

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 972 511**

51 Int. Cl.:

C12N 9/24 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.02.2017 PCT/EP2017/052792**

87 Fecha y número de publicación internacional: **17.08.2017 WO17137459**

96 Fecha de presentación y número de la solicitud europea: **08.02.2017 E 17704236 (3)**

97 Fecha y número de publicación de la concesión europea: **10.01.2024 EP 3414328**

54 Título: **Enzimas para el corte de glicoproteínas**

30 Prioridad:

08.02.2016 EP 16154712
08.02.2016 EP 16154739
08.06.2016 EP 16173595
08.06.2016 EP 16173599
23.12.2016 EP 16206867

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
13.06.2024

73 Titular/es:

SYNAFFIX B.V. (100.0%)
Kloosterstraat 9
5349 AB Oss, NL

72 Inventor/es:

VAN GEEL, REMON;
WIJDEVEN, MARIA ANTONIA;
HURKMANS, INGE CATHARINA JOSEPHINA;
VAN DELFT, FLORIS LOUIS y
VAN BERKEL, SANDER SEBASTIAAN

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

Observaciones:

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

ES 2 972 511 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 corte de glicoproteínas

5 **Campo de la invención**

[0001] La presente invención se sitúa en el campo de la hidrólisis enzimática de oligosacáridos, más en particular del corte de glicoproteínas. La invención se refiere a enzimas mejoradas para tal corte para liberar la GlcNAc central y a un proceso para cortar glicoproteínas usando las enzimas según la invención.

10

Antecedentes de la invención

[0002] Las glicoproteínas existen en muchas variantes glicosiladas, o glicofomas, que pueden diferir sustancialmente en sus propiedades bioquímicas y funciones (biológicas). Los glicanos son estructuralmente diversos e incorporan una amplia gama de residuos de monosacáridos y enlaces glicosídicos.

15

[0003] Muchas proteínas terapéuticas son glicoproteínas y, aunque algunas se purifican a partir de fuentes naturales, la mayoría se expresan de forma recombinante. La elección del sistema de expresión influye mucho en la glicosilación. Se han realizado esfuerzos notables para controlar la glicosilación de los sistemas de producción de glicoproteínas motivados por el impacto en la funcionalidad *in vivo*. Por ejemplo, los anticuerpos monoclonales con glicosilación modificada muestran una farmacocinética y una función efectora mejoradas. Los glicopéptidos ofrecen posibilidades interesantes en el desarrollo de vacunas contra el cáncer dada su capacidad para estimular la inmunidad tanto humoral como celular. Además, el escudo de glicanos del VIH es una diana eficaz para la neutralización de anticuerpos y una diana emergente para el diseño de vacunas.

20

25

[0004] Por otro lado, la eliminación de N-glicanos de las glicoproteínas proporciona oportunidades terapéuticas complementarias. La desglicosilación de IgG disminuye significativamente la unión de anticuerpos a los receptores Fc-gamma, evitando así la captación inespecífica de anticuerpos por, por ejemplo, macrófagos o megacariocitos, lo que puede provocar trombocitopenia. Este último fenómeno biológico es responsable de la toxicidad limitante de dosis (DLT) de Kadcyla[®], un conjugado de anticuerpo y fármaco para tratar el cáncer de mama con regulación positiva de HER2. La desglicosilación selectiva de anticuerpos *in vivo* brinda oportunidades para tratar a pacientes con autoinmunidad mediada por anticuerpos. La eliminación de la glicofoma con alto contenido en manosa en una glicoproteína terapéutica recombinante puede ser beneficiosa, ya que se sabe que las glicofomas con alto contenido en manosa ponen en riesgo la eficacia terapéutica mediante una captación inespecífica por parte de los receptores de manosa endógenos y conducen a una eliminación rápida, como se describe, por ejemplo, en Gorovits y Krinos-Fiorotti, Cancer Immunol. Immunother. 2013, 62, 217-223 y Goetze et al, Glycobiology 2011, 21, 949-959. Además, Van de Bovenkamp et al., J. Immunol. 2016, 196, 1435-1441 describen cómo los glicanos con alto contenido en manosa pueden influir en la inmunidad. Reusch y Tejada, Glycobiology 2015, 25, 1325-1334, describen que una glicosilación inadecuada en anticuerpos monoclonales podría contribuir a una producción ineficaz a partir de genes de Ig expresados. En algunos casos, una secuencia de adición de carbohidratos generada por un reordenamiento de la región V o una hipermutación somática puede dar como resultado un anticuerpo que no puede plegarse ni secretarse adecuadamente, como se describe en Gala y Morrison, J. Immunol. 2004, 172, 5489-5494.

30

35

40

45

[0005] Una ventaja adicional de las proteínas terapéuticas desglicosiladas es la regularidad entre lotes mucho más facilitada y la homogeneidad significativamente mejorada, lo cual es muy ventajoso para la aprobación por las autoridades.

50

55

[0006] Un enfoque muy útil y sencillo para obtener proteínas recombinantes desglicosiladas, ofreciendo así una ruta para mejorar la eficacia de los anticuerpos terapéuticos y otras N-glicoproteínas, es mediante la eliminación enzimática de glicanos. Afortunadamente, se han descubierto endoglicosidasas que son capaces de escindir N-glicanos, lo que ofrece la posibilidad de eliminación selectiva a partir de una glicoproteína recombinante. Las endoglicosidasas también han encontrado un uso en la preparación de conjugados a partir de glicoproteínas, liberando selectivamente los restos básicos de GlcNAc tras el corte, seguido de la bioconjugación. Otro campo de uso de las endoglicosidasas es la espectrometría de masas, una de las herramientas analíticas clave para caracterizar proteínas (terapéuticas), incluidas las glicoproteínas y los anticuerpos monoclonales en particular. Mediante la escisión enzimática del glicano complejo y heterogéneo de la proteína, se facilita significativamente el análisis por espectrometría de masas.

60

65

[0007] La bioconjugación es el proceso de unir dos o más moléculas, de las cuales al menos una es una biomolécula y la otra molécula puede denominarse "molécula diana" o "molécula de interés". Se ha descubierto que muchos compuestos diferentes son útiles para conjugarse con glicoproteínas. Por ejemplo, la modulación de la estructura y función de las proteínas mediante modificación covalente con una sonda química para su detección y/o aislamiento se ha convertido en una poderosa herramienta en la investigación basada en proteomas y en aplicaciones biomédicas. El etiquetado fluorescente o de afinidad de proteínas es clave para estudiar el tráfico de proteínas en su hábitat nativo, y las vacunas basadas en conjugados de proteínas y

carbohidratos han ganado importancia en la lucha contra el VIH, el cáncer, la malaria y las bacterias patógenas. La PEGilación de proteínas o la unión de una proteína a la albúmina sérica son estrategias útiles para mejorar el perfil farmacocinético al reducir las tasas de eliminación, mientras que la funcionalización de una proteína transportadora, como un anticuerpo monoclonal con una carga útil tóxica, es una estrategia prometedora en el tratamiento dirigido de enfermedades (en particular del cáncer).

[0008] En general, se pueden reconocer dos conceptos estratégicos en el campo de la bioconjugación: (a) conjugación basada en un grupo funcional nativo (en otras palabras: un grupo funcional ya presente en la biomolécula de interés, como por ejemplo un tiol, una amina, un alcohol o una unidad de hidroxifenol) o (b) un proceso en dos pasos que implica la modificación de uno (o más) grupos reactivos únicos en una biomolécula antes del proceso de conjugación real.

[0009] El primer enfoque normalmente implica una cadena lateral de aminoácido reactivo en una proteína (por ejemplo, cisteína o lisina), o un grupo funcional en un glicano (por ejemplo, amina, aldehído) o ácido nucleico (por ejemplo, funcionalidad purina o pirimidina o alcohol). Como se resume, entre otros, en G.T. Hermanson, "Bioconjugate Techniques", Elsevier, 3ª ed. 2013. Cabe destacar que la conjugación de cisteína-maleimida destaca para la conjugación de proteínas en virtud de su alta velocidad de reacción y quimioselectividad. Sin embargo, cuando no hay cisteína libre disponible para la conjugación, como ocurre con muchas proteínas, a menudo se requieren otros métodos, cada uno de los cuales adolece de sus propias deficiencias, especialmente en términos de especificidad al sitio. Además, una desventaja general de los conjugados de proteínas obtenidos mediante alquilación con maleimidas es que, en general, los conjugados de succinimida resultantes pueden ser inestables debido a la reacción inversa de la alquilación, es decir, una reacción retro-Michael.

[0010] Una buena solución y de amplia aplicación para la bioconjugación implica el enfoque en dos pasos.

[0011] Aunque es más laboriosa, la conjugación en dos pasos mediante una funcionalidad modificada generalmente conduce a una mayor selectividad (especificidad al sitio) que la conjugación con una funcionalidad natural. Además de eso, se puede lograr una estabilidad total mediante la elección adecuada de la construcción. Los ejemplos típicos de un grupo funcional que se puede impartir a la biomolécula incluyen alquino (tensionado), alqueno (tensionado), norborneno, tetrazina, azida, fosfina, óxido de nitrilo, nitrona, nitrilo imina, compuesto diazo, compuesto carbonilo, (O-alquil)hidroxilamina e hidracina, que pueden lograrse mediante un enfoque de biología molecular o química. Se sabe que cada uno de los grupos funcionales anteriores tiene al menos un compañero de reacción, que en muchos casos implica una reactividad mutua completa. Por ejemplo, los ciclooctinos reaccionan selectiva y exclusivamente con los 1,3-dipolos, los alquenos tensionados con las tetrazinas y las fosfinas con las azidas, dando lugar a enlaces covalentes totalmente estables.

[0012] Una ruta eficaz hacia la introducción de funcionalidades modificadas, como azidas, en glicoproteínas específicas es mediante la funcionalización selectiva de los glicanos presentes en la glicoproteína. Todos los anticuerpos recombinantes, generados en sistemas hospedadores de mamíferos, contienen el sitio de N-glicosilación conservado en el residuo de asparagina en la posición 297 o cerca de ella. Estos glicanos siempre se forman como una mezcla compleja de isoformas, véanse, por ejemplo, las figuras 1 y 2, que consisten en una mezcla altamente heterogénea de glicanos complejos, híbridos y con alto contenido en manosa. El corte de estos glicanos mediante una endoglicosidasa deja solo el resto central de GlcNAc (unido a N297), opcionalmente fucosilado en el grupo 6-OH. El resto central de GlcNAc liberado proporciona un punto de anclaje adecuado para las moléculas diana, proporcionando un producto con una homogeneidad mucho mayor en comparación con los productos obtenidos por conjugación con restos de azúcar terminales presentes en la estructura de glicanos original. Sin embargo, una desventaja de este enfoque es que diferentes glicanos pueden requerir diferentes endoglicosidasas, cada una con sus propias condiciones óptimas, de modo que pueden ser necesarios múltiples tratamientos enzimáticos para un corte adecuado y completo de la glicoproteína. Por ejemplo, se sabe que EndoH corta glicofomas híbridas y con alto contenido en manosa, pero no glicanos de tipo complejo, mientras que EndoS es capaz de cortar glicanos de tipo complejo y, hasta cierto punto, glicanos híbridos, pero no formas con alto contenido en manosa. EndoF2 puede cortar glicanos complejos (pero no híbridos), mientras que endoF3 solo puede cortar glicanos complejos que también están 1,6-fucosilados. Otra endoglicosidasa, EndoD, es capaz de hidrolizar únicamente el glicano Man₅ (M5). Una descripción general de las actividades específicas de diferentes endoglicosidasas se divulga en Freeze et al. *Curr. Protoc. Mol. Biol.*, 2010, 89:17.13A.1-17. Yamamoto et al. divulgan en *Glycoconjugate J.* 2005, 22, 35-42, una construcción química de EndoD y EndoBH, que estaba completamente inactiva. La construcción química se diseñó para investigar la homología de ambas endoglicosidasas en el corte de glicanos. En el contexto de la conjugación de glicoproteínas, en WO 2007/095506 y WO 2008/029281 se describe que el corte del glicano puede tener lugar con EndoH, hidrolizando así un enlace glicosídico GlcNAc-GlcNAc y liberando una GlcNAc para la introducción enzimática de GalNAz. Van Geel et al. divulgan en *Bioconjugate Chem.* 2015, 26, 2233 que la transferencia de una variedad de restos de galactosa modificados con azida al residuo central de GlcNAc de un anticuerpo monoclonal, obtenido mediante corte con una endoglicosidasa, seguido de la unión de una carga útil tóxica mediante química click sin cobre, es un método eficaz para obtener conjugados anticuerpo-fármaco con un perfil de eficacia y seguridad mejorado demostrado en comparación con el fármaco comercializado Kadcyła®.

[0013] Como producto de la tecnología del ADN recombinante, las proteínas de fusión se han desarrollado como una clase de biomoléculas novedosas con propiedades multifuncionales. Al fusionar genéticamente dos o más proteínas o dominios proteicos, se genera un producto de proteína de fusión que puede mostrar funciones similares o claramente diferentes a las de los restos componentes. Las proteínas de fusión han encontrado aplicaciones en estrategias de purificación, inmovilización, adquisición de imágenes y productos biofarmacéuticos. Por ejemplo, muchos fármacos proteicos se fusionan con dominios Fc de anticuerpos, como la Fc-inmunoglobulina G1 (Fc-IgG1), o con proteínas transportadoras como la albúmina sérica humana (HSA) o la transferrina (Tf) para extender sus semividas plasmáticas y lograr efectos terapéuticos mejorados. Varios medicamentos de proteínas de fusión, incluido Enbrel® (factor de necrosis tumoral/Fc-IgG1), Ontak® (interleucina-2/toxina diftérica), Orencia® (antígeno 4 asociado al linfocito T citotóxico/Fc-IgG1), Amevive® (antígeno 3 de función leucocitaria/Fc-IgG1), Arcalyst® (dominio extracelular del receptor de interleucina 1/Fc-IgG1) y Nplate® (trombopoyetina/Fc-IgG1) han sido aprobados por la FDA para su aplicación terapéutica. Un ejemplo relevante de una proteína de fusión de una endoglicosidasa se puede encontrar en Warren et al., Prot. Ing. Design Select. 2005, 18, 497-501, donde se describe una fusión del dominio de unión a carbohidratos (CBM) con EndoF1 o PNGasaF.

[0014] La construcción exitosa de una proteína de fusión recombinante requiere las proteínas que la componen, pero también las porciones conectoras pueden desempeñar un papel importante. Las porciones conectoras pueden ser cortas o largas, flexibles o rígidas y de naturaleza escindible o no escindible. En algunos casos, la porción conectora puede aumentar la estabilidad o el plegamiento, mejorar la expresión o la actividad biológica o alterar la farmacocinética. La naturaleza típica de las porciones conectoras conocidas en la técnica son oligómeros de glicina, por ejemplo, G₈, oligómeros de GGGGS, oligómeros de EAAAK y variantes de estos. Se puede encontrar una descripción general reciente de las porciones conectoras para proteínas de fusión en Chen et al., Adv. Drug Deliv. Rev.2013, 65, 1357-1369.

Resumen de la invención

[0015] La invención se refiere a proteínas de fusión, en las que dos endoglicosidasas están fusionadas, posiblemente mediante una porción conectora. Las enzimas de fusión según la invención son capaces convenientemente de cortar glicoproteínas que comprenden al menos dos glicofomas distintas en un solo paso. Todos los glicanos de glicoproteínas que no pueden eliminarse mediante una única enzima convencional se cortan completamente para liberar la GlcNAc central mediante la enzima de fusión según la invención. Sorprendentemente, ambas actividades de la enzima de fusión funcionan bien al pH óptimo de una de las endoglicosidasas, mientras que la otra endoglicosidasa normalmente puede requerir un pH diferente para funcionar de manera óptima. Además, se ha descubierto que la actividad de una endoglicosidasa particular en una proteína de fusión puede mostrar una mayor eficiencia de corte en comparación con la misma endoglicosidasa como una sola enzima. La invención se refiere además al uso de la enzima de fusión según la invención para cortar glicoproteínas. En otro aspecto, la invención se refiere al proceso de producción de la enzima de fusión. En un aspecto adicional, la invención se refiere a un proceso para cortar glicoproteínas, que comprende cortar la glicoproteína con una enzima de fusión según la invención, para obtener una glicoproteína cortada.

Descripción de las figuras

[0016]

En la Figura 1 se muestran glicanos ejemplares con alto contenido en manosa (Man₃ (M3), Man₅ (M5) y Man₉ (M9)), de tipo complejo (biantenarico), tipo bisecado, tipo triantenarico y tetraantenarico.

En la Figura 2 se muestra un anticuerpo que comprende un glicano en cada cadena pesada. Los patrones de glicosilación complejos más típicos de un anticuerpo recombinante son G1F, G0F y G1. También puede estar presente cierta glicosilación de tipo manosa (M5) y, en algunos casos, un glicano de tipo híbrido (por ejemplo, M5G1).

En la Figura 3 se muestran los resultados del corte de la glicoproteína RNasaB con alto contenido en manosa en Tris pH 7,5 mediante las endoglicosidasas EndoS, EndoS2 y EndoSH. Serie de concentraciones: 0,025, 0,125, 0,125 (duplicado) y 0,25 mg/ml. Banda superior = RNasaB intacta, banda inferior = RNasaB cortada.

En la Figura 4 se muestra una gráfica del porcentaje de conversión (corte) de cAC10 a pH 6 mediante las endoglicosidasas EndoS, EndoS2 y EndoSH obtenidas en el Ejemplo 14. Línea de tendencia para EndoSH: $y = 2,0784x + 1,027$ ($R^2 = 0,9982$). Línea de tendencia para EndoS: $y = 2,0103x$ ($R^2 = 0,9838$). Línea de tendencia para EndoS2: $y = 1,0297x - 0,1622$ ($R^2 = 0,9998$).

En la Figura 5 se muestra una gráfica del porcentaje de conversión (corte) de trastuzumab con alto contenido en manosa a pH 6 mediante las endoglicosidasas EndoS, EndoS2 y EndoSH obtenidas en el Ejemplo 15.

Descripción detallada de la invención

Definiciones

[0017] El verbo "comprender", y sus conjugaciones, tal como se usa en esta descripción y en las reivindicaciones, se usa en su sentido no limitativo para significar que los elementos que siguen a la palabra están incluidos, pero los elementos no mencionados específicamente no están excluidos.

5 [0018] Además, la referencia a un elemento por el artículo indefinido "un" o "una" no excluye la posibilidad de que esté presente más de uno de los elementos, a menos que el contexto requiera claramente que exista uno y solo uno de los elementos. Por tanto, el artículo indefinido "un" o "una" suele significar "al menos un/a".

10 [0019] El término general "azúcar" se utiliza en este documento para indicar un monosacárido, por ejemplo, glucosa (Glc), galactosa (Gal), manosa (Man) y fucosa (Fuc). El término "derivado de azúcar" se utiliza en el presente documento para indicar un derivado de un azúcar monosacárido, es decir, un azúcar monosacárido que comprende sustituyentes y/o grupos funcionales. Entre los ejemplos de un derivado de azúcar se incluyen los aminoazúcares y ácidos de azúcar, por ejemplo, glucosamina (GlcNH₂), galactosamina (GalNH₂), N-acetilglucosamina (GlcNAc), N-acetilgalactosamina (GalNAc), ácido siálico (Sia), también conocido como ácido N-acetilneuramínico (NeuNAc), y ácido N-acetilmurámico (MurNAc), ácido glucurónico (GlcA) y ácido idurónico (IdoA).

20 [0020] El término "nucleótido" se utiliza en este documento con su significado científico normal. El término "nucleótido" se refiere a una molécula que está compuesta por una nucleobase, un azúcar de cinco carbonos (ya sea ribosa o 2-desoxirribosa) y uno, dos o tres grupos fosfato. Sin el grupo fosfato, la nucleobase y el azúcar componen un nucleósido. Por tanto, un nucleótido también puede denominarse nucleósido monofosfato, nucleósido difosfato o nucleósido trifosfato. La nucleobase puede ser adenina, guanina, citosina, uracilo o timina. Entre los ejemplos de nucleótidos se incluyen uridina difosfato (UDP), guanosina (GDP), timidina difosfato (TDP), citidina difosfato (CDP) y citidina monofosfato (CMP).

25 [0021] El término "proteína" se utiliza en este documento en su significado científico normal. En el presente documento, los polipéptidos que comprenden aproximadamente 10 o más aminoácidos se consideran proteínas. Una proteína puede comprender aminoácidos naturales, pero también no naturales.

30 [0022] Las proteínas y enzimas incluyen mutantes de las mismas. Por ejemplo, "endoglicosidasa" incluye tanto endoglicosidasas nativas (de tipo salvaje) como endoglicosidasas mutantes, siempre que la actividad endoglicosidasa se mantenga sustancialmente. Un dominio que tiene una secuencia de aminoácidos que es diferente de una secuencia de aminoácidos de tipo salvaje se denomina aquí dominio mutante. La mutación puede, por ejemplo, comprender un único cambio de aminoácido (una mutación puntual), pero también múltiples cambios de aminoácidos (por ejemplo, de 1 a 10, preferiblemente de 1 a 6, más preferiblemente de 1, 2, 3 o 4, incluso más preferiblemente de 1 o 2 aminoácidos), o una delección o inserción de uno o más (por ejemplo, de 1 a 10, preferiblemente de 1 a 6, tal como 1, 2, 3 o 4, preferiblemente de 1 o 2) aminoácidos. Alternativamente, se pueden aplicar delecciones o inserciones más grandes a la enzima. Por ejemplo, se ha descubierto que un fragmento de la endoglicosidasa D (delección de 599 aminoácidos de su porción C-terminal) conserva su actividad endoglicosidasa (Yamamoto et al. en *Glycoconjugate J.* 2005, 22, 35-42). El experto conoce las posibilidades a este respecto y, siempre que se conserve sustancialmente la actividad endoglicosidasa, la enzima puede contener cualquier tipo de mutación.

45 [0023] El término "glicoproteína" se utiliza en este documento en su significado científico normal y se refiere a una proteína que comprende una o más cadenas de monosacáridos u oligosacáridos ("glicanos") unidas covalentemente a la proteína. Un glicano puede estar unido a un grupo hidroxilo de la proteína (glicano unido a O), por ejemplo al grupo hidroxilo de la serina, treonina, tirosina, hidroxilisina o hidroxiprolina, o a una función amida de la proteína (N-glicoproteína), por ejemplo, asparagina o arginina, o a un carbono de la proteína (C-glicoproteína), por ejemplo, triptófano. Una glicoproteína puede comprender más de un glicano, puede comprender una combinación de uno o más monosacáridos y uno o más oligosacáridos glicanos, y puede comprender una combinación de glicanos unidos a N, unidos a O y unidos a C. Se estima que más del 50 % de todas las proteínas tienen algún tipo de glicosilación y, por tanto, se consideran glicoproteínas. Entre los ejemplos de glicoproteínas se incluyen PSMA (antígeno prostático específico de membrana), CAL (lipasa de *Candida antartica*), gp41, gp120, EPO (eritropoyetina), proteína anticongelante y anticuerpos.

55 [0024] El término "glicano" se utiliza en este documentocoen su significado científico normal y se refiere a una cadena de monosacáridos u oligosacáridos que está unida a una proteína. Por tanto, el término glicano se refiere a la parte carbohidrato de una glicoproteína. El glicano está unido a una proteína a través del carbono C-1 de un azúcar, que puede no tener más sustituciones adicionales (monosacárido) o puede tener sustituyentes adicionales en uno o más de sus grupos hidroxilo (oligosacárido). Un glicano de origen natural normalmente comprende de 1 a aproximadamente 10 restos de sacárido. Sin embargo, cuando una cadena de sacáridos más larga está unida a una proteína, dicha cadena de sacáridos también se considera aquí un glicano. Un glicano de una glicoproteína puede ser un monosacárido. Normalmente, un glicano monosacárido de una glicoproteína consiste en una única N-acetilglucosamina (GlcNAc), glucosa (Glc), manosa (Man) o fucosa (Fuc) unida covalentemente a la proteína. Un glicano también puede ser un oligosacárido. Una cadena de oligosacáridos de una glicoproteína puede ser lineal o ramificada. En un oligosacárido, el azúcar que está directamente unido a la

proteína se llama azúcar central. En un oligosacárido, un azúcar que no está directamente unido a la proteína y está unido a al menos otros dos azúcares se llama azúcar interno. En un oligosacárido, un azúcar que no está directamente unido a la proteína, sino a otro azúcar, es decir, que no lleva más sustituyentes azúcar en uno o más de sus otros grupos hidroxilo, se llama azúcar terminal. Para evitar dudas, pueden existir múltiples azúcares terminales en un oligosacárido de una glicoproteína, pero solo un azúcar central. El extremo de un oligosacárido que está directamente unido a la proteína se llama extremo reductor de un glicano. El otro extremo del oligosacárido se llama extremo no reductor de un glicano. Un glicano puede ser un glicano unido a O, un glicano unido a N o un glicano unido a C.

[0025] En un glicano unido a O, un glicano monosacárido u oligosacárido está unido a un átomo de O en un aminoácido de la proteína, típicamente a través de un grupo hidroxilo de serina (Ser) o treonina (Thr). Para los glicanos unidos a O, existe una amplia diversidad de cadenas. Los glicanos unidos a O de origen natural suelen presentar un resto α -O-GalNAc unido a serina o treonina, además sustituido con galactosa, ácido siálico y/o fucosa. El aminoácido hidroxilado que lleva la sustitución de glicano puede formar parte de cualquier secuencia de aminoácidos en la proteína.

[0026] En un glicano unido a N, un glicano monosacárido u oligosacárido está unido a la proteína a través de un átomo de N en un aminoácido de la proteína, típicamente a través de una amida de nitrógeno en la cadena lateral de la asparagina (Asn) o la arginina (Arg). Para los glicanos unidos a N, existe una amplia diversidad de glicanos. Los glicanos unidos a N de origen natural presentan un resto β -N-GlcNAc unido a asparagina, a su vez con un sustituyente adicional β -GlcNAc en su 4-OH, a su vez con un sustituyente adicional β -Man en su 4-OH, a su vez con un sustituyente adicional α -Man en su 3-OH y 6-OH, lo que lleva al glicano pentasacárido $\text{Man}_3\text{GlcNAc}_2$. El resto central GlcNAc puede tener incorporado α -Fuc en la posición 6-OH. $\text{Man}_3\text{GlcNAc}_2$ es la estructura base de oligosacáridos común de casi todas las glicoproteínas unidas a N y puede transportar una amplia variedad de otros sustituyentes, incluidos, entre otros, Man, GlcNAc, Gal y ácido siálico. La asparagina que está reemplazada con el glicano en su cadena lateral suele ser parte de la secuencia Asn-X-Y, siendo X cualquier aminoácido excepto prolina e Y siendo serina o treonina.

[0027] En un glicano unido a C, un glicano monosacárido u oligosacárido está unido a un átomo de C en un aminoácido de la proteína, típicamente a un átomo de C del triptófano (Trp).

[0028] El término "anticuerpo" se utiliza en este documento en su significado científico normal. Un anticuerpo es una proteína generada por el sistema inmunitario que es capaz de reconocer y unirse a un antígeno específico. Un anticuerpo es un ejemplo de glicoproteína. El término anticuerpo en el presente documento se usa en su sentido más amplio e incluye específicamente anticuerpos monoclonales, anticuerpos policlonales, dímeros, multímeros, anticuerpos multispecíficos (por ejemplo, anticuerpos biespecíficos), fragmentos de anticuerpos y anticuerpos de cadena simple y doble. En el presente documento también se pretende que el término "anticuerpo" incluya anticuerpos humanos, anticuerpos humanizados, anticuerpos quiméricos y anticuerpos que se unen específicamente al antígeno del cáncer. El término "anticuerpo" pretende incluir anticuerpos completos, pero también fragmentos de un anticuerpo que se unen al antígeno, por ejemplo, un fragmento Fab de un anticuerpo, F(ab')_2 , fragmento Fv o fragmento Fc de un anticuerpo escindido, un fragmento scFv-Fc, un minicuerpo, un diacuerpo o un scFv. Además, el término incluye anticuerpos modificados genéticamente y derivados de un anticuerpo. Los anticuerpos, fragmentos de anticuerpos y anticuerpos modificados genéticamente pueden obtenerse mediante métodos conocidos en la técnica. Entre los ejemplos típicos de anticuerpos se incluyen, entre otros, abciximab, rituximab, basiliximab, palivizumab, infliximab, trastuzumab, alemtuzumab, adalimumab, tositumomab-I131, cetuximab, ibrituximab tiuxetan, omalizumab, bevacizumab, natalizumab, ranibizumab, panitumumab, eculizumab, certolizumab pegol, golimumab, canakinumab, catumaxomab, ustekinumab, tocilizumab, ofatumumab, denosumab, belimumab, ipilimumab y brentuximab.

[0029] Una "porción conectora" se define en este caso como un resto que conecta dos o más elementos de un compuesto. La enzima de fusión según la invención puede contener, por ejemplo, una porción conectora que conecta las dos unidades de endoglicosidasa. En el contexto de las enzimas de fusión según la presente invención, las porciones conectoras normalmente contienen al menos un aminoácido y, de la manera más preferible, consisten en uno o más aminoácidos.

[0030] Un "bioconjugado" se define en el presente documento como un compuesto en el que una biomolécula está unida covalentemente a una molécula diana a través de una porción conectora. Un bioconjugado comprende una o más biomoléculas y/o una o más moléculas diana. La porción conectora puede comprender uno o más restos espaciadores. Una molécula diana puede ser una sustancia activa, una molécula indicadora, un polímero, una superficie sólida, un hidrogel, una nanopartícula, una micropartícula o una biomolécula.

[0031] El término "enzima de fusión" en el presente documento se refiere a una enzima en la que las secuencias de aminoácidos de dos o más enzimas que originalmente pertenecían a enzimas separadas se unen entre sí, opcionalmente mediante una porción conectora. Las enzimas de fusión son conocidas en la técnica y pueden crearse mediante la unión de dos o más genes que originalmente codifican enzimas separadas. La traducción de

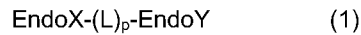
este gen da como resultado un único polipéptido con propiedades funcionales derivadas de cada una de las enzimas originales.

Enzima de fusión

5

[0032] En un primer aspecto, la invención se refiere a una enzima de fusión que comprende dos endoglicosidasas, opcionalmente conectadas mediante una porción conectora. La enzima de fusión según la invención puede representarse mediante la estructura (1):

10



15

[0033] En el presente documento, EndoX y EndoY son individualmente una endoglicosidasa, L es una porción conectora y p es 0 o 1. EndoX es EndoS y EndoY es una endoglicosidasa distinta de EndoX. En el contexto de la presente invención, la "enzima de fusión" también puede denominarse "proteína de fusión". La enzima de fusión según la invención es preferiblemente una fusión de extremo a extremo, ya sea directa o mediante una porción conectora L.

