



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 336 790**

(51) Int. Cl.:

C12P 21/00 (2006.01)

C12N 9/10 (2006.01)

C12N 5/10 (2006.01)

C12N 15/81 (2006.01)

C12R 1/84 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **04713388 .9**

(96) Fecha de presentación : **20.02.2004**

(97) Número de publicación de la solicitud: **1597381**

(97) Fecha de publicación de la solicitud: **23.11.2005**

(54) Título: **Producción de glucoproteínas modificadas con múltiples estructuras antenarias.**

(30) Prioridad: **20.02.2003 US 371877**
07.10.2003 US 680963

(45) Fecha de publicación de la mención BOPI:
16.04.2010

(45) Fecha de la publicación del folleto de la patente:
16.04.2010

(73) Titular/es: **GlycoFi, Inc.**
21 Lafayette Street, Suite 200
Lebanon, New Hampshire 03766, US

(72) Inventor/es: **Bobrowicz, Piotr;**
Hamilton, Stephen R.;
Gerngross, Tillman U.;
Wildt, Stefan;
Choi, Byung-Kwon;
Nett, Juergen Hermann y
Davidson, Robert C.

(74) Agente: **Carpintero López, Mario**

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

Producción de glucoproteínas modificadas con múltiples estructuras antenarias.

5 **Campo de la invención**

La presente invención se refiere a procedimientos y composiciones por los que células huésped eucariotas no humanas, como por ejemplo células de hongos unicelulares y multicelulares, pueden modificarse genéticamente para que produzcan proteínas glucosiladas (glucoproteínas) con patrones de glucosilación similares a los de las glucoproteínas producidas por células animales, especialmente células humanas, que son útiles como agentes terapéuticos en seres humanos o animales.

Antecedentes de la invención15 *Rutas de glucosilación en seres humanos y eucariotas inferiores*

Después de que el ADN se transcribe y traduce a proteína, el procesamiento postraduccional posterior supone la unión de restos de azúcares, un procedimiento conocido como glucosilación. Los diferentes organismos producen diferentes enzimas de glucosilación (glucosiltransferasas y glucosidasas), y tienen diferentes sustratos (azúcares de nucleótidos) disponibles, de forma que los patrones de glucosilación así como la composición de los oligosacáridos individuales, incluso en la misma proteína, serán diferentes dependiendo del sistema huésped en el que se esté expresando la proteína particular. Las bacterias habitualmente no glucosilan las proteínas y, si lo hacen, únicamente de forma muy inespecífica (Moens y Vanderleyden (1997) Arch Microbiol. 168 (3):169-175). Los eucariotas inferiores como por ejemplo los hongos filamentosos y las levaduras añaden principalmente las azúcares manosa y manosilfosfato. El glucano resultante se conoce como glucano de tipo "alto en manosa" o manano. Las células vegetales y las células de insectos (como por ejemplo las células Sf9) glucosilan las proteínas de otra forma distinta adicional. Por el contrario, en los eucariotas superiores como por ejemplo los seres humanos, la cadena lateral del oligosacárido en desarrollo puede recortarse para eliminar diversos restos de manosa y alargarse con restos de azúcares adicionales que habitualmente no se encuentran en los *N*-glucanos de eucariotas inferiores. Véase, por ejemplo, Bretthauer *et al.* (1999) Biotechnology and Applied Biochemistry 30:193-200; Martinet, *et al.* (1998) Biotechnology Letters 20:1171-1177; Weikert, *et al.* (1999) Nature Biotechnology, 17:1116-1121; M. Malissard, *et al.* (2000) Biochemical and Biophysical Research Communications 267:169-173; Jarvis, *et al.*, (1998) Current Opinion in Biotechnology 9:528-533; y Takeuchi (1997) Trends in Glycoscience and Glycotechnology 9:S29-S35.

La síntesis de una estructura de oligosacárido de tipo mamífero se inicia con un conjunto de reacciones secuenciadas en cuyo transcurso se añaden y se eliminan restos de azúcares mientras la proteína progresa por la ruta secretora del organismo huésped. Las enzimas que se encuentran a lo largo de la ruta de glucosilación del organismo o célula huésped determinan los patrones de glucosilación resultantes de las proteínas segregadas. Así, el patrón de glucosilación de proteínas resultante expresado en las células huésped de eucariotas inferiores difieren sustancialmente del patrón de glucosilación de las proteínas expresadas en los eucariotas superiores como por ejemplo seres humanos y otros mamíferos (Bretthauer, 1999). La estructura de un *N*-glucano fúngico típico se muestra en la Figura 1A.

Las etapas tempranas de la glucosilación humana pueden dividirse al menos en dos fases diferentes: (i) los oligosacáridos ligados a lípidos Glc₃Man₉GlcNAc₂ se ensamblan mediante un conjunto de reacciones secuenciales en la membrana del retículo endoplasmático (RE) (Figura 13) y (ii) la transferencia de este oligosacárido desde el ancla lipídica doliquilpirofosfato sobre la proteína sintetizada *de novo*. El sitio de la transferencia específica se define mediante un resto asparagina (Asn) de la secuencia Asn-Xaa-Ser/Thr (SEC. ID. N.º: 1 y 2) donde Xaa puede ser cualquier aminoácido excepto prolina (Gavel y von Heijne (1990) Protein Eng. 3:433-42). Se produce un procesamiento adicional mediante glucosidasas y manosidasas en el RE antes de que la glucoproteína en formación sea transferida al aparato de Golgi cis, donde los restos adicionales de manosa son eliminados por las alfa (α)-1,2-manosidasas específicas del aparato de Golgi. El procesamiento continúa mientras la proteína pasa a través del aparato de Golgi. En el aparato de Golgi medio, un número de enzimas modificadoras, que incluyen *N*-acetilglucosaminil transferasas (GnTI, GnTII, GnTIII, GnTIV y GnTV), manosidasa II y fucosiltransferasas, añaden y eliminan restos de azúcares específicos. Finalmente, en el aparato de Golgi trans, las galactosiltransferasas (GalT) y sialiltransferasas (ST) producen una estructura de glucoproteína que se libera desde el aparato de Golgi. Es esta estructura, que se caracteriza por estructuras biantenarias, triantenarias y tetraantenarias, que contienen galactosa, fucosa, *N*-acetilglucosamina y un alto grado de ácido siálico terminal, la que proporciona a las glucoproteínas sus características humanas. La estructura de un *N*-glucano humano típico se muestra en la Figura 1B. Véanse también las Figuras 14 y 15 para las etapas del procesamiento de *N*-glucanos de tipo mamífero.

En todos los eucariotas estudiados hasta la fecha, las glucoproteínas se derivan de un precursor oligosacárido ligado a lípidos común: Gl₃Man₉GlcNAc₂-dolicol-pirofosfato. Dentro del retículo endoplasmático, la síntesis y procesamiento de los oligosacáridos unidos a dolicol pirofosfato son idénticos entre todos los eucariotas conocidos. Sin embargo, el procesamiento posterior del oligosacárido nuclear por las células fúngicas, por ejemplo, las levaduras, difiere significativamente del de los seres humanos al desplazarse por la ruta secretora.

En las levaduras, estas etapas se catalizan mediante manosiltransferasas que residen en el aparato de Golgi, como OChlp, Mntlp y Mnnlp, que secuencialmente añaden azúcares manosa al oligosacárido nuclear. La estructura resultan-

te es indeseable para la producción de proteínas humanoides y por lo tanto es deseable reducir o eliminar la actividad de la manosil transferasa. Los mutantes de *S. cerevisiae*, deficientes en actividad de manosil transferasa (por ejemplo mutantes de *och1* o *mn9*) han demostrado que no son letales y presentan un menor contenido en manosa en el oligosacárido de las glucoproteínas de las levaduras. También podrían tener que eliminarse otros enzimas para el procesamiento de oligosacáridos, como por ejemplo manosilfosfato transferasa dependiendo del patrón de glucosilación particular del huésped.

Precursores de nucleótidos con azúcares

Los *N*-glucanos de las glucoproteínas animales habitualmente incluyen galactosa, fucosa, y ácido siálico terminal. Estas azúcares no se encuentran en las glucoproteínas producidas en levaduras ni en hongos filamentosos. En los seres humanos, se sintetiza todo el abanico de azúcares de nucleótidos (por ejemplo UDP-*N*-acetilglucosamina, UDP-*N*-acetilgalactosamina, CMP-ácido *N*-acetilneuramínico, UDP-galactosa, GDP-fucosa, etc.) en el citosol y se transportan al aparato de Golgi, donde se unen al oligosacárido nuclear mediante glucosiltransferasas. (Sommers y Hirschberg (1981) *J. Cell Biol.* 91(2):A406-A406; Sommers y Hirschberg (1982) *J. Biol. Chem.* 257(18):811-817; Perez and Hirschberg (1987) *Methods in Enzymology* 138:709-715).

Las reacciones de transferencia de glucosilo habitualmente proporcionan un producto secundario que es un nucleósido difosfato o monofosfato. Mientras que los monofosfatos pueden exportarse directamente en un intercambio con las azúcares de nucleósido trifosfato mediante un mecanismo antiporte, los difosfonucleósidos (por ejemplo GDP) tienen que ser escindidos por las fosfatasas (por ejemplo GDPasa) proporcionando nucleósido monofosfatos y fosfato inorgánico antes de ser exportados. Esta reacción es importante para la glucosilación eficaz; por ejemplo, se ha encontrado que la GDPasa de *Saccharomyces cerevisiae* (*S. cerevisiae*) es necesaria para la manosilación. Sin embargo, esa GDPasa tiene una actividad un 90% menor sobre UDP (Berninsone *et al.*, 1994 *J. Biol. Chem.* 269(1):207-211). Los eucariotas inferiores habitualmente carecen de actividad de difosfatasa específica de UDP en el aparato de Golgi dado que no utilizan precursores de azúcares UDP para la síntesis de las glucoproteínas con base en el aparato de Golgi. Se ha encontrado que *Schizosaccharomyces pombe*, una levadura que añade restos de galactosa a los polisacáridos de la pared celular (de UDP-galactosa) tiene una actividad específica de UDPasa, lo que indica la potencial necesidad de dicha enzima (Berninsone *et al.*, 1994). (1994) *J. Biol. Chem.* 269(1):207-211). Es sabido que el UDP es un inhibidor potente de las glucosiltransferasas y la eliminación de este subproducto de la glucosilación puede ser importante para evitar la inhibición de la glucosiltransferasa en la luz del aparato de Golgi (Khatara *et al.* (1974) *Eur. J. Biochem.* 44:537-560). Véase Berninsone *et al.* (1995) *J. Biol. Chem.* 270(24):14564-14567; Beaudet *et al.* (1998) *ABC Transporters: Biochemical, Cellular, and Molecular Aspects* 292: 397-413.

*Procesamiento secuencial de *N*-glucanos mediante actividades enzimáticas compartimentalizadas*

Las azúcares transferasas y las glucosidasas tapizan la superficie interna (luz) del RE y del aparato de Golgi y así proporcionan una superficie "catalítica" que permite el procesamiento secuencial de las glucoproteínas al ir pasando a través de la red del RE y del aparato de Golgi. Los múltiples compartimientos del aparato de Golgi cis, medio, y trans y la red trans del aparato de Golgi (TGN), proporcionan las diferentes localizaciones en las que puede tener lugar la secuencia ordenada de reacciones de glucosilación. Cuando una glucoproteína va pasando por la síntesis en el RE a la maduración completa en el aparato de Golgi trans o TGN, se ve expuesta secuencialmente a diferentes glucosidasas, manosidasas y glucosiltransferasas de tal forma que puede sintetizarse una estructura de carbohidrato específica. Se ha dedicado mucho trabajo a revelar el mecanismo exacto por el cual estas enzimas son retenidas y ancladas a su organelo respectivo. La imagen que se está obteniendo es compleja, pero los indicios sugieren que la región del tallo, la región que atraviesa la membrana y la cola citoplasmática de forma individual o conjunta dirigen las enzimas a la membrana de los organelos individuales y así localizan el dominio catalítico asociado a ese sitio (véase, por ejemplo, Gleeson (1998) *Histochem. Cell Biol.* 109:517-532).

En algunos casos, se encontró que estas interacciones específicas funcionaban entre especies. Por ejemplo, se demostró que el dominio que atraviesa la membrana de $\alpha 2,6$ -ST de las ratas, una enzima que se sabe está localizada en el trans-aparato de Golgi del animal, localiza también un gen testigo (invertasa) en el aparato de Golgi de las levaduras (Schwientek *et al.* (1995) *J. Biol. Chem.* 270(10):5483-9). Sin embargo, este mismo dominio que atraviesa la membrana como parte de una $\alpha 2,6$ -ST de longitud completa se mantenía en el RE y no se transportaba al aparato de Golgi de la levadura (Kretdorn *et al.* (1994) *Eur. J. Biochem.* 220(3):809-17). Una GalT de longitud completa humana ni siquiera se sintetiza en las levaduras, a pesar de los elevados niveles de transcripción que podían demostrarse. Por el contrario, la región transmembrana de la misma GalT humana fusionada a un testigo de invertasa fue capaz de dirigir la localización al aparato de Golgi de la levadura, aunque con niveles bajos. Schwientek y colaboradores han demostrado que fusionar 28 aminoácidos de una manosiltransferasa (Mnt1) de levadura, una región que contiene una cola citoplasmática, una región transmembrana y ocho aminoácidos de la región del tallo, al dominio catalítico de una GalT humana es suficiente para la localización en el aparato de Golgi de una GalT activa. Otras galactosiltransferasas parecen depender de interacciones con enzimas que residen en organelos particulares dado que después de eliminar su región transmembrana, todavía son capaces de localizarse adecuadamente.

- La localización inadecuada de una enzima de glucosilación puede evitar el funcionamiento adecuado de la enzima en la ruta. Por ejemplo *Aspergillus nidulans*, que tiene numerosas α -1,2-manosidasas (Eades y Hintz, 2000 gen 255 (1):25-34), no añade GlcNAc a $\text{Man}_5\text{GlcNAc}_2$ cuando está transformado con el gen GnTI de conejo, a pesar de al nivel elevado global de actividad de GnTI (Kalsner *et al.* (1995) Glycoconj. J. 12(3):360-370). GnTI, aunque se exprese de forma activa, puede estar localizado incorrectamente de tal forma que la enzima no esté en contacto con ambos de sus sustratos: UDP-GlcNAc₂ y un sustrato $\text{Man}_5\text{GlcNAc}_2$ productivo (no todas las estructuras $\text{Man}_5\text{GlcNAc}_2$ son productivas; véase más adelante). De forma alternativa, el organismo huésped puede no proporcionar un nivel adecuado de UDP-GlcNAc en el aparato de Golgi o la enzima puede estar localizada apropiadamente pero sin embargo ser inactiva en su nuevo entorno. Además, las estructuras $\text{Man}_5\text{GlcNAc}_2$ presentes en las células huésped pueden diferir en estructura de la $\text{Man}_5\text{GlcNAc}_2$ que se encuentra en los mamíferos. Maras y colaboradores encontraron que aproximadamente un tercio de los *N*-glucanos de la celobiohidrolasa I (CBHI) obtenidos de *T. reesei* pueden ser recortados para dar $\text{Man}_5\text{GlcNAc}_2$ mediante la 1,2 manosidasa de *A. saitoi in vitro*. Menos del 1% de esos *N*-glucanos, sin embargo, podrían valer de sustrato productivo para GnTI. Maras *et al.* (1997) Eur. J. Biochem. 249:701-707. La mera presencia de $\text{Man}_5\text{GlcNAc}_2$, por lo tanto, no asegura que pueda lograrse un procesamiento adicional *in vivo* de $\text{Man}_5\text{GlcNAc}_2$. Lo que se requiere es la formación de una estructura de $\text{Man}_5\text{GlcNAc}_2$ productiva, que reaccione con GnTI. Aunque $\text{Man}_5\text{GlcNAc}_2$ podría producirse en la célula (aproximadamente 27 mol %), sólo una pequeña parte podría convertirse en $\text{Man}_5\text{GlcNAc}_2$ (menos de aproximadamente 5%, véase Chiba *et al.* documento WO 01/14522).
- Hasta la fecha, no existe ninguna forma fiable de predecir si una glucosiltransferasa particular o manosidasa particular expresadas de forma heteróloga en un eucariota inferior será (1), traducida suficientemente (2), catalíticamente activa o (3) estará localizada en el organelo apropiado en la ruta de secreción. Debido a que los tres son necesarios para afectar los patrones de glucosilación en los eucariotas inferiores, sería deseable un esquema sistemático para lograr la función catalítica deseada y la retención apropiada de enzimas en ausencia de herramientas de predicción, que actualmente no están disponibles.

Producción de glucoproteínas terapéuticas

- Un número significativo de proteínas aisladas en los seres humanos o animales son modificadas postraduccionalmente, siendo la glucosilación una de las modificaciones más significativas. Se estima que un 70% de todas las proteínas terapéuticas están glucosiladas y así actualmente se basan en un sistema de producción (es decir, células huésped) que sea capaz de glucosilar de un modo similar a los seres humanos. Diversos estudios han demostrado que la glucosilación desempeña un papel importante para determinar (1) la capacidad inmunógena, (2) propiedades farmacocinéticas, (3) desplazamiento, y (4) eficacia de proteínas terapéuticas. Por lo tanto no es sorprendente que la industria farmacéutica haya realizado un esfuerzo sustancial para desarrollar procedimientos para obtener glucoproteínas que lo más “humanoides” o “humanizadas” posible. Hasta la fecha, la mayoría de las glucoproteínas se producen en un sistema huésped mamífero. Esto puede suponer la genomanipulación de dichas células de mamífero para mejorar el grado de sialilación (*es decir*, la adición terminal de ácido siálico) de las proteínas expresadas por las células, que es sabido que mejora las propiedades farmacocinéticas de dichas proteínas. De forma alternativa, se puede mejorar el grado de sialilación mediante la adición *in vitro* de dichas usando glucosiltransferasas conocidas y sus respectivas azúcares nucleotídicas (por ejemplo, 2,3-sialiltransferasa y CMP-ácido siálico).

- Aunque la mayoría de los eucariotas superiores realizan reacciones de glucosilación que son similares a las que se encuentran en los seres humanos, las proteínas humanas recombinantes expresadas en los sistemas huésped mencionados anteriormente de forma invariable difieren de sus homólogos humanos “naturales” (Raju *et al.* (2000) Glycobiology 10(5):477-486). Por lo tanto, se ha desarrollado un extenso trabajo para encontrar formas de mejorar el “carácter humano” de las proteínas creadas en estos sistemas de expresión. Esto incluye la optimización de las condiciones de fermentación y la modificación genética de los huéspedes para la expresión de proteínas introduciendo los genes que codifican las enzimas implicadas en la formación de glucoformas humanoides. Goochee *et al.* (1999) Biotechnology 9(12):1347-55; Andersen y Goochee (1994) Curr Opin Biotechnol. 5(5):546-49; Wemer *et al.* (1998) Arzneimittelforschung. 48(8):870-80; Weikert *et al.* (1999) Nat Biotechnol. 17(11):1116-21; Yang y Butler (2000) Biotech. Bioeng. 68:370-80. Los problemas inherentes asociados a todos los sistemas de expresión mamíferos no se han resuelto.

Producción de glucoproteínas usando microorganismos eucariotas

- Aunque la estructura de oligosacárido transferida a una proteína del retículo endoplasmático es básicamente idéntica en los mamíferos y eucariotas inferiores, se han encontrado diferencias sustanciales en las reacciones de procesamiento posteriores que se producen en el aparato de Golgi de los hongos y mamíferos. De hecho, incluso entre diferentes eucariotas inferiores existe una gran variedad de estructuras de glucosilación. Esto ha evitado históricamente el uso de eucariotas inferiores como huéspedes para la producción de glucoproteínas humanas recombinantes a pesar de ventajas por lo demás notables con respecto a los sistemas de expresión mamíferos.

- Las glucoproteínas terapéuticas producidas en un microorganismo huésped como por ejemplo levaduras utilizando la ruta de glucosilación endógenas del huésped difieren estructuralmente de las producidas en las células de mamífero y habitualmente muestran una eficacia terapéutica muy reducida. Dichas glucoproteínas son habitualmente inmunó-

genas en los seres humanos y demuestran una semivida reducida (y por lo tanto bioactividad) *in vivo* después de la administración (Takeuchi (1997) Trends in Glycoscience and Glycotechnology 9:S29-S35). Los receptores específicos en los seres humanos y animales (es decir, los receptores de manosa de los macrófagos) pueden reconocer los restos de manosa terminales y promover el rápido aclaramiento de la glucoproteína exógena en el torrente sanguíneo.

5 Los efectos adversos adicionales pueden incluir cambios en el plegamiento de las proteínas, solubilidad, susceptibilidad a proteasas, tráfico, transporte, compartimentalización, secreción, reconocimiento por otras proteínas o factores, capacidad antígena o capacidad alérgica.

Las levaduras y los hongos filamentosos se han usado ambos con éxito para la producción de proteínas recombinantes, tanto intracelulares como segregadas (Cereghino y Cregg (2000) FEMS Microbiology Reviews 24(1):45-66; Harkki *et al.* (1989) BioTechnology 7(6):596; Berka *et al.* (1992) Abstr. Papers Amer. Chem. Soc. 203:121-BIOT; Svetina *et al.* (2000) J. Biotechnol. 76(2-3):245-251). Diversas levaduras, como por ejemplo *K. lactis*, *Pichia pastoris*, *Pichia methanolica*, y *Hansenula polymorpha*, han desempeñado papeles particularmente importantes como sistemas de expresión eucariotas porque son capaces de crecer a altas densidades celulares y/o de segregar grandes cantidades de proteína recombinante. Del mismo modo, se han usado hongos filamentosos, como por ejemplo *Aspergillus niger*, *Fusarium sp*, *Neurospora crassa* y otros, eficazmente para producir glucoproteínas a escala industrial. Sin embargo, como se ha indicado anteriormente, las glucoproteínas que se expresan en cualquiera de estos microorganismos eucariotas difieren sustancialmente de las de los animales en la estructura de los *N*-glucanos. Esto ha evitado que usara levadura u hongos filamentosos como huéspedes para la producción de muchas glucoproteínas terapéuticas.

Aunque la glucosilación en las levaduras y hongos es muy diferente a la de los seres humanos, comparten algunos elementos comunes. La primera etapa, la transferencia de la estructura nuclear del oligosacárido a la proteína en formación, está muy conservada en todos los eucariotas entre otros en las levaduras, hongos, plantas y los seres humanos (compárense las Figuras 1A y 1B). El procesamiento posterior del oligosacárido nuclear, sin embargo, difiere considerablemente en las levaduras y supone la adición de varios azúcares de manosa. Esta etapa es catalizada por manosiltransferasas situadas en el aparato de Golgi (por ejemplo: *OCH1*, *MNT1*, *MNN1*, etc.), que secuencialmente añaden azúcares de manosa al oligosacárido nuclear. La estructura resultante es indeseable para la producción de proteínas humanoides y por lo tanto es deseable reducir o eliminar la actividad de la manosiltransferasa. Los mutantes de *S. cerevisiae* deficientes en actividad de manosiltransferasa (por ejemplo, *mutantes en och1* o *mnn9*) han demostrado no ser letales y presentan un menor contenido en manosa en el oligosacárido de levaduras glucoproteínas. También podrían tener que eliminarse otras enzimas para el procesamiento de oligosacáridos, como por ejemplo la manosilfosfato transferasa dependiendo del patrón de glucosilación endógeno particular del huésped. Después de reducir las reacciones de glucosilación endógenas no deseadas, debe genomanipularse la función de los *N*-glucanos complejos en el sistema del huésped. Esto requiere la expresión estable de varias enzimas y transportadores de azúcares-nucleótidos.

35 Además, deben localizarse estas enzimas de forma tal que se asegure un procesamiento secuencial de la estructura con glucosilación para la maduración.

Se han realizado varias tentativas para modificar las rutas de glucosilación de los microorganismos eucariotas para proporcionar glucoproteínas más adecuadas para usar como agentes terapéuticos en mamíferos. Por ejemplo, se han clonado y expresado varias glucosiltransferasas por separado en *S. cerevisiae* (GalT, Gn TI), *Aspergillus nidulans* (GnTI) y otros hongos (Yoshida *et al.* (1999) Glycobiology 9(1):53-B, Kalsner *et al.* (1995) Glycoconj. J. 12(3):360-370). Sin embargo, no se obtuvieron *N*-glucanos que imitan los formados en las células humanas.

Las levaduras producen una variedad de manosiltransferasas por ejemplo 1,3-manosiltransferasas como *MNN1* en *S. cerevisiae*; Graham y Emr (1991) J. Cell. Biol. 114(2):207-21 B), 1,2-manosiltransferasas (por ejemplo, de la familia KTR/KRE de *S. cerevisiae*), 1,6-manosiltransferasas (por ejemplo, *OCH1* de *S. cerevisiae*), manosilfosfato transferasas y sus reguladores (por ejemplo, *MNN4* y *MNN6* de *S. cerevisiae*), y enzimas adicionales que están implicadas en las reacciones de glucosilación endógenas. Muchos de estos genes se han suprimido de forma individual, dando lugar a organismos viables que tienen perfiles de glucosilación alterados. Se muestran ejemplos en la Tabla 1.

Tabla 1. Ejemplos de cepas de levadura que tienen manosilación alterada

Cepa	<i>N</i> -glucano (tipo silvestre)	Mutación	<i>N</i> -glucano (mutante)	Referencia
<i>S. pombe</i>	Man _{>9} GlcNAc ₂	<i>OCH1</i>	Man ₅ GlcNAc ₂	Yoko-o et al. (2001) FEBS Lett.

				489(1): 75-80
5	<i>S. cerevisiae</i>	Man _{>9} GlcNAc ₂	<i>OCH1</i> /MNN1	Man ₈ GlcNAc ₂
10				Nakanishi-Shindo et al. (1993) J. Biol. Chem. 268(35):26338-26345
15	<i>S. cerevisiae</i>	Man _{>9} GlcNAc ₂	<i>OCH1</i> /MNN1/MNN4	Man ₅ GlcNAc ₂
20				Chiba et al. (1998) J. Biol. Chem. 273, 26298-26304
25	<i>P. pastoris</i>	Hiperglucosiladas	<i>OCH1</i> (delección completa)	No hiperglucosiladas
30				Welfide, publicación de solicitud de patente japonesa n.º 8-336387
35	<i>P. pastoris</i>	Man _{>8} GlcNAc ₂	<i>OCH1</i> (alteración)	Man _{>8} GlcNAc ₂
				Contreras et al. WO 02/00856 A2

La publicación de solicitud de patente japonesa n.º 8-336387 describe la delección de un homólogo de *OCH1* en *Pichia pastoris*. En *S. cerevisiae*, *OCH1* codifica una 1,6-manosiltransferasa, que añade una manosa a la estructura del glucano Man₈GlcNAc₂ proporcionando Man₉GlcNAc₂. La estructura de Man₉GlcNAc₂ que contiene tres restos 1,6 manosa después se convierte en sustrato para más 1,2-, 1,6-, y 1,3- manosiltransferasas *in vivo*, produciendo glucoproteínas hipermanosiladas que son características de *S. cerevisiae* y que habitualmente pueden tener 30-40 restos de manosa por *N*-glucano. Debido a que Ochlp inicia la transferencia de 1,6 manosa al núcleo de Man₃GlcNAc₂, a menudo se denomina "1,6-manosiltransferasa de iniciación" para diferenciarla de otras 1,6-manosiltransferasas que actúan posteriormente en el aparato de Golgi. En una cepa mutante para *och1*, *mnn1* y *mnn4* de *S. cerevisiae*, se acumulan las proteínas glucosiladas con Man₈GlcNAc₂ y no se produce hipermanosilación. Sin embargo, Man₃GlcNAc₂ no es sustrato para las glucosiltransferasas de mamíferos, como por ejemplo UDPGlcNAc transferasa I human, y por consiguiente, el uso de esa cepa mutante, en sí, no es útil para producir proteínas de tipo mamífero, *es decir*, las que presentan patrones de glucosilación complejos o híbridos.

