis. Se incubó TNFaR (1 µg) solo con la mezcla de sialidasa GVS_Smix (carriles marcados con TNFaR (Smix)) como control, o con la mezcla de sialidasa GVS_Smix en combinación con las diferentes cantidades de O-glucosidasa de *S. oralis* (So) que se muestran (" +5 µg So etc"). Las cantidades indicadas de O-glucosidasas añadidas se basan en los valores de las nanogotas. Sin embargo, debido a la fragmentación, la cantidad real de proteína de longitud completa añadida está más cerca de aproximadamente el 10-20% del valor escrito. Estas muestras se incubaron en Tris-HCl 20 mM pH 8,0, complementado con CaCl₂ 2 mM, pero sin MgCl₂.

Figura 12: Las O-glucosidasas pueden actuar sobre todas las glucoproteínas naturales investigadas. Se incubaron diferentes proteínas naturales (P) con la mezcla GVS_Smix de sialidasas (Smix) y O-glucosidasas de *S. oralis* (So-glyk) y *B. bifidum* (Bb-glyk). Las incubaciones fueron o/n, después de lo cual los productos se analizaron en SDS-PAGE.

Figura 13: La O-glucosidasa de *S. oralis* se compara favorablemente con los productos disponibles comercialmente para la acción contra las glucoproteínas naturales. El TNFaR natural se incubó solo con la mezcla de sialidasa GVS_Smix (carriles etiquetados como TNFaR (Smix)) como control, o con la mezcla de sialidasa GVS_Smix en combinación con las diferentes cantidades de O-glucosidasa de *S. oralis* (So) que se muestran (" + So (0,5µg)", etc.) o con una O-glucosidasa (NEB) disponible comercialmente en la cantidad mostrada (" +NEB (3ul)", etc.) en combinación con una sialidasa suministrada por el fabricante. Todas las enzimas se incubaron con TNFaR en sus respectivos tampones

durante 1 hora o 16 horas. La dosis más alta de la glucosidasa de *S. oralis* ("So (0,5ug)") es aproximadamente igual a 0.3 ul de la O-glucosidasa comercial (NEB) en molaridad. La banda distintiva de bajo peso molecular (parte superior del gel) en las muestras tratadas con NEB es la sialidasa acompañante.

Figura 14: La composición de los glucanos influye en la actividad. Incubación del sustrato con sialidasa (S) y/o galactosidasa (G) antes de la adición de O-glucosidasa (So); incubación durante la noche a 37 °C en tampones óptimos (Tris-HCl 20 mM pH 8,0).

Figura 15: El producto final es competitivo. Usando una relación enzima:sustrato 1:40 de la O-glucosidasa, así como una relación de sustrato enzimático de 1:40+1:40 para Am1757 +Am0707 en la mezcla de sialidasa, las tres enzimas en Tris-HCl 20 mM pH 6,8 (carriles etiquetados Smix / O-glyk), la actividad de la O-glucosidasa de *S. oralis* se comparó con la de la combinación comercial de O-glucosidasa + sialidasa (carriles etiquetados NEB (kit de O-glyk)). La eliminación de O-glucano de TNFaR por la mezcla de O-glucosidasa + sialidasa de *S. oralis* ("GVS") se verificó con transferencia de lectina, usando ConA (glucanos unidos a N) y jacalina (glucanos enlazados a O), TNFaR no tratado como control.

Figura 16: El haz de O-glucosidasa + sialidasa (Smix / O-glyk) es activo contra una amplia variedad de glucoproteínas. Las enzimas añadidas en su concentración final y formulación (1:40) en Tris-HCl 20 mM pH 6,8 se incubaron con las glucoproteínas durante 4 h a 37 °C y luego se separaron en SDS-PAGE.

Figura 17: Los enlaces de ácido siálico específicos influyen en la actividad de la O-glucosidasa. Las enzimas (Am1757, Am0707 o Smix) añadidas en su concentración final y formulación (1:40) en Tris-HCl 20 mM pH 6,8 se incubaron con las glucoproteínas durante 4 h a 37 °C y luego se separaron en SDS-PAGE.

Figura 18: La O-glucosidasa de *S. oralis* sobresale al menos en parte debido a las potentes sialidasas. Se incubaron O-glucosidasa de *S. oralis* (GVS O-glyk) u O-glucosidasa comercial (NEB O-glyk) con la mezcla de sialidasa GVS_Smix (sialidasa GVS) o con sialidasas comerciales (NEB A, NEB S, NEB O). Se utilizaron tampones apropiados (por ejemplo, Tris-HCl 20 mM pH 6,8 para GVS), y cada conjunto de enzimas se incubó con Enbrel (Etanercept) durante 4 h a 37 °C, antes de la separación de los productos en SDS-PAGE.

Figura 19: 2-3 ácidos siálicos unidos limitan la eficiencia de una endoproteasa específica de O-glucoproteína (LS). La incubación simultánea de LS con un conjunto de diversas sialidasas durante 30 min - 20 h, utilizando Enbrel como sustrato de glucoproteína, reveló la mayor eficiencia en presencia de la sialidasa específica 2-3 1757, o con la mezcla (0707 + 1757), mientras que la sialidasa de amplio espectro 0707 no fue necesaria para la actividad aparentemente completa de LS, lo que sugiere que los enlaces 2-6 (y 2-8) no son una preocupación para la actividad de LS.

Breve descripción de las secuencias

Las SEQ ID NO: 1, 2 y 3, cada una es una secuencia de aminoácidos de una sialidasa aislada de *Akkermansia muciniphila*. La SEQ ID NO: 1 es la secuencia de tipo silvestre que incluye un motivo señal en el extremo N. La SEQ ID NO: 2 es la secuencia de tipo silvestre con el motivo señal eliminado. La SEQ ID NO: 3 es idéntica a la SEQ ID NO: 2, excepto que incluye una metionina N terminal adicional y un enlazador C terminal + etiqueta His₆. Cualquier secuencia que comprenda la secuencia de la SEQ ID NO: 2 (que incluye cada una de las SEQ ID NO: 1 a 3) puede denominarse Am0707 en el presente documento.

Las SEQ ID NO: 4, 5 y 6, cada una es una secuencia de aminoácidos de otra sialidasa aislada de *Akkermansia muciniphila*. La SEQ ID NO: 4 es la secuencia de tipo silvestre que incluye un motivo señal en el extremo N. La SEQ ID NO: 5 es la secuencia de tipo silvestre con el motivo señal eliminado. La SEQ ID NO: 6 es idéntica a la SEQ ID NO: 5, excepto que incluye una metionina N terminal adicional y un enlazador C terminal + etiqueta His₆. Cualquier secuencia que comprenda la secuencia de la SEQ ID NO: 5 (que incluye cada una de las SEQ ID NO: 4 a 6) puede denominarse Am1757 en el presente documento.

Las SEQ ID NO: 7, 8, 9 y 10 son cada una secuencia de aminoácidos de una O-glucosidasa aislada de *S. oralis*. La SEQ ID NO: 7 es la secuencia de tipo silvestre que incluye un motivo de señal en el extremo N y un motivo de anclaje de pared LPXTG en el extremo C. La SEQ ID NO: 8 es la secuencia de tipo silvestre con el motivo señal eliminado. La SEQ ID NO: 9 es la secuencia de tipo silvestre con el motivo de señal y el motivo de anclaje a la pared eliminado. La SEQ ID NO: 10 es idéntica a la SEQ ID NO: 9, excepto que incluye una metionina N terminal adicional y un enlazador C terminal + etiqueta His₆. Cualquier secuencia que comprenda la secuencia de la SEQ ID NO: 9 (que incluye cada una de las SEQ ID NO: 7 a 10) puede denominarse en el presente documento "O-glyk" o "So".

La SEQ ID NO: 11 es una secuencia de aminoácidos de una endoproteasa específica de O-glucoproteína.

La SEQ ID NO: 12 es la secuencia de aminoácidos de un polipéptido ejemplar que tiene actividad de endoproteasa específica de O-glucoproteína. Relativo a la SEQ ID NO: 11 que incluye una metionina N terminal adicional y un enlazador C terminal + etiqueta His6. El polipéptido que consiste en esta secuencia puede denominarse en la presente LS.

La SEQ ID NO: 13 es una secuencia de ácido nucleico ejemplar que codifica el polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 10.

Descripción detallada de la invención

Debe entenderse que las diferentes aplicaciones de los productos y métodos descritos pueden adaptarse a las necesidades específicas de la técnica. También debe entenderse que la terminología utilizada en la presente tiene el propósito de describir realizaciones particulares de la invención únicamente, y no pretende ser limitativa.

Tal como se usa en la presente memoria descriptiva y las reivindicaciones adjuntas, las formas singulares "un", "una" y "el/la" incluyen referentes plurales a menos que el contenido indique claramente lo contrario. Por lo tanto, por ejemplo, la referencia a "un polipéptido" incluye "polipéptidos" y similares.

Esta memoria descriptiva se refiere particularmente a polipéptidos que son sialidasas, O-glucosidasas y endoproteasas específicas de O-glucoproteína. Por lo tanto, los usos generales del término polipéptido se pueden aplicar a cada uno de estos tipos de enzimas.

Características generales del polipéptido

Un "polipéptido" se usa en la presente en su sentido más amplio para referirse a un compuesto de dos o más subunidades de aminoácidos, análogos de aminoácidos u otros peptidomiméticos. Por lo tanto, el término "polipéptido" incluye secuencias peptídicas cortas y también polipéptidos y proteínas más largos. Los términos "proteína", "péptido" y "polipéptido" se pueden usar indistintamente. Tal como se usa en la presente, el término "aminoácido" se refiere a aminoácidos naturales y/o no naturales o sintéticos, que incluyen isómeros ópticos D o L, y análogos de aminoácidos y peptidomiméticos.

Un polipéptido puede producirse mediante un método adecuado, que incluye métodos recombinantes o sintéticos. Por ejemplo, el polipéptido se puede sintetizar directamente usando técnicas estándar conocidas en la técnica, tales como química de fase sólida Fmoc, química de fase sólida Boc o mediante síntesis de péptidos en fase de solución. Alternativamente, un polipéptido puede producirse transformando una célula, típicamente una célula bacteriana, con una molécula de ácido nucleico o vector que codifica dicho polipéptido. La producción de polipéptidos mediante expresión en células hospedadoras bacterianas se describe a continuación y se ejemplifica en los Ejemplos. En la presente se proporcionan moléculas de ácido nucleico y vectores que codifican un polipéptido. También se proporciona en la presente una célula hospedadora que comprende tal ácido nucleico o vector. Un ejemplo de moléculas polinucleótidas que codifican un polipéptido descrito en la presente se proporciona como la SEQ ID NO: 13. Esta secuencia incluye en el extremo 5' un codón para la metionina N terminal (ATG) y, antes del codón de terminación (TAA) en el extremo 3', codones para un enlazador GSGLE y una etiqueta His 6x, que pueden excluirse opcionalmente. La inclusión opcional de una metionina adicional y una etiqueta se analizan con más detalle a continuación.

Los términos "molécula de ácido nucleico" y "polinucleótido" se usan indistintamente en la presente y se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sea desoxirribonucleótidos o ribonucleótidos, o análogos de estos. Los ejemplos no limitativos de polinucleótidos incluyen un gen, un fragmento de gen, ARN mensajero (ARNm), ADNc, polinucleótidos recombinantes, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. Un polinucleótido como se describe en la presente codifica un polipéptido y puede proporcionarse en forma aislada o sustancialmente aislada. Por sustancialmente aislada, se entiende que puede haber un aislamiento sustancial, pero no total, del polipéptido de cualquier medio circundante. Los polinucleótidos se pueden mezclar con portadores o diluyentes que no interferirán con su uso previsto y aún se considerarán sustancialmente aislados. Una secuencia de ácido nucleico que "codifica" un polipéptido seleccionado es una molécula de ácido nucleico que se transcribe (en el caso del ADN) y se traduce (en el caso del ARNm) en un polipéptido in vivo cuando se coloca bajo el control de secuencias reguladoras apropiadas, por ejemplo, en un vector de expresión. Los límites de la secuencia codificante están determinados por un codón de inicio en el extremo 5' (amino) y un codón de terminación de la traducción en el extremo 3' (carboxi). Tales secuencias de ácido nucleico pueden incluir, pero sin limitarse a ADNc de ARNm viral, procariota o eucariota, secuencias genómicas de ADN o ARN viral o procariota e incluso secuencias de ADN sintéticas. Una secuencia de terminación de la transcripción puede localizarse en 3' con respecto a la secuencia codificante.

Los polinucleótidos se pueden sintetizar de acuerdo con métodos bien conocidos en la técnica, como se describe a modo de ejemplo en Sambrook et al (1989, Molecular Cloning - a laboratory manual; Cold Spring Harbor Press). Las moléculas de ácido nucleico descritas en la presente pueden proporcionarse en forma de un casete de

expresión que incluye secuencias de control unidas operativamente a la secuencia insertada, permitiendo así la expresión del polipéptido in vivo. Estos casetes de expresión, a su vez, se proporcionan típicamente dentro de vectores (por ejemplo, plásmidos o vectores virales recombinantes). Tal casete de expresión se puede administrar directamente a un sujeto hospedador. Alternativamente, un vector que comprende un polinucleótido descrito en la presente se puede administrar a un sujeto hospedador. Preferentemente, el polinucleótido se prepara y/o administra utilizando un vector genético. Un vector adecuado puede ser cualquier vector que sea capaz de portar una cantidad suficiente de información genética y permitir la expresión de un polipéptido.

En la presente se describen vectores de expresión que comprenden tales secuencias de polinucleótidos. Tales vectores de expresión se construyen de forma rutinaria en la técnica de la biología molecular y pueden implicar, por ejemplo, el uso de ADN plasmídico e iniciadores, promotores, potenciadores y otros elementos apropiados, tales como, por ejemplo, señales de poliadenilación que pueden ser necesarias y que se colocan en la orientación correcta, para permitir la expresión de un péptido. Otros vectores adecuados serían evidentes para los expertos en la técnica. A modo de ejemplo adicional a este respecto, nos referimos a Sambrook et al.

También se describen en la presente células que se han modificado para expresar un polipéptido. Tales células típicamente incluyen células procariotas tales como células bacterianas, por ejemplo, *E. coli*. Tales células se pueden cultivar usando métodos rutinarios para producir un polipéptido.

Un polipéptido puede derivatizarse o modificarse para ayudar con su producción, aislamiento o purificación. Por ejemplo, cuando se produce un polipéptido mediante expresión recombinante en una célula hospedadora bacteriana, la secuencia del polipéptido puede incluir un residuo de metionina (M) adicional en el extremo N para mejorar la expresión. Como otro ejemplo, el polipéptido puede derivatizarse o modificarse mediante la adición de un ligando que es capaz de unirse directa y específicamente a un medio de separación. Alternativamente, el polipéptido puede derivatizarse o modificarse mediante la adición de un miembro de un par de unión y el medio de separación comprende un reactivo que se deriva o modifica mediante la adición del otro miembro de un par de unión. Se puede utilizar cualquier par de unión adecuado. En una realización preferida donde el polipéptido se derivatiza o modifica mediante la adición de un miembro de un par de unión, el polipéptido está preferentemente etiquetado con histidina o etiquetado con biotina. Típicamente, la secuencia codificante de aminoácidos de la etiqueta de histidina o biotina se incluye a nivel génico y el polipéptido se expresa de forma recombinante en *E. coli*. La etiqueta de histidina o biotina está típicamente presente en cualquier extremo del polipéptido, preferentemente en el extremo C. Puede unirse directamente al polipéptido o unirse indirectamente mediante cualquier secuencia enlazadora adecuada, tal como 3, 4 o 5 aminoácidos. El enlazador típicamente puede consistir predominantemente en residuos de glicina y serina. La etiqueta de histidina típicamente consiste en seis residuos de histidina, aunque puede ser más larga que esto, típicamente hasta 7, 8, 9, 10 o 20 aminoácidos o más corta, por ejemplo 5, 4, 3, 2 o 1 aminoácidos.

Un polipéptido puede proporcionarse en una forma sustancialmente aislada o purificada. Es decir, aislado de la mayoría de los otros componentes presentes en un extracto celular de una célula en la que se expresó el polipéptido. Por sustancialmente purificado, se entenderá que el polipéptido se purifica hasta al menos 50%, 60%, 70%, 80% o preferentemente al menos 90% de homogeneidad. El nivel de pureza se puede evaluar por cualquier medio adecuado, pero normalmente implica el análisis SDS-PAGE de una muestra, seguido de la detección con azul de Coomassie. Un polipéptido puede mezclarse con portadores, diluyentes o conservantes que no interfieran con el propósito previsto del polipéptido y que aún se consideren sustancialmente aislados o purificados. Cuando se proporciona un polipéptido en una composición con un componente activo adicional, tal como otro polipéptido, cada uno de dichos polipéptidos se purificó individualmente hasta un alto nivel de homogeneidad antes de mezclarse a una relación apropiada para el propósito previsto de cada uno. Por ejemplo, cada uno de dos polipéptidos puede purificarse hasta al menos 90% de homogeneidad antes de combinarse en una relación 1:1.

Se puede proporcionar un polipéptido (o mezcla de estos) en forma liofilizada, adecuado para la reconstitución en solución acuosa antes de su uso. La composición liofilizada tiene una estabilidad mejorada que permite un almacenamiento más prolongado del polipéptido. En la presente se proporciona un método para preparar un polipéptido (o mezcla de estos) en forma liofilizada, que comprende liofilizar dicho polipéptido (o mezcla) en un tampón adecuado, tal como solución salina tamponada con Tris (TBS). Típicamente, un polipéptido se purifica sustancialmente antes de la liofilización. El polipéptido (o mezcla) resultante en forma liofilizada también se proporciona. Un método para preparar una solución de un polipéptido (o mezcla), que comprende proporcionar el polipéptido (o mezcla) en forma liofilizada y reconstituir con un portador o diluyente adecuado, tal como agua, también se proporciona.

Un polipéptido puede inmovilizarse usando métodos conocidos en la técnica, por ejemplo, como se describe en Datta S et al., *Enzyme immobilization: an overview on techniques and support materials*, 3 Biotech, 3(1):1-9 (2013). Por ejemplo, el polipéptido puede inmovilizarse por adsorción, unión covalente, inmovilización por afinidad o atrapamiento. Los materiales que se pueden usar como soportes incluyen, pero no se limitan a, por ejemplo, soportes naturales tales como agarosa, colágeno, gelatina, celulosa, pectina, sefarosa, materiales inorgánicos tales como cerámica, sílice, vidrio, carbón activado o carbón vegetal, o polímeros sintéticos. Por ejemplo, el polipéptido puede inmovilizarse en sefarosa.

Sialidasas

Características funcionales

Además de utilizar enfoques químicos y genéticos para modificar las cadenas de glucanos, varias enzimas (glucosidasas) pueden actuar sobre los enlaces que unen los ácidos siálicos a otros glucanos. Estas enzimas, denominadas sialidasas o neuraminidasas, muestran un alto grado de especificidad para tipos particulares de enlace de ácido siálico. Comúnmente se encuentran tres tipos distintos de enlaces dentro de las glucoproteínas humanas, los enlaces alfa(2-3) son la forma dominante, seguidos por alfa(2-6) y alfa(2-8). Estos tipos de enlaces pueden denominarse en la presente 2-3, 2-6 y 2-8 enlaces con fines de simplicidad. Un enlace 2-3 significa que el átomo de carbono en la posición número 2 de la hexosa de ácido siálico está unido, a través de un átomo de oxígeno, al carbono en la posición 3 de la hexosa del glucano enlazado. En consecuencia, un enlace 2-6 o un enlace 2-8 significa que la unión es a la posición 6 o la posición 8 de la hexosa del glucano unido, respectivamente.

La mayoría de las sialidasas conocidas son específicas para el enlace 2-3 (escindiéndolo con una actividad muy alta), o son capaces de escindir una gama más amplia de enlaces, típicamente todos los enlaces 2-3, 2-6 y 2-8. Estos diferentes tipos de sialidasa pueden denominarse de espectro estrecho o de amplio espectro, respectivamente. Las sialidasas de amplio espectro generalmente exhiben una alta actividad contra 2-3 enlaces, con una actividad decreciente contra 2-6 y una actividad muy baja contra 2-8 enlaces. Las enzimas que escinden 2-8 enlaces de manera eficiente son comparativamente raras (incluso desconocidas) en el campo.

La actividad enzimática de una sialidasa se puede evaluar mediante cualquier método adecuado, tal como los descritos en los Ejemplos. Un método adecuado puede incluir incubar una sialidasa conocida o sospechada con un sustrato de sialidasa estándar, tal como una o más moléculas pequeñas que comprenden colectivamente 2-3, 2-6 y 2-8 enlaces de tipo. Tales moléculas pequeñas incluyen 2-3'-sialil-lactosa, 2-6'-sialil-lactosa y ácido colomínico (2-8'). La actividad de sialidasa en tales moléculas dará como resultado ácidos siálicos libres, que pueden cuantificarse mediante métodos rutinarios. Alternativamente, la actividad de sialidasa se puede evaluar utilizando una glucoproteína como sustrato. Cualquier producto de escisión resultante se puede detectar y cuantificar usando métodos rutinarios tales como SDS-PAGE o transferencias de lectina.

Características estructurales

Los presentes inventores han identificado y caracterizado varias sialidasas de la bacteria intestinal comensal *Akkermansia muciniphila*. Una de las sialidasas, denominada en la presente Am0707, tiene una actividad inesperadamente alta contra 2-8 enlaces, pero también puede escindir 2-3 y 2-6 enlaces. Por lo tanto, puede considerarse una sialidasa de amplio espectro. Otra sialidasa, denominada en la presente Am1757, tiene una actividad exclusivamente alta contra 2-3 enlaces. Por lo tanto, puede considerarse una sialidasa de espectro estrecho.

La estructura primaria de tipo silvestre completa (secuencia de aminoácidos) de la primera sialidasa (Am0707) se muestra en la SEQ ID NO: 1. La secuencia con el motivo de señal eliminado se muestra en la SEQ ID NO: 2. La primera sialidasa puede comprender, consistir esencialmente o consistir en la secuencia de la SEQ ID NO: 2, y normalmente no tiene más de 400 aminoácidos.

La estructura primaria de tipo silvestre completa (secuencia de aminoácidos) de la segunda sialidasa (Am1757) se muestra en la SEQ ID NO: 4. La secuencia con el motivo de señal eliminado se muestra en la SEQ ID NO: 5. La segunda sialidasa puede comprender, consistir esencialmente o consistir en la secuencia de la SEQ ID NO: 5, y normalmente no tiene más de 600 aminoácidos.