15

Endoglicosidasa

20

[0034] Las endoglicosidasas se conocen en la técnica como enzimas que escinden oligosacáridos entre dos enlaces glicosídicos, liberándolos como tales de glicoproteínas, glicopéptidos o glicolípidos. Estos oligosacáridos suelen denominarse glicanos. En el contexto de la presente invención, "Endo" se refiere a endoglicosidasa. Las endoglicosidasas hidrolizan el enlace entre dos unidades de azúcar en un oligosacárido o polisacárido, pero no entre la unidad central de azúcar, que está directamente unida a la parte peptídica de una glicoproteína, y el aminoácido al que está conectada. Las endoglicosidasas normalmente hidrolizan el enlace entre los dos residuos centrales de N-acetilglucosamina (GlcNAc) en glicanos unidos a N, dejando así el residuo central de GlcNAc conectado a la parte peptídica de la glicoproteína.

25

30

[0035] En el contexto de la presente invención, el término endoglicosidasa abarca todos los miembros de la familia de las endoglicosidasas que liberan oligosacáridos a partir de glicoproteínas, glicopéptidos o glicolípidos. La endoglicosidasa también puede escindir cadenas de polisacáridos entre residuos que no son el residuo terminal, aunque es más común liberar oligosacáridos de moléculas de lípidos y proteínas conjugadas.

35

[0036] En el contexto de la presente invención, el término endoglicosidasa abarca tanto las endoglicosidasas sin desnaturar como las endoglicosidasas truncadas y sus mutantes, siempre que se conserve sustancialmente la actividad endoglicosidasa. En otras palabras, la secuencia de aminoácidos de EndoX y EndoY puede comprender una secuencia de aminoácidos diferente en comparación con la endoglicosidasa sin desnaturar. En una forma de realización, la secuencia de aminoácidos de EndoX y EndoY comprende un mutante. En una forma de realización, la secuencia de aminoácidos de EndoX y EndoY no comprende un mutante. En una forma de realización, la secuencia de aminoácidos de EndoX y EndoY comprende una secuencia truncada. En una forma de realización, la secuencia de aminoácidos de EndoX y EndoY no comprende una secuencia truncada. Cuando se observa la secuencia de EndoX y EndoY individualmente, se prefiere que cada una de EndoX y EndoY tenga al menos un 80 % de identidad de secuencia con la correspondiente secuencia de aminoácidos nativa del dominio catalítico de la endoglicosidasa, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de aminoácidos nativa correspondiente. De la manera más preferible, cada una de EndoX y EndoY tiene una identidad de secuencia del 100 % con la secuencia de aminoácidos correspondiente del dominio catalítico de la endoglicosidasa. Alternativa o adicionalmente, se prefiere que cada una de EndoX y EndoY tenga al menos un 80 % de similitud de secuencia con la secuencia de aminoácidos nativa correspondiente, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de similitud de secuencia con la correspondiente secuencia de aminoácidos nativa del dominio catalizador de la endoglicosidasa. De la manera más preferible, cada una de EndoX y EndoY tiene un 100 % de similitud de secuencia con la correspondiente secuencia de aminoácidos nativa del dominio catalítico de la endoglicosidasa.

55

60

[0037] La identidad de secuencia y las similitudes se pueden calcular fácilmente mediante métodos conocidos y/o métodos de programas informáticos conocidos en la técnica, tales como BLASTP, disponible públicamente en NCBI y otras fuentes (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990), incorporados por referencia.

65

[0038] Los glicanos que pueden ser escindidos por las glicosidasas existen en varias glicofomas, que generalmente se agrupan en tres tipos: con alto contenido en manosa, complejas e híbridas. Los tres tipos tienen un núcleo de β 1,4-N,N'-diacetilquitobiosa (GlcNAc₂), conectado a un trisacárido formado por manosas (Man₃). Opcionalmente, la GlcNAc central puede estar fucosilada, pero no es siempre así. Los glicanos con alto contenido en manosa contienen al menos 2 residuos de manosa adicionales, lo que normalmente da como resultado de 5 a 9 residuos de manosa. Los glicanos complejos tienen uno o más monómeros de azúcar, que no

son manosa, conectados a dos de los residuos de manosa de la unidad central Man₅. Estos monómeros de azúcar adicionales normalmente se seleccionan de entre GlcNAc, galactosa (Gal) y ácido siálico (Neu5Ac). Los glicanos complejos existen en formas bi, tri y tetraantenarias, dependiendo del número de (oligo)sacáridos que están conectados a la unidad central Man₅. Los glicanos híbridos tienen un oligosacárido con alto contenido en manosa conectado a uno de los residuos de manosa de la unidad central Man₅, y un oligosacárido de tipo complejo conectado al otro residuo de manosa. En la Figura 1 se ofrece una descripción general de los tipos de glicanos. Incluso dentro de un tipo de glicano específico, existen muchas posibilidades, lo que aumenta la heterogenicidad de las glicoproteínas. Por ejemplo, los glicanos complejos biantenarios unidos al residuo N297 de los anticuerpos pueden existir en varias glicofomas distintas, incluidas, entre otras, G1, G1F, G0F y SG1F, como se muestra en la Figura 2.

[0039] Las endo-β-N-acetilglucosaminidasas (ENGasas, también denominadas endoglicosidasas o Endos), normalmente hidrolizan el N-glicano de una glicoproteína en el enlace β-1,4-glicosídico de la quitobiosa central. Estas enzimas, que se encuentran principalmente en GH18 y GH85 en la clasificación CAZy (Lombard et al. en Nucleic Acid Res. 2014, 42, D490) se distribuyen ampliamente desde bacterias hasta animales y participan en diversas funciones biológicas, como el metabolismo de los glicanos o la patogénesis bacteriana (Karamanos, Adv. Biochem. 2013, 1, 81). Las endoglicosidasas suelen ser específicas para la hidrólisis de uno o dos tipos de glicanos. Algunas son específicas para la hidrólisis de glicanos con alto contenido en manosa (por ejemplo, EndoA, EndoD, EndoT), mientras que otras, además de los glicanos con alto contenido en manosa, también escinden glicanos híbridos (por ejemplo, EndoF1, EndoH). También existen endoglicosidasas específicas para la hidrólisis de glicanos complejos en varias variantes. Una descripción general de las diferentes actividades se divulga en Freeze et al. en Curr. Protoc. Mol. Biol., 2010, 89:17.13A.1-17. La EndoS escinde glicanos complejos biantenarios, triantenarios y tetraantenarios, pero su actividad se limita casi exclusivamente a los anticuerpos, en particular la cadena pesada de IgG.

[0040] EndoX y EndoY son dos endoglicosidasas distintas. EndoX es EndoS y EndoY se selecciona preferiblemente del grupo que consiste en EndoA, EndoBi, EndoBH, EndoBT, EndoCE, EndoD, EndoE, EfEndo18A, EndoF1, EndoF2, EndoF3, EndoH, EndoLL, EndoM, EndoOm y EndoT. Estas endoglicosidasas y sus secuencias de aminoácidos son conocidas por los expertos. A continuación, se proporcionan algunas secuencias de aminoácidos preferidas para endoglicosidasas específicas.

[0041] En una forma de realización preferida, EndoS tiene al menos un 80 %, preferiblemente al menos un 90 %, más preferiblemente al menos un 95 % de identidad de secuencia con la SEQ ID N.º 4 o la SEQ ID N.º 5, de la manera más preferible, EndoS tiene un 100 % de identidad de secuencia con la SEQ ID N.º 4 o la SEQ ID N.º 5. En una forma de realización, EndoS tiene la SEQ ID N.º 4 o la SEQ ID N.º 5. Preferiblemente, EndoS tiene las identidades de secuencia indicadas con la SEQ ID N.º 4.

SEQ ID n° 4:

MPSIDSLHYLSSENSKKEFKKEELSKAGQESQKVKEILAKAQQADKQAQELAKMKIPEKI PMKPLHGPLYGGYF
 RTWHDKTS DPTEKDKVNSMGELPKVEVDLAFIFHDWTKDYSLFWKELATKHVPKLNKQGT RVIRTI PWRFLAG
 GDNSGIAEDTSKYPNTPEGNKALAKAIVDEYVYKYNLDGLDVDVEHDSIPKVDKKEDTAGVERSIQVFEEIG
 KLI GPKGVDKSRLFIMDSTYMADKNPLIERGAPYINLLLVQVYGSQGEKGGWEPVSNRPEKTMEERWQGYSK
 YIRPEQYMI GFSFYEENAQEGNLWYDINSRKDEDKANGINTDI TGTRAERYARWOPKTGGVKGGIFSYAIDR
 DGVAHQPKKYAKQKEFKDATDNI FHS DYSVSKALKTVMLKDKSYDLIDEKDFPDKALREAVMAQVGT RKGDL
 40 ERFNGFLRLDNP AIQSLEGLNKFKKLAQLDLIGLSRI TKLDRSVLPANMKPGKDTLETVLETYKKDNKEEPA
 TIPVSLKVSGLTGLKELDL SGFDRETLAGLDAATLTSLEKVDISGNKLDLAPGTENRQIFDTMLSTISNHV
 GSNEQTVKFDKQKPTGHYPDTYGKTS LR L P VANEKVVDLQSQ LLEFGVTNQGTLINSEADYKAYQNHKIAGRS
 FVDSNYHYNNFKVSYENYTVKVT DSTLGTTT DKT LATDKEETYKVDFFS PADKTKAVHTAKVI VGDEKTM MV
 NLAEGATVIGGSADPVNARKVFDGQLGSETDNISLGWDSKQSIIFK LKEDGLIKHWRFFNDSARNPETTNKP
 IQEASLQIFNIKDYNDNLLENPNKFDDEKYWITVDTYSAQGERATAFSNTLNNITSKYWRVVFDTKGD RYS
 SPVPELQILGYPLFNADTIMKTVTTAKELSQQKDKFSQKMLDELKIKEMALETSLNSKIFDVTAINANAGV
 LKDCIEKRQLLKK

SEQ ID nº 5:

MGSSHHHHHHSSGLVPRGSHMPSIDSLHYLSENSKKEFKEELS KAGQESQKVKEI LAKAQQADKQAQELAKM
KI PEKI PMKPLHGPLYGGYFRTWHDKTS DPTEKDKVNSMGELPKEVDLAFI FHDWTKDYSLFWKELATKHVP
KLNKQGTRVIRTI PWRFLAGGDN SGIAEDTSKYPNTPEGNKALAKAI VDEYVYKYNLDGLDVDVEHDSI PKV
DKKEDTAGVERS IQVFEEIGKLI GPKGVDKSRLFIMDSTYMADKNPLI ERGAPYINLLLLVQVYGSQGEKGGW
EPVSNRPEKTMEERWQGY SKYIRPEQYMI GFS FYEENAQEGNLWYDINSRKDEDKANGINTDITGTRAERYA
RWQPKTGGVKGGI FSYAIDRDGVAHQPKKYAKQKEFKDATDNI FHS DYSVSKALKTVMLKDKSYDLI DEKDF
PDKALREAVMAQVGT RKGDLERFNGTLRLDNPAIQSLEGLNKFKKLAQLDLIGLSRITKLD RSVLEANMKPG
KDTLETVLETYKKDNKEEPATIPVSLKVSGLTGLKELDL SGFDRETLAGLDAATLTSLEKVDISGNKDLA
PGTENRQIFDTMLSTI SNHVGSNEQTVKFDKQKPTGHYPDTYGKTS LR LFPVANEKVLDLQSOLLFGTVTNQGT
LINSEADYKAYQNHKIAGRSFVDSNYHYNNFKVSYENYTVKVTDSTLGTTTTDKTLATDKEETYKVDFFS PAD
KTKAVHTAKVIVGDEKTMVMNLAEGATVIGGSADPVNARKVFDGQLGSETDNI SLGWDSKQSI I FKLKEDGL
IKHWRFFNDSARNPETTNKPIQEASLQIFNI KDYNLDNLLNPNKFDDEKYWITVDTYSAQGERATAFSNTL
NNITSKYWRVVFDTKGD RYSSPVPELQILGYPLPNADTIMKT VTTAKELSQQKDKFSQKMLDELKI KEMAL
ETSLNSKIFDVTAINANAGVLKDCIEKRQLLKK

- 5 [0042] En una forma de realización preferida, EndoH tiene al menos un 80 %, preferiblemente al menos un 90 %, más preferiblemente al menos un 95 % de identidad de secuencia con la SEQ ID N.º 6, más preferiblemente EndoH tiene un 100 % de identidad de secuencia con la SEQ ID N.º 6. En una forma de realización, EndoH tiene SEQ ID N.º 6. SEQ ID nº 6:

APAPVKQGPTSVAYVEVNNNSMLNVGKYTLADGGGNAFDVAVI FAANINYDTGKTAYLHFNENVQRVLDNA
VTQIRPLQQQGIVKVL LSVLGNHQGAGFANFP SQQAASAFKQLSDAVAKYGLDGVDFDDEYAEYGNNGTAQP
NDSSFVHLVTALRANMPDKI I SLYNIGPAASRLSYGGVDVSDKFDYAWNPPYGTWQVPGIALPKAQLSPA
10 EIGRTSRSTVADLARRTVDEGYGVYLTYNLDGGDRTADVSAFTRELYGSEAVRTP

- 15 [0043] En una forma de realización preferida, EndoF1 tiene al menos un 80 %, preferiblemente al menos un 90 %, más preferiblemente al menos un 95 % de identidad de secuencia con la SEQ ID N.º 7, más preferiblemente EndoF1 tiene un 100 % de identidad de secuencia con la SEQ ID N.º 7. En una forma de realización, EndoF1 tiene la SEQ ID N.º 7. SEQ ID nº 7:

AVTGTTKANIKLFSFTEVNDTNPLNLLNFTLKN SGKPLVDMVVLFSANINYDAANDKVFSNNPNVQHLLTN
RAKYLKPLQDKGIKVILSILGNHDRSGIANLSTARAKAFAQELKNTCDLYNLDGVVFFDDEYSAYQTPPPSGF
VTPSNNAARLAYETKQAMPNKLVTVYVYSRTSSFFTAVDGVNAGSYVDYAIH DYGGSYDLATNYPGLAKSG
MVMSSQEFNQGRYATAQALRNIVTKGYGGHMI FAMPNRSNFTSGQLPALKLIAKELYGDELVYSNTPYSKD
W

- 20 [0044] En una forma de realización preferida, EndoF2 tiene al menos un 80 %, preferiblemente al menos un 90 %, más preferiblemente al menos un 95 % de identidad de secuencia con la SEQ ID N.º 8, más preferiblemente EndoF2 tiene un 100 % de identidad de secuencia con la SEQ ID N.º 8. En una forma de realización, EndoF2 tiene la SEQ ID N.º 8. SEQ ID nº 8:

MAVNLSNLIAYKNSDHQISAGYYRTWRDSATASGNLPSMRWLPDSLDMVMVFPDYTPPENAYWNTLKTNYVF
YLHKRGTKVIITLGD LNSATTTGGQDSIGYSSWAKGIYDKWVGEYNLDGIDIDI ESSPSGATLTKEVAATKA
LSKYFGPKSGTGKTFVYDTNQNPTNFFIQTAPRYNYVFLQAYGRSTTNLT TVSGLYAPYI SMKQFLPGFSFY
EENGYPGNYWNDVRY PQNGTGRAYDYARWQPATGKKGGVFSYAIERDAPLTS SNDNTLRAPNFRVTKDLIKI
MNP

25

[0045] En una forma de realización preferida, EndoF3 tiene al menos un 80 %, preferiblemente al menos un 90 %, más preferiblemente al menos un 95 % de identidad de secuencia con la SEQ ID N.º 9, más preferiblemente EndoF3 tiene un 100 % de identidad de secuencia con la SEQ ID N.º 9. En una forma de realización, EndoF3 tiene la SEQ ID N.º 9. SEQ ID nº 9:

5 MATALAGSNGVCIAYYITDGRNPTFKLKDI PDKVDMVILFGLKYWSLQDFTTKLPGGTGMMGSFKSYKDLDTQ
 I RSLQSRGIKVLQNI DDDVSWQSSKPGGFASAAAYGDAI KSIVIDKWKLDGISLDIEHSGAKPNPIPTFFPGY
 AATGYNGWYSGSMAATPAFLNVI SELTKYFGTTAPNNKQLQIASGIDVYAWN KIMENFRNNFNFIQLQSYGA
 NVSRTQLMMNYATGTNKI PASKMVF GAYAEGGTNQANDVEVAKWPTPTQGA KGGMMI YTYNSNVSYANAVRDA
 VKN

[0046] En una forma de realización preferida, EfEndo18A tiene al menos un 80 %, preferiblemente al menos un 90 %, más preferiblemente al menos un 95 % de identidad de secuencia con la SEQ ID N.º 10, de la manera más preferible, EfEndo18A tiene un 100 % de identidad de secuencia con la SEQ ID N.º 10. En una forma de realización, EfEndo18A tiene SEQ ID N.º 10. SEQ ID nº 10:

10 ASTVTPKTVMYVEVNNHDFNNVVGKYTLAGTNQPAFDMGII FAANINYDTVNKKPYLYLNERVQQT LNEAETQ
 I RVPVQARGTKVLLSILGNHEGAGFANFPTYESADAFAAQLEQVVNTYHLDGIDFDDEYAEY GKNGT P QPNNS
 SFIWLLQALRNRLGNDKLITFYNI GPAAANS SANPQMS SLIDYAWN PYYSTWNP PQIAGMPASRLGASAVEV
 GVNQNLAAQYAKRTKAEQYGI YLMYNLPGKDS SAYI SAATQELYGRKTNYSPTVPTP

15 [0047] Estas secuencias preferidas para las endoglicosidasas individuales también son válidas para la enzima de fusión según la invención. Por lo tanto, para que EndoX sea EndoS, se prefiere que la secuencia de aminoácidos de EndoS sea como se ha definido antes en este documento. El experto es capaz de aplicar las secuencias proporcionadas anteriormente a la enzima de fusión según la fórmula (1).

20 [0048] En una forma de realización, la enzima según la invención comprende una secuencia de aminoácidos seleccionada entre SEQ ID N.º: 4 - SEQ ID N.º: 5, conectada mediante una secuencia de aminoácidos seleccionada entre SEQ ID N.º: 11 y SEQ ID N.º: 12 a otra secuencia de aminoácidos seleccionada entre SEQ ID N.º: 6 - SEQ ID N.º: 10, que tiene individualmente al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia con las SEQ ID
 25 individuales, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con respecto a cada una de SEQ ID N.º: 4 - SEQ ID N.º: 12. En una forma de realización preferida, estas identidades de secuencia se aplican a la combinación de SEQ ID en la enzima de fusión según la invención.

30 [0049] Preferiblemente, la enzima de la invención, que tiene las identidades de secuencia indicadas anteriormente con respecto a la SEQ ID N.º: 2, tiene actividad EndoS y EndoH. De la manera más preferible, la enzima según la invención tiene una identidad de secuencia del 100 % con la SEQ ID N.º: 2. En una forma de realización, la enzima según la invención comprende la SEQ ID N.º: 4 conectada a través de la SEQ ID N.º: 11 a
 35 la SEQ ID N.º: 6, tiene individualmente al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia con las SEQ ID individuales, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con respecto a la SEQ ID N.º: 4 , SEQ ID N.º: 11 y SEQ ID N.º: 6. En una forma de realización preferida, estas identidades de secuencia se aplican a la combinación de la SEQ ID N.º: 4, SEQ ID N.º: 11 y SEQ ID N.º: 6. Preferiblemente, la enzima de la invención,
 40 que tiene las identidades de secuencia indicadas anteriormente con respecto a la SEQ ID N.º: 1, tiene actividad EndoS y EndoH. De la manera más preferible, la enzima según la invención tiene una identidad de secuencia del 100 % con la SEQ ID N.º: 1. En una forma de realización, la enzima según la invención que comprende la SEQ ID N.º: 4 conectada a través de la SEQ ID N.º: 12 a la SEQ ID N.º: 6, que tiene individualmente al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de
 45 identidad de secuencia. % de identidad de secuencia con las SEQ ID individuales, tales como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con respecto a la SEQ ID N.º: 4, SEQ ID N.º: 12 y SEQ ID N.º: 6.

[0050] En una forma de realización preferida, la enzima de fusión según la invención tiene al menos un 50 % de
 50 identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % o, de la manera más preferible, un 100 % de identidad de secuencia con respecto a cualquiera de las SEQ ID N.º: 1, 2, 15, 18 y 21. Las ID de secuencia más preferidas se seleccionan de entre SEQ ID N.º: 1, 2 y 21.

[0051] En una forma de realización, la enzima de fusión según la invención tiene al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % o, de la manera más preferible, un 100 % identidad de secuencia con respecto a la SEQ ID N.º: 1. En una forma de realización, la enzima de fusión según la invención tiene al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % o, de la manera más preferible, un 100 % identidad de secuencia con respecto a la SEQ ID N.º: 2. En una forma de realización, la enzima de fusión según la invención tiene al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % o, de la manera más preferible, un 100 % identidad de secuencia con respecto a la SEQ ID N.º: 15. En una forma de realización, la enzima de fusión según la invención tiene al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % o, de la manera más preferible, un 100 % identidad de secuencia con respecto a la SEQ ID N.º: 18. En una forma de realización, la enzima de fusión según la invención tiene al menos un 50 % de identidad de secuencia, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % o, de la manera más preferible, un 100 % de identidad de secuencia con respecto a la SEQ ID N.º: 21.

[0052] En una forma de realización, EndoY se selecciona del grupo que consiste en EndoA, EndoBi, EndoBH, EndoBT, EndoCE, EndoD, EndoE, EfEndo18A, EndoF1, EndoF2, EndoF3, EndoH, EndoLL, EndoM, EndoOm y EndoT. Un grupo preferido de endoglicosidasas para usar como EndoY consiste en EndoE, EfEndo18A, EndoF1, EndoF2, EndoF3, EndoH, EndoS y EndoT, más preferiblemente en EndoF1, EndoF2, EndoF3, EfEndo18A y EndoH. En una forma de realización, al menos una de entre EndoX y EndoY, preferiblemente EndoY, se selecciona del grupo que consiste en EfEndo18A EndoF1 y EndoH.

[0053] En una forma de realización, EndoY es una endoglicosidasa capaz de escindir un glicano del tipo con alto contenido en manosa, tal como EndoA, EndoE, EfEndo18A, EndoF1, EndoH, EndoM o EndoT. Preferiblemente, la endoglicosidasa capaz de escindir un glicano del tipo alto en manosa se selecciona del grupo que consiste en EndoE, EfEndo18A, EndoF1, EndoH y EndoT, más preferiblemente se selecciona del grupo que consiste en EfEndo18A, EndoF1 y EndoH. De la manera más preferible, la endoglicosidasa capaz de escindir un glicano del tipo con alto contenido en manosa es EndoH.

[0054] EndoX es una endoglicosidasa capaz de escindir un glicano de tipo complejo, es decir, EndoS. Preferiblemente, EndoY es una endoglicosidasa que tiene una actividad diferente, preferiblemente una endoglicosidasa capaz de escindir un glicano del tipo con alto contenido en manosa.

[0055] Se prefiere especialmente que la enzima de fusión según la invención contenga dos endoglicosidasas distintas que difieran en su actividad endoglicosidasa, ya que dos actividades endoglicosidasa distintas pueden combinarse como tales en una única enzima. Por lo tanto, EndoX y EndoY preferiblemente tienen cada una una actividad endoglicosidasa distinta seleccionada entre la capacidad de hidrolizar glicanos con alto contenido en manosa, la capacidad de hidrolizar glicanos complejos y la capacidad de hidrolizar glicanos híbridos, más preferiblemente seleccionada entre la capacidad de hidrolizar glicanos con alto contenido en manosa y la capacidad de hidrolizar glicanos complejos. Preferiblemente, EndoY es una endoglicosidasa capaz de hidrolizar glicanos con alto contenido en manosa. Preferiblemente, la endoglicosidasa que es capaz de hidrolizar glicanos con alto contenido en manosa también es capaz de hidrolizar glicanos híbridos. Preferiblemente, la endoglicosidasa que es capaz de hidrolizar glicanos complejos es capaz de hidrolizar glicanos complejos biantenarios y/o triantenarios, más preferiblemente todos los glicanos complejos.

[0056] Por ejemplo, cuando EndoX es EndoS y EndoY es EndoH, la enzima de fusión resultante tiene actividad tanto EndoS como EndoH, y es capaz de cortar glicanos complejos en glicoproteínas (como anticuerpos) en la unidad central GlcNAc, dejando solo el residuo central de GlcNAc en la glicoproteína (actividad EndoS), así como el corte (escisión) de glicanos con alto contenido en manosa (actividad EndoH).

[0057] Sorprendentemente, ambas actividades de la enzima de fusión funcionan sin problemas a un pH de alrededor de 7 - 8, mientras que la EndoH monomérica requiere un pH en el rango de 5 - 6, o incluso un pH de 6 para funcionar de manera óptima.

[0058] En una forma de realización, EndoX y EndoY son dos endoglicosidasas distintas que difieren en el pH óptimo de al menos 1 unidad de pH, preferiblemente al menos 1,5 unidad de pH, de la manera más preferible, al menos 2 unidades de pH. El experto conoce el pH óptimo que pertenece a endoglicosidasas específicas. Dichas

enzimas de fusión pueden ser activas a un pH específico, que no es el pH óptimo de al menos una de EndoX y EndoY.

[0059] En una forma de realización preferida, EndoX es EndoS y EndoY se selecciona de entre EndoD, EndoH, EndoE, EfEndo18A, EndoT o EndoF1. De este modo, la enzima de fusión es capaz de hidrolizar glicanos complejos (actividad EndoS), así como de hidrolizar glicanos con alto contenido en manosa (actividad EndoD, EndoF1, EndoH, EndoE, EfEndo18A, EndoT o EndoF1). En una forma de realización, EndoX es EndoS y EndoY es preferiblemente EndoF1, EndoH o EfEndo18A, más preferiblemente EndoY es EndoH. De la manera más preferible, EndoX es EndoS y EndoY es EndoH.

[0060] En una forma de realización, EndoY se selecciona de entre EndoF1, EndoF2, EndoF3, EfEndo18A y EndoH.

[0061] Preferiblemente, EndoX es EndoS y EndoY se selecciona de entre EndoF1, EfEndo18A y EndoH.

[0062] En una forma de realización, EndoX es EndoS y EndoY es EndoF1 o EndoH.

[0063] En una forma de realización preferida, la enzima de fusión según la invención se selecciona del grupo formado por enzimas de estructura (1), en donde EndoX = EndoS y EndoY = EndoH; EndoX = EndoS y EndoY = EndoE; EndoX = EndoS y EndoY = EfEndo18A; EndoX = EndoS y EndoY = EndoT; EndoX = EndoS y EndoY = EndoF1. Más preferiblemente, las enzimas de fusión según la invención se seleccionan del grupo que consiste en enzimas con la estructura (1), en donde EndoX = EndoS y EndoY = EndoH; EndoX = EndoS y EndoY = EfEndo18A; EndoX = EndoS y EndoY = EndoF1. Aún más preferiblemente, las enzimas de fusión según la invención se seleccionan del grupo que consiste en enzimas con la estructura (1), en donde EndoX = EndoS y EndoY = EndoH; EndoX = EndoS y EndoY = EfEndo18A; EndoX = EndoS y EndoY = EndoF1. De la manera más preferible, las enzimas de fusión según la invención son una enzima de estructura (1), donde EndoX = EndoS y EndoY = EndoH.

Porción conectora

[0064] En la enzima según la invención, EndoX y EndoY están unidas preferiblemente mediante una porción conectora. En caso de que esté presente una porción conectora, $p = 1$. En caso de que no esté presente ninguna porción conectora, $p = 0$. Preferiblemente, $p = 1$. Las porciones conectoras para enzimas de fusión son conocidas en la técnica y se puede utilizar cualquier porción conectora adecuada, incluidas porciones conectoras flexibles y rígidas. Se puede encontrar más información en Chen et al., Adv. Drug Deliv. Rev.2013, 65, 1357-1369 y Fusion Protein Technologies for Biopharmaceuticals: Application and Challenges, Cap. 4: Fusion Protein Linkers: Effects on Production, Bioactivity, and Pharmacokinetics, 2013, John Wiley & Sons, Inc.

[0065] Preferiblemente, dicha porción conectora es una porción conectora flexible que permite que la proteína adyacente se mueva con relativa libertad.

[0066] En una forma de realización, la porción conectora, preferiblemente la porción conectora flexible, está compuesta de residuos amino y tiene una longitud de 1 a 100 residuos de aminoácidos, preferiblemente de 3 a 59, de 10 a 45 o de 15 a 40 residuos de aminoácidos, tal como 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40 residuos de aminoácidos.

[0067] En una forma de realización, la porción conectora, preferiblemente la porción conectora flexible, está compuesta de residuos amino como glicina, serina, histidina y/o alanina y tiene una longitud de 3 a 59 residuos de aminoácidos, preferiblemente de 10 a 45 o 15 a 40 residuos de aminoácidos, como 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40 residuos de aminoácidos.

[0068] La porción conectora comprende preferiblemente uno o más dominios flexibles que proporcionan flexibilidad a la porción conectora. Preferiblemente, uno o dos, más preferiblemente dos de dichos dominios flexibles están comprendidos en la porción conectora. Dichos dominios flexibles son conocidos en la técnica y normalmente están compuestos de glicina, serina y/o treonina. En una forma de realización, la porción conectora comprende al menos un residuo de glicina, serina y/o treonina. Preferiblemente, al menos el 40 % de los aminoácidos de la porción conectora se seleccionan de entre glicina, serina y treonina, más preferiblemente del 50 al 90 %, de la manera más preferible, del 70 al 85 % de los aminoácidos de la porción conectora se seleccionan de entre glicina, serina y treonina. En una forma de realización, la porción conectora no comprende treonina y los rangos anteriores se aplican a glicina y serina.

[0069] Los dominios flexibles adecuados específicos incluyen dominios GS (como $(G_4S)_n$, en donde n es un número entero en el rango de 1 - 10, preferiblemente 1 - 6, de la manera más preferible, 2 - 4), poli-G (tal como G_m , en donde m es un número entero en el rango de 1 - 30, preferiblemente 3 - 20, de la manera más preferible, 5 - 10), porciones conectoras GSAGSAAGSGEF, EGKSSGSGSESKST, PAS (porciones conectoras basadas en Pro, Ala, Ser; véase Schlapschy et al., Protein Eng Des Sel. 2013, 26, 489-501), y porciones conectoras de

polipéptido recombinante extendido (XTEN) (véase Podust et al., Protein Eng Des Sel. 2013, 26, 743-753). Los dominios GS, que consisten en tramos de residuos de glicina y serina, son los más preferidos. Por lo tanto, en una forma de realización, la porción conectora comprende uno o más dominios $(G_4S)_n$, preferiblemente uno o dos, de la manera más preferible, dos dominios.

5

[0070] Alternativa o adicionalmente, la porción conectora puede comprender uno o más dominios rígidos, tales como dominios que forman hélice α , tales como $(EAAAk)_o$ o $A(EAAAk)_oA$ (donde o es un número entero en el intervalo de 1 a 10, preferiblemente de 2 a 5, de la manera más preferible, 3 o 4), y dominios ricos en prolina, tal como $(XP)_q$ (donde X es cualquier aminoácido, preferiblemente seleccionado de entre alanina, lisina y glutamina, y q es un número entero en el intervalo 2 - 25, preferiblemente 5 - 17).