Se pueden recortar las estructuras de Man₃GlcNAc₂ para obtener el isómero Man₅GlcNAc₂ en *S. cerevisiae* (aunque todavía queda por demostrar una eficacia de recorte alta superior a 50% *in vivo*) genomaniplando una manosidasa fúngica de *A. saitoi* en el retículo endoplasmático (RE). Los inconvenientes de estas estrategia son dobles: (1) no queda claro si las estructuras Man₅GlcNAc₂ formadas de hecho se forman *in vivo* (o si han sido segregadas y modificadas posteriormente por las manosidasas presentes en el exterior de la célula); y (2) no queda claro si cualquier estructura de Man₅GlcNAc₂ formada, de hecho *in vivo*, es la isoforma correcta para que sea un sustrato productivo para la posterior modificación con *N*-glucanos por la GlcNAc transferasa I (Maras *et al.* (1997) Eur. J. Biochem. 249:701-707).

Con el objetivo de proporcionar una glucoproteína más humanoide derivada de un huésped fúngico, la patente de Estados Unidos n.º 5.834.251 describe un procedimiento para producir una glucoproteína híbrida derivada de *Trichoderma reesei*. Un *N*-glucano híbrido tiene sólo restos de manosa en la ramificación Man_α1-6 de la estructura de manosa nuclear y una o dos antenas complejas en la ramificación Man_α1-3. Aunque esta estructura es útil, el procedimiento tiene la desventaja de que deben realizarse numerosas etapas enzimáticas *in vitro*, lo que es costoso y lleva mucho tiempo. Las enzimas aisladas son caras de preparar y requieren sustratos costosos (por ejemplo, UDP-GlcNAc). El procedimiento tampoco permite producir glucanos complejos sobre una proteína deseada.

Actividad intracelular de manosidasa implicada en el recorte de N-glucanos

La actividad de alfa-1,2-manosidasa es necesaria para recortar Man₃GlcNAc₂ para formar Man₅GlcNAc₂, que es un intermedio principal para la formación de N-glucanos complejos en los mamíferos. Trabajos anteriores han demostrado que la α -1,2-manosidasa murina, fúngica y humana puede expresarse en la levadura metilotrófica *P. pastoris* y demuestran actividad de recorte de Man₃GlcNAc₂ a Man₅GlcNAc₂ (Lal *et al.* (1998) *Glycobiology* 8(10):981-95; Tremblay *et al.* (1998) *Glycobiology* 8(6):585-95, Callewaert *et al.* (2001) *FEBS Lett.* 503(2-3):173-8). Sin embargo, hasta la fecha, no se ha reseñado un gran nivel de recorte *in vivo* de Man₃GlcNAc₂ a Man₅GlcNAc₂ en una glucoproteína segregada de *P. pastoris*.

Además, la mera presencia de una α -1,2-manosidasa en la célula no asegura, por sí sola, un recorte intracelular adecuado de Man₃GlcNAc₂ a Man₅GlcNAc₂. (Véase, por ejemplo, Contreras *et al.* documento WO 02/00856 A2, en el que una manosidasa marcada con HDEL de *T. reesei* se localiza principalmente en el RE y se coexpresa con una proteína testigo de hemaglutinina (HA) de influenza, sobre la que prácticamente no podía detectarse Man₅GlcNAc₂. Véase también Chiba *et al.* (1998) *J. Biol. Chem.* 273(41):26298-26304, en el que una fusión de los dominios transmembrana de α -1,2-manosidasa/*Och1p* quimérica localizada en el RE, aparato de Golgi cis y el citosol de *S. cerevisiae*, no presentaba actividad de recorte de manosidasas). Por consiguiente, la mera localización de una manosidasa en el RE o aparato de Golgi es insuficiente para asegurar la actividad de la respectiva enzima en ese organelo diana. (Véase también, Martinet *et al.* (1998) *Biotech. Letters* 20(12):1171-1177, que demuestra que α -1,2-manosidasa de *T. reesei*, aunque localizada intracelularmente, aumentaba en lugar de disminuir el grado de manosilación). Hasta la fecha, no hay informes que demuestren la localización intracelular de una α -1,2-manosidasa heteróloga ni en levaduras ni en hongos usando una secuencia de localización transmembrana.

Aunque es útil diseñar cepas que sean capaces de producir Man₅GlcNAc₂ como estructura de N-glucano primaria, cualquier intento de modificar adicionalmente estas estructuras precursoras altas en manosa para que se parezcan más a los glucanos humanos requiere etapas adicionales *in vivo* o *in vitro*. Los procedimientos para humanizar más los glucanos y levaduras fuentes *in vitro* se describen en la patente de Estados Unidos N.º 5.834.251 (referencia anterior). Si Man₅GlcNAc₂ debe humanizarse más *in vivo*, se debe asegurar que las estructuras de Man₅GlcNAc₂ generadas se generen, de hecho, intracelularmente y no sean producto de la actividad de las manosidasas del medio. La formación de N-glucanos complejos en levaduras u hongos requerirá que se generen niveles altos de Man₅GlcNAc₂ dentro de la célula porque sólo los glucanos Man₅GlcNAc₂ intracelulares pueden procesarse ulteriormente para producir N-glucanos híbridos y complejos *in vivo*. Además, se debe demostrar que la mayoría de las estructuras de Man₅GlcNAc₂ generadas de hecho son un sustrato para GnTI y así permiten formar N-glucanos híbridos y complejos.

Por consiguiente, existe la necesidad de procedimientos para producir glucoproteínas caracterizadas por un alto contenido en Man₅GlcNAc₂ intracelular que puede procesarse posteriormente para producir estructuras de glucoproteínas humanoides en células huésped eucariotas no humanas, y en particular en levaduras y hongos filamentosos.

N-Acetilglucosaminiltransferasas

Las N-acetilglucosaminiltransferasas ("GnT") pertenecen a otra clase de enzimas de glucosilación que modifican los oligosacáridos ligados en N en la ruta secretora. Dichas glucosiltransferasas catalizan la transferencia de un monosacárido de donantes de azúcares de nucleótidos específicos a una posición de hidroxilo particular de un monosacárido en una cadena de glucano que se desarrolla en uno de los dos posibles enlaces anoméricos (o α o β). Dennis *et al.* (1999) *Bioessays* 21 (5):412-21. Las GnT añaden N-acetilglucosamina ("GlcNAc₂") sobre la ramificación Man α 1,6 o la ramificación Man α 1,3 ramificación de un sustrato de N-glucano (por ejemplo, Man₅GlcNAc₂ ("núcleo de manosa-5") y Man₃GlcNAc₂ (una "estructura nuclear interna"). El producto de reacción (por ejemplo, GlcNAcMan₅GlcNAc₂ o GlcNAc₂Man₃GlcNAc₂) después puede ser modificado en estructuras de oligosacáridos bi-, tri-, tetra- y penta-antennarias ligados en N.

La N-acetilglucosaminiltransferasa en ("GnTIII") es una enzima que cataliza la adición de una GlcNAc, en la manosa del medio del núcleo de trimanosa (Man α 1,6 (Man α 1,3) Man β 1,4 - GlcNAc β 1,4 - GlcNAc β 1,4 - Asn) de un oligosacárido ligado en N. La adición por GnTIII de una GlcNAc diseccionadora a un sustrato receptor (por ejemplo un núcleo de trimanosa) proporciona lo que se denomina N-glucano diseccionado. Por ejemplo, la adición por GnTIII de una GlcNAc diseccionadora a la estructura de GlcNAcMan₃GlcNAc₂ puede proporcionar un N-glucano diseccionado, GlcNAc₂Man₃GlcNAc₂. De forma similar, la adición por GnTIII de una GlcNAc diseccionadora a una estructura de GlcNAc₂Man₃GlcNAc₂ proporciona otro un N-glucano diseccionado, GlcNAc₂Man₃GlcNAc₂. Esta última estructura ha sido implicada en una mayor citotoxicidad celular dependiente de anticuerpos (ADCC).

Umana *et al.* (1999) *Nat Biotechnol.* 17(2):176-80. Pueden formarse otros N-glucanos biseccionados por la acción de GnTIII. Por ejemplo, GlcNAcMan₄GlcNAc₂ puede convertirse en GlcNAc₂Man₄GlcNAc₂ diseccionado, Man₅GlcNAc₂ puede convertirse en GlcNAcMan₅GlcNAc₂ diseccionado y GlcNAcMan₅GlcNAc₂ puede convertirse en GlcNAc₂Man₅GlcNAc₂ diseccionado. Véase, por ejemplo, Narasimhan (1982) *J. Biol. Chem.* 257:10235-42. Hasta la fecha, la actividad de GnTIII sólo se ha demostrado en las células de mamífero.

El rediseño de glucoformas de inmunoglobulinas expresadas por las células de mamífero es una tarea tediosa e incómoda. Especialmente en el caso de GnTIII, donde la hiperexpresión de esta enzima se ha implicado en la inhibición del crecimiento, y han tenido que emplearse procedimientos que suponen la expresión génica regulada

(inducible) para producir inmunoglobulinas con *N*-glucanos diseccionados. Umana *et al.* (1999) *Biotechnol Bioeng.* 65(5):542-9; Umana *et al.* (1999) *Nat. Biotechnol.* 17(2):176-80; Umana *et al.* documento WO 031011878; patente de Estados Unidos n.º 6.602.684. Dicho efecto de inhibición del crecimiento complica la capacidad de coexpresar la proteína diana y GnTIII y puede imponer un límite superior sobre la hiperexpresión de GnTIII. Patente de Estados Unidos n.º 6.602.684. Puede ser necesaria la optimización cuidadosa de los niveles de expresión de GnTIII. *Id.* Como se describe anteriormente, sin embargo, el desarrollo de células huésped de eucariotas inferiores que se usan en dicho sistema de producción de proteínas requiere modificar aún más las rutas de glucosilación endógenas de las células huésped.

Es sabido que las enzimas GnTIV, GnTV y GnTIX expresadas en células mamífero catalizan la transferencia de restos de GlcNAc en particular la conformación sobre sustratos de oligosacáridos que producen estructuras de glucanos multiantenarias. UDP-*N*-acetilglucosamina:α1,3-D-manosida/β1,4-*N*-acetilglucosaminil-transferasa (GnTIV; EC 2.4.1.145) cataliza la transferencia de GlcNAc a partir de UDP-GlcNAc en los restos en enlace β1,4 a α1,3-D-manosida de GlcNAcβ1-2Manα1-6 (GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn (Gleeson y Schachter, *J Biol Chem.* 25 de mayo de 1983; 258(10):6162-73; Schachter *et al.*, (1989) *Methods Enzymol.*, 179, 351-397). UDP-*N*-acetilglucosamina:α-6-D-manosida β1,6-*N*-acetilglucosaminiltransferasa (GnTVI; EC 4.1.155) cataliza la adición de una *N*-acetilglucosamina al núcleo de α1,6 manosilo en un enlace β1,6 que forma *N*-glucanos tri- y tetraantenarios.

De forma similar, la expresión de GnTVI en células aviarias cataliza la transferencia de los restos de GlcNAc sobre sustratos de oligosacárido. Específicamente, UDP *N*-acetil-D-glucosamina (GlcNAc):GlcNAcβ1-6(GlcNAcβ1-2)Manα1-R[GlcNAc a Man] β,4-*N*-acetilglucosaminiltransferasa VI(GnTVI) cataliza la formación de *N*-glucanos pentaantenarios (Sakamoto *et al.*, *J. Biol. Chem.* 17 de Nov de 2000; 275(46):36029-34). El gen que codifica GnTVI ha sido purificado y aislado recientemente. Taniguchi *et al.*, documento JP2002209587A2.

Los sustratos necesarios para producir estructuras multiantenarias complejas no se han sintetizado en huéspedes fúngicos hasta últimamente (Hamilton *et al.*, *Science.* 29 de agosto de 2003; 301(5637):1244-6). Las células de mamífero habitualmente producen un abanico de glucanos complejos como por ejemplo glucoformas biantenarias, triantenarias, tetraantenarias e incluso pentaantenarias mediante reacción secuencia de GnT específicas. En el aparato de Golgi de dichas células, el procesamiento con *N*-glucano de glucoproteínas produce estructuras biantenarias predominantemente, además de la formación de estructuras triantenarias y tetraantenarias. Actualmente se entiende que en la formación de glucanos complejos, las GnT específicas catalizan enlaces β-GlcNAc específicos (por ejemplo, β1,2; β1,4; β1,6), produciendo glucanos multiantenarios en células de mamífero. Estas células, sin embargo, son incapaces de producir ninguna glucoforma homogénea con un rendimiento alto.

Recientemente, se han diseñado eucariotas inferiores para producir glucanos complejos en formas homogéneas a niveles significativos (Hamilton *et al.*, *Science.* 29 de agosto de 2003; 301(5637):1244-6). La capacidad de producir glucanos complejos multiantenarios en los eucariotas inferiores proporcionaría grandes cantidades de proteínas plegadas de forma adecuada y glucosiladas a escala industrial a bajo coste, en menos tiempo, de forma más segura y con mayor calidad. Lo que se necesita, por lo tanto, es un sistema de producción de proteínas que utiliza la capacidad inherente de títulos de producto fuertes como por ejemplo los que se producen en las células huésped de eucariotas inferiores (por ejemplo levaduras y hongos filamentosos), que es capaz de producir glucanos multiantenarios (y opcionalmente, diseccionados) ligados en N sobre proteínas, especialmente proteínas terapéuticas, expresadas en estas células.

45 Sumario de la invención

Se han desarrollado células huésped y líneas celulares que tienen rutas de glucosilación genéticamente modificadas que les permite llevar a cabo una secuencia de reacciones enzimáticas, que imitan el procesamiento de las glucoproteínas mamíferos, especialmente en los seres humanos. Las proteínas recombinantes expresadas en estos huéspedes genomanipulados proporcionan glucoproteínas más similares, si no sustancialmente idénticas, a sus homólogas mamíferas, por ejemplo humanas. Las células huésped de la invención, por ejemplo, células huésped de hongos unicelulares y multicelulares que se desarrollan en cultivo, se modifican para producir *N*-glucanos, como por ejemplo *N*-glucanos diseccionados, u otras estructuras producidas en rutas de glucosilación humanas. Este resultado se logra usando una combinación de diseño y/o selección de cepas que, por ejemplo, no expresan enzimas que crean las indeseables estructuras características de las glucoproteínas fúngicas y que, por ejemplo, sí expresan las enzimas heterólogas capaces de producir una glucoproteína "humanizada".

La presente invención así proporciona un procedimiento de producción de glucoproteínas usando un huésped eucariota inferior como por ejemplo un hongo unicelular o filamentosos, que tiene un patrón de glucosilación diferente del de los seres humanos, para modificar la composición y las estructuras de diferentes de las proteínas que se forman en un organismo huésped ("células huésped") de forma que se parezcan más a las estructuras de carbohidrato que se encuentran en las proteínas de los mamíferos, por ejemplo, de los seres humanos. El procedimiento permite obtener unas células huésped genomanipuladas que pueden usarse para expresar y dirigirse contra cualquier gen(es) deseable(s), por ejemplo, uno implicado en la glucosilación, mediante procedimientos que están bien establecidos en la bibliografía científica y que son conocidos de forma general por el experto en el campo de la expresión de proteínas. Las células huésped con oligosacáridos modificados se crean o se seleccionan. Para la producción de proteínas terapéuticas, este procedimiento puede adaptarse para diseñar líneas celulares en las que pueda obtenerse cualquier estructura de glucosilación que se desee.

Por consiguiente, en una realización, la invención proporciona procedimientos para preparar una glucoproteína humanoide en una célula huésped de la invención introduciendo en la célula una actividad de *N*-acetilglucosaminiltransferasa III. En una realización preferente, la actividad de *N*-acetilglucosaminiltransferasa III se expresa en las células huésped de la invención, y en una realización todavía más preferente, esta expresión conlleva la producción de *N*-glucanos que comprenden estructuras $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ o $\text{GlcNAc}_2\text{Man}_5\text{GlcNAc}_2$ diseccionadas. En otra realización preferente, la actividad de *N*-acetilglucosaminiltransferasa III es sustancialmente intracelular. En otra realización preferente de la invención, la glucoproteína que incluye los *N*-glucanos con estructuras diseccionadas se aísla de las células huésped de eucariotas inferiores. En una realización todavía más preferente, la glucoproteína producida en las células huésped es una proteína terapéutica.

En otro aspecto, la invención proporciona una célula huésped de un eucariota inferior que incluye tanto una actividad de *N*-acetilglucosaminiltransferasa III como una actividad de *N*-acetilglucosaminiltransferasa II. En una realización preferente, las células huésped que incluyen la actividad de *N*-acetilglucosaminiltransferasa III producen *N*-glucanos que comprenden estructuras de $\text{GlcNAcMan}_3\text{GlcNAc}_2$ que pueden reaccionar con esta actividad. En una realización más preferente, la actividad produce un glucano diseccionado. Las células huésped de eucariotas inferiores de algunas realizaciones de la invención así pueden incluir un *N*-glucano con un glucano diseccionado. En una realización preferente, el *N*-glucano incluye más de 10% molar del glucano diseccionado. En algunas realizaciones, las células huésped incluyen un *N*-glucano que comprende estructuras diseccionadas de $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$, o $\text{GlcNAc}_2\text{Man}_5\text{GlcNAc}_2$. En una realización preferente, las células huésped incluyen una estructura nuclear de $\text{Man}_5\text{GlcNAc}_2$ o una estructura nuclear de $\text{Man}_3\text{GlcNAc}_2$ que se modifica mediante un GlcNAc biseccionador. En una realización todavía más preferente, la célula produce más de 10% molar de la estructura modificada.

En otra realización de la invención, las células huésped de eucariotas inferiores contienen una actividad de *N*-acetilglucosaminiltransferasa I además de la actividad de *N*-acetilglucosaminiltransferasa III. En una realización preferente, las actividades son sustancialmente intracelulares. En otra realización preferente, la célula produce *N*-glucanos que comprenden $\text{GlcNAcMan}_3\text{GlcNAc}_2$ que son capaces de reaccionar con la actividad de GnTIII. En una realización todavía más preferente, la actividad de GnTIII de la célula produce un glucano diseccionado.

En otra realización, las células huésped de eucariotas inferiores de la invención contienen tanto una actividad de *N*-acetilglucosaminiltransferasa III como actividad de manosidasa II. En una realización preferente, las células huésped además contienen una actividad de *N*-acetilglucosaminiltransferasa I. En otra realización preferente, las células huésped además contienen una actividad de *N*-acetilglucosaminiltransferasa II. En otra realización preferente, las células huésped además contienen tanto una actividad de *N*-acetilglucosaminiltransferasa I como una actividad de *N*-acetilglucosaminiltransferasa II.

La presente invención también proporciona procedimientos para preparar una glucoproteína humanoide en una célula huésped de un eucariota inferior introduciendo en la célula un ácido nucleico que codifica el dominio catalítico de GnTIVB humana (A 104-53) *N*-acetilglucosaminiltransferasa IV exógeno fusionado a los nucleótidos 1-108 del gen *MNN2* de *S. cerevisiae* que actúa como señal de direccionamiento celular no asociada normalmente al dominio catalítico. En una realización preferente, la actividad de *N*-acetilglucosaminiltransferasa VI de GnT IVB (Δ 104-53) se expresa en la célula, y en una realización todavía más preferente, esta expresión consigue la producción de *N*-glucanos que comprenden la estructura $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$. En otra realización preferente la actividad de *N*-acetilglucosaminiltransferasa VI de GnT IVB (Δ 104-53) es sustancialmente intracelular. En otra realización preferente de la invención, la glucoproteína que incluye los *N*-glucanos con estructuras triantenarias se aísla de las células huésped de eucariotas inferiores. En una realización más preferente, el *N*-glucano incluye más de 90% molar del glucano triantenario. En una realización todavía más preferente, la glucoproteína que se produce en las células huésped es una proteína terapéutica.

Así, en otro aspecto, la invención proporciona una célula huésped de un eucariota inferior que incluye tanto una actividad de *N*-acetilglucosaminiltransferasa IV como una actividad de *N*-acetilglucosaminiltransferasa V. En una realización preferente, las células huésped que incluyen dichas actividades de *N*-acetilglucosaminiltransferasa IV y V producen *N*-glucanos que comprenden estructuras de $\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$. En una realización más preferente, la actividad produce un glucano tetraantenario. Las células huésped de eucariotas inferiores de algunas realizaciones de la invención así pueden incluir un *N*-glucano con un glucano tetraantenario. En algunas realizaciones, las células huésped incluyen un *N*-glucano que comprende estructuras $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ y $\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$. En una realización preferente, las células huésped incluyen una estructura nuclear de $\text{GlcNAcMan}_5\text{GlcNAc}_2$ que es modificada por dicha GnT IV o una estructura nuclear de $\text{GlcNAc}_2\text{MafIaGlcNAc}_2$ que es modificada por dichas GnT IV y GnT V. En una realización preferente, el *N*-glucano incluye más de 70% molar del glucano tetraantenario. En una realización todavía más preferente, la célula produce más de 75% molar de glucanos tetraantenarios.

También se describe en el presente documento una célula huésped de un eucariota inferior que incluye una actividad de *N*-acetilglucosaminiltransferasa VI. Las células huésped descritas que expresan la actividad de *N*-acetilglucosaminiltransferasa VI producen *N*-glucanos que comprenden estructuras de $\text{GlcNAc}_5\text{Man}_3\text{GlcNAc}_2$ (por ejemplo, glucanos pentaantenarios). Las células huésped unicelulares o multicelulares que se describen en el presente documento así pueden incluir un *N*-glucano con un glucano pentaantenario. En algunos casos, las células huésped incluyen un *N*-glucano que comprende estructuras de $\text{GlcNAc}_5\text{Man}_3\text{GlcNAc}_2$. Como se describe en el presente documento, las células huésped incluyen una estructura nuclear de $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ que es modificada por una GnT VI.

También se describen en el presente documento procedimientos para preparar una glucoproteína humanoide en células huésped de hongos unicelulares o multicelulares introduciendo en la célula una actividad de *N*-acetilglucosaminiltransferasa IX. Como se describe en el presente documento, la actividad de *N*-acetilglucosaminiltransferasa IX se expresa en la célula, y esta expresión consigue la producción de *N*-glucanos que comprenden estructuras de GlcNAc₃Man₃GlcNAc₂ y GlcNAc₄Man₃GlcNAc₂. Como se describe en el presente documento, la actividad de *N*-acetilglucosaminiltransferasa IX es sustancialmente intracelular. En otra realización preferente de la invención, la glucoproteína que incluye los *N*-glucanos con estructuras multiantenarias se aísla de las células huésped de hongos unicelulares o multicelulares. En una realización todavía más preferente, la glucoproteína producida en las células huésped es una proteína terapéutica.

En otra realización de la invención, las células huésped de hongos unicelulares o multicelulares contienen una actividad de *N*-acetilglucosaminiltransferasa I y una actividad de *N*-acetilglucosaminiltransferasa II. En una realización preferente, la actividades son sustancialmente intracelulares. En otra realización preferente, la célula produce *N*-glucanos que comprenden GlcNAc₂Man₃GlcNAc₂ que son capaces de reaccionar con dicha actividad de GnTIV produciendo glucanos triantenarios. En una realización todavía más preferente, la actividad de GnTV de la célula produce un glucano tetraantenario.

En otra realización, las células huésped de eucariotas inferiores de la invención contienen tanto dicha actividad de *N*-acetilglucosaminiltransferasa IV como una actividad de manosidasa II. En una realización preferente, las células huésped además contienen una actividad de *N*-acetilglucosaminiltransferasa I. En otra realización preferente, las células huésped además contienen una actividad de *N*-acetilglucosaminiltransferasa II. En otra realización preferente, las células huésped además contienen una actividad de *N*-acetilglucosaminiltransferasa V.

En ciertas realizaciones preferentes, la célula huésped de la invención es deficiente en una actividad de *OCH1* manosiltransferasa. Dicha célula, por ejemplo, puede ser deficiente en una actividad de Dol-P-Man:Man₅GlcNAc₂-PP-Dol manosiltransferasa. Todavía en otra realización, las células huésped de la invención pueden comprender además una actividad de α -1,2-manosidasa I. En otra realización, las células huésped además pueden comprender un transportador de azúcar de nucleótido. Preferentemente, las células huésped comprenden un transportador de UDP-GlcNAc en el que la transferencia de restos de GlcNAc es facilitada por cualquiera de las actividades de *N*-acetilglucosaminiltransferasa mencionadas anteriormente.

La presente invención también proporciona glucoproteínas que se preparan mediante los procedimientos de la invención. En una realización, la glucoproteína incluye una GlcNAc diseccionadora en una estructura nuclear de Man₅GlcNAc₂ o de Man₃GlcNAc₂ y se produce en una célula huésped de un eucariota inferior. En otra realización, la glucoproteína incluye una GlcNAc diseccionadora unida a una estructura nuclear de Man₅GlcNAc₂, Man₄GlcNAc₂, Man₃GlcNAc₂, GlcNAcMan₃GlcNAc₂, GlcNAcMan₅GlcNAc₂, o GlcNAc₂Man₃GlcNAc₂ y se produce en las células huésped de hongos unicelulares o multicelulares. En una realización preferente, más de 10% molar de las estructuras nucleares de la glucoproteína de la invención son modificadas por la GlcNAc diseccionadora.

Todavía en otra realización, la invención proporciona una glucoproteína que incluye una estructura triantenaria como por ejemplo GlcNAc₂Man₃GlcNAc₂ y se produce en células huésped de hongos unicelulares o multicelulares. En una realización preferente, más de 90% molar de las estructuras nucleares de la glucoproteína de la invención son modificadas por la GnTIV de la invención. En otra realización, la glucoproteína incluye una estructura tetraantenaria como por ejemplo GlcNAc₄Man₃GlcNAc₂ y se produce en una célula huésped de un eucariota inferior. En una realización más preferente, más de 75% molar de las estructuras nucleares de la glucoproteína de la invención son modificadas por la GnTV de la invención.

En otro aspecto, la invención proporciona composiciones farmacéuticas que contienen las glucoproteínas humanoides producidas en una célula huésped de un eucariota inferior. También se describen en el presente documento vectores que codifican proteínas que tienen una o más actividades de *N*-acetilglucosaminiltransferasa III, IV, V, VI y IX y que contienen secuencias peptídicas directoras unidas. En una realización preferente, las proteínas codificadas por los vectores están localizadas en células huésped de hongos unicelulares o multicelulares de tal forma que producen *N*-glucanos que tienen estructuras diseccionadas y/o multiantenarias.

La invención también está definida por las reivindicaciones.

Breve descripción de los dibujos

La Figura 1A es un diagrama esquemático de una ruta de *N*-glucosilación fúngica típica. La Figura 1B es un diagrama esquemático de una ruta de *N*-glucosilación humana típica.

La Figura 2 representa la construcción de una adenoteca de combinaciones de construcciones fusionadas. La Figura 2A es un diagrama de la inserción de un fragmento peptídico director en pCR2.1- TOPO (Invitrogen, Carlsbad, CA). La Figura 2B muestra la subadenoteca de péptidos directores generado con los sitios de restricción *NotI*- *AscI*. La Figura 2C es un diagrama de la inserción de una región de dominio catalítico en pJN347, un vector pUC19 modificado. La Figura 2D muestra la subadenoteca de dominios catalíticos generados con los sitios de restricción *NotI*, *AscI* y *FacI*. La Figura 2E representa una construcción fusionada particular generada a partir de la subadenoteca de péptidos directores y la subadenoteca de dominios catalíticos.