De acuerdo con la descripción, cada una de dicha primera y/o dicha segunda sialidasa puede reemplazarse independientemente por una variante de cada una de ellas, siempre que se conserve la actividad enzimática. Una variante de dicha sialidasa puede comprender, consistir esencialmente o consistir en una variante de la secuencia de aminoácidos de la secuencia de la SEQ ID NO: 2 o 5, respectivamente, que es al menos 50% idéntica a dicha secuencia de aminoácidos. La secuencia variante puede ser al menos 60%, al menos 70%, al menos 80%, al menos 85%, al menos 90%, al menos 95%, al menos 98% o al menos 99% idéntica a dicha secuencia de aminoácidos. El nivel de identidad es preferentemente de al menos el 85% o superior. La identidad con respecto a una secuencia se puede medir en una región de al menos 100, al menos 200, al menos 300, al menos 350, al menos 400 o al menos 500 o más aminoácidos contiguos de la secuencia, o más preferentemente en toda la longitud de la secuencia. Una variante es típicamente de una longitud que no es más de 50 aminoácidos más larga o más corta que la secuencia de referencia, y es preferiblemente de aproximadamente (o exactamente) la misma longitud que la secuencia de referencia.

La identidad de aminoácidos se puede calcular usando cualquier algoritmo adecuado. Por ejemplo, los algoritmos PILEUP y BLAST se pueden utilizar para calcular secuencias de identidad o alineación (tales como la identificación de secuencias equivalentes o correspondientes (típicamente en sus configuraciones predeterminadas), por ejemplo, como se describe en Altschul S. F. (1993) J Mol Evol 36:290-300; Altschul, S, F et al (1990) J Mol Biol

215:403-10. El software para realizar análisis BLAST está disponible públicamente a través del Centro Nacional de Información Biotecnológica (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo implica identificar primero el par de secuencias de alta puntuación (HSP) mediante la identificación de palabras cortas de longitud W en la secuencia de consulta que coinciden o satisfacen alguna puntuación umbral T de valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de base de datos. T se conoce como el umbral de puntuación de la palabra vecina (Altschul *et al*, supra). Estos aciertos iniciales de palabras vecinas actúan como semillas para iniciar búsquedas para encontrar HSP que los contengan. Los aciertos de palabras se extienden en ambas direcciones a lo largo de cada secuencia en la medida en que se puede aumentar la puntuación de alineación acumulativa. Las extensiones para los aciertos de palabras en cada dirección se detienen cuando: la puntuación de alineación acumulativa cae en la cantidad X desde su valor máximo alcanzado; la puntuación acumulativa va a cero o menos, debido a la acumulación de una o más alineaciones de residuos de puntuación negativa; o se alcanza el final de cualquiera de las secuencias. Los parámetros del algoritmo BLAST W , T y X determinan la sensibilidad y la velocidad de la alineación. El programa BLAST utiliza por defecto una longitud de palabra (W) de 11, la matriz de puntuación BLOSUM62 (véase Henikoff and Henikoff (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919) alineaciones (B) de 50, expectativa (E) de 10, $M=5$, $N=4$ y una comparación de ambas cadenas.

El algoritmo BLAST realiza un análisis estadístico de la similitud entre dos secuencias; véase, por ejemplo, Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90: 5873-5877. Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña ($P(N)$), que proporciona una indicación de la probabilidad por la cual una coincidencia entre dos secuencias de polinucleótidos o aminoácidos ocurriría por casualidad. Por ejemplo, una secuencia se considera similar a otra secuencia si la probabilidad de suma más pequeña en comparación de la primera secuencia con la segunda secuencia es menor que aproximadamente 1, preferentemente menor que aproximadamente 0,1, más preferentemente menor que aproximadamente 0,01 y lo más preferentemente menor que aproximadamente 0,001. Alternativamente, el paquete UWGCG proporciona el programa BESTFIT que se puede utilizar para calcular la identidad (por ejemplo, se utiliza en su configuración predeterminada) (Devereux et al (1984) Nucleic Acids Research 12, 387-395).

Tal como se describe en la presente, la secuencia de una sialidasa puede comprender una variante de la SEQ ID NO respectiva en la que se realizan modificaciones, tales como adiciones, deleciones o sustituciones de aminoácidos con respecto a la secuencia de dicha SEQ ID NO. A menos que se especifique lo contrario, las modificaciones son preferentemente sustituciones de aminoácidos conservativas. Las sustituciones conservadoras reemplazan los aminoácidos con otros aminoácidos de estructura química similar, propiedades químicas similares o volumen de cadena lateral similar. Los aminoácidos introducidos pueden tener polaridad, hidrofilia, hidrofobicidad, basicidad, acidez, neutralidad o carga similares a los aminoácidos que reemplazan. Alternativamente, la sustitución conservadora puede introducir otro aminoácido que es aromático o alifático en lugar de un aminoácido aromático o alifático preexistente. Los cambios conservadores de aminoácidos son bien conocidos en la técnica y se pueden seleccionar de acuerdo con las propiedades de los 20 aminoácidos principales como se define a continuación en la Tabla A1. Cuando los aminoácidos tienen una polaridad similar, esto se puede determinar por referencia a la escala de hidropatía para las cadenas laterales de aminoácidos en la Tabla A2. Una secuencia de una sialidasa puede comprender una variante de la SEQ ID NO respectiva en la que se realizan hasta 10, 20, 30, 40, 50 o 60 sustituciones conservadoras.

Tabla A1 - Propiedades químicas de los aminoácidos

Ala (A)	alifático, hidrófobo, neutro	Met (M)	hidrófobo, neutro
Cys (C)	polar, hidrófobo, neutro	Asn (N)	polar, hidrófilo, neutro
Asp (D)	polar, hidrófilo, cargado (-)	Pro (P)	hidrófobo, neutro
Glu (E)	polar, hidrófilo, cargado (-)	Gln (Q)	polar, hidrófilo, neutro
Phe (F)	aromático, hidrófobo, neutro	Arg (R)	polar, hidrófilo, cargado (+)
Gly (G)	alifático, neutro	Ser (S)	polar, hidrófilo, neutro
His (H)	aromático, polar, hidrófilo, cargado (+)	Thr (T)	polar, hidrófilo, neutro
Ile (I)	alifático, hidrófobo, neutro	Val (V)	alifático, hidrófobo, neutro
Lys (K)	polar, hidrófilo, cargado(+)	Trp (W)	aromático, hidrófobo, neutro
Leu (L)	alifático, hidrófobo, neutro	Tyr (Y)	aromático, polar, hidrófobo

Tabla A2 - Escala de hidropatía

Cadena lateral	Hidropatía
Ile	4,5

Cadena lateral	Hidropatía
Val	4,2
Leu	3,8
Phe	2,8
Cys	2,5
Met	1,9
Ala	1,8
Gly	-0,4
Thr	-0,7
Ser	-0,8
Trp	-0,9
Tyr	-1,3
Pro	-1,6
His	-3,2
Glu	-3,5
Gin	-3,5
Asp	-3,5
Asn	-3,5
Lys	-3,9
Arg	-4,5

Solo como referencia, se entenderá que una sialidasa puede reemplazarse por un fragmento más corto de la SEQ ID NO respectiva o de una variante de esta como se describió anteriormente. Los fragmentos pueden describirse como una forma truncada de dicha SEQ ID NO que retiene la actividad enzimática. Tales fragmentos son más cortos que la SEQ ID NO correspondiente y típicamente tienen al menos 100, 150, 200, 250, 300, 350, 400, 450 o 500 aminoácidos de longitud.

Cualquier sialidasa descrita en la presente puede incluir opcionalmente una metionina adicional en el extremo N y/o una histidina u otra etiqueta en el extremo C. Tales secuencias adicionales pueden ayudar con la expresión y/o purificación. Una etiqueta de histidina consiste preferentemente en seis residuos de histidina. La etiqueta de histidina se une preferentemente al extremo C mediante un enlazador, que es típicamente una secuencia corta de aminoácidos, tal como 3 - 5 aminoácidos. El enlazador típicamente consiste predominantemente en residuos de glicina y serina, y puede incluir preferentemente la secuencia GSG. Por ejemplo, GSG y GSGLE son enlazadores adecuados.

En resumen, por lo tanto, la primera sialidasa de acuerdo con la invención es:

(a) un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 2.

También se describe en el presente documento como referencia una primera sialidasa, que es:

(a) un polipéptido que comprende o consiste en una secuencia de aminoácidos que es al menos 85% idéntica a la secuencia de aminoácidos de la SEQ ID NO: 2 o

(b) un polipéptido que comprende o consiste en una secuencia de aminoácidos que es un fragmento de la secuencia de la SEQ ID NO: 2 o un fragmento de un aminoácido que es 85% idéntico a la secuencia de aminoácidos de la SEQ ID NO: 2;

La segunda sialidasa de acuerdo con la invención es:

(a) un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 5;

También se describe en el presente documento como referencia una segunda sialidasa, que es:

(b) un polipéptido que comprende o consiste en una secuencia de aminoácidos que es al menos 85% idéntica a la secuencia de aminoácidos de la SEQ ID NO: 5 o

(c) un polipéptido que comprende o consiste en una secuencia de aminoácidos que es un fragmento de la secuencia de la SEQ ID NO: 5 o un fragmento de un aminoácido que es 85% idéntico a la secuencia de aminoácidos de la SEQ ID NO: 5;

opcionalmente en donde dicha primera y/o segunda sialidasa incluye una metionina adicional en el extremo N y/o una etiqueta de histidina en el extremo C, cuya etiqueta puede unirse al extremo C mediante un enlazador.

Una primera sialidasa ejemplar es el polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 3. Una segunda sialidasa ejemplar es el polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 6.

Composiciones de sialidasa

Tal como se describe en la presente, una composición de sialidasa comprende al menos una sialidasa, preferentemente en forma sustancialmente aislada o purificada. Como en la descripción general relacionada con los polipéptidos establecidos anteriormente, esto típicamente significa aislado de la mayoría de los otros componentes presentes en un extracto celular de una célula en la que se expresó la sialidasa. Por sustancialmente purificado, se entenderá que la sialidasa se purifica hasta al menos 50%, 60%, 70%, 80% o preferentemente al menos 90% de homogeneidad. El nivel de pureza se puede evaluar por cualquier medio adecuado, pero normalmente implica el análisis SDS-PAGE de una muestra, seguido de la detección con azul de Coomassie. Una sialidasa puede mezclarse con portadores, diluyentes o conservantes que no interfieran con el propósito previsto del polipéptido y que aún se consideren sustancialmente aislados o purificados. La composición de sialidasa puede comprender un componente activo adicional, tal como otra sialidasa u otra enzima, en cuyo caso cada uno de dichos componentes se purificó individualmente a un alto nivel de homogeneidad antes de mezclarse a una relación adecuada para el propósito previsto de cada uno. En una composición de sialidasa preferida de la invención, la composición comprende una primera sialidasa y una segunda sialidasa que se purifican cada una hasta al menos 90% de homogeneidad y están presentes en una relación 1:1 entre sí. Tal composición puede incluir un componente activo adicional, tal como otra enzima que no es una sialidasa. La otra enzima puede ser una proteasa y/o una glucosidasa. La proteasa es preferentemente una endoproteasa específica de O-glucoproteína. La glucosidasa es preferentemente una O-glucosidasa. Ambos tipos de enzimas se analizan con más detalle a continuación.

Cuando una composición de sialidasa incluye un componente activo que no es una sialidasa, la relación preferida del contenido total de sialidasa (por ejemplo, primera más segunda sialidasa) con respecto a la otra enzima será de 1:1. Por ejemplo, si una composición incluye 2000 unidades de otra enzima, también incluirá 2000 unidades de sialidasa, en la que si hay dos sialidasas, entonces dichas 2000 unidades comprenden 1000 unidades de la primera sialidasa y 1000 unidades de la segunda sialidasa.

Se puede proporcionar una composición de sialidasa (como con un polipéptido en general) en forma liofilizada, adecuada para la reconstitución en solución acuosa antes de su uso. La composición liofilizada tiene una estabilidad mejorada que permite un almacenamiento más prolongado de la(s) sialidasa(s). En la presente se proporciona un método para preparar una composición de sialidasa en forma liofilizada, que comprende liofilizar una o más sialidasas en un tampón adecuado, tal como solución salina tamponada con Tris (TBS). El tampón comprende preferentemente una baja concentración de NaCl, típicamente hasta 300 mM, 250 mM, 200 mM o 150 mM. La concentración de NaCl es preferentemente de aproximadamente 150 mM, tal como entre 125 mM y 175 mM. Típicamente, una sialidase se purifica sustancialmente antes de la liofilización. También se proporciona la forma liofilizada resultante de la composición. También se proporciona un método para preparar una composición de sialidasa que es una solución, que comprende proporcionar la composición en forma liofilizada y reconstituirla con un portador o diluyente adecuado, tal como agua.

Los presentes inventores determinaron que una primera sialidasa tiene una actividad inusualmente alta contra 2-8 enlaces. Por consiguiente, la presente invención proporciona una composición que comprende una primera sialidasa, una segunda sialidasa y una O-glucosidasa, en donde la primera sialidasa es un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 2; en donde la segunda sialidasa es un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 5; y la O-glucosidasa comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 9;

opcionalmente, en donde dicha primera y/o segunda sialidasa incluye una metionina adicional en el extremo N y/o una etiqueta de histidina en el extremo C, cuya etiqueta puede unirse al extremo C mediante un enlazador.

Dicha composición puede ser para su uso en un método de escisión de 2-8 enlaces de ácido siálico, preferentemente con alta eficacia. Un ejemplo de tal composición comprende la sialidasa que consiste en la secuencia de aminoácidos de la SEQ ID NO: 3.

Los presentes inventores también determinaron que una combinación de la primera sialidasa (Am0707) y la segunda sialidasa (Am1757) hidroliza 2-3, 2-6 y 2-8 enlaces con una eficiencia inusualmente alta, permitiendo así la eliminación eficiente de sustancialmente todos (típicamente >90%) los ácidos siálicos de cualquier glucoproteína. La combinación también fue sorprendentemente eficaz contra las glucoproteínas en un estado natural (es decir, no desnaturalizado).

Dicha composición puede ser para su uso en un método para asilar completamente una glucoproteína, o para escindir >90% de los enlaces siálicos en una glucoproteína, preferentemente con alta eficacia. La glucoproteína se encuentra preferentemente en un estado natural. Es decir, que no ha sido sometida a ningún tipo de condiciones desnaturalizantes.

Un ejemplo de una composición de la invención comprende la sialidasa que consiste en la secuencia de aminoácidos de la SEQ ID NO: 3 y la sialidasa que consiste en la secuencia de aminoácidos de la SEQ ID NO: 6, preferentemente en una relación 1:1, y una O-glucosidasa que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 9.

La actividad de sialidasa de una composición de sialidasa se puede evaluar usando los mismos métodos que se describieron anteriormente para sialidasas individuales. Sin embargo, se evalúa preferentemente utilizando una glucoproteína no desnaturalizada como sustrato. Los resultados se pueden comparar con los obtenidos en el mismo ensayo cuando el sustrato se pone en contacto con una sialidasa ejemplar o mezcla de la misma, tal como una mezcla 1:1 del polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 3 y el polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 6. Una unidad de tal mezcla de sialidasa es típicamente la cantidad requerida para hidrolizar ácidos siálicos de $\geq 90\%$ de 1 μg de glucoproteína (fetuína) cuando se incubaba en Tris 20 mM pH 6,8 a 37 °C durante 2 h a 37 °C como se monitorea por SDS-PAGE. Se considera que esto representa una alta eficiencia.

O-glucosidasa

Los presentes inventores también han identificado y caracterizado una O-glucosidasa de la bacteria comensal *Streptococcus oralis*, que reside en el tracto oral de mamíferos, que hidroliza eficazmente los glucanos enlazados a O particularmente cuando se usa en combinación con una composición de sialidasa como se describió anteriormente. La O-glucosidasa puede denominarse en la presente "O-glyk" o "So". La secuencia de tipo silvestre de O-glyk se proporciona como SEQ ID NO: 7, que incluye una secuencia señal y un motivo de anclaje de pared celular LPTXG. La secuencia de tipo silvestre de O-glyk que carece de la secuencia señal se proporciona como SEQ ID NO: 8. La secuencia de tipo silvestre de O-glyk que carece de la secuencia señal y la parte C terminal del motivo de anclaje de la pared celular se proporciona como SEQ ID NO: 9. Estas secuencias pueden modificarse opcionalmente para incluir una metionina adicional en el extremo N y/o una histidina u otra etiqueta en el extremo C. Tales secuencias adicionales pueden ayudar con la expresión (por ejemplo, en *E. coli*) y/o purificación. Una etiqueta de histidina consiste preferentemente en seis residuos de histidina. La etiqueta de histidina se une preferentemente al extremo C mediante un enlazador, que es típicamente una secuencia corta de aminoácidos, tal como 3 - 5 aminoácidos. El enlazador típicamente consiste predominantemente en residuos de glicina y serina, y puede incluir preferentemente la secuencia GSG. Por ejemplo, GSG y GSGLE son enlazadores adecuados. Una secuencia de O-glyk ejemplar que tiene una metionina adicional en el extremo N y un enlazador GSGLE y etiqueta His6 en el extremo C se proporciona como SEQ ID NO: 10. Cualquier referencia a "O-glyk" o "So" en la presente descripción puede significar cualquiera de las SEQ ID NO: 7, 8, 9 o 10, pero preferentemente se refiere a un polipéptido que comprende o consiste en la secuencia de aminoácidos de la SEQ ID NO: 9, y normalmente no tiene más de 2070 aminoácidos. Lo más preferido es un polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 10.

Los presentes inventores también descubrieron que la acción de una composición de sialidasa como se describió anteriormente también mejora la actividad de otras O-glucosidasas. Por lo tanto, la presente invención también proporciona un método para modificar una glucoproteína que comprende poner en contacto una muestra de glucoproteína tanto con una composición de sialidasa como se describió anteriormente como con una O-glucosidasa. La presente invención también proporciona una composición de sialidasa como se describió anteriormente que comprende una O-glucosidasa que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 9. En dicho método y dicha composición, dicha O-glucosidasa comprende o consiste en la secuencia de aminoácidos de la SEQ ID NO: 9. Aunque no es una realización de la presente invención, se puede usar cualquier otra O-glucosidasa, tal como una enzima obtenida de una bacteria intestinal, por ejemplo, *Enterococcus faecalis*. Una O-glucosidasa preferida de *E. faecalis* es un polipéptido que tiene la secuencia de aminoácidos de la entrada Uniprot B5UB72 versión 22. Otras O-glucosidasas adecuadas incluyen las descritas en WO2009129086, particularmente EngEF, EngPA y EngAA truncado como se describe en la página 7 y se muestra en la Figura 5 de WO2009129086.

Una variante de la O-glucosidasa puede comprender, consistir esencialmente o consistir en una variante de la secuencia de aminoácidos de la secuencia de la SEQ ID NO: 9. La variante de dicha SEQ ID NO puede definirse

como se estableció anteriormente con respecto a las sialidasas, excepto que la actividad enzimática relevante a retener es la actividad hidrolítica contra O-glucanos.

Solo como referencia, se entenderá que la O-glucosidasa puede reemplazarse por un fragmento más corto de la SEQ ID NO: 9 o de una variante de la misma como se describió anteriormente. Los fragmentos pueden describirse como una forma truncada de dicha SEQ ID NO que retiene la actividad enzimática. Tales fragmentos son más cortos que la SEQ ID NO: 3 y tienen típicamente al menos 300, 400, 500, 600, 800, 1000, 1200, 1300, 1400 o 1500 aminoácidos de longitud.

Cualquier O-glucosidasa descrita en la presente puede incluir opcionalmente una metionina adicional en el extremo N y/o una histidina u otra etiqueta en el extremo C. Tales secuencias adicionales pueden ayudar con la expresión y/o purificación. Una etiqueta de histidina consiste preferentemente en seis residuos de histidina. La etiqueta de histidina se une preferentemente al extremo C mediante un enlazador, que es típicamente una secuencia corta de aminoácidos, tal como 3 - 5 aminoácidos. El enlazador típicamente consiste predominantemente en residuos de glicina y serina, y puede incluir preferentemente la secuencia GSG. Por ejemplo, GSG y GSGLE son enlazadores adecuados. Un ejemplo de O-glucosidasa de este tipo consiste en la secuencia de aminoácidos de la SEQ ID NO: 10.

La actividad enzimática de una O-glucosidasa se puede evaluar mediante cualquier método adecuado, tal como los descritos en los Ejemplos. Un método adecuado puede incluir incubar una O-glucosidasa conocida o sospechosa con un sustrato estándar, tal como una o más moléculas pequeñas que comprenden colectivamente regiones centrales de O-glucanos. Tales moléculas pequeñas incluyen sustratos de 4-metilumbeliferona (4MU) y sustratos de pNP, con la liberación de pNP que indica actividad. Alternativamente, la actividad se puede evaluar utilizando una glucoproteína como sustrato. Cualquier producto de escisión resultante se puede detectar y cuantificar usando métodos rutinarios tales como SDS-PAGE o transferencias de lectina. Cuando se usa una glucoproteína como sustrato, puede ser necesario un pretratamiento (o tratamiento simultáneo) con una composición de sialidasa como se describió anteriormente. Los resultados se pueden comparar con los obtenidos en el mismo ensayo cuando el sustrato se pone en contacto con una O-glucosidasa ejemplar, tal como un polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 10. Una unidad del polipéptido de la SEQ ID NO: 10 se define como la cantidad requerida para eliminar los O-glucanos de > 90% de 1 µg de TNFaR en combinación con una unidad de una mezcla de sialidasa en tampón Tris 20 mM pH 6,8, en 2 horas a 37 °C como se monitorea por SDS-PAGE (las mezclas de sialidasa preferidas son como se describió anteriormente). Un polipéptido de prueba preferiblemente logra un nivel similar de actividad cuando está presente en la misma cantidad. Los ensayos de ejemplo también se describen en los Ejemplos.

Una composición que comprende una O-glucosidasa como se describió anteriormente se puede proporcionar en solución o en forma liofilizada para la reconstitución en solución. La O-glucosidasa se puede liofilizar en solución salina tampón con pH de 7,6.

Endoproteasa específica para O-glucano

Los presentes inventores también descubrieron que la acción de una composición de sialidasa como se describió anteriormente mejora la actividad de las endoproteasas específicas de O-glucano, en particular una endoproteasa específica para O-glucano que comprende la secuencia de aminoácidos de la SEQ ID NO: 11 que normalmente no tiene más de 375 aminoácidos, y que es preferentemente un polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 12.

Por lo tanto, la presente invención también proporciona un método para modificar una glucoproteína que comprende poner en contacto una muestra de glucoproteína tanto con una composición de sialidasa como se describió anteriormente como con una endoproteasa específica para O-glucano. La presente invención también proporciona una composición de sialidasa de la invención que opcionalmente también comprende una endoproteasa específica para O-glucano. En dicho método y dicha composición, dicha endoproteasa específica para O-glucano puede ser la de la SEQ ID NO: 12.

En cualquiera de las descripciones en la presente, las endoproteasas específicas de O-glucano de la SEQ ID NO: 11 puede reemplazarse por una variante de la misma, siempre que se conserve la actividad enzimática. Una variante de la endoproteasa puede comprender, consistir esencialmente o consistir en una variante de la secuencia de aminoácidos de la secuencia de la SEQ ID NO: 11. La variante de dicha SEQ ID NO puede definirse como se estableció anteriormente con respecto a las sialidasas, excepto que la actividad enzimática relevante a retener es la actividad hidrolítica contra O-glucoproteínas.