10

[0071] Opcionalmente, la porción conectora comprende una etiqueta para facilitar la purificación y/o detección como se conoce en la técnica, tal como una etiqueta Fc, una etiqueta FLAG, una etiqueta poly(His), etiqueta $(RP)_6R$, etiqueta HA y etiqueta Myc. Una etiqueta de este tipo también puede estar presente en otra parte de la porción conectora según la invención. Por tanto, en una forma de realización, la enzima de fusión según la invención comprende una etiqueta para facilitar la purificación y/o detección, tal como una etiqueta Fc, una etiqueta FLAG, una etiqueta poly(His), una etiqueta $(RP)_6R$, una etiqueta HA y una etiqueta Myc, más preferiblemente una etiqueta poly(His). En una forma de realización, la enzima de fusión según la invención comprende una porción conectora, es decir, $p = 1$, y la porción conectora comprende una etiqueta para facilitar la purificación y/o detección, tal como una etiqueta Fc, una etiqueta FLAG, una etiqueta poly(His), una etiqueta $(RP)_6R$, una etiqueta HA y una etiqueta Myc, más preferiblemente una etiqueta poly(His). La etiqueta puede estar situada en el extremo C de la porción conectora, en el extremo N de la porción conectora o puede estar incrustada en la porción conectora con aminoácidos adicionales a cada lado de la etiqueta. Se prefiere esta última conformación, especialmente cuando los dominios flexibles están ubicados a ambos lados de la etiqueta, ya que proporciona una accesibilidad óptima de la etiqueta para unirse a una matriz de afinidad.

15

20

25

[0072] En una forma de realización, la porción conectora tiene la estructura $(G_4S)_{n1}(H)_r(EF)_s(G_4S)_{n2}$, en donde $n1$ y $n2$ individualmente son números enteros en el rango de 1 - 10, preferiblemente 1 - 6, incluso más preferiblemente 2 - 4, de la manera más preferible, 3, y r es un número entero en el rango de 2 - 10, preferiblemente 4 - 8, de la manera más preferible, preferiblemente 6, y s = 0 o 1. En una forma de realización, la porción conectora tiene la estructura $(G_4S)_3(H)_6(G_4S)_3$, es decir, en donde $n1 = 3$, $n2 = 3$, $r = 6$ y $s = 0$ (aminoácidos 950 a 985 de SEQ ID N.º 2). En una forma de realización, la porción conectora tiene la estructura $(G_4S)_3(H)_6EF(G_4S)_3$, es decir, en donde $n1 = 3$, $n2 = 3$, $r = 6$ y $s = 1$ (aminoácidos 950 a 987 de SEQ ID N.º 1).

30

35

[0073] En una forma de realización preferida, la porción conectora tiene al menos un 80 %, preferiblemente al menos un 90 %, más preferiblemente al menos un 95 % de identidad de secuencia con la SEQ ID N.º 11 o 12, de la manera más preferible, la porción conectora tiene un 100 % de identidad de secuencia con la SEQ ID N.º 11 o SEQ ID N.º 12. En una forma de realización, la porción conectora tiene la SEQ ID N.º 11. En una forma de realización, la porción conectora tiene la SEQ ID N.º 12.

40

SEQ ID n° 11: GGGGSGGGGSGGGSHHHHHHEFGGGSGGGSGGGGS
 SEQ ID n° 12: GGGGSGGGGSGGGSHHHHHHHGGGGSGGGSGGGGS

[0074] La enzima de fusión según la invención se puede preparar mediante técnicas rutinarias conocidas en la técnica, tales como introducir un vector de expresión (por ejemplo, plásmido) que comprende la secuencia codificante de la enzima en una célula huésped (por ejemplo, *E. coli*) para la expresión, a partir del cual se puede aislar la enzima. Alternativamente, la enzima se produce mediante expresión transitoria en CHO. Un posible enfoque para la preparación y purificación de la enzima de fusión según la invención se proporciona en los ejemplos 1 - 4 y 16 - 24, y su funcionamiento se demuestra en los ejemplos 5, 6, 8, 13 - 15 y 25 - 37, en los que varias glicoproteínas, incluido el trastuzumab y el trastuzumab con alto contenido en manosa, se cortan de manera eficiente en un solo paso.

45

50

Enzima de fusión preferida

[0075] En una forma de realización especialmente preferida, la invención se refiere a una enzima de fusión que comprende las dos endoglicosidasas EndoS y EndoH. En un ejemplo particular, las dos endoglicosidasas EndoS y EndoH están conectadas mediante una porción conectora, preferiblemente una porción conectora $-(Gly_4Ser)_3-(His)_6-(Gly_4Ser)_3-$. La enzima de fusión según la invención también se denomina EndoSH. En una forma de realización, la enzima según la invención tiene al menos un 50 % de identidad de secuencia con la SEQ ID N.º: 1, preferiblemente al menos un 70 %, más preferiblemente al menos un 80 % de identidad de secuencia con la SEQ ID N.º: 1, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con la SEQ ID N.º: 1. Preferiblemente, la enzima de la invención, que tiene la identidad de secuencia indicada anteriormente con la SEQ ID N.º: 1, tiene actividad EndoS y EndoH. De la manera más preferible, la enzima según la invención tiene un 100 % de identidad de secuencia con la SEQ ID N.º: 1. En una forma de realización, la enzima según la invención tiene al menos un 50 % de identidad de secuencia con la SEQ ID N.º: 2, preferiblemente al menos un

55

60

65

70 %, más preferiblemente al menos un 80 % de identidad de secuencia con SEQ ID N.º: 2, tal como al menos un 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con la SEQ ID N.º: 2. Preferiblemente, la enzima de la invención, que tiene la identidad de secuencia indicada anteriormente con la SEQ ID N.º: 2, tiene actividad EndoS y EndoH. de la manera más preferible, la enzima según la invención tiene una identidad de secuencia del 100 % con la SEQ ID N.º: 2.

[0076] También se incluyen enzimas de fusión de EndoS y EndoH, en las que la porción conectora se reemplaza por otra porción conectora adecuada conocida en la técnica, donde dicha porción conectora puede ser rígida o flexible. Preferiblemente, dicha porción conectora es una porción conectora flexible que permite que los dominios proteicos adyacentes se muevan con relativa libertad entre sí. Preferiblemente, dicha porción conectora flexible está compuesta de residuos amino como glicina, serina, histidina y/o alanina y tiene una longitud de 3 a 59 residuos de aminoácidos, preferiblemente 10 a 45 o 15 a 40 residuos de aminoácidos, tales como 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40 residuos de aminoácidos, o de 20 a 38, de 25 a 37 o de 30 a 36 residuos de aminoácidos. Opcionalmente, la enzima de fusión está unida covalentemente a, o comprende, una etiqueta para facilitar la purificación y/o la detección como se conoce en la técnica, tal como una etiqueta Fc, una etiqueta FLAG, una etiqueta poly(His), una etiqueta HA y una etiqueta Myc. El corte de glicoproteínas se conoce en la técnica, por ejemplo, en Yamamoto, Biotechnol. Lett. 2013, 35, 1733, WO 2007/133855 o WO 2014/065661. La enzima según esta forma de realización tiene actividad tanto EndoS como EndoH, y es capaz de cortar glicanos en glicoproteínas (tales como anticuerpos) por la unidad central GlcNAc, dejando solo el residuo central de GlcNAc en la glicoproteína (actividad EndoS), así como también glicanos con alto contenido en manosa (actividad EndoH). Sorprendentemente, ambas actividades de la enzima de fusión funcionan sin problemas a un pH de alrededor de 7 - 8, mientras que la EndoH monomérica requiere un pH de 6 para funcionar de manera óptima.

[0077] Uso

[0078] Otro aspecto de la invención se refiere al uso de la enzima de fusión según la invención para cortar glicoproteínas, preferiblemente para cortar anticuerpos. El corte también puede denominarse desglicosilación y se define con más detalle a continuación, en el contexto del procedimiento según la invención. El uso según este aspecto puede ocurrir *in vitro* o *in vivo*.

Proceso de corte de glicoproteínas

[0079] La enzima de fusión según la invención es especialmente adecuada para el corte de glicoproteínas. Por tanto, en un aspecto adicional, la invención se refiere a un proceso para cortar glicoproteínas. El proceso según este aspecto puede ocurrir *in vitro* o *in vivo*. El corte de glicoproteínas se conoce en la técnica, por ejemplo, Yamamoto, Biotechnol. Lett. 2013, 35, 1733, WO 2007/133855 o WO 2014/065661. Las glicoproteínas, como los anticuerpos, normalmente contienen diferentes glicofomas, que requieren diferentes endoglicosidasas para eliminarse. Las enzimas de fusión de la invención son especialmente adecuadas para desglicosilar en un solo paso una glicoproteína que tiene dos cadenas de glicano diferentes. Así, en una forma de realización, la glicoproteína que se somete al proceso según la invención comprende al menos dos glicanos distintos, preferiblemente dos glicanos distintos. Preferiblemente, la glicoproteína comprende al menos un glicano con alto contenido en manosa y al menos un glicano complejo, más preferiblemente la glicoproteína comprende al menos un glicano con alto contenido en manosa, al menos un glicano híbrido y al menos un glicano complejo. El glicano complejo puede ser un glicano bi-, tri- o tetraantenario.

[0080] En una forma de realización especialmente preferida, la glicoproteína es un anticuerpo.

[0081] El proceso según el presente aspecto también puede denominarse proceso para modificar una glicoproteína. El proceso comprendió poner en contacto la glicoproteína con una enzima de fusión según la invención, para obtener una glicoproteína cortada. El proceso también puede denominarse proceso para cortar una glicoproteína o desglicosilación de una glicoproteína. El corte o desglicosilación de una glicoproteína se refiere a la eliminación de un glicano de dicha glicoproteína. La estructura exacta del glicano que se elimina puede variar dependiendo de la naturaleza exacta de las endoglicosidasas que están presentes en la enzima de fusión, pero el residuo central de GlcNAc se mantiene en la glicoproteína en todo momento. El experto apreciará qué enzima de fusión, es decir, qué combinación de endoglicosidasas, es adecuada para cortar qué patrón de glicosilación de la glicoproteína.

[0082] Con las endoglicosidasas convencionales, las glicoproteínas que contienen una combinación de un glicano con alto contenido en manosa y un glicano bi, tri o tetraantenario complejo requerirían dos enzimas distintas para el corte, lo que a menudo requiere diferentes condiciones de solución amortiguadora y rangos de pH. Estas glicoproteínas ahora se pueden cortar eficientemente en un solo paso, sin necesidad de aplicar intercambio de solución amortiguadora para lograr el pH óptimo, con la enzima de fusión según la invención. Así, en una forma de realización, la glicoproteína, preferiblemente el anticuerpo, comprende al menos un glicano con alto contenido en manosa y al menos un glicano bi-, tri- o tetraantenario complejo, más preferiblemente al menos

un glicano con alto contenido en manosa, al menos un glicano híbrido y/o complejo bi-, tri- o tetraantenario. Por ejemplo, la enzima de fusión en la que EndoX es EndoS, y EndoY se selecciona de entre EndoH, EndoE, EfEndo18A, EndoT o EndoF1, es adecuada para cortar una glicoproteína que comprende un complejo de glicano complejo unido a N y un glicano con alto contenido en manosa, para obtener una glicoproteína cortada que comprende solo el/los sustituyente(s) de N-acetilglucosamina central opcionalmente fucosilada.

[0083] El experto conoce las condiciones adecuadas para realizar el corte de glicoproteínas. Por ejemplo, el proceso se lleva a cabo en un medio y a una temperatura que sea eficaz para cortar las glicoproteínas. Normalmente, se pueden aplicar los medios y condiciones que se aplican para una de las enzimas endoglicosidasas individuales. Como el pH óptimo de las endoglicosidasas individuales puede diferir, el proceso se puede llevar a cabo en una forma de realización a un pH que es de 0,5 a 3 unidades de pH, preferiblemente de 1 a 2 unidades de pH, diferente del pH óptimo de una o ambas, preferiblemente una de EndoX y EndoY. Por ejemplo, en el caso de que EndoY sea EndoH, que tiene un pH óptimo de 5 - 6, el proceso se puede llevar a cabo a pH 7 - 8. En una forma de realización, el corte se realiza a un pH en el intervalo de 4 - 9, preferiblemente en el intervalo de 6 - 8, de la manera más preferible, en el intervalo de 7 - 8. Los inventores han descubierto sorprendentemente que la enzima de fusión según la invención en la que EndoX = EndoS y EndoY = EndoH es completamente operativa a un pH superior a 7, mientras que el rango de pH funcional para EndoH es de 5,0 a 6,0, con el pH óptimo a 5,5.

[0084] Además, los inventores han descubierto que la actividad de una endoglicosidasa particular en una proteína de fusión puede mostrar una mayor eficiencia de corte en comparación con la misma endoglicosidasa como una sola enzima.

[0085] El corte proporciona glicoproteínas cortadas, en las que todos los restos de glicano presentes en la glicoproteína original, independientemente de su tipo y glicofoma, se cortan y solo quedan los sustituyentes de N-acetilglucosamina central opcionalmente fucosilados. Dicho sustituyente central de N-acetilglucosamina opcionalmente fucosilado se une normalmente mediante un enlace N-glicosídico al átomo de nitrógeno de la amida en la cadena lateral de un aminoácido de asparagina de la glicoproteína, tal como N297 cuando la glicoproteína es un anticuerpo.

[0086] La glicoproteína cortada obtenida de este modo se puede utilizar según se considere adecuado. Por ejemplo, cuando la glicoproteína es el producto de interés, la glicoproteína cortada según la invención es homogénea con respecto a los patrones de glicosilación. Esto puede ser particularmente importante cuando la glicoproteína se usa como medicamento, ya que la eficacia terapéutica y/o la toxicidad pueden variar para diferentes glicofomas de la glicoproteína. Tales variaciones impredecibles en eficacia y toxicidad se erradican cuando se utiliza el proceso según la invención.

[0087] Alternativamente, la glicoproteína cortada se puede usar para una funcionalización adicional, tal como mediante la introducción de un resto de azúcar opcionalmente sustituido que se conoce en la técnica, por ejemplo, en van Geel et.al, Bioconjugate Chem, 2015, 26, 2233. La glicoproteína cortada se puede poner en contacto con un compuesto con la fórmula S-P, en la que S es un resto de azúcar opcionalmente sustituido y P es un nucleótido, en presencia de un catalizador adecuado, tal como una glicosiltransferasa o N-acetilglicosiltransferasa. La glicoproteína modificada así obtenida comprende un resto de azúcar S conectado al extremo no reductor del glicano cortado. Usando un resto de azúcar S sustituido, las posibilidades de modificación o funcionalización adicional de la glicoproteína a través de dicho sustituyente son infinitas. Esta secuencia de pasos de reacción encuentra un uso particular en la preparación de bioconjugados, tales como conjugados anticuerpo-fármaco. Dichos pasos son conocidos por el experto en la materia, por ejemplo, a partir de WO 2014/065661.

Ejemplos

[0088] Análisis RP-HPLC de anticuerpos monoclonales reducidos: Antes del análisis por RP-HPLC, las muestras se redujeron incubando una solución de 10 µg de IgG (modificada) durante 15 minutos a 37 °C con DTT 10 mM y Tris 100 mM, pH 8,0 en un volumen total de 50 µl. A la muestra reducida se le añadió una solución de ACN al 49 %, MQ al 49 % y ácido fórmico al 2 % (50 µl). La HPLC de fase inversa se realizó en una HPLC Agilent 1100 utilizando una columna ZORBAX Poroshell 300SB-C8 de 1 × 75 mM, 5 µm (Agilent Technologies) a 1 ml/min a 70 °C utilizando un gradiente lineal de 16,9 minutos del 25 al 50 % de solución amortiguadora B (con solución amortiguadora A = 90 % MQ, 10 % ACN, 0,1 % TFA y solución amortiguadora B = 90 % ACN, 10 % MQ, 0,1 % TFA).

[0089] Análisis por espectrometría de masas de anticuerpos monoclonales: Antes del análisis por espectrometría de masas, las IgG se trataron con DTT, que permite el análisis de cadenas ligeras y pesadas, o se trataron con Fabricator™ (disponible comercialmente en Genovis, Lund, Suecia), que permite el análisis del fragmento Fc/2. Para el análisis de las cadenas ligera y pesada, se incubó una solución de 20 µg de IgG (modificada) durante 5 minutos a 37 °C con DTT 100 mM en un volumen total de 4 µl. Si están presentes, las funcionalidades azida se reducen a aminas en estas condiciones. Para el análisis del fragmento Fc/2, se incubó una solución de 20 µg de

IgG (modificada) durante 1 hora a 37 °C con Fabricator™ (1.25 U/μL) en solución salina regulada con fosfato (PBS) a pH 6,6 en un volumen total de 10 μL. Después de la reducción o digestión con el Fabricator, las muestras se lavaron tres veces con milliQ utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Millipore), lo que dio como resultado un volumen de muestra final de aproximadamente 40 μL. A continuación, las muestras se analizaron mediante tiempo de vuelo con ionización por electroespray (ESI-TOF) en un AccuTOF de JEOL. Los espectros desconvolucionados se obtuvieron utilizando el software Magtran.

Ejemplo 1: clonación de la proteína de fusión EndoSH en el vector de expresión (pET22B)

[0090] Se obtuvo de Genscript un vector pET22B que contiene una secuencia codificante EndoS-(G₄S)₃-(His)₆-EF-(G₄S)₃-EndoH (EndoSH) (EndoSH se identifica mediante la SEQ ID N.º: 1) entre los sitios EcoRI-HindIII. La secuencia de ADN para la proteína de fusión EndoSH consiste en los residuos codificantes 48-995 de EndoS fusionados mediante una porción conectora N-terminal de glicina-serina (GS) a los residuos codificantes 41-313 de EndoH. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

Ejemplo 2: expresión en E. coli de la proteína de fusión EndoSH

[0091] La expresión de la proteína de fusión EndoSH (identificada por SEQ ID N.º: 1) comienza con la transformación del plásmido (pET22b-EndoSH) en células BL21. El siguiente paso es la inoculación de 500 ml de cultivo (medio LB + ampicilina) con células BL21. Cuando la DO₆₀₀ alcanzó 0,7, los cultivos se indujeron con IPTG 1 mM (500 μl de solución madre 1 M).

Ejemplo 3: purificación de la proteína de fusión EndoSH de E. coli

[0092] Después de inducción durante la noche a 16 °C, el cultivo se sedimentó mediante centrifugación. El sedimento se resuspendió en 40 ml de PBS y se incubó en hielo con 5 ml de lisozima (10 mg/ml) durante 30 minutos. Después de media hora, se añadieron 5 ml de Triton-X-100 al 10 % y se sonicaron (10 minutos) en hielo. Después de la sonicación, los restos celulares se eliminaron mediante centrifugación (10 minutos a 8000 xg) seguidax de filtración a través de un filtro con un diámetro de poro de 0,22 μm. Alternativamente, la lisis del sedimento que contiene EndoSH se puede realizar mediante prensa de French. En este caso, el sedimento se resuspende en 10 ml de PBS/G de sedimento. La suspensión celular se lisa tres veces bajo presión (20000-25000 psi) mediante una prensa de French usando Emulsiflex C3, Avestin. Después de la prensa de French, los restos celulares se eliminaron mediante centrifugación (20 minutos, 10 000 xg). El extracto/fracción soluble se cargó en una columna HisTrap HP de 5 ml (GE Healthcare). La columna se lavó primero con solución amortiguadora A (solución amortiguadora Tris 20 mM, imidazol 20 mM, NaCl 500 mM, pH 7,5). La proteína retenida se eluyó con solución amortiguadora B (Tris 20 mM, NaCl 500 mM, imidazol 250 mM, pH 7,5, 10 ml). Las fracciones se analizaron mediante SDS-PAGE en geles de poliacrilamida (12 %). Las fracciones que contenían proteína diana purificada se combinaron y la solución amortiguadora se intercambió con Tris 20 mM, pH 7,5 y NaCl 150 mM mediante diálisis realizada durante la noche a 4 °C. La proteína purificada se concentró hasta al menos 2 mg/ml usando una membrana Amicon Ultra-0.5, Ultracel-10 (Millipore). El producto se almacena a -80 °C antes de volverse a usar.

Ejemplo 4: expresión de CHO y purificación de la proteína de fusión EndoSH de CHO

[0093] EndoSH (identificada por SEQ ID N.º: 2) se expresó transitoriamente en células CHO K1 de Evitria (Zurich, Suiza) a una escala de 20 ml. El sobrenadante, que contenía la proteína de fusión EndoSH, se diluyó con solución amortiguadora de elución (2 ml, Tris 20 mM, NaCl 500 mM, imidazol 500 mM) y solución amortiguadora de unión (18 ml, Tris 20 mM, NaCl 500 mM, imidazol 5 mM, pH = 7.4) hasta una concentración final de imidazol de 10 mM. La mezcla se cargó en una columna de Ni-NTA (GE Healthcare) y el producto se eluyó siguiendo un protocolo de elución estándar. Las fracciones recogidas (5 ml) se analizaron en un gel SDS-PAGE (10 %). El producto que contenía la fracción se concentró parcialmente (~2 ml) y se dializó con solución amortiguadora TBS. La concentración de proteína, determinada mediante análisis con nanodrop, se fijó en 0,5 mg/ml.

Ejemplo 5: corte de trastuzumab por EndoSH

[0094] Se cortó trastuzumab (obtenido de Epirus biopharma (Utrecht, Países Bajos); 14 mg/ml) en solución amortiguadora Tris 25 mM, pH 8, utilizando una concentración de EndoSH de 0,1 o 1 % p/p. Las reacciones, 350 μg de trastuzumab (25 μl) y la cantidad adecuada de EndoSH se agitaron a 37 °C y se analizaron mediante análisis de EM en tiempo extra, de 1 a 3 horas. Las muestras se sometieron al tratamiento con Fabricator antes del análisis. Se observaron conversiones completas al producto cortado, que se corta hasta el residuo de azúcar GlcNAc central, después de 1 hora a 37 °C con EndoSH al 0,1 % p/p.

Ejemplo 6: corte de trastuzumab con alto contenido en manosa mediante la proteína de fusión EndoSH

[0095] Se cortó trastuzumab con glicanos con alto contenido en manosa (obtenido mediante expresión transitoria en células CHO K1 en presencia de kifunensina realizada por Evitria (Zurich, Suiza)) (14 mg/mL) en solución amortiguadora Tris 25 mM pH 8 usando una concentración de 0,1 o 1 % p/p de EndoSH. Las reacciones, 350 µg de trastuzumab con alto contenido en manosa (25 µl) y la cantidad adecuada de EndoSH se agitaron a 37 °C y se analizaron mediante análisis de EM en tiempo extra, de 1 a 3 horas. Las muestras se sometieron al tratamiento con Fabricator antes del análisis. Se observaron conversiones completas al producto cortado, que se corta hasta el residuo de azúcar GlcNAc central, después de 3 horas a 37 °C con EndoSH al 1 % p/p.

Ejemplo 7: expresión transitoria y purificación de cAC10

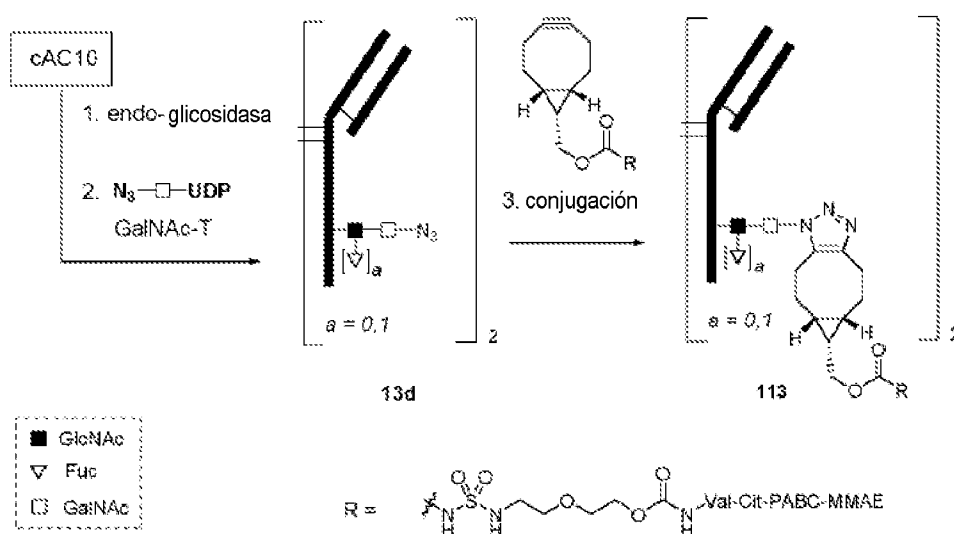
[0096] cAC10 se expresó transitoriamente en células CHO K1 por Evitria (Zurich, Suiza) a una escala de 5 L. El sobrenadante se purificó usando una columna XK 26/20 rellena con 50 ml de proteína A sefarosa. En una sola serie, se cargaron 5 litros de sobrenadante en la columna seguido de lavado con al menos 10 volúmenes de columna de Tris 25 mM, pH 7,5, NaCl 150 mM. La proteína retenida se eluyó con glicina 0,1 M, pH 2,7. El cAC10 eluido se neutralizó inmediatamente con Tris-HCl 1,5 M, pH 8,8, y se dializó con Tris 25 mM, pH 8,0. A continuación, la IgG se concentró hasta aproximadamente 20 mg/ml usando una unidad de ultrafiltración Vivaspin Turbo 15 (Sartorius) y se almacenó a -80 °C antes de volverse a usar.

Ejemplo 8: corte de cAC10 por EndoSH

[0097] El corte de glicanos de cAC10 (obtenido mediante expresión transitoria en células CHO K1 realizada por Evitria (Zurich, Suiza)) se realizó con la proteína de fusión EndoSH. Así, se incubó cAC10 (14,5 mg/ml) con EndoSH (1 p/p %) en Tris 25 mM, pH 7,5 con NaCl 150 mM durante aproximadamente 16 horas a 37 °C. La IgG cortada se dializó con 3 x 1 l de Tris-HCl 25 mM, pH 8,0. El análisis por espectrometría de masas de una muestra digerida con Fabricator mostró tres picos del fragmento Fc/2 perteneciente a un producto principal (masa observada de 24105 Da, aproximadamente el 80 % del fragmento Fc/2 total), correspondiente a cAC10 con sustituyentes en GlcNAc(Fuc) central, y dos productos menores (masas observadas de 23959 y 24233 Da, aproximadamente 5 y 15 % del fragmento Fc/2 total), correspondientes a cAC10 con sustituyentes en GlcNAc central y cAC10 con sustituyentes en GlcNAc(Fuc) central de lisina C-terminal.

Ejemplos 9 - 12: preparación de bioconjugado cAC10

[0098] Para demostrar que los anticuerpos cortados por la enzima de fusión según la invención pueden modificarse adicionalmente, se preparó el conjugado anticuerpo-fármaco 113 a partir del anticuerpo cortado del Ejemplo 8. El compuesto 99 se preparó mediante la activación del compuesto 58 como se describe y se prepara de acuerdo con el Ejemplo 50 de WO 2016/053107 (PCT/NL2015/050697). En el segundo paso, el cAC10 cortado se convirtió en el mAb 13d modificado con azido mediante la acción de His-TnGalNAcT en presencia de 6-N₃-GalNAc-UDP (disponible comercialmente en GlycoHub) como sustrato. La preparación de los bioconjugados de cAC10 se representa esquemáticamente a continuación:



Ejemplo 9: preparación del compuesto 100

[0099]

aproximadamente el 80 % del fragmento Fc/2 total), correspondiente al fragmento Fc/2 conjugado. El análisis RP-HPLC de la muestra reducida indicó una DAR media de 1,88.

Ejemplo 13: comparación de la eficiencia de corte de EndoS, EndoS2 y EndoSH en RNasaB a diferentes concentraciones

[0104] Primero, se preparan diluciones enzimáticas de las tres enzimas (EndoS y EndoS2 de Genovis, Lund, Suecia; EndoSH obtenida como en el Ejemplo 3) para obtener soluciones madre con 0,25 mg/mL (dil 1), 0,125 mg/mL (dil 2) y 0,025 mg/ml (dil 3). A continuación, se cargaron 12 viales con 2,5 µl de RNasa B (5 mg/ml) seguido de 0,5 µl de dilución 1-3 (dil 2 por duplicado) para cada enzima. Las reacciones se incubaron durante 30 minutos seguido de la adición de 36 µl de agua. De estas soluciones diluidas, se agregaron de 6 µl a 6 µl de solución amortiguadora de muestra para el análisis de SDS-PAGE. Se cargaron doce muestras en gel SDS-PAGE (4 por enzima) y se procesaron durante 70 minutos, se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua (véase la Figura 3 para ver el gel resultante). Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

Tabla 1. Porcentajes de corte (conversión) de RNasaB tras el tratamiento con diferentes endoglicosidasas a diferentes concentraciones de enzima.

[E] mg/mL	EndoS2	EndoS	EndoSH
0,25	0	0	45
0,125	0	0	53
0,125	0	0	50
0,025	0	0	66

Ejemplo 14: comparación de la eficiencia de corte de EndoS, EndoS2 y EndoSH en cAC10

[0105] Se trató cAC10 (4 mg, 20 mg/ml en Tris pH 8,0) con Fabricator™ (Genovis, Lund, Suecia, 4 µl, 66 U/ µl) durante 1 h a 37 °C. A continuación, el cAC10 escindido se intercambió con solución amortiguadora Tris pH 6,0 (50 mM, 3 x) utilizando una membrana Amicon Ultra-0,5, Ultracel-10 (Merck Millipore) hasta una concentración de 20 mg/ml. Posteriormente, se iniciaron tres reacciones que contenían cada una cAC10 (8,3 mg/ml) y una endoglicosidasa (EndoS y EndoS2 de Genovis, Lund, Suecia; EndoSH obtenida como en el Ejemplo 3) a 0,83 µg/ml en Tris pH 6,0 50 mM. Se tomaron muestras de 2 µl después de 15 min y 35 min, se diluyeron con 70 µl de MiliQ y se analizaron directamente mediante tiempo de vuelo con ionización por electroespray (ESI-TOF) en un AccuTOF de JEOL. Los porcentajes de conversión se calcularon en función de las intensidades de los picos de masa cortados y no cortados (véase el gráfico de la Figura 4).

Tabla 2. Porcentajes de corte (conversión) de cAC10 tras el tratamiento con diferentes endoglicosidasas en diferentes momentos.

Tiempo	EndoS2	EndoS	EndoSH
0 min.	0	0	0
15 min.	15	24	34
35 min.	36	73	73

Ejemplo 15: comparación de la eficacia de corte de EndoS, EndoS2 y EndoSH en trastuzumab con alto contenido en manosa

[0106] Se trató trastuzumab con alto contenido en manosa (1,3 mg, 8,8 mg/ml en Tris pH 8,0), obtenido mediante la expresión de trastuzumab en presencia de kifunensina, con Fabricator™ (3 µl, 66 U/µL) durante 1 h a 37 °C. A continuación, el trastuzumab con alto contenido en manosa escindido se intercambió con solución amortiguadora Tris a pH 6,0 (50 mM, 3 x) utilizando una membrana Amicon Ultra-0,5, Ultracel-10 (Merck Millipore) hasta una concentración de 20 mg/ml. Se iniciaron tres reacciones que contenían trastuzumab con alto contenido en manosa (10 mg/ml) y una endoglicosidasa (EndoS y EndoS2 de Genovis, Lund, Suecia; EndoSH obtenida como en el Ejemplo 3) a 4,4 µg/ml en Tris pH 6,0 50 mM. Se tomaron muestras de 2 µl después de 30, 60 y 120 minutos, se diluyeron con 70 µl de MiliQ y se analizaron directamente mediante tiempo de vuelo con ionización por electroespray (ESI-TOF) en un AccuTOF de JEOL. Los porcentajes de conversión se calcularon en función de las intensidades de los picos de masa cortados y no cortados (véase el gráfico de la Figura 5).