La Figura 3 ilustra la secuencia de ácidos nucleicos con marco de lectura abierto (SEC. ID. N.º: 48) y la secuencia del polipeptido codificado (SEC. ID. N.º: 49) de α -1,2-manosidasa IA de *M. musculus*. Las secuencias de los cebadores peR que se usaron para generar truncados en el extremo N están subrayadas.

- 5 La Figura 4 ilustra el diseño de vectores con marcadores auxotrópicos múltiples e integración genética de proteínas diana en el locus *OCH1* de *P. pastoris*.

Las Figuras 5A - 5E muestran el análisis por MALDI-TOF que demuestra la producción del dominio kringle 3 de las glucoproteínas de plasminógeno humano (K3) que tienen Man₅GlcNAc₂ como estructura de *N*-glucano predominante en *P. pastoris*. La Figura 5A representa el patrón del glucano Man₅GlcNAc₂ [a] (Glyko, Novato, CA) y Man₅GlcNAc₂ + Na⁺ [b]. La Figura 5B muestra los glucanos liberados por la PNGasa a partir de K3 de tipo silvestre. Los *N*-glucanos que se muestran son los siguientes: Man₉GlcNAc₂ [d]; Man₁₀GlcNAc₂ [e]; Man₁₁GlcNAc₂ [f]; Man₁₂GlcNAc₂ [g]. La Figura 5C representa la delección de *och1* por la que se obtiene la producción de Man₃GlcNAc₂ [c] como *N*-glucano predominante. Las Figuras 5D y 5E muestran la producción de Man₅GlcNAc₂ [b] después del recorte *in vivo* de Man₅GlcNAc₂ con una α -1,2-manosidasa quimérica. El *N*-glucano predominante se indica mediante un pico con una masa (*m/z*) de 1253 coherente con su identificación como Man₅GlcNAc₂ [b].

Las Figuras 6A - 6F muestran el análisis por MALDI-TOF que demuestra la producción de glucoproteínas de IFN- β que tienen Man₅GlcNAc₂ como estructura de *N*-glucano predominante en *P. pastoris*. La Figura 6A muestra el patrón de Man₅GlcNAc₂ [a] y Man₅GlcNAc₂ + Na⁺ [b] como patrón (Glyko, Novato, CA). La Figura 6B muestra los glucanos liberados por PNGasa a partir de IFN- β de tipo silvestre. La Figura 6C representa el *och1* inactivado que produce Man₃GlcNAc₂ [c]; Man₉GlcNAc₂ [d]; Man₁₀GlcNAc₂ [e]; Man₁₁GlcNAc₂ [f]; Man₁₂GlcNAc₂ [g]; y la no producción de Man₅GlcNAc₂ [b].

25 La Figura 6D muestra una cantidad relativamente pequeña de Man₅GlcNAc₂ [b] entre otros *N*-glucanos Man₃GlcNAc₂ intermedios [c] a Man₁₂GlcNAc₂ [g]. La Figura 6E muestra una cantidad significativa de Man₅GlcNAc₂ [b] con respecto a los otros glucanos Man₆GlcNAc₂ [c] y Man₉GlcNAc₂ [d] producidos por pGC5 (*MNS1(m)* de *Saccharomyces/manosidasa* de ratón IB Δ 99). La Figura 6F muestra la producción predominante de Man₅GlcNAc₂ [b] sobre la glucoproteína segregada FN- β por pFB8 (*SEC12(m)* de *Saccharomyces/manosidasa* IA de ratón Δ 187). El *N*-glucano predominante se indica mediante un pico con una masa (*m/z*) de 1254 coherente con su identificación como Man₅GlcNAc₂ [b].

35 La Figura 7 muestra un cromatograma de líquidos de alta resolución para: (A) patrón Man₉GlcNAc₂ marcado con 2-AB (control negativo); (B) sobrenadante de medio de *P. pastoris*, Δ *loch1* transformado con pFBB manosidasa, que demuestra una falta de actividad extracelular de manosidasa en el sobrenadante; y (C) patrón de Man₉GlcNAc₂ marcado con 2AB después de exposición a manosidasa de *T. reesei* (control positivo).

40 La Figura 8 muestra un cromatograma de líquidos de alta resolución para: (A) patrón de Man₉GlcNAc₂ marcado con 2-AB (control negativo); (B) sobrenadante de medio de *P. pastoris*, Δ *loch1* transformado con pGC5 manosidasa, que demuestra una carencia de actividad extracelular de manosidasa en el sobrenadante; y (C) patrón de Man₉GlcNAc₂ marcado con 2AB después de exposición a manosidasa de *T. reesei* (control positivo).

45 La Figura 9 muestra un cromatograma de líquidos de alta resolución para: (A) patrón de Man₉GlcNAc₂ marcado con 2-AB (control negativo); (B) sobrenadante de medio de *P. pastoris*, Δ *loch1* transformado con pBC18-5 manosidasa, que demuestra una carencia de actividad extracelular de manosidasa en el sobrenadante; y (C) sobrenadante de medio de *P. pastoris*, Δ *loch1* transformado con pDD28-3, que demuestra actividad en el sobrenadante (control positivo).

50 Las Figuras 10A - 10B demuestran la actividad de un transportador de UDP-GlcNAc en la producción de GlcNAcMan₅GlcNAc₂ en *P. pastoris*. La Figura 10A representa una cepa de *P. pastoris* (YSH-3) con una GnTI humana pero sin el transportador de UDP-GlcNAc que consiguió alguna producción de GlcNAcMan₅GlcNAc₂ [b] pero una producción predominante de Man₅GlcNAc₂ [a]. La Figura 10B representa la adición del transportador de UDP-GlcNAc de *K. lactis* en una cepa (PBP-3) con la GnTI humana, que consiguió la producción predominante de GlcNAcMan₅GlcNAc₂ [b]. El único pico prominente de masa (*m/z*) a 1457 es coherente con su identificación como GlcNAcMan₅GlcNAc₂ [b] como se muestra en la Figura 10B.

La Figura 11 muestra un óptimo de pH de una enzima manosidasa heteróloga codificada por pBB27 -2 (manosidasa IB Δ 31 de *Saccharomyces MNN10* (s)/*C. elegans*) expresada en *P. pastoris*.

60 Las Figuras 12A - 12C muestran el análisis por MALDI-TOF de *N*-glucanos liberados de un extracto acelular de *K. lactis*. La Figura 12A muestra los *N*-glucanos liberados por células de tipo silvestre, que incluyen *N*-glucanos de tipo altos en manosa. La Figura 12B muestra los *N*-glucanos liberados de células sin *och1 mnn1*, que revela un pico destacado de masa (*m/z*) a 1908 coherente con su identificación como Man₉GlcNAc₂ [d]. La Figura 12C muestra los *N*-glucanos liberados de células sin *och1 mnn1* después de la digestión de α -1,2-manosidasa *in vitro* que corresponde a un pico coherente con Man₅GlcNAc₂.

La Figura 13 es un esquema de la estructura del oligosacárido ligado con doliquil pirofosfato.

ES 2 336 790 T3

La Figura 14 es un esquema de la generación de *N*-glucanos de GlcNAc₂Man₃GlcNAc₂ de células huésped fúngicas que son deficientes en las actividades de *Alg3*, *alg9*, o *alg12*.

La Figura 15 es un esquema de reacciones de procesamiento necesarias para producir estructuras de oligosacárido de tipo mamífero en una célula huésped fúngica con un genotipo *Alg3*, *och1*.

La Figura 16 muestra las comparaciones entre las secuencias de *Alg3* de *S. cerevisiae* (Blast) (SEC. ID. N.º: 9 - 20, respectivamente, por orden de aparición)

La Figura 17 muestra las secuencias de *ALG3* (SEC. ID. N.º: 21) y *Alg3p* (SEC. ID. N.º:22) de *S. cerevisiae*

La Figura 18 muestra las secuencias *ALG3* (SEC. ID. N.º: 23) y *Alg3p* (SEC. ID. N.º: 24) de *P. pastoris*.

La Figura 19 muestra las comparaciones de las secuencias de *ALG3* (Blast) (SEC. ID. N.º: 23 - 31, respectivamente, por orden de aparición) de *P. pastoris*.

La Figura 20 muestra las secuencias de *ALG3* (SEC. ID. N.º:33) y *Alg3p* (SEC. ID. N.º: 34) de *K. lactis*.

La Figura 21 muestra las comparaciones de las secuencias de *ALG3* (Blast) (SEC. ID. N.º: 35 - 40, respectivamente, por orden de aparición) de *K. lactis*.

La Figura 22 muestra un modelo de una inmunoglobulina IgG. La cadena pesada y la cadena ligera pueden subdividirse en dominios, basándose en una estructura secundaria y terciaria similar. Las dos cadenas pesadas (dominios V_H, C_H 1, C_H2 y C_H3) están ligadas a través de tres puentes disulfuro. Las cadenas ligeras (dominios V_L y C_L) están ligadas a otro puente disulfuro en la porción C_H1 de la cadena pesada y, junto con los fragmentos C_H1 y V_H, conforman la región Fab. Los antígenos se unen a la porción terminal de la región Fab. Las funciones efectoras, como por ejemplo la unión al receptor Fc-gamma se han localizado en el dominio C_H2, justo aguas abajo de la región de bisagra y se ven influenciadas por la *N*-glucosilación de asparragina 297 en la cadena pesada.

La Figura 23 es un resumen esquematizado de un vector de expresión 19G1 modular.

La Figura 24 muestra las secuencias de ácido nucleico (SEC. ID. N.º: 45) y de aminoácidos (SEC. ID. N.º: 46) de GnTIII de *M. musculus*.

La Figura 25 (arriba) es un análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 producidos en una *P. pastoris* YSH-1 que presenta picos predominantes en 1461 m/z que corresponden a la masa de GlcNAcMan₅GlcNAc₂ [d]; La Figura 25 (abajo) muestra un análisis por MALDI-TOF-MS de *N*-glucanos aislados a partir de una glucoproteína kringle 3 producida en una *P. pastoris* YSH-1 transformada con manosidasa II Δ74 de *D. melanogaster* /MNN2(s) de *S. cerevisiae* que demuestran un pico predominante en 1140 m/z que corresponde a la masa de GlcNAcMan₃GlcNAc₂ [b] y otros picos que corresponden a GlcNAcMan₄GlcNAc₂ [e] en 1303 m/z y GlcNAcMan₅GlcNAc₂ [d] en 1465 m/z. Esta cepa se denominó YSH-37.

La Figura 26 (arriba) es el análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 producidos en *P. pastoris* YSH-1 como se muestra en la Figura 25 (arriba); La Figura 26 (abajo) es el análisis por MALDI-TOF-MS de *N*-glucanos aislados a partir de una glucoproteína kringle 3 expresada en células de *P. pastoris* YSH-1 transformadas con una construcción pVA53 (MNN2(s)/mGnTIII de *S. cerevisiae*). The pico en 1463 m/z corresponde a la masa de GlcNAcMan₅GlcNAc₂ [d] y el pico en 1666 m/z corresponde a la masa de GlcNAc₂Man₅GlcNAc₂ [a].

La Figura 27 (arriba) es el análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 producidos en *P. pastoris* YSH-1 como se muestra en la Figura 25 (arriba); La Figura 27 (abajo) es el análisis por MALDI-TOF-MS de *N*-glucanos aislados a partir de una glucoproteína kringle 3 expresada en células de *P. pastoris* YSH-1 transformadas con una construcción pVB51 (GNT1(s)/mGnTIII de *K. lactis*). El pico predominante en 1463 m/z corresponde a la masa de GlcNAcMan₅GlcNAc₂ [d] y el pico en 1667 m/z corresponde a la masa de GlcNAc₂Man₅GlcNAc₂ [a].

La Figura 28 (arriba) es el análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 producidos en *P. pastoris* YSH-1 como se muestra en la Figura 25 (arriba); La Figura 28 (abajo) es el análisis por MALDI-TOF-MS de *N*-glucanos aislados a partir de una glucoproteína kringle 3 expresada en células de *P. pastoris* YSH-1 transformadas con una construcción pVB51 (GNT1(s)/mGnTIII de *K. lactis*). El pico predominante en 1463 m/z corresponde a la masa de GlcNAcMan₅GlcNAc₂ [d] y se observa un segundo pico en 1726 m/z [e], que no corresponde a la masa de GlcNAc₂Man₅GlcNAc₂.

La Figura 29 es el análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* YSH-44. El pico predominante en 1356 m/z corresponde a la masa de GlcNAc₂Man1:3GlcNAc₂ [x].

ES 2 336 790 T3

La Figura 30 es un análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* YSH-44 transformadas con una construcción pVA53 (*S. cerevisiae*). El pico en 1340 m/z corresponde a la masa de GlcNAc₂Man₃GlcNAc₂ [x] y el pico en 1542 m/z corresponde a la masa de GlcNAc₂Man₃GlcNAc₂ [y].

La Figura 31 es el análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* PBP6-5. El pico predominante en 1340 m/z corresponde a la masa de GlcNAc₂Man₃GlcNAc₂ [x].

La Figura 32 es el análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* PBP6-5 transformadas con una construcción pVA53 (*MNN2(s)/mGnTIII* de *S. cerevisiae*). El pico en 1340 m/z corresponde a la masa de GlcNAc₂Man₃GlcNAc₂ [x] y el pico en 1543 m/z corresponde a la masa de GlcNAc₂Man₃GlcNAc₂ [y].

La Figura 33 muestra un cromatograma de líquidos de alta resolución, que demuestra la falta de actividad extracelular de GnTIII (pVA53) en el sobrenadante. El *N*-glucano GlcNAcMan₅GlcNAc₂ purificado a partir de K3 expresada en la cepa PBP-3 se añadió a: BMMY (A); UDP-GlcNAc 1 mM (Sigma Chemical Co., St. Louis, MO) en BMMY (B); el sobrenadante de YSH-44 transformado con pVA53 [YSH-57] (C); y el sobrenadante de YSH-57 + UDP-GlcNAc 1 mM (D).

La Figura 34 muestra un cromatograma de líquidos de alta resolución, que demuestra la falta de actividad extracelular de GnTIII (pVA53) en el sobrenadante. El *N*-glucano GlcNAc₂Man₃GlcNAc₂ purificado a partir de K3 expresada en la cepa YSH-44 se añadió a: BMMY (A); UDP-GlcNAc 1 mM (Sigma Chemical Co., St. Louis, MO) En BMMY (B); y el sobrenadante de YSH-44 transformado con pVA53 [YSH-57] (C).

La Figura 35 es un diagrama esquemático que compara las rutas de glucosilación normales en los seres humanos y en *P. pastoris* (Panel A) con una ruta de *N*-glucosilación humanizada genomanipulada en eucariotas inferiores (Panel B). La ruta genomanipulada representa la construcción de la cepa PBP6-5 de *P. pastoris*, que después de la modificación con GnTIII se convierte en la cepa PBP38 de *P. pastoris*.

La Figura 36 es un diagrama esquemático que muestra la glucoforma predominante segregada producida por cada una de las cepas designadas de *P. pastoris* y la modificación génica que se usa para diseñar cada una de las cepas.

La Figura 37 es una representación estructural de la transferencia de una GlcNAc al intermedio de oligosacárido, GlcNAcMan₅GlcNAc₂, producido en glucoproteínas en una célula huésped de un eucariota inferior, catalizada por GnTIII.

La Figura 38 es una representación estructural de la transferencia de una GlcNAc al intermedio de oligosacárido, GlcNAcMan₃GlcNAc₂, producido en glucoproteínas en una célula huésped de un eucariota inferior, catalizada por GnTII, y la posterior transferencia de una GlcNAc al producto de esa reacción, GlcNAc₂Man₃GlcNAc₂, catalizada por GnTIII.

La Figura 39 es un diagrama esquemático que muestra la transferencia de restos de GlcNAc sobre intermedios de oligosacáridos catalizada por GnTIV, GnTV, GnTVI y GnTIX en *P. pastoris*.

La Figura 40 muestra tres mapas de plásmidos representativos. La Figura 40A: pPB144 que contiene un fragmento de un gen que codifica GnTIV de ratón; La Figura 40B: pPB140 que contiene un fragmento de un gen que codifica GnTV humana; y La Figura 40C: pPB176 que contiene un fragmento de un gen que codifica GnTIX de ratón que se usaron para la transformación en un huésped *P. pastoris*.

La Figura 41 muestra el gen que codifica la manosil (alfa-1,3-)-glucoproteína beta1,4-*N*-acetilglucosaminiltransferasa, isoenzima A (MGAT4A) de *Homo sapiens* Número de acceso NM_012214.

La Figura 42 muestra el gen que codifica la manosil (alfa-1,3-)-glucoproteína beta1,4-*N*-acetilglucosaminiltransferasa, isoenzima A (MGAT4B) de *Homo sapiens* Número de acceso NM_014275.

La Figura 43 muestra el gen que codifica la *N*-acetilglucosaminiltransferasa V (Mgat5) de *Mus musculus* con Número de acceso AF474154.

La Figura 44 muestra el gen que codifica la *N*-acetilglucosaminiltransferasa VI de *Gallus gallus* con Número de acceso AB040608.

La Figura 45 muestra el gen que codifica la *N*-acetilglucosaminiltransferasa IX de *Homo sapiens* con Número de acceso AB109185.1.

La Figura 46 muestra el fragmento de ADN con codones optimizados que codifica parte de la *N*-acetilglucosaminiltransferasa IX humana que carece del dominio TM (Δ 43).

La Figura 47 es un análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* PBP43. La cepa de la levadura *P. pastoris* YSH-44 se transformó con pPB 144 que contenía la construcción de la fusión de *MNN2*(s) de *S. cerevisiae*/GnTIV humana. El pico en 1543 m/z corresponde a la masa de GlcNAc₃Man₃GlcNAc₂ [V].

La Figura 48 es un análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* PBP32. La cepa de la levadura *P. pastoris* YSH-44 se transformó con pPB 144 que contenía la construcción de la fusión *MNN2*(s) de *S. cerevisiae*/GnTV de ratón. El pico en 1559 m/z corresponde a la masa de GlcNAc₂Man₃GlcNAc₂ [V].

La Figura 49 es un análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* PBP46. La cepa de la levadura *P. pastoris* YSH-44 se transformó con pPB140 y pPB144. El pico en 1543 m/z corresponde a la masa de GlcNAc₂Man₃GlcNAc₂ [V] y el pico en 1747 m/z corresponde a la masa de GlcNAc₄Man₃GlcNAc₂ [z].

La Figura 50 es un análisis por MALDI-TOF-MS de *N*-glucanos aislados de una glucoproteína Kringle 3 expresados en células de *P. pastoris* PBP94. La cepa de la levadura *P. pastoris* YSH-44 se transformó con pPB128 que contenía la construcción fusionada *MNN2* (s) de *S. cerevisiae*/GnTIVA de ratón y pPB140 que contenía la construcción fusionada *MNN2* (s) de *S. cerevisiae*/GnTV de ratón. El pico predominante en 1743 m/z corresponde a la masa de GlcNAc₄Man₃GlcNAc₂ [z].

Descripción detallada de la invención

A no ser que se defina de otro modo en el presente documento, los términos científicos y técnicos que se usan referidos a la presente invención tendrán los significados que entienden habitualmente las personas de experiencia ordinaria en la técnica. Además, a no ser que el contexto requiera lo contrario, los términos en singular incluirán el plural y los términos en plural incluirán el singular. Los procedimientos y técnicas de la presente invención de forma general se realizan de acuerdo con procedimientos convencionales notorios en la técnica de forma general, las nomenclaturas que se usan con respecto a y las técnicas de bioquímica, enzimología, biología molecular y celular, microbiología, genética y química de proteínas y ácidos nucleicos e hibridación que se describen en el presente documento son las notorias y que se usan habitualmente en la técnica.

Los procedimientos y técnicas de la presente invención de forma general se realizan de acuerdo con procedimientos convencionales notorios en la técnica y como se describe en diversas referencias generales y más específicas que se citan y describen por toda la presente memoria descriptiva a no ser que se indique lo contrario. Véase, por ejemplo, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992, y suplementos hasta 2002); Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1990); Introduction to Glycobiology, Maureen E. Taylor, Kurt Drickamer, Oxford Univ. Press (2003); Worthington Enzyme Manual, Worthington Biochemical Corp. Freehold, NJ; Handbook of Biochemistry: Section A Proteins, Vol I 1976 CRC Press; Handbook of Biochemistry: Section A Proteins, Vol II 1976 CRC Press; Essentials of Glycobiology, Cold Spring Harbor Laboratory Press (1999). Las nomenclaturas que se usan con respecto a y los procedimientos y técnicas de laboratorio de, biología molecular y celular, bioquímica de proteínas, enzimología y química medicinal y farmacéutica que se describen en el presente documento son los notorios y habitualmente empleados en la técnica.

Los siguientes términos, a no ser que se indique lo contrario, se entenderá que tienen los siguientes significados:

Tal como se usa en el presente documento, el término “*N*-glucano” se refiere a un oligosacárido ligado en N, por ejemplo, uno que está unido mediante un enlace asparagina-*N*-acetilglucosamina a un resto asparagina de un polipéptido. Los *N*-glucanos tienen un núcleo pentasacárido en común de Man₃GlcNAc₂ (“Man” se refiere a manosa; “Glc” se refiere a glucosa; y “NAc” se refiere a *N*-acetil; GlcNAc se refiere a *N*-acetilglucosamina). El término “núcleo de trimanosa” empleado con respecto al *N*-glucano también se refiere a la estructura Man₃GlcNAc₂ (“Man₃”). El término “núcleo pentamanosa” o “núcleo de manosa-5” o “Man₅” empleado con respecto al *N*-glucano se refiere a la estructura Man₅GlcNAc₂. Los *N*-glucanos difieren con respecto al número de ramas (antenas) que comprenden azúcares periféricas (por ejemplo, GlcNAc, fucosa, y ácido siálico) que están unidas a la estructura nuclear de Man₃. Los *N*-glucanos se clasifican de acuerdo con sus constituyentes ramificados (por ejemplo, alto en manosa, complejo o híbrido).

Un *N*-glucano de tipo “alto en manosa” tiene cinco o más restos de manosa. Un *N*-glucano de tipo “complejo” habitualmente tiene al menos una GlcNAc unida a la ramificación 1,3 manosa y al menos una GlcNAc unida a la ramificación 1,6-manosa del núcleo de trimanosa. Los *N*-glucanos complejos también pueden tener restos galactosa (“Gal”) que están opcionalmente modificados con ácido siálico o sus derivados (“NeuAc”, donde “Neu” se refiere a ácido neuramínico y “Ac” se refiere a acetilo). Un *N*-glucano complejo habitualmente tiene al menos una ramificación que termina en un oligosacárido como por ejemplo, NeuNAc-; NeuAca2-6GalNAca1-; NeuAca2-3Galβ1-3GalNAca1-; NeuAca2-3/6Galβ1-4GlcNAcβ1-; GlcNAca1-4Galβ1-(sólo las mucinas); Fucα1-2Galβ1-(grupo de sangre H). Los ésteres de sulfatos pueden aparecer sobre restos de galactosa, GalNAc, y GlcNAc, y los ésteres de fosfatos pueden aparecer sobre restos de manosa. NeuAc (Neu: ácido neuramínico; Ac:acetilo) puede estar acetilado

en O o ser reemplazado por NeuGI (ácido *N*-glicolilneuramínico). Los *N*-glucanos complejos también pueden tener sustituciones intracatenarias que comprenden GlcNAc “diseccionadora” y la fucosa nuclear (“Fuc”). Un *N*-glucano “híbrido” tiene al menos un GlcNAc en el extremo de la ramificación de 1,3 manosa del núcleo de trimanosa y cero o más manosas en la ramificación de 1,6 manosa del núcleo de trimanosa.

5

El término “predominante” o “predominantemente” usado con respecto a la producción de *N*-glucanos se refiere a una estructura que representa el pico principal detectado mediante análisis por espectrometría de masas con tiempo de vuelo por ionización con desorción asistida por matriz (MALDI-TOF).

10

Las abreviaturas que se usan en el presente documento son de uso habitual en la técnica, véase, por ejemplo, las abreviaturas de los azúcares, más arriba. Otras abreviaturas habituales incluyen “PNGasa”, que se refiere a *N*-glucosidasa peptídica F (EC 3.2.2.18); “GlcNAc Tr” o “GnT,” que se refiere a las enzimas *N*-acetilglucosaminil transferasa; “NANA” se refiere a ácido *N*-acetilneuramínico.

15

Tal como se usa en el presente documento, una “glucoproteína humanizada” o “glucoproteína humanoide” se refiere alternativamente a una proteína que tiene unida a ella *N*-glucanos que incluyen menos de cuatro restos de manosa, y las glucoproteínas intermedias sintéticas (que también son de utilidad y pueden manipularse adicionalmente *in vitro* o *in vivo*) que tiene al menos cinco restos de manosa. Preferentemente, las glucoproteínas producidas de acuerdo con la invención contienen al menos 30% molar, preferentemente al menos 40% molar y más preferentemente 50, 60, 70, 80, 90, o incluso 100% molar del intermedio Man₅GlcNAc₂, al menos de forma transitoria. Esto puede lograrse, por ejemplo, diseñando una célula huésped de la invención para que exprese una enzima de glucosilación “mejor”, *es decir*, más eficaz. Por ejemplo, se selecciona una manosidasa de tal forma que tenga una actividad óptima en las condiciones presentes en el sitio de las células huésped donde las proteínas son glucosiladas y se introduce en las células huésped preferentemente dirigiendo la enzima a un organelo de una célula huésped donde se desea la actividad.

25

El término “enzima”, cuando se usa en el presente documento con respecto a la alteración de la glucosilación de células huésped, se refiere a una molécula que tiene al menos una actividad enzimática, e incluye enzimas de longitud completa, fragmentos catalíticamente activos, quimeras, complejos, y similares. Un “fragmento catalíticamente activo” de una enzima se refiere a un polipéptido que tiene un nivel de actividad funcional (enzimática) detectable. La actividad enzimática es “sustancialmente intracelular” cuando menos del 10% de la actividad enzimática puede medirse fuera de la célula comparado con la que puede medirse en células lisadas.

30

Una célula huésped de un eucariota inferior, cuando se usa en el presente documento con respecto a los perfiles de glucosilación, se refiere a cualquier célula eucariota que habitualmente produce *N*-glucanos con alto contenido en manosa, y así se pretende que incluya algunas células animales o vegetales y las células de eucariotas inferiores más habituales, que incluyen células de hongos y algas unicelulares y multicelulares.

35

Tal como se usa en el presente documento, el término “ruta de secreción” se refiere al ensamblaje de diversas enzimas de glucosilación a las que un precursor oligosacárido ligado a lípidos y un sustrato *N*-glucano se ven expuestos secuencialmente, siguiendo el flujo molecular de una cadena polipeptídica en formación desde el citoplasma al retículo endoplasmático (RE) y los compartimientos del aparato de Golgi. Se dice que las enzimas están localizadas en esta ruta. Una enzima X que actúa sobre un glucano ligado a enzimas o un *N*-glucano antes que la enzima Y se dice que está o que actúa “aguas arriba” de una enzima Y; de forma similar, la enzima Y está o actúa “aguas abajo” de la enzima X.

45

El término “péptido director” tal como se usa en el presente documento se refiere a secuencias de nucleótidos o aminoácidos que codifican un péptido de señal director celular que actúa de mediador en la localización (o retención) de una secuencia asociada en emplazamientos subcelulares, por ejemplo, organelos.