Alternativamente, la endoproteasa puede reemplazarse por un fragmento más corto de la SEQ ID NO: 3 o de una variante de la misma como se describió anteriormente. Los fragmentos pueden describirse como una forma truncada de dicha SEQ ID NO que retiene la actividad enzimática. Tales fragmentos son más cortos que la SEQ ID NO: 11 y tienen típicamente al menos 100, 150, 200, 250, 300, 350, 400, 450 o 500 aminoácidos de longitud.

Cualquier endoproteasa descrita en la presente puede incluir opcionalmente una metionina adicional en el extremo N y/o una histidina u otra etiqueta en el extremo C. Tales secuencias adicionales pueden ayudar con la expresión y/o purificación. Una etiqueta de histidina consiste preferentemente en seis residuos de histidina. La etiqueta de histidina se une preferentemente al extremo C mediante un enlazador, que es típicamente una secuencia corta de aminoácidos, tal como 3 - 5 aminoácidos. El enlazador típicamente consiste predominantemente en residuos de glicina y serina, y puede incluir preferentemente la secuencia GSG. Por ejemplo, GSG y GSGLE son enlazadores adecuados. Un ejemplo de endoproteasa de este tipo consiste en la secuencia de aminoácidos de la SEQ ID NO: 12.

La actividad enzimática de una endoproteasa puede evaluarse mediante cualquier ensayo adecuado. Por ejemplo, un sustrato de O-glucoproteína estándar, tal como una molécula de IgA, puede incubarse con un polipéptido de prueba. Los materiales de partida y los productos de reacción pueden analizarse mediante SDS-PAGE y/o espectrometría de masas para determinar la presencia de productos de escisión (si los hay) y, si es necesario, también para caracterizar aún más esos productos. Un sustrato de glucoproteína que no está O-glicosilado, tal como una molécula de IgG1, se puede utilizar como control negativo. Los resultados se pueden comparar con los obtenidos en el mismo ensayo cuando el sustrato se pone en contacto con un polipéptido ejemplar, tal como un polipéptido que consiste en la secuencia de aminoácidos de la SEQ ID NO: 12.

Una composición que comprende una endoproteasa como se describió anteriormente se puede proporcionar en solución o en forma liofilizada para la reconstitución en solución. La endoproteasa específica de O-O-glucano se puede liofilizar en solución salina tampón con pH de 7,6.

Métodos de uso

En la presente se describen métodos en los que una muestra se incuba con una composición de la invención en condiciones adecuadas para que las enzimas en dicha composición actúen sobre cualquier sustrato que esté presente. Dichos métodos pueden incluir opcionalmente un análisis de los productos resultantes. Dicho análisis puede incluir la separación y/o detección y/o aislamiento de los productos por cualquier medio adecuado, incluidos SDS-PAGE, HPLC, transferencia de lectina, ELISA o espectrometría de masas.

Las condiciones adecuadas incluyen la incubación con una composición de la invención durante al menos 20 minutos, 30 minutos, 40 minutos, 50 minutos, 60 minutos, 70 minutos, 80 minutos, 90 minutos o 120 minutos, 3 horas, 5 horas, 10 horas, 12 horas o durante la noche. La incubación se lleva a cabo preferentemente a temperatura ambiente, más preferentemente a aproximadamente 20 °C, 25 °C, 30 °C, 35 °C, 40 °C o 45 °C, y lo más preferentemente a aproximadamente 37 °C. Los métodos se pueden llevar a cabo bajo cualquier pH adecuado. Los valores de pH adecuados incluyen, por ejemplo, un pH de aproximadamente 3,0, 3,5, 4, 4,5, 5, 5,5, 6, 6,5, 7, 7,5, 8, 8,5, 9 o 9,5. El pH preferido está en el intervalo de 5,6 a 6,8. Los métodos se pueden llevar a cabo en cualquier tampón adecuado, tal como solución salina tamponada con tris (TBS) o solución salina tamponada con fosfato (PBS). El tampón comprende preferentemente una baja concentración de NaCl, típicamente no mayor que 300 mM, 250 mM, 200 mM o 150 mM. La concentración de NaCl es preferentemente de aproximadamente 150 mM, tal como entre 125 mM y 175 mM. La relación aproximada de las enzimas en la composición de la invención al contenido de proteína de la muestra puede ser 1:1, 2:1, 4:1, 6:1, 10:1, 15:1, 20:1, 1:2, 1:4 o 1:6, 1:10, 1:15, 1:20, 1:40, 1:50 o 1:100.

Los siguientes son métodos particularmente preferidos de la invención:

Un método para la modificación de una glucoproteína, el método comprende poner en contacto una muestra que contiene la glucoproteína con una composición de la invención y opcionalmente analizar los productos resultantes. Tal análisis puede incluir la separación y/o detección y/o aislamiento de los productos por cualquier medio adecuado, incluidos SDS-PAGE, HPLC, transferencia de lectina, ELISA o espectrometría de masas.

En una realización particular, el método puede comprender poner en contacto una muestra con una composición de la invención que comprende solo sialidasas, opcionalmente separar los productos y luego poner en contacto dichos productos con otra enzima, tal como una proteasa y/o una glucosidasa. Este método puede describirse como un "pretratamiento" de la muestra para eliminar el ácido siálico antes de añadir la otra enzima. En una variación de esta realización, la otra enzima puede añadirse a la muestra por separado pero simultáneamente con la composición de sialidasa, y puede describirse como un "tratamiento conjunto". En otra variación, la otra enzima está presente en la composición de sialidasa. La otra enzima es preferentemente una endoproteasa específica para O-glucano o una O-glucosidasa, por ejemplo, como se describe en la presente.

Los siguientes ejemplos ilustran la invención:

Ejemplo 1 - Sialidasas

Materiales y métodos

Expresión y purificación de sialidasas

Los genes (Am0705, Am0707, Am1757, Am2058) identificados en *Akkermansia muciniphila* se optimizaron por codones para expresarse bien en *E. coli* en el vector pET21a(+). El vector se transformó en células Star BL21(DE3). *E. coli* se cultivó de forma rutinaria en LB a 37 °C, 200 rpm. En presencia del plásmido, se añadieron 100 µg/mL de ampicilina. Después de la incubación durante la noche, los cultivos se diluyeron 1:20 en LB (amp) recién elaborado y se cultivaron hasta OD₆₂₀ ~ 0,7-0,8, después de lo cual se indujo la expresión de proteína recombinante mediante la adición de IPTG 1 mM, y la expresión continuó durante 5 horas antes de que las células se recolectaran y congelaran. Las células congeladas se descongelaron y se resolvieron en tampón de unión a His (NaP 20 mM pH 7,4, NaCl 500 mM, imidazol 20 mM) y se sometieron a ultrasonido para la liberación de proteínas intracelulares. Los residuos celulares se eliminaron por centrifugación. El sobrenadante filtrado estéril se purificó por afinidad en una columna de níquel y se volvió a tamponar a Tris-HCl 20 mM pH 8,0 en una columna PD-25. La concentración de las proteínas se determinó utilizando el Nanodrop, y la pureza se estimó a través de SDS-PAGE.

Evaluación de la actividad utilizando moléculas pequeñas

2-3'-sialilactosa, 2-6'-sialilactosa y ácido colomínico (2-8'; Sigma-Aldrich) se utilizaron como sustratos para determinar la especificidad hidrolítica de las enzimas. Las enzimas (0.05 µg) se mezclaron con un sustrato (25 µM) en Tris-HCl 20 mM pH 6,8, y se incubaron durante 30 minutos a 37 °C, después de lo cual se cuantificaron los ácidos siálicos libres en la mezcla, de acuerdo con las instrucciones del fabricante (Kit de cuantificación de ácido siálico, Abcam).

Evaluación de la actividad utilizando sustrato proteico

Se mezclaron TNFαR, ePO, Enbrel y fetuina (0,5 µg) junto con concentraciones variables de sialidasas probadas (1:40) o sialidasa comercial (de NEB, según el fabricante) y se incubaron durante 30 minutos, después de lo cual las proteínas se separaron en SDS-PAGE con gradiente Novex al 4-20% y/o se analizaron con transferencias de lectina SNA.

Transferencias de lectina

Las proteínas separadas se transfirieron a membranas de PVDF utilizando el sistema de transferencia Trans-Blot Turbo (BioRad). Las membranas se bloquearon con tampón de lectina y posteriormente se incubaron con el aglutinante primario (SNA-biotina) y el aglutinante secundario (HRP-estreptavidina; VectorLabs) con etapas de lavado intermedias. Chemi-luminescense fue desarrollado por el kit West Pico SuperSignal (ThermoFisher) y detectado en un ChemiDoc (BioRad).

Condiciones enzimáticas óptimas

Las sialidasas se incubaron con 2-3' sialilactosa (25 µM) en Tris-HCl 20 mM pH 8,0 para investigar el impacto del NaCl (0-1,5 M) y iones (CaCl₂ 2 mM, ZnCl₂ 2 mM, EDTA 5 mM). Para el pH óptimo, las sialidasas se incubaron en tampones de ácido acético (4,6 y 5,6) y tampones de Tris-HCl (6,8, 7,4, 8,0 y 8,8). Todas las muestras se incubaron durante 15 minutos a temperatura ambiente, después de lo cual la mezcla se añadió 1:1 a un kit de cuantificación de ácido siálico (Abcam), de acuerdo con las instrucciones del fabricante. Todos los valores se expresaron como actividad relativa en relación con la actividad más alta dentro de cada grupo.

Comparación con sialidasas conocidas de empresas biotecnológicas establecidas

Las sialidasas de *A. muciniphila*, las mezclas de sialidasas y las sialidasas compradas de marcas establecidas (NEB P0743S, P0720S y P0722S) se incubaron con 2-3' y 2-6' sialilactosa y ácido colomínico durante 20 minutos en su respectivo tampón óptimo junto con el tampón de cuantificación de ácido siálico (relación 1:1), de acuerdo con las instrucciones del fabricante.

Resultados**Todas las sialidasas de *A. muciniphila* se expresan bien y se pueden purificar en las columnas His**

Las sialidasas anotadas de *Akkermansia* Am0705, Am0707, Am1758 y Am2058 se expresaron de forma recombinante en *E. coli* y se purificaron hasta una alta pureza en las columnas His (Fig. 1). Las proteínas se expresaron bien, con niveles que varían entre 70-150 mg de enzima purificada por cultivo de *E. coli*. Además, las sialidasas eran altamente solubles en tampones Tris y podían concentrarse a > 3 mg/mL sin precipitación significativa.

Solo tres de las sialidasas muestran actividad hacia una variedad de sustratos de proteínas

Durante la selección inicial de las sialidasas purificadas hacia diferentes glucoproteínas unidas a O y N, todas las sialidasas excepto Am0705 mostraron una potente actividad (datos no mostrados). Debido a la actividad inconsistente de Am2058, se continuó la caracterización de Am0707 y una mezcla de Am0707/1757 (Mezcla, 1:1).

5 **Las sialidasas tienen diferentes especificidades**

Varias sialidasas muestran especificidad de enlace, con capacidad limitada para hidrolizar ciertos enlaces de ácido siálico (por ejemplo, 2-3, 2-6 y/o 2-8). Para investigar la capacidad de las cuatro sialidasas de *A. muciniphila* para actuar sobre diferentes enlaces, incubamos las sialidasas con sustratos específicos que solo tienen un tipo de enlaces presentes (2-3'-sialilactosa, 2-6'-sialilactosa y ácido colomínico), y cuantificamos los ácidos siálicos libres (Fig. 2). Am1757 tuvo una alta actividad específica contra 2-3 enlaces, mientras que Am0707 tuvo una actividad más amplia, aunque menor, contra todos los enlaces probados. La combinación de Am0707 y Am1757 (mezcla) dio como resultado un producto de amplio espectro superior (Fig. 2).

15 **Las sialidasas son sensibles al NaCl y dependen de cationes divalentes**

Para conocer más a fondo las condiciones necesarias para optimizar la actividad de la sialidasa, se investigó la dependencia de los iones, el pH y el NaCl para las sialidasas. Las dos sialidasas se comportaron de manera similar, con una alta sensibilidad a EDTA y Zn^{2+} , mientras que dependían de Ca^{2+} . Las enzimas tuvieron una mayor actividad en pH neutro a básico y perdieron gran parte de su actividad en presencia de NaCl (Fig. 3).

20 **Una mezcla de sialidasas aumenta la eficiencia general de la hidrólisis**

Dado que las sialidasas caracterizadas tenían actividades complementarias, con Am1757 teniendo una alta actividad hidrolítica de 2-3 enlaces, y Am0707 también actuando sobre 2-6, 8 enlaces, se investigó si una mezcla de las dos enzimas podría demostrar una alta eficiencia hacia todos los enlaces de ácidos siálicos en glucoproteínas naturales al variar la relación de las dos enzimas. Una mezcla que contenía 1:40 Am1757 junto con 1:40 Am0707 hidrolizó rápidamente (<15 min) todos los enlaces de ácido siálico en fetuina (Fig. 4).

30 **Evaluación comparativa de la sialidasa frente a sustratos moleculares pequeños**

Con el fin de comparar la eficiencia de la mezcla de sialidasa (GVS_Smix) con las enzimas individuales, así como con los competidores de la marca, se compararon las enzimas de estudio con tres sialidasas disponibles comercialmente de New England Biolabs (NEB). Estas fueron una sialidasa de amplio espectro de *Arthrobacter ureafaciens* (NEB A, la enzima α 2-3,6,8,9 Neuraminidasa A de *Arthrobacter ureafaciens*, catálogo # P0722S; escinde 2-3, 2-6, 2-8 y 2-9 enlaces), una sialidasa de espectro estrecho de *Streptococcus pneumoniae* (NEB S, la enzima α 2-3 Neuraminidasa S de *Streptococcus pneumoniae*, catálogo # P0743S; escinde solo 2-3 enlaces) y una sialidasa general de *Clostridium perfringens* (NEB O, la enzima α 2-3,6,8 Neuraminidasa O de *Clostridium perfringens*, catálogo # P0720S, escinde 2-3, 2-6, 2-8 enlaces). Se añadieron enzimas según lo sugerido por los fabricantes y se incubaron con el sustrato durante 30 minutos a 37 °C (Fig. 5).

A juzgar por una SDS-PAGE, la cantidad de sialidasa en las diferentes muestras es menor en el GVS_Smix que en el NEB, la neuraminidasa A fue posiblemente la excepción. La Neuraminidasa S mostró una capacidad limitada para hidrolizar todos los enlaces de ácido siálico en la glucoproteína, mientras que la Neuraminidasa A y la Neuraminidasa O hidrolizaron todos los ácidos siálicos presentes en la fetuina (Fig. 6). Del mismo modo, como se observó para la actividad contra sustratos sintéticos de moléculas pequeñas, Smix hidrolizó todos los enlaces de ácido siálico en la glucoproteína, generando una banda de proteína estrecha, indicativa de una asialilación completa. Las muestras tratadas con Am1757, no capaces de hidrolizar 2-6 o 2-8 enlaces, migraron como un peso molecular más alto que las muestras tratadas con Smix. La hidrólisis se confirmó con transferencia de SNA, marcando los ácidos siálicos, lo que sugiere que todas las sialidasas de amplio espectro eliminaron eficazmente todos los ácidos siálicos presentes en la fetuina, mientras que 2-3 enzimas específicas (por ejemplo, neuraminidasa S y Am1757) solo eliminaron una fracción de los ácidos siálicos (Fig. 6).

55 **Smix puede liberar eficientemente ácidos siálicos de proteínas naturales**

Aunque Smix podía actuar sobre sustratos semisintéticos pequeños con una eficiencia similar o mejor que los productos NEB (Fig. 5), era imperativo investigar también su actividad contra proteínas naturales con diferentes enlaces de ácido siálico. Como proteína modelo, se eligió la fetuina, ya que tiene 2-3 y 2-6 ácidos siálicos unidos. Usando la transferencia de SNA, quedó claro que Smix a la concentración elegida podría eliminar completamente todos los ácidos siálicos de la fetuina, de manera similar a NEB A y O (Fig. 6). Como se esperaba, las dos 2-3 enzimas específicas Am1757 y NEB S no pudieron asialilar completamente la fetuina.

Para determinar más cuantitativamente la capacidad de Smix y Am1757 para liberar ácidos siálicos de las proteínas naturales en comparación con las tres sialidasas NEB, se incubaron diferentes sustratos de glucoproteínas (TNFaR, IgA, Plasminógeno y Abatacept [Orencia]) con cada sialidasa durante 15 + 15 minutos y se cuantificaron los ácidos siálicos liberados. Si bien ciertos sustratos resultaron más difíciles de hidrolizar que otros, el GVS_Smix

fue al menos comparable a los productos NEB en todos los casos (Fig. 7) y demostró la mayor consistencia, ya que mostró una alta actividad contra todos los sustratos. Aunque las enzimas individuales de Am1757 y cada uno de los productos NEB tuvieron mayores actividades contra uno u otro sustrato individual, solo el GVS_Smix mostró una actividad consistentemente alta contra todos los sustratos, lo que lo convierte en una herramienta altamente atractiva para el análisis de glucanos.

Ejemplo 2 - O-glucosidasa

Materiales y métodos

Expresión y purificación de una endo-O-glucosidasa

La endo-N-acetil-galactosaminidasa de *Streptococcus oralis* se optimizó por codones para expresarse bien en *E. coli* en el vector pET21a(+). El vector se transformó en células Star BL21(DE3). *E. coli* se cultivó de forma rutinaria en LB a 37 °C, 200 rpm. En presencia del plásmido, se añadieron 100 µg/mL de ampicilina. Después de la incubación durante la noche, los cultivos se diluyeron 1:20 en LB (amp) recién elaborado y se cultivaron hasta OD₆₂₀ ~ 0,7-0,8, después de lo cual se indujo la expresión de proteína recombinante mediante la adición de IPTG 1 mM, y la expresión continuó durante 5 horas antes de que las células se recolectaran y congelaran. Las células congeladas se descongelaron y se resolvieron en tampón de unión a His (NaP 20 mM pH 7,4, NaCl 500 mM, imidazol 20 mM) y se sometieron a ultrasonido para la liberación de proteínas intracelulares. Los residuos celulares se eliminaron por centrifugación. El sobrenadante filtrado estéril se purificó por afinidad en una columna de níquel y se volvió a tamponar en PBS en una columna PD-25. La concentración de la proteína se determinó utilizando el Nanodrop, y la pureza se estimó a través de SDS-PAGE. La secuencia de la proteína se proporciona como SEQ ID NO: 3.

Evaluación de la actividad utilizando moléculas pequeñas

Los sustratos de 4-metilumbeliferona (4MU) y los sustratos de pNP de las regiones centrales de O-glucano se utilizaron como sustratos para determinar la actividad hidrolítica de las enzimas. Las enzimas (1 µg) se mezclaron con un sustrato (2 mM) y se incubaron durante 15-120 minutos a 37 °C, tiempo durante el cual se registró la fluorescencia y la absorbancia (405 nm), respectivamente.

Evaluación de la actividad utilizando sustrato proteico

TNFαR, ePO, Enbrel, fetuina, IgA, Orenia y plasminógeno (0,5 µg) se mezclaron con O-glucosidasa (1:40) con o sin la presencia de Smix (1:40 + 1:40) o Am1757 (1:40) durante 0-24 horas. Las proteínas se separaron en SDS-PAGE con gradiente de Novex al 4-20%.

Condiciones enzimáticas óptimas

Las enzimas se incubaron con su sustrato respectivo (4MU o pNP) para la investigación de la dependencia de NaCl (0-1,5 M) y los iones (CaCl₂ 2 mM, ZnCl₂ 2 mM, EDTA 5 mM). Para un pH óptimo, las sialidasas se incubaron en tampones de ácido acético (50 mM, pH de 4,6 y 5,6) y tampones de Tris-HCl (20 mM, pH de 6,8, 7,4, 8,0 y 8,8). Todas las muestras se incubaron durante 15 minutos a 37 °C. Todos los valores se expresaron como actividad relativa en relación con la actividad más alta dentro de cada grupo.

Comparación con enzimas disponibles comercialmente

La enzima identificada, así como las enzimas disponibles comercialmente de marcas establecidas (NEB O-glucosidasa de *Enterococcus faecalis*, catálogo# P0733S, también como un paquete # E0540S) se incubaron con sus respectivos sustratos y tampones optimizados, y se incubaron durante 0-24 horas con diferentes glucoproteínas en condiciones naturales antes de separarse en SDS-PAGE.

Resultados

Expresión/purificación de endo-α-N-acetil-galactosaminidasa de Streptococcus oralis

Durante la fase de descubrimiento, se consideraron dos endo-α-N-acetil-galactosaminidasas diferentes de *S. oralis* y *Bifidobacterium bifidum*, ambas proteínas bacterianas de alto peso molecular (>200 kDa). La glucosidasa de *B. bifidum* fue altamente inestable, o al menos dio como resultado un alto grado de partes fragmentadas después de la expresión y la purificación por afinidad. Si bien también se pudo visualizar cierta fragmentación para *S. oralis*, fue mucho menos pronunciada (Fig. 8A). Además, esta enzima fue estable a 4 °C por hasta una semana sin ninguna degradación adicional (Fig. 8B). Sin embargo, el nivel de expresión (ca 5-10 mg/L) es bastante bajo. Un análisis adicional sugirió que la fragmentación se debía al sometimiento a ultrasonido.

Las O-glucosidasas pueden actuar sobre O-glucanos Núcleo 1-3 marcados con pNP sintéticos.

El análisis continuo de las dos O-glucosidasas reveló una sorprendente preferencia por los glucanos Núcleo 1, con una actividad mucho menor hacia el núcleo 2 y 3 (Fig. 9). Es importante destacar que la glucosidasa de *S. oralis* mostró una actividad significativamente mayor que el gen correspondiente de *B. bifidum*, por lo que fue decisión centrarse principalmente en la O-glucosidasa de *S. oralis*.

La adición de $MgCl_2$ aumenta significativamente la actividad de la O-glucosidasa de *S. oralis*

Con el fin de determinar las condiciones óptimas para la O-glucosidasa, la enzima se incubó con el sustrato pNP Núcleo 1 en diferentes condiciones, a un intervalo de pH y iones (Fig. 10). La O-glucosidasa mostró una alta actividad a pH neutro a ligeramente básico, se inhibió completamente por la presencia de Zn^{2+} , pero tuvo una mayor actividad en presencia de $MgCl_2$ hasta 8 mM (Fig. 10C).

La O-glucosidasa de *S. oralis* puede hidrolizar los glucanos de las glucoproteínas naturales

Para investigar la cinética y las dosis necesarias para mediar la hidrólisis de todos los O-glucanos, se incubó TNF α R natural con cantidades variables de O-glucosidasa durante 1-12 h en combinación con la mezcla de sialidasa caracterizada en el Ejemplo 1 (mezcla 1:1 de Am1757:Am0707), o solo con mezcla de sialidasa. Como se muestra en la Figura 11, incluso concentraciones comparativamente bajas de la enzima (por ejemplo, aproximadamente 0,1 μ g, 1:10) podrían hidrolizar el sustrato en una hora. El aumento del tiempo de incubación a 12 horas dio como resultado una capacidad de la enzima para hidrolizar completamente el sustrato a una relación de 1:50 con una alta concentración de sialidasa. Es importante señalar que la enzima es, en algunos aspectos, una enzima "todo o nada", que apenas da como resultado glucoproteínas semihidrolizadas, sino más bien proteínas no hidrolizadas o completamente hidrolizadas.