Tabla 3. Porcentajes de corte (conversión) de trastuzumab con alto contenido en manosa tras el tratamiento con diferentes endoglicosidasas en diferentes momentos.

Tiempo	EndoS2	EndoS	EndoSH
0 min.	0	0	0
30 min.	14	14	27
60 min.	25	14	40
120 min.	33	15	79

[0107] Estos experimentos demuestran que EndoSH es más eficaz para cortar trastuzumab y cAC10 con alto contenido en manosa que EndoS2, y EndoSH permite cortar otras glicoproteínas (por ejemplo, RNasaB), lo que no es posible con EndoS2, ya que la actividad está restringida al sitio N297. Por tanto, si un anticuerpo, por ejemplo, un anticuerpo monoclonal, tiene un alto nivel indeseable de manosa en un sitio de N-glicosilación diferente, EndoSH podría cortarlo, mientras que EndoS2 no puede.

Ejemplo 16: clonación de proteínas de fusión en un vector de expresión

[0108] Un vector pET22B que contiene EndoF3-(G₄S)₃-(His)₆-EF-(G₄S)₃-EfEndo18A (EndoF3-EfEndo18A), identificándose la secuencia codificante EndoF3-EfEndo18A mediante SEQ ID N.º: 13; o EndoF2-(G₄S)₃-His₆-EF-(G₄S)₃-EfEndo18A (EndoF2-EfEndo18A), identificándose la secuencia codificante EndoF2-EfEndo18A mediante SEQ ID N.º: 14; o EndoS-(G₄S)₃-His₆-EF-(G₄S)₃-EfEndo18A (EndoS-EfEndo18A), identificándose la secuencia codificante EndoS-EfEndo18A mediante SEQ ID N.º: 15; o EndoF3-(G₄S)₃-His₆-EF-(G₄S)₃-EndoF1 (EndoF3-EndoF1), identificándose la secuencia codificante EndoF3-EndoF1 mediante SEQ ID N.º: 16; o EndoF2-(G₄S)₃-His₆-EF-(G₄S)₃-EndoF1 (EndoF2-EndoF1), identificándose la secuencia codificante EndoF2-EndoF1 mediante SEQ ID N.º: 17; o EndoS-(G₄S)₃-His₆-EF-(G₄S)₃-EndoF1 (EndoS-EndoF1), identificándose la secuencia codificante EndoS-EndoF1 mediante SEQ ID N.º: 18; o EndoF3-(G₄S)₃-His₆-EF-(G₄S)₃-EndoH (EndoF3-EndoH), identificándose la secuencia codificante EndoF3-EndoH mediante SEQ ID N.º: 19; o EndoF2-(G₄S)₃-His₆-EF-(G₄S)₃-EndoH (EndoF2-EndoH), la secuencia codificante EndoF2-EndoH identificada por SEQ ID N.º: 20, entre los sitios NdeI-HindIII se obtuvo en Genscript, Piscataway, EE. UU.

[0109] La secuencia de ADN para la proteína de fusión EndoF3-EfEndo18A consiste en los residuos codificantes 40-329 de EndoF3 fusionados mediante una porción conectora N-terminal de glicina-serina (GS) a los residuos codificantes 42-314 de EfEndo18A. La secuencia de ADN se identifica mediante la SEQ ID N.º: 22. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

[0110] La secuencia de ADN de la proteína de fusión EndoF2-EfEndo18A consiste en los residuos codificantes 46-335 de EndoF2 fusionados mediante una porción conectora N-terminal de glicina-serina (GS) a los residuos codificantes 42-314 de EfEndo18A. La secuencia de ADN se identifica mediante la SEQ ID N.º: 23. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

[0111] La secuencia de ADN de la proteína de fusión EndoS-EfEndo18A consiste en los residuos codificantes 48-995 de EndoS fusionados mediante una porción conectora N-terminal de glicina-serina (GS) a los residuos codificantes 42-314 de EfEndo18A. La secuencia de ADN se identifica mediante la SEQ ID N.º: 24. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

[0112] La secuencia de ADN de la proteína de fusión EndoF3-EndoF1 consiste en los residuos codificantes 40-329 de EndoF3 fusionados mediante una porción conectora de glicina-serina (GS) N-terminal unida a los residuos codificantes 51-339 de EndoF1. La secuencia de ADN se identifica mediante la SEQ ID N.º: 25. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

[0113] La secuencia de ADN de la proteína de fusión EndoF2-EndoF1 consiste en los residuos codificantes 46-335 de EndoF2 fusionados mediante una porción conectora de glicina-serina (GS) N-terminal unida a los residuos codificantes 51-339 de EndoF1. La secuencia de ADN se identifica mediante la SEQ ID N.º: 26. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

[0114] La secuencia de ADN de la proteína de fusión EndoS-EndoF1 consiste en los residuos codificantes 48-995 de EndoS fusionados mediante una porción conectora de glicina-serina (GS) N-terminal unida a los residuos codificantes 51-339 de EndoF1. La secuencia de ADN se identifica mediante la SEQ ID N.º: 27. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

[0115] La secuencia de ADN para la proteína de fusión EndoF3-EndoH consiste en los residuos codificantes 40-329 de EndoF3 fusionados mediante una porción conectora N-terminal de glicina-serina (GS) a los residuos codificantes 41-313 de EndoH. La secuencia de ADN se identifica mediante la SEQ ID N.º: 28. La porción conectora glicina-serina (GS) comprende un formato -(G₄S)₃-(His)₆-EF-(G₄S)₃-, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

[0116] La secuencia de ADN de la proteína de fusión EndoF2-EndoH consiste en los residuos codificantes 46-335 de EndoF2 fusionados mediante una porción conectora N-terminal de glicina-serina (GS) a los residuos

codificantes 41-313 de EndoH. La secuencia de ADN se identifica mediante la SEQ ID N.º: 29. La porción conectora glicina-serina (GS) comprende un formato $-(G_4S)_3-(His)_6-EF-(G_4S)_3-$, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

5 [0117] La secuencia de ADN de la proteína de fusión EndoS-EndoH consiste en los residuos codificantes 48-995 de EndoS fusionados mediante una porción conectora N-terminal de glicina-serina (GS) a los residuos codificantes 41-313 de EndoH. La secuencia de ADN se identifica mediante la SEQ ID N.º: 30. La porción conectora glicina-serina (GS) comprende un formato $-(G_4S)_3-(His)_6-EF-(G_4S)_3-$, que permite espaciar las dos enzimas y, al mismo tiempo, introduce una etiqueta de purificación IMAC.

10 [0118] La secuencia de ADN de la proteína de fusión His₆-EndoS-EndoH consiste en los residuos codificantes 48-995 de EndoS fusionados directamente (es decir, sin porción conectora (GS)) a los residuos codificantes 41-313 de EndoH. La secuencia de ADN se identifica mediante la SEQ ID N.º: 31. La etiqueta His₆ permite la purificación mediante purificación IMAC.

15 [0119] La secuencia de ADN para los residuos codificantes 33-421 de His₆-TnGalNacT con una etiqueta N-terminal His₆ se identifica con la SEQ ID N.º: 32. La etiqueta His₆ permite la purificación mediante purificación IMAC.

20 *Ejemplo 17: expresión de proteína de fusión en E. coli a pequeña escala*

[0120] La expresión de las proteínas de fusión EndoF3-EndoF1 (SEQ ID N.º: 16), EndoS-EndoF1 (SEQ ID N.º: 18), EndoF2-EndoH (SEQ ID N.º: 20) comienza con la transformación del plásmido en células BL21 (DE3). El siguiente paso es la inoculación de 50 ml de cultivo (medio LB + ampicilina; 100 g/ml) con células BL21(DE3). En caso de His₆-EndoS-EndoH, se utilizó kanamicina (50 µg/ml). Cuando la DO₆₀₀ alcanzó un valor de 0,5-0,7, los cultivos se indujeron con IPTG 1 mM (50 µl de solución madre 1 M). Las expresiones de EndoF3-EndoF1 (SEQ ID N.º: 16) y EndoF2-EndoH (SEQ ID N.º: 20) se repitieron a gran escala como se describe en los Ejemplos 19 y 21.

30 *Ejemplo 18: purificación a pequeña escala de proteína de fusión de E. coli mediante NiNTA*

[0121] Después de inducción durante la noche a 16 °C, los cultivos de las expresiones EndoF3-EndoF1 (SEQ ID N.º: 16), EndoS-EndoF1 (SEQ ID N.º: 18), EndoF2-EndoH (SEQ ID N.º: 20) se sedimentaron mediante centrifugación. Los sedimentos se resuspendieron en 3-8 ml de PBS y se sonicaron con Sonopuls Mini20, Bandelin (usando micropunta MS 2,5) al 70 % (3 x 1 min) en hielo. Después de la sonicación, los restos celulares se eliminaron mediante centrifugación (10 min, 10 000 x g). El extracto soluble se cargó en una columna de Ni sefrosa hecha a mano (obtenida de ThermoFisher Scientific y Ni sefrosa de GE Healthcare). La columna se lavó primero con solución amortiguadora A (solución amortiguadora Tris 20 mM, imidazol 5 mM, NaCl 500 mM, pH 7,5). La proteína retenida se eluyó con solución amortiguadora B (Tris 20 mM, NaCl 500 mM, imidazol 250 mM, pH 7,5, 5 ml). Las fracciones se analizaron mediante SDS-PAGE en geles de poliacrilamida (12 %). Las fracciones que contenían la proteína diana purificada se combinaron y la solución amortiguadora se intercambió por TBS pH 7,5 mediante diálisis realizada durante la noche a 4 °C. Los rendimientos se muestran en la Tabla 4. Las proteínas se congelaron rápidamente y se almacenaron a -80 °C antes de volverse a usar.

45 Tabla 4. Rendimientos para las purificaciones a pequeña escala de proteínas de fusión de *E. coli* mediante NiNTA.

Proteína	Rendimiento (mg)
EndoF3-EndoF1	0,36
EndoS-EndoF1	2,7
EndoF2-EndoH	0,39

Ejemplo 19: expresión a gran escala de proteína de fusión en E. coli

50 [0122] La expresión de las proteínas de fusión EndoF3-EndoF1 (SEQ ID N.º: 16), EndoF2-EndoH (SEQ ID N.º: 20), EndoS-EfEndo18A (SEQ ID N.º: 15), EndoF2-EndoF1 (SEQ ID N.º: 17) y EndoF3 -EndoH (SEQ ID N.º: 19) comenzó con la transformación del plásmido en células BL21(DE3). El siguiente paso fue la inoculación de 500 ml de cultivo (medio LB + ampicilina; 100 µg/ml) con células BL21(DE3). Cuando la DO₆₀₀ alcanzó 0,5-0,7 los cultivos se indujeron con IPTG 1 mM (500 µl de solución madre 1 M).

55 *Ejemplo 20: purificación a gran escala de proteína de fusión de E. coli mediante NiNTA*

[0123] Después de inducción durante la noche a 16 °C, los cultivos de proteínas EndoS-EfEndo18A (SEQ ID N.º: 15), EndoF2-EndoF1 (SEQ ID N.º: 17), EndoF3-EndoH (SEQ ID N.º: 19), EndoF3-EndoF1 (SEQ ID N.º: NO: 16) y EndoF2-EndoH (SEQ ID N.º: 20) se sedimentaron mediante centrifugación. Los sedimentos se resuspendieron en 10 ml de PBS/G de sedimento. La suspensión celular se lisó tres veces bajo presión (20000-25000 psi) mediante una prensa de French usando Emulsiflex C3, Avestin. Después de la prensa de French, los restos

celulares se eliminaron mediante centrifugación (20 minutos a 10 000 xg). El extracto soluble se cargó en una columna HisTrap HP de 5 ml (GE Healthcare). La columna se lavó primero con solución amortiguadora A (solución amortiguadora Tris 20 mM, imidazol 5 mM, NaCl 500 mM, pH 7,5). La proteína retenida se eluyó con solución amortiguadora B (Tris 20 mM, NaCl 500 mM, imidazol 250 mM, pH 7,5, 10 ml). Las fracciones se analizaron mediante SDS-PAGE en geles de poliacrilamida (12 %). Las fracciones que contenían proteína diana purificada se combinaron y la solución amortiguadora se intercambió con Tris pH 7,5 mediante diálisis realizada durante la noche a 4 °C. Los rendimientos se muestran en la tabla 5 a continuación. Las proteínas se congelaron rápidamente y se almacenaron a -80 °C antes de volverse a usar.

Tabla 5. Rendimientos para las purificaciones a gran escala de proteínas de fusión de *E. coli* mediante NiNTA.

Proteína	Rendimiento (mg)
EndoS-EfEndo18A	55
EndoF2-EndoF1	22,1
EndoF3-EndoH	28
EndoF3-EndoF1	5,1
EndoF2-EndoH	3,5

Ejemplo 21: purificación de proteína de fusión de E. coli mediante SEC

[0124] Para EndoF3-EndoF1 (SEQ ID N.º: 16), EndoF2-EndoF1 (SEQ ID N.º: 17), EndoF3-EndoH (SEQ ID N.º: 19) y EndoF2-EndoH (SEQ ID N.º: 20) la purificación NiNTA, que se describe en el ejemplo 20, fue seguida de cromatografía de exclusión por tamaño (SEC) para aislar el monómero. Se instaló una Superdex 75 10/300 GL en el purificador Akta. La columna se enjuagó con MilliQ (20 ml) seguido de acondicionamiento con TBS pH 7,5 (25 ml, 0,8 ml/min). Se cargaron aproximadamente 1 - 3 mg de proteína purificada con NiNTA y se procesaron a 0,8 ml/min usando TBS pH 7,5. Se recogió la proteína monomérica y se analizaron las fracciones mediante SDS-PAGE en geles de poliacrilamida (12 %) o en masa en AccuTOF. Los rendimientos se muestran a continuación en la tabla 6. Las proteínas se congelaron rápidamente y se almacenaron a -80 °C antes de volverse a usar.

Tabla 6. Resumen de la cantidad de proteína de fusión de endoglicosidasa purificada con NiNTA que se cargó en la columna SEC y los rendimientos de la fracción monomérica.

Proteína	Cantidad cargada (mg)	Rendimiento (mg)
EndoF2-EndoF1	1,75	0,10
EndoF3-EndoH	0,80	0,02
EndoF3-EndoF1	2,30	0,79
EndoF2-EndoH	1,40	0,12

Ejemplo 22: clonación de la proteína de fusión His₆-EndoS-EndoH (sin porción conectora GS) en el vector de expresión pET28B

[0125] Un vector pET28B que contenía His₆-EndoS-EndoH (His₆-EndoSH sin porción conectora GS), con la secuencia de codificación de His₆-EndoS-EndoH identificada por la SEQ ID N.º: 21, entre los sitios NcoI-HindIII se obtuvo en Genscript Piscataway, EE. UU.

[0126] La secuencia de ADN de la proteína de fusión His₆-EndoSH codifica una etiqueta de purificación IMAC N-terminal y un sitio de escisión de trombina fusionado a los residuos codificantes 48-995 de EndoS fusionados a los residuos codificantes 41-313 de EndoH.

Ejemplo 23: expresión a pequeña escala de E. coli de la proteína de fusión His₆-EndoS-EndoH (sin porción conectora GS)

[0127] La expresión de la proteína de fusión His₆-EndoS-EndoH (SEQ ID N.º: 21) comienza con la transformación del plásmido en células BL21(DE3). El siguiente paso es la inoculación de 50 ml de cultivo (medio LB + kanamicina; 50 g/ml) con células BL21(DE3). Cuando la DO₆₀₀ alcanzó un valor de 0,5, el cultivo se indujo con IPTG 1 mM (50 µl de solución madre 1 M).

Ejemplo 24: purificación a pequeña escala de proteína de fusión de E. coli mediante NiNTA

[0128] Después de inducción durante la noche a 16 °C, el cultivo de la expresión del Ejemplo 23 se sedimentó mediante centrifugación. El sedimento se resuspendió en 7 ml de PBS y se sonicó con Sonopuls Mini20, Bandelin (usando micropunta MS 2,5) al 70 % (3 x 1 min) en hielo. Después de la sonicación, los restos celulares se eliminaron mediante centrifugación (10 min, 10 000 x g). El extracto soluble se cargó en una columna de Ni sefarosa hecha a mano (obtenida de ThermoFisher Scientific y Ni sefarosa de GE Healthcare). La columna se lavó primero con solución amortiguadora A (solución amortiguadora Tris 20 mM, imidazol 5 mM, NaCl 500 mM, pH 7,5). La proteína retenida se eluyó con solución amortiguadora B (Tris 20 mM, NaCl 500 mM, imidazol 250 mM, pH 7,5, 5 ml). Las fracciones se analizaron mediante SDS-PAGE en geles de poliacrilamida (12 %). Las

fracciones que contenían proteína diana purificada se combinaron y la solución amortiguadora se intercambi6 por TBS pH 7,5 mediante di6lisis realizada durante la noche a 4 °C. El rendimiento despu6s de la di6lisis fue de 9 mg. El producto se congel6 r6pidamente y se almacen6 a -80 °C antes de volverse a usar.

5 *Ejemplo 25: comparaci6n de la eficacia de corte de EndoSH, EndoF3-EndoH, EndoS-EfEndo18A y EndoF2-EndoF1 en trastuzumab con alto contenido en manosa*

[0129] Se trat6 trastuzumab con alto contenido en manosa (0,7 ml, 6,0 mg, 8,8 mg/ml en Tris pH 8,0) con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuaci6n, el trastuzumab con alto contenido en manosa digerido con Fabricator™ se dividi6 en tres porciones iguales y se intercambi6 la soluci6n amortiguadora por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, respectivamente. El intercambio de soluci6n amortiguadora se realiz6 utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentraci6n final de 10 mg/ml. Para EndoSH (identificada por la SEQ ID N.º: 1), EndoF3-EndoH (identificada por la EQ ID N.º: 19), EndoS-EfEndo18A (identificada por la SEQ ID N.º: 15) y EndoF2-EndoF1 (identificada por la SEQ ID N.º: 17), se prepararon series de diluciones de 10, 50 y 250 nM en cada una de las soluciones amortiguadoras de reacci6n mencionadas anteriormente. Las reacciones se iniciaron a6nadiendo 2 µl de trastuzumab con alto contenido en manosa (10 mg/mL) digerido con Fabricator™ a 2 µl de la prote6na de fusi6n de endoglicosidasa diluida en la soluci6n amortiguadora correspondiente, lo que da como resultado una concentraci6n final de 5 mg/mL de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con endoglicosidasa 5, 25 y 125 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Las reacciones se neutralizaron mediante la adici6n de 16 µl de soluci6n amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol seguida de una incubaci6n durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se ti6nieron en coomassie coloidal durante la noche y finalmente se quit6 la tinci6n en agua. Los porcentajes de conversi6n se calcularon bas6ndose en el escaneo del gel SDS-PAGE con un esc6ner plano normal y la cuantificaci6n con una herramienta de software (CLIQS v1.1).

Tabla 7. Porcentajes de corte (conversi6n) de trastuzumab con alto contenido en manosa digerido con Fabricator™ tras el tratamiento de varias prote6nas de fusi6n de endoglicosidasas a pH 4,5, 6,0 y 7,5.

pH	Enzima (nM)	EndoSH	EndoF3-EndoH	EndoS-EfEndo18A	EndoF2-EndoF1
4,5	5	23*	26*	26*	26*
25	35	31	38	26*	
125	60	60	61	83	
6,0	5	27*	25*	27*	28*
25	36	37	31	32	
125	64	57	69	61	
7,5	5	24*	26*	24*	26*
25	29*	31	31	30	
125	61	53	58	58	

La conversi6n se calcul6 utilizando la siguiente f6rmula: conversi6n (%) = 100 × (Fc/2_{cortado}) / (Fc/2_{cortado} + Fc/2_{glicosilado}). *Nota: Los resultados son dif6ciles de interpretar con precisi6n debido a la se6al de fondo a la altura del trastuzumab con alto contenido en manosa digerido con Fabricator™ sin tratar. No se aplic6 ninguna correcci6n de fondo para la cuantificaci6n de la conversi6n.

30

Ejemplo 26: comparaci6n de la eficacia de corte de EndoSH, EndoF3-EndoH, EndoS-EfEndo18A y EndoF2-EndoF1 en trastuzumab

[0130] Trastuzumab (obtenido de Epirus biopharma (Utrecht, Pa6ses Bajos); 287 µl, 6,0 mg, 21 mg/ml en Tris pH 8,0), se trat6 con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuaci6n, el trastuzumab digerido con Fabricator™ se dividi6 en tres porciones iguales y la soluci6n amortiguadora se intercambi6 por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, respectivamente. El intercambio de soluci6n amortiguadora se realiz6 utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentraci6n final de 10 mg/ml. Para EndoSH (identificada por la SEQ ID N.º: 1), EndoF3-EndoH identificada por la SEQ ID N.º: 19), EndoS-EfEndo18A (identificada por la SEQ ID N.º: 15) y EndoF2-EndoF1 (identificada por la SEQ ID N.º: 17), se prepararon diluciones de 50 y 250 nM en cada una de las soluciones amortiguadoras mencionadas anteriormente. Las reacciones se iniciaron a6nadiendo 5 µl de trastuzumab digerido con Fabricator™ (10 mg/ml) en 5 µl de la prote6na de fusi6n de endoglicosidasa diluida en la soluci6n amortiguadora correspondiente, lo que da como resultado una concentraci6n final de 5 mg/ml de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con endoglicosidasa 25 y 125 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacci6n se tom6 una muestra (4 µl) y se a6nadi6 a 16 µl de soluci6n amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol y se incub6 durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se ti6nieron en coomassie coloidal durante la noche y

45

finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

5 Tabla 8. Porcentajes de corte (conversión) de trastuzumab digerido con Fabricator™ tras el tratamiento de varias proteínas de fusión de endoglicosidasas a pH 4,5, 6,0 y 7,5.

pH	Enzima (nM)	EndoSH	EndoF3-EndoH	EndoS-EfEndo18A	EndoF2-EndoF1
4,5	25	29*	29*	29*	29*
125	28*	40	29*	50	
6,0	25	72	28*	77	28*
125	85	31	84	42	
7,5	25	81	27*	82	26*
125	87	31	86	41	

La conversión se calculó utilizando la siguiente fórmula: conversión (%) = 100 × (Fc/2_{cortado}) / (Fc/2_{cortado} + Fc/2_{glicosilado}). *Nota: Los resultados son difíciles de interpretar con precisión debido a la señal de fondo a la altura del trastuzumab digerido con Fabricator™ sin tratar. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

Ejemplo 27: corte de trastuzumab mediante EndoF3-EndoH (comparativo)

10 [0131] Para demostrar que se puede lograr una conversión completa para el corte de trastuzumab mediante EndoF3-EndoH (identificada por la SEQ ID N.º: 19), se incubó trastuzumab digerido con Fabricator™ (4 µl, 40 µg, 10 mg/ml en citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM) con citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM (8,73 µl) y EndoF3-EndoH (7,27 µl), 4 µg, 0,55 mg/ml en citrato de sodio 50 mM (pH 4,5 con NaCl 150 mM) durante 60 minutos a 37 °C. El análisis por espectrometría de masas de una muestra digerida con Fabricator mostró un pico principal del fragmento Fc/2 (masa observada 24139 Da, aproximadamente el 95 % del fragmento Fc/2 total), correspondiente al trastuzumab con sustituyentes en la GlcNAc(Fuc) central.

Ejemplo 28: corte de trastuzumab mediante EndoF2-EndoF1 (comparativo)

20 [0132] Para demostrar que se puede lograr una conversión completa para el corte de trastuzumab mediante EndoF2-EndoF1 (identificada por la SEQ ID N.º: 17), se incubó trastuzumab digerido con Fabricator™ (4 µl, 40 µg, 10 mg/ml en citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM) en citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM (11,06 µl) y EndoF2-EndoF1 (4,94 µl), 4 µg, 0,81 mg/ml en citrato de sodio 50 mM (pH 4,5 con NaCl 150 mM) durante 60 minutos a 37 °C. El análisis por espectrometría de masas de una muestra digerida con Fabricator mostró un pico principal del fragmento Fc/2 (masa observada 24139 Da, aproximadamente el 95 % del fragmento Fc/2 total), correspondiente al trastuzumab con sustituyentes en la GlcNAc(Fuc) central.

Ejemplo 29: comparación de la eficacia de corte de EndoSH, EndoF3-EndoF1, EndoS-EndoF1, EndoF2-EndoH en trastuzumab con alto contenido en manosa

30 [0133] Se trató trastuzumab con alto contenido en manosa (0,7 ml, 6,0 mg, 8,8 mg/ml en Tris pH 8,0) con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuación, se dividió trastuzumab con alto contenido en manosa digerido con Fabricator™ en tres porciones iguales y se intercambió la solución amortiguadora por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, respectivamente. El intercambio de solución amortiguadora se realizó utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentración final de 10 mg/ml. EndoSH (identificada por la SEQ ID N.º: 1), EndoF3-EndoF1 (identificada por la SEQ ID N.º: 16), EndoS-EndoF1 (identificada por la SEQ ID N.º: 18) y EndoF2-EndoH (identificada por la SEQ ID N.º: 20) se diluyeron a una concentración de 250 nM en cada una de las soluciones amortiguadoras de reacción mencionadas anteriormente. Las reacciones se iniciaron añadiendo 2 µl de trastuzumab con alto contenido en manosa (10 mg/ml) digerido con Fabricator™ en 2 µl de la proteína de fusión de endoglicosidasa (250 nM) en la solución amortiguadora correspondiente, lo que da como resultado una concentración final de 5 mg/ml de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con endoglicosidasa 125 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Las reacciones se inactivaron mediante la adición de 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol seguida de una incubación durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

50 Tabla 9. Porcentajes de corte (conversión) de trastuzumab con alto contenido en manosa digerido con Fabricator™ tras el tratamiento de varias proteínas de fusión de endoglicosidasas (125 nM) a pH 4,5, 6,0 y 7,5.

pH	EndoSH	EndoF3-EndoF1	EndoS-EndoF1	EndoF2-EndoH
----	--------	---------------	--------------	--------------

4,5	61	23*	60	33
6,0	69	19*	73	28*
7,5	67	17*	72	25*

La conversión se calculó utilizando la siguiente fórmula: conversión (%) = $100 \times (Fc/2_{cortado}) / (Fc/2_{cortado} + Fc/2_{glicosilado})$. *Nota: Los resultados son difíciles de interpretar con precisión debido a la señal de fondo a la altura del trastuzumab con alto contenido en manosa digerido con Fabricator™ sin tratar. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

Ejemplo 30: comparación de la eficacia de corte de EndoSH, EndoF3-EndoH y EndoS-EfEndo18A, EndoF2-EndoF1 en trastuzumab

5 [0134] Se trató trastuzumab (obtenido en Epirus biopharma (Utrecht, Países Bajos); 287 µl, 6,0 mg, 21 mg/ml en Tris pH 8,0), con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuación, el trastuzumab digerido con Fabricator™ se dividió en tres porciones iguales y la solución amortiguadora se intercambió por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, respectivamente. El intercambio de solución amortiguadora se realizó utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentración final de 10 mg/ml. EndoSH (identificada por la SEQ ID N.º: 1), EndoF3-EndoF1 (identificada por la SEQ ID N.º: 16), EndoS-EndoF1 (identificada por la SEQ ID N.º: 18) y EndoF2-EndoH (identificada por la SEQ ID N.º: 20) se diluyeron a una concentración de 250 nM en cada una de las soluciones amortiguadoras de reacción mencionadas anteriormente. Las reacciones se iniciaron añadiendo 5 µl de trastuzumab digerido con Fabricator™ (10 mg/mL) a 5 µl de la proteína de fusión de endoglicosidasa (250 nM) en la solución amortiguadora correspondiente, lo que da como resultado una concentración final de 5 mg/mL de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con endoglicosidasa 125 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacción, se tomó una muestra (4 µl) y se añadió a 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol y se incubó durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

25 Tabla 10. Porcentajes de corte (conversión) de trastuzumab tras el tratamiento con diferentes endoglicosidasas (125 nM) a pH 4,5, 6,0 y 7,5.

pH	EndoSH	EndoF3-EndoF1	EndoS-EndoF1	EndoF2-EndoH
4,5	22*	19*	23*	20*
6,0	77	16*	79	16*
7,5	80	14*	83	13*

La conversión se calculó utilizando la siguiente fórmula: conversión (%) = $100 \times (Fc/2_{cortado}) / (Fc/2_{cortado} + Fc/2_{glicosilado})$. *Nota: Los resultados son difíciles de interpretar con precisión debido a la señal de fondo a la altura deltrastuzumab digerido con Fabricator™ sin tratar. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

Ejemplo 31: comparación de la eficacia de corte de proteínas de fusión EndoS-EndoH con y sin porción conectora en trastuzumab con alto contenido en manosa

30 [0135] Se trató trastuzumab con alto contenido en manosa (0,7 ml, 6,0 mg, 8,8 mg/ml en Tris pH 8,0) con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuación, se dividió trastuzumab con alto contenido en manosa digerido con Fabricator™ en tres porciones iguales y se intercambió la solución amortiguadora por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, respectivamente. El intercambio de solución amortiguadora se realizó utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentración final de 10 mg/ml. Para EndoSH (identificada por la SEQ ID N.º: 1) e His₆-EndoSH (His₆-EndoS-EndoH sin porción conectora GS; identificada por la SEQ ID N.º: 21) se prepararon series de diluciones de 10, 50 y 250 nM en cada una de las soluciones amortiguadoras de reacción mencionadas anteriormente. Las reacciones se iniciaron añadiendo 2 µl de trastuzumab con alto contenido en manosa digerido con Fabricator™ (10 mg/mL) en 2 µl de la proteína de fusión de endoglicosidasa diluida en la solución amortiguadora correspondiente, lo que da como resultado una concentración final de 5 mg/mL de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con endoglicosidasa 5, 25 y 125 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Las reacciones se inactivaron mediante la adición de 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol seguida de una incubación durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

50

Tabla 11. Porcentajes de corte (conversión) de trastuzumab con alto contenido en manosa digerido con Fabricator™ tras el tratamiento con proteínas de fusión EndoS-EndoH con y sin porción conectora a pH 4,5, 6,0 y 7,5.

pH	Enzima (nM)	EndoSH (con porción conectora GS)	His ₆ -EndoSH (sin porción conectora GS)
4,5	5	23*	23*
	25	35	31
	125	60	63
6,0	5	27*	28*
	25	36	42
	125	64	66
7,5	5	24*	28*
	25	29*	38
	125	61	63

La conversión se calculó utilizando la siguiente fórmula: conversión (%) = $100 \times (Fc/2_{cortado}) / (Fc/2_{cortado} + Fc/2_{glicosilado})$. *Nota: Los resultados son difíciles de interpretar con precisión debido a la señal de fondo a la altura del trastuzumab con alto contenido en manosa digerido con Fabricator™ sin tratar. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

5 *Ejemplo 32: comparación de la eficacia de corte de proteínas de fusión EndoS-EndoH con y sin porción conectora en trastuzumab*

[0136] Se trató trastuzumab (obtenido en Epirus biopharma (Utrecht, Países Bajos); 287 µl, 6,0 mg, 21 mg/ml en Tris pH 8,0) con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuación, el trastuzumab digerido con Fabricator™ se dividió en tres porciones iguales y la solución amortiguadora se intercambió por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, respectivamente. El intercambio de solución amortiguadora se realizó utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentración final de 10 mg/ml. Para EndoSH (identificada por la SEQ ID N.º: 1) e His₆ se prepararon series de diluciones -EndoSH (EndoS-EndoH sin porción conectora GS; identificada por la SEQ ID N.º: 21) de 2, 10, 50 y 250 nM en cada una de las soluciones amortiguadoras mencionadas anteriormente. Las reacciones se iniciaron añadiendo 5 µl de trastuzumab digerido con Fabricator™ (10 mg/ml) en 5 µl de la proteína de fusión de endoglicosidasa diluida en la solución amortiguadora correspondiente, lo que da como resultado una concentración final de 5 mg/ml de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con endoglicosidasa 1, 5, 25 y 125 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacción, se tomó una muestra (4 µl) y se añadió a 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol y se incubó durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

Tabla 12. Porcentajes de corte (conversión) de trastuzumab digerido con Fabricator™ tras tratamiento con proteínas de fusión EndoS-EndoH con y sin porción conectora a pH 4,5, 6,0 y 7,5.

pH	Enzima (nM)	EndoSH	EndoSH, sin porción conectora
4,5	1	30	28*
	5	29*	29*
	25	29*	29*
	125	28*	29*
6,0	1	36	35
	5	45	46
	25	72	73
	125	85	83
7,5	1	34	40
	5	47	51
	25	81	80
	125	87	84

La conversión se calculó utilizando la siguiente fórmula: conversión (%) = $100 \times (Fc/2_{cortado}) / (Fc/2_{cortado} + Fc/2_{glicosilado})$. *Nota: Los resultados son difíciles de interpretar con precisión debido a la señal de fondo a la altura del trastuzumab digerido con Fabricator™ sin tratar. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

Ejemplo 33: comparación de la eficiencia de corte de EndoSH, EndoF3-EndoH y EndoS-EfEndo18A, EndoF2-EndoF1 en cAC10

[0137] Se trató cAC10 (300 µl, 6,0 mg, 20,1 mg/ml en Tris pH 8,0), obtenido como se describe anteriormente en el Ejemplo 7, con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuación, el cAC10 digerido con Fabricator™ se dividió en tres porciones iguales y la solución amortiguadora se intercambió por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM. El intercambio de solución amortiguadora se realizó utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentración final de 10 mg/ml. El Ejemplo 26 mostró un pH óptimo para el corte de glicanos de tipo complejo de pH 7,5 para EndoSH (identificada por la SEQ ID N.º: 1), pH 4,5 para EndoF3-EndoH (identificada por la SEQ ID N.º: 19), pH 7,5 para EndoS-EfEndo18A (identificada por la SEQ ID N.º: 15) y pH 4,5 para EndoF2-EndoF1 (identificada por la SEQ ID N.º: 17). Para cada una de las proteínas de fusión mencionadas anteriormente se preparó una serie de diluciones de 5, 50 y 500 nM en la solución amortiguadora de reacción con el pH óptimo como se ha mencionado anteriormente. Las reacciones se iniciaron añadiendo 5 µl de cAC10 digerido con Fabricator™ (10 mg/mL) a 5 µl de la proteína de fusión de endoglicosidasa diluida en la solución amortiguadora correspondiente, lo que da como resultado una concentración final de 5 mg/mL de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con proteína de fusión de endoglicosidasa 2,5, 25 y 250 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacción, se tomó una muestra (4 µl) y se añadió a 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol y se incubó durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

Tabla 13. Porcentajes de corte (conversión) de cAC10 digerido con Fabricator™ tras el tratamiento de varias proteínas de fusión de endoglicosidasa a un pH óptimo específico para cada proteína de fusión.