50

El término “polinucleótido” o “molécula de ácido nucleico” se refiere a una forma polimérica de nucleótidos de al menos 10 bases de longitud. El término incluye moléculas de ADN (por ejemplo, ADNc o ADN genómico o sintético) y a moléculas de ARN (por ejemplo, ARNm o ARN sintético), así como análogos de ADN o ARN que contienen análogos de nucleótidos no naturales, enlaces internucleosídicos no nativos, o ambos. El ácido nucleico puede estar en cualquier conformación topológica. Por ejemplo, el ácido nucleico puede ser monocatenario, bicatenario, tricatenario, cuádruplexado, parcialmente bicatenario, ramificado, con formación de horquilla, circular o con una formación de candado. El término incluye formas de ADN monocatenarias y bicatenarias. Una molécula de ácido nucleico de esta invención puede incluir hebras codificantes y no codificantes de ARN, ADNc, ADN genómico, y formas sintéticas y polímeros mixtos de los anteriores. Pueden estar modificadas químicamente o bioquímicamente o pueden contener bases de nucleótidos no naturales o derivadas, como fácilmente apreciarán los expertos en la técnica. Dichas modificaciones incluyen, por ejemplo, marcadores, metilación, sustitución de uno o más de los nucleótidos naturales con un análogo, modificaciones internucleotídicas como por ejemplo enlaces sin carga (por ejemplo, metil fosfonatos, fosfotriésteres, fosforamidatos, carbamatos, etc.), enlaces cargados (por ejemplo, fosforotioatos, fosforoditioatos, etc.), restos pendientes (por ejemplo, polipéptidos), intercaladores (por ejemplo, acridina, psoraleno, etc.), quelantes, alquiladores y enlaces modificados (por ejemplo, ácidos nucleicos alfa anoméricos, etc.) También se incluyen las moléculas sintéticas que imitan a los polinucleótidos en su capacidad de unirse a una secuencia designada mediante enlaces de hidrógeno y otras interacciones químicas. Dichas moléculas son conocidas en la técnica e incluyen, por ejemplo, aquellas en las que los enlaces peptídicos sustituyen a enlaces fosfato en el esqueleto de la molécula.

65

A no ser que se indique lo contrario, un “ácido nucleico que comprende la SEC. ID. N.º: X” se refiere a un ácido nucleico, al menos una porción de cual tiene o bien (i) la secuencia de la SEC. ID. n.º: X, o (ii) una secuencia complementaria a la SEC. ID. N.º: X. La elección entre las dos viene dictada por el contexto. Por ejemplo, si el ácido nucleico se usa como sonda, la elección entre las dos viene dictada por el requisito de que la sonda sea complementaria a la diana deseada.

Un ácido nucleico o polinucleótido “aislado” o “sustancialmente puro” (por ejemplo, un ARN, ADN o un polímero mixto) es uno que está sustancialmente separado de otros componentes celulares que naturalmente acompañan al polinucleótido nativo en su célula huésped natural, por ejemplo, ribosomas, polimerasas, y secuencias genómicas con las que está asociado de forma natural. El término engloba un ácido nucleico o polinucleótido que (1) ha sido retirado de su entorno natural, (2) no está asociado a todo ni a una porción de un polinucleótido en el que se encuentra el “polinucleótido aislado” en la naturaleza, (3) está ligado operablemente a un polinucleótido al que no está ligado en la naturaleza, o (4) no se encuentra en la naturaleza. El término “aislado” o “sustancialmente puro” también puede usarse en referencia a moléculas de ADN aisladas recombinantes o clonadas, análogos de polinucleótidos sintetizados químicamente o análogos de polinucleótidos que se sintetizan biológicamente mediante sistemas heterólogos.

Sin embargo, “aislado” no requiere necesariamente que el ácido nucleico o polinucleótido así descrito haya sido retirado físicamente de su entorno nativo. Por ejemplo, una secuencia de ácidos nucleicos endógena del genoma de un organismo se considera “aislada” en el presente documento si una secuencia heteróloga (*es decir*, una secuencia que no está adyacente de forma natural a esta secuencia de ácidos nucleicos endógena) está situada adyacente a la secuencia de ácidos nucleicos endógena, de tal forma que la expresión de este secuencia de ácidos nucleicos endógena se vea alterada. A modo de ejemplo, una secuencia promotora no nativa puede sustituirse (por ejemplo, mediante recombinación homóloga) por el promotor nativo de un gen del genoma de una célula humana, de tal forma que este gen tenga un patrón de expresión alterado. Este gen no estaría “aislado” porque está separado de al menos algunas de las secuencias que la flanquean de forma natural.

Un ácido nucleico también se considera “aislado” si contiene cualquier modificación que no se produzca naturalmente en el ácido nucleico correspondiente en un genoma. Por ejemplo, una secuencia codificante endógena se considera “aislada” si contiene una inserción, delección o una mutación puntual introducida artificialmente, por ejemplo, mediante intervención humana. Un “ácido nucleico aislado” también incluye un ácido nucleico integrado en el cromosoma de una célula huésped en un sitio heterólogo, una construcción de ácido nucleico presente como episoma. Además, un “ácido nucleico aislado” puede estar sustancialmente libre de otro material celular, o sustancialmente libre de medio de cultivo cuando se produce mediante técnicas recombinantes, o sustancialmente libre de precursores químicos u otras sustancias químicas cuando se sintetiza químicamente.

Tal como se usa en el presente documento, la frase “variante degenerativa” de una secuencia de ácidos nucleicos de referencia engloba secuencias de ácido nucleico que pueden traducirse, de acuerdo con el código genético estándar, proporcionando una secuencia de aminoácidos idéntica a la traducida a partir de la secuencia de ácidos nucleicos de referencia.

El término “identidad porcentual entre secuencias” o “idéntico” en el contexto de las secuencias de ácido nucleico se refiere a los restos de dos secuencias que son iguales cuando se alinean para conseguir una correspondencia máxima. La longitud de la comparación de la identidad de las secuencias puede ser en un tramo de al menos aproximadamente nueve nucleótidos, habitualmente al menos aproximadamente 20 nucleótidos, más habitualmente al menos aproximadamente 24 nucleótidos, habitualmente al menos aproximadamente 28 nucleótidos, más habitualmente al menos aproximadamente 32 nucleótidos, y preferentemente al menos aproximadamente 36 o más nucleótidos. Existe un número de algoritmos diferentes conocidos en la técnica que puede usarse para medir la identidad entre las secuencias de nucleótidos. Por ejemplo, pueden compararse las secuencias de polinucleótidos usando FASTA, Gap o Bestfit, que son programas del Wisconsin Package Versión 10.0, Genetics Computer Group (GCG), Madison, Wisconsin. FASTA proporciona alineaciones e identidad porcentual entre secuencias de las regiones con mejor coincidencia entre las secuencias de interrogación y de búsqueda (Pearson (1990) *Methods Enzymol.* 183:63-98).

Por ejemplo, puede determinarse la identidad porcentual entre secuencias entre las secuencias de ácido nucleico usando FASTA con sus parámetros por defecto (un tamaño de palabra de 6 y el factor NOPAM para la matriz de puntuación) o usando Gap con sus parámetros por defecto que vienen en GCG Versión 6.1.

El término “homología sustancial” o “similitud sustancial”, cuando se refiere a un ácido nucleico o fragmento del mismo, indica que, cuando está alineado de forma óptima con las inserciones o delecciones de nucleótidos apropiadas con otro ácido nucleico (o su hebra complementaria), existe una identidad entre las secuencias de nucleótidos en al menos aproximadamente 50%, más preferentemente 60% de las bases de nucleótido, habitualmente al menos aproximadamente 70%, más habitualmente al menos aproximadamente 80%, preferentemente al menos aproximadamente 90%, y más preferentemente al menos aproximadamente 95%, 96%, 97%, 98% o 99% de las bases de nucleótidos, medido por cualquier algoritmo conocido de identidad entre las secuencias, como por ejemplo FASTA, BLAST o Gap, como se ha descrito anteriormente.

De forma alternativa, existe una homología o similitud sustancial cuando un ácido nucleico o fragmento del mismo se hibrida a otro ácido nucleico, a una hebra de otro ácido nucleico, o a su hebra complementaria, en condiciones de hibridación rigurosas. Las “condiciones de hibridación rigurosas” y “condiciones de lavado rigurosas” en el contexto

de los experimentos de hibridación de ácido nucleico dependen de un número de diferentes parámetros físicos. La hibridación del ácido nucleico se verá afectada por condiciones tales como concentración salina, temperatura, disolventes, la composición de las bases de las especies a hibridar, la longitud de las regiones complementarias y el número de emparejamientos erróneos entre las bases de nucleótidos entre los ácidos nucleicos que se hibridan, tal como apreciarán los expertos en la técnica. Una persona de experiencia ordinaria en la técnica sabe cómo variar estos parámetros para lograr una rigurosidad particular en la hibridación.

En general, la “hibridación rigurosa” se realiza a aproximadamente 25°C por debajo del punto térmico de fusión (T_f) para el híbrido de ADN específico con un conjunto de condiciones particular. El “lavado riguroso” se realiza a aproximadamente 5°C por debajo de la T_f para el híbrido de ADN específico con un conjunto de condiciones particular. La T_f es la temperatura a la que el 50% de la secuencia diana se hibrida a una sonda con la que corresponde perfectamente. Véase Sambrook *et al.*, referencia anterior, página 9.51.

Para los fines del presente documento, las “condiciones muy rigurosas” se definen para una hibridación en fase de solución como hibridación acuosa (es decir, libre de formamida) en 6X SSC (donde 20X SSC contiene NaCl 3,0 M y citrato sódico 0,3 M), 1% SDS a 65°C durante 8-12 horas, seguido de dos lavados en 0,2 X SSC, 0,1% de SDS a 65°C durante 20 minutos. El experto en la técnica apreciará que la hibridación a 65°C se producirá a una velocidad diferente dependiendo de un número de factores que incluyen la longitud e identidad porcentual de las secuencias que se hibridan.

El término “mutado” cuando se aplica a secuencias de ácido nucleico quiere decir que los nucleótidos de una secuencia de ácidos nucleicos puede insertarse, delecionarse o compararse con una secuencia de ácidos nucleicos de referencia. Puede producirse una única alteración en un locus (una mutación puntual) o pueden insertarse, borrarse o cambiarse múltiples nucleótidos en un único locus. Además, pueden producirse una o más alteraciones en cualquier número de locus dentro de una secuencia de ácidos nucleicos. Una secuencia de ácidos nucleicos puede mutarse mediante cualquier procedimiento conocido en la técnica que incluye, pero sin limitación, técnicas de mutagénesis como por ejemplo “PCR con tendencia a errores” (un procedimiento para realizar PCR en condiciones en las que la fidelidad de copiado de la ADN polimerasa es bajo, de tal forma que se obtiene un porcentaje elevado de mutaciones puntuales por toda la longitud del producto de la PCR. Véase, por ejemplo, Leung *et al.* (1989) Technique 1:11-15 y Caldwell y Joyce (1992) PCR Methods Applic. 2:28-33); y “mutagénesis dirigida a los oligonucleótidos” (un procedimiento que permite generar mutaciones específicas de sitio en cualquier segmento de ADN de interés clonado. Véase, por ejemplo, Reidhaar-Olson *et al.* (1988) Science 241:53-57).

El término “vector” tal como se usa en el presente documento se pretende que se refiera a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se haya ligado. Un tipo de vector es un “plásmido”, que se refiere a un bucle de ADN bicatenario en el que pueden ligarse segmentos adicionales de ADN. Otros vectores incluyen cósmidos, cromosomas bacterianos artificiales (BAC) y cromosomas artificiales para levaduras (YAC). Otro tipo de vector es un vector vírico, en el que pueden ligarse segmentos adicionales de ADN en el genoma vírico (se describe en más detalle más adelante). Ciertos vectores pueden replicarse de forma autónoma en una célula huésped en la que se introducen (por ejemplo, vectores que tienen un origen de replicación que funciona en las células huésped). Pueden integrarse otros vectores en el genoma de una célula huésped después de introducirlos en las células huésped, y así se replican junto con el genoma del huésped. Además, ciertos vectores preferentes pueden dirigir la expresión de genes a los que están ligados operablemente. Dichos vectores se denominan en el presente documento “vectores de expresión recombinantes” (o simplemente, “vectores de expresión”).

Secuencias de control de la expresión “ligadas operablemente” se refiere a un enlace en el que la secuencia de control de la expresión está contigua al gen de interés para controlar el gen de interés, así como a secuencias de control de la expresión que actúan en trans o a distancia para controlar el gen de interés.

La expresión “secuencia de control de la expresión” tal como se usa en el presente documento se refiere a secuencias de polinucleótidos que son necesarias para afectar a la expresión de secuencias codificantes a las que están ligadas operablemente. Las secuencias de control de la expresión son secuencias que controlan los eventos de transcripción, postraducción y traducción de secuencias de ácido nucleico. Las secuencias de control de la expresión incluyen secuencias apropiadas de inicio, terminación, de la transcripción y secuencias promotoras y potenciadoras; señales para el procesamiento eficaz de ARN como por ejemplo señales de ajuste y poliadenilación; secuencias que estabilizan el ARNm citoplasmático; secuencias que potencian la eficiencia de la traducción (por ejemplo, sitios de unión de ribosomas); secuencias que potencian la estabilidad de la proteína; y cuando se desee, secuencias que potencian la secreción de la proteína. La naturaleza de dichas secuencias de control difiere dependiendo del organismo huésped; en los procariontes, dichas secuencias de control de forma general incluyen secuencias promotoras, sitio de unión del ribosoma y secuencia de terminación de la transcripción. El término “secuencias de control” se pretende que incluya, como mínimo, todos los componentes cuya presencia es esencial para la expresión, y también puede incluir componentes adicionales cuya presencia es ventajosa, por ejemplo, secuencias líder y secuencias de fusión de parejas.

La expresión “células huésped recombinantes” (o simplemente “células huésped”), tal como se usa en el presente documento, se pretende que se refiera a una célula en la que se ha introducido un ácido nucleico como por ejemplo un vector recombinante. Debería entenderse que dichos términos se pretende que se refieran no únicamente a la célula sujeto particular sino a la progenie de dicha célula. Debido a que pueden producirse ciertas modificaciones en las generaciones posteriores debido o bien a mutación o a influencias ambientales, dicha progenie puede ser idéntica, o

no, a la célula parental, pero aún así están incluidas dentro del alcance del término “célula huésped” tal como se usa en el presente documento. Una célula huésped recombinante puede ser una célula o línea celular aislada cultivada o puede ser una célula que reside en un tejido u organismo vivo.

5 El término “péptido” tal como se usa en el presente documento se refiere a un polipéptido corto, por ejemplo, uno que habitualmente tiene menos de aproximadamente 50 aminoácidos de longitud y más habitualmente menos de aproximadamente 30 aminoácidos de longitud. El término tal como se usa en el presente documento engloba análogos y miméticos que imitan la estructura y por lo tanto la función biológica.

10 El término “polipéptido” tal como se usa en el presente documento engloba tanto proteínas naturales como no naturales, y fragmentos, mutantes, derivados y análogos de las mismas. Un polipéptido puede ser monomérico o polimérico. Además, un polipéptido puede comprender un número de diferentes dominios cada uno de los cuales tiene una o más actividades diferentes.

15 El término “proteína aislada” o “polipéptido aislado” es un proteína o polipéptido que en virtud de su origen o fuente de derivación (1) no está asociado a los componentes con los que naturalmente esta asociado que lo acompañan en su estado nativo, (2) cuando existe en una pureza que no se encuentra en la naturaleza, donde la pureza puede determinarse con respecto a la presencia de otro material celular (por ejemplo, está libre de otras proteínas de la misma especie) (3) es expresado por una célula de una especie diferentes, o (4) no existe en la naturaleza (por ejemplo, es un
20 fragmento de un polipéptido que se encuentra en la naturaleza o incluye análogos o derivados de aminoácidos que no se encuentran en la naturaleza o con enlaces distintos de los enlaces peptídicos estándar). Así, un polipéptido que se sintetiza químicamente o que se sintetiza en un sistema celular diferente de la célula de la que se origina estará “aislado” de sus componentes naturalmente asociados. Un polipéptido o proteína también puede hacerse sustancialmente libre de los componentes naturalmente asociados aislándolo, usando técnicas de purificación de proteínas notorias en la
25 técnica. Definido como tal, “aislado” no requiere necesariamente que la proteína, polipéptido, péptido u oligopéptido que se describe así haya sido retirado físicamente de su entorno nativo.

El término “fragmento polipeptídico” tal como se usa en el presente documento se refiere a un polipéptido que tiene una delección en el extremo amino y/o en el extremo carboxi comparado con un polipéptido de longitud completa. En una realización preferente, el fragmento polipeptídico es una secuencia contigua en la que la secuencia de
30 aminoácidos del fragmento es idéntica a la de las posiciones correspondientes en la secuencia natural. Los fragmentos habitualmente son de al menos 5, 6, 7, 8, 9 ó 10 aminoácidos de longitud, preferentemente al menos 12, 14, 16 ó 18 aminoácidos de longitud, más preferentemente al menos 20 aminoácidos de longitud, más preferentemente al menos 25, 30, 35, 40 ó 45, aminoácidos, incluso más preferentemente al menos 50 ó 60 aminoácidos de longitud, e incluso
35 más preferentemente al menos 70 aminoácidos de longitud.

Un “derivado modificado” se refiere a polipéptidos o sus fragmentos que son sustancialmente homólogos en la secuencia estructural primaria pero que incluyen, por ejemplo, modificaciones químicas y bioquímicas *in vivo* o *in vitro* o que incorporan aminoácidos que no se encuentran en el polipéptido nativo. Dichas modificaciones incluyen,
40 por ejemplo, acetilación, carboxilación, fosforilación, glucosilación, ubiquitinación, marcado, por ejemplo, con radio-núclidos, y diversas modificaciones enzimáticas, como fácilmente entenderán los expertos en la técnica. Una variedad de procedimientos para marcar los polipéptidos y de sustituyentes o marcadores útiles para dichos fines son notorios en la técnica, e incluyen isótopos radioactivos como por ejemplo ¹²⁵I, ³²P, ³⁵S, y ³H, ligandos que se unen a antiligandos marcados (por ejemplo, anticuerpos), fluoróforos, agentes quimioluminiscentes, enzimas, y antiligandos que pueden
45 servir como miembros de pares de unión específicos para un ligando marcado. La elección del marcador depende de la sensibilidad necesaria, facilidad de conjugación con el cebador, requisitos de estabilidad, e instrumentación disponible. Los procedimientos para marcar polipéptidos son notorios en la técnica. Véase Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992, y suplementos hasta 2002).

50 Un “polipéptido mutante” o “muteína” se refiere a un polipéptido cuya secuencia contiene una inserción, duplicación, delección, reordenación o sustitución de uno o más aminoácidos comparados con la secuencia de aminoácidos de una proteína nativa o de tipo silvestre. Una muteína puede tener una o más sustituciones puntuales de aminoácidos, en las que un único aminoácido en una posición se ha cambiado a otro aminoácido, una o más inserciones y/o delecciones, en las que uno o más aminoácidos se insertan o se eliminan, respectivamente, de la secuencia de la proteína natural y/o
55 truncados de la secuencia de aminoácidos en cualquiera o en ambos extremos amino o carboxi. Una muteína puede tener la misma actividad biológica pero preferentemente tiene una a diferente de la proteína natural.

Una muteína tiene al menos 70% de homología global entre las secuencias con su homólogo de tipo silvestre. Incluso más preferente son muteínas que tienen 80%, 85% o 90% homología global entre las secuencias con la proteína de tipo silvestre. En una realización todavía más preferente, una muteína muestra un 95% de identidad entre las
60 secuencias, incluso más preferentemente 97%, incluso más preferentemente 98% e incluso más preferentemente 99% de homología global entre las secuencias. La homología entre las secuencias puede medirse mediante algoritmo de análisis de las secuencias habitual, como por ejemplo Gap o Bestfit.

65 Preferente las sustituciones de aminoácidos son aquellas que: (1) reducen la susceptibilidad a la proteólisis, (2) reducen la susceptibilidad a la oxidación, (3) alteran la afinidad de unión para formar complejos de proteínas, (4) alteran la afinidad de unión o la actividad enzimática, y (5) confieren o modifican otras propiedades fisicoquímicas o funcionales de dichos análogos.

Tal como se usa en el presente documento, los veinte aminoácidos convencionales y sus abreviaturas siguen el uso convencional. Véase Immunology - A Synthesis (2ª Edición, ES. Golub y D. R. Gren, Eds., Sinauer Associates, Sunderland, Mass. (1991)), estereoisómeros (por ejemplo, D-aminoácidos) de los veinte aminoácidos convencionales, aminoácidos no naturales como por ejemplo aminoácidos α -, α -disustituidos, *N*-alquil aminoácidos, y otros aminoácidos no convencionales también pueden ser componentes adecuados para los polipéptidos de la presente invención. Los ejemplos de aminoácidos no convencionales incluyen: 4-hidroxiprolina, γ -carboxiglutamano, ϵ -N,N,N-trimetil-lisina, ϵ -N-acetil-lisina, O-fosfoserina, *N*-acetilserina, *N*-formilmetionina, 3-metilhistidina, 5-hidroxilisina, *N*-metilarginina, y otros aminoácidos e iminoácidos similares (por ejemplo, 4-hidroxiprolina). En la notación de los polipéptidos que se usa en el presente documento, la mano izquierda es la dirección del extremo amino y la mano derecha es la dirección del extremo carboxi, de acuerdo con el uso y la convención estándar.

Una proteína tiene “homología” o es “homóloga” a una segunda proteína si la secuencia de ácidos nucleicos que codifica la proteína tiene una secuencia similar a la secuencia de ácidos nucleicos que codifica la segunda proteína. De forma alternativa, una proteína presenta homología con una segunda proteína si las dos proteínas tienen secuencias de aminoácidos “similares”. (Así, el término “proteínas homólogas” se define como que significa que las dos proteínas tienen secuencias de aminoácidos similares). En una realización preferente, una proteína homóloga es una que muestran una homología del 60% entre las secuencias con la proteína de tipo silvestre, más preferente es un 70% de homología entre las secuencias. Incluso más preferentes son las proteínas homólogas que muestran un 80%, 85% o 90% de homología entre las secuencias con la proteína de tipo silvestre. Todavía en otra realización más preferente, una proteína homóloga muestra un 95%, 97%, 98% o 99% de identidad entre las secuencias. Tal como se usa en el presente documento, homología entre dos regiones de una secuencia de aminoácidos (especialmente con respecto similitudes estructurales previstas) se interpreta como que implica similitud en la función.

Cuando “homóloga” se usa con respecto a proteínas o péptidos, se reconoce que las posiciones de los restos que no son idénticas a menudo difieren en sustituciones conservadoras de aminoácidos. Una “sustitución conservadora de aminoácidos” es una en la que un resto aminoacídico es sustituida por otro resto aminoacídico que tiene una cadena lateral (grupo R) con propiedades química similares (por ejemplo, carga o hidrofobicidad). En general, una sustitución conservadora de aminoácidos no cambiará sustancialmente las propiedades funcionales de una proteína. En los casos en los que dos o más secuencias de aminoácidos difieren entre sí por sustituciones conservadoras, la identidad porcentual entre secuencias o el grado de homología puede ajustarse hacia arriba para corregir la naturaleza conservadora de la sustitución. Los medios para preparar este ajuste son notorios para los expertos en la técnica (véase, por ejemplo, Pearson (1990) *Methods Enzymol.* 183:63-98).

Los seis grupos siguientes contienen cada uno aminoácidos que son sustituciones conservadoras los unos de los otros: 1) serina (S), treonina (T); 2) ácido aspártico (D), ácido glutámico (E); 3) asparragina (N), glutamina (Q); 4) arginina (R), lisina (K); 5) isoleucina (I), leucina (L), metionina (M), alanina (A), valina (V), y 6) fenilalanina (F), tirosina (Y), triptófano (W).

La homología entre las secuencias para los polipéptidos, que también se denomina identidad porcentual entre secuencias, habitualmente se mide usando programas para el análisis de secuencias. Véase, por ejemplo, el Sequence Analysis Software Package del Genetics Computer Group (GCG), University of Wisconsin Biotechnology Center, 910 University Avenue, Madison, Wisconsin 53705. El programa de análisis de proteínas empareja secuencias similares usando la medición de la homología asignada a diversas sustituciones, delecciones y otras modificaciones, que incluyen sustituciones conservadoras de aminoácidos. Por ejemplo, GCG contiene programas como por ejemplo “Gap” y “Bestfit” que pueden usarse con los parámetros por defecto para determinar la homología entre las secuencias o la identidad entre las secuencias entre polipéptidos similares, como por ejemplo polipéptidos homólogos de diferentes especies de organismos o entre una proteína de tipo silvestre y una muteína de la misma. Véase, por ejemplo, GCG Versión 6.1.

Un algoritmo preferente cuando se compara una secuencia de una molécula inhibidora con una base de datos que contiene un gran número de secuencias de diferentes organismos es el programa de ordenador BLAST (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410; Gish and States (1993) *Nature Genet.* 3:266-272; Madden *et al.* (1996) *Meth. Enzymol.* 266:131-141; Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402; Zhang y Madden (1997) *Genome Res.* 7:649-656), especialmente blastp o tblastn (Altschul *et al.*, 1997). Los parámetros preferentes para BLASTp son: Valor de expectación: 10 (por defecto); Filtro: seg (por defecto); Coste por abrir un espacio: 11 (por defecto); Coste por extender un espacio: 1 (por defecto); Max. de alineaciones: 100 (por defecto); Tamaño de la palabra: 11 (por defecto); N.º de descripciones: 100 (por defecto); Matriz de penalizaciones: BLOWSUM62.

La longitud de las secuencias de polipéptidos que se comparan para determinar la homología de forma general será al menos aproximadamente de 16 restos aminoacídicos, habitualmente al menos aproximadamente 20 restos, más habitualmente al menos aproximadamente 24 restos, habitualmente al menos aproximadamente 28 restos, y preferentemente más de aproximadamente 35 restos. Cuando se busca en una base de datos que contiene secuencias de un gran número de organismos diferentes, es preferible comparar secuencias de aminoácidos. Las búsquedas en bases de datos usando secuencias de aminoácidos puede medirse mediante algoritmos distintos de blastp conocidos en la técnica. Por ejemplo, pueden compararse las secuencias de polipéptidos usando FASTA, un programa de GCG Versión 6.1. FASTA proporciona alineaciones e identidad porcentual entre secuencias de las regiones con mejor coincidencia entre las secuencias de interrogación y de búsqueda (Véase Pearson (1990) *Methods Enzymol.* 183:63-98). Por ejemplo, la identidad porcentual de las secuencias entre secuencias de aminoácidos puede determinarse usando FASTA con sus

parámetros por defecto (un tamaño de palabra de 2 y la matriz de puntuación Palm250), que se proporciona en GCG Versión 6.1.

El término “proteína de fusión” se refiere a un polipéptido que comprende un polipéptido o fragmento acoplado a secuencias de aminoácidos heterólogas. Las proteínas de fusión son útiles porque puede interpretarse que contienen dos o más elementos funcionales deseados de dos o más proteínas diferentes. Una proteína de fusión comprende al menos 10 aminoácidos contiguos de un polipéptido de interés, más preferentemente al menos 20 ó 30 aminoácidos, incluso más preferentemente al menos 40, 50 ó 60 aminoácidos, todavía más preferentemente al menos 75, 100 ó 125 aminoácidos. Las proteínas de fusión pueden producirse de forma recombinante construyendo una secuencia de ácidos nucleicos que codifica el polipéptido o un fragmento del mismo en marco con una secuencia de ácidos nucleicos que codifica una proteína o péptido diferente y después expresando la proteína de fusión. De forma alternativa, puede producirse una proteína de fusión químicamente ligando el polipéptido o un fragmento del mismo a otra proteína.