Las O-glucosidasas de *S. oralis* pueden actuar sobre diferentes glucoproteínas naturales

Para investigar más a fondo si la O-glucosidasa solo actúa sobre TNF α R o puede actuar sobre varias glucoproteínas naturales, incubamos siete glucoproteínas diferentes con combinaciones de enzimas hidrolíticas (Fig. 12). La O-glucosidasa de *S. oralis*, así como la enzima de *B. bifidum* (en combinación con sialidasas) fueron capaces de hidrolizar todas las proteínas naturales después de una incubación durante la noche, a pesar de que la actividad sobre el plasminógeno fue difícil de evaluar debido al alto peso molecular de la proteína. Sin embargo, ambas enzimas dependían en gran medida de la presencia de sialidasas, ya que los ácidos siálicos terminales inhibían la actividad.

La O-glucosidasa de *S. oralis* es superior en la hidrólisis de glucoproteínas naturales en comparación con la cartera de productos NEB.

Para comparar la actividad de la glucosidasa de *S. oralis* con los productos comerciales existentes, comparamos su capacidad para hidrolizar el TNF α R utilizando cantidades variables de glucosidasas durante 1 h o 16 horas. Usando una relación enzima: sustrato de aproximadamente 1: 5, la glucosidasa de *S. oralis* fue capaz de hidrolizar completamente su sustrato en 1 h. Incluso a una relación enzima: sustrato de 1: 1, la NEB O-glucosidasa no hidrolizó todos los glucanos enlazados a O, sino que solo actuó sobre unos pocos glucanos fácilmente accesibles. Una incubación adicional (por ejemplo, 16 h) permitió una relación enzima: sustrato de 1:50 mientras se mantenía el efecto completo de la glucosidasa de *S. oralis*. Sin embargo, la O-glucosidasa NEB aún no desglucosiló completamente el producto, lo que indica que la desnaturalización es fundamental para su función, mientras que el producto de O-glucosidasa de *S. oralis* tiene una alta actividad también contra las proteínas naturales (Fig. 13). La O-glucosidasa de *S. oralis* no solo puede actuar sobre las proteínas naturales, sino que también es capaz de hidrolizar los glucanos de las proteínas desnaturalizadas (datos no mostrados).

La composición de los glucanos influye en la actividad de la O-glucosidasa

Para estudiar más a fondo la necesidad de glucanos específicos para la actividad de O-glucosidasa, se preincubó TNF α R con diferentes enzimas para eliminar los glucanos individuales antes de la adición de la O-glucosidasa. La eliminación de los ácidos siálicos terminales fue crítica para la actividad, al igual que la presencia de galactosas, lo que indica que la O-glucosidasa no puede eliminar GalNAcs individuales (terminales) (Fig. 14).

La O-glucosidasa es altamente eficiente en la eliminación de O-glucanos de las proteínas naturales

Una vez que se determinó la concentración final de la composición combinada de O-glyk + Smix (1:40), repetimos la comparación con el producto de la marca de la competencia (NEB). Si bien se puede detectar cierta hidrólisis en el TNF α R utilizando los productos NEB, es evidente que la hidrólisis no es completa, ni siquiera después de 12 h. Contrariamente a esto, la composición combinada de O-glyk + Smix dio como resultado una hidrólisis completa de la glucoproteína en 4 horas, respaldada por la transferencia de lectina (Fig. 15).

De manera similar, Enbrel mostró un patrón idéntico, con la composición combinada de O-glyk + Smix capaz de hidrolizarla completamente, mientras que el producto NEB no lo hizo. Sin embargo, para la fetuína, ambos productos fueron aparentemente igualmente eficientes (Fig. 16).

5 **Tanto los ácidos siálicos 2-3 como 2-6 unidos inhiben la actividad de la O-glucosidasa**

Con el fin de evaluar el efecto mutuo de 2-3 o 2-6 ácidos siálicos unidos para la actividad de O-glucosidasa, incubamos la O-glucosidasa con 2-3 sialidasas específicas o de amplio espectro (por ejemplo, Am1757 y Smix, respectivamente). Aunque los datos preliminares sugieren que Am1757 y Am0707 pueden liberar cantidades iguales de ácidos siálicos de las glucoproteínas, el tratamiento de la glucoproteína con Am1757 (o Smix) dio como resultado una hidrólisis más rápida del sustrato (TNFaR). Sin embargo, para la hidrólisis completa fue fundamental tratar la glucoproteína con una sialidasa de amplio espectro, para eliminar también los ácidos siálicos 2-6 (o 2-8) (Fig. 17).

15 **La actividad natural del paquete de O-glucosidasa GVS se debe a una actividad de sialidasa altamente eficiente**

Para determinar el impacto de los componentes individuales en la composición combinada de O-glyk + Smix, se incubaron O-glyk y NEB con cuatro productos de sialidasa diferentes (GVS_Smix, NEB A, NEB S y NEB O). La composición combinada de O-glyk + Smix hidrolizó de manera eficiente todos los O-glucanos de Enbrel (Fig. 18), mientras que el cambio de la sialidasa a cualquiera de los productos NEB dio como resultado una actividad significativamente menor de O-glyk, siendo la sialidasa A NEB la más potente. Sin embargo, si bien el haz de NEB O-glucosidasa (NEB O-glucosidasa con NEB sialidasa O) no pudo hidrolizar el Enbrel o TNFaR (Fig. 15, 16), al usar GVS_Smix en lugar de la NEB sialidasa O, se observó una hidrólisis completa de la glucoproteína, lo que sugiere que la eficiencia de la composición combinada de O-glyk + Smix se basa al menos en parte en la capacidad de las sialidasas para hidrolizar todos los enlaces 2-3 y 2-6 en la glucoproteína.

Ejemplo 3 - Endoproteasa específica de O-glucoproteína

Recientemente se determinó que la actividad endoproteasa del polipéptido que consiste en la SEQ ID NO: 12 al menos en parte se basó en enlaces de ácido siálico específicos, lo que requirió la eliminación de 2-3 y 2-6 ácidos siálicos unidos para un efecto completo. Para determinar el papel individual de los enlaces de ácido siálico específicos para la actividad de la endoproteasa, se incubó Etanercept con diferentes sialidasas en combinación con la endoproteasa durante 30 min - 20 h. La eliminación de 2-3 enlaces pareció suficiente para mejorar la actividad de la endoproteasa (Fig. 19). Las sialidasas utilizadas fueron (1) Am0707 = polipéptido de la SEQ ID NO: 3; (2) Am1757 = polipéptido de la SEQ ID NO: 6; Mezcla = combinación 1:1 de (1) y (2).

Secuencias

40 SEQ ID NO: 1 - sialidasa, Am0707 - tipo silvestre (secuencia señal subrayada)

MTWLLCGRGKWNKVKRMNSVFKCLMSAVCAVALPAFGQEEKTGFTDRAVTVFSAGEGNPYASIRIPALLSIGKQLLAFAEGRYKNTDQGENDIIMSVSKNGGKTWSRPAIAKAHGATFNNPCPVYDAKTRTVTVVFQRYPAKVKERQPNIPDGWDDEKCI RNFMIQSRNGGSSWTKPQEIITKTTKRPSGVDIMASGPNAGTQLKSGAHKGR LVIPMNEGPF GKWVISC IYSDDGKSWKLGQPTANMKGMVNETSIAETDNGGVVMVARHWGAGNCRRIAWSQDGETWGQVEDAPELFC DSTQNSLMTYSLSDQPAYGKSRILFSGPSAGRRIKGQVAMS YDNGKTWPVKLLGEGGFAYSSLAMVEPGIVGVLYEENQEHIKKLFVPITMEWLTDGEDTGLAPGKKAPVLK

45 SEQ ID NO: 2 - sialidasa, Am0707 - tipo silvestre sin señal

QEEKTGFTDRAVTVFSAGEGNPYASIRIPALLSIGKQLLAFAEGRYKNTDQGENDIIMSVSKNGGKTWSRPAIAKAHGATFNNPCPVYDAKTRTVTVVFQRYPAKVKERQPNIPDGWDDEKCI RNFMIQSRNGGSSWTKPQEIITKTTKRPSGVDIMASGPNAGTQLKSGAHKGR LVIPMNEGPF GKWVISC IYSDDGKSWKLGQPTANMKGMVNETSIAETDNGGVVMVARHWGAGNCRRIAWSQDGETWGQVEDAPELFC DSTQNSLMTYSLSDQPAYGKSRILFSGPSAGRRIKGQVAMS YDNGKTWPVKLLGEGGFAYSSLAMVEPGIVGVLYEENQEHIKKLFVPITMEWLTDGEDTGLAPGKKAPVLK

50 SEQ ID NO: 3 - sialidasa, Am0707 - con metionina N terminal adicional y enlazador C terminal + etiqueta His6 (en negrillas y subrayado)

MQEEKTGFTDRAVTVFSAGEGNPYASIRIPALLSIGKQLLAFAEGRYKNTDQGENDIIMSVSKNGGKTWSRPAIAKAHGATFNNPCPVYDAKTRTVTVVFQRYPAKVKERQPNIPDGWDDEKCI RNFMIQSRNGGSSWTKPQEIITKTTKRPSGVDIMASGPNAGTQLKSGAHKGR LVIPMNEGPF GKWVISC IYSDDGKSWKLGQPTANMKGMVNETSIAETDNGGVVMVARHWGAGNCRRIAWSQDGETWGQVEDAPELFC DSTQNSLMTYSLSDQPAYGKSRILFSGPSAGRRIKGQVAMS YDNGKTWPVKLLGEGGFAYSSLAMVEPGIVGVLYEENQEHIKKLFVPITMEWLTDGEDTGLAPGKKAPVLK**GSGL EHHHHH**

SEQ ID NO: 4 - sialidasa, Am1757 - tipo silvestre (secuencia señal subrayada)

MKNLLFALLTGSFCCCYAQQKAAPVPEPEVVATPPADAGRGLIRVDSREIRHYSGRKEPDYLVSRDNGKTWE
MKAAPAGYPPNYGGIPKESPAIVRNPLTREFIRVQPIGGFVFLSRGGLDGKWLAVTNDGKLEEDWKDPEKRKN
LKKLGGIMRTPVFNKGRRVIVPFHNMGGGTFKFI SDDGGLTWHVSRNGVTSRHEARPPHQGVRWFNNAVEA
TVLEMKDGTWALARTSQDQAWQAFSKDYGETWSKPEPSRFFGTLTMNTLGRLLDDGTIVSLWTNTMALPENAT
AGNGTWEDVFTNRDSHHIAMSGDEGKTWYGFREIILDEHRNHPGYATLDGPEDRGKHQSEMVQLDKNRILISL
GQHKNHRRLLVIVDRRWVGAKTATQTGKDLDSQWTIHTYIPQKKGHC SYNRPKPSAELVQDPSGGTKKVLQIKR
LDDPELVNEKSNVDYRNGGATWNFPNGTTGLVKFRFRVVDGEQADDSGLQVSLTDRLFNACDSTTKDYALFTF
PIRLKPAPHLLGMKKVPFPTGAWHEISLLWQGGQAVVSLDGKKAGTLKMANKSPNGASYIHFI STGSQPDAG
ILLDTVNARVK

5 SEQ ID NO: 5 - sialidasa, Am1757 - tipo silvestre sin señal

QQKAAPVPEPEVVATPPADAGRGLIRVDSREIRHYSGRKEPDYLVSRDNGKTWEMKAAPAGYPPNYGGIPK
SPAIVRNPLTREFIRVQPIGGFVFLSRGGLDGKWLAVTNDGKLEEDWKDPEKRKNLKKLGGIMRTPVFNKGR
RVIVPFHNMGGGTFKFI SDDGGLTWHVSRNGVTSRHEARPPHQGVRWFNNAVEATVLEMKDGTWALARTSQ
DQAWQAFSKDYGETWSKPEPSRFFGTLTMNTLGRLLDDGTIVSLWTNTMALPENATAGNGTWEDVFTNRDSHH
IAMSGDEGKTWYGFREIILDEHRNHPGYATLDGPEDRGKHQSEMVQLDKNRILISLQGHKNHRRLLVIVDRRWV
GAKTRATQTGKDLDSQWTIHTYIPQKKGHC SYNRPKPSAELVQDPSGGTKKVLQIKRLDDPELVNEKSNVDYR
NGGATWNFPNGTTGLVKFRFRVVDGEQADDSGLQVSLTDRLFNACDSTTKDYALFTFPIRLKPAPHLLGMKKVP
FPTGAWHEISLLWQGGQAVVSLDGKKAGTLKMANKSPNGASYIHFI STGSQPDAGILLDTVNARVK

10 SEQ ID NO: 6 - sialidasa, Am1757 - con metionina N terminal adicional y enlazador C terminal + etiqueta His6 (en negrillas y subrayado)

MQQKAAPVPEPEVVATPPADAGRGLIRVDSREIRHYSGRKEPDYLVSRDNGKTWEMKAAPAGYPPNYGGIPK
ESPAIVRNPLTREFIRVQPIGGFVFLSRGGLDGKWLAVTNDGKLEEDWKDPEKRKNLKKLGGIMRTPVFNKGR
RRVIVPFHNMGGGTFKFI SDDGGLTWHVSRNGVTSRHEARPPHQGVRWFNNAVEATVLEMKDGTWALARTSQ
DQAWQAFSKDYGETWSKPEPSRFFGTLTMNTLGRLLDDGTIVSLWTNTMALPENATAGNGTWEDVFTNRDSHH
IAMSGDEGKTWYGFREIILDEHRNHPGYATLDGPEDRGKHQSEMVQLDKNRILISLQGHKNHRRLLVIVDRRWV
GAKTRATQTGKDLDSQWTIHTYIPQKKGHC SYNRPKPSAELVQDPSGGTKKVLQIKRLDDPELVNEKSNVDYRN
GGATWNFPNGTTGLVKFRFRVVDGEQADDSGLQVSLTDRLFNACDSTTKDYALFTFPIRLKPAPHLLGMKKVP
PFTGAWHEISLLWQGGQAVVSLDGKKAGTLKMANKSPNGASYIHFI STGSQPDAGILLDTVNARVK**GSGLEH**
HHHHH

15 SEQ ID NO: 7 - O-glucosidasa de *S. oralis*

Tipo silvestre (secuencia de señal subrayada; elemento C terminal del motivo de anclaje de la pared celular LPXTG subrayado en negrillas)

MDKRFFEKRCKFSIRKFTLGVASVMIGATFFAASPVLAQARVGSTDNLPSELADLDKKASDEGHDFDKEAAA
QNPQSAETTEGPQTEEEELLAQEKEKSEKPSNLPKELEDKLEKAEDNGREVQDQLAQDTGKLVPEVAKTTNG
ELNYGATVKIKTPSGEGSGIIVAKDLVLTVSHNFIKDSQEGNIRKVVNDNDQGDGDIYSISYPGLPDVKFSKDD
IIHWDREGYLKGFKNDLALVRLRTVLENTPEVTKKPVVKIGDKLHVFGYPEGKLNPIVNTTVDFAEYPYEG
VQIGYQGGKPGASGGGIFDTEGKLVGVHONGVVGKRSKGILFSPAQLKWIQDHMQGISSVKPADLEEKEKPA
EEKPKEDKPAAPKPTPKAVTPEWQTVANKEQQGTVTIREEKGVRYNQLSSTAQNNDNGKPALFEKQGLTVDA
NGNATVDLTFKDDSEKGSRFVFLKFKDTKNNVFGYDQGGWFWEYKTPGNSTWYKGNRVAAPPGSVNRLS
ITLKSQGLNANNDVNLFDVTPLPQAVNENLKNKKILLKAGTYSNDRVTVSVKTDNQEGVKADDTPAQKET
GPAVDDSKVTYDTIQSKVLKAVIDQAFPRVKEYTLNGHTLPGQVQVFNQVFINNHRTPEVTYKINETTAEY
LMKLRDDAHLINAEMTVRLQVVDNQLHFDVTKIVNHNQVTPGQKIDDERKLLSTISFLGNALVSVDQAGAK
FDGATMSNNTHVSGDDHIDVTNPMKDLAKGYMYGFVSTDKLAAGVWSNSQNSYGGGSNDWTRLTAYKETVGNA
NYVGIHSSWEQWEKAYKGI VFPYTKELPSAKVITEDANADNKVDWQDGAIAYSIMNPNQGWKVKDITAY
RIAMNFGSQAQNPFLMTLDGKIKINLHTDGLGQGVLLKGYGSEGHDSGHLNYADIGKRIGGVDFKTLIEKAK
KYGAHLGIHVNASETYPEKYFNENILRKNPDGSYSYGNWLDQGINIDAAYDLAHLARWEDLKKKLGEGL
DFIYVDVWNGQSGDNGAWATHVLAKEINKQGRFAIEWGHGGEYDSTFQHWAADLTGYGYTNKGINSAITRF
IRNHQKDSWVGDIYSYGAANYPLGGYSMKDFEGWQGRSDYNGYVTNLFADHVMTKYFQHFVTSKENGTPV
TMTDNGSTYKWTPEMKVELVDAAGNKVVVTRKSNVDNSPQYRERTVTNLGRVQDGSAYLTPWNWDANGKKLP
TEKEKMYFNTQAGATTWTLPSDWANSKYVLYKLTQDGKTEEQELTVTDGKITLDLLANQPYVLYRSKQTNPE
MSWSEGMIHYDQGFNSGTLKHWTISGDASKAEIVKSQGANEMLRIQGNKSKVSLTQKLTGLKPNTKYAVYVGV
DNRNNAKASITVNTGEKEVTYTNKSLALNYIKAYAHNNRRENATVDDTSYFQNMAYFTTGSQSDVSNVTLLS

20

REAGDEATYFDEIRTFENNSSMYGDKHDTGQGTFFKQDFENVAQGIFFPVVGGVEGVEDNRTHLSEKHDPYTQR
 GWNGKKVDDVIEGNWSLKTNGLVSRRLVYQTI PQNFRFEAGKTYRVTFFEYAGSDNTYAFVVGKGEFQSGRR
 GTQASNLEMHELPNTWTDSKKAKKVTFLVTGAETGDTWVGIIYSTGNASNTRGDAGGNANFRGYNDFMMDNLQI
 EEITLTGKMLTENALKNYLPTVAMTNYTKESMDALKEAVFNLSQADDDISVEEARAEIAKIEALKNALVQKKT
 ALVAEDFESLDAPAQPGEGLNADFNGVSSLWHTSWNGGDVGKPATMVLKEPTEITGLRYVPRASDSNGNLRD
 VKLVVTDSESGEHTFNVTDPNNNNPKDIDFGKTIKAKKIVLTGKTYGDDGDKYQSAELIFTRPQVAETPL
 DLSCGYEALAKAQKLTDKDNQEEVASVQASMKYATDNHLLTERMVAYFADYLNQLKDSATKPDAPTSSKGEEQ
 PPVLDVPEFKGGVNATEAAVHEVPEFKGGVNAVQALVHELPEYKGGANAVLAAANEVPEYKGGANAVEALVNE
 KPAYTGVLATAGDQAAPTVEKPEYPLTPSPVADTKTPGAKDEEKLPATGEHSSEVALFLASVSIALSAAVLAT
KRKEE

SEQ ID NO: 8 - O-glucosidasa de *S. oralis*

- 5 Tipo silvestre con secuencia de señal eliminada (elemento C terminal del motivo de anclaje de la pared celular LPXTG subrayado en negrillas)

DQARVGSTDNLPSELADLDKKASDEGHDFDKEAAAQNPGSAETTEGPQTEEELLAQEKEKSEKPSNLPKELED
 KLEKAEDNGREVDKQDLAQDTCKLVPEDVAKTTNGELNYGATVKIKTPSGEGSGIVVAKDLVLTVSHNFIKDS
 QEGNIRKVVDDNDQGDGDIYSISYPGLPDVKFSKKDIHWDREGYLGFKNDLALVRLRTVLENTPVEVTKKPV
 VKKIGDKLHVFGYPEGKLNPIVNTTVDFAEPYGEGVQGIQYGGKPGASGGGIFDTEGKLVGVHQNQVVGKRS
 GGILFSPAQLKWIQDHMQGIISSVKPADLEEKEKPAEEKPKEDKPAAAKPETPKAVTPEWQTVANKEQQGTVTI
 REEKGVRYNQLSSTAQNDNDGKPALEKQGLTVDANGNATVDLTFKDDSEKKSFRFGVFLKFKDTKNNVFGY
 DQGGWFWEYKTPGNSTWYKGNRVAAPPEGSVNRLSITLKSQGLNASNNDVNLFDVTPLGAVNENLKNKKI
 LLKAGTYSNDRTVSVKTDNQEGLVADDTPAQKETGPAVDDSKVTYDTIQSKVLKAVIDQAFPRVKEYTLNGH
 TLPQGVQQFNQVFINNHRTPEVTYKKINETTAEYLMKLRRDDAHLINAEMTVRLQVVDNLHFDVTIKVNHQ
 VTPGQKIDDERKLLSTISFLGNALVSVSSDQAGAKFDGATMSNNTHVSGDDHIDVTNPMKDLAKGYMYGFVST
 DKLAAGVWSNSQNSYGGGSNDWTRLTAYKETVGNANYVGIHSSSEWQWEKAYKGVFPPEYTKELPSAKVVITED
 ANADNKVDWQDGAIAYSIMNPNQGWKVKDITAYRIAMNFGSQAQNPFLMTLDGIKKINLHTDGLGQGVLLK
 GYGSEGHDSGHLNYADIGKRIGGVEDFKTLIEKAKKYGAHLGIHVNASETYPESKYFNENILRKNPDGSYSYG
 WNWLDQGINIDAAAYDLAHGRLARWEDLKKKLGEGLDIYVDVWNGNQSGDNGAWATHVLAKEINKQGRFAIE
 WGHGGEYDSTFQHWAADLTYGGYTNGKINSATIRFIRNHQKDSWVGDRSYGGAANYPLGGYSMKDFEGWQG
 RSDYNGVVTNLFAHDVMTKYQHFTVSKWENGTPVMTDNGSTYKWTPEMKVELVDAAGNKVVVTRKSNVDNS
 PQYRERTVTLNQRVIQDGSAYLTPWNWDANGKKLPTEKEKMYFNTQAGATTWTLPSDWANSKVLYKLTQDQ
 KTEEQELTVTDGKITLDDLQNPVLYRSKQTNPEMSWSEGMHIYDQGFNSGTLKHWITSGDASKAEIVKSQG
 ANEMLRIQGNKSKVSLTQKLTGLKPNKYAVYVGVNDRSNKASITVNTGEKEVTYTNKSLALNYIKAYAHN
 NRRENATVDDTSYFQNMAYFTTGSDVSNVTLTSLREAGDEATYFDEIRTFENNSSMYGDKHDTGQGTFFKQDF
 ENVAQGIFFPVVGGVEGVEDNRTHLSEKHDPYTQRGWNGKKVDDVIEGNWSLKTNGLVSRRLVYQTI PQNFR
 FEAGKTYRVTFFEYAGSDNTYAFVVGKGEFQSGRRGTQASNLEMHELPNTWTDSKKAKKVTFLVTGAETGDTW
 VGIYSTGNASNTRGDAGGNANFRGYNDFMMDNLQIEEITLTGKMLTENALKNYLPTVAMTNYTKESMDALKEA
 VFNLSQADDDISVEEARAEIAKIEALKNALVQKKTALVAEDFESLDAPAQPGEGLNADFNGVSSLWHTSWNG
 GDVGKPATMVLKEPTEITGLRYVPRASDSNGNLRDVKLVVTDSESGEHTFNVTDPNNNNPKDIDFGKTIKAK
 KIVLTGKTYGDDGDKYQSAELIFTRPQVAETPLDLSCGYEALAKAQKLTDKDNQEEVASVQASMKYATDNH
 LLTERMVAYFADYLNQLKDSATKPDAPTSSKGEEQPPVLDVPEFKGGVNATEAAVHEVPEFKGGVNAVQALVH
 ELPEYKGGANAVLAAANEVPEYKGGANAVEALVNEKPAYTGVLATAGDQAAPTVEKPEYPLTPSPVADTKTPG
 AKDEEKLPATGEHSSEVALFLASVSIALSAAVLATKRKEE