Concentración de enzima (nM)	EndoSH (a pH 7,5)	EndoF3-EndoH (a pH 4,5)	EndoS-EfEndo18A (a pH 7,5)	EndoF2-EndoF1 (a pH 4,5)
2,5	27*	70	26*	71
25	100	71	100	71
250	100	70	100	73

La conversión se calculó utilizando la siguiente fórmula: conversión (%) = 100 × (Fc/2_{cortado}) / (Fc/2_{cortado} + Fc/2_{glicosilado}). *Nota: Los resultados son difíciles de interpretar con precisión debido a la señal de fondo a la altura del cAC10 digerido con Fabricator sin tratar™. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

Ejemplo 34: comparación de la eficiencia de corte de EndoSH, EndoF3-EndoH y EndoS-EfEndo18A, EndoF2-EndoF1 en RNasaB

[0138] Se preparó una solución madre de RNasaB (2 mg/ml) en citrato de sodio 50 mM, pH 4,5, con NaCl 150 mM y en Tris.HCl 50 mM, pH 6,0, con NaCl 150 mM. El Ejemplo 26 mostró un pH óptimo para el corte de glicanos con alto contenido en manosa de pH 6,0 para EndoSH (identificada por la SEQ ID N.º: 1), pH 6,0 para EndoF3-EndoH (identificada por la SEQ ID N.º: 19), pH 6,0 para EndoS-EfEndo18A (identificada por la SEQ ID N.º: 15) y pH 4,5 para EndoF2-EndoF1 (identificada por la SEQ ID N.º: 17). Para cada una de las proteínas de fusión mencionadas anteriormente se preparó una serie de diluciones de 10, 50 y 250 nM en la solución amortiguadora de reacción con el pH óptimo como se ha mencionado anteriormente. Las reacciones se iniciaron añadiendo 5 µl de RNasa B (2 mg/mL en la solución amortiguadora correspondiente) a 5 µl de la proteína de fusión de endoglicosidasa diluida en la solución amortiguadora de reacción óptima como se ha mencionado anteriormente. Esto da como resultado una concentración final de 1 mg/ml de RNasa B con proteína de fusión de endoglicosidasa 5, 25 y 125 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacción, se tomó una muestra (4 µl) y se añadió a 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol y se incubó durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

Tabla 14. Porcentajes de corte (conversión) de RNasa B tras el tratamiento de diversas proteínas de fusión de endoglicosidasas al valor de pH óptimo específico para cada proteína de fusión.

Concentración de enzima (nM)	EndoSH (a pH 6,0)	EndoF3-EndoH (a pH 6,0)	EndoS-EfEndo18A (a pH 6,0)	EndoF2-EndoF1 (a pH 4,5)
5	34	37	52	43
25	100	100	100	100
125	100	100	100	100

La conversión se calculó utilizando la siguiente fórmula: conversión (%) = 100 × (RNasaB_{cortada}) /

(RNasaB_{cortada} + RNasaB_{glicosilada}).

Ejemplo 35: comparación de la eficiencia de corte de EndoSH, EndoF3-EndoH y EndoS-EfEndo18A, EndoF2-EndoF1 en fibrinógeno

5 [0139] Se disolvió fibrinógeno de plasma humano (disponible comercialmente en Sigma), que contiene un sitio de glicosilación en las cadenas alfa, beta y gamma, hasta una concentración final de 10 mg/ml en Tris.HCl 50 mM, pH 6,0 con 150 NaCl mM y en Tris.HCl 50 mM pH 7,5 con NaCl 150 mM girando a 300 r.p.m. a 37 °C durante 15 minutos. El fibrinógeno no se pudo disolver en citrato de sodio 50 mM, pH 4,5, con NaCl 150 mM usando el procedimiento mencionado anteriormente. El Ejemplo 26 mostró un pH óptimo para el corte de glicanos de tipo complejo de pH 7,5 para EndoSH (identificada por la SEQ ID N.º: 1), pH 4,5 para EndoF3-EndoH (identificada por la SEQ ID N.º: 19), pH 7,5 para EndoS- EfEndo18A (identificada por la SEQ ID N.º: 15) y pH 4,5 para EndoF2-EndoF1 (identificada por la SEQ ID N.º: 17). Para EndoSH y EndoS-EfEndo18A se preparó una serie de diluciones de 5, 50 y 500 nM en Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, que es la solución amortiguadora de reacción óptima para estas enzimas. Para EndoF3-EndoH y EndoF2-EndoF1 se preparó una serie de diluciones de 5, 50 y 500 nM en Tris.HCl 50 mM pH 6,0 con NaCl 150 mM, que es el pH más óptimo en el que se puede solubilizar el fibrinógeno. Las reacciones se iniciaron añadiendo 5 µl de fibrinógeno (10 mg/mL) a 5 µl de la proteína de fusión de endoglicosidasa diluida en la solución amortiguadora correspondiente, dando como resultado una concentración final de 5 mg/mL de fibrinógeno con proteína de fusión endoglicosidasa de 2,5, 25 y 250 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacción, se tomó una muestra (4 µl) y se añadió a 16 µl de solución amortiguadora de muestra 1x Laemmli con 2-mercaptoetanol y se incubó durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante aproximadamente 120 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon para las cadenas beta y gamma basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

Tabla 15. Porcentajes de corte (conversión) de fibrinógeno tras el tratamiento de diversas proteínas de fusión de endoglicosidasas al valor de pH óptimo específico para cada proteína de fusión.

Cadena de fibrinógeno	Concentración de enzima (nM)	EndoSH (a pH 7,5)	EndoF3-EndoH (a pH 6,0)	EndoS-EfEndo18A (a pH 7,5)	EndoF2-EndoF1 (a pH 6,0)
cadena beta	2,5	9*	10*	1*	23
	25	9*	10*	10*	34
	250	8*	8*	7*	54
cadena gamma	2,5	5*	7*	6*	13
	25	5*	4*	5*	25
	250	4*	6*	4*	50

La conversión se calculó por separado para las cadenas beta y gamma utilizando la siguiente fórmula: conversión (%) = 100 × (fibrinógeno_{cortado}) / (fibrinógeno_{cortado} + fibrinógeno_{glicosilado}). *Nota: Los resultados son difíciles de interpretar con precisión debido a las señales de fondo en el nivel máximo de fibrinógeno no tratado. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

30 *Ejemplo 36: comparación de la eficacia de corte de las proteínas de fusión EndoSH y EndoF3-EndoH con las proteínas individuales EndoS, EndoF3 y EndoH en trastuzumab*

35 [0140] Se trató trastuzumab (obtenido en Epirus biopharma (Utrecht, Países Bajos); 287 µl, 6,0 mg, 21 mg/ml en Tris pH 8,0) con Fabricator™ (9 µl, 50 U/µL) durante 1 h a 37 °C. A continuación, el trastuzumab digerido con Fabricator™ se dividió en tres porciones iguales y la solución amortiguadora se intercambió por citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM, respectivamente. El intercambio de solución amortiguadora se realizó utilizando una membrana Amicon Ultra-0.5, Ultracel-10 (Merck Millipore) y las muestras se concentraron hasta una concentración final de 10 mg/ml. EndoSH (identificada por la SEQ ID N.º: 1), EndoF3-EndoH (identificada por la SEQ ID N.º: 19), EndoS (disponible comercialmente en Genovis, Lund, Suecia), EndoF3 (disponible comercialmente en Sigma-Aldrich, UE) y EndoH (disponible comercialmente en New England Biolabs, Ipswich, EE. UU.) se diluyeron a 50 y 500 nM en cada una de las soluciones amortiguadoras mencionadas anteriormente. Las reacciones se iniciaron añadiendo 5 µl de trastuzumab digerido con Fabricator™ (10 mg/mL) a 5 µl de endoglicosidasas diluidas en la solución amortiguadora correspondiente, lo que da como resultado una concentración final de 5 mg/mL de IgG digerida con Fabricator™ (fragmento Fc/2 67 µM) con endoglicosidasa 25 y 250 nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacción se tomó una muestra (4 µl) y se añadió a 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol y se incubó durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión

se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

5 Tabla 16. Porcentajes de corte (conversión) de trastuzumab digerido con Fabricator™ tras el tratamiento de diversas endoglicosidasas y proteínas de fusión de endoglicosidasas a pH 4,5, 6,0 y 7,5.

pH	Enzima (nM)	EndoSH	EndoF3-EndoH	EndoS	EndoF3	EndoH
4,5	25	21*	21*	22*	21*	21*
	250	27*	50	29*	21*	19*
6,0	25	68	23*	76	24*	21*
	250	83	39	82	22*	19*
7,5	25	77	26*	82	15*	21*
	250	100	36	100	12*	20*

La conversión se calculó utilizando la siguiente fórmula: $\text{conversión (\%)} = 100 \times (\text{Fc}/2_{\text{cortado}}) / (\text{Fc}/2_{\text{cortado}} + \text{Fc}/2_{\text{glicosilado}})$. *Nota: Los resultados son difíciles de interpretar con precisión debido a la señal de fondo a la altura del trastuzumab digerido con Fabricator™ sin tratar. No se aplicó ninguna corrección de fondo para la cuantificación de la conversión.

Ejemplo 37: comparación de la eficacia de corte de las proteínas de fusión EndoSH y EndoF3-EndoH con las proteínas individuales EndoS, EndoF3 y EndoH en RNasaB

10 [0141] Se preparó una solución madre de RNasaB (2 mg/ml) en citrato de sodio 50 mM, pH 4,5 con NaCl 150 mM, en Tris.HCl 50 mM, pH 6,0 con NaCl 150 mM y en Tris.HCl 50 mM, pH 7,5 con NaCl 150 mM. EndoSH (identificada por la SEQ ID N.º: 1), EndoF3-EndoH (identificada por la SEQ ID N.º: 19), EndoS (disponible comercialmente en Genovis, Lund, Suecia), EndoF3 (disponible comercialmente en Sigma-Aldrich, UE) y EndoH (disponibles comercialmente en New England Biolabs, Ipswich, EE. UU.) se diluyeron hasta una concentración de 50 nM en cada una de las soluciones amortiguadoras mencionadas anteriormente. Las reacciones se iniciaron añadiendo 5 µl de RNasa B (2 mg/mL) a 5 µl de la proteína de fusión de endoglicosidasa diluida (50 nM) en la solución amortiguadora de reacción correspondiente, dando como resultado una concentración final de 1 mg/mL de RNasa B y 25 Endoglicosidasa nM. Las reacciones se incubaron durante 60 minutos a 37 °C. Para cada reacción, se tomó una muestra (4 µl) y se añadió a 16 µl de solución amortiguadora de muestra 1x Laemmli sin 2-mercaptoetanol y se incubó durante 5 minutos a 95 °C. Las muestras (5 µl/muestra) se cargaron en gel SDS-PAGE y se procesaron durante 70 minutos (20 mA), se tiñeron en coomassie coloidal durante la noche y finalmente se quitó la tinción en agua. Los porcentajes de conversión se calcularon basándose en el escaneo del gel SDS-PAGE con un escáner plano normal y la cuantificación con una herramienta de software (CLIQS v1.1).

25 Tabla 17. Porcentajes de corte (conversión) de RNasa B tras el tratamiento de diversas endoglicosidasas y proteínas de fusión de endoglicosidasas (25 nM) a pH 4,5, 6,0 y 7,5.

pH	EndoSH	EndoF3-EndoH	EndoS	EndoF3	EndoH
4,5	74	75	0	0	30
6,0	70	100	0	0	39
7,5	42	48	0	0	33

La conversión se calculó utilizando la siguiente fórmula: $\text{conversión (\%)} = 100 \times (\text{RNasaB}_{\text{cortada}}) / (\text{RNasaB}_{\text{cortada}} + \text{RNasaB}_{\text{glicosilada}})$.

Secuencias

30 [0142]

Identificación de secuencia de proteína de fusión EndoS-EndoH (o EndoSH) tal como se expresa en E coli (SEQ. ID N.º: 1):

ES 2 972 511 T3

1 MPSIDSLHYL SENSKKEFKE ELSKAGQESQ KVKEILAKAQ QADKQAQELA
51 KMKIPEKIPM KPLHGPLYGG YFRTWHDKTS DPTEKDKVNS MGELPKEVDL
101 AFIFHDWTKD YSLFWKELAT KHVPKLNKQG TRVIRTIPWR FLAGGDNSGI
151 AEDTSKYPT PEGNKALAKA IVDEYVYKYN LDGLDVDVEH DSIPKVDKKE
201 DTAGVERSIQ VFEEIGKLIG PKGVDKSRLF IMDSTYMADK NPLIERGAPY
251 INLLLVQVYG SQGEKGGWEP VSNRPEKTME ERWQGYISKYI RPEQYMI GFS
301 FYEENAQEGN LWYDINSRKD EDKANGINTD ITGTRAERYA RWQPKTGGVK
351 GGIFSYAIDR DGVAHQPKKY AKQKEFKDAT DNIFHSDYSV SKALKTVMLK
401 DKSVDLIDEK DFPDKALREA VMAQVGTRKG DLERFNGTLR LDNPAIQSLE
451 GLNKFKKLAQ LDLIGLSRIT KLDRSVLPAN MKPGKDTLET VLETYKKDNK
501 EEPATIPVVS LKVSGLTGLK ELDLSGFDRE TLAGLDAATL TSLEKVDISG
551 NKLDLAPGTE NRQIFDTMLS TISNHVGSNE QTVKFDKQKP TGHYPDTY GK
601 TSLRLPVANE KVDLQSQLLF GTVTNQGTLI NSEADYKAYQ NHKIAGRSFV
651 DSNYHYNNFK VSYENYTVKV TDSTLGT TTD KTLATDKEET YKVDFFS PAD
701 KTKAVHTAKV IVGDEKTMV NLAEGATVIG GSADPVNARK VFDGQLGSET
751 DNISLGWDSK QSIIFKLKED GLIKHWRFFN DSARNPETTN KPIQEASLQI
801 FNIKDYNDN LLENPNKFDD EKYWITVDY SAQGERATAF SNTLNNITSK
851 YWRVVFDTKG DRYSSPVVPE LQILGYPLPN ADTIMKTVTT AKELSQQKDK
901 FSQKMLDELK IKEMALETSL NSKIFDVTAI NANAGVLKDC IEKRQLLKKG
951 GGGSGGGSGG GGGSHHHHHH EFGGGGSGGG GSGGGGSAPA PVKQGPTSVA
1001 *YVEVNNNSML NVGKYTLADG GGNAFDVAVI FAANINYDTG TKTAYLHFNE*
1051 *NVQRVLDNAV TQIRPLQQQG IKVLLSVLGN HQGAGFANFP SQQAASAFK*
1101 *QLSDAVAKYG LDGVDFDDEY AEYGNNGTAQ PNDSSFVHLV TALRANMPDK*
1151 *IISLYNIGPA ASRLSYGGVD VSDKFDYAWN PYYGTWQVPG IALPKAQLSP*
1201 *AAVEIGRTSR STVADLARRT VDEGYGVYLT YNLDGGDRTA DVSAFTRELY*
1251 *GSEAVRTP*

(la porción conectora está subrayada, la secuencia de EndoH se indica en cursiva)

Identificación de secuencia de la proteína de fusión EndoS-EndoH (o EndoSH) tal como se expresa en CHO (SEQ. ID N.º: 2):

1 MPSIDSLHYL SENSKKEFKE ELSKAGQESQ KVKEILAKAQ QADKQAQELA
51 KMKIPEKIPM KPLHGPLYGG YFRTWHDKTS DPTEKDKVNS MGELPKEVDL
101 AFIFHDWTKD YSLFWKELAT KHVPKLNKQG TRVIRTIPWR FLAGGDNSGI

ES 2 972 511 T3

151 AEDTSKYPNT PEGNKALAKA IVDEYVYKYN LDGLDVDVEH DSIPKVDKKE
 201 DTAGVERSIQ VFEEIGKLIQ PKGVDKSRLF IMDSTYMADK NPLIERGAPY
 251 INLLLVQVYG SQGEKGGWEP VSNRPEKTME ERWQGYSKYI RPEQYMI GFS
 301 FYEENAQEGN LWYDINSRKD EDKANGINTD ITGTRAERYA RWQPKTGGVK
 351 GGIFSYAIDR DGVAHQPKKY AKQKEFKDAT DNI FHSYDYSV SKALKTVMLK
 401 DKSVDLIDEK DFPDKALREA VMAQVGTRKG DLERFNGTLR LDNPAIQSLE
 451 GLNKFKKLAQ LDLIGLSRIT KLDRSVLPAN MKPGKDTLET VLETYKKNK
 501 EEPATIPPVS LKVSGLTGLK ELDLSGFDRE TLAGLDAATL TSLEKVDISG
 551 NKLDLAPGTE NRQIFDTMLS TISNHVGSNE QTVKFDKQKP TGHYPDTYGK
 601 TSLRLPVANE KVDLQSQLLF GTVTNQGTLI NSEADYKAYQ NHKIAGRSFV
 651 DSNYHYNNFK VSYENYTVKV TDSTLGTTTD KTLATDKEET YKVDFFS PAD
 701 KTKAVHTAKV IVGDEKTMV NLAEGATVIG GSADPVNARK VFDGQLGSET
 751 DNISLGWDSK QSIIFKLKED GLIKHWRFFN DSARNPETTN KPIQEASLQI
 801 FNIKDYNLN LLENPNKFDD EKYWITVDTY SAQGERATAF SNTLNNITSK
 851 YWRVVFDTKG DRYSSPVVPE LQILGYPLPN ADTIMKTVTT AKELSQQKDK
 901 FSQKMLDELK IKEMALETSL NSKIFDVTAI NANAGVLKDC IEKRQLLKKG
 951 GGGSGGGGSG GGGSHHHHHH GGGGSGGGGS GGGGSAPAPV *KQGPTSVAYV*
 1001 *EVNNSMLNV GKYTLADGGG NAFDVAVIFA ANINYDTGTK TAYLHFNENV*
 1051 *QRVLDNAV TQ IRPLQQQGIK VLLSVLGNHQ GAGFANFPSQ QAASAFKQL*
 1101 *SDAVAKYGLD GVDFDDEYAE YGNNGTAQPN DSSFVHLVTA LRAMPKKII*
 1151 *SLYNIGPAAS RLSYGGVDVS DKFDYAWN PY YGTWQVPGIA LPKAQLSPAA*
 1201 *VEIGRTSRST VADLARRTV D EGYGVYLTYN LDGGDRTADV SAFTRELYGS*
 1251 *EAVRTP*

(la porción conectora está subrayada, la secuencia EndoH se indica en cursiva)

Secuencia de His₆-TnGalNAcT(33-421) tal como se expresa en *E coli* (SEQ. ID N.º: 3):

1 MGSSHHHHHH SSGLVPRGSH MSPLRITYLT PLYNATQPTL RNVERLAANW PPKIPSNIYE
 61 DSEEYSIKNI SLSNHTTRAS VVHPPSSITE TASKLDKNMT IQDGAFAMIS PTPLLITKLM
 121 DSIKSYVTTE DGVKKAEAVV TLPLCDMPP DLGPITLNKT ELELEWVEKK FPEVEWGGRY
 181 SPPNCTARHR VAIIVPYRDR QQHLAIFLNH MHPFLMKQQI EYGI FIVEQE GNKDFNRAKL
 241 MNVGFVESQK LVAEGWQCFV FHDIDLPLD TRNLYSCPRQ PRHMSASIDK LHFKLPYEDI
 301 FGGVSAMTLE QFTRVNGFSN KYWG WGGEDD DMSYRLKIN YHIARYKMSI ARYAML DHHK
 361 STPNPKRYQL LSQTSKTFQK DGLSTLEYEL VQVVQYHLYT HILVNIDERS

Identificación de secuencia de la proteína de fusión EndoF3-EfEndo18A tal como se expresa en *E coli* (SEQ. ID N.º: 13):

ES 2 972 511 T3

1 MATALAGSNG VCIAYYITDG RNPTFKLKDI PDKVDMVILF GLKYWSLQDT
51 TKLPGGTGMM GSFKSYKDLT TQIRSLQSRG IKVLQNIIDD VSWQSSKPGG
101 FASAAAYGDA IKSIVIDKWK LDGSLDIEH SGAKPNPIPT FPGYAATGYN
151 GWYSGSMAAT PAFLNVISEL TKYFGTTAPN NKQLQIASGI DVYAWNKIME
201 NFRNNFNFIQ LQSYGANVSR TQLMMNYATG TNKIPASKMV FGAYAEGGTN
251 QANDVEVAKW TPTQGAKEGGM MIYTYNSNVS YANAVRDAVK NGGGGSGGGG
301 SGGGGSHHHH HHEFGGGGSG GGGSGGGGSA *STVTPKTVMY* *VEVNNHDFNN*
351 *VGKYTLAGTN QPAFDMGIIF AANINYDTVN KKPYYLYLNER VQOTLNEAET*
401 *QIRPVQARGT KVLISILGNH EGAGFANFPT YESADAFAAQ LEQVVNTYHL*
451 *DGIDFDDEYA EYGKNGTPQP NNSSEIWLLO ALRNRLGNDK LITFYNIGPA*
501 *AANSSANPQM SSLIDYAWNP YYSTWNPPQI AGMPASRLGA SAVEVGVNQN*
551 *LAAQYAKRTK AEQYGIYLMY NLPKGDSSAY ISAATQELYG RKTNYSPVTP*
601 TP

- 5 (la porción conectora está subrayada, la secuencia EfEndo18A está indicada en cursiva)

Identificación de secuencia de la proteína de fusión EndoF2-EfEndo18A tal como se expresa en E coli (SEQ. ID N.º: 14):

1 MAVNLSNLIA YKNSDHQISA GYYRTWRDSA TASGNLPSMR WLPDSLDMVM
51 VFPDYTPPEN AYWNTLKTNY VPYLHHRGK VIITLGLDLS ATTTGGQDSI
101 GYSSWAKGIY DKWVGEYNLD GIDIDIESP SGATLTKFVA ATKALSKYFG
151 PKSGTGKTFV YDTNQNPTNF FIQTAPRYNY VFLQAYGRST TNLTTVSGLY
201 APYISMKQFL PGFSFYEENG YPGNYWNVDR YPQNGTGRAY DYARWQPATG
251 KKGGVFSYAI ERDAPLTSSN DNTLRAPNFR VTKDLIKIMN PGGGGSGGGG
301 SGGGGSHHHH HHEFGGGGSG GGGSGGGGSA *STVTPKTVMY* *VEVNNHDFNN*
351 *VGKYTLAGTN QPAFDMGIIF AANINYDTVN KKPYYLYLNER VQOTLNEAET*
401 *QIRPVQARGT KVLISILGNH EGAGFANFPT YESADAFAAQ LEQVVNTYHL*
451 *DGIDFDDEYA EYGKNGTPQP NNSSEIWLLO ALRNRLGNDK LITFYNIGPA*
501 *AANSSANPQM SSLIDYAWNP YYSTWNPPQI AGMPASRLGA SAVEVGVNQN*
551 *LAAQYAKRTK AEQYGIYLMY NLPKGDSSAY ISAATQELYG RKTNYSPVTP*
601 TP

- 10 (la porción conectora está subrayada, la secuencia EfEndo18A está indicada en cursiva)

Identificación de secuencia de la proteína de fusión EndoS-EfEndo18A tal como se expresa en E coli (SEQ. ID N.º: 15):

ES 2 972 511 T3

1 MPSIDSLHYL SENSKKEFKE ELSKAGQESQ KVKEILAKAQ QADKQAQELA
51 KMKIPEKIPM KPLHGPLYGG YFRTWHDKTS DPTEKDKVNS MGELPKEVDL
101 AFIFHDWTKD YSLFWKELAT KHVPKLNKQG TRVIRTIPWR FLAGGDNSGI
151 AEDTSKYPNT PEGNKALAKA IVDEYVYKYN LDGLDVDVEH DSIPKVDKKE
201 DTAGVERSIQ VFEEIGKLIG PKGVDKSRLF IMDSTYMADK NPLIERGAPY
251 INLLLVQVYG SQGEKGGWEP VSNRPEKTME ERWQGYSKYI RPEQYMIGFS
301 FYEENAQEGN LWYDINSRKD EDKANGINTD ITGTRAERYA RWQPKTGGVK
351 GGIFSYAIDR DGVAHQPKKY AKQKEFKDAT DNIFHSDYSV SKALKTVMLK
401 DKSVDLIDEK DFPDKALREA VMAQVGTRKG DLERFNGTLR LDNPAIQSLE
451 GLNKFKKLAQ LDLIGLSRIT KLDRSVLPAN MKPGKDTLET VLETYKKNK
501 EEPATIPVVS LKVSGLTGLK ELDLSGFDRE TLAGLDAATL TSLEKVDISG
551 NKLDLAPGTE NRQIFDTMLS TISNHVGSNE QTVKFDKQKP TGHYPDTY GK

601 TSLRRLPVANE KVDLQSQLLF GTVTNQGTLI NSEADYKAYQ NHKIAGRSFV
651 DSNYHYNNFK VSYENYTVKV TDSTLGTTTD KTLATDKEET YKVDFESPAD
701 KTKAVHTAKV IVGDEKTMV NLAEGATVIG GSADPVNARK VFDGQLGSET
751 DNISLGWDSK QSIIFKLKED GLIKHWRFN DSARNPETTN KPIQEASLQI
801 FNIKDYNDLN LLENPNKFDD EKYWITVDY SAQGERATAF SNTLNNITSK
851 YWRVVFDTKG DRYSSPVVPE LQILGYPLPN ADTIMKTVTT AKELSQQKDK
901 FSQKMLDELK IKEMALETSL NSKIFDVTAI NANAGVLKDC IEKRQLLKKG
951 GGGSGGGGSG GGGSHHHHHH EFGGGGSGGG GSGGGGSAST *VTPKTVMYVE*
1001 *VNNHDFNNVG KYTLAGTNQP AFDMGIIFAA NINYDTVNKK PYLYLNERVQ*
1051 *QTLNEAETQI RPVQARGTKV LLSILGNHEG AGFANFPTYE SADAFAAQLE*
1101 *QVVNTYHLDG IDFDDEYAEY GKNGTPQPNN SSFIWLLQAL RNRLGNDKLI*
1151 *TFYNIGPAAA NSSANPQMSS LIDYAWNPYY STWNPPQIAG MPASRLGASA*
1201 *VEVGVNQNLA AQYAKRTKAE QYGIYLMYNL PGKDSSAYIS AATQELYGRK*
1251 *TNYSPTVPTP*

(la porción conectora está subrayada, la secuencia EfEndo18A está indicada en cursiva)

Identificación de secuencia de la proteína de fusión EndoF3-EndoF1 tal como se expresa en *E coli* (SEQ. ID N.º: 16):