El término “región” tal como se usa en el presente documento se refiere a una porción físicamente contigua de la estructura primaria de una biomolécula. En el caso de las proteínas, una región viene definida por una porción contigua de la secuencia de aminoácidos de esa proteína.

El término “dominio” tal como se usa en el presente documento se refiere a una estructura de una biomolécula que contribuye a una función conocida o que se sospecha de la biomolécula. Los dominios pueden tener la misma extensión que regiones o sus porciones; los dominios también pueden incluir regiones diferentes, no contiguas de una biomolécula. Los ejemplos de dominios de proteínas incluyen, pero sin limitación, un dominio de Ig, un dominio extracelular, un dominio transmembrana, y un dominio citoplasmático.

Tal como se usa en el presente documento, el término “molécula” quiere decir cualquier compuesto, que incluye, pero sin limitación, una molécula pequeña, péptido, proteína, azúcar, nucleótido, ácido nucleico, lípido, etc., y dicho compuesto puede ser natural o sintético.

A no ser que se defina de otro modo, todos los términos incluidos los técnicos y científicos que se usan en la presente memoria descriptiva tienen el mismo significado que entiende habitualmente una persona de experiencia ordinaria en la técnica a la que pertenece esta invención. Más adelante se describen procedimientos y materiales ejemplares, aunque también pueden usarse procedimientos y materiales similares o equivalentes a los que se describen en el presente documento en la práctica de la presente invención y serán obvios para los expertos en la técnica.

En caso de conflicto, prevalecerá la presente memoria descriptiva, que incluye las definiciones. Los materiales, procedimientos, y ejemplos son únicamente ilustrativos y no se pretende que sean limitantes.

En toda esta memoria descriptiva y en las reivindicaciones, la palabra “comprenden” o sus variaciones como por ejemplo “comprende” o “que comprende”, se entenderá que implican la inclusión de un número o de grupo de números citados pero no la exclusión de cualquier otro número o grupo de números.

Procedimientos para producir glucoproteínas humanoides en células huésped de eucariotas inferiores

La invención proporciona procedimientos para producir una glucoproteína que tiene glucosilación humanoide en células huésped de hongos unicelulares o multicelulares. Como se describe en más detalle más adelante, se selecciona como célula huésped inicial una célula huésped de hongos unicelulares o multicelulares que no expresa de forma natural, o que está genomanipulada para que no exprese, una o más enzimas implicadas en la producción de estructuras altas en manosa. Dicha célula huésped seleccionada se genomanipula para que exprese una o más enzimas u otros factores necesarios para producir glucoproteínas humanoides. Una cepa huésped deseada puede genomanipularse con una enzima o más de una enzima de cada vez. Además, puede usarse una molécula de ácido nucleico que codifica una o más enzimas o actividades para diseñar una cepa huésped de la invención. Preferentemente, se crea un conjunto de moléculas de ácido nucleico que codifican enzimas potencialmente útiles (por ejemplo, enzimas quiméricas que comprenden un fragmento de enzima catalíticamente activo ligado en marco a una secuencia directora subcelular heteróloga) (por ejemplo, ligando subadenotecas que comprenden fragmentos enzimáticos y secuencias directoras subcelulares), y puede seleccionarse una cepa que tiene una o más enzimas con actividades óptimas o que producen la mayor cantidad de glucoproteínas “humanoides” transformando células huésped diana con uno o más miembros del conjunto.

En particular, los procedimientos que se describen en el presente documento permiten obtener, *in vivo*, estructuras $\text{Man}_5\text{GlcNAc}_2$ con alto rendimiento, al menos de forma transitoria, con el fin de modificarlas posteriormente proporcionando *N*-glucanos complejos. Un esquema con éxito para obtener estructuras $\text{Man}_5\text{GlcNAc}_2$ adecuadas con rendimientos apropiado en una célula huésped, como por ejemplo un organismo eucariota inferior, de forma general supone dos estrategias paralelas: (1) reducir las elevadas estructuras de manosa preparadas mediante las actividades endógenas de las manosiltransferasas, si las hubiera, y (2) eliminar 1,2- α -manosa mediante las manosidasas proporcionando niveles altos de estructuras $\text{Man}_5\text{GlcNAc}_2$ adecuadas que pueden hacerse reaccionar posteriormente en la célula huésped formando glucoformas humanoides complejas.

Por consiguiente, una primera etapa supone la selección o creación de una célula huésped eucariota, por ejemplo, de un eucariota inferior, capaz de producir una estructura precursora específica de $\text{Man}_5\text{GlcNAc}_2$ que sea capaz de

aceptar GlcNAc *in vivo* mediante la acción de una GlcNAc transferasa I ("GnTI"). En una realización, el procedimiento supone preparar o usa una célula huésped eucariota no humana carente de una actividad de 1,6 manosiltransferasa con respecto a *N*-glucano de una glucoproteína. Preferentemente, la célula huésped carente de una actividad de iniciación de 1,6 manosiltransferasa (véase más adelante). Dicha célula huésped carecerá de una o más enzimas implicadas en la producción de estructuras con manosa elevada que no son deseables para producir glucoproteínas humanoides.

Después se introducen en dicha célula huésped una o más actividades enzimáticas para producir *N*-glucanos dentro de las células huésped que se caracteriza porque tiene al menos 30% molar de las estructuras de carbohidrato Man₅GlcNAc₂ ("Man₅"). Las estructuras Man₅GlcNAc₂ son necesarias para la formación de *N*-glucanos complejos: Man₅GlcNAc₂ debe formarse *in vivo* con a alto rendimiento (por ejemplo, más del 30%), al menos de forma transitoria, ya que las posteriores reacciones de glucosilación de tipo mamífero y humanoide requieren Man₅GlcNAc₂ o un derivado del mismo.

Esta etapa también requiere la formación de una estructura isomérica particular de Man₅GlcNAc₂ en el interior de la célula con alto rendimiento. Aunque las estructuras Man₅GlcNAc₂ son necesarias para la formación de *N*-glucanos complejos, su presencia en modo alguno es suficiente. Esto es debido a que Man₅GlcNAc₂ puede darse en diferentes formas isoméricas, que pueden o no servir como sustrato para GlcNAc transferasa I. Como la mayoría de las reacciones de glucosilación no están completas, una proteína glucosilada particular de forma general contiene un abanico de diferentes estructuras de carbohidratos (*es decir*, glucoformas) sobre su superficie. Así, la mera presencia de cantidades traza (es decir, de menos de 5%) de una estructura particular como Man₅GlcNAc₂ tiene poca relevancia práctica para producir glucoproteínas de tipo mamífero o humanoides. Lo que es necesario es la formación de un intermedio de Man₅GlcNAc₂ que acepte GlcNAc transferasa I (Figura 1 B) con alto rendimiento (*es decir*, superior al 30%). La formación de este intermedio es necesaria para permitir la posterior síntesis *in vivo* de *N*-glucanos complejos sobre proteínas glucosiladas de interés (proteínas diana).

Por consiguiente, parte o todo el Man₅GlcNAc₂ producido por las células huésped seleccionadas debe ser un sustrato productivo para las actividades enzimáticas a lo largo de una ruta de glucosilación de mamífero, por ejemplo, puede servir como sustrato para una actividad de GlcNAc transferasa I *in vivo*, formando así el *N*-glucano intermedio humanoide GlcNAcMan₅GlcNAc₂ en las células huésped. En una realización preferente, al menos 10%, más preferentemente al menos 30% y lo más preferentemente 50% o más del Man₅GlcNAc₂ intermedio producido en las células huésped de la invención es un sustrato productivo para GnTI *in vivo*. Se entiende que si, por ejemplo, GlcNAcMan₅GlcNAc₂ se produce al 10% y Man₅GlcNAc₂ se produce al 25% en una proteína diana, que la cantidad total del Man₅GlcNAc₂ producido de forma transitoria es del 35% porque GlcNAcMan₅GlcNAc₂ es un producto de Man₅GlcNAc₂.

Una persona de experiencia ordinaria en la técnica puede seleccionar células huésped de la naturaleza, por ejemplo, hongos u otros eucariotas inferiores existentes que produzcan niveles significativos de Man₅GlcNAc₂ *in vivo*. Hasta la fecha, sin embargo, no se ha demostrado que ningún eucariota inferior proporcione dichas estructuras *in vivo* más allá del 1,8% de los *N*-glucanos totales (véase por ejemplo Maras *et al.* (1997) Eur. J. Biochem. 249:701-707). De forma alternativa, dichas células huésped pueden ser genomanipuladas para que produzcan la estructura Man₅GlcNAc₂ *in vivo*. Pueden emplearse procedimientos como por ejemplo los que se describe en la patente de Estados Unidos n.º 5.595.900 para identificar la ausencia o presencia de glucosiltransferasas, manosidasas y transportadores de azúcares de nucleótidos particulares en una célula huésped u organismo de interés diana.

Inactivación de enzimas de glucosilación indeseables en células huésped

Los procedimientos de la invención se dirigen a preparar células huésped que producen glucoproteínas que tienen estructuras de *N*-glucano alteradas y preferentemente humanoides. En una realización preferente, los procedimientos se dirigen a preparar células huésped en las que los precursores oligosacáridos están enriquecidos en Man₅GlcNAc₂. Preferentemente, se usa una célula huésped eucariota que no expresa una o más enzimas implicadas en la producción de estructuras con alto contenido en manosa. Dicha célula huésped puede encontrarse en la naturaleza o puede ser genomanipulada, por ejemplo, empezando o derivando a partir de una de muchos de dichos mutantes que ya se han descrito en las levaduras. Así, dependiendo de la célula huésped seleccionada, habrá que eliminar uno o un número de genes que codifican enzimas que se sabe son características de las reacciones de glucosilación no humanas. Dichos genes y sus proteínas correspondientes han sido caracterizados extensamente en un número de eucariotas inferiores (por ejemplo, *S. cerevisiae*, *T. reesei*, *A. nidulans*, etc.), proporcionando así una lista de glucosiltransferasas conocidas en eucariotas inferiores, sus actividades y su secuencia genética respectiva. Estos genes probablemente se seleccionarán a partir del grupo de las manosiltransferasas, por ejemplo 1,3 manosiltransferasas (por ejemplo MNN1 de *S. cerevisiae*) (Graham, 1991), 1,2 manosiltransferasas (por ejemplo la familia KTRI KRE de *S. cerevisiae*), 1,6 manosiltransferasas (OCH1 de *S. cerevisiae*), manosilfosfato transferasas y sus reguladores (MNN4 y MNN6 de *S. cerevisiae*) y enzimas adicionales que están implicadas en reacciones de glucosilación aberrantes, es decir, no humanas. Muchos de estos genes, de hecho, se han suprimido de forma individual, dando lugar a genotipos viables con perfiles de glucosilación alterados. Se muestran ejemplos en la Tabla 1.

Las células huésped de eucariotas inferiores preferentes de la invención, como se describe en el presente documento para ejemplificar las etapas de manipulación necesarias, son los mutantes sin hipermanosilación (*och1*) de *Pichia pastoris* o *K. lactis*. Al igual que otros eucariotas inferiores, *P. pastoris* procesa las estructuras Man₅GlcNAc₂ en el RE con una α -1,2-manosidasa proporcionando Man₃GlcNAc₂ (Figura 1A). A través de la acción de varias manosiltrans-

ferasas, esta estructura se convierte después en estructuras hipermanosiladas ($\text{Man}_{>9}\text{GlcNAc}_2$), también conocidas como mananos (Figura 35A). Además, se ha encontrado que *P. pastoris* es capaz de añadir grupos fosfato no terminales, a través de la acción de manosilfosfato transferasas a la estructura de carbohidratos. Esto difiere de las reacciones realizadas en células de mamífero, que suponen la eliminación, en lugar de la adición de azúcares de manosa (Figura 35A). Es de importancia particular eliminar la capacidad de la célula huésped eucariota de hipermanosilar la estructura de $\text{Man}_8\text{GlcNAc}_2$ existente. Esto puede lograrse o bien seleccionando una célula huésped que no hipermanosile, o diseñando genéticamente una célula así.

Se han identificado los genes que están implicados en el procedimiento de hipermanosilación, por ejemplo, en *Pichia pastoris*, y al crear mutaciones en estos genes, se puede reducir la producción de glucoformas "indeseables". Dichos genes pueden identificarse por homología con manosiltransferasas o sus reguladores existentes (por ejemplo *OCH1*, *MNN4*, *MNN6*, *MNN1*), que se encuentran en otros eucariotas inferiores como por ejemplo *C. albicans*, *Pichia angusta* o *S. cerevisiae* o mutando la cepa huésped y seleccionando un fenotipo de glucosilación con manosilación reducida. Basándose en las homologías entre manosiltransferasas y manosilfosfato transferasas conocidas, se pueden diseñar cebadores de PCR, ejemplos de los cuales se muestran en la Tabla 2, o usar genes o fragmentos de genes que codifican dichas enzimas como sondas para identificar homólogos en adenotecas del organismo diana o de uno relacionado. De forma alternativa, se puede identificar un homólogo funcional que tiene actividad de manosiltransferasa por su capacidad de complementar fenotipos de glucosilación manosiltransferasa en organismos relacionados.

TABLA 2
Cebadores de PCR

Cebador A de PCR	Cebador B de PCR	Gen(es) diana en <i>P. pastoris</i>	Homólogos
		1,6-manosiltransferasa	<i>OCH1 S. cerevisiae</i> , <i>Pichia albicans</i>
		1,2-manosiltransferasas	familia KTR/KRE, <i>S. cerevisiae</i>
Legenda: M = A o C, R = A o G, W = A o T, S = C o G, Y = C o T, K = G o T, V = A o C o G, H = A o C o T, D = A o G o T, B = C o G o T, N = G o A o T o C.			

Para obtener el gen o genes que codifican la actividad de 1,6-manosiltransferasa en *P. pastoris*, por ejemplo, se realizarían las siguientes etapas: los mutantes en *OCH1* de *S. cerevisiae* son sensibles a la temperatura y se desarrollan despacio a temperaturas elevadas. Así se puede identificar a los homólogos funcionales de *OCH1* en *P. pastoris* complementando un mutante en *OCH1* de *S. cerevisiae* con un ADN o adenoteca de *P. pastoris*. Los mutantes de *S. cerevisiae* pueden obtenerse, por ejemplo, en la Universidad de Stanford y están comercialmente disponibles en ResGen, Invitrogen Corp. (Carlsbad, CA). Los mutantes que presentan un fenotipo de crecimiento normal a temperatura elevada, después de haber sido transformados con una adenoteca de *P. pastoris*, con probabilidad portan un homólogo de *OCH1* de *P. pastoris*. Dicha adenoteca puede crearse digiriendo parcialmente ADN cromosómico de *P. pastoris* con una enzima de restricción adecuada y después inactivando la enzima de restricción que liga el ADN digerido en un vector adecuado, que se ha digerido con enzimas de restricción compatibles.

Los vectores adecuados incluyen, por ejemplo, pRS314, un plásmido (CEN6/ARS4) con pocas copias basado en pBluescript que contiene el marcador Trp1 (Sikorski y Hieter (1989) Genetics 122:19-27) y pFL44S, un plásmido (2μ) de muchas copias basado en un pUC19 modificado que contiene el marcador URA3 (Bonneaud *et al.* (1991) Yeast 7:609-615). Dichos vectores habitualmente son empleados por investigadores académicos y hay vectores similares disponibles en numerosos proveedores (por ejemplo Invitrogen (Carlsbad, CA), Pharmacia (Piscataway, NJ), New England Biolabs (Beverly, MA)). Otros ejemplos incluyen el plásmido de expresión en levaduras basado en el origen de replicación 2μ pYES/GS, de Invitrogen, o el vehículo de clonación Yep24 de New England Biolabs.

Después de ligar el ADN cromosómico y el vector, se puede transformar la adenoteca en la cepa de *S. cerevisiae* con una mutación específica y seleccionar la corrección del fenotipo correspondiente. Después de subclonar y secuenciar el fragmento de ADN que es capaz de restaurar el fenotipo silvestre, puede usarse este fragmento para eliminar la actividad del producto génico codificado por *OCH1* en *P. pastoris* usando mutagénesis *in vivo* y/o técnicas de recombinación notorias para los expertos en la técnica.

De forma alternativa, si se conoce toda la secuencia genómica de una célula huésped particular, por ejemplo un hongo, de interés, se pueden identificar dichos genes simplemente buscando en bases de datos de ADN disponibles al público, que están disponibles en varias fuentes como por ejemplo NCBI, Swissprot. Por ejemplo, buscando una secuencia genómica dada o base de datos con secuencias de un gen de 1,6 manosiltransferasa conocidos (por ejemplo, *OCH1* de *S. cerevisiae*), se pueden identificar los genes de elevada homología en dicho genoma de la célula huésped que podría (pero no necesariamente) codificar proteínas que presentan actividad de 1,6-manosiltransferasa. La homología de la secuencia de ácidos nucleicos por sí sola no es suficiente como prueba, sin embargo, de que se ha identificado y aislado un homólogo que codifica una enzima que tiene la misma actividad. Hasta la fecha, por ejemplo, no existen datos que demuestren que una delección en *OCH1* en *P. pastoris* elimina la clave que inicia la actividad de 1,6-manosiltransferasa (Martinet *et al.* (1998) Biotech. Letters 20(12):1171-1177; Contreras *et al.* documento WO 02/00856 A2). Así, no hay datos que prueben que el homólogo del gen *OCH1* de *P. pastoris* realmente codifica esa función. Esa demostración se proporciona por primera vez en el presente documento.

Se han identificado homólogos de varias manosiltransferasas de *S. cerevisiae* en *P. pastoris* usando estas estrategias. Los genes homólogos a menudo tienen funciones similares a las de los genes implicados en la manosilación de las proteínas de *S. cerevisiae* y así su delección puede usarse para manipular el patrón de glucosilación de *P. pastoris* o, por analogía, en cualquier otra célula huésped, por ejemplo, células fúngicas, vegetales, de insectos o animales, con rutas de glucosilación similares.

La creación de inactivaciones génicas, una vez se ha determinado una secuencia génica diana, es una técnica bien establecida en la técnica y puede realizarla una persona de experiencia ordinaria en la técnica (véase, por ejemplo, Rothstein (1991) Methods in Enzymology 194:281). La elección de un organismo huésped puede verse influenciada por la disponibilidad de buenas técnicas de transformación y de alteración génica.

Si deben inactivarse varias manosiltransferasas, el procedimiento desarrollado por Alani y Kleckner (1987) Genetics 116:541-545, por ejemplo, permite usar repetidamente un marcador seleccionable, por ejemplo, el marcador *URA3* en levaduras, para eliminar secuencialmente toda la actividad de las manosiltransferasas endógenas deseable. Esta técnica ha sido refinada por otros, pero básicamente supone el uso de dos secuencias de ADN repetidas, flanqueando un marcador contraseleccionable. Por ejemplo: puede usarse *URA3* como marcador para asegurar la selección de transformantes que tienen una construcción integrada. Al flanquear el marcador *URA3* con repeticiones directas, uno podría seleccionar primero los transformantes que tengan la construcción integrada y que hayan alterado así el gen diana. Después de aislar los transformantes, y de su caracterización, se puede volver a realizar una selección en una segunda ronda los que son resistentes a ácido 5-fluoroorótico (5-FOA). Las colonias que pueden sobrevivir sobre las placas que contienen 5-FOA han perdido su marcador *URA3* de nuevo por un evento de cruzamiento que afecta a las repeticiones mencionadas anteriormente. Esta estrategia permite así el uso repetido del mismo marcador y facilita la alteración de múltiples genes sin requerir marcadores adicionales. También pueden usarse técnicas similares para la eliminación secuencial de los genes adaptados para usar en otras células huésped eucariotas con otros marcadores seleccionables y contraseleccionables.

La eliminación de manosiltransferasas específicas, como por ejemplo 1,6-manosiltransferasa (*OCH1*) o manosilfosfato transferasas (MNN6, o genes que complementan los mutantes en *Ibd*) o reguladores (MNN4) en *P. pastoris* permite crear cepas genomanipuladas de este organismo que sintetizan Man₃GlcNAc₂ primario y que pueden usarse para modificar el patrón de glucosilación para que imite estructuras de glucoformas más complejas, por ejemplo, las producidas en células de mamífero, por ejemplo, células humanas. Una realización preferente de este procedimiento utiliza secuencias de ADN que codifican actividades de glucosilación bioquímica para eliminar funciones bioquímicas similares o idénticas en *P. pastoris* para modificar la estructura de glucosilación de las glucoproteínas producidas en la cepa genéticamente alterada de *P. pastoris*.

Los procedimientos que se usan para diseñar la ruta de glucosilación en levaduras como se presenta en los ejemplos del presente documento pueden usarse en hongos filamentosos para producir un sustrato preferente para la posterior modificación. Pueden desarrollarse estrategias para modificar las rutas de glucosilación en *A. niger* y otros hongos filamentosos, por ejemplo, usando protocolos análogos a los que se describen en el presente documento para diseñar cepas para producir glucoproteínas humanoides en levaduras. Las actividades de genes indeseadas implicadas en la actividad de 1,2-manosiltransferasa, por ejemplo, homólogos de *KTRIKRE*, se modifican o se eliminan. Un hongo filamentos, como por ejemplo *Aspergillus*, es un huésped preferente porque carece de la actividad de 1,6-manosiltransferasa y, por lo tanto, no se puede esperar una actividad de hipermanosilación, por ejemplo *OCH1*, en este huésped. Por el contrario, se introducen en el huésped otras actividades deseadas (por ejemplo, α -1,2-manosidasa, transportador de UDP-GlcNAc, glucosiltransferasa (GnT), galactosiltransferasa (GalT) y sialiltransferasa (ST)) implicadas en la glucosilación usando los procedimientos de direccionamiento de la invención.

Genomanipulación o selección de huéspedes que tienen una menor actividad de α -1,6-manosiltransferasa de iniciación

En una realización preferente, el procedimiento de la invención supone preparar o usar una célula huésped que tiene una actividad menor o anulada de una α -1,6-manosiltransferasa de iniciación, es decir, una enzima específica de iniciación que inicia la manosilación de la cadena externa en la ramificación α -1,3 de la estructura nuclear de Man₃GlcNAc₂. En *S. cerevisiae*, esta enzima está codificada por el gen *OCH1*. La alteración del gen *OCH1* en *S.*

cerevisiae produce un fenotipo en el que los azúcares ligados en N carecen completamente de la cadena exterior de polimannosa. Las estrategias previas para obtener una glucosilación de tipo mamífero en cepas fúngicas ha necesitado de la inactivación de *OCH1* (véase, por ejemplo, Chiba *et al.* (1998) J. Biol. Chem. 273:26298-304). La alteración de la actividad de α -1,6-manosiltransferasa de iniciación en una célula huésped de la invención puede ser opcional, sin embargo (dependiendo de la célula huésped seleccionada), dado que la enzima Och1p necesita un Man₃GlcNAc₂ intacto para la iniciación eficaz de la cadena externa de manosa. Así, las células huésped seleccionadas o producidas de acuerdo con esta invención que acumulan oligosacáridos que tienen siete o menos restos de manosa puede producir N-glucanos hipoglucosilados que probablemente sean sustratos deficientes para Och1p (véase, por ejemplo, Nakayama *et al.* (1997) FEBS Lett. 412(3):547-50).

El gen *OCH1* se clonó a partir de *P. pastoris* (Ejemplo 1) y *K. lactis* (Ejemplo 9), como se ha descrito. Las secuencias de ácido nucleico y de aminoácidos del gen *OCH1* de *K. lactis* se describen en las SEC. ID. N.º: 7 y 8. Usando cebadores específicos de gen, se preparó una construcción a partir de cada clon para eliminar el gen *OCH1* del genoma de *P. pastoris* y *K. lactis* (Ejemplos 1 y 9, respectivamente). Así se obtuvieron células huésped sin actividad de α -1,6-manosiltransferasa de iniciación y genomanipuladas para que produzcan N-glucanos que tienen una estructura de carbohidrato Man₅GlcNAc₂ (véase, por ejemplo, Las Figuras 5, 6, y 12; Ejemplos 4 y 9).

Así, en el presente documento se describe una molécula de ácido nucleico aislada que tiene una secuencia de ácidos nucleicos que comprende o constituida por al menos 45, preferentemente al menos 50, más preferentemente al menos 60 y lo más preferentemente 75 o más restos nucleotídicos del gen *OCH1* de *K. lactis* (SEC. ID. N.º: 7), y sus homólogos, variantes y derivados. También se describen en el presente documento moléculas de ácido nucleico que se hibridan en condiciones rigurosas a las moléculas de ácido nucleico descritas anteriormente. De forma similar, en el presente documento se describen polipéptidos aislados (que incluyen muteínas, variantes de alelos, fragmentos, derivados, y análogos) codificados por las moléculas de ácido nucleico de la invención. También se describen en el presente documento vectores, que incluyen vectores de expresión, que comprenden las moléculas de ácido nucleico de la invención anteriores, como se describen adicionalmente en el presente documento. De forma similar, se describen en el presente documento células huésped transformadas con las moléculas de ácido nucleico o vectores de la invención.

La invención proporciona además procedimientos de preparar o usar una célula huésped de hongo unicelular o multicelular con una actividad menor o anulada del gen *alg* (*es decir*, actividades de *alg*, que incluyen actividades enzimáticas equivalentes en células huésped no fúngicas) y de introducir en las células huésped al menos una actividad de glucosidasa. En una realización preferente, la actividad de glucosidasa se introduce provocando la expresión de una o más actividades de manosidasa en las células huésped, por ejemplo, por activación de una actividad de manosidasa, o por expresión a partir de una molécula de ácido nucleico de una actividad de manosidasa en las células huésped.

En otra realización, el procedimiento supone preparar o usar una célula huésped con una actividad menor o anulada de una o más enzimas que transfieren un resto de azúcar a la ramificación 1,6 de los precursores de oligosacáridos ligados a lípidos (Figura 13). Una célula huésped de la invención se selecciona o se genomanipula introduciendo una mutación en uno o más de los genes que codifican una enzima que transfiere un resto de azúcar (por ejemplo, imanosilatos) a la ramificación 1,6 de un precursor de oligosacárido ligado con lípidos. El resto de azúcar es más preferentemente manosa, es preferentemente un resto glucosa, GlcNAc, galactosa, ácido siálico, fucosa o GlcNAc fosfato. En una realización preferente, la actividad de una o más enzimas que manosilan la ramificación 1,6 de un precursor de oligosacáridos ligado con lípidos se disminuye o se anula. El procedimiento puede comprender además la etapa de introducir en las células huésped al menos una actividad de glucosidasa (véase más adelante).

Todavía en otra realización, la invención proporciona un procedimiento para introducir una glucoproteína humana en un huésped no humano, en el que la glucoproteína comprende un N-glucano que tiene al menos dos GlcNAc unidas a una estructura nuclear de trimanosa.

En cada una de las realizaciones anteriores, el procedimiento se refiere a preparar una célula huésped en la que los precursores de oligosacáridos ligados con lípidos están enriquecidos en estructuras Man_xGlcNAc₂, donde X es 3, 4 ó 5 (Figura 14). Estas estructuras se transfieren en el RE de las células huésped sobre las cadenas de un polipéptido en formación mediante una oligosacaril-transferasa y después pueden procesarse mediante tratamiento con glucosidasas (por ejemplo, α -manosidasas) y glucosiltransferasas (por ejemplo, GnT1) para producir N-glucanos que tienen estructuras nucleares de GlcNAcMan_xGlcNAc₂, en las que X es 3, 4 ó 5, y es preferentemente 3 (Figuras 14 y 15). Como se muestra en la Figura 14, los N-glucanos que tienen una estructura nuclear de GlcNAcMan_xGlcNAc₂ donde X es más de 3 puede convertirse en GlcNAcMan₃GlcNAc₂, por ejemplo, mediante tratamiento con una actividad de α -1,3 y/o α -1,2,1,3 manosidasa, donde fuera aplicable.