- 10 SEQ ID NO: 9 - O-glucosidasa de *S. oralis*
- Tipo silvestre con secuencia de señal y elemento C terminal del motivo de anclaje de pared celular LPXTG, ambos eliminados

DQARVGSTDNLPSELADLDKKASDEGHDFDKEAAAQNPGSAETTEGPQTEEELLAQEKEKSEKPSNLPKELED
 KLEKAEDNGREVDKQDLAQDTCKLVPEDVAKTTNGELNYGATVKIKTPSGEGSGIVVAKDLVLTVSHNFIKDS
 QEGNIRKVVDDNDQGDGDIYSISYPGLPDVKFSKKDIHWDREGYLGFKNDLALVRLRTVLENTPVEVTKKPV
 VKKIGDKLHVFGYPEGKLNPIVNTTVDFAEPYGEGVQGIQYGGKPGASGGGIFDTEGKLVGVHQNQVVGKRS
 GGILFSPAQLKWIQDHMQGIISSVKPADLEEKEKPAEEKPKEDKPAAAKPETPKAVTPEWQTVANKEQQGTVTI
 REEKGVRYNQLSSTAQNDNDGKPALEKQGLTVDANGNATVDLTFKDDSEKKSFRFGVFLKFKDTKNNVFGY
 DQGGWFWEYKTPGNSTWYKGNRVAAPPEGSVNRLSITLKSQGLNASNNDVNLFDVTPLGAVNENLKNKKI
 LLKAGTYSNDRTVSVKTDNQEGLVADDTPAQKETGPAVDDSKVTYDTIQSKVLKAVIDQAFPRVKEYTLNGH
 TLPQGVQQFNQVFINNHRTPEVTYKKINETTAEYLMKLRRDDAHLINAEMTVRLQVVDNLHFDVTIKVNHQ

15

VTPGQKIDDERKLLSTISFLGNALVSVSSDQAGAKFDGATMSNNTHVSGDDHIDVTNPMKDLAKGYMYGFVST
DKLAAGVWSNSQNSYGGGSNDWTRLTAYKETVGNANVYVGIHSSSEWQWEKAYKGVFPEYTKELPSAKVVITED
ANADNKVDWQDGAIAYSIMNPNQGWKVDITAYRIAMNFGSQAQNPFLMTLDGIIKINLHTDGLGQGVLLK
GYGSEGHDSGHLNYADIGKRIGGVDFKTLIEKAKKYGAHLGIHVNASETYESPKEYFNENILRKNPDGSSYSY
WNWLDQGINIDAAAYDLAHLARWEDLKKLGEGLDFIYVDVWNGQSGDNGAWATHVLAKINKQWRFAIE
WGHGGEYDSTFQHWAADLTYYGYTNKGINSATIRFIRNHQKDSWVGDIRSYGGAANYPLLGGYSMKDFEGWQ
RSDYNGYVTNLFAHDVMTKYFQHFTVSKWENGTPVTMTDNGSTYKWTPEMKVELVDAAGNKVVVTRKSNVDNS
PQYRERTVTNLGRVIQDGSAYLTPWNWDANGKKLPTEKEKMYFYNTQAGATTWTLPSDWANSKVLYKLTQDQ
KTEEQELTVDGKITLDDLANQPYVLYRSKQTNPEMSWSEGMHIYDQGFNSGTLKHWITISGDASKAEIVKSQ
ANEMLRIOGNKSKVSLTQKLTGLKPNTKYAVYVGVNDRSNKASITVNTGEKEVTTYTNKSLALNYIKAYAHN
NRRENATVDDTSYFQNMAYAFFTTGSDVSNVTLLTSREAGDEATYFDEIRTFENNSSMYGDKHDTGQGTFFKQD
ENVAQGIFFPVVGGVEGVEDNRTHLSEKHDPYTQRGWNGKKVDDVIEGNWSLKTNGLVSRRLVYQTIPONFR
FEAGKTYRVTFFEYAGSDNTYAFVVGKGEFQSGRRGTQASNLEMHELPTNTWTDSSKAKKVTFLVTGAETGDTW
VGIYSTGNASNTRGDAGGNANFRGYNDFMMDNLQIEEITLTGKMLTENALKNYLPTVAMTNYTKESMDALKEA
VFNLSQLADDDISVEEARAEIAKIEALKNALVQKKTALVAEDFESLDAPAQPGEGLNFAFGNVSSSLWHTSWNG
GDVGKPMATMVLKEPTEITGLRYVPRASDSNGNLRDVKLVTDESKEHTFNVTDPNNNPKDIDFGKTIKAK
KIVLTGTKTYGDDGDKYQSAELIFTRPQVAETPLDLSGYEALAKAQKLTDKDNQEEVASVQASMKYATDNH
LLTERMVAYFADYLNQLKDSATKPDAPTSKSGEEQPPVLDVPEFKGGVNATEAAVHEVPEFKGGVNAVQALVH
ELPEYKGGANAVLAAANEVPEYKGGANAVEALVNEKPAYTGVLATAGDQAAPTVEKPEYPLTPSPVADTKTPG
AKDEEKLPA

SEQ ID NO: 10 - O-glucosidasa de *S. oralis* con Met N terminal adicional, GSGLE-His6tag C terminal (subrayado en negrillas) y secuencia de señal y elemento C terminal del motivo de anclaje de pared celular LPXTG eliminado

MDQARVGSTDNLPSELADLDKKASDEGHDFDKEAAAQNPESAETTEGPQTEEEELLAQEKEKSEKPSNLKPELE
DKLEKAEDNGREVDDKQLAQDTGKLVPEVAKTTNGELNYGATVKIKTPSGEGSGIVVAKDLVLTVSHNFIDK
SQEGNIRKVVNDNDQGDGDIYSISYPGLPDVKFSKKDIHWDREGYLKGFKNLALVRLRVLNTPEVETKKP
VVKKIGDKLHVFCYPEGKLNPIVNTTVDFAEPYGEGVQIGYQGGKPGASGGGIFDTEGKLVGVHONGVVGKR
SGGILFSPAQLKWIQDHMQGSSVKPADLEEKEKPAEEKPKEDKPAAKPKAVTPEWQTVANKEQQGTVT
IREEKGVRYNQLSSSTAQNNDNGKPALFEKQGLTVDANGNATVDLTFKDDSEKKSFRGFVFLFKKDTKNVVFV
YDQGGWFWEYKTPGNSTWYKGNRVAAPEPGSVNRLSITLKSQGLNASNNDVNLFDTVTLPGAVNENLKNK
ILLKAGTYSNDRTVSVKTDNQGEGVKAADDTPAQKETGPAVDDSKVYDITIQSKVLKAVIDQAFPRVKYTLNG
HTLPGQVQGFQNFVFINNHRIPEVYKKNINETTAEYLMKLRDDAHLINAEMTVRLQVVDNQLHFDVTKIVNHN
QVTPGQKIDDERKLLSTISFLGNALVSVSSDQAGAKFDGATMSNNTHVSGDDHIDVTNPMKDLAKGYMYGFVS
TDKLAAGVWSNSQNSYGGGSNDWTRLTAYKETVGNANVYVGIHSSSEWQWEKAYKGVFPEYTKELPSAKVVITE
DANADNKVDWQDGAIAYSIMNPNQGWKVDITAYRIAMNFGSQAQNPFLMTLDGIIKINLHTDGLGQGVLL
KGYGSEGHDSGHLNYADIGKRIGGVDFKTLIEKAKKYGAHLGIHVNASETYESPKEYFNENILRKNPDGSSYSY
GWNWLDQGINIDAAAYDLAHLARWEDLKKLGEGLDFIYVDVWNGQSGDNGAWATHVLAKINKQWRFAIE
EWGHGGEYDSTFQHWAADLTYYGYTNKGINSATIRFIRNHQKDSWVGDIRSYGGAANYPLLGGYSMKDFEGWQ
GRSDYNGYVTNLFAHDVMTKYFQHFTVSKWENGTPVTMTDNGSTYKWTPEMKVELVDAAGNKVVVTRKSNVDN
SPQYRERTVTNLGRVIQDGSAYLTPWNWDANGKKLPTEKEKMYFYNTQAGATTWTLPSDWANSKVLYKLTQDQ
GKTEEQELTVDGKITLDDLANQPYVLYRSKQTNPEMSWSEGMHIYDQGFNSGTLKHWITISGDASKAEIVKSQ
GANEMLRIOGNKSKVSLTQKLTGLKPNTKYAVYVGVNDRSNKASITVNTGEKEVTTYTNKSLALNYIKAYAHN
NRRENATVDDTSYFQNMAYAFFTTGSDVSNVTLLTSREAGDEATYFDEIRTFENNSSMYGDKHDTGQGTFFKQD
FENVAQGIFFPVVGGVEGVEDNRTHLSEKHDPYTQRGWNGKKVDDVIEGNWSLKTNGLVSRRLVYQTIPONFR
RFEAGKTYRVTFFEYAGSDNTYAFVVGKGEFQSGRRGTQASNLEMHELPTNTWTDSSKAKKVTFLVTGAETGDT
WVGIYSTGNASNTRGDAGGNANFRGYNDFMMDNLQIEEITLTGKMLTENALKNYLPTVAMTNYTKESMDALKE
AVFNLSQLADDDISVEEARAEIAKIEALKNALVQKKTALVAEDFESLDAPAQPGEGLNFAFGNVSSSLWHTSWN
GGDVGKPMATMVLKEPTEITGLRYVPRASDSNGNLRDVKLVTDESKEHTFNVTDPNNNPKDIDFGKTIKAK
KIVLTGTKTYGDDGDKYQSAELIFTRPQVAETPLDLSGYEALAKAQKLTDKDNQEEVASVQASMKYATDN
HLLTERMVAYFADYLNQLKDSATKPDAPTSKSGEEQPPVLDVPEFKGGVNATEAAVHEVPEFKGGVNAVQALV
HELPEYKGGANAVLAAANEVPEYKGGANAVEALVNEKPAYTGVLATAGDQAAPTVEKPEYPLTPSPVADTKTP
GAKDEEKLPA**GSGLEHHHHH**

SEQ ID NO: 11 - Endoproteasa específica de O-glucoproteína

EVTVPDALKDRIALKKTARQLNIVYFLGSDTEPVPDYERRLSELLLYLQQFYGKEMQRHGYGARSFGLDIKSP
GRVNIIEYKAKNPAAHPYENGCGWKAQELDEFFKAHPDRKKSQHTLIIMPTWNDEKNGPDNPGGVFPFYGMG
RNCFALDYPAFDIKHLGQKTREGRLTKWYGGMAHELGHGLNLPNNHQTASDGKKYGTALMGSGNYTFGTSP
FLTPASCALLDACEVFSVTPSQQFYEGKPEVEVGDAISFKGDQILVSGNYKSPQTVKALNVYIQDPYAVNQ
DYDAVSFSRRLGKKSFKSMKIDKKELEGLNNEFRISLMFILLANGLHMQKHFTFHWDALQDYRDGSKS

SEQ ID NO: 12 - Endoproteasa específica de O-glucoproteína (LS)

MEVTVPDALKDRIALKKTARQLNIVYFLGSDTEPVPDYERRLSELLLYLQQFYGKEMQRHGYGARSEFLDIKS
PGRVNIIEYKAKNPAAHYPYENGSGWKAAQELDEFFKAHPDRKKSQHTLIIMPTWNDEKNGPDNPGGVPPFYGM
GRNCFALDYPAFDIKHLGQKTREGRLLTKWYGGMAHELGHGLNLP3NHQTA5DGKKYGTALMGSGNYTFGTSP
TFLTPASCALLDACEVFSVTPSQQFYEGKPEVEVGDAISFKGDQILVSGNYKSPQTVKALNVYIQDPPYAVN
QDYDAVFSRRLGKKSFKFSMKIDKKELEGLNNNEFRISLMFILANGLHMQKHFTTFHWDALQDYRDGSKSGSG
HHHHHH

SEQ ID NO: 13 - codifica a la SEQ ID NO: 10

ATGGACCAAGCGCGTGTGGGTAGCACCGATAACCTGCCGAGCGAGCTGGCGGATCTGGACAAGAAAGCGAGCG
ACGAAGGCCACGATTTTGACAAAGAGGCGGCGCGCAGAACCCGGGTAGCGCGGAAACCACCGAAGGTCCGCA
GACCGAGGAAGAGCTGCTGGCGCAAGAAAAAGAGAAGAGCGAGAAGCCGAGCAACCTGCCGAAAGAACTGGAG
GATAAAGTGGAAAAAGGCGGAGGACAACGCTGCTGAAGTGGATAAAGACCAGCTGGCGCAAGACACCGGCAAGC
TGGTGCCGAGGATGTTGCGAAAACCAACCGGTGAAGTGAAGTACGGCGCGACCGTTAAAAATTAAGACCCC
GAGCGGCGAGGGTAGCGGTATTGTGGTTGCGAAGGACCTGGTGCTGACCGTTAGCCACAACCTTCATTAAGGAT
AGCCAGGAAGTAATATCCGTAAAGTGGTTGATAACGACCAAGGCGATGGTGACATCTACAGCATTAGCTATC
CGGGCCTGCCGACGTTAAGTTCAGCAAGAAAGATATCATCCACTGGGACCGTGAGGGTTACCTGAAAGGCTT
CAAGAACGATCTGGCGCTGGTGCGTCTGCGTACCGTTCTGGAAAACACCCCGGTTGAGGTGACCAAGAAACCG
GTGGTTAAGAAAATTTGGTGACAAGCTGCACGTGTTTGGTTATCCGGAGGGCAAACTGAACCCGATCGTGAACA
CCACCGTTGATTTGCGGAACCGTACGGCGAGGGTGTTTCAAGGCATTGGTTATCAAGGTGGCAAAACCGGGCGC
GAGCGGTGGCGGTATCTTTGACACCGAAGGCAAGCTGGTTGGCGTGCACCAAGACCGGTGGTTGGCAAAACGT
AGCGGCGGTATTCTGTTTACGCGCGGCGCAACTGAAGTGGATTACGAGCACATGCAAGGTATCAGCAGCGTGA
AACCGGCGGATCTGGAAGAGAAAGAGAAGCCGCGGAAGAGAAACCGAAGGAAGACAAGCCGGCGGCGGCGGAA
GCCGGAACCCCGAAGCGGTACCCCGGAGTGGCAACCGTGGCGAACAAGGAACAGCAAGGTACCGTTACC
ATCCGTGAAGAGAAAGCGTTCGCTTACAACCGAGTGAAGCAGCACCGCGCAAAACGATAACGACGGCAAGCCGG
CGCTGTTTGAGAAACAGGGTCTGACCGTTGACGCGAACCGCAACCGGACCGTGGATCTGACCTTCAAGGACGA
TAGCGAAAAAGGCAAGAGCCGTTTTCGGCGTTTTCTGAAAATTCAAGGACACCAAAACACGTTTTTGTGGGT
TACGATCAAGGCGGTGGTTCTGGGAGTATAAGACCCCGGTAACAGCACCTGGTACAAGGGTAACCGTGTGG
CGGCGCGGAACCGGTTAGCGTGAACCGTCTGAGCATTACCTGAAAAGCGACGGCCAGCTGAACGCGAGCAA
CAACGATGTGAACCTGTTTCGACACCGTTACCTGCGGGTGCGGTGAACGAAAACCTGAAGAACGAGAAGAAA
ATCCTGCTGAAAGCGGGCACCTACAGCAACGACCGTACCGTGGTTAGCGTTAAGACCGATAACAGGAAGGTG
TGAAAGCGGACGATACCCCGCGCAAAAGGAAACCGGTCCGGCGGTGGACGATAGCAAGGTACCTACGACAC
CATTACAGAGCAAAGTGCTGAAGGCGGTATCGATCAAGCGTTTCCGCGTGTGAAAGAGTATACCTGAAACGGT
CACACCTGCCGGGTACAGTTACGCAATTTAACCAAGTGTTCATTAACAACACCGTATCACCCCGGAAGTGA
CCTATAAGAAAATTAACGAAACACCGCGGAGTACCTGATGAAGCTGCGTGACGATGCGCACCTGATCAACGC
GGAATGACCGTGCGTCTGACAGTGGTTGATAACCAACTGCACCTTCGACGTGACCAAAATTTGTTAACCACAAC
CAGGTTACCCCGGTTCAAAAGATTGACGATGAGCGTAAACTGCTGAGCACCATCAGCTTCTGGGCAACGCGC
TGGTTAGCGTGAGCAGCGATCAAGCGGGTGCGAAGTTGATGGTGCGACCATGAGCAACAACACCCACGTTAG
CGGTGACGATCACATCGATGTGACCAACCCGATGAAAGACCTGGCGAAGGGTTACATGTATGGCTTTGTTAGC
ACCGACAAGCTGGCGGCGGGTGTGTGGAGCAACAGCCAAAACAGCTACGGCGGTGGCAGCAACGATTGGACCC
GTCTGACCGCGTATAAAGAAACCGTTGGTAACGCGAACTACGTGGGCATTACAGCAGCGAATGGCAGTGGGA
GAAAGCGTACAAGGGTATCGTGTTCCCGGAATATACCAAGGAGCTGCCGAGCGGAAAGTGGTTATACCCGAG
GATGCGAACGCGGACAAACAAAGTGGATTGGCAGGACGGTGCGATTGCGTACCGTAGCATCATGAACAACCCG
AAGGCTGGGAAAAAGTTAAGGACATTACCGCGTATCGTATCGCGATGAACCTTGGTAGCCAGGCGCAAAACCC
GTTCCGTGATGACCTGGACGGCATCAAGAAAATTAACCTGCACACCGATGGCCTGGGTACGGCGGTTCTGCTG
AAGGTTATGGTAGCGAGGGTCATGACAGCGGTACCTGAACTACGCGGATATCGGTAACGTTATGGTGGCG
TGGAAGACTTTAAGACCTGATTGAGAAAGCGAAGAAATACGGTGCGCACCTGGGCATCCACGTTAACCGGAG
CGAAACCTACCCGGAGAGCAAGTATTTCAACGAAAACATTTCTGCGTAAAAACCCGGACGGTAGCTACAGCTAT
GGCTGGAAGTGGCTGGATCAGGGTATCAACATTGATGCGGCGTACGACCTGGCGCATGGCCGTCTGGCGCGTT
GGGAGCACCTGAAGAAAACCTGGGTGAAGGCCTGGATTTTATCTATGTTGACGTGTGGGGTAACGGTCAGAG
CGGTGATAACGGTGCCTGGGCGACCATGTGCTGGCGAAAGAGATTAAACAAGCAAGGTTGGCGTTTTGCGATC
GAATGGGGCCACGGTGGCGAGTACGACAGCACCTTCCAGCACTGGGCGGCGGATCTGACCTACGGTGGCTATA