ES 2 972 511 T3

1 MATALAGSNG VCIAYITDG RNPTFKLKI PDKVDMVILF GLKYWSLQDT
51 TKLPGGTGMM GSFKSYKDLL TQIRSLQSRG IKVLQNIIDD VSWQSSKPGG
101 FASAAAYGDA IKSIVIDKWK LDGISLDIEH SGAKPNPIPT FPGYAATGYN
151 GWYSGSMAAT PAFLNVISEL TKYFGTTAPN NKQLQIASGI DVYAWNKIME
201 NFRNNFNFIQ LQSYGANVSR TQLMMNYATG TNKIPASKMV FGAYAEGGTN
251 QANDVEVAKW TPTQGAKGGM MIYTYNSNVS YANAVRDAVK NGGGGSGGGG
301 SGGGGSHHHH HHEFGGGGSG GGGSGGGGSA VTGTTKANIK LFSFTEVNDT
351 NPLNNLNFTL KNSGKPLVDM VVLFANINNY DAANDKVFVS NNPVQHLLT
401 NRAKYLKPLQ DKGIVILSI LGNHDRSGIA NLSTARAKAF AQELKNTCDL
451 YNLDGVFFDD EYSAYQTPPP SGFVTPSNNA AARLAYETKQ AMPNKLVTVY
501 VYSRTSSFPT AVDGVNAGSY VDYAIHDYGG SYDLATNYPG LAKSGMVMSS
551 QEFNQGRYAT AQALRNIVTK GYGGHMIFAM DPNRSNFTSG QLPALKLIAK
601 ELYGDELVYS NTPYSKDW

(la porción conectora está subrayada, la secuencia EndoF1 se indica en cursiva)

5 Identificación de secuencia de la proteína de fusión EndoF2-EndoF1 tal como se expresa en *E coli* (SEQ. ID N.º: 17):

1 MAVNLSNLIA YKNSDHQISA GYYRTWRDSA TASGNLPSMR WLPDSLDMVM
51 VFPDYTPPEN AYWNTLKTNY VPYLHKRGTK VIITLGLDLS ATTTGGQDSI
101 GYSSWAKGIY DKWVGEYNLD GIDIDIESSP SGATLTKFVA ATKALSKYFG
151 PKSGTGKTFV YDTNQNPNTNF FIQTAPRYNY VFLQAYGRST TNLTTVSGLY
201 APYISMKQFL PGFSFYEENG YPGNYWNDVR YPQNGTGRAY DYARWQPATG
251 KKGGVFSYAI ERDAPLTSSN DNTLRAPNFR VTKDLIKIMN PGGGGSGGGG
301 SGGGGSHHHH HHEFGGGGSG GGGSGGGGSA VTGTTKANIK LFSFTEVNDT
351 NPLNNLNFTL KNSGKPLVDM VVLFANINNY DAANDKVFVS NNPVQHLLT
401 NRAKYLKPLQ DKGIVILSI LGNHDRSGIA NLSTARAKAF AQELKNTCDL
451 YNLDGVFFDD EYSAYQTPPP SGFVTPSNNA AARLAYETKQ AMPNKLVTVY
501 VYSRTSSFPT AVDGVNAGSY VDYAIHDYGG SYDLATNYPG LAKSGMVMSS
551 QEFNQGRYAT AQALRNIVTK GYGGHMIFAM DPNRSNFTSG QLPALKLIAK
601 ELYGDELVYS NTPYSKDW

10 (la porción conectora está subrayada, la secuencia EndoF1 se indica en cursiva)

Identificación de secuencia de proteína de fusión EndoS-EndoF1 tal como se expresa en *E coli* (SEQ. ID N.º: 18):

ES 2 972 511 T3

1 MPSIDSLHYL SENSKKEFKE ELSKAGQESQ KVKEILAKAQ QADKQAQELA
 51 KMKIPEKIPM KPLHGPLYGG YFRTWHDKTS DPTEKDKVNS MGELPKVEVDL
 101 AFIFHDWTKD YSLEFWKELAT KHVPKLNKQG TRVIRTIPWR FLAGGDNSGI
 151 AEDTSKYPNT PEGNKALAKA IVDEYVYKYN LDGLDVDVEH DSI PKVDKKE
 201 DTAGVERSIQ VFEEIGKLIG PKGVDKSRLF IMDSTYMADK NPLIERGAPY
 251 INLLLVQVYG SQGEKGGWEP VSNRPEKTME ERWQGYSKYI RPEQYMI GFS
 301 FYEENAQEGN LWYDINSRKD EDKANGINTD ITGTRAERYA RWQPKTGGVK
 351 GGIFSYAIDR DGVAHQPKKY AKQKEFKDAT DNI FHS DYSV SKALKTVMLK
 401 DKS YDLIDEK DFPDKALREA VMAQVGTTRKG DLERFNGTLR LDNPAIQSLE
 451 GLNKFKKLAQ LDLIGLSRIT KLDRSVLPAN MKPGKDTLET VLETYKKDNK
 501 EEPATIPVVS LKVSGLTGLK ELDLSGFDRE TLAGLDAATL TSLEKVDISG
 551 NKLDLAPGTE NRQIFDTMLS TISNHVGSNE QTVKFDKQKP TGHYPDTY GK
 601 TSLRLPVANE KVDLQSQLLF GTVTNQGTLI NSEADYKAYQ NHKIAGRSFV
 651 DSNYHYNNFK VSYENYTVKV TDSTLGT TTD KTLATDKEET YKVDF FSPAD
 701 KTKAVHTAKV IVGDEKTMV NLAEGATVIG GSADPVNARK VFDGQLGSET
 751 DNISLGWDSK QSIIIFKLKED GLIKHWRFN DSARNPETTN KPIQEASLQI
 801 FNIKDYNDN LLENPNKFDD EKYWITVDY SAQGERATAF SNTLNNITSK
 851 YWRVVFDTKG DRYSSPVVPE LQILGYPLPN ADTIMKT VTT AKELSQQKDK
 901 FSQKMLDELK IKEMALETSL NSKIFDVTAI NANAGVLKDC IEKRQLLKKG
 951 GGSGGGGGSG GGGSHHHHHH EFGGGGSGGG GSGGGGSAVT *GTTKANIKLF*
 1001 *SFTEVNDTNP LNNLNFTLKN SGKPLVDMVV LFSANINYDA ANDKVFVSNN*
 1051 *PNVQHLLTNR AKYLKPLQDK GIKVILSILG NHDRSGIANL STARAKAFAQ*
 1101 *ELKNTCDLYN LDGVFFDDEY SAYQTPPPSG FVTPSNAAA RLAYETKQAM*
 1151 *PNKLVTVVYVY SRTSSFPTAV DGVNAGSYVD YAIHDYGGSY DLATNYPGLA*
 1201 *KSGMVMSSQE FNQGRYATAQ ALRNIVTKGY GGHMIFAMDP NRSNFTSGQL*
 1251 *PALKLIAKEL YGDELVYSNT PYSKDW*

(la porción conectora está subrayada, la secuencia EndoF1 se indica en cursiva)

Identificación de secuencia de la proteína de fusión EndoF3-EndoH tal como se expresa en *E coli* (SEQ. ID N.º: 19):

1 MATALAGSNG VCIAYYITDG RNPTFKLKDI PDKVDMVILF GLKYWSLQDT
 51 TKLPGGTGMM GSFKSYKDLT TQIRSLQSRG IKVLQNI DDD VSWQSSKPGG
 101 FASAAAYGDA IKSIVIDKWK LDGISLDIEH SGAKPNPIPT FPGYAATGYN
 151 GWYSGSMAAT PAFLNVISEL TKYFGTTAPN NKQLQIASGI DVYAWNKIME

ES 2 972 511 T3

201 NFRNNFNFIQ LQSYGANVSR TQLMMNYATG TNKIPASKMV FGAYAEGGTN
251 QANDVEVAKW TPTQGAKGGM MIYTYNSNVS YANAVRDAVK NGGGGSGGGG
301 SGGGGSHHHH HHEFGGGGSG GGGSGGGGSA PAVKQGPTS VAYVEVNNNS
351 *MLNVGKYTLA* *DGGGNAFDVA* *VIFAANINYD* *TGKTAYLHF* *NENVQRVLDN*
401 *AVTQIRPLQQ* *QGIKVLSSVL* *GNHQGAGFAN* *FPSQQAASAF* *AKQLSDAVAK*
451 *YGLDGVDVFD* *EYAEYGNGT* *AQPNDSSFVH* *LVTALRANMP* *DKIISLYNIG*
501 *PAASRLSYGG* *VDVSDKFDYA* *WNPYYGTWQV* *PGIALPKAQL* *SPAAVEIGRT*
551 *SRSTVADLAR* *RTVDEGYGVY* *LTYNLDGGDR* *TADVSAFTRE* *LYGSEAVRTP*

(la porción conectora está subrayada, la secuencia EndoH se indica en cursiva)

Identificación de secuencia de la proteína de fusión EndoF2-EndoH tal como se expresa en E coli (SEQ. ID N.º: 20):

1 MAVNLSNLIA YKNSDHQISA GYYRTWRDSA TASGNLPSMR WLPDSLDMVM
51 VFPDYTPPEN AYWNTLKTNY VPYLHKRGTK VIITLGLDLS ATTTGGQDSI
101 GYSSWAKGIY DKWVGEYNLD GIDIDIESP SGATLTKFVA ATKALSKYFG
151 PKSGTGKTFV YDTNQNPNTNF FIQTAPRYNY VFLQAYGRST TNLTTVSGLY
201 APYISMKQFL PGFSFYEENG YPGNYWINDVR YPQNGTGRAY DYARWQPATG
251 KKGGVFSYAI ERDAPLTSSN DNTLRAPNFR VTKDLIKIMN PGGGGSGGGG
301 SGGGGSHHHH HHEFGGGGSG GGGSGGGGSA PAVKQGPTS VAYVEVNNNS
351 *MLNVGKYTLA* *DGGGNAFDVA* *VIFAANINYD* *TGKTAYLHF* *NENVQRVLDN*
401 *AVTQIRPLQQ* *QGIKVLSSVL* *GNHQGAGFAN* *FPSQQAASAF* *AKQLSDAVAK*
451 *YGLDGVDVFD* *EYAEYGNGT* *AQPNDSSFVH* *LVTALRANMP* *DKIISLYNIG*
501 *PAASRLSYGG* *VDVSDKFDYA* *WNPYYGTWQV* *PGIALPKAQL* *SPAAVEIGRT*
551 *SRSTVADLAR* *RTVDEGYGVY* *LTYNLDGGDR* *TADVSAFTRE* *LYGSEAVRTP*

(la porción conectora está subrayada, la secuencia EndoH se indica en cursiva)

Identificación de secuencia de la proteína de fusión His₆-EndoS-EndoH (EndoS-EndoH sin porción conectora GS) tal como se expresa en E coli (SEQ. ID N.º: 21):

ES 2 972 511 T3

1 MGSSHHHHHH SSGLVPRGSH MPSIDSLHYL SENSKKEFKE ELSKAGQESQ
 51 KVKEILAKAQ QADKQAQELA KMKIPEKIPM KPLHGPLYGG YFRTWHDKTS
 101 DPTEKDKVNS MGELPKEVDL AFIFHDWTKD YSLFWKELAT KHVPKLNKQG
 151 TRVIRTIPWR FLAGGDNSGI AEDTSKYPNT PEGNKALAKA IVDEYVYKYN
 201 LDGLDVEDVEH DSIPKVDKKE DTAGVERSIQ VFEEIGKLIG PKGVDKSRLF
 251 IMDSTYMADK NPLIERGAPY INLLLVQVYG SQGEKGGWEP VSNRPEKTME
 301 ERWQGYSKYI RPEQYMIGFS FYEENAQEGN LWYDINSRKD EDKANGINTD
 351 ITGTRAERYA RWQPKTGGVK GGIFSYAIDR DGVAHQPKKY AKQKEFKDAT
 401 DNI FHS DYSV SKALKTVMLK DKS YDLIDEK DFPDKALREA VMAQVGTRKG
 451 DLERFNGTLR LDNPAIQSLE GLNKFKKLAQ LDLIGLSRIT KLD RSVLPAN
 501 MKPGKDTLET VLETYKDKNK EEPATIPVVS LKVSGLTGLK ELDLSGFDRE
 551 TLAGLDAATL TSLEKVDISG NKLDLAPGTE NRQIFDTMLS TISNHVGSNE
 601 QTVKFDKQKP TGHYPDTY GK TSLR L PVANE KVDLQSQLLF GTVTNQGTLI
 651 NSEADYKAYQ NHKIAGRSFV DSNYHYNNFK VSYENYTVKV TDSTLGTTTTD
 701 KTLATDKEET YKVDF F SPAD KTKAVHTAKV IVGDEKTMMV NLAEGATVIG
 751 GSADPVNARK VFDGQLGSET DNISLGWDSK QSII FKLKED GLIKHWRFFN
 801 DSARNPETTN KPIQEASLQI FNIKDY NLDN LLENPNKFDD EKYWITVDTY
 851 SAQGERATAF SNTLNNITSK YWRVVFDTKG DRYSSPVVPE LQILGYPLPN
 901 ADTIMKTVTT AKELSQQKDK FSQKMLDELK IKEMALETSL NSKIFDVTAI
 951 NANAGVLKDC IEKRQLLKKA *PAPVKQGPTS* VAYVEVNNNS MLNVGKYTLA
 1001 DGGGNAFDVA *VIFAANINYD* *TGKTAYLHF* *NENVQRVLDN* *AVTQIRPLQQ*
 1051 *QGIKVL LSVL* *GNHQGAGFAN* *FPSQQAASAF* *AKQLSDAVAK* *YGLDGVDFDD*
 1101 *EYAEYGNGT* *AQPNDSSFVH* *LVTALRANMP* *DKIISLYNIG* *PAASRLSYGG*
 1151 *VDVSDKFDYA* *WNPYYGTWQV* *PGIALPKAQL* *SPA AVEIGRT* *SRSTVADLAR*
 1201 *RTVDEGYGVY* *LTYNLDGGDR* *TADVSAFTRE* *LYGSEAVRTP*

(La secuencia N-terminal que incluye la etiqueta His y el sitio de escisión de trombina está subrayada, la secuencia EndoH está en cursiva)

ES 2 972 511 T3

Identificación de secuencia de ADN que codifica la proteína de fusión EndoF3-EfEndo18A tal como se expresa en E coli (SEQ. ID N.º: 22):

ATGGCTACAGCGCTGGCTGGTTCTAACGGGGTCTGCATCGCGTATTACATCACCGATGGGCGTAATCCGACG
TTCAAATTGAAAGACATCCCGGATAAAGTAGACATGGTAATCTTTTTGGTCTTAAGTATTGGTCATTGCAG
GATACAACCAAATTGCCAGGGGGTACTGGTATGATGGGTTTCGTTTAAATCCTACAAGGACCTGGACACCCAG
ATTTCGTAGTCTTCAAAGCCGTGGAATCAAAGTGTTCGAGAACATTGACGACGACGTCTCATGGCAGTCCCTCG
AAGCCGGGTGGGTTTCGCTTCCGCCGCTGCTTACGGGGATGCTATTAAGAGTATCGTAATTGATAAGTGAAG
CTGGACGGGATTAGCTTGGATATTGAGCATTCGGGGGCTAAACCCAACCCATCCCAACTTTTCCTGGATAT
GCCGCGACAGGATATAATGGCTGGTATTCAGGATCTATGGCAGCCACGCCTGCCTTCTTAATGTTATCTCA
GAGCTTACTAAATACTTTGGTACAACGGCACCGAATAATAAGCAACTTCAGATTGCTTCGGGTATTGACGTA
TATGCCCTGGAATAAAAATCATGGAGAACTTTCGTAATAACTTCAACTACATCCAATTACAGTCATACGGAGCT
AATGTCTCTCGTACTCAACTTATGATGAATTACGCAACGGGAACATAAAAATTCGCCCTCTAAAATGGTT
TTCGGCGCCTACGCAGAGGGTGGCACATAACCAGGCAAATGACGTGGAGGTCGCCAAGTGGACACCTACGCAG
GGCGCAAAGGGCGGTATGATGATCTATACTTACAATTCGAACGTGAGCTATGCAAATGCGGTTTCGCGACGCA
GTGAAAAATGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTCACCACCACCACCACCAC
GAATTCGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGCTTCAACCGTAACCCCTAAA
ACGGTTATGTACGTAGAAGTAAATAACCACGATTTCAACAATGTCGGGAAATACACTCTTGCCGGTACTAAT
CAGCCGGCGTTCGATATGGGTATTATTTTTGCCGCCAACATCAATTATGACACCGTCAATAAGAAACCATAC
CTGTACTTGAACGAGCGCGTACAGCAAACACTGAATGAAGCGGAGACGCAGATCCGTCCGGTCCAGGCACGT
GGAACGAAGGTTTTGCTTTCCATCTTGGGTAATCACGAAGGCGCAGGATTTGCCAATTTTCTACGTATGAG
TCGGCGGACGCTTTCGCCGCGCAACTTGAGCAGGTTGTCAATACGTACCATTTAGACGGGATTGATTTTCGAT
GATGAGTACGCCGAGTACGGAAAAACGGGACCCCTCAGCCGAACAACCTCATCCTTCATCTGGTTACTGCAA
GCTCTTCGCAACCGTCTGGGAAATGATAAACTTATCACTTCTACAACATTGGCCCGGCAGCCGCTAACAGC
AGCGCAAACCCCTCAAATGTCATCTTTGATTGACTATGCCTGGAATCCCTATTATTCGACATGGAACCCCCCA
CAAATTGCAGGTATGCCTGCCTCCCGCCTGGGGGCTTCTGCGGTTGAAGTGGGCGTTAACAGAATCTTGCA
GCACAGTATGCCAAGCGTACTAAGGCTGAGCAGTATGGAATCTATCTGATGTACAATCTGCCAGGAAAAGAT

TCTAGCGCTTATATCTCAGCAGCGACTCAGGAGCTGTATGGGCGCAAGACGAACTATAGCCCCACGGTCCCC
ACTCCGTGATAA

ES 2 972 511 T3

Identificación de secuencia de ADN que codifica la proteína de fusión EndoF2-EfEndo18A tal como se expresa en E coli (SEQ. ID N.º: 23):

ATGGCGGTAAACCTTAGTAATCTTATCGCTTATAAAAAATAGTGACCATCAGATCAGTGCGGGATATTACCGT
ACATGGCGTGACAGCGCCACAGCCAGTGGTAATCTTCCTAGTATGCGTTGGTTGCCAGACTCATTGGACATG
GTAATGGTATTTCCAGACTATACTCCTCCGGAAAATGCGTATTGGAACACACTGAAGACTAACTACGTACCA
TACCTGCATAAGCGTGGCACGAAAGTTATTATCACATTGGGGGACCTTAACTCTGCAACGACCACGGGAGGG
CAAGATTCATTTGGGTATTCATCGTGGGCCAAAGGAATCTATGATAAAATGGGTGGGCGAGTATAATCTTGAT
GGAATCGATATTGACATCGAATCGTCACCGTCCGGTGCGACCTTAACGAAGTTTGTGCGGCAACAAAAGCG
TTGTCAAAGTATTTTGGACCAAAGAGTGGGACAGGCAAGACCTTTGTATACGATACCAATCAGAATCCGACT
AATTTCTTTATCCAACTGCCCCACGCTACAACACTACGTATTTCTTCAAGCATAAGGGCGCTCGACCCTAAT
CTGACGACGGTCTCTGGATTATACGCCCCCTATATTTCAATGAAACAATTTCTGCCCGGCTTCTCTTTTTAC
GAAGAAAACGGTTACCCAGGTAATTATTGGAATGATGTGCGTTACCCCCAGAAGGTACAGGCCGTGCCCTAC
GACTACGCGCGCTGGCAGCCCCGCCACGGGAAAAAAGGAGGGGTGTTTCAGTTATGCCATCGAGCGCGACGCC
CCTCTTACATCGTCAAACGACAATACCCTGCGTGCGCCTAACTTTTCGTGTAACGAAGGACTTAATCAAATTT
ATGAATCCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTCACCACCACCACCACCAC
GAATTCGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGCTTCAACCGTAACCCCTAAA
ACGGTTATGTACGTAGAAGTAAATAACCACGATTTCAACAATGTGGGAAATACACTCTTGCCGGTACTAAT
CAGCCGGCGTTCGATATGGGTATTATTTTTGCCGCCAACATCAATTATGACACCGTCAATAAGAAACCATAC
CTGTACTTGAACGAGCGCGTACAGCAAACACTGAATGAAGCGGAGACGCAGATCCGTCCGGTCCAGGCAGT
GGAACGAAGTTTTTGCTTTCCATCTTGGGTAATCACGAAGGCGCAGGATTTGCCAATTTTCTACGTATGAG
TCGGCGGACGCTTTCCGCGCGCAACTTGAGCAGGTTGTCAATACGTACCATTTAGACGGGATTGATTTGAT
GATGAGTACGCCGAGTACGGAAAAACGGGACCCCTCAGCCGAACAACACTCATCTTCATCTGGTTACTGCAA
GCTCTTCGCAACCGTCTGGGAAATGATAAACTTATCACTTTCTACAACATTTGGCCCCGCGAGCCGCTAACAGC
AGCGCAAACCCCTCAAATGTCATCTTTGATTGACTATGCCTGGAATCCCTATTATTCGACATGGAACCCCCCA
CAAATTCAGGTATGCCTGCCTCCCGCCTGGGGGCTTCTGCGGTTGAAGTGGGCGTTAACCAGAATCTTGCA
GCACAGTATGCCAAGCGTACTAAGGCTGAGCAGTATGGAATCTATCTGATGTACAATCTGCCAGGAAAAGAT
TCTAGCGCTTATATCTCAGCAGCGACTCAGGAGCTGTATGGGCGCAAGACGAACTATAGCCCCACGGTCCCG
ACTCCGTGATAA

Identificación de secuencia de ADN que codifica la proteína de fusión EndoS-EfEndo18A tal como se expresa en E coli (SEQ. ID N.º: 24):

ATGCCGTCAATCGATTGCTGCATTATCTGAGCGAAAACCTCTAAAAAGAATTTAAAGAAGAACTGAGCAAA
GCGGGCCAGGAATCTCAAAAAGTTAAAGAAATCCTGGCAAAAGCTCAGCAAGCCGATAAACAGGCACAAGAA
CTGGCTAAAATGAAAATTCGGAAAAAATCCCAGTAAACCGCTGCATGGTCCGCTGTACGGCGGTTATTTTC
CGTACCTGGCACGATAAAACGTCAGACCCGACCGAAAAAGACAAAGTCAACTCGATGGGCGAACTGCCGAAA
GAAGTGGATCTGGCTTTTATTTCCATGATTGGACCAAAGACTACTCTCTGTTTTGGAAAGAACTGGCAACG
AAACACGTTCCGAAACTGAACAAACAGGGTACGCGTGTCAATTCGTACCATTCCGTGGCGCTTCCCTGGCTGGC
GGTGATAATTCAGGCATCGCGGAAGACACCTCGAAATATCCGAACACGCCGGAAGGTAATAAAGCGCTGGCC

ES 2 972 511 T3

AAAGCAATCGTCGATGAATACGTGTACAAATACAATCTGGACGGCCTGGATGTGGACGTTGAACATGATTCA
ATTCCGAAAAGTGGATAAAAAAGAAGACACCGCCGGCGTGGAAACGTTTCGATCCAGGTTTTTGAAGAAATTGGT
AAACTGATCGGCCGAAAAGGTGTTGATAAAAGCCGTCTGTTTCATCATGGATTCTACCTATATGGCCGACAAA
AATCCGCTGATTGAACCGGGTGCACCGTACATCAACCTGCTGCTGGTCCAGGTGTATGGCAGCCAAGGTGAA
AAAGGCGGTTGGGAACCGGTGTCTAACCGTCCGGAAAAACCATGGAAGAACGCTGGCAGGGCTACTCAAAA
TATATTCGTCGGAAACAATACATGATCGGCTTTTTCGTTCTATGAAGAAAACGCGCAGGAAGGTAATCTGTGG
TACGATATTAATAGTCGCAAAGATGAAGACAAAGCCAACGGCATTAAATACCGATATCACGGGTACCCGTGCG
GAACGCTATGCCCGTTGGCAGCCGAAAACCGGCGGTGTTAAAGGCGGTATTTTTAGCTACCGGATCGATCGT
GACGGTGTGCCCATCAGCCGAAAAAATACGCAAAACAAAAAGAGTTCAAAGATGCTACCACAACATCTTC
CACAGCGATTACAGTGTCTCAAAGCGCTGAAAACCGTGATGCTGAAAGATAAATCTTACGATCTGATCGAC
GAAAAAGATTTTTCCGGACAAAGCGCTGCGCGAAGCCGTTATGGCACAGGTCCGGCACC CGCAAAGGTGACCTG
GAACGTTTTAATGGCACGCTGCGCCTGGATAACCCGGCCATTAGAGCCTGGAAGGTCTGAATAAATTCAAA
AAACTGGCACAACTGGACCTGATTGGCCTGAGCCGTATCACCAAACCTGGATCGCTCTGTGCTGCCGGCCAAC
ATGAAACCGGGTAAAGACACGCTGAAACCGTTCTGGAAACCTACAAAAAGATAACAAAGAAGAACCGGCA
ACGATCCCGCCGGTGTCTCTGAAAGTTTTCCGGCCTGACCGGTCTGAAAGAACCTGGATCTGAGCGGCTTTGAC
CGTGAAACGCTGGCAGGTCTGGATGCGGCCACGCTGACCAGTCTGAAAAAGTTGATATTTCCGGCAATAAA
CTGGACCTGGCGCCGGGTACC GAAAACCGCCAGATTTTTGATACGATGCTGAGTACCATCTCCAACCATGTT
GGCAGCAATGAACAGACCGTCAAATTCGACAAACAAAAACCGACGGGCCACTACCCGGATACGTATGGTAAA
ACCAGCCTGCGTCTGCCGGTCCGCAACGAAAAAGTGGATCTGCAGTCTCAACTGCTGTTTGGCACGGTTACC
AATCAGGGTACCCTGATTAACAGCGAAGCAGATTACAAGGCTTACCAAACCATAAAATCGCGGGTTCGCTCA
TTTGTGGATTGAACTACCCTACAACAACCTTCAAAGTTAGTTACGAAAACCTACCCGTTAAAGTCACGGAT
TCCACCCTGGGCACCACGACCGATAAAACGCTGGCCACCGACAAAGAAGAAACCTACAAAGTCGATTTCTTT
AGCCCGGCAGACAAAACGAAAGCGGTGCATACCGCCAAAGTGATTGTTGGCGATGAAAAACCATGATGGTG
AACCTGGCTGAAGGTGCGACGGTTATCGGCGGTTCCGCGACCCCGGTTAACGCTCGCAAAGTCTTTGATGGC
CAGCTGGGTAGTGAAACCGATAATATTTCCCTGGGTTGGGACTCAAACAGTCGATTATCTTCAAACCTGAAA
GAAGACGGCCTGATCAAACACTGGCGTTTTCTTTAACGATAGTGCCCGCAATCCGGAAACGACCAACAAACCG
ATTCAGGAAGCATCCCTGCAAATCTTCAACATCAAAGATTACAACCTGGACAATCTGCTGGAAAACCCGAAT
AAATTCGATGACGAAAAATACTGGATCACGGTGGATACCTATAGCGCGCAGGGCGAACGTGCTACGGCGTTT
AGTAACACCCTGAAACAATATTACGTCCAAATACTGGCGTGTGGTTTTTCGATACCAAAGGTGACCGCTATAGC
TCTCCGGTCTGCGGAACTGCAGATTCTGGGCTATCCGCTGCCGAATGCTGATACGATCATGAAAACCGTG
ACGACCGCGAAAGAACTGTCACAGCAAAAAGATAAATTTCTCGCAGAAAATGCTGGACGAACGAAAATTAAA
GAAATGGCTCTGGAAACAGCCTGAACAGTAAAATTTTTCGATGTTACGGCGATCAATGCTAACGCTGGTGTG
CTGAAAGACTGTATGAAAAACGCCAACTGCTGAAAAAAGGCGGCGGGCTCTGGCGGCGGCGGCTCTGGC
GGCGGCGGCTCTCACCACCACCACCACGAATTCGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGC
GGCGGCTCTGCTTCAACCGTAACCCCTAAAACGGTTATGTACGTAGAAGTAAATAACCACGATTTCAACAAT
GTCGGGAAATACACTCTTGCCGGTACTAATCAGCCGGCGTTCGATATGGGTATTATTTTTGCCGCCAACATC
AATTATGACACCGTCAATAAGAAACCATACCTGTACTTGAACGAGCGCGTACAGCAAACACTGAATGAAGCG
GAGACGCAGATCCGTCCGGTCCAGGCACGTGGAACGAAGGTTTTGCTTTCCATCTTGGGTAATCACGAAGGC
GCAGGATTTGCCAATTTTCTACGTATGAGTCGGCGGACGCTTTCCCGCGCAACTTGAAGCAGGTTGTCAAT
ACGTACCATTAGACGGGATTGATTTTCGATGATGAGTACGCCGAGTACGGAAAAACGGGACCCCTCAGCCG
AACAACTCATCCTTCATCTGGTTACTGCAAGCTCTTCGCAACCGTCTGGGAAATGATAAATTTACTCTTTT

ES 2 972 511 T3

TACAACATTGGCCCGGCAGCCGCTAACAGCAGCGCAAACCCCTCAAATGTCATCTTTGATTGACTATGCCTGG
AATCCCTATTATTCGACATGGAACCCCCACAAATTGCAGGTATGCCTGCCTCCCGCCTGGGGGCTTCTGCG
GTTGAAGTGGGCGTTAACCCAGAATCTTGCAGCACAGTATGCCAAGCGTACTAAGGCTGAGCAGTATGGAATC
TATCTGATGTACAATCTGCCAGGAAAAGATTCTAGCGCTTATATCTCAGCAGCGACTCAGGAGCTGTATGGG
CGCAAGACGAACTATAGCCCCACGGTCCCGACTCCGTGATAA

Identificación de secuencia de ADN que codifica la proteína de fusión EndoF3-EndoF1 tal como se expresa en E coli (SEQ. ID N.º: 25):