El procesamiento adicional de GlcNAcMan₃GlcNAc₂ mediante tratamiento con glucosiltransferasas (por ejemplo, GnTII) produce estructuras nucleares de GlcNAc₂Man₃GlcNAc₂ que después pueden modificarse, según se desee, por ejemplo, mediante tratamiento ex vivo o mediante expresión heteróloga en las células huésped de un grupo de enzimas de glucosilación, que incluyen glucosiltransferasas, transportadores de azúcares y manosidasas (véase más adelante), para que se conviertan en N-glucanos humanoides. Las glucoproteínas humanoides preferentes que pueden producirse de acuerdo con la invención incluyen las que comprenden N-glucanos que tienen siete o menos, o tres o menos, restos de manosa; comprenden una o más azúcares, que se seleccionan a partir del grupo constituido por galactosa, GlcNAc, ácido siálico, y fucosa; y comprenden al menos una ramificación de oligosacárido que comprende la estructura NeuNAc-Gal-GlcNAc-Man.

En una realización, las células huésped presentan una actividad menor o anulada de Dol-P-Man:Man₅GlcNAc₂-PP-Dol Manosiltransferasa, que es una actividad implicada en la primera etapa de manosilación de Man₅GlcNAc₂-PP-Dol a Man₆GlcNAc₂-PP-Dol en la parte de la luz del RE (por ejemplo, *ALG3* Figura 13; Figura 14). En *S. cerevisiae*, esta enzima es codificada por el gen *ALG3*. Como se ha descrito anteriormente, las células de *S. cerevisiae* que portan una mutación *Alg3-1* acumulan Man₅GlcNAc₂-PP-Dol y las células que tienen una delección en *Alg3* parecen transferir las estructuras Man₅GlcNAc₂ sobre cadenas polipeptídicas en formación dentro del RE. Por consiguiente, en esta realización, las células huésped acumularán *N*-glucanos enriquecidos en estructuras Man₅GlcNAc₂ que después pueden convertirse en GlcNAc₂Man₃GlcNAc₂ mediante tratamiento con glucosidasas (por ejemplo, con actividades de α -1,2 manosidasa, α -1,3 manosidasa, o α -1,2-1,3 manosidasa) y glucosiltransferasa (por ejemplo, GnTI, GnTII) (Figura 14; Figura 35B).

Como se describe en el Ejemplo 10, se diseñaron cebadores degenerados basándose en una alineación de secuencias de la proteína *Alg3* de *S. cerevisiae*, *D. melanogaster* y seres humanos (*H. sapiens*) (Figuras 16 y 17), y se usaron para amplificar un producto del ADN genómico de *P. pastoris*. El producto de PCR resultante se usó como sonda para identificar y aislar un clon genómico de *P. pastoris* que comprende un marco de lectura abierto (ORF) que codifica una proteína que tiene 35% de homología global entre las secuencias y 53% de similitud entre las secuencias con el gen *ALG3* de *S. cerevisiae* (Figuras 18 y 19). Este gen de *P. pastoris* se denomina en el presente documento "*PpALG3*". El gen *ALG3* se identificó de forma similar y se aisló a partir de *K. lactis* (Ejemplo 10; Figuras 20 y 21).

Así, en el presente documento se describe una molécula de ácido nucleico aislada que tiene una secuencia de ácidos nucleicos que comprende o constituida por al menos 45, preferentemente al menos 50, más preferentemente al menos 60 y lo más preferentemente 75 o más restos nucleotídicos del gen *OCH1* del gen *ALG3* de *P. pastoris* (Figura 18) y en el presente documento se describe el gen *ALG3* de *K. lactis* (Figura 20), y sus homólogos, variantes y derivados. También se describen en el presente documento moléculas de ácido nucleico que se hibridan en condiciones rigurosas a las moléculas de ácido nucleico descritas anteriormente. De forma similar, en el presente documento se proporcionan polipéptidos aislados (que incluyen muteínas, variantes de alelos, fragmentos, derivados, y análogos) codificados por las moléculas de ácido nucleico (en las Figuras 18 y 20 se muestran los productos del gen *ALG3* de *P. pastoris* y *K. lactis* *ALG3*). Además, también se describen vectores, que incluyen vectores de expresión, que comprenden dichas moléculas de ácido nucleico, como se describe adicionalmente en el presente documento.

Usando cebadores específicos de gen, se preparó una construcción para eliminar el gen *PpALG3* del genoma de *P. pastoris* (Ejemplo 10). Esta cepa se usó para generar una célula huésped con actividad de manosiltransferasa Dol-P-Man:Man₅GlcNAc₂-PP-Dol anulada y para producir precursores de Man₅GlcNAc₂-PP-Dol ligados con lípidos que se transfieren sobre cadenas de polipéptido en desarrollo para producir *N*-glucanos que tienen una estructura de carbohidrato de Man₅GlcNAc₂.

Como se describe en el Ejemplo 11, dicha célula huésped puede ser genomanipulada mediante la expresión de manosidasas apropiadas para que produzcan *N*-glucanos que tienen la estructura de carbohidrato nuclear de Man₃GlcNAc₂ deseada. La expresión de GnTs en las células huésped de la invención (por ejemplo, dirigidas a una molécula de ácido nucleico o a un conjunto de moléculas de ácido nucleico como se describe más adelante) permite a las células huésped modificadas producir *N*-glucanos que tienen una o dos estructuras de GlcNAc unidas a cada ramificación de la estructura nuclear de Man₃ (es decir, GlcNAc₁Man₃GlcNAc₂, GlcNAc₂Man₃GlcNAc₂, o GlcNAc₃Man₃GlcNAc₂; véase la Figura 15). Estas estructuras pueden procesarse adicionalmente usando los procedimientos de la invención para producir *N*-glucanos humanoides sobre proteínas que entran en la ruta de secreción de las células huésped.

En una realización preferente, el procedimiento de la invención supone preparar o usar una célula huésped que tiene (a) una actividad menor o anulada de un gen *alg* o de una o más actividades que manosilan *N*-glucanos sobre la ramificación α -1,6 del la estructura de carbohidrato nuclear Man₃GlcNAc₂ ("Man₃"); y (b) una actividad menor o anulada una α -1,6-manosiltransferasa de iniciación, es decir, una enzima específica de iniciación que inicia la mannosilación de la cadena externa (de la ramificación α -1,3 de la estructura nuclear de Man₃). En *S. cerevisiae*, esta enzima está codificada por el gen *OCH1*. La alteración del gen *OCH1* en *S. cerevisiae* produce un fenotipo en el que los azúcares ligados en *N* carecen completamente de la cadena exterior de polimannosa. Las estrategias previas para obtener una glucosilación de tipo mamífero en cepas fúngicas ha necesitado de la inactivación de *OCH1* (véase, por ejemplo, Chiba *et al.* (1998) J. Biol. Chem. 273:26298-304). La alteración de la actividad de α -1,6-manosiltransferasa de iniciación en una célula huésped de la invención es opcional, sin embargo (dependiendo de la célula huésped seleccionada), dado que la enzima Ochlp necesita un Man₃GlcNAc₂ intacto para la iniciación eficaz de la cadena externa de manosa. Así, las células huésped seleccionada o producidas de acuerdo con esta invención, que acumulan oligosacáridos ligados con lípidos que tienen siete o menos restos de manosa, después de la transferencia, producirán *N*-glucanos hipoglucosilados que probablemente serán sustratos deficientes para Ochlp (véase, por ejemplo, Nakayama *et al.* (1997) FEBS Lett. 412(3):547-50).

Genomanipulación o selección de huéspedes que tienen actividad de *N*-acetilglucosaminiltransferasa III

La invención además proporciona un procedimiento para producir una glucoproteína humanoide en las células huésped de la invención expresando una actividad de *N*-acetilglucosaminiltransferasa III (que incluye una enzima de longitud completa, y sus homólogos, variantes, derivados, y fragmentos catalíticamente activos). En una realización, las células huésped de la invención (por ejemplo, *P. pastoris*) se genomanipula para producir más *N*-glucanos huma-

noides, por ejemplo, mediante activación de una actividad de *N*-acetilglucosaminiltransferasa III o mediante expresión a partir de una molécula de ácido nucleico de una actividad de *N*-acetilglucosaminiltransferasa III. Usando técnicas notorias en la técnica, se diseñan cebadores específicos de genes para complementar las regiones homólogas de un gen de GnTIII, preferentemente un gen de GnTIII de mamífero (por ejemplo, GnTIII de ratón) (Figura 24), cuyas secuencias están fácilmente disponibles en la técnica (por ejemplo, n.º de acceso de Genbank L39373) y se amplifican por PCR.

En una realización, la invención proporciona un procedimiento para producir una glucoproteína humanoide en una célula huésped de la invención (por ejemplo, *P. pastoris*), en la que la glucoproteína comprende un *N*-glucano que muestran una diseccionadora GlcNAc sobre una estructura nuclear de trimanosa o trimanosilo (Man₃GlcNAc₂). En esta realización, GlcNAcMan₃GlcNAc₂ (que puede producirse haciendo reaccionar un núcleo de trimanosa con la actividad de *N*-acetilglucosaminiltransferasa I ("GnTI"), pero que habitualmente se produce recortando de GlcNAcMan₅GlcNAc₂ mediante una actividad de α -1,3/ α 1,6-manosidasa, como por ejemplo manosidasa II (Hamilton *et al.* (2003) Science 301:1244-46)) se hace reaccionar con una actividad de *N*-acetilglucosaminiltransferasa III produciendo una GlcNAc₂Man₃GlcNAc₂ diseccionada. Por consiguiente, la invención proporciona actividad de GnTIII, que transfiere α -1,4 GlcNAc sobre sustratos que tienen capacidad de aceptar la GlcNAc diseccionadora en eucariotas inferiores.

En otra realización, la invención proporciona un procedimiento para producir una glucoproteína humanoide en una célula huésped de la invención (por ejemplo, *P. pastoris*), en la que la glucoproteína comprende un *N*-glucano que muestran una diseccionadora GlcNAc sobre una estructura nuclear de trimanosa o trimanosilo (Man₃GlcNAc₂) que tiene al menos dos GlcNAc unidas al núcleo de trimanosa. En esta realización, Man₃GlcNAc₂ se hace reaccionar con una actividad de GnTI y después con una actividad de *N*-acetilglucosaminiltransferasa II ("GnTII") y una actividad de GnTIII (en cualquier orden) produciendo una GlcNAc₃Man₃GlcNAc₂ diseccionada (Figura 38). Debería entenderse que la estructura nuclear de trimanosilo de esta realización también puede contener un grupo manosilo adicional en lugar de un resto GlcNAc. Por ejemplo, GlcNAcMan₄GlcNAc₂ puede hacerse reaccionar con una actividad de GnTIII para producir una GlcNAc₂Man₄GlcNAc₂ diseccionada.

La invención también proporciona un procedimiento para producir una glucoproteína más humanoide en una célula huésped de la invención (por ejemplo *P. pastoris*), en la que la glucoproteína producida comprende un *N*-glucano que tiene al menos dos GlcNAc unidas a una estructura nuclear de pentamanosa (Man₅GlcNAc₂) y que exhibe un *N*-glucano diseccionado. Por consiguiente, en esta realización, se hace reaccionar una estructura nuclear de pentamanosa (Man₅GlcNAc₂) con una actividad de GnTIII para producir una estructura diseccionada de GlcNAcMan₅GlcNAc₂ y GlcNAc₂Man₅GlcNAc₂.

En una realización alternativa, se hace reaccionar una estructura nuclear de pentamanosa producida mediante la mutación de los genes *och1* y *Alg3* con actividades de α 1,2-manosidasa, GnTI, GnTII y GnTIII y UDP-GlcNAc para producir un glucano GlcNAc₂Man₃GlcNAc₂ diseccionado (Figura 35B). En otra realización, se hace reaccionar una estructura nuclear de pentamanosa con actividades de GnTI y de GnTIII (en cualquier orden o combinadas) para producir una estructura de GlcNAc₂Man₅GlcNAc₂ diseccionada (Figura 37).

Como se describe en el presente documento, usando el procedimiento de la adenoteca de combinaciones, como se describe más adelante, se expresa una construcción pVA53 que comprende la secuencia líder *MNN2*(s) de *S. cerevisiae* (n.º de acceso de Genbank NP_009571) fusionada a un dominio GnTIII de ratón catalíticamente activo (GnTIII A32) en la cepa YSH-1 de *P. pastoris* (Ejemplo 13) produciendo así *N*-glucanos que tienen una estructura GlcNAc₂Man₅GlcNAc₂ diseccionada (Ejemplo 20). La Figura 26 (abajo) presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en la cepa mencionada anteriormente, que se denomina PBP26 (Figura 36), y muestra un pico predominante en 1666 m/z [a], que corresponde a GlcNAc₂Man₅GlcNAc₂ diseccionado. (Para comparar, la Figura 26 (arriba) presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en la cepa YSH-1 que carece de la construcción pVA53. El pico predominante en 1461 m/z [d] corresponde al glucano no modificado: GlcNAcMmsGlcNAc₂). Por consiguiente, en una realización, un huésped de la presente invención se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que exhiben al menos 50% molar de una estructura de GlcNAc₂Man₅GlcNAc₂ o al menos 50% molar de una GlcNAc₂Man₃GlcNAc₂ que tiene una GlcNAc diseccionadora. El porcentaje molar de los glucanos hace referencia al porcentaje de los glucanos neutros totales detectados por MALDI-TOF. Se entiende que si, por ejemplo, GlcNAc₂Man₃GlcNAc₂ que tiene una GlcNAc diseccionadora se produce al 20% y GlcNAc₃Man₃GlcNAc₂ se produce al 25% en una proteína diana, la cantidad total del GlcNAc₂Man₃GlcNAc₂ producido de forma transitoria que tiene una GlcNAc diseccionadora es 45%, porque GlcNAc₃Man₃GlcNAc₂ es un producto de un GlcNAc₂Man₃GlcNAc₂ que tiene una GlcNAc diseccionadora que se hace reaccionar después con GnTII.

De forma similar, como también se describe en el presente documento, una construcción pVA55 que comprende la secuencia líder *MNN2*(1) de *S. cerevisiae* (n.º de acceso de Genbank NP_009571) fusionada a un dominio GnTIII de ratón catalíticamente activo (GnTIII A32) se expresa en una cepa de *P. pastoris* (YSH-1), produciendo así una estructura de *N*-glucanos GlcNAcMan₅GlcNAc₂ y *N*-glucanos GlcNAc₂Man₅GlcNAc₂ diseccionados. Como se muestra en la Figura 27 (abajo), estas estructuras corresponden a los picos en 1463 m/z y 1667 m/z, respectivamente. (Para comparar, la Figura 27 (arriba) presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en la cepa YSH-1 que carece de la construcción pVA53. El pico predominante corresponde a GlcNAcMan₅GlcNAc₂ sin modificar en 1461 m/z [d]). Por consiguiente, en otra realización, un huésped de la presente invención se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que exhiben al menos

20% molar de una estructura $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ o al menos 20% molar de una $\text{GlcNAc}_2\text{Man}_5\text{GlcNAc}_2$ que tiene una GlcNAc diseccionadora.

Como también se describe en el presente documento, una construcción pVA53 que comprende la secuencia líder *MNN2(s)* de *S. cerevisiae* (N.º de acceso de GenBank NP_009571) fusionada a un dominio GnTIII de ratón catalíticamente activo (GnTIII A32) se expresa en una cepa YSH-44 de *P. pastoris* (Ejemplo 15) produciendo así *N*-glucanos que tienen una estructura $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ diseccionada (Ejemplo 20). La Figura 30 presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en la cepa mencionada anteriormente, que se denomina YSH-57, y muestra un pico predominante en 1542 m/z [y], que corresponde a $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ diseccionado. (Para comparar, la Figura 29 presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en la cepa YSH-44 que carece de la construcción pVA53. El pico predominante en 1356 m/z [x] en la Figura 29 corresponde al glucano no modificado: $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$.) Por consiguiente, en una realización, un huésped de la presente invención se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que exhiben al menos 80% molar de una estructura $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ que tiene una GlcNAc diseccionadora. El porcentaje molar de los glucanos hace referencia al porcentaje de los glucanos neutros totales detectados por MALDI-TOF.

De forma alternativa, como se describe en el presente documento, una construcción pVA53 que comprende la secuencia líder *MNN2(s)* de *S. cerevisiae* (N.º de acceso de GenBank NP_009571) fusionada a un dominio GnTIII de ratón catalíticamente activo (GnTIII Δ32) se expresa en una cepa de *P. pastoris* (PBP6-5) (Ejemplo 11) produciendo así *N*-glucanos que tienen una estructura $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ y una $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ diseccionada. Como se muestra en la Figura 32, estas estructuras corresponden a picos a 1340 m/z y 1543 m/z, respectivamente. Por consiguiente, en otra realización, un huésped de la presente invención se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que exhiben al menos 20% molar de una estructura $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ que tiene una GlcNAc diseccionadora en una célula huésped mutante para *Alg3*.

La invención proporciona procedimientos para producir una glucoproteína humanoide en una célula huésped de la invención, en la que la glucoproteína comprende una estructura nuclear de $\text{Man}_5\text{GlcNAc}_2$ o una estructura nuclear de $\text{Man}_3\text{GlcNAc}_2$, y en la que la estructura nuclear se modifica además mediante dos o más GlcNAc. En algunas realizaciones de la invención, 10% o más de las estructuras nucleares son modificadas por las dos o más GlcNAc. En otras realizaciones preferentes, se modifican 20%, 30%, 40%, 50%, 60%, 70%, 80% o incluso más de las estructuras nucleares. En una realización muy preferente, una de las GlcNAc es una GlcNAc diseccionadora.

En otro aspecto de la invención, se usa una adenoteca de combinaciones que codifica al menos un dominio catalítico GnTIII para expresar una actividad de GnTIII en una célula huésped de la invención (Ejemplo 18). Preferentemente, una adenoteca de la invención comprende una subadenoteca de secuencias líder fusionadas en marco a una única molécula de ácido nucleico o a una subadenoteca de moléculas de ácido nucleico que comprenden secuencias de GnTIII, una o más de las cuales codifica un dominio catalítico que tiene actividad de GnTIII en las células huésped. De forma alternativa, una única molécula de ácido nucleico o una subadenoteca de moléculas de ácido nucleico que comprenden secuencias líder están fusionadas en marco a una subadenoteca de moléculas de ácido nucleico que comprenden secuencias de GnTIII, una o más de las cuales codifica un dominio catalítico que tiene actividad de GnTIII en las células huésped de la invención. (Véase más adelante). La expresión de estas y otras adenotecas de combinaciones se realiza en una célula huésped que expresa una glucoproteína diana cuyas estructuras de *N*-glucano se analizan para determinar si y cuanta GnTIII se expresa. Puede producirse una amplia variedad de enzimas GnTIII catalíticamente activas en una célula huésped usando los procedimientos y adenotecas de la invención. Es este aspecto de la invención el que permite a los expertos crear y elegir entre enzimas GnTIII que tienen poca o ninguna actividad y las enzimas que se expresan activamente y que producen niveles predominantes de un intermedio de oligosacárido diseccionado deseado como por ejemplo $\text{GlcNAc}_2\text{Man}_5\text{GlcNAc}_2$, $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ o $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ en las células huésped.

Como se describe en más detalle más adelante, el direccionamiento apropiado de una enzima responsable de una etapa dada en la ruta de glucosilación a la localización subcelular apropiada y la suficiencia de la actividad de la enzima al pH particular de esa localización subcelular son factores importantes en la producción de glucoproteínas que tienen *N*-glucanos con las estructuras deseadas. El uso de adenotecas de combinaciones de proteínas de fusión para generar diversas poblaciones de quimeras enzimáticas y el cribado de estas adenotecas en células transformadas proporciona un procedimiento poderoso para identificar cepas huésped con la actividad de interés en la localización apropiada. En realizaciones preferentes de la invención, la actividad enzimática está localizada de tal forma que un *N*-glucano que contiene glucoproteína expresada en la célula es capaz de reaccionar con la actividad durante el procedimiento de secreción.

Sin embargo, no todas las combinaciones de secuencias líder y dominios catalíticos producen las actividades enzimáticas deseadas. Se crea una amplia variedad de combinaciones de secuencias líder y dominios catalíticos, y sólo unas pocas de ellas pueden ser útiles para producir los intermedios actualmente deseados. La presente invención, sin embargo, engloba incluso aquellas combinaciones que actualmente no exhiben una actividad enzimática deseada en las células huésped ejemplificadas. La Figura 28 (abajo) muestra una construcción pVB51 que comprende la secuencia líder de *K. lactis* *GNT(s)* (n.º de acceso de GenBank AF106080) fusionada a un dominio GnTIII de ratón catalíticamente activo (GnTIII, 1.32) expresada en una cepa YSH-1 de *P. pastoris*, que no muestra fácilmente actividad de GnTIII. (Para comparar, la Figura 28 (arriba) presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir

de una proteína kringle 3 expresada en la cepa YSH-1 que carece de la construcción pVA53. El pico predominante corresponde a GlcNAcMan₅GlcNAc₂ sin modificar en 1461 (*m/z*). El pico predominante que se observa en la Figura 28 (abajo) en 1463 *m/z*, se correlaciona con la masa de GlcNAcMan₅GlcNAc₂.

También se observa un segundo pico en 1726 *m/z*, que no se correlaciona con la masa de GlcNAc₂Man₅GlcNAc₂. Se contempla que estas y otras combinaciones de ese tipo pueden ser útiles, con o sin ligeras modificaciones usando técnicas notorias en la técnica, cuando se expresan, por ejemplo, en otras células huésped que incluyen las que han sido modificadas para producir glucoformas humanoides.

El uso de adenotecas de combinaciones para generar diversas poblaciones de quimeras enzimáticas y el cribado de estas adenotecas en células transformadas permite además identificar cepas en las que la actividad enzimática es sustancialmente intracelular. El Ejemplo 6, más adelante, proporciona un ejemplo de condiciones de ensayo útiles para medir la actividad de α -1,2-manosidasa extracelular. Los Ejemplos 22 y 23 también proporcionan ejemplos de ensayos para la actividad de glucosiltransferasa (GnTIII) en el medio. Véase también la Tabla 9, más adelante, y Choi *et al.* (2003) Proc. Natl. Acad. Sci. U.S.A. 100(9):5022-27. Para los fines de la invención, una actividad enzimática es sustancialmente intracelular cuando menos de 10% de la actividad enzimática puede cuantificarse en el medio extracelular.

Como se describe en los Ejemplos 11, 12, 13, 14, 15, y 19-21, una célula huésped puede genomanipularse mediante la expresión de glucosiltransferasas apropiadas (por ejemplo, *N*-acetilglucosaminiltransferasa) para producir *N*-glucanos que tienen las estructuras de carbohidratos deseadas (por ejemplo, GlcNAc₂Man₃GlcNAc₂, GlcNAc₃Man₃GlcNAc₂). La expresión de GnTs en las células huésped (por ejemplo, dirigiendo una molécula de ácido nucleico o una adenoteca de moléculas de ácido nucleico como se describe más adelante y en Choi *et al.* (2003) Proc. Natl. Acad. Sci. U.S.A. 100(9):5022-27 y el documento WO 02100879) permite a las células huésped modificadas producir *N*-glucanos que tiene la GlcNAc diseccionadora en mitad de la manosa. Estas estructuras pueden procesarse adicionalmente usando los procedimientos de la invención para producir *N*-glucanos humanoides sobre proteínas que entran en la ruta de secreción de las células huésped.

En una realización más preferente, la coexpresión de transportador(es) de UDP-azúcar apropiados y transferasa(s) harán de caperuza de los restos α -1,6 y α -1,3 terminales así como la manosa media con GlcNAc, que resultaría en el precursor del complejo de tipo mamífero (por ejemplo GlcNAcMan₃GlcNAc₂) y *N*-glucosilación híbrida. Estas cadenas de oligosacáridos ligadas en N unidas a péptidos sirven como precursor para la modificación adicional a una estructura de oligosacárido de tipo mamífero. La expresión posterior de galactosil-transferasas y la genomanipulación de la capacidad de transferir ácido sialílico a los extremos (véase la Figura 1B) producirá una estructura de *N*-glucano de tipo mamífero (por ejemplo, humanoide).

Genomanipulación o selección de huéspedes que tienen actividad de *N*-acetilglucosaminiltransferasa IV, V, VI o IX

La presente invención proporciona huéspedes fúngicos unicelulares o multicelulares que tienen actividad de *N*-acetilglucosaminiltransferasa que cataliza la formación de un enlace glucosídico GlcNAc β 1,4 o GlcNAc β 1,6 la ramificación Man_a1,6 y/o en la ramificación Man_a1,6 de un sustrato oligosacárido (por ejemplo, GlcNAc₂Man₃GlcNAc₂) en presencia de un azúcar de nucleótido UDP-GlcNAc. La transferencia de los restos GlcNAc generalmente es preferente en presencia de un transportador de UDP-GlcNAc. La presente invención proporciona moléculas recombinantes de ácido nucleico que codifican proteínas que tienen actividad de *N*-acetilglucosaminiltransferasa y procedimientos para expresar la enzima activa en las rutas secretoras de las levaduras. Además, la presente invención proporciona estructuras de oligosacárido producidas a partir de los huéspedes transformados que son útiles para la administración terapéutica. Al catalizar la transferencia del azúcar GlcNAc de UDP-GlcNAc a los sustratos oligosacáridos mediante una actividad de *N*-acetilglucosaminiltransferasa, se forman glucoformas multiantenarias en una proteína, que después se extienden mediante galactosiltransferasa y sialiltransferasas.

La invención además proporciona un procedimiento para producir una glucoproteína humanoide en una célula huésped de un eucariota inferior expresando un ácido nucleico que codifica la GnTIVB humana (Δ 104-53) dominio catalítico de *N*-acetilglucosaminiltransferasa IV exógeno fusionado a los nucleótidos 1-108 del gen *MNN2* de *S. cerevisiae* junto con actividades de *N*-acetilglucosaminiltransferasa V, VI o IX (que incluye una enzima de longitud completa, sus homólogos, variantes, derivados, y fragmentos catalíticamente activos). En una realización, una célula huésped de la invención (por ejemplo, *P. pastoris*) es genomanipulada para que produzca más *N*-glucanos humanoides, por ejemplo, por activación de dichas actividades de *N*-acetilglucosaminiltransferasa IV, V, VI, IX o por expresión a partir de una molécula de ácido nucleico de la invención que codifica actividades de *N*-acetilglucosaminiltransferasa IV, V, VI, IX de la invención (Figura 39). Usando técnicas notorias en la técnica, se diseñan cebadores específicos de gen para hibridarse con regiones homólogas de miembros de la familia de las glucosiltransferasas, como por ejemplo las secuencias génicas de GnTIV, V, VI, IX, que son fácilmente accesibles en la técnica (por ejemplo, las bases de datos GenBank, SWISSProt) y son se amplifican por PCR.

Expresión de *N*-acetilglucosaminiltransferasa IV

De acuerdo con la invención, una célula huésped de hongo unicelular o multicelular se transforma con una secuencia de nucleótidos que codifican el dominio catalítico de *N*-acetilglucosaminiltransferasa IV humana exógeno GnT IVB (Δ 104-53) fusionado a los nucleótidos 1-108 del gen *MNN2* de *S. cerevisiae* que cataliza la adición

de un resto de azúcar $\beta(1,4)$ *N*-acetilglucosamina ("GlcNAc") sobre la ramificación $\text{Man}\alpha 1,3$ del GlcNAc $\beta 1,2$ - $\text{Man}\alpha 1,6$ (GlcNAc $\beta 1,2$ $\text{Man}\alpha 1,3$) $\text{Man}\beta 1,4$ - GlcNAc $\beta 1,4$ - GlcNAc $\beta 1,4$ - Asn de un sustrato oligosacárido. La adición de un GlcNAc $\beta 1,4$ mediante GnTIV sobre la ramificación $\text{Man}\alpha 1,3$ del sustrato receptor (por ejemplo GlcNAc₂Man₃GlcNAc₂) proporciona lo que se denomina un *N*-glucano triantenario.