CCAACAAGGGTATCAACAGCGCGATTACCCGTTTCATCCGTAACCACCAGAAAAGATAGCTGGGTTGGCGACTA
 CCGTAGCTATGGTGGCGCGGCGAACTACCCGCTGCTGGGTGGCTATAGCATGAAGGACTTTGAGGGTTGGCAA
 GGCCGTAGCGATTACAACGGTTATGTTACCAACCTGTTCGCGCACGACGTGATGACCAAGTACTTTACGACACT
 TCACCGTTAGCAAATGGGAAAACGGTACCCCGGTGACCATGACCGATAACGGCAGCACCTATAAGTGGACCCC
 GGAAATGAAAGTGGAGCTGGTTGACGCGCGGGTAACAAGGTGGTTGTGACCCGTAAAAGCAACGATGTGAAC
 AGCCCGCAGTACCGTGAGCGTACCGTTACCTGAACGGTCTGTGATCCAAGACGGCAGCGCGTATCTGACCC
 CGTGGAACCTGGGATGCGAACGCTAAAAAGCTGCCGACCGAAAAAGAGAAGATGTACTATTTTAACACCCAAGC
 GGGTGCGACACCTGGACCCCTGCCGAGCGACTGGGCGAACAGCAAGGTTTACCTGTATAAACTGACCGATCAG
 GGCAAGACCGAGGAGCAAGAACTGACCGTGACCGATGGCAAAATTACCTGGACCTGCTGGCGAACCCAGCCGT
 ACGTTCTGTATCGTAGCAAGCAAAACCAACCCGGAATGAGCTGGAGCGAGGGTATGCACATCTACGACCAAGG
 TTTCACAGCGGCACCCCTGAAACACTGGACCATTAGCGGCGATGCGAGCAAGGCGGAGATCGTGAAAAGCCAG
 GGTGCGAACGAAATGCTGCGTATCCAAGGCAACAAAAGCAAGGTTAGCCTGACCCAGAAGCTGACCGGTCTGA
 AACCGAACACCAAGTACGCGTTTATGTGGGCGTTGACAACCGTAGCAACGCGAAAGCGAGCATTACCGTTAA
 CACCGGTGAAAAAGAGGTGACCACCTACACCAACAAGAGCCTGGCGCTGAACTACATCAAAGCGTATGCGCAC
 AACACCGCTCGTGAGAACGCGACCGTGGACGATACAGCTACTTCCAGAACATGTATGCGTCTTTTACCACCG
 GTAGCGAGTGAAGCAACGTTACCCCTGACCTGAGCCGTGAAGCGGGCGATGAGGCGACCTATTTTGACGAAAT
 TCGTACCTTCGAGAACAACAGCAGCATGTACGGTGATAAGCACGACACCGGTCAGGGCACCTTTAAACAAGAT
 TTCGAAAACGTTGCGCAAGGTATCTTCCCGTTTGTGTGGGTGGCGTGGAAGGCGTTGAGGACAACCGTACCC
 ACCTGAGCGAGAAGCACGATCCGTACACCCAGCGTGGTTGGAACGGCAAAAAGGTGGACGATGTTATTGAGGG
 TAACTGGAGCCTGAAAACCAACGGCCTGGTTAGCCGTGTAACCTGGTGTACCAGACCATCCCGCAAAACTTC
 CGTTTTGAGGCGGGCAAGACCTACCGTGTGACCTTTGAATATGAGGCGGGCAGCGACACACCTATGCGTTTG
 TTGTGGGTAAAGGCGAATTCCAGAGCGGTGCTGCTGGCACCCAAAGCGAGCAACCTGGAATGCACGAGCTGCC
 GAACACCTGGACCGATAGCAAAAAGGCGAAAAAGGTGACCTTCTGTTTACCGGTGCGGAAACCGGTGACACC
 TGGGTGGGTATCTACAGCACCGGCAACGCGAGCAACACCCGTGGTGATGCGGGTGGCAACGCGAAGCTTTCGTG
 GCTATAACGATTTTCATGATGGACAACCTGCAAAATCGAAGAGATTACCTGACCGGCAAGATGCTGACCGAAAA
 CGCGCTGAAAAACTATCTGCCGACCGTTGCGATGACCAACTACACCAAGGAAAGCATGGACGCGCTGAAAGAG
 GCGGTTTTCAACCTGAGCCAGGCGGACGATGACATCAGCGTGGAAGAGGCGCGTGCGGAAATCGCGAAGATTG
 AGGCGCTGAAAAACGCGCTGGTTTCAGAAAAAGACCGCGCTGGTTGCGGAAGATTTTGAGAGCCTGGATGCGCC
 GCGCAACCGGGTGAAGGCCCTGGAGAACCGTTTCGACGGTAACGTTAGCAGCCTGTGGCACACCAGCTGGAAC
 GGTGGCGATGTTGGCAAGCCGCGGACCATGGTGCTGAAAGAACCAGCGAGATCACCGGTCTGCGTTATGTGC
 CGCGTGCGAGCGATAGCAACGGCAACCTGCGTGACGTTAAGCTGGTTGTGACCGATGAAAGCGGTAAAGAGCA
 CACCTTTAACGTGACCGACTGGCCGAACAACAACCAACCGAAGGATATTGACTTCGGCAAAACCATTAAGGCG
 AAAAAGATCGTTCTGACCGGTACCAAGACCTACGGCGATGGTGGCGACAAATATCAGAGCGCGGCGGAGCTGA
 TCTTTACCCGTCCGCAAGTGGCGGAAACCCGCTGGATCTGAGCGGTTACGAAGCGGCGCTGGCGAAAGCGCA
 GAAGCTGACCGATAAGGACAACCAGGAAGAGGTGGCGAGCGTTCAAGCGAGCATGAAATATGCGACCGACAAC
 CACCTGCTGACCGAACGTATGGTTGCGTACTTCGCGGATTATCTGAACCAACTGAAGGATAGCGCGACCAAC
 CGGATGCGCCGACCAGCAGCAAGGGTGAAGAACAGCCGCGCGGTGCTGGATGTTCCGGAGTTTAAAGGTGGCGT
 GAACGCGACCGAGGCGGCGGTGCACGAAGTTCCGGAGTTCAAGGGTGGCGTGAACGCGGTTACAGGCGCTGGTT
 CACGAACCTGCCGAGTATAAAGGTGGCGGAACGCGGTTCTGGCGGCGGCAACGAAGTGCCGAGTACAAGG
 GTGGCGGAACGCGGTGGAAGCGCTGGTTAACGAGAAACCGGCGTATACCGGTGTTCTGGCGACCGCGGCGCA
 CCAGGCGGCGCGACCGGTGGA AAAACCGGAGTACCCGCTGACCCGAGCCCGGTTGCGGACACCAAAACCCCG
 GGTGCGAAAGATGAAGAGAAGCTGCCGCGGGTAGCGCCTCGAGCACCACCACCACCACTGA

REIVINDICACIONES

1. Una composición que comprende una primera y una segunda sialidasa,
 5 en donde la primera sialidasa es un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 2; y
 en donde la segunda sialidasa es un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 5; y
 en donde la composición además comprende una O-glucosidasa que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 9,
 10 opcionalmente en donde dicha primera y/o segunda sialidasa incluye una metionina adicional en el extremo N y/o una etiqueta de histidina en el extremo C, cuya etiqueta puede unirse al extremo C mediante un enlazador.
2. La composición de acuerdo con la reivindicación 1, que es capaz de hidrolizar >90% de los enlaces siálicos en una glucoproteína, preferentemente una glucoproteína en su estado natural, no desnaturalizado.
- 15 3. La composición de acuerdo con la reivindicación 1 o 2, que comprende la sialidasa que consiste en la secuencia de aminoácidos de la SEQ ID NO: 3 y la sialidasa que consiste en la secuencia de aminoácidos de la SEQ ID NO: 6, preferentemente en una relación 1:1.
- 20 4. La composición de acuerdo con cualquiera de las reivindicaciones anteriores, en donde la primera sialidasa y/o la segunda sialidasa están presentes en forma altamente purificada o aislada.
5. La composición de acuerdo con cualquiera de las reivindicaciones anteriores, en donde no más de dos polipéptidos en la composición son sialidasas obtenidas de *Akkermansia muciniphila*.
- 25 6. La composición de acuerdo con cualquiera de las reivindicaciones anteriores, que comprende adicionalmente una endoproteasa específica de O-glucoproteína, opcionalmente en donde la O-glucosidasa y/o endoproteasa específica de O-glucoproteína está presente en forma altamente purificada o aislada.
- 30 7. La composición de acuerdo con cualquiera de las reivindicaciones anteriores, en donde dicha O-glucosidasa consiste en la secuencia de aminoácidos de la SEQ ID NO: 10.
8. La composición de acuerdo con la reivindicación 6 o 7, en donde dicha endoproteasa específica de O-glucoproteína es una endoproteasa específica de O-glucoproteína seleccionada de:
 35 a. un polipéptido que comprende o consiste en una secuencia de aminoácidos de la SEQ ID NO: 11;
 b. un polipéptido que comprende o consiste en una secuencia de aminoácidos que es al menos 85% idéntica a la secuencia de aminoácidos de la SEQ ID NO: 11 o
 c. un polipéptido que comprende o consiste en una secuencia de aminoácidos que es un fragmento de la secuencia de la SEQ ID NO: 11 o un fragmento de una secuencia que es 85% idéntica a la secuencia de aminoácidos de la
 40 SEQ ID NO: 11;
 opcionalmente en donde dicha endoproteasa específica de O-glucoproteína consiste en la secuencia de aminoácidos de la SEQ ID NO: 12.
9. Un método para modificar una glucoproteína que comprende poner en contacto una muestra que contiene la glucoproteína con una composición de acuerdo con cualquiera de las reivindicaciones anteriores, opcionalmente en donde se analizan los productos resultantes.
- 45 10. El método de acuerdo con la reivindicación 9, en donde los productos resultantes se analizan, y en donde dicho análisis incluye la separación y/o detección y/o aislamiento de los productos por cualquier medio adecuado, incluyendo SDS-PAGE, HPLC, transferencia de lectina, ELISA o espectrometría de masas.
- 50 11. Un método de acuerdo con la reivindicación 9 o 10, en donde la composición es como se define en cualquiera de las reivindicaciones 1 a 4, y en donde el método comprende poner en contacto dicha muestra con una endoproteasa específica de O-glucoproteína simultáneamente, antes o después de que la glucoproteína se ponga en contacto con dicha composición.
- 55 12. Un método de acuerdo con la reivindicación 9, en donde la composición comprende adicionalmente una endoproteasa específica de O-glucoproteína, en donde dicha endoproteasa específica de O-glucoproteína es opcionalmente como se define en la reivindicación 8.

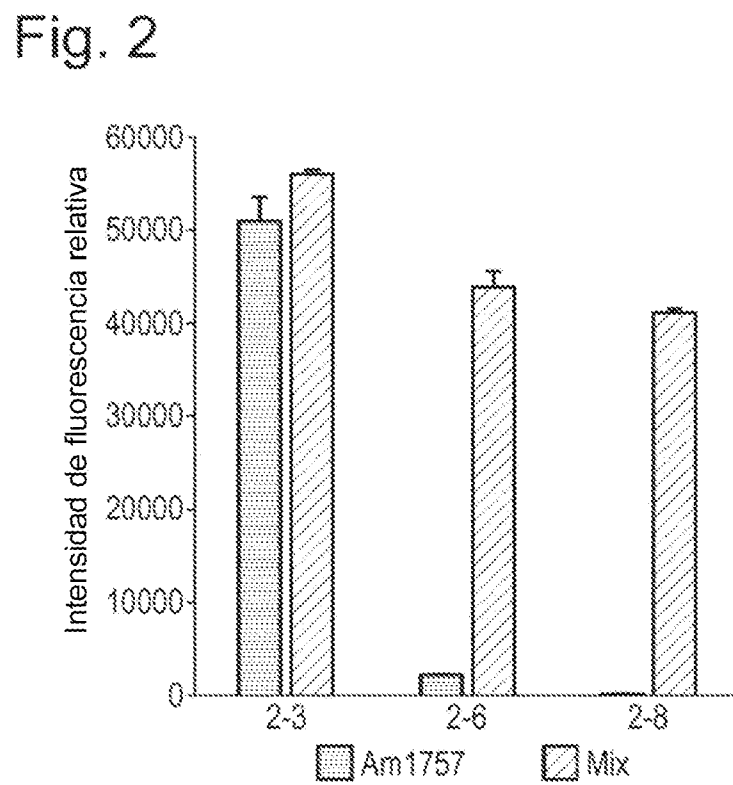
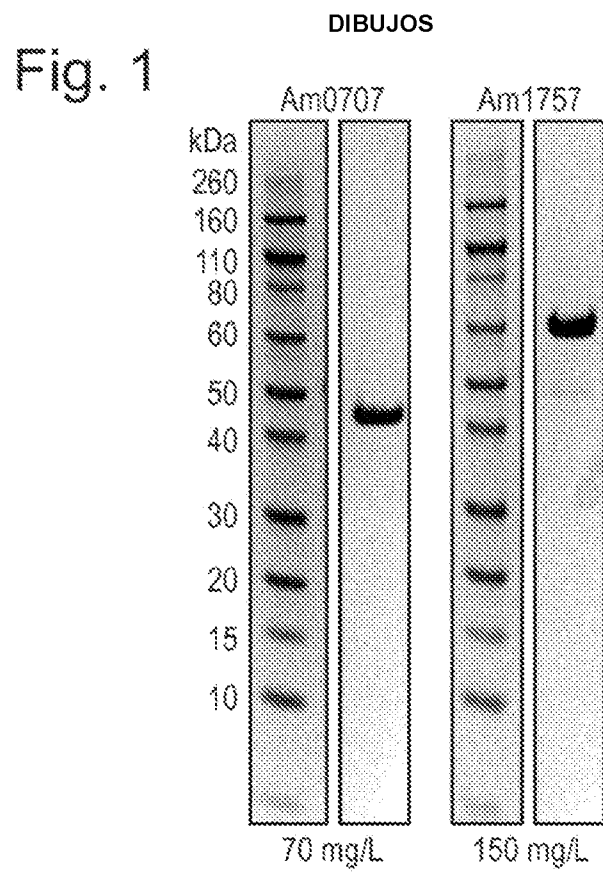


Fig. 3

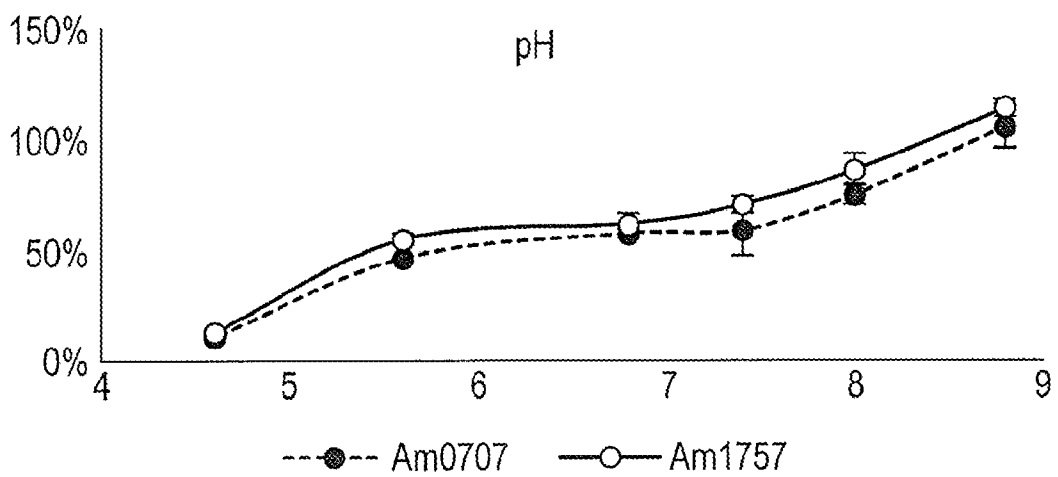
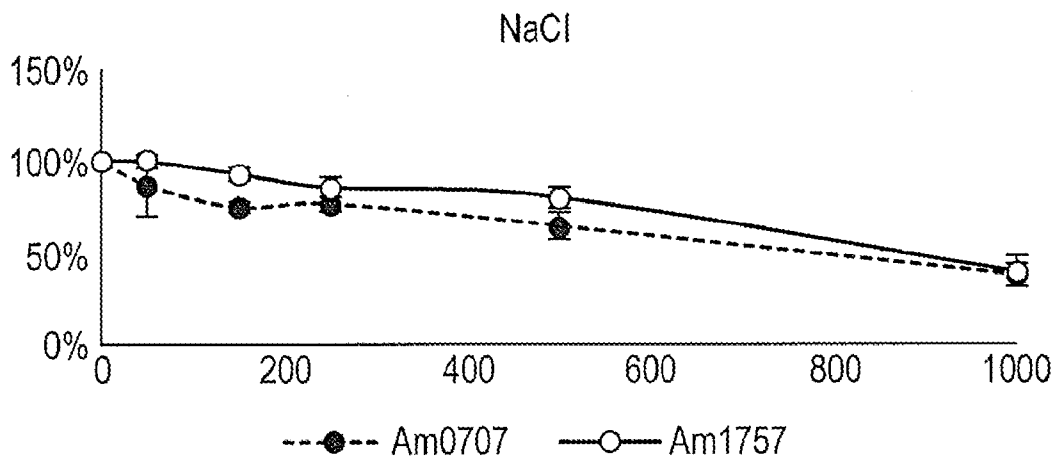
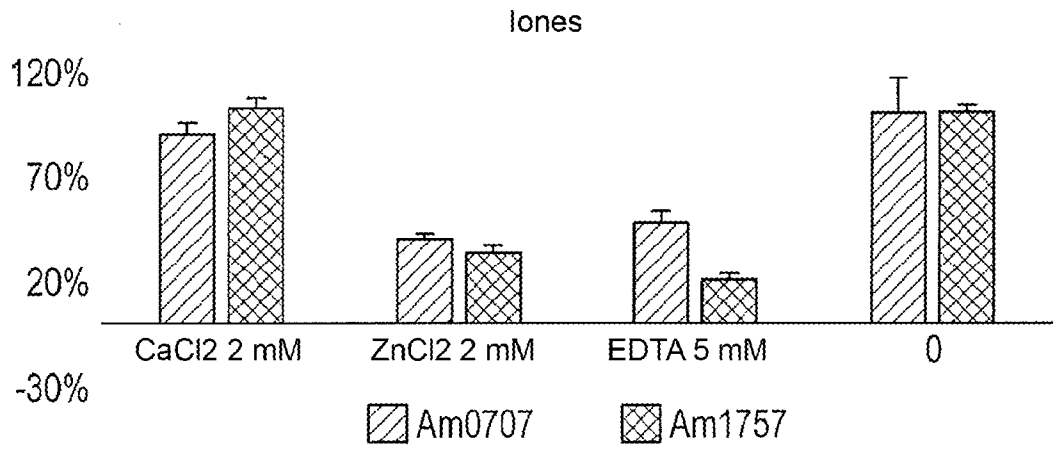