ATGGCTACAGCGCTGGCTGGTTCTAACGGGGTCTGCATCGCGTATTACATCACCGATGGGCGTAATCCGACG
TTCAAATTGAAAGACATCCCGGATAAGTAGACATGGTAATTCTTTTTGGTCTTAAGTATTGGTCATTGCAG
GATACAACCAAATTGCCAGGGGGTACTGGTATGATGGGTTTCGTTTAAATCCTACAAGGACCTGGACACCCAG
ATTTCGTAGTCTTCAAAGCCGTGGAATCAAAGTGTTCAGAACATTGACGACGACGTCTCATGGCAGTCCTCG
AAGCCGGGTGGGTTTCGCTTCCGCCGCTGCTTACGGGGATGCTATTAAGAGTATCGTAATTGATAAGTGAAG
CTGGACGGGATTAGCTTGGATATTGAGCATTTCGGGGGCTAAACCCAACCCTATCCCAACTTTTCTGGATAT
GCCGCGACAGGATATAATGGCTGGTATTTCAGGATCTATGGCAGCCACGCCTGCCTTTCTTAATGTTATCTCA
GAGCTTACTAAATACTTTGGTACAACGGCACCGAATAATAAGCAACTTCAGATTGCTTCGGGTATTGACGTA
TATGCCTGGAATAAAATCATGGGAACTTTTCGTAATAACTTCAACTACATCCAATTACAGTCATACGGAGCT
AATGTCTCTCGTACTCAACTTATGATGAATTACGCAACGGGAACTAATAAAATTCCTCGCCTCTAAAATGGTT
TTCGGCGCCTACGCAGAGGGTGGCACTAACAGGCAAATGACGTGGAGGTCGCCAAGTGGACACCTACGCAG
GGCGCAAAGGGCGGTATGATGATCTATACTTACAATTCGAACGTGAGCTATGCAAATGCGGTTTCGCGACGCA
GTGAAAATGGCGGGCGGCTCTGGCGGGCGGCTCTGGCGGGCGGCTCTCACCACCACCACCACCAC
GAATTCGGCGGGCGGCTCTGGCGGGCGGCTCTGGCGGGCGGCTCTGCGGTAACCGGGACAACGAAG
GCTAACATCAAACCTTTTTAGTTTTACAGAGGTAAACGACACTAATCCGTTGAACAATCTGAACTTTACCTTA
AAAACTCGGGAAAACCTTAGTAGATATGGTAGTGTTATTTTTCCGCGAACATTAACTATGACGCGGCAAC
GATAAGGTCTTCGTATCGAATAATCCGAACGTACAGCATCTTTTGACCAATCGTGCGAAGTACCTTAAGCCG
TTACAAGACAAGGGGATCAAGGTGATTTTTGTCAATCTTAGGGAACCATGATCGCTCCGGGATCGCCAATTTG
AGTACGGCTCGTGCGAAGGCATTTGCTCAGGAACTGAAGAATACTTGCATTTGTATAATTTAGACGGGGTA
TTCTTTGATGATGAGTACTCTGCTTACCAAACGCCACCGCCGAGCGGCTTCGTGACACCCAGTAATAACGCC
GCAGCTCGCCTTGCTTATGAAACAAAGCAGGCTATGCCAAACAAGCTGGTCACGGTGTACGTCTATTCCTCGC
ACTTCGAGTTTTCCACAGCGGTAGACGGGGTCAACGCCGGTCCCTACGTAGACTATGCGATTTCATGACTAC
GGTGGCTCATACGACTTGGCTACTAATTATCCGGGGTTGGCTAAGTCTGGGATGGTGTCTAGTCAGGAG
TTTAACCAGGGCCGTTACGCGACTGCACAAGCATTGCGCAACATTGTGACCAAGGGCTATGGAGGCCACATG
ATCTTTGCCATGGACCCCAATCGTTCTAATTTACGTCAGGGCAAAGTCCCGCACTGAAGCTGATTGCCAAG
GAGCTTTACGGGGATGAGCTTGTGTACAGCAACACTCCTTACAGTAAGGATTGGTGATAA

ES 2 972 511 T3

Identificación de secuencia de ADN que codifica la proteína de fusión EndoF2-EndoF1 tal como se expresa en E coli (SEQ. ID N.º: 26):

ATGGCGGTAAACCTTAGTAATCTTATCGCTTATAAAAATAGTGACCATCAGATCAGTGCGGGATATTACCGT
ACATGGCGTGACAGCGCCACAGCCAGTGGTAATCTTCTAGTATGCGTTGGTTGCCAGACTCATTGGACATG
GTAATGGTATTTCCAGACTATACTCCTCCGGAAAATGCGTATTGGAACACACTGAAGACTAACTACGTACCA
TACCTGCATAAGCGTGGCACGAAAGTTATTATCACATTGGGGACCTTAACTCTGCAACGACCACGGGAGGG
CAAGATTCATTGGGTATTCATCGTGGGCCAAAGGAATCTATGATAAATGGGTGGGCGAGTATAAATCTTGAT
GGAATCGATATTGACATCGAATCGTCACCGTCCGGTGCGACCTTAACGAAGTTTGTGCGGCAACAAAAGCG
TTGTCAAAGTATTTTGGACCAAAGAGTGGGACAGGCAAGACCTTTGTATACGATACCAATCAGAATCCGACT
AATTTCTTTATCCAAACTGCCCCACGCTACAACCTACGTATTTCTTCAAGCATAACGGGCGCTCGACCCTAAT
CTGACGACGGTCTCTGGATTATACGCCCCCTATATTTCAATGAAACAATTTCTGCCCGGCTTCTCTTTTAC
GAAGAAAACGGTTACCCAGGTAATTATTGGAATGATGTGCGTTACCCCAAGACGGTACAGGCCGTGCCCTAC
GACTACGCGCGCTGGCAGCCCCGCCACGGGAAAAAAGGAGGGGTGTTTCAAGTATGCCATCGAGCGCGACGCC
CCTCTTACATCGTCAAACGACAATACCCTGCGTGCGCCTAACTTTTCGTGTAACGAAGGACTTAATCAAAAT
ATGAATCCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTCACCACCACCACCACC
GAATTCGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGCGGTAACGGGACACGAAG
GCTAACATCAAACTTTTTAGTTTTACAGAGGTAAACGACACTAATCCGTTGAACAATCTGAACTTTACCTTA
AAAACTCGGGAAAACCTTAGTAGATATGGTAGTGTATTTTTCCGCGAACATTAACATGACGCGGCCAAC
GATAAGGTCTTCGTATCGAATAATCCGAACGTACAGCATTTTTGACCAATCGTGCGAAGTACCTTAAGCCG
TTACAAGACAAGGGGATCAAGGTGATTTTGTCAATCTTAGGGAACCATGATCGCTCCGGGATCGCCAATTTG
AGTACGGCTCGTGCGAAGGCATTTGCTCAGGAACTGAAGAATACTTGCGATTTGTATAATTTAGACGGGGTA
TTCTTTGATGATGAGTACTCTGCTTACCAAACGCCACCGCCGAGCGGCTTCGTGACACCCAGTAATAACGCC
GCAGCTCGCCTTGCTTATGAAACAAAGCAGGCTATGCCAAACAAGCTGGTACGGTGTACGTCTATTCCC
ACTTCGAGTTTTCCACAGCGGTAGACGGGGTCAACGCCGGGTCTACGTAGACTATGCGATTCATGACTAC
GGTGGCTCATACTGGCTACTAATTATCCGGGGTTGGCTAAGTCTGGGATGGTGATGTCTAGTCAGGAG
TTTAACCAGGGCCGTTACGCGACTGCACAAGCATTGCGCAACATTGTGACCAAGGGCTATGGAGGCCACATG
ATCTTTGCCATGGACCCCAATCGTTCTAATTTACGTCAGGGCAACTGCCCGCACTGAAGCTGATTGCCAAG
GAGCTTTACGGGGATGAGCTTGTGTACAGCAACACTCCTTACAGTAAGGATTGGTGATAA

ES 2 972 511 T3

Identificación de secuencia de ADN que codifica la proteína de fusión EndoS-EndoF1 tal como se expresa en E coli (SEQ. ID N.º: 27):

```
ATGCCGTCAATCGATTTCGCTGCATTATCTGAGCGAAAACCTCTAAAAAAGAATTTAAAGAAGAACTGAGCAAA
GCGGGCCAGGAATCTCAAAAAGTTAAAGAAATCCTGGCAAAAGCTCAGCAAGCCGATAAACAGGCACAAGAA
CTGGCTAAAATGAAAATTCCGGAAAAAATCCCGATGAAACCGCTGCATGGTCCGCTGTACGGCGGTTATTTTC
CGTACCTGGCAGGATAAAAACGTGACACCCGACCGAAAAAGACAAAGTCAACTCGATGGGCGAACTGCCGAAA
GAAGTGGATCTGGCTTTTATTTTCCATGATTGGACCAAAGACTACTCTCTGTTTTGGAAAGAACTGGCAACG
AAACACGTTCCGAAACTGAACAAACAGGGTACGCGTGTTCATTTCGTACCATTCCGTGGCGCTTCCTGGCTGGC
GGTGATAATTCAGGCATCGCGGAAGACACCTCGAAATATCCGAACACGCCGGAAGGTAATAAAGCGCTGGCC
AAAGCAATCGTCGATGAATACGTGTACAAATACAATCTGGACGGCCTGGATGTGGACGTTGAACATGATTCA
ATTCCGAAAGTGGATAAAAAAGAAGACACCGCCGGCGTGGAACGTTTCGATCCAGGTTTTTGAAGAAATTGGT
AAACTGATCGGCCCCGAAAGGTGTTGATAAAAAGCCGTCTGTTTCATCATGGATTCTACCTATATGGCCGACAAA
AATCCGCTGATTGAACGCGGTGCACCGTACATCAACCTGCTGCTGGTCCAGGTGTATGGCAGCCAAGGTGAA
AAAGGCGGTTGGGAACCGGTGTCTAACCGTCCGGAAAAAACCATGGAAGAACGCTGGCAGGGCTACTCAAAA
TATATTCGTCCGGAACAATACATGATCGGCTTTTTCGTTCTATGAAGAAAACGCGCAGGAAGGTAATCTGTGG
TACGATATTAATAGTCGCAAAGATGAAGACAAAGCCAACGGCATTAAATACCGATATCACGGGTACCCGTGCG
GAACGCTATGCCCGTTGGCAGCCGAAAACCGCGGTGTTAAAGGCGGTATTTTTAGCTACGGATCGATCGT
```

ES 2 972 511 T3

GACGGTGTGCGCCATCAGCCGAAAAAATACGCAAAACAAAAAGAGTTCAAAGATGCTACCGACAACATCTTC
CACAGCGATTACAGTGTCTCCAAAGCGCTGAAAACCGTGATGCTGAAAGATAAATCTTACGATCTGATCGAC
GAAAAAGATTTTCCGGACAAAGCGCTGCGCGAAGCCGTTATGGCACAGGTCGGCACCCGCAAAGGTGACCTG
GAACGTTTTAATGGCACGCTGCGCCTGGATAACCCGGCCATTCAGAGCCTGGAAGGTCTGAATAAATTCAAA
AACTGGCACAACCTGGACCTGATTGGCCTGAGCCGTATCACCAAACCTGGATCGCTCTGTGCTGCCGGCCAAC
ATGAAACCGGGTAAAGACACGCTGGAAACCGTTCTGGAAACCTACAAAAAAGATAACAAAGAAGAACCGGCA
ACGATCCCGCCGGTGTCTCTGAAAGTTTCCGGCCTGACCGGTCTGAAAGAAGCTGGATCTGAGCGGCTTTGAC
CGTGAACGCTGGCAGGTCTGGATGCGGCCACGCTGACCAGTCTGGAAAAAGTTGATATTTCCGGCAATAAA
CTGGACCTGGCGCCGGGTACCGAAAACCGCCAGATTTTTGATACGATGCTGAGTACCATCTCCAACCATGTT
GGCAGCAATGAACAGACCGTCAAATTCGACAAAACAAAAACCGACGGGCCACTACCCGGATACGTATGGTAAA
ACCAGCCTGCGTCTGCCGGTGCACAACGAAAAAGTGGATCTGCAGTCTCAACTGCTGTTTGGCACGGTTACC
AATCAGGGTACCCTGATTAACAGCGAAGCAGATTACAAGGCTTACCAAACCATAAAATCGCGGGTTCGCTCA
TTTGTGGATTGAACTACCACTACAACAACCTCAAAGTTAGTTACGAAAACCTACACCGTTAAAGTCACGGAT
TCCACCCTGGGCACCACGACCGATAAAAACGCTGGCCACCGACAAAAGAAGAAACCTACAAAGTCGATTTCTTT
AGCCCGGCAGACAAAACGAAAGCGGTGCATACCGCCAAAGTGATTGTTGGCGATGAAAAACCATGATGGTG
AACCTGGCTGAAGGTGCGACGGTTATCGGCGGTTCCGCAGACCCGGTTAACGCTCGCAAAGTCTTTGATGGC
CAGCTGGGTAGTGAAACCGATAATATTTCCCTGGGTTGGGACTCAAACAGTCGATTATCTTCAAACCTGAAA
GAAGACGGCCTGATCAAACACTGGCGTTTCTTTAACGATAGTGCCCGCAATCCGGAAACGACCAACAAACCG
ATTCAGGAAGCATCCCTGCAAATCTTCAACATCAAAGATTACAACCTGGACAATCTGCTGGAAAACCCGAAT
AAATTCGATGACGAAAAATACTGGATCACGGTGGATACCTATAGCGCGCAGGGCGAACGTGCTACGGCGTTT
AGTAACACCCTGAACAATATTACGTCAAATACTGGCGTGTGGTTTTCGATACCAAAGGTGACCGCTATAGC
TCTCCGGTTCGTGCCGGAACCTGCAGATTCTGGGCTATCCGCTGCCGAATGCTGATACGATCATGAAAACCGTG
ACGACCGCGAAAGAACTGTACAGCAAAAAGATAAATTTCTCGCAGAAAATGCTGGACGAACTGAAAATTTAAA
GAAATGGCTCTGGAAACCAGCCTGAACAGTAAAATTTTCGATGTTACGGCGATCAATGCTAACGCTGGTGTG
CTGAAAGACTGTATTGAAAACGCCAACTGCTGAAAAAAGGCGGCGGGCTCTGGCGGCGGGCTCTGGC
GGCGGCGGCTCTCACCACCACCACCACCACGAATTCGGCGGCGGGCTCTGGCGGCGGGCTCTGGCGGC
GGCGGCTCTGCGGTAACCGGGACAACGAAGGCTAACATCAAACTTTTTAGTTTTACAGAGGTAAACGACACT
AATCCGTTGAACAATCTGAACTTTACCTTAAAAAACTCGGGAAAACCCCTTAGTAGATATGGTAGTGTATTT
TCCGCGAACATTAACATGACCGGCCAACGATAAGGTCTTCGTATCGAATAATCCGAACGTACAGCATCTT
TTGACCAATCGTGCGAAGTACCTTAAGCCGTTACAAGACAAGGGGATCAAGGTGATTTTGTCAATCTTAGGG
AACCATGATCGCTCCGGGATCGCCAATTTGAGTACGGCTCGTGCGAAGGCATTTGCTCAGGAAGTGAAGAAT
ACTTGGGATTTGTATAATTTAGACGGGGTATTCTTTGATGATGAGTACTCTGCTTACCAAACGCCACCGCCG
AGCGGCTTCGTGACACCCAGTAATAACGCCCGCAGCTCGCCTTGCTTATGAAACAAAGCAGGCTATGCCAAAC
AAGCTGGTCACGGTGTACGTCTATTCCCGCACTTCGAGTTTTCCACAGCGGTAGACGGGGTCAACGCCGGG
TCCTACGTAGACTATGCGATTGACTACGGTGGCTCATAAGACTTGGCTACTAATTATCCGGGGTTGGCT
AAGTCTGGGATGGTGTCTAGTCAGGAGTTTAAACAGGGCCGTTACCGGACTGCACAAGCATTGCGCAAC
ATTGTGACCAAGGGCTATGGAGGCCACATGATCTTTGCCATGGACCCCAATCGTTCTAATTTACGTCAGGG
CAACTGCCCGCACTGAAGCTGATTGCCAAGGAGCTTTACGGGGATGAGCTTGTGTACAGCAACACTCCTTAC
AGTAAGGATTTGGTGATAA

ES 2 972 511 T3

Identificación de secuencia de ADN que codifica la proteína de fusión EndoF3-EndoH tal como se expresa en E coli (SEQ. ID N.º: 28):

ATGGCTACAGCGCTGGCTGGTTCTAACGGGGTCTGCATCGCGTATTACATCACCGATGGGCGTAATCCGACG
TTCAAATTGAAAGACATCCCGGATAAAGTAGACATGGTAATCTTTTTGGTCTTAAGTATTGGTCATTGCAG
GATACAACCAAATTGCCAGGGGTTACTGGTATGATGGGTTTCGTTTAAATCCTACAAGGACCTGGACACCCAG
ATTTCGTAGTCTTCAAAGCCGTGGAATCAAAGTGTTCAGAACATTGACGACGACGTCTCATGGCAGTCCTCG
AAGCCGGGTGGGTTTCGCTTCCGCCGCTGCTTACGGGGATGCTATTAAGAGTATCGTAATTGATAAGTGGAAAG
CTGGACGGGATTAGCTTGGATATTGAGCATTTCGGGGGCTAAACCCCAACCCTATCCCAACTTTTCTGGATAT
GCCGCGACAGGATATAATGGCTGGTATTCAGGATCTATGGCAGCCACGCCTGCCTTTCTTAATGTTATCTCA
GAGCTTACTAAATACTTTGGTACAACGGCACCGAATAATAAGCAACTTCAGATTGCTTCGGGTATTGACGTA
TATGCCTGGAATAAAATCATGGAGAACTTTCGTAATAACTTCAACTACATCCAATTACAGTCATACGGAGCT
AATGTCTCTCGTACTCAACTTATGATGAATTACGCAACGGGAATAATAAAATCCCAGCTTAAAATGGTT
TTCGGCGCCTACGCAGAGGGTGGCACTAACAGGCAAATGACGTGGAGGTCGCCAAGTGGACACCTACGCAG
GGCGCAAAGGGCGGTATGATGATCTATACTTACAATTCGAACGTGAGCTATGCAAATGCGGTTTCGCGACGCA
GTGAAAAATGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTCACCACCACCACCACCAC
GAATTCGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGGCGGCGGCGGCTCTGCCCCGGCCCCGGTGAAGCAG
GGGCCGACCTCGGTGGCCTACGTGAGGTGAACAACAACAGCATGCTCAACGTGGCAAGTACACCCTGGCG
GACGGAGGCGGCAACGCCTTCGACGTAGCCGTGATCTTCGCGGCGAACATCAACTACGACACCGGCACGAAG
ACGGCCTACCTGCACTTCAACGAGAACGTGCAGCGCGTCTTGACAACGCTGTCACGCAGATACGGCCGTTG
CAGCAACAGGGCATCAAGGTCCTCCTCTCGGTGCTCGGCAACCACCAGGGCGCCGGGTTTCGCGAACTTCCCC
TCACAGCAGGCGGCTTCGGCGTTCGCGAAGCAGCTCTCGGACGCCGTGGCGAAGTACGGCCTCGACGGCGTC
GACTTCGACGACGAATACGCCGAGTACGGCAACAACGGCACCGCGCAGCCCAACGACAGTTTCGTTTCGTGCAC
CTGGTGACGGCACTGCGCGCGAACATGCCCCACAAGATCATCAGCCTCTACAACATCGGCCCGGCGCGTCC
CGCCTGTCGTACGGCGGTGTCGACGTCTCCGACAAGTTTCGACTACGCCTGGAATCCCTACTACGGCACCTGG
CAGGTCCCCGGCATCGCACTGCCAAGGCGCAGCTGTCGCCGGCGGCGTTCGAGATCGGCCGGACCTCACGG
AGCACCGTTCGCCGACCTCGCCCGTTCGACCGTTCGACGAGGGGTACGGCGTCTATCTGACGTACAACCTCGAC
GGCGGCGATCGCACCGCCGACGTCTCCGCGTTCACCAGGGAGCTGTACGGCAGCGAGGCGGTCCGGACGCCG
TGATAA

Identificación de secuencia de ADN que codifica la proteína de fusión EndoF2-EndoH tal como se expresa en E coli (SEQ. ID N.º: 29):

ATGGCGGTAAACCTTAGTAATCTTATCGCTTATAAAAATAGTGACCATCAGATCAGTGCGGGATATTACCGT
ACATGGCGTGACAGCGCCACAGCCAGTGGTAATCTTCTAGTATGCGTTGGTTGCCAGACTCATTTGGACATG
GTAATGGTATTCAGACTATACTCCTCCGAAAATGCGTATTGGAACACACTGAAGACTAACTACGTACCA
TACCTGCATAAGCGTGGCACGAAAGTTATTATCACATTGGGGGACCTTAACTCTGCAACGACCACGGGAGGG
CAAGATTCTATTGGGTATTCATCGTGGGCCAAAGGAATCTATGATAAATGGGTGGGCGAGTATAATCTTGAT
GGAATCGATATTGACATCGAATCGTCACCGTCCGGTTCGACCTTAACGAAGTTTGTTCGGCAACAAAAGCG
TTGTCAAAGTATTTTGGACCAAGAGTGGGACAGGCAAGACCTTTGTATACGATACCAATCAGAATCCGACT
AATTTCTTTATCCAACTGCCCCAGCTACAACCTACGATTTCTTCAAGCATAACGGCGCTCGACCCTAAT
CTGACGACGGTCTCTGGATTATACGCCCCCTATATTTCAATGAAACAATTTCTGCCCGGCTTCTCTTTTAC

ES 2 972 511 T3

GAAGAAAACGGTTACCCAGGTAATTATTGGAATGATGTGCGTTACCCCCAGAACGGTACAGGCCGTGCCTAC
GACTACGCGCGCTGGCAGCCCGCCACGGGAAAAAAGGAGGGGTGTTTTCAGTTATGCCATCGAGCGCGACGCC
CCTCTTACATCGTCAAACGACAATACCCTGCGTGCGCCTAACTTTTCGTGTAACGAAGGACTTAATCAAAT
ATGAATCCTGGCGGCGCGGCTCTGGCGGCGCGGCTCTGGCGGCGCGGCTCTCACCACCACCACCACCAC
GAATTCGGCGGCGCGGCTCTGGCGGCGCGGCTCTGGCGGCGCGGCTCTGCCCCGGCCCCGGTGAAGCAG
GGGCCGACCTCGGTGGCTACGTGAGGTGAACAACAACAGCATGCTCAACGTGGCAAGTACACCCTGGCG
GACGGAGGCGCAACGCCTTCGACGTAGCCGTGATCTTCGCGGCGAACATCAACTACGACACCGGCACGAAG
ACGGCTACCTGCACTTCAACGAGAACGTGCAGCGCGTCTTGCACAACGTGTACGCAGATACGGCCGTTG
CAGCAACAGGGCATCAAGGTCTCCTCTCGGTGCTCGGCAACCACCAGGGCGCCGGGTTTCGCGAACTTCCCC
TCACAGCAGCGGCTTCGGCGTTTCGCGAAGCAGCTCTCGGACGCCGTGGCGAAGTACGGCCTCGACGGCGTC
GACTTCGACGACGAATACGCCGAGTACGGCAACAACGGCACCGCGCAGCCCAACGACAGTTTCGTTTCGTGCAC
CTGGTGACGGCACTGCGCGGAACATGCCCCACAAGATCATCAGCCTCTACAACATCGCCCCGGCCGCGTCC
CGCTGTCGTACGGCGGTGTCGACGTCTCCGACAAGTTCGACTACGCCTGGAATCCCTACTACGGCACCTGG
CAGGTCCCCGGCATCGCACTGCCCAAGGCGCAGCTGTGCGCGCGGCGGCTCGAGATCGGCCGGACCTCACGG
AGCACCGTCGCCGACCTCGCCCCGTGCGACCGTCGACGAGGGGTACGGCGTCTATCTGACGTACAACCTCGAC
GGCGGCGATCGCACCGCCGACGTCTCCGCGTTACCAGGGAGCTGTACGGCAGCGAGGCGGTCCGGACGCCG
TGATAA

Identificación de secuencia de ADN que codifica la proteína de fusión EndoS-EndoH (o EndoSH) tal como se expresa en E coli (SEQ. ID N.º: 30):

ATGCCGTCAATCGATTTCGCTGCATTATCTGAGCGAAAACCTAAAAAAGAATTTAAAGAAGAACTGAGCAAA
GCGGGCCAGGAATCTCAAAAAGTTAAAGAAATCCTGGCAAAAGCTCAGCAAGCCGATAAACAGGCACAAGAA
CTGGCTAAAATGAAAATTCGGAAAAAATCCCGATGAAACCGCTGCATGGTCCGCTGTACGGCGGTTATTTT
CGTACCTGGCACGATAAAACGTGACACCCGACCGAAAAAGACAAAAGTCAACTCGATGGGCGAACTGCCGAAA
GAAGTGGATCTGGCTTTTATTTTCCATGATTGGACCAAAGACTACTCTCTGTTTTGGAAAGAACTGGCAACG
AAACACGTTCCGAAACTGAACAAACAGGGTACGCGTGTATTTCGTACCATTCGTTGGCGCTTCCTGGCTGGC
GGTGATAATTCAGGCATCGCGGAAGACACCTCGAAATATCCGAACACGCCGGAAGGTAATAAAGCGCTGGCC
AAAGCAATCGTCGATGAATACGTGTACAAATACAATCTGGACGGCCTGGATGTGGACGTTGAACATGATTC
ATTCCGAAAGTGGATAAAAAAGAAGACACCGCCGGCGTGGAAACGTTTCGATCCAGGTTTTTGAAGAAATGGT
AAACTGATCGGCCCGAAAGGTGTTGATAAAAGCCGTCTGTTTCATCATGGATTCTACCTATATGGCCGACAAA
AATCCGCTGATTGAACGCGGTGCACCGTACATCAACCTGCTGCTGGTCCAGGTGTATGGCAGCCAAGGTGAA
AAAGGCGGTTGGGAACCGGTGTCTAACCGTCCGGAAAAAACCATGGAAGAACGCTGGCAGGGCTACTCAAAA
TATATTCGTCCGGAACAATACATGATCGGCTTTTTCGTTCTATGAAGAAAACCGCAGGAAGGTAATCTGTGG
TACGATATTAATAGTCGCAAAGATGAAGACAAAGCCAACGGCATTAATACCGATATCACGGGTACCCGTGCG
GAACGCTATGCCCGTTGGCAGCCGAAAACCGGCGGTGTTAAAGGCGGTATTTTTAGCTACGGATCGATCGT
GACGGTGTGCCCATCAGCCGAAAAAATACGCAAAACAAAAGAGTTCAAAGATGCTACCGACAACATCTTC
CACAGCGATTACAGTGTCTCAAAGCGCTGAAAACCGTGATGCTGAAAGATAAATCTTACGATCTGATCGAC
GAAAAAGATTTTCGGACAAAGCGCTGCGCGAAGCCGTTATGGCACAGGTTCGGCACCCGCAAGGTGACCTG
GAACGTTTTAATGGCACGCTGCGCCTGGATAACCCGGCCATTACAGAGCCTGGAAGGTCTGAATAAATTCAAA
AAACTGGCACAACCTGGACCTGATTGGCCTGAGCCGTATACCAAACCTGGATCGCTCTGTGCTGCCGGCCAAC
ATGAAACCGGGTAAAGACACGCTGGAAACCGTTCTGGAAACCTACAAAAAGATAACAAAGAAGAACCGGCA

ES 2 972 511 T3

ACGATCCCGCCGGTGTCTCTGAAAGTTTCCGGCCTGACCGGTCTGAAAGAAGTGGATCTGAGCGGCTTTGAC
CGTGAAACGCTGGCAGGTCTGGATGCGGCCACGCTGACCAGTCTGGAAAAAGTTGATATTTCCGGCAATAAA
CTGGACCTGGCGCCGGGTACCGAAAACCGCCAGATTTTTGATACGATGCTGAGTACCATCTCCAACCATGTT
GGCAGCAATGAACAGACCGTCAAATTCGACAAACAAAAACCGACGGGCCACTACCCGGATACGTATGGTAAA
ACCAGCCTGCGTCTGCCGGTGCACAACGAAAAAGTGGATCTGCAGTCTCAACTGCTGTTTGGCACGGTTACC
AATCAGGGTACCCTGATTAACAGCGAAGCAGATTACAAGGCTTACCAAACCATAAAATCGCGGGTTCGCTCA
TTTGTGGATTGAACTACCCTACAACAACCTCAAAGTTAGTTACGAAAACCTACACCGTTAAAGTCACGGAT
TCCACCCTGGGCACCACGACCGATAAAAACGCTGGCCACCACGAAAAGAAACCTACAAAGTCGATTTCTTT
AGCCCCGCGACAAAACGAAAGCGGTGCATACCGCCAAAGTGATTGTTGGCGATGAAAAACCATGATGGTG
AACCTGGCTGAAGGTGCGACGGTTATCGGCGGTTCCGCGACCCGGTTAACGCTCGCAAAGTCTTTGATGGC
CAGCTGGGTAGTGAAACCGATAATATTTCCCTGGGTTGGGACTCAAACAGTCGATTATCTTCAAACCTGAAA
GAAGACGGCCTGATCAAACACTGGCGTTTTCTTTAACGATAGTGCCCGCAATCCGGAAACGACCAACAAACCG
ATTCAGGAAGCATCCCTGCAAATCTTCAACATCAAAGATTACAACCTGGACAATCTGCTGGAAAACCCGAAT
AAATTCGATGACGAAAAATACTGGATCACGGTGGATACCTATAGCGCGCAGGGCGAACGTGCTACGGCGTTT
AGTAACACCCTGAACAATATTACGTCAAATACTGGCGTGTGGTTTTTCGATACCAAAGGTGACCGCTATAGC
TCTCCGGTTCGTGCCGGAACCTGCAGATTCTGGGCTATCCGCTGCCGAATGCTGATACGATCATGAAAACCGTG
ACGACCGCGAAAGAAGTGTACAGCAAAAAGATAAATTTCTCGCAGAAAATGCTGGACGAACTGAAAATTA
GAAATGGCTCTGGAAACCAGCCTGAACAGTAAAATTTTCGATGTTACGGCGATCAATGCTAACGCTGGTGTG
CTGAAAGACTGTATTGAAAACGCCAAGTCTGAAAAAGGCGGCGGGCTCTGGCGGGCGGGCTCTGGC
GGCGGGCTCTCACCACCACCACCACGAATTCGGCGGGCGGGCTCTGGCGGGCGGGCTCTGGCGGGC
GGCGGGCTCTGCCCCGGCCCCGGTGAAGCAGGGGCCGACCTCGGTGGCCTACGTGAGGTGAACAACAACAGC
ATGCTCAACGTCCGCAAGTACACCCTGGCGGACGGAGGGCGCAACGCCTTCGACGTAGCCGTGATCTTCGG
GCGAACATCAACTACGACACCGGCACGAAGACGGCCTACCTGCACTTCAACGAGAAGTGCAGCGCGTCTT
GACAACGCTGTACGCAGATACGGCCGTTGCAGCAACAGGGCATCAAGGTCCTCCTCTCGGTGCTCGGCAAC
CACCAGGGCGCCGGGTTTCGCGAAGTCTCCCCTCACAGCAGGCGGCTTCGGCGTTTCGCGAAGCAGCTCTCGGAC
GCCGTGGCGAAGTACGGCCTCGACGGCGTCTGACTTCGACGACGAATACGCCGAGTACGGCAACAACGGCACC
GCGCAGCCCAACGACAGTTCGTTTCGTGCACCTGGTACGGCACTGCGCGCGAACATGCCCGACAAGATCATC
AGCCTCTACAACATCGGCCCGGCCGCTCCCGCCTGTCTGACGGCGGTGTGACGCTCTCCGACAAGTTTCGAC
TAGCCTGGAATCCCTACTACGGCACCTGGCAGGTCCCGGCATCGCACTGCCAAGGCGCAGCTGTCCGCC
GCGGCCGTGAGATCGGCCGACCTCACGGAGCACCGTCCCGACCTCGCCGTGCGACCGTGCAGAGGGG
TAGGGCGTCTATCTGACGTACAACCTCGACGGCGGCGATCGCACCCGCGACGTCTCCGCGTTTACCAGGGAG
CTGTACGGCAGCGAGGCGGTCCGGACGCCGTGATAA