En una realización, la invención proporciona un procedimiento para producir una glucoproteína humanoide en un eucariota inferior (por ejemplo, *P. pastoris*), en la que la glucoproteína comprende una estructura de *N*-glucano triantenario sobre una estructura de oligosacárido (por ejemplo, GlcNAc₃Man₃GlcNAc₂). En esta realización, el sustrato oligosacárido GlcNAcMan₅GlcNAc₂ (Choi *et al.*, (2003) Proc Natl Acad Sci USA, 29 de abril de 2003; 100(9):5022-7) es recortado por una actividad de α -1,3/ α -1,6-manosidasa, como por ejemplo manosidasa II (Hamilton *et al.* (2003) Science 301:1244 produciendo el sustrato GlcNAc₂Man₃GlcNAc₂, que a su vez se hace reaccionar con dicha actividad de *N*-acetilglucosaminiltransferasa IV para producir una estructura triantenaria: GlcNAc₃Man₃GlcNAc₂ (Figura 39). En las Figuras 41 y 42 se presentan dos isozimas A y B de GnTIV. En una realización preferente, un fragmento de un gen que codifica GnTIV humana (Figura 42) ligada en marco a los nucleótidos 1-108 de la secuencia del péptido director MNN2(s) de *S. cerevisiae* se introduce y se expresa en YSH-44 de *P. pastoris* (Ejemplo 15). Así, en ciertas realizaciones, una célula huésped de la presente invención se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que exhiben preferentemente al menos 50, 60, 70, 80, 90% molar o más de la estructura de *N*-glucano deseada GlcNAc₃Man₃GlcNAc₂. En una realización todavía más preferente, el sustrato oligosacárido GlcNAc₃Man₃GlcNAc₂ es un sustrato para una reacción de galactosiltransferencia. Por consiguiente, la invención proporciona una actividad de GnTIV, que cataliza la transferencia de restos GlcNAc sobre sustratos oligosacáridos que forman un enlace GlcNAc β -1,4 glucosídico sobre la ramificación $\text{Man}\alpha 1,3$ de un sustrato oligosacárido (por ejemplo, GlcNAc₂Man₃GlcNAc₂) en eucariotas inferiores.

En una realización más preferente, usando una adenoteca de combinaciones y procedimientos de la invención, se expresa la construcción pPB144 (Figura 40A) que comprende la secuencia líder MNN2(s) de *S. cerevisiae* (n.º de acceso de Genbank NP_009571) fusionada a un dominio GnTIV catalíticamente activo de (GnTIVB Δ 104 se expresa en una cepa YSH-44 de *P. pastoris* (Ejemplo 15) produciendo así estructuras de *N*-glucano triantenario GlcNAc₃Man₃GlcNAc₂ (Ejemplo 25). La Figura 47 presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en una cepa transformada denominada PBP43, que exhibe un pico predominante en 1543 m/z [y], que corresponde a una estructura de *N*-glucano triantenario GlcNAc₃Man₃GlcNAc₂. (Véase la Figura 29 para comparar, que presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en células FSH-44 produciendo GlcNAc₂Man₃GlcNAc₂.) Por consiguiente, en ciertas realizaciones, una célula huésped de la presente invención se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que muestran preferentemente al menos 50% molar de una estructura triantenaria GlcNAc₃Man₃GlcNAc₂. En otra realización, las células huésped de la presente invención producen la estructura triantenaria GlcNAc₃Man₃GlcNAc₂ catalizada por GnTIV en una cantidad que es al menos 60, 70, 80, más preferentemente 90% molar o superior. Aquí y en todas partes, el porcentaje molar de los glucanos hace referencia al porcentaje de los glucanos neutros totales detectados por MALDI-TOF MS.

40 Expresión de *N*-acetilglucosaminiltransferasa IV y V

En otro aspecto de la invención, una célula huésped de hongo unicelular o multicelular que es genomaniplada o seleccionada para producir un sustrato oligosacárido triantenario (por ejemplo, GlcNAc₃Man₃GlcNAc₂), se transforma con un ácido nucleico que codifica una actividad de *N*-acetilglucosaminiltransferasa V ("GnTV") que cataliza la adición de un resto de azúcar $\beta(1,6)$ *N*-acetilglucosamina ("GlcNAc") sobre la ramificación $\text{Man}\alpha 1,6$ del GlcNAc $\beta 1,2$ - $\text{Man}\alpha 1,6$ (GlcNAc $\beta 1,4$ (GlcNAc $\beta 1,2$) $\text{Man}\alpha 1,3$) $\text{Man}\beta 1,4$ - GlcNAc $\beta 1,4$ - GlcNAc $\beta 1,4$ - Asn de un sustrato oligosacárido. La adición de un GlcNAc $\beta 1,6$ por GnTV sobre el sustrato receptor (por ejemplo GlcNAc₃Man₃GlcNAc₂) en presencia de restos GlcNAc proporciona un *N*-glucano: tetraantenario GlcNAc₄Man₃GlcNAc₂. Preferentemente, la actividad de GnTV de la presente invención se expresa en células huésped de hongos unicelulares o multicelulares produciendo glucanos triantenarios, por ejemplo en *P. pastoris* PBP43 en los que la actividad de GnTV cataliza la transferencia de un resto GlcNAc sobre la ramificación $\text{Man}\alpha 1,6$ del sustrato oligosacárido GlcNAc₃Man₃GlcNAc₂ formando un enlace GlcNAc β 1,6 glucosídico.

En una realización, usando el procedimiento de adenoteca de combinaciones de la invención, *P. pastoris* YSH-44 se transforma con las construcciones fusionadas GnTIVB/MNN2(s) de *S. cerevisiae* y GnTVI/MNN2(s) de *S. cerevisiae*. La Figura 49 presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en una cepa transformada denominada PBP46, que muestra un pico predominante en 1747 m/z [z], que corresponde a estructura de *N*-glucano tetraantenaria GlcNAc₄Man₃GlcNAc₂ y un pico residual en 1543 m/z [y], que corresponde a estructura de *N*-glucano tetraantenaria GlcNAcMan₃GlcNAc₂.

En otro ejemplo, usando el procedimiento de adenoteca de combinaciones de la invención, *P. pastoris* YSH-44 se transforma con una combinación diferente de fusión de enzima y líder: el plásmido pPB128 que contiene GnTIVA (Δ 82)/MNN2(s) de *S. cerevisiae* y plásmido pPB140 que contiene construcciones fusionadas de GnTV (Δ 145)/MNN2(s) de *S. cerevisiae*. La Figura 50 presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en una cepa transformada denominada PBP94, que muestra a pico predominante en 1743 m/z [z], que corresponde a la estructura de *N*-glucano tetraantenaria GlcNAc₄Man₃GlcNAc₂.

Por consiguiente, en ciertas realizaciones, la presente invención proporciona una célula huésped que se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que exhiben preferentemente al menos 50, 60, 70, 80, 90% molar o más de la estructura de *N*-glucano deseada $\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$. En una realización más preferente, al expresar las actividades de GnTIV y GnTV de la invención, las células huésped producen la estructura de *N*-glucano multiantenaria deseada.

No todas las combinaciones de secuencias líder y dominios catalíticos producen las actividades enzimáticas deseadas en la producción de estructuras de glucanos multiantenarias. Se crea una amplia variedad de combinaciones de secuencias líder y dominios catalíticos, usando los procedimientos y adenotecas de la invención, y sólo unas pocas de ellas pueden ser útiles para producir los intermedios actualmente deseados. La presente invención, sin embargo, engloba incluso aquellas combinaciones que actualmente no exhiben una actividad enzimática deseada en las células huésped ejemplificadas. Se contempla que estas y otras combinaciones de ese tipo pueden ser útiles, con o sin ligeras modificaciones usando técnicas notorias en la técnica, cuando se expresan, por ejemplo, en otras células huésped que incluyen las que han sido modificadas para producir glucoformas humanoides.

Expresión de *N*-acetilglucosaminiltransferasa V

En otro ejemplo que se describe en el presente documento, un ácido nucleico que codifica actividad de GnTV se expresa en células huésped produciendo las estructuras nucleares $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ que consiguen la formación de estructuras triantenarias $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$. La secuencia de GnTV conocida se describe en la Figura 43. En un ejemplo, usando el procedimiento de adenoteca de combinaciones que se describe en el presente documento, se expresa la construcción pPB140 (Figura 40B) que comprende la secuencia líder *MNN2(s)* de *S. cerevisiae* (n.º de acceso de Genbank NP_009571) fusionada a un dominio GnTV de ratón catalíticamente activo (GnTV 8145) en una cepa YSH44 de *P. pastoris* (Ejemplo 15) produciendo así las estructuras del *N*-glucano triantenario $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (Ejemplo 25). La Figura 48 presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en un cepa transformada denominada PBP32, que muestra un pico predominante en 1559 m/z [V], que corresponde a una estructura de *N*-glucano tetraantenaria, y un segundo pico en 1355 m/z [u] que corresponde a la estructura de *N*-glucano tetraantenaria $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$. (Véase la Figura 29 para comparar, que presenta el espectro por MALDI-TOF de *N*-glucanos liberados a partir de una proteína kringle 3 expresada en células YSH-44 produciendo $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$.) Por consiguiente, dicho huésped se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que muestran al menos 40% molar de la estructura triantenaria: $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$. Preferentemente, dicho huésped produce al menos 50, 60, 70, 80, 90% molar o más de la estructura triantenaria catalizada por GnTV.

Expresión de *N*-Acetilglucosaminiltransferasa VI

La invención también proporciona una célula huésped de hongo unicelular o multicelular transformado con un ácido nucleico que codifica la actividad de *N*-acetilglucosaminiltransferasa VI ("GnTVI"), que cataliza la transferencia de un resto de azúcar $\beta(1,4)$ *N*-acetilglucosamina sobre la ramificación $\text{Man}\alpha 1,6$ de la $\text{GlcNAc } \beta 1,4$ ($\text{GlcNAc } \beta 1,2$) $\text{Man}\alpha 1,6$ ($\text{GlcNAc } \beta 1,4$ ($\text{GlcNAc } \beta 1,2$) $\text{Man}\alpha 1,3$) $\text{Man } \beta 1,4$ - $\text{GlcNAc } \beta 1,4$ - $\text{GlcNAc } \beta 1,4$ - Asn de un sustrato oligosacárido. La adición de una $\text{GlcNAc } \beta 1,4$ por GnTVI a un sustrato receptor (por ejemplo $\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$) proporciona lo que se denomina *N*-glucano pentaantenario. La adición del resto GlcNAc forma un enlace glucosídico $\text{GlcNAc } \beta 1,4$ en el sustrato oligosacárido. En una realización, se expresa un ácido nucleico que codifica actividad de GnTVI en una célula huésped produciendo *N*-glucanos pentaantenarios como por ejemplo $\text{GlcNAc}_5\text{Man}_3\text{GlcNAc}_2$. En otra realización, usando un fragmento de ADN que codifica la actividad de GnTVI como por ejemplo la que se describe en la Figura 44, se expresa un construcción de plásmido que comprende la secuencia líder *MNN2(s)* de *S. cerevisiae* (n.º de acceso de Genbank NP_009571) fusionado a un dominio GnTVI catalíticamente activo de *Gallus gallus* en una cepa YSH-44 de *P. pastoris* produciendo así estructuras de *N*-glucano pentaantenarias $\text{GlcNAc}_5\text{Man}_3\text{GlcNAc}_2$. Por consiguiente, una célula huésped de la invención se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que producen *N*-glucanos pentaantenarios en un resto detectable.

Expresión de *N*-acetilglucosaminiltransferasa IX

En otra realización, el ácido nucleico que codifica la actividad de GnTIX (por ejemplo, Genbank AN NP_945193) se expresa en una célula huésped de la invención en la que la actividad de GnTIX cataliza la transferencia de restos de GlcNAc sobre los sustratos receptores de glucanos complejos (por ejemplo, $\text{GlcNAc}_2\text{Man}_3\text{NAc}_2$) en ausencia de GnTIV, GnTV o GnTVI. El ácido nucleico que codifica la actividad de GnTIX, que normalmente parece que se expresa exclusivamente en el cerebro, se ha demostrado que cataliza la síntesis de un oligosacárido único ligado en N en células CHO mutantes (Raju *et al.*, (1998) J. Biol. Chem 273, 14090-14098). La expresión de una GnTIX humana recombinante demostró actividad de GnTV, catalizando la transferencia de GlcNAc a la posición 6-OH de la ramificación de manosa ligada en α -1,6 del oligosacárido $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ -PA mediante un enlace $\beta 1,6$ además de actuar sobre la ramificación de manosa ligada en α -1,3 (J Biol Chem. 31 de octubre de 2003; 278(44):43102-9). La GnTIX es capaz de catalizar la transferencia de GlcNAc a la posición 6-OH de la manosa en la secuencia $\text{GlcNAc } \beta 1,2$ - $\text{Man}\alpha 1$.

Por consiguiente, la presente invención proporciona un procedimiento de producir glucano estructuras tetraantenarias en *P. pastoris* usando un ácido nucleico que codifica la actividad de GnTIX (Figura 45). Preferentemente, la introducción y expresión de la actividad de GnTIX cataliza la transferencia de restos GlcNAc sobre el sustrato receptor

GlcNAc₂Man₃GlcNAc₂ produciendo GlcNAc₄Man₃GlcNAc₂ y GlcNAc₃Man₃GlcNAc₂. La expresión por las células huésped de la actividad de GnTIX cataliza la transferencia de restos GlcNAc a la posición 6-OH preferentemente tanto sobre las ramificaciones Man α 1,3 y Man α 1,6 del sustrato oligosacárido nuclear GlcNAc₂Man₃GlcNAc₂ produciendo la estructura tetraantenaria GlcNAc₄Man₃GlcNAc₂. En una realización, una célula huésped que produce el sustrato receptor GlcNAc₂Man₃GlcNAc₂, como por ejemplo *P. pastoris* YSH-44, se transforma con un plásmido (como por ejemplo pPB176) que comprende un gen que codifica la actividad de GnTIX (Δ 43) fusionado en marco a la secuencia líder MNN2(s) de *S. cerevisiae* (Ejemplo 29). La célula huésped se caracteriza por su capacidad de producir, al menos de forma transitoria, *N*-glucanos que muestran una actividad al menos 5% molar o preferentemente superior de GnTIX que cataliza la transferencia de GlcNAc a la posición 6-OH de una ramificación de manosa ligada en α -1,6 del oligosacárido GlcNAc₂Man₃GlcNAc₂ mediante el enlace α -1,6 además de actuar sobre la ramificación de manosa ligada en α 1,3 del sustrato oligosacárido.

Además, el ácido nucleico que codifica la actividad de GnTIX puede tener codones optimizados para la eficacia de traducción en levaduras usando procedimientos notorios. La Figura 46 muestra el fragmento de ADN con codones optimizados que codifica parte de la GnTIX humana que carece del dominio TM (Δ 43) sintetizada a partir de oligonucleótidos usando PCR.

Variaciones de múltiples glucanos antenarios

La presente invención está adaptada para producir una variedad de glucanos antenarios múltiples que son adecuados para fines terapéuticos. Se contempla que los glucanos que tienen los mismos enlaces GlcNAc β pueden aumentar la semivida de las proteínas. Por consiguiente, la invención proporciona una célula huésped de un eucariota inferior que comprende *N*-glucanos que tienen al menos dos restos GlcNAc o bien en la ramificación Man α 1,3 o en Man α 1,6 del intermedio de oligosacárido del núcleo de trimanosa (por ejemplo, Man₃GlcNAc₂). En una realización, las células huésped de eucariotas inferiores comprenden al menos dos restos GlcNAc β 1,4 en la ramificación Man α 1,3 y Man α 1,6 del intermedio de oligosacárido con núcleo de trimanosa (por ejemplo, Man₃GlcNAc₂). En otra realización, las células huésped de eucariotas inferiores comprenden al menos dos restos GlcNAc β 1,6 en la ramificación Man α 1,3 y Man α 1,6 del intermedio de oligosacárido con núcleo de trimanosa (por ejemplo, Man₃GlcNAc₂). Todavía en otra realización, las células huésped de eucariotas inferiores comprenden al menos dos restos GlcNAc β 1,2 en la ramificación Man α 1,3 y Man α 1,6 del intermedio de oligosacárido con núcleo de trimanosa (por ejemplo, Man₃GlcNAc₂).

Como se indica más adelante, aunque las células huésped de hongos unicelulares o multicelulares son huéspedes preferentes para producir proteínas terapéuticas usando los procedimientos de la invención, la presente invención también es útil para modificar los perfiles de *N*-glucano de glucoproteínas formadas en cualquier célula huésped eucariota, preferentemente en una célula huésped no humana (por ejemplo, mamífero).

Células huésped de la invención

Una célula huésped preferente de la invención es una célula de hongo unicelular o multicelular, por ejemplo, levaduras, de un hongo filamentosos unicelular y multicelular.

Las células fúngicas unicelulares y multicelulares que pueden producir glucoproteínas que tienen el *N*-glucano Man₅GlcNAc₂ unido son de particular utilidad dado que (a) carecen de un alto grado de manosilación (por ejemplo más de 8 manosas por *N*-glucano, o especialmente 30-40 manosas), muestran una capacidad inmunógena reducida en los seres humanos; y (b) el *N*-glucano es un sustrato para reacciones de glucosilación posteriores formando una glucoforma incluso más humanoide, por ejemplo mediante la acción de GlcNAc transferasa I (Figura 1B; β 1,2 GnTI) formando GlcNAcMan₅GlcNAc₂. Se obtiene un rendimiento de más de 30% molar, más preferentemente un rendimiento de 50, 60, 70, 80, 90, o incluso 100% molar, de glucoproteínas con *N*-glucanos que tienen una estructura Man₅GlcNAc₂. En una realización preferente, se demuestra que más del 50% de la estructura Man₅GlcNAc₂ es sustrato para una actividad de GnTI y puede servir como tal sustrato *in vivo*.

Los eucariotas inferiores preferentes de la invención incluyen pero sin limitación: *Pichia pastoris*, *Pichia finlandica*, *Pichia trehalophila*, *Pichia koclamae*, *Pichia membranaefaciens*, *Pichia minutia* (por ejemplo, *Ogataea minute*, *Pichia lindneri*), *Pichia opuntiae*, *Pichia thermotolerans*, *Pichia salictaria*, *Pichia guercuum*, *Pichia pijperi*, *Pichia stiptis*, *Pichia methanolica*, *Pichia sp.*, *Saccharomyces cerevisiae*, *Saccharomyces sp.*, *Hansenula polymorpha*, *Kluyveromyces sp.*, *Candida albicans*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei*, *Chrysosporium lucknowense*, *Fusarium sp.*, *Fusarium gramineum*, *Fusarium venenatum*, y *Neurospora crassa*.

En cada una de las realizaciones anteriores, el procedimiento se refiere a preparar una célula huésped en la que los precursores de oligosacáridos están enriquecidos en Man₅GlcNAc₂. Estas estructuras son deseables porque después pueden procesarse mediante tratamiento *in vitro*, por ejemplo, usando el procedimiento de Maras y Contreras, patente de Estados Unidos n.º 5.834.251. En una realización preferente, sin embargo, los precursores enriquecidos en Man₅GlcNAc₂ se procesan mediante al menos una reacción de glucosilación adicional *in vivo* con glucosidasas (por ejemplo, α -manosidasas) y glucosiltransferasas (por ejemplo, GnTI) - para producir *N*-glucanos humanoides. Los precursor de oligosacáridos enriquecidos en Man₅GlcNAc₂, por ejemplo, preferentemente se procesan para obtener estructuras nucleares que tienen GlcNAcMan_xGlcNAc₂, en las que X es 3, 4 ó 5, y es preferentemente 3. Los

N-glucanos que tienen una estructura nuclear de $\text{GlcNAcMan}_x\text{GlcNAc}_2$ donde x es más de 3 puede convertirse en $\text{GlcNAcMan}_3\text{GlcNAc}_2$, por ejemplo, mediante tratamiento con una actividad de α -1,3 y/o α -1,6 manosidasa, donde fuera aplicable. El procesamiento adicional de $\text{GlcNAcMan}_3\text{GlcNAc}_2$ mediante tratamiento con glucosiltransferasas (por ejemplo, GnTII) produce estructuras nucleares de $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ que después pueden modificarse, según se desee, por ejemplo, mediante tratamiento ex vivo o mediante expresión heteróloga en las células huésped de enzimas de glucosilación adicionales, que incluyen glucosiltransferasas, transportadores de azúcares y manosidasas (véase más adelante), para que se conviertan en *N*-glucanos humanoide.

Las glucoproteínas humanoides preferentes que pueden producirse de acuerdo con la invención incluyen las que comprenden *N*-glucanos que tienen siete o menos, o tres o menos, restos de manosa; y que comprenden una o más azúcares, que se seleccionan a partir del grupo constituido por galactosa, GlcNAc, ácido siálico, y fucosa.

Otra célula huésped preferente no humana de la invención es una célula de hongo unicelular o multicelular que tiene una actividad menor o anulada de una o más actividades del gen *alg* (que incluye una actividad enzimática que es un homólogo o equivalente a una actividad de *alg*). Otra célula huésped actividad de la invención tiene una actividad menor o anulada en la actividad de una o más enzimas (además de las actividades de *alg*) que manosilan la ramificación α -1,6 de una estructura de oligosacáridos ligada con lípidos.

Aunque son preferentes las células huésped de hongos unicelulares o multicelulares, se prevé una amplia variedad de células huésped que tienen las propiedades mencionadas anteriormente como útiles en los procedimientos de la invención. Las células vegetales, por ejemplo, pueden genomanipularse para que expresen una glucoproteína humanoide de acuerdo con la invención. Del mismo modo, puede alterarse una variedad de células de mamífero huésped no humanas para que expresen más glucoproteínas humanoides usando los procedimientos de la invención. Puede genomanipularse una célula huésped apropiada, o puede usarse una de las muchas mutantes ya descritas en las levaduras. Una célula huésped preferente de la invención, como se ejemplifica en el presente documento, es un mutante con hipermanosilación negativa (*OCH1*) en *Pichia pastoris* que ha sido modificada además para eliminar el gen *Alg3*.

La invención además proporciona células huésped de hongos unicelulares o multicelulares capaces de producir glucoproteínas que tienen *N*-glucanos diseccionados, como por ejemplo $\text{GlcNAcMan}_5\text{GlcNAc}_2$, $\text{GlcNAc}_2\text{Man}_5\text{GlcNAc}_2$, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$, y, preferentemente, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ diseccionados. En una realización preferente de la invención, las células huésped tal como se definen en las reivindicaciones comprenden una actividad de GnTIII. En una realización más preferente, las células huésped tal como se definen en las reivindicaciones además comprenden una o más actividades que se seleccionan a partir de: GnTI, GnTII, y GnTV. Las células huésped preferente expresan GnTI, GnTII, y GnTIII y además expresan la GnTIV tal como se define en las reivindicaciones y/o GnTV. Incluso más preferentemente, la una o más actividades de GnT de las células huésped son sustancialmente intracelulares.

Así, en realizaciones preferentes de la invención, las células huésped que comprenden la una o más actividades de GnT producen *N*-glucanos que comprenden estructuras que incluyen, pero sin limitación $\text{GlcNAcMan}_3\text{GlcNAc}_2$, $\text{GlcNAcMan}_4\text{GlcNAc}_2$, o $\text{GlcNAcMan}_5\text{GlcNAc}_2$, que son capaces de reaccionar con una actividad enzimática de GnTIII para producir *N*-glucanos diseccionados correspondientes. Las actividades enzimáticas por lo tanto convierten las glucoproteínas que contienen estos *N*-glucanos en formas con propiedades nuevas y más deseables. Debido a que actualmente se entiende que GnTIII inhibe la actividad de GnT adicional en células de mamífero, el experto debería entender que la reacción de glucosilación secuencial puede tener importancia o no. La presente invención contempla, sin embargo, la adición de GnTI y GnTIII en cualquier orden o juntas. También debería entenderse que pueden actuar otras actividades enzimáticas dentro de la célula, como por ejemplo, una o más actividades de manosidasa deseadas (por ejemplo, α 1,2 manosidasa, manosidasa I, manosidasa II), en concierto con las actividades de GnT para generar todavía otras glucoproteínas humanoides de interés (véase la Figura 1B).

En una realización preferente, se introduce una manosidasa II o un fragmento catalíticamente activo de la misma en las células huésped de la invención para recortar las ramificaciones que contienen α 1,3 y α 1,6 manosa de una estructura nuclear pentamansosa diseccionada como por ejemplo $\text{GlcNAc}_2\text{Man}_5\text{GlcNAc}_2$. Los glucanos resultantes (por ejemplo, $\text{GlcNAc}_2\text{Man}_4\text{GlcNAc}_2$ y $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ diseccionados) son preferente sustratos para la posterior modificación con *N*-glucanos humanoides.

En otra realización de la invención, las células huésped comprenden una estructura nuclear de $\text{Man}_5\text{GlcNAc}_2$ o una estructura nuclear de $\text{Man}_3\text{GlcNAc}_2$ modificadas por dos o más GlcNAc. Debería entenderse que cualquiera de las dos estructuras nucleares pueden incluir otras modificaciones además de la modificación por GlcNAc. Preferentemente, 10% o más de las estructuras nucleares son modificadas por GlcNAc. Los más preferentemente, 20%, 30%, 40%, 50%, 60%, 70%, 80% o incluso más de las estructuras nucleares contienen la modificación por GlcNAc.

Formación de *N*-glucanos complejos

La formación de la síntesis de *N*-glucanos complejos es un procedimiento secuencial por el que se eliminan restos de azúcares específicos y se unen a la estructura nuclear del oligosacárido. En los eucariotas superiores, esto se logra exponiendo el sustrato secuencialmente a diversas enzimas de procesamiento. Estas enzimas realizan reacciones específicas dependiendo de su localización particular dentro de la cascada completa de procesamiento. Esta "línea de ensamblaje" está formada por el RE, aparato de Golgi cis, medio y trans y la red del aparato de Golgi trans con todo su entorno de procesamiento específico. Para recrear el procesamiento de glucoproteínas humanas en el aparato de Golgi

y RE de eucariotas inferiores, deben expresarse numerosas enzimas (por ejemplo, glucosiltransferasas, glucosidasas, fosfatasa y transportadores) y dirigirse especialmente a estos organelos, y preferentemente, en una localización de forma que trabajen lo más eficientemente posible con respecto a su entorno, así como a otras enzimas de la ruta.

- 5 Dado que un objetivo de los procedimientos que se describen en el presente documento es una cepa con una producción potente de proteínas que sea capaz de funcionar bien en un proceso de fermentación industrial, la integración de los múltiples genes en el cromosoma de la célula huésped supone una planificación cuidadosa. Como se describe anteriormente, preferentemente habrá que eliminar uno o más genes que codifican enzimas que se sabe que son características de las reacciones de glucosilación no humanas. La cepa de células genomanipulada se transforma con un abanico de diferentes genes que codifican actividades deseadas, y estos genes se transforman de forma estable, asegurando así que se mantiene la actividad deseada durante todo el procedimiento de fermentación.