Fig. 4

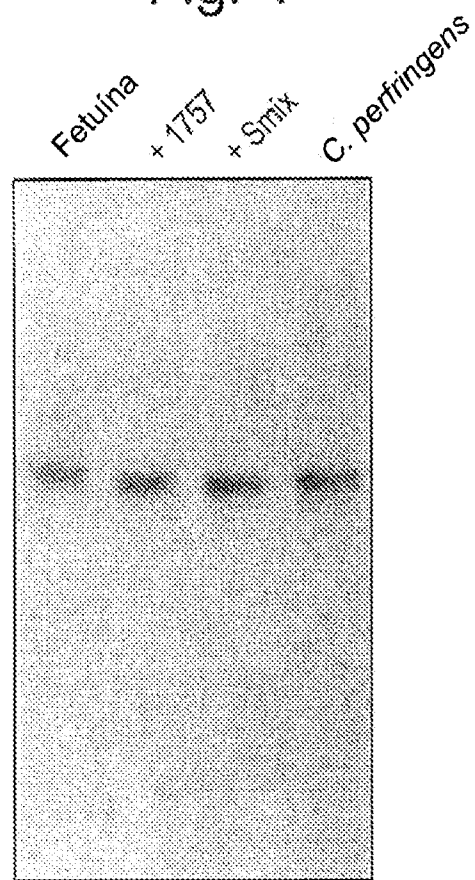


Fig. 5

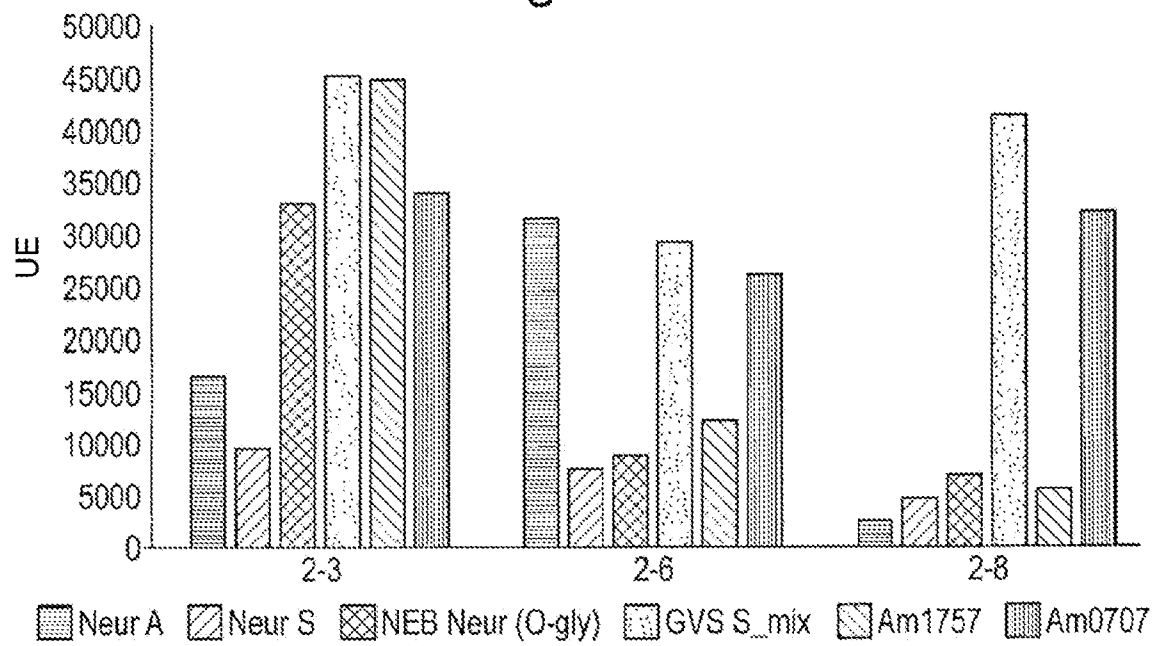


Fig. 6

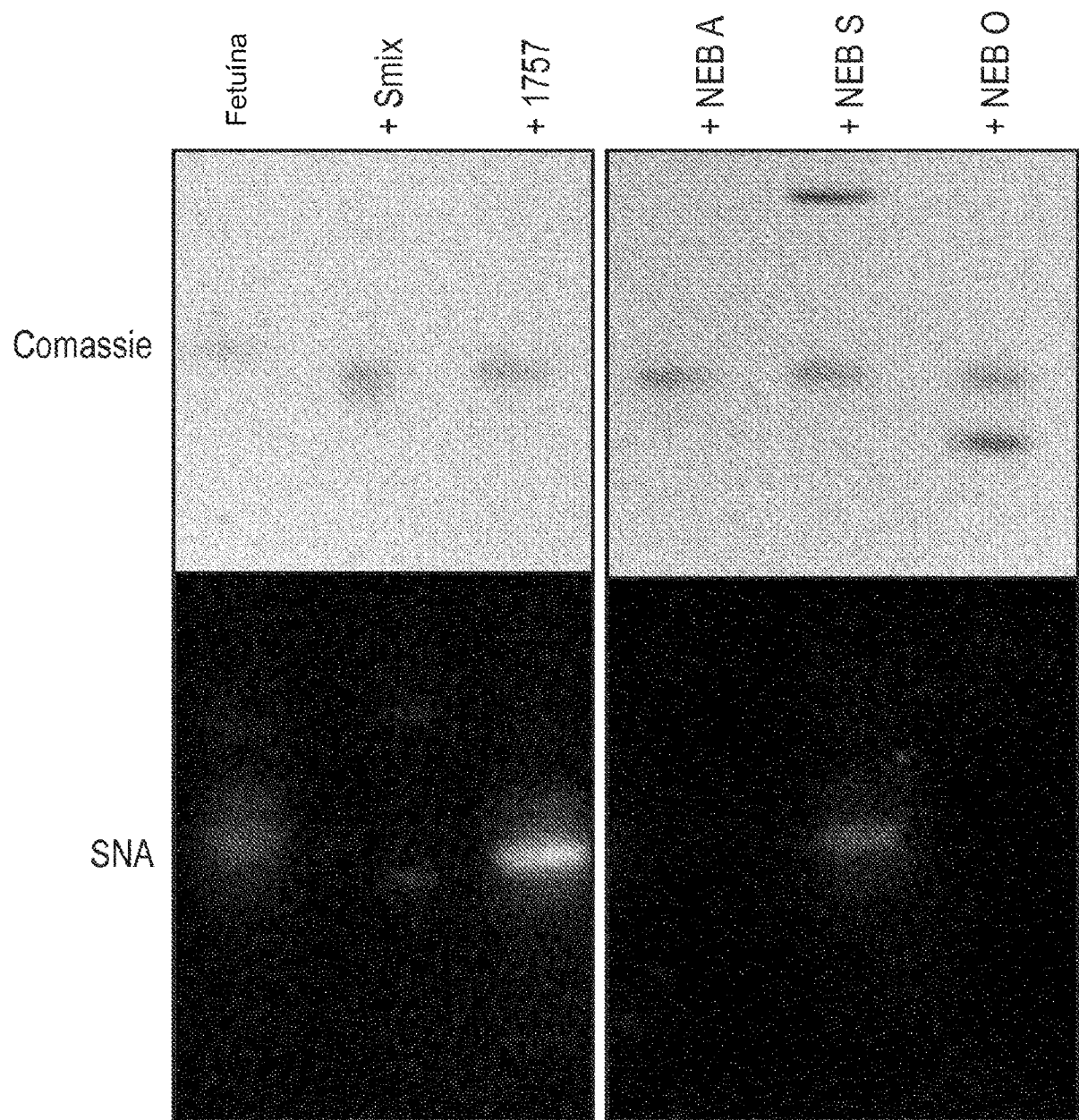


Fig. 7

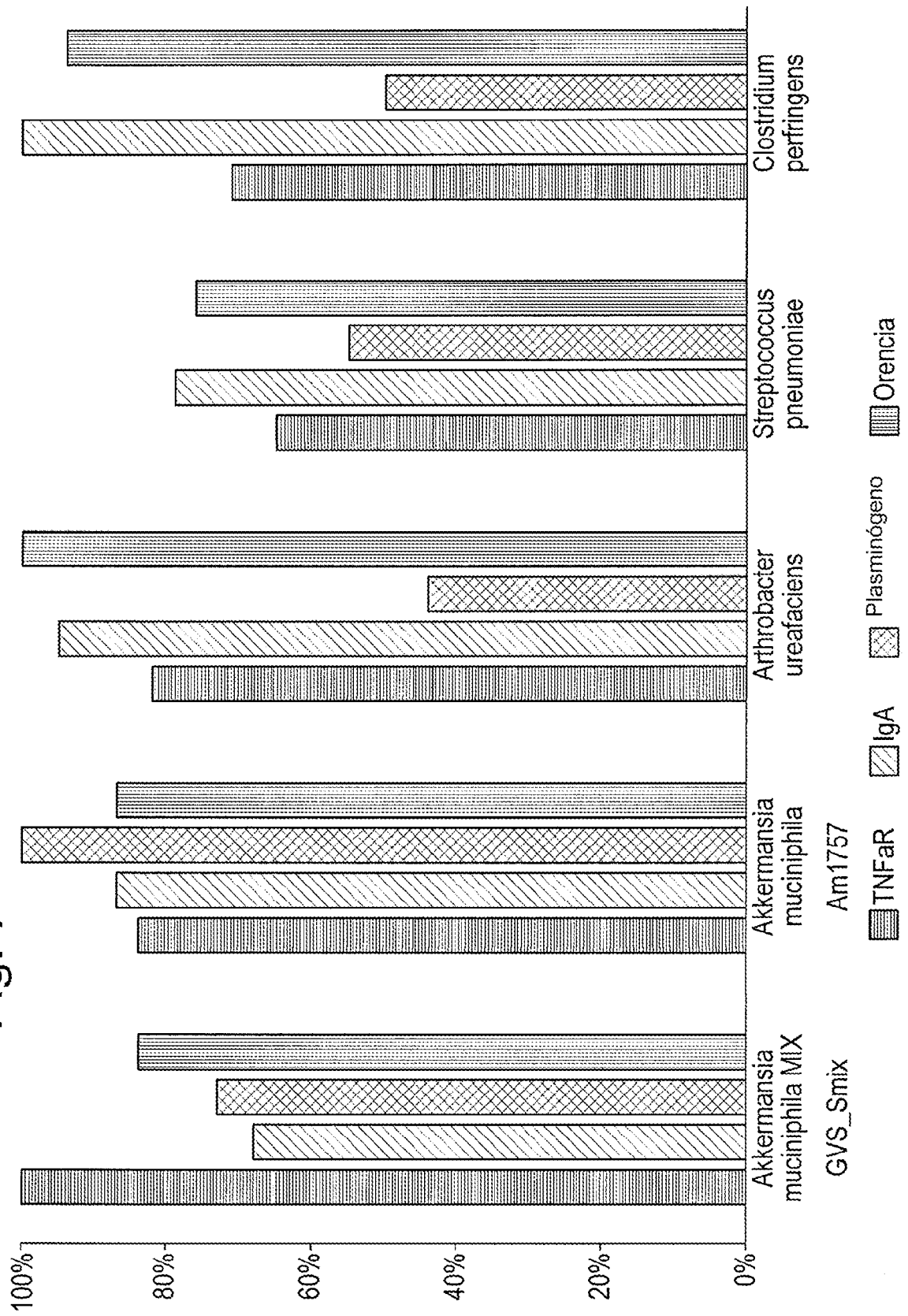


Fig. 8A

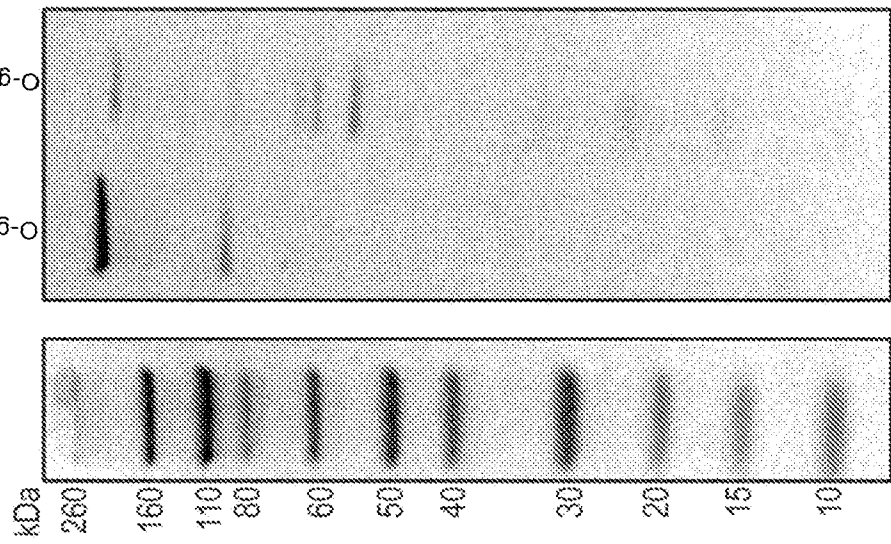


Fig. 8B

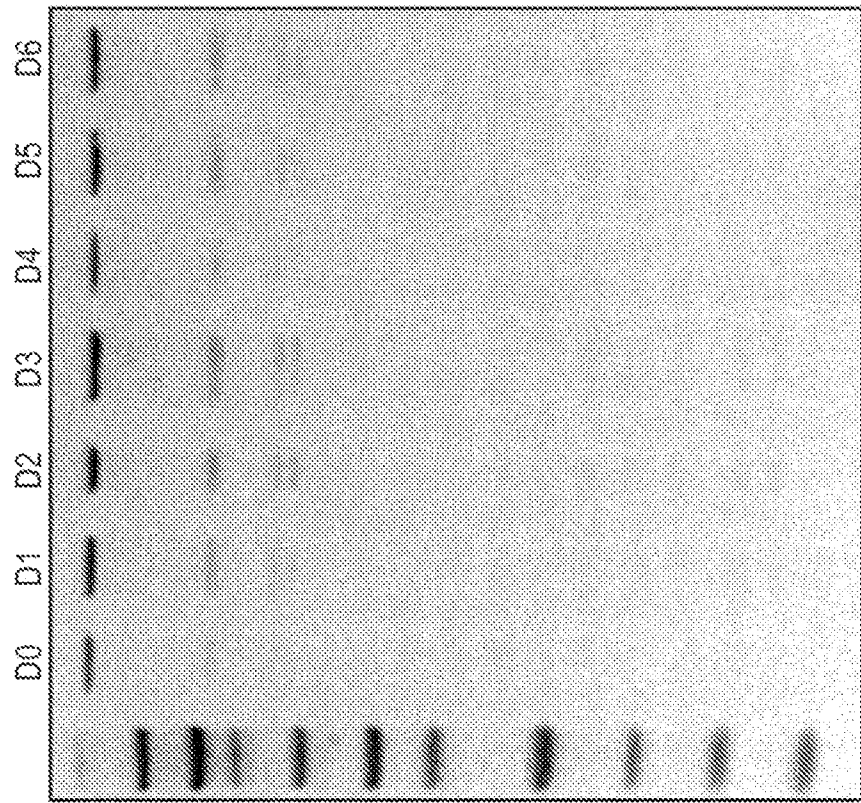


Fig. 9A

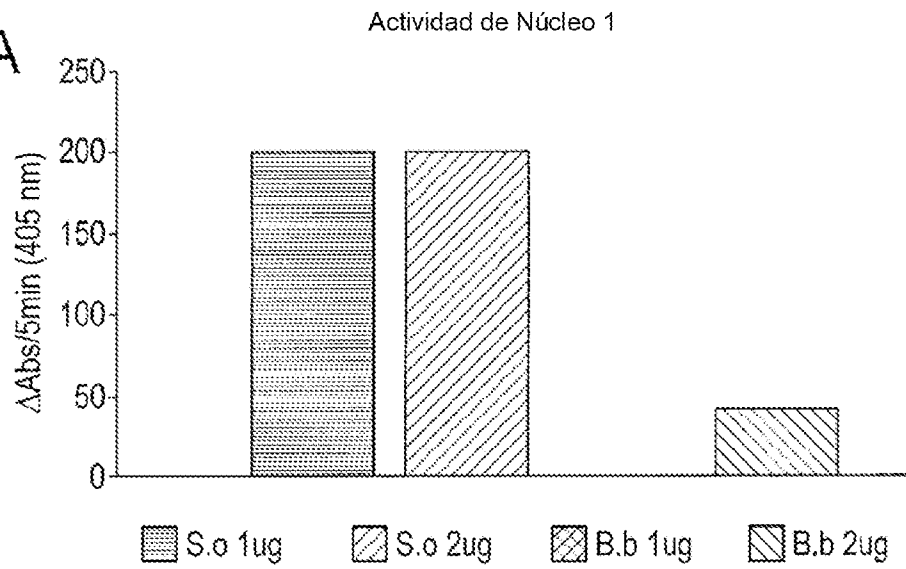


Fig. 9B

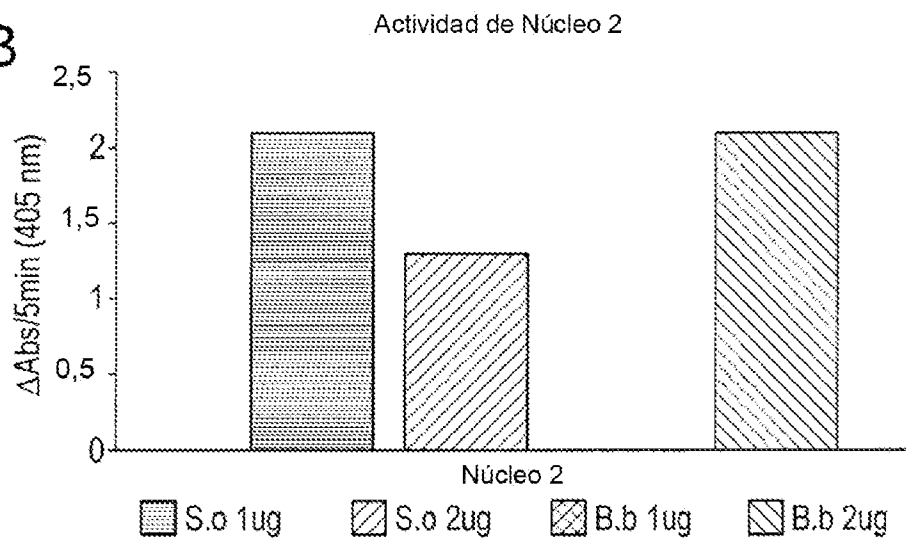


Fig. 9C

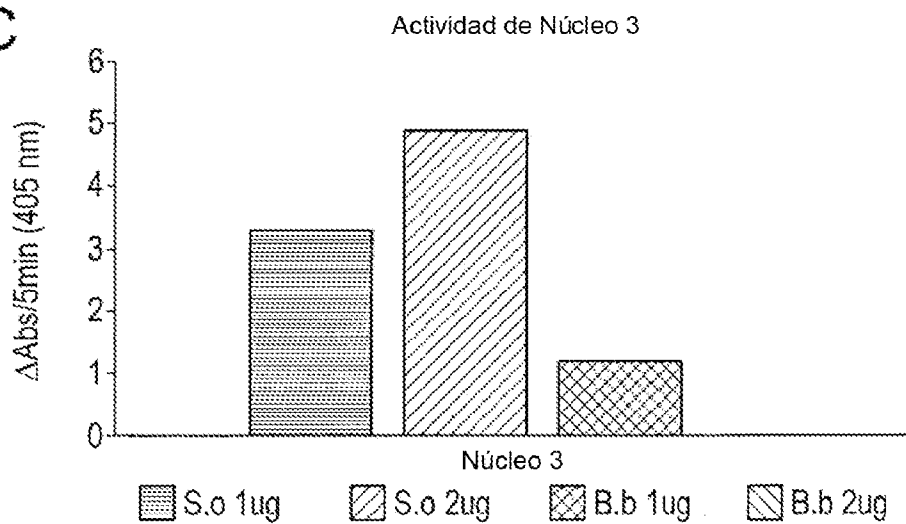


Fig. 10A

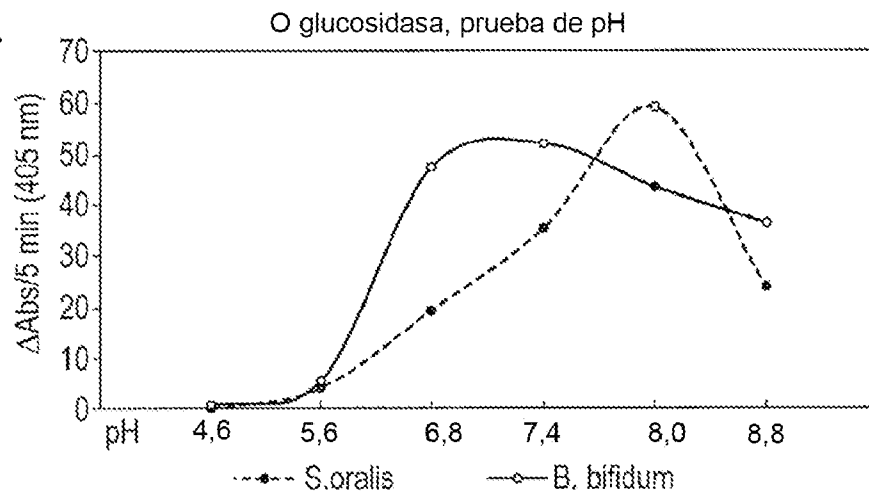


Fig. 10B

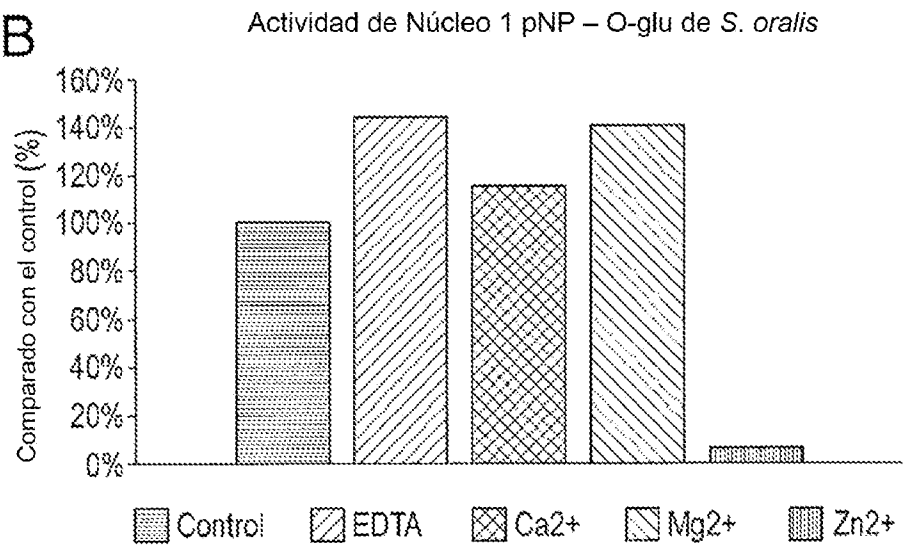