Identificación de secuencia de ADN que codifica la proteína de fusión His₆-EndoS-EndoH (EndoS-EndoH sin porción conectora GS) tal como se expresa en E coli (SEQ. ID N.º: 31):

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCCTGGTGCCGCGCGGCAGCCATATGCCGTCAATC
GATTCGCTGCATTATCTGAGCGAAAACCTCAAAAAAGAATTTAAAGAAGAAGTGCAGCAAGCGGGCCAGGAA
TCTCAAAAAGTTAAAGAAATCCTGGCAAAAGCTCAGCAAGCCGATAAACAGGCACAAGAAGTGGCTAAAATG
AAAATTCGGAAAAAATCCCGATGAAACCGCTGCATGGTCCGCTGTACGGCGGTTATTTCCGTACCTGGCAC
GATAAACGTCAGACCCGACCGAAAAAGACAAAGTCAACTCGATGGGCGAAGTCCGAAAGAAGTGGATCTG
GCTTTTATTTTCCATGATTGGACCAAAGACTACTCTCTGTTTTGGAAAGAAGTGGCAACGAAACACGTTCCG

ES 2 972 511 T3

AAACTGAACAAAACAGGGTACGCGTGTCAATTCGTACCATTCCGTGGCGCTTCCTGGCTGGCGGTGATAAATCA
GGCATCGCGGAAGACACCTCGAAATATCCGAACACGCCGGAAGGTAATAAAGCGCTGGCCAAAGCAATCGTC
GATGAATACGTGTACAAAACAATCTGGACGGCTGGATGTGGACGTTGAACATGATTCAATTCGAAAGTG
GATAAAAAAGAGACACCGCCGGCGTGGAACGTTGATCCAGGTTTTTGAAGAAATTGGTAAACTGATCGGC
CCGAAAGGTGTTGATAAAAGCCGTCTGTTTCATCATGGATTCTACCTATATGGCCGACAAAATCCGCTGATT
GAACGCGGTGCACCGTACATCAACCTGCTGCTGGTCCAGGTGTATGGCAGCCAAGGTGAAAAAGGCGGTTGG
GAACCGGTGTCTAACCGTCCGAAAAAACCATGGAAGAACGCTGGCAGGGCTACTCAAAATATATTCTGCCG
GAACAATACATGATCGGCTTTTTCGTTCTATGAAGAAAACGCGCAGGAAGGTAATCTGTGGTACGATATTAAT
AGTCGCAAAGATGAAGACAAAGCCAACGGCATTAATACCGATATCACGGGTACCCGTGCGGAACGCTATGCC
CGTTGGCAGCCGAAAAACCGGCGGTGTTAAAGGCGGTATTTTTAGCTACGCGATCGATCGTGACGGTGTGCC
CATCAGCCGAAAAATACGCAAAAACAAAAGAGTTCAAAGATGCTACCGACAACATCTTCCACAGCGATTAC
AGTGTCTCCAAAGCGCTGAAAACCGTGATGCTGAAAGATAAATCTTACGATCTGATCGACGAAAAAGATTTT
CCGGACAAAGCGCTGCGCGAAGCCGTTATGGCACAGGTCCGCACCCGCAAAGGTGACCTGGAACGTTTTAAT
GGCACGCTGCGCCTGGATAACC CGCCATT CAGAGCCTGGAAGGTCTGAATAAATTCAAAAACTGGCACAA
CTGGACCTGATTGGCCTGAGCCGTATCACAAAACCTGGATCGCTCTGTGCTGCCGGCCAACATGAAAACCGGT
AAAGACACGCTGAAAACCGTTCTGAAAACCTACAAAAAAGATAACAAAGAAGAACC GGCAACGATCCC GCCG
GTGTCTCTGAAAGTTTTCCGGCTGACC GGCTCTGAAAAGAACTGGATCTGAGCGGCTTTGACCGTGAAACGCTG
GCAGGTCTGGATGCGGCCACGCTGACCAGTCTGGAAAAAGTTGATATTTCCGGCAATAAACTGGACCTGGCG
CCGGGTACCGAAAACCGCCAGATTTTTGATACGATGCTGAGTACCATCTCCAACCATGTTGGCAGCAATGAA
CAGACCGTCAAATTCGACAAAACAAAACCGACGGGCCACTACCCGGATACGTATGGTAAAACCGCCTGCGT
CTGCCGGTCCGCAACGAAAAAGTGATCTGCAGTCTCAACTGCTGTTTTGGCACGGTTACCAATCAGGGTACC
CTGATTAACAGCGAAGCAGATTACAAGGCTTACAAAACCATAAAATCGCGGGTTCGCTCATTTGTGGATTCTG
AACTACCACTACAACAACCTTCAAAGTTAGTTACGAAAACCTACACCGTTAAAGTCACGGATTCCACCCTGGGC
ACCACGACCGATAAAAACGCTGGCCACCGACAAAAGAAGAAACCTACAAAGTCGATTTCTTTAGCCCGGCAGAC
AAAACGAAAGCGGTGCATACCGCCAAAAGTGATTGTTGGCGATGAAAAAACCATGATGGTGAACCTGGCTGAA
GGTGCAGCGTTATCGGCGGTTCCGCAGACCCGGTTAACGCTCGCAAAGTCTTTGATGGCCAGCTGGGTAGT
GAAACCGATAATATTTCCCTGGGTGGGACTCAAACAGTCGATTATCTTCAAACCTGAAAGAAGACGGCCTG
ATCAAACACTGGCGTTTCTTTAACGATAGTCCCCGCAATCCGGAAACGACCAACAAACCGATT CAGGAAGCA
TCCCTGCAATCTTCAACATCAAAGATTACAACCTGGACAATCTGCTGGAAAACCCGAATAAATTCGATGAC
GAAAAATACTGGATCACGGTGGATACCTATAGCGCGCAGGGCGAACGTGCTACGGCGTTTTAGTAACACCCTG
AACAAATATTACGTCCAAATACTGGCGTGTGGTTTTCGATACCAAAGGTGACCGCTATAGCTCTCCGGTCTGTG
CCGAACTGCAGATTTCTGGGCTATCCGCTGCCGAATGCTGATACGATCATGAAAACCGTGACGACCGCGAAA
GAACTGTCACAGCAAAAAGATAAATTTCTCGCAGAAAATGCTGGACGAACTGAAAATTAAGAAATGGCTCTG
GAAACCAGCCTGAACAGTAAAAATTTTCGATGTTACGGCGATCAATGCTAACGCTGGTGTGCTGAAAGACTGT
ATTGAAAAACGCCAACTGCTGAAAAAAGCCCCGGCCCCGGTGAAGCAGGGGCGACCTCGGTGGCCTACGTC
GAGGTGAACAACAACAGCATGCTCAACGTCCGCAAGTACACCCTGGCAGACGGAGGCGCAACGCCTTCGAC
GTAGCCGTGATCTTTCGCGCGAACATCAACTACGACACCGGCACGAAGACGGCCTACCTGCACCTTCAACGAG
AACGTGCAGCGCTCTTTGACAACGCTGT CACGCAGATACGGCCGTTGCAGCAACAGGGCATCAAGGTCTC
CTCTCGGTGCTCGGCAACCACCAGGGCGCCGGTTTCGCGAACTTCCCCTCACAGCAGGCGGCTTCGGCGTTC
GCGAAGCAGCTCTCGACCGCTGGCGAAGTACGGCCTCGACGGCGTGCAGCTTCGACGACGAATACGCCGAG
TACGGCAACAACGGCACCGCGCAGCCCAACGACAGTTTCGTTTCGTGCACCTGGTGACGGCACTGCGCGCGAAC

ES 2 972 511 T3

ATGCCCGACAAGATCATCAGCCTCTACAACATCGGCCCGCCGCGTCCCGCCTGTGCTACGGCGGTGTCGAC
GTCTCCGACAAGTTCGACTACGCCTGGAATCCCTACTACGGCACCTGGCAGGTCCCCGGCATCGCACTGCCC
AAGGCGCAGCTGTCGCCGGCGGCCGTGAGATCGGCCGGACCTCACGGAGCACCGTCCGCCGACCTCGCCCCG
CGCACCGTTCGACGAGGGGTACGGCGTCTATCTGACGTACAACCTCGACGGCGGGGATCGCACCGCCGACGTC
TCCGCGTTCACCAGGGAGCTGTACGGCAGCGAGGCGGTCCGGACGCCGTGATAA

Identificación de secuencia de ADN que codifica His₆-TnGalNAcT(33-421) como se expresa en CHO (SEQ. ID N.º: 32):

ATGAATTTTGGACTGAGGCTGATTTTCTTGGTGTGACCCTGAAAGGCGTCCAGTGTCATCACCATCACCAT
CACTCCCCGCTTCGCACATATCTTTTACTCCATTATACAATGCCACCCAGCCCACTCAGAAACGTGAG
AGGCTGGCAGCTAACTGGCCAAAGAAGATCCCTAGTAATTATATAGAAGATAGCGAAGAGTATAGCATCAAG
AATATTTCTTTGAGCAACCACAACTAGAGCATCTGTGGTACATCCTCCTTCTCTATCACCGAAACGGCA
AGCAAACCTGGATAAGAATATGACCATCCAAGACGGCGCCTTTGCTATGATTAGCCCCGACGCCCTTGCTTATC
ACCAAATTGATGGATAGCATCAAATCTTATGTTACTACCGAGGATGGGGTTAAGAAAGCCGAAGCCGTCGTA
ACTCTCCCCCTCTGTGATAGCATGCCTCCTGACCTTGGTCCTATTACTCTTAACAAAACCGAGCTCGAGCTC
GAATGGGTTGAGAAAAGTTCCCTGAGGTCGAGTGGGGTGGACGTTATAGTCCCCCAACTGCACAGCTAGG
CATCGCGTAGCAATCATAGTCCCGTACCGAGACAGACAGCAACACCTGGCAATCTTCTTAAATCACATGCAC
CCCTTCTGATGAAACAGCAGATCGAATATGGCATCTTTATCGTGGAGCAGGAAGGAAACAAGGACTTTAAC
CGTGCGAAACTTATGAACGTCGGCTTTGTTGAAAGTCAAAAACCTCGTTGCCGAGGGATGGCAGTGTTCGTT
TTTCATGACATAGACCTGCTCCCACTGGACACTAGAAACCTCTATAGCTGCCCGAGACAGCCACGCCACATG
AGCGCTTCCATTGACAACTTCACTTTAAGCTGCCTTACGAAGACATCTTCGGTGGCGTGTGAGCCATGACT
CTGGAACAGTTCACCCGAGTGAATGGATTTTCAAATAAATACTGGGGATGGGGGGGAGAGGACGACGATATG
AGTTATCGGCTTAAGAAAATCACTACCATATTGCAAGATATAAAATGTCCATCGCCCCGATACGCCATGTTG
GACCACAAGAAGTCAACACCCAATCCTAAGCGGTACCAATTACTCTCACAGACCTCAAAGACATTCAGAAA
GACGGGCTGAGCACCCCTGGAATATGAGCTGGTGCAAGTCGTTCAATATCATCTGTATACTCACATCCTGGTT
AATATTGACGAGAGGTCCTGATAA

(la secuencia señal para la secreción está subrayada)

Identificación de secuencia de His₆-TnGalNAcT(33-421) como se expresa en CHO (SEQ. ID N.º: 33):

HHHHHHSPLRITYLYTPLYNATQPTLRNVERLAANWPKKIPSNYIEDSEEYSIKNISLSNHTTRASVVHPPSS
ITETASKLDKNMTIQDGAFAMISPTPLLITKLMDSIKSYVTTEDGVKKAEEAVVTLPLCDSMPDDLGPITLNK
TELELEWVEKKFPEVEWGGGRYSPNCTARHRVAIIVPYRDRQQHLAIFLNHMHPFLMKQQIEYGI FIVEQEG
NKDFNRAKLMNVGFVESQKLVAEGWQC FVFHDI DLLPLDTRNLYSCPRQPRHMSASIDKLFKLPYEDI FGG
VSAMTLEQFTRVNGFSNKYWG WGGEDDDMSYRLKKINYHIARYKMSIARYAML DHHKSTPNPKRYQLLSQTS
KTFQKDGLSTLEYELVQVVQYHLYTHILVNIDERS

REIVINDICACIONES

1. Enzima de fusión con la estructura (1):

5 EndoX-(L)_p-EndoY (1)

en donde EndoX es EndoS, EndoY es una endoglicosidasa distinta de EndoX, L es una porción conectora y p es 0 o 1.

10 2. Enzima según la reivindicación 1, en donde EndoY se selecciona del grupo formado por EndoA, EndoBi, EndoBH, EndoBT, EndoCE, EndoD, EndoE, EfEndo18A, EndoF1, EndoF2, EndoF3, EndoH, EndoLL, EndoM, EndoOm y EndoT, preferiblemente se seleccionada del grupo formado por EndoF1, EndoF2, EndoF3, EfEndo18A y EndoH.

15 3. Enzima según la reivindicación 1 o 2, en donde las endoglicosidasas representadas por EndoX y EndoY tienen actividad endoglicosidasa distinta.

20 4. Enzima según cualquiera de las reivindicaciones 1 - 3, en donde EndoY es EndoF1, EndoD, EndoE, EndoH, EfEndo18A o EndoT, preferiblemente EndoH.

5. Enzima según cualquiera de las reivindicaciones 1 a 4, en donde EndoX tiene al menos un 80 % de identidad de secuencia con cualquiera de las SEQ ID N.º 4 y 5, y EndoY tiene al menos un 80 % de identidad de secuencia con cualquiera de las SEQ ID N.º 6-10.

25 6. Enzima según cualquiera de las reivindicaciones 1 a 5, que tiene al menos un 50 % de identidad de secuencia con cualquiera de las SEQ ID N.º 1, 2, 15, 18 o 21.

7. Enzima según cualquiera de las reivindicaciones 1 - 6, en donde p = 0.

30 8. Enzima según cualquiera de las reivindicaciones 1 - 6, en donde p = 1 y L está compuesto de residuos amino y tiene una longitud de 1 a 100 residuos aminoácidos.

35 9. Enzima según la reivindicación 8, en la que la porción conectora tiene la secuencia (G₄S)_{n1}(H)_r(EF)_s(G₄S)_{n2}, en donde n1 y n2 individualmente son números enteros en el rango de 1 - 10, r es un número entero en el rango de 2 - 10 y s = 0 o 1.

10. Uso de la enzima de fusión según cualquiera de las reivindicaciones 1 a 9 para cortar una glicoproteína.

40 11. Uso según la reivindicación 10, en donde la glicoproteína es un anticuerpo.

12. Método para cortar una glicoproteína, poniendo en contacto la glicoproteína con una enzima de fusión según cualquiera de las reivindicaciones 1 - 9.

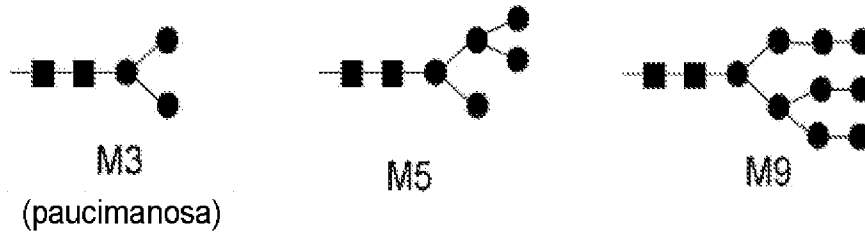
45 13. Método según la reivindicación 12, en donde la glicoproteína comprende al menos un glicano con alto contenido en manosa y al menos un glicano complejo, preferiblemente en el que la glicoproteína comprende además al menos un glicano híbrido.

14. Método según cualquiera de las reivindicaciones 12 - 13, en donde la glicoproteína es un anticuerpo.

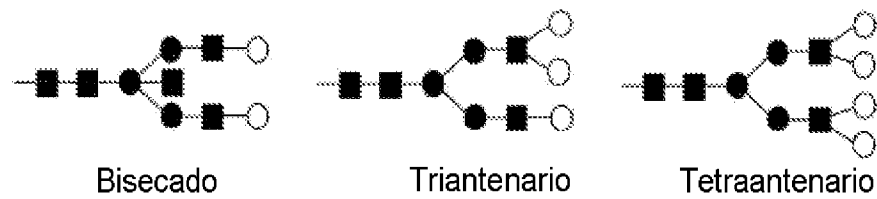
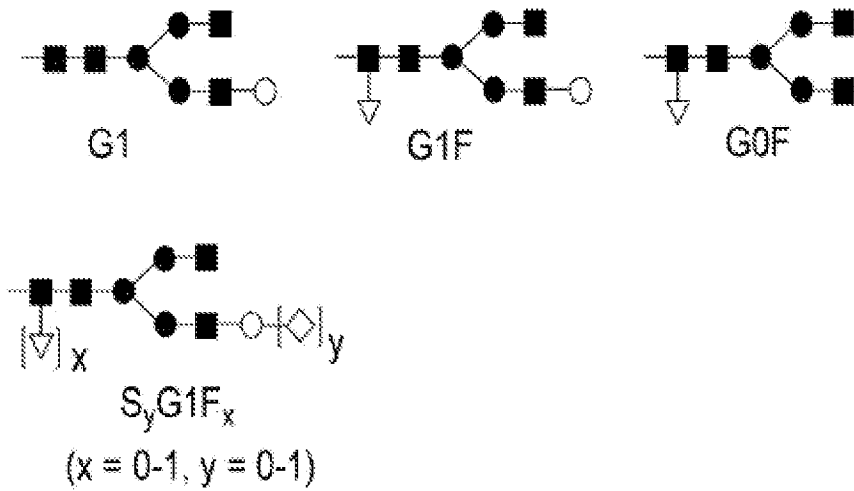
50 15. Método según cualquiera de las reivindicaciones 12 a 14, en donde el contacto se realiza a un pH que es de 0,5 a 3 unidades de pH, preferiblemente de 1 a 2 unidades de pH, diferente del pH óptimo de una o ambas de EndoX y EndoY.

Fig. 1

Manosilado



Complejo



Híbrido

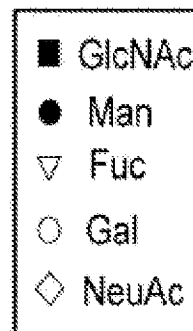
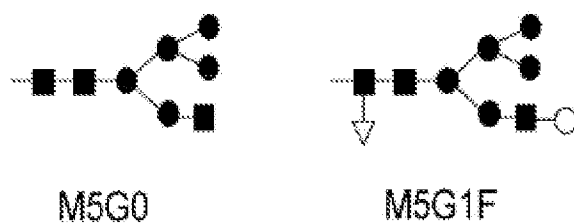


Fig. 2

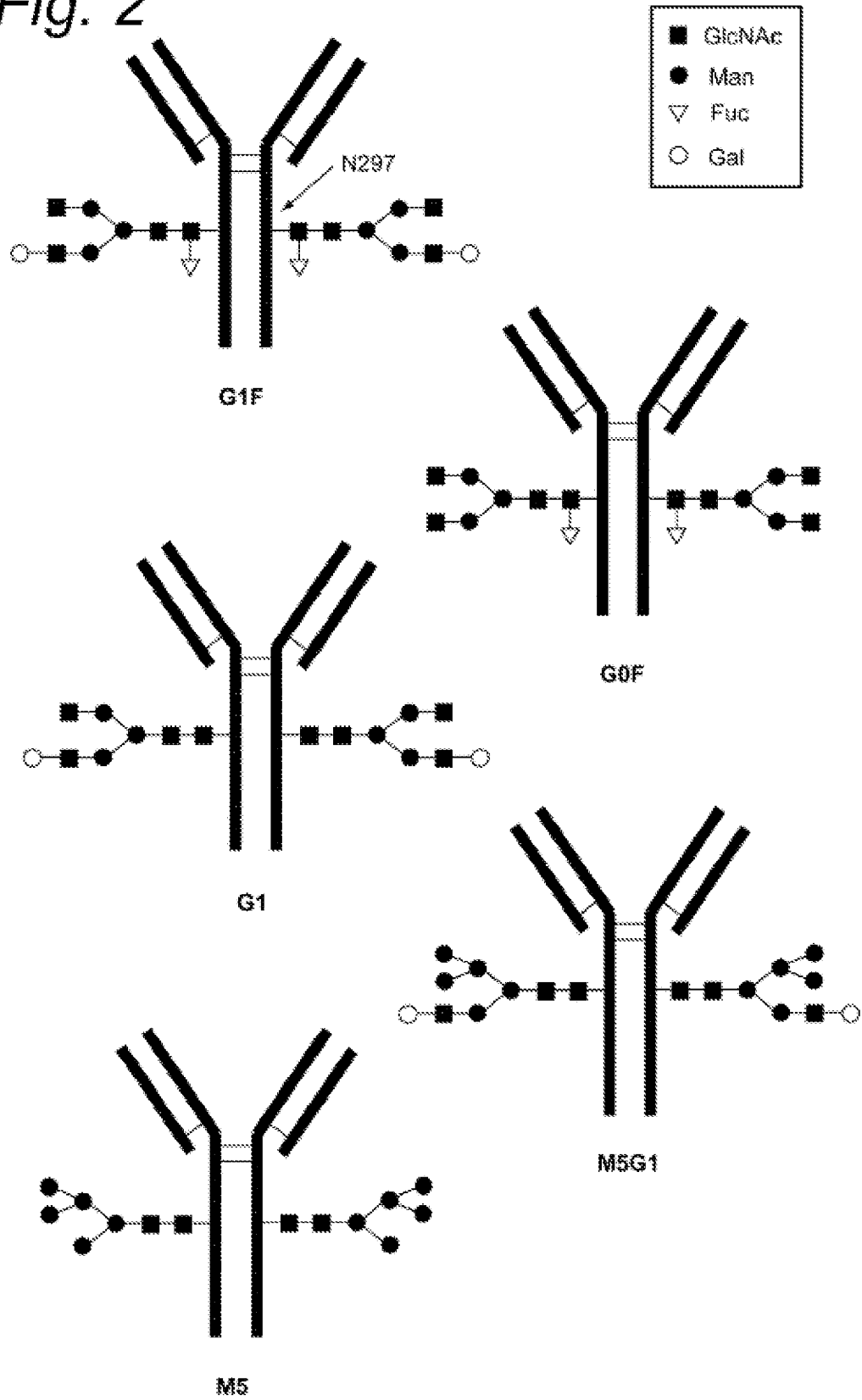


Fig. 3

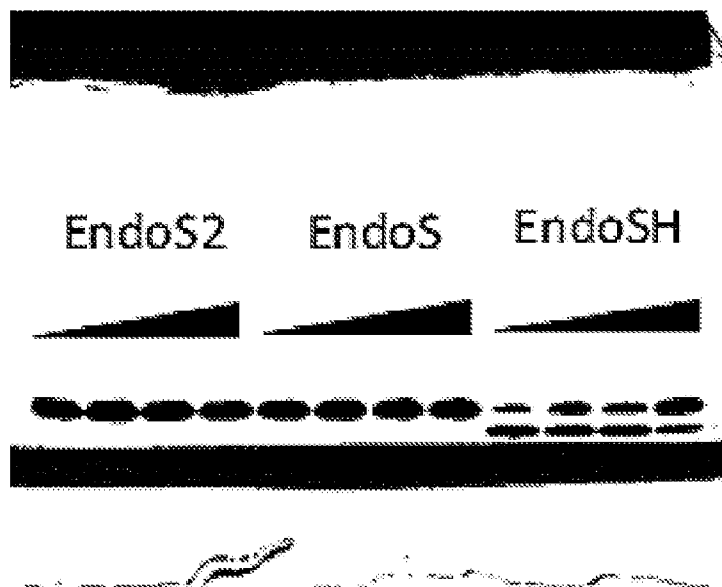
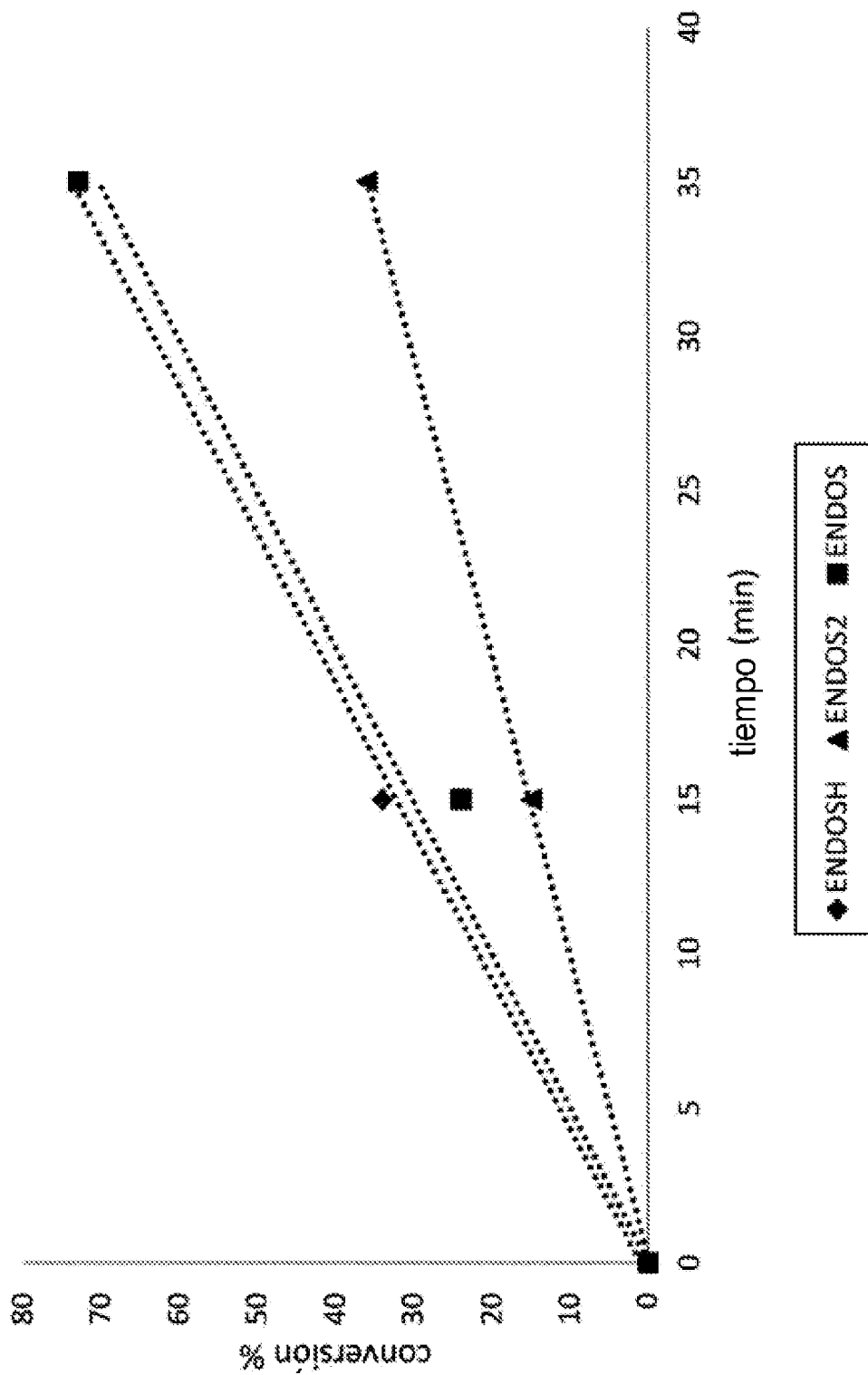


Fig 4



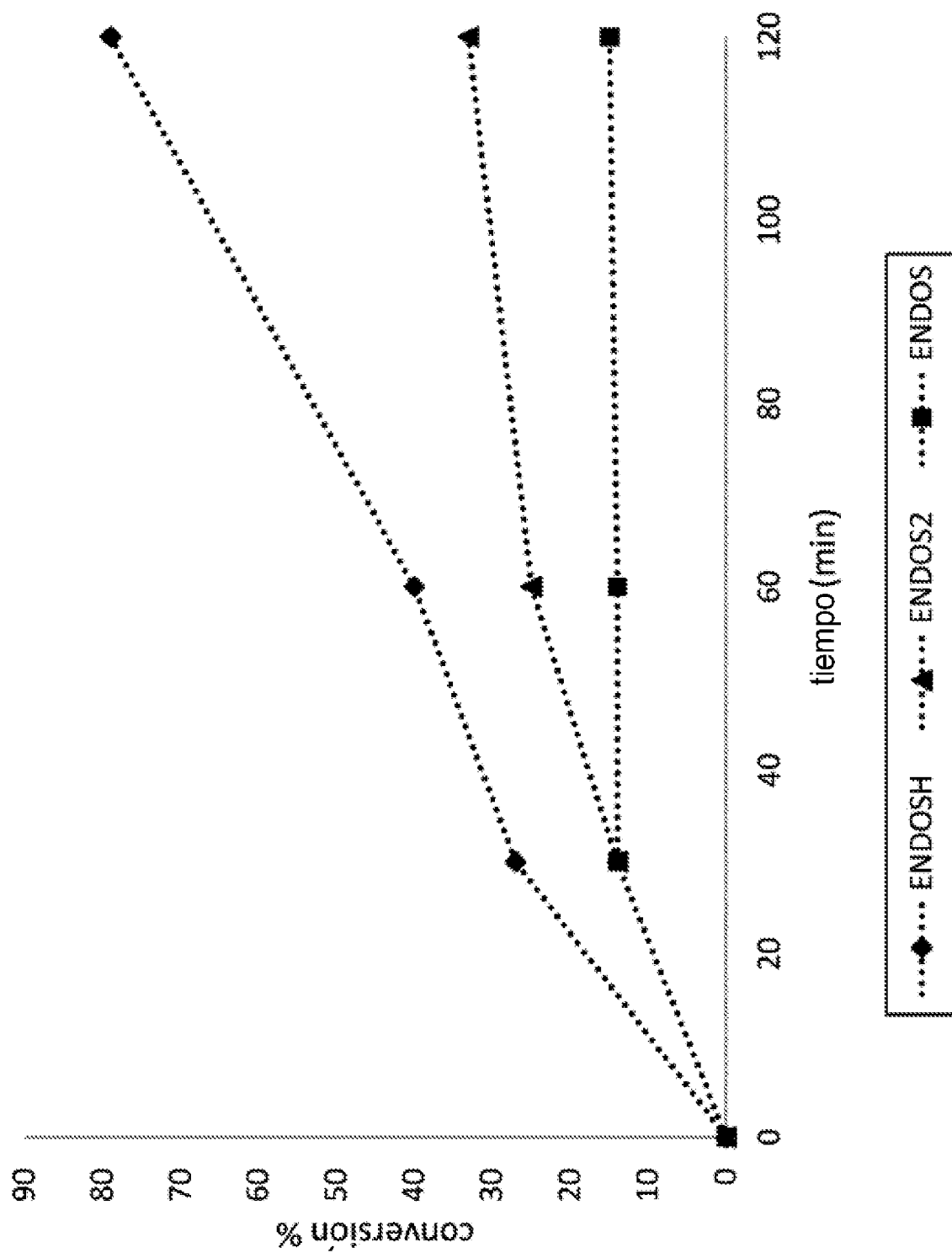


Fig 5