Fig. 10C

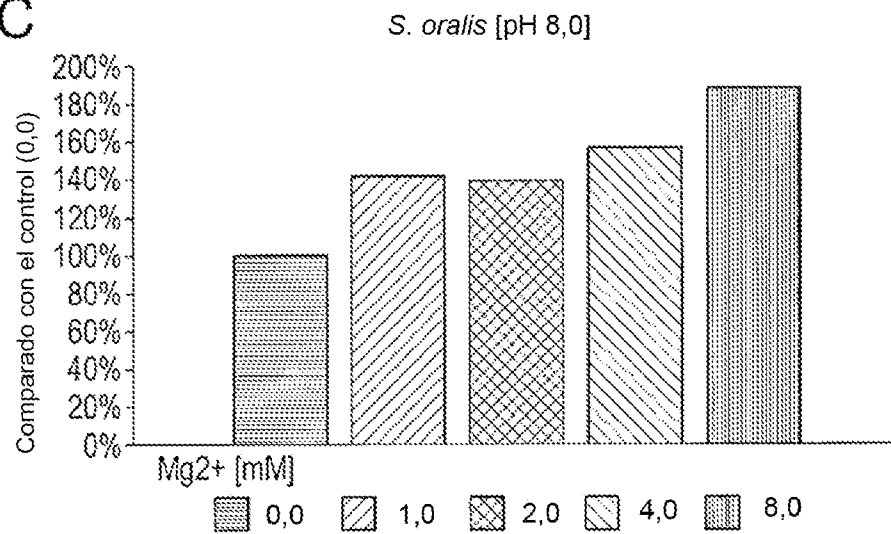


Fig. 11

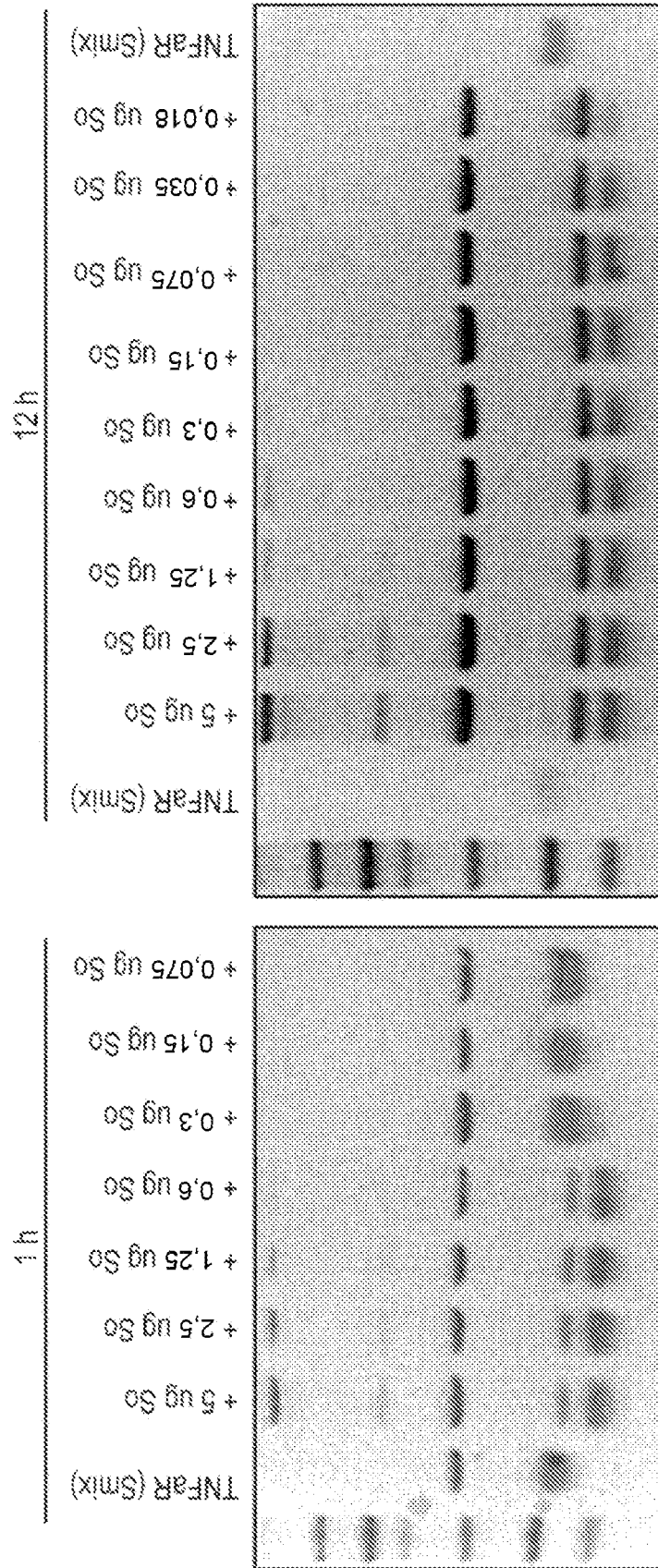


Fig. 12B

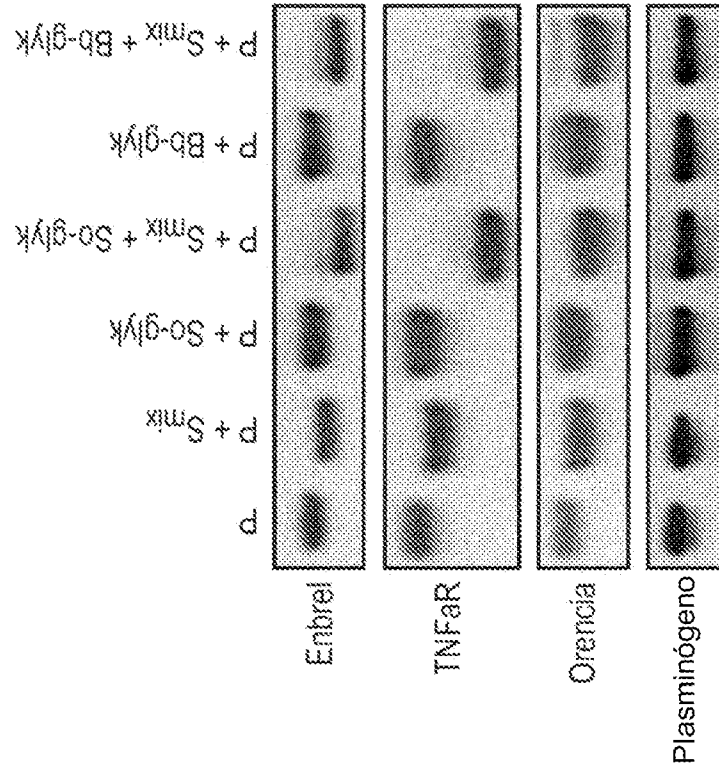


Fig. 12A

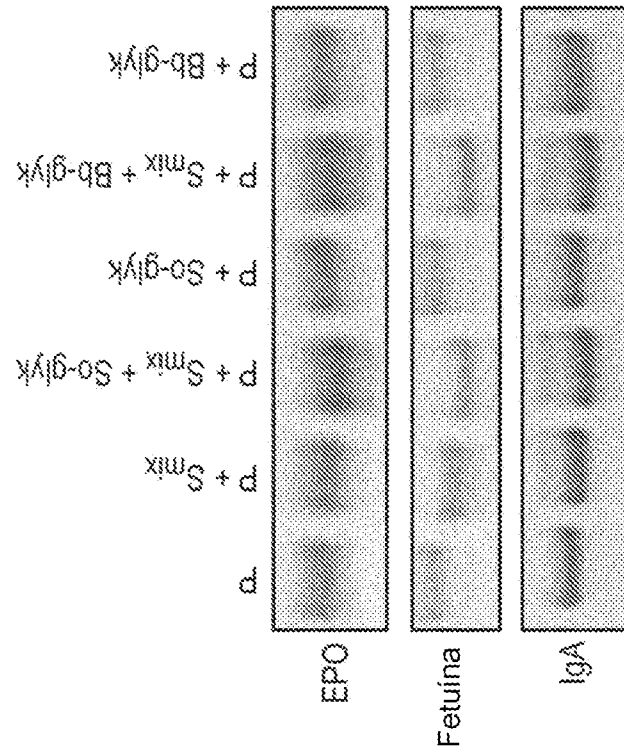


Fig. 13

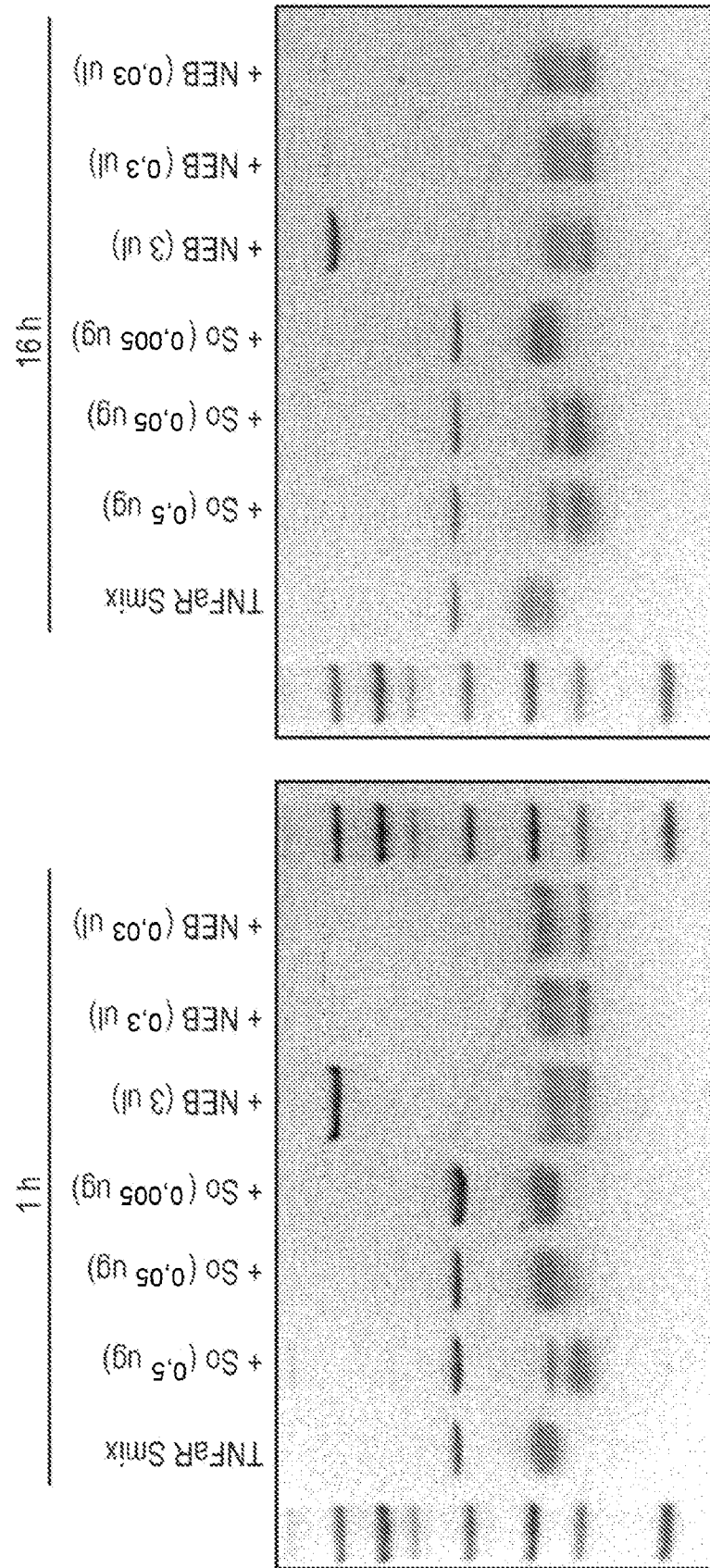


Fig. 14

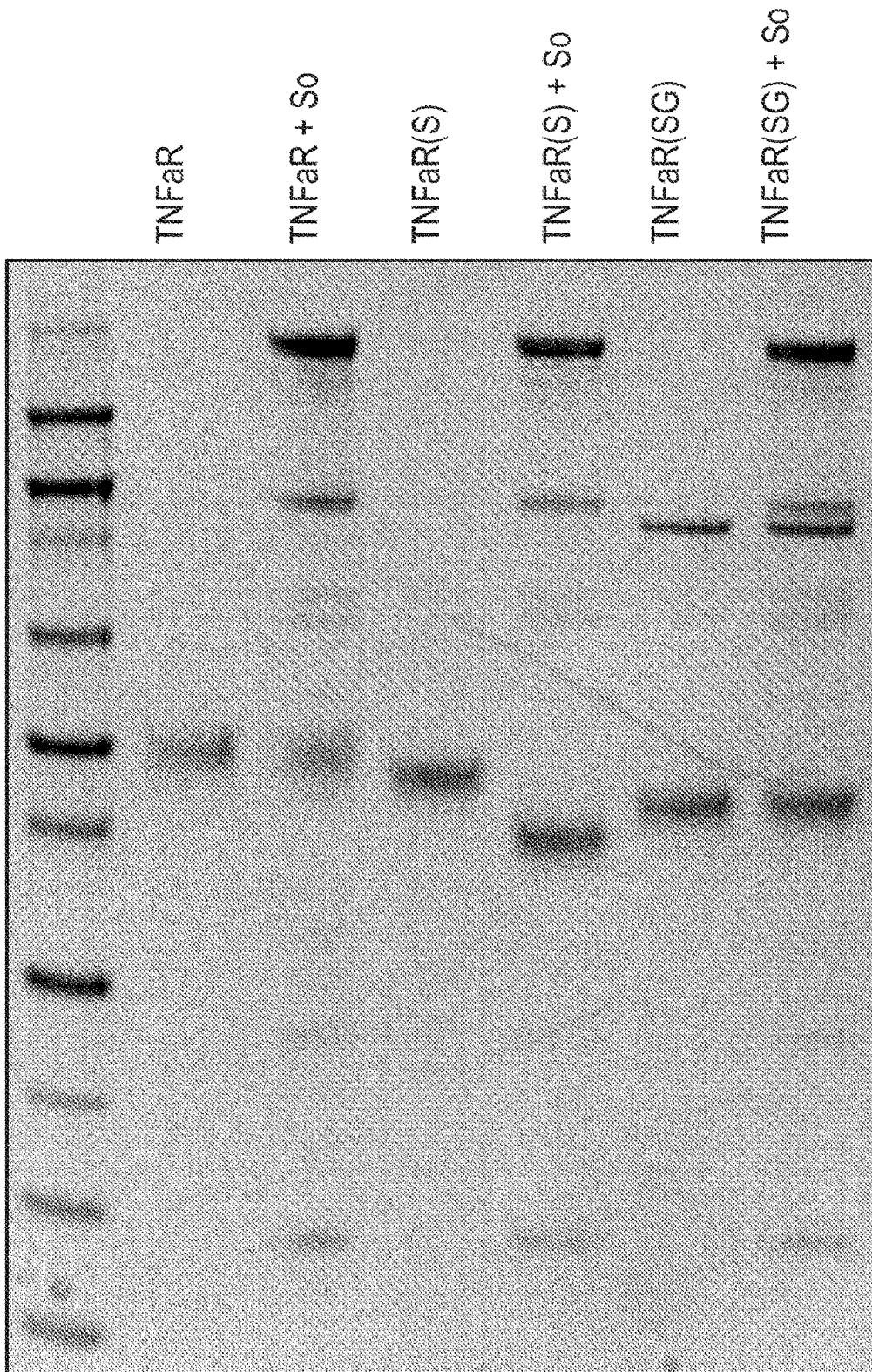


Fig. 15

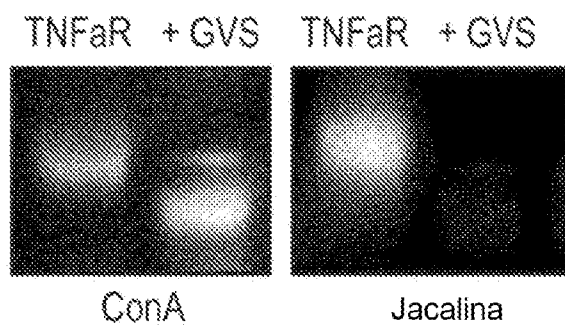
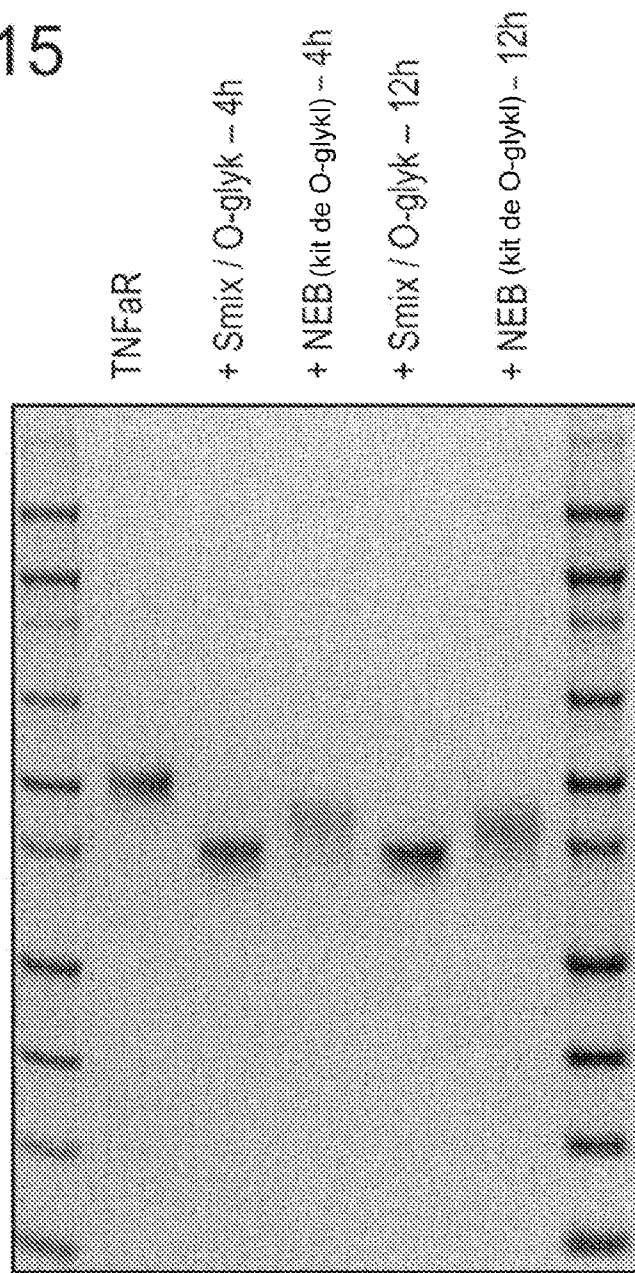


Fig. 16

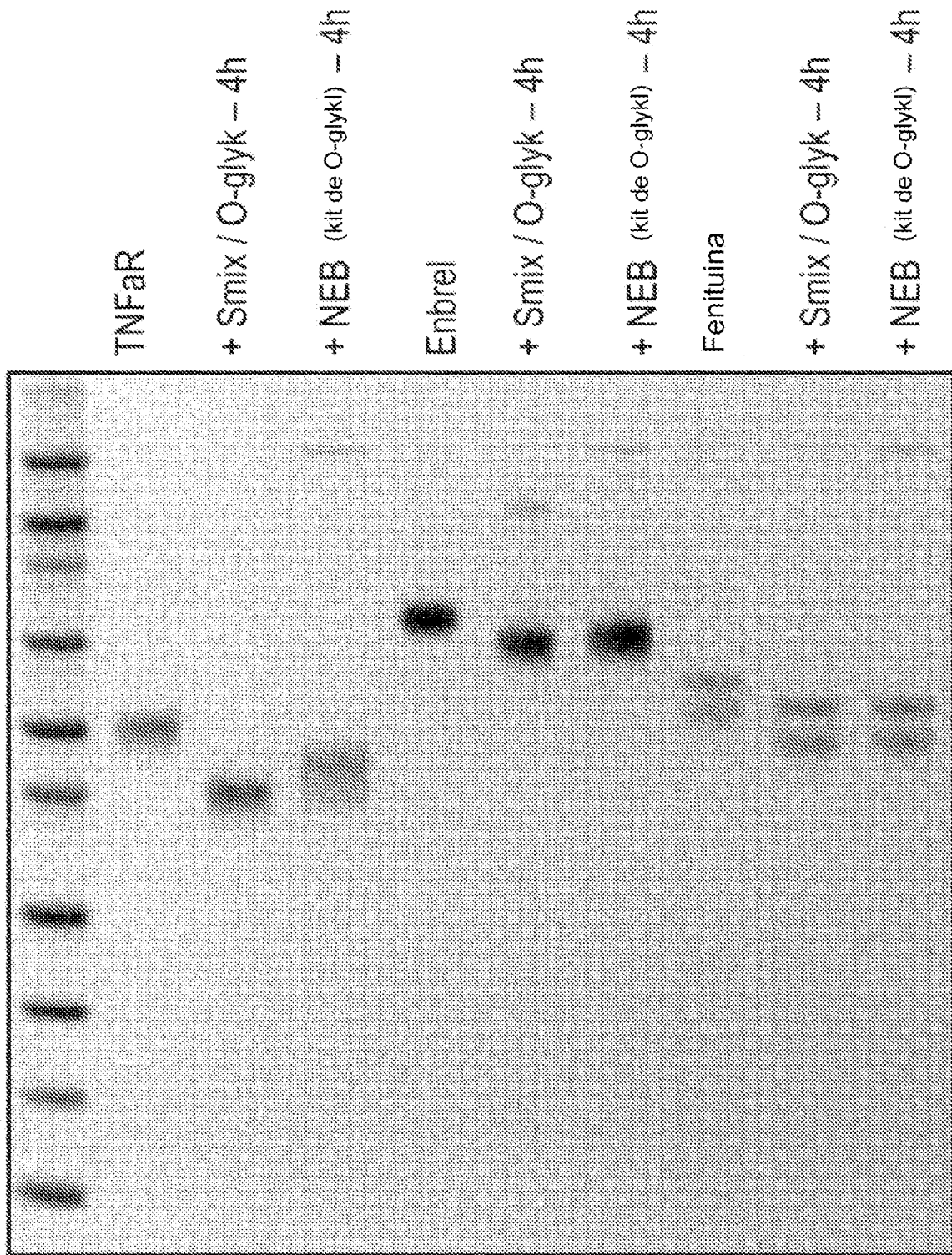


Fig. 17

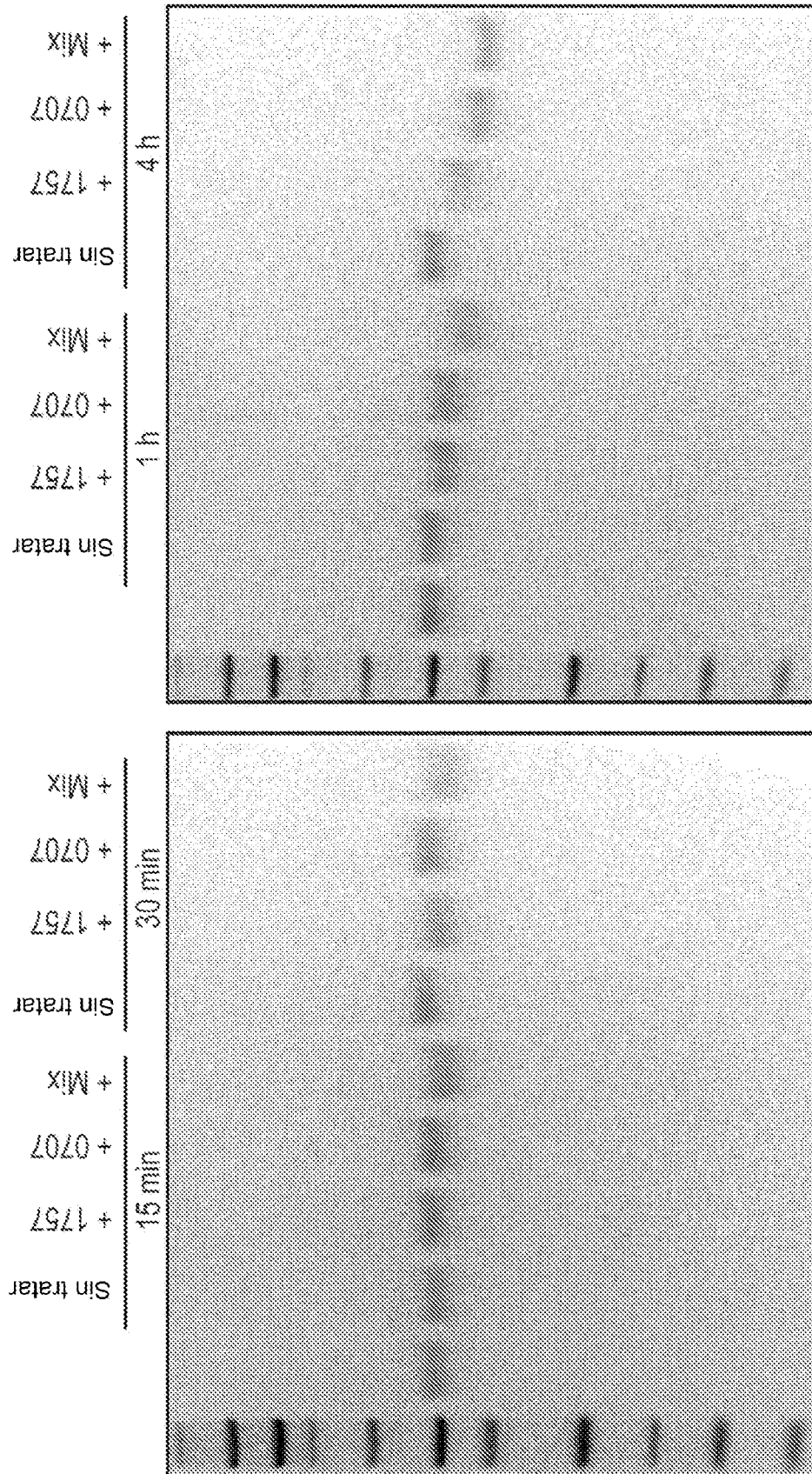


Fig. 18

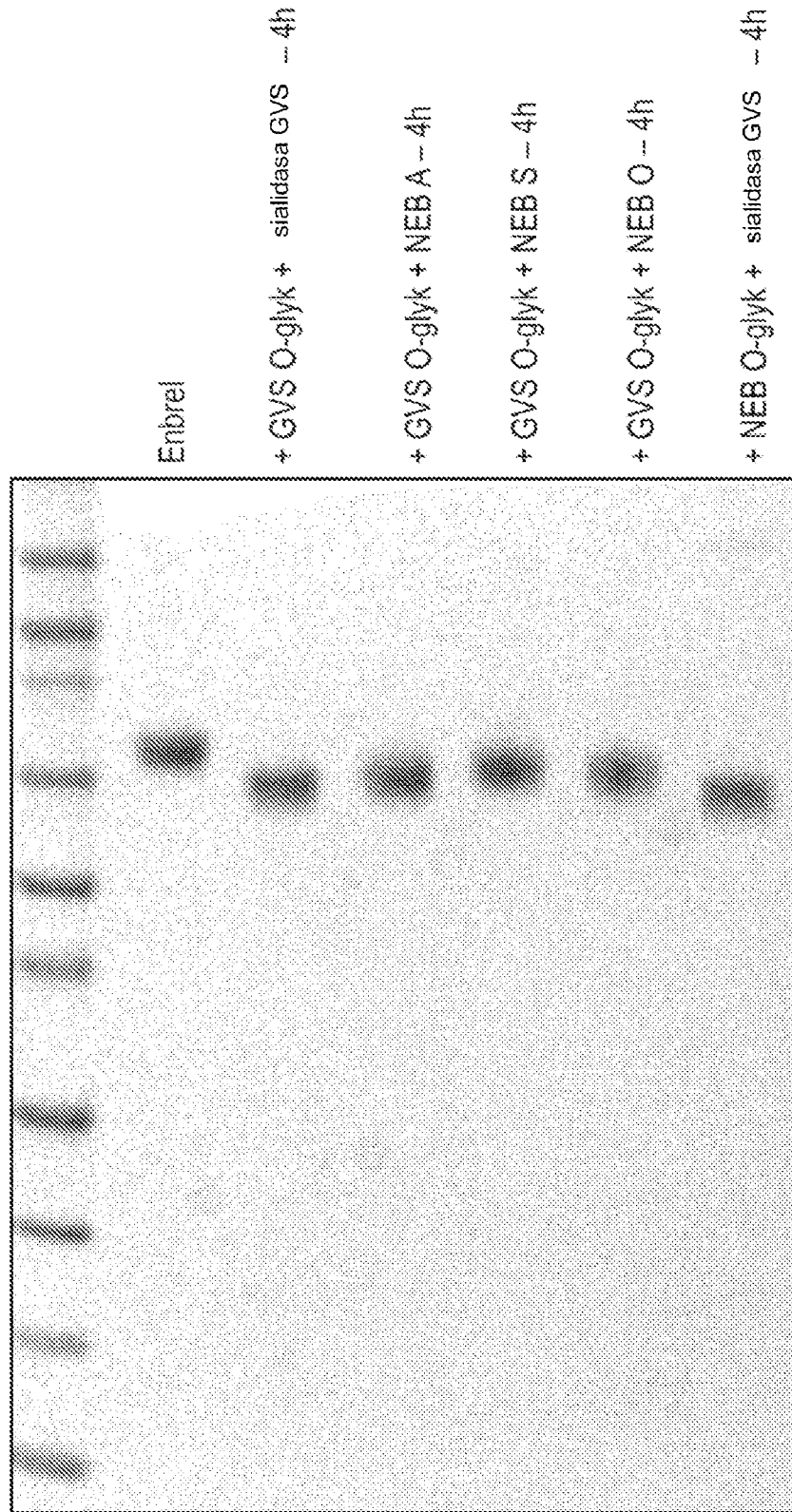


Fig. 19