- Puede genomanipularse cualquier combinación de las siguientes actividades enzimáticas de una en una o de forma múltiple en el huésped usando los procedimientos de la invención: sialiltransferasas, manosidasas, fucosiltransferasas, galactosiltransferasas, GlcNAc transferasas, transportadores específicos del RE y del aparato de Golgi (por ejemplo transportadores simport y antiport para UDP-galactosa y otros precursores), otras enzimas implicadas en el procesamiento de los oligosacáridos, y enzimas implicadas en la síntesis de precursores de oligosacáridos activados como por ejemplo UDP-galactosa y CMP-*N*-ácido acetilneuramínico. Preferentemente, las actividades enzimáticas se introducen en una o más moléculas de ácido nucleico (véase también más adelante). Las moléculas de ácido nucleico pueden introducirse de una en una o de forma múltiple, por ejemplo, en el contexto de una adenoteca de ácido nucleico como por ejemplo una adenoteca de combinaciones de la invención. Debe entenderse, sin embargo, que pueden introducirse actividades enzimáticas únicas o múltiples en una célula huésped de cualquier manera, que incluye pero sin limitación procedimientos de administración de proteínas y/o el uso de una o más moléculas de ácido nucleico sin usar necesariamente una adenoteca de ácido nucleico o adenoteca de combinaciones de la invención.

25 *Expresión de glucosiltransferasas para producir N-glucanos complejos*

- Con la información de las secuencias de ADN, el experto puede clonar moléculas de ADN que codifican actividades de GnT (por ejemplo, los Ejemplos 3, 8, 11, 15, y 18). Usando técnicas estándar notorias para los expertos en la técnica, pueden insertarse moléculas de ácido nucleico que codifican GnTI, II, III, IV o V (o que codifican sus fragmentos catalíticamente activos) en vectores de expresión apropiados con transcripción controlada por promotores y otras secuencias de control de la expresión capaces de dirigir la transcripción en una célula huésped seleccionada de la invención, por ejemplo, un huésped fúngico como por ejemplo *Pichia* sp., *Kluyveromyces* sp. y *Aspergillus* sp., como se describe en el presente documento, de tal forma que una o más de estas enzimas GnT de mamífero puedan expresarse activamente en una célula huésped de elección para la producción de una glucoproteína compleja humanoide (por ejemplo, los Ejemplos 8, 20, y 21).

- Se han clonado y expresado varias glucosiltransferasas en *S. cerevisiae* (GalT, GnTI), *Aspergillus nidulans* (GnTI) y otros hongos, sin demostrar sin embargo el resultado de "humanización" deseado en el patrón de glucosilación de los organismos (Yoshida *et al.* (1999) *Glycobiology* 9(1):53-8; Kalsner *et al.* (1995) *Glycoconj. J.* 12(3):360-370). Se especuló que la estructura de carbohidrato que debía aceptar azúcares mediante la acción de dichas glucosiltransferasas no estaba presente en una cantidad suficiente, lo que con toda probabilidad contribuía a la falta de formación de *N*-glucanos complejos.

- 45 Un procedimiento preferente de la invención proporciona la expresión funcional de una GnT, como por ejemplo GnTI, GnTII, y GnTIII, en el aparato de Golgi cis, medio o trans, y además asegura un suministro suficiente de UDP-GlcNAc (por ejemplo, por expresión de un transportador de UDP-GlcNAc; véanse los Ejemplos más adelante).

Procedimientos para proporcionar precursores de azúcares-nucleótidos al aparato de Golgi

- 50 Para que una glucosiltransferasa funcione de forma satisfactoria en el aparato de Golgi, es necesaria una concentración suficiente de un azúcar nucleotídica apropiada, que es el donante con gran cantidad de energía del resto de azúcar añadido a una glucoproteína en formación. En los seres humanos, generalmente se sintetiza todo el abanico de azúcares de nucleótidos (por ejemplo UDP-*N*-acetilglucosamina, UDP-*N*-acetilgalactosamina, CMP-ácido *N*-acetilneuramínico, UDP-galactosa, etc.) en el citosol y se transportan al aparato de Golgi, donde se unen al oligosacárido nuclear mediante glucosiltransferasas.

- Para replicar este procedimiento en células huésped no humanas, como por ejemplo eucariotas inferiores, deben expresarse transportadores específicos de azúcares de nucleósidos en el aparato de Golgi para asegurar niveles adecuados de precursores de azúcares de nucleósidos (Sommers y Hirschberg (1981) *J. Cell Biol.* 91(2):A406-A406; Sommers y Hirschberg (1982) *J. Biol. Chem.* 257(18):811-817; Perez and Hirschberg (1987) *Methods in Enzymology* 138:709-715). Los azúcares de nucleósidos pueden proporcionarse a los compartimientos apropiados, por ejemplo, expresando en el microorganismo huésped un gen exógeno que codifica un transportador de azúcar de nucleótido. La elección de la enzima transportadora se ve influenciada por la naturaleza de la glucosiltransferasa exógena que se esté usando. Por ejemplo, una GlcNAc transferasa puede requerir un transportador de UDP-GlcNAc, una fucosiltransferasa puede requerir un transportador de GDP-fucosa, una galactosiltransferasa puede requerir un transportador de UDP-galactosa, y una sialiltransferasa puede requerir un transportador de CMP-ácido siálico.

La proteína transportadora añadida transporta un azúcar nucleotídica del citosol al aparato de Golgi, donde el azúcar nucleotídica puede hacerse reaccionar mediante la glucosiltransferasa, por ejemplo para alargar un *N*-glucano. La reacción libera un nucleósido difosfato o monofosfato, por ejemplo UDP, GDP, o CMP. Los nucleósido monofosfatos pueden exportarse directamente desde el aparato de Golgi en un intercambio por azúcares de nucleósido trifosfato mediante un mecanismo de antiport. La acumulación de un nucleósido difosfato, sin embargo, inhibe la posterior actividad de una glucosiltransferasa. Dado que esta reacción parece ser importante para una glucosilación eficaz, con frecuencia es deseable proporcionar una copia expresada de un gen que codifica una nucleótido difosfatasa. La difosfatasa (específica para UDP o GDP según sea apropiado) hidroliza el difosfonucleósido proporcionando un nucleósido monofosfato y fosfato inorgánico.

A continuación se describen las enzimas transportadoras adecuadas, que habitualmente son de origen mamífero. Dichas enzimas pueden ser genomanipuladas en una célula huésped seleccionada usando los procedimientos de la invención.

En otro ejemplo, α 2,3- o α 2,6-sialiltransferasa forma una caperuza sobre los restos de galactosa con ácido siálico en el aparato de Golgi trans y TGN de los seres humanos produciendo una forma madura de la glucoproteína (Figura 1B). Volver a diseñar esta etapa de procesamiento en levaduras u hongos metabólicamente genomanipuladas requerirá (1) actividad de α 2,3- o α 2,6-sialiltransferasa y (2) un suministro suficiente de ácido CMEP-*N*-acetil neuramínico, en el aparato de Golgi trans de las levaduras. Para obtener una suficiente actividad de α 2,3-sialiltransferasa en el aparato de Golgi trans, por ejemplo, debe dirigirse el dominio catalítico de una sialiltransferasa conocida (por ejemplo de seres humanos) al aparato de Golgi trans en los hongos (véase anteriormente). Del mismo modo, deben genomanipularse transportadores para permitir el transporte de ácido CMP-*N*-acetil neuramínico en el aparato de Golgi trans. Actualmente no existe ninguna indicación de que los hongos sinteticen ni incluso de que puedan transportar cantidades suficientes de ácido CMP-*N*-acetil neuramínico en el aparato de Golgi. En consecuencia, para asegurar que se proporciona el suministro adecuado de sustrato para las glucosiltransferasas correspondientes, debe diseñarse metabólicamente la producción de CMP-ácido siálico en el hongo.

UDP-*N*-acetilglucosamina

El ADNc del transportador de UDP-*N*-acetilglucosamina humano, que se reconoció mediante búsqueda de homología en la base de datos de marcadores de las secuencias expresadas (dbEST), ha sido clonado (Ishida (1999) J. Biochem, 126(1):68-77). El transportador de la membrana del aparato de Golgi en mamíferos para UDP-*N*-acetilglucosamina se clonó mediante corrección fenotípica con ADNc de células renales caninas (MDCK) de un mutante de *Kluyveromyces lactis* caracterizado recientemente deficiente en transporte en el aparato de Golgi de la anterior azúcar de nucleótido (Guillen *et al.* (1998) Proc. Natl. Acad. Sci. USA 95(14):7888-7892). Los resultados demuestran que el gen transportador de UDP-GlcNAc en el aparato de Golgi de mamífero tiene toda la información necesaria para que la proteína se exprese y se dirija funcionalmente al aparato de Golgi de levaduras y que dos proteínas con secuencias de aminoácidos muy diferentes pueden transportar el mismo soluto dentro de la misma membrana del aparato de Golgi (Guillen *et al.* (1998) Proc. Natl. Acad. Sci. USA 95(14):7888-7892).

Por consiguiente, se puede incorporar la expresión de un transportador de UDP-GlcNAc a una célula huésped por medio de una construcción de ácido nucleico que puede contener, por ejemplo: (1) una región mediante la cual la construcción transformada se mantiene en la célula (por ejemplo, origen de replicación o una región que actúa de mediadora en la integración cromosómica), (2) un gen marcador que permite seleccionar células que han sido transformadas, que incluye marcadores contraseleccionables y reciclables como por ejemplo *ura3* o *T-urf13* (Soderholm *et al.* (2001) Biotechniques 31(2):306-10) u otros marcadores de selección bien caracterizados (por ejemplo, *his4*, *bla*, *Sh ble* etc.), (3) un gen o fragmento del mismo que codifica un transportador de UDP-GlcNAc funcional (por ejemplo, de *K. lactis*, (Abeijon, (1996) Proc. Natl. Acad. Sci. U.S.A. 93:5963-5968), o de *H. sapiens* (Ishida *et al.* (1996) J. Biochem. (Tokio) 120(6):1074-8), y (4) un promotor que activa la expresión de la anterior adenoteca de construcciones de fusiones de localización/dominio catalítico.

GDP-Fucosa

El transportador de GDP-fucosa de la membrana del aparato de Golgi en el hígado de la rata, ha sido identificado y purificado por Puglielli y Hirschberg (1999) J. Biol. Chem. 274(50):35596-35600. El gen correspondiente no ha sido identificado, sin embargo, puede usarse secuenciación del extremo N para el diseño de sondas de oligonucleótidos específicas para el gen correspondiente. Estos oligonucleótidos pueden usarse como sondas para clonar el gen que codifica el transportador de GDP-fucosa.

UDP-Galactosa

Se han expresado funcionalmente dos genes heterólogos, *gmal2(+)* que codifican alfa 1,2-galactosiltransferasa (alfa 1,2 GalT) de *Schizosaccharomyces pombe* y (*hUGT2*) que codifican el transportador de UDP-galactosa (UDP-Gal) humano, en *S. cerevisiae* para examinar las condiciones intracelulares necesarias para la galactosilación. La correlación entre la galactosilación de proteína y la actividad de transporte de UDP-galactosa indicaron que un aporte exógeno de transportador de UDP-Gal y no de alfa 1,2 GalT desempeñaba un papel clave para la galactosilación eficaz en *S. cerevisiae* (Kainuma (1999) Glycobiology 9(2):133-141). Del mismo modo, se clonó un transportador de UDP-galactosa de *S. pombe* (Segawa (1999) FEBS Letters 451(3):295-298).

Ácido CMP-N-acetilneuramínico (CMP-ácido siálico)

El transportador de CMP-ácido siálico humano (hCST) ha sido clonado y expresado en células CHO Lec 8 por (Aoki *et al.* (1999) J. Biochem. (Tokio) 126(5):940-50; Eckhardt *et al.* (1997) Eur. J. Biochem. 248(1):187-92). Se logró la expresión funcional del transportador de CMP-ácido siálico murino en *Saccharomyces cerevisiae* (Berninsone, *et al.* (1997) J. Biol. Chem. 272(19):12616-9). Se ha encontrado ácido siálico en algunos hongos, sin embargo no queda claro si el sistema huésped elegido será capaz de proporcionar niveles suficientes de CMP-ácido siálico. El ácido siálico puede o bien proporcionarse en el medio o, de forma alternativa, también pueden integrarse rutas fúngicas implicadas en la síntesis de ácido siálico en el genoma huésped.

Expresión de difosfatasas

Cuando se transfieren azúcares sobre una glucoproteína, se libera o bien un nucleósido difosfato o monofosfato de los precursores de azúcares de nucleótidos. Mientras que los monofosfatos pueden exportarse directamente en un intercambio con los azúcares de nucleósido trifosfatos mediante un mecanismo antiporte, los difosfonucleósidos (por ejemplo GDP) tienen que ser escindidos por las fosfatasas (por ejemplo GDPasa) liberando nucleósido monofosfatos y fosfato inorgánico antes de ser exportados. Esta reacción parece ser importante para la glucosilación eficaz, ya que se ha encontrado que la GDPasa de *S. cerevisiae* es necesaria para la manosilación. Sin embargo, la enzima sólo tiene 10% de la actividad hacia UDP (Berninsone *et al.* (1994) J. Biol. Chem. 269(1):207-211). Los eucariotas inferiores a menudo carecen de actividad de difosfatasa específica de UDP en el aparato de Golgi dado que no utilizan precursores de azúcares UDP para la síntesis de las glucoproteínas en el aparato de Golgi. Se encontró que *Schizosaccharomyces pombe*, una levadura que añade restos de galactosa a los polisacáridos de la pared celular (de UDP-galactosa), tenía actividad específica de UDPasa, lo que sugiere también la necesidad de dicha enzima (Berninsone *et al.* (1994) J. Biol. Chem. 269(1):207-211). Es sabido que el UDP es un inhibidor potente de las glucosiltransferasas y la eliminación de este subproducto de la glucosilación es importante para evitar la inhibición de la glucosiltransferasa en la luz del aparato de Golgi (Khatara *et al.*, (1974) Eur. J. Biochem. 44:537-560).

Procedimientos para alterar los N-glucanos en un huésped expresando una actividad enzimática dirigida a partir de una molécula de ácido nucleico

También se describen en el presente documento procedimientos para producir una glucoproteína humanoide en una célula huésped no humana que comprende la etapa de introducir en la célula una o más moléculas de ácido nucleico que codifican una enzima o enzimas para la producción de la estructura de carbohidrato Man₅GlcNAc₂. Como se describe en el presente documento, se introduce en el huésped una molécula de ácido nucleico que codifica una o más actividades de manosidasa implicada en la producción de Man₅GlcNAc₂ a partir de Man₈GlcNAc₂ o Man₉GlcNAc₂. La invención además se refiere a procedimientos para preparar glucoproteínas alteradas en una célula huésped que comprende la etapa de introducir en las células huésped una molécula de ácido nucleico que codifica una o más enzimas o actividades de glucosilación. Las actividades enzimáticas preferentes se seleccionan a partir del grupo constituido por UDP-GlcNAc transferasa, UDP-galactosiltransferasa, GDP-fucosiltransferasa, CMP-sialiltransferasa, transportador de UDP-GlcNAc, transportador de UDP-galactosa, transportador de GDP-fucosa, transportador de CMP-ácido siálico y nucleótido difosfatasas. En una realización particularmente preferente, se selecciona el huésped o se genomanipula para que exprese dos o más actividades enzimáticas en las que el producto de una actividad aumenta los niveles de los sustratos de otra actividad, por ejemplo, una glucosiltransferasa y un transportador de azúcares correspondiente, por ejemplo, actividades de GnTI y de transportador de UDP-GlcNAc. En otra realización preferente, el huésped se selecciona o se genomanipula para que exprese una actividad para eliminar productos que puedan inhibir las posteriores reacciones de glucosilación, por ejemplo una actividad de difosfatasa específica de UDP o GDP.

Los procedimientos preferentes de la invención suponen expresar una o más actividades enzimáticas a partir de una molécula de ácido nucleico en una célula huésped y comprenden la etapa de dirigir al menos una actividad enzimática a una localización subcelular deseada (por ejemplo, un organelo) formando una proteína de fusión que comprende un dominio catalítico de la enzima y un péptido de señalización de direccionamiento celular, por ejemplo, un péptido de señal heterólogo que normalmente no está ligado ni asociado al dominio catalítico. La proteína de fusión es codificada por al menos una construcción genética ("construcción fusionada") que comprende un fragmento de ácido nucleico que codifica un péptido de señalización de direccionamiento celular ligado en el mismo marco de lectura de traducción ("en marco") a un fragmento de ácido nucleico que codifica una enzima (por ejemplo, una enzima de glucosilación), o un fragmento catalíticamente activo de la misma.

El componente de péptido de señalización de direccionamiento de la construcción fusionada o proteína preferentemente se deriva a partir de un miembro del grupo constituido por: proteínas unidas a la membrana del RE o aparato de Golgi, señales de recuperación, proteínas de la membrana de tipo I, transportadores de azúcares de nucleótidos que atraviesan la membrana, manosidasas, sialiltransferasas, glucosidasas, manosiltransferasas y fosfomanosiltransferasas.

El componente del dominio catalítico de la construcción fusionada o proteína preferentemente se deriva de una actividad de glucosidasa, manosidasa o glucosiltransferasa que se deriva a partir de un miembro del grupo constituido por GnTI, GnTII, GnTIII, GnTIV, GnTV, GnTVI, GaIT, fucosiltransferasa y sialiltransferasa. El dominio catalítico preferentemente tiene un pH óptimo a 1,4 unidades de pH del pH medio óptimo de otras enzimas representativas en el organelo en el que la enzima está localizada o tiene una actividad óptima a un pH entre 5,1 y 8,0. En una realización preferente, el dominio catalítico codifica una manosidasa que se selecciona a partir del grupo constituido

por manosidasa IA de *C. elegans*, manosidasa IB de *C. elegans*, manosidasa IA de *D. melanogaster*, manosidasa IB de *H. sapiens*, manosidasa I de *P. citrinum*, manosidasa IA de ratón, manosidasa IB de ratón, manosidasa IA de *A. nidulans*, manosidasa IB de *A. nidulans*, manosidasa IC de *A. nidulans*, manosidasa II de ratón, manosidasa II de *C. elegans*, manosidasa II, manosidasa fix, y manosidasa III de *H. sapiens*.

Selección de una enzima de glucosilación: pH óptimos y localización subcelular

En una realización de la invención, se prepara de forma eficaz una glucoproteína humanoide en una célula huésped eucariota no humana introduciendo en un compartimiento subcelular de la célula una enzima de glucosilación que se selecciona para que tenga un pH óptimo similar a los pH óptimos de otras enzimas en el compartimiento subcelular objetivo. Por ejemplo, la mayoría de las enzimas que son activas en el RE y aparato de Golgi de *S. cerevisiae* tienen un pH óptimo que está entre 6,5 y 7,5 (véase la Tabla 3). Dado que la glucosilación de proteínas es un proceso muy evolucionado y eficiente, el pH interno en el RE y en el aparato de Golgi probablemente también esté comprendido en el intervalo de aproximadamente 6-8. Todas las estrategias previas para reducir la manosilación mediante la acción de manosidasas recombinantes en huéspedes finales, sin embargo, han introducido enzimas que tienen un óptimo de pH de aproximadamente pH 5,0 (Martinet *et al.* (1998) Biotech. Letters 20(12):1171-1177, y Chiba *et al.* (1998) J. Biol. Chem. 273(41):26298-26304). A pH 7,0, la actividad determinada *in vitro* de esas manosidasas se reduce a menos de 10%, lo que probablemente es una actividad insuficiente en el lugar en que se usa, en concreto el RE y el aparato de Golgi cis, para la producción eficaz *in vivo* de Man₅GlcNAc₂ sobre *N*-glucanos.

Por consiguiente, una realización preferente de esta invención dirige una enzima de glucosilación seleccionada (o su dominio catalítico), por ejemplo, una (1-manosidasa, a una localización subcelular en las células huésped (por ejemplo, un organelo) donde el pH óptimo de la enzima o dominio está comprendida en el intervalo de 1,4 unidades de pH del pH óptimo medio de otras enzimas marcadoras representativas localizadas en el mismo organelo(s). El pH óptimo de la enzima que va a dirigirse a un organelo específico debería corresponder al pH óptimo de otras enzimas que se encuentran en el mismo organelo, para maximizar la actividad por unidad de enzima. La Tabla 3 resume la actividad de las manosidasas de diversas fuentes y sus pH óptimos respectivos. La Tabla 4 resume su localización subcelular habitual.

Tabla 3. Manosidasas y su pH óptimo.

Fuente	Enzima	pH óptimo	Referencia
<i>Aspergillus saitoi</i>	α -1,2-manosidasa	5,0	Ichishima <i>et al.</i> (1999) <i>Biochem. J.</i> 339 (Pt 3): 589-597
<i>Trichoderma reesei</i>	α -1,2-manosidasa	5,0	Maras <i>et al.</i> (2000) <i>J. Biotechnol.</i> 77 (2-3): 255-263
<i>Penicillium citrinum</i>	α -D-1,2-manosidasa	5,0	Yoshida <i>et al.</i> (1993) <i>Biochem. J.</i> 290 (Pt 2): 349-354
<i>C. elegans</i>	α -1,2-manosidasa	5,5	véase la Figura 11
<i>Aspergillus nidulans</i>	α -1,2-manosidasa	6,0	Eades y Hintz (2000) <i>Gene</i> 255 (1): 25-34
<i>Homo sapiens</i> IA(aparato de Golgi)	α -1,2-manosidasa	6,0	
<i>Homo sapiens</i> IB (aparato de Golgi)	α -1,2-manosidasa	6,0	

5	Células de insecto lepidóptero	Manosidasa α -1,2-Man ₆ de tipo I	6,0	Ren <i>et al.</i> (1995) <i>Biochem.</i> 34(8): 2489-2495
10	<i>Homo sapiens</i>	α -D-manosidasa	6,0	Chandrasekaran <i>et al.</i> (1984) <i>Cancer Res.</i> 44(9): 4059-68
15	<i>Xanthomonas manihotis</i>	α -1,2,3-manosidasa	6,0	patente de Estados Unidos N.º 6.300.113
20	IB de ratón (aparato de Golgi)	α -1,2-manosidasa	6,5	Schneikert y Herscovics (1994) <i>Glycobiology</i> , 4(4): 445-50
25	<i>Bacillus</i> sp. (segregada)	α -D-1,2-manosidasa	7,0	Maruyama <i>et al.</i> (1994) <i>Carbohydrate Res.</i> 251: 89-98

En una realización preferente, una enzima o dominio catalítico particular se dirige a una localización subcelular en las células huésped por medio de una construcción quimérica por fusión que codifica una proteína que comprende un péptido de señalización de direccionamiento celular no asociado normalmente con el dominio enzimático. Preferentemente, un enzima o dominio se dirige al RE, al aparato de Golgi cis, medio o trans, o al aparato de Golgi trans de las células huésped.

En una realización más preferente, la enzima de glucosilación dirigida es una manosidasa, glucosiltransferasa o una glucosidasa. En una realización especialmente preferente, la actividad de manosidasa se dirige al RE o al aparato de Golgi cis, donde se producen las reacciones de glucosilación tempranas. Aunque este procedimiento es útil para producir una glucoproteína humanoide en una célula huésped no humana, se entenderá que el procedimiento también es útil de forma más general para modificar los perfiles de carbohidratos de una glucoproteína en cualquier célula huésped eucariota, que incluye las células huésped humanas.

Dirigir las secuencias que actúan de mediadoras de la retención de proteínas en ciertos organelos de la ruta de secreción de las células huésped es notoria y se ha descrito en la bibliografía científica y en bases de datos públicas, que se describen en más detalle a continuación con respecto a adenotecas para la selección de secuencias de direccionamiento y enzimas dirigidas. Dichas secuencias de direccionamiento subcelular pueden usarse solas o combinadas para dirigir una enzima de glucosilación seleccionada (o su dominio catalítico) a una localización subcelular particular en una célula huésped, *es decir*, especialmente a una en la que la enzima tenga una actividad mejorada u óptima basándose en los pH óptimos o en la presencia de otros factores estimuladores.

Cuando se intenta recortar las estructuras con alto contenido en manosa proporcionando Man₅GlcNAc₂ en el RE o en el aparato de Golgi de una célula huésped como por ejemplo *S. cerevisiae*, por ejemplo, se puede elegir cualquier enzima o combinación de enzimas que (1) tenga un pH óptimo lo suficientemente parecido (es decir, entre pH 5,2 y pH 7,8), y (2) se sepa que genera, sólo o combinada, la estructura isomérica específica Man₅GlcNAc₂ necesaria para aceptar la posterior adición de GlcNAc por GnTI. Cualquier enzima o combinación de enzimas que haya demostrado que genera una estructura que puede convertirse en GlcNAcMan₅GlcNAc₂ mediante GnTI *in vitro* constituiría una elección apropiada. Este conocimiento puede obtenerse en la bibliografía científica o de forma experimental. Por ejemplo, se puede determinar si una manosidasa potencial puede convertir Man₃GlcNAc₂-2AB (2-aminobenzamida) en Man₅GlcNAc₂-AB y después verificar que la estructura Man₅GlcNAc₂-2AB obtenida puede servir como sustrato para GnTI y UDP-GlcNAc para obtener GlcNAcMan₅GlcNAc₂ *in vitro*. Por ejemplo, la manosidasa IA de una fuente humana o murina sería una elección apropiada (véase, por ejemplo, Ejemplo 4). Los ejemplos que se describen en el presente documento utilizan oligomanosa marcada con 2-aminobenzamida ligada en N seguida de análisis por HPLC para realizar esta determinación.

Tabla 4. Localización celular y pH óptimo de diversas enzimas relacionadas con la glucosilación de *S. cerevisiae*.

Gen	Actividad	Localización	pH óptimo	Referencia(s)
KTR1	α -1,2-manosiltransferasa	Golgi	7,0	Romero <i>et al.</i> (1997) <i>Biochem. J.</i> 321 (Pt 2): 289-295
MNS1	α -1,2- manosidasa	RE	6,5	Lipari <i>et al.</i> (1994) <i>Glycobology.</i> Oct; 4(5): 697-702
GWH41--	glucosidasa I	ER Golgi	6,8 7-8	Lehele y Tanner (1974) <i>Biochim. Biophys. Acta</i> 350 (1): 225-235
KRE2	α -1,2 manosiltransferasa	Golgi	6,5-9,0	Romero <i>et al.</i> (1997) <i>Biochem. J.</i> 321 (Pt 2): 289-295

Por consiguiente, una enzima de glucosilación como por ejemplo una enzima α -1,2-manosidasa que se usa de acuerdo con la invención tiene una actividad óptima a un pH de entre 5,1 y 8,0. En una realización preferente, la enzima tiene una actividad óptima a un pH de entre 5,5 y 7,5. La enzima manosidasa de *C. elegans*, por ejemplo, funciona bien en los procedimientos de la invención y tiene un pH óptimo aparente de aproximadamente 5,5). Las manosidasas preferidas incluyen las que se recogen en la Tabla 3 que tienen pH óptimos apropiados, por ejemplo IA de *Aspergillus nidulans*, *Homo sapiens* (aparato de Golgi), IB de *Homo sapiens* (aparato de Golgi), células de insectos lepidópteros (IPLB-SF21AE), Ib de *Homo sapiens*, y de ratón (aparato de Golgi), y *Xanthomonas manihotis*, *Drosophila melanogaster* y *C. elegans*.

En el Ejemplo 7 se describe un experimento que ilustra el pH óptimo para una enzima α -1,2-manosidasa. Una proteína de fusión quimérica BB27-2 (MNN10 (s) de *Saccharomyces*/manosidasa IB Δ 3,1 de *C. elegans*), que sale al medio se sometió a diversos intervalos de pH para determinar la actividad óptima de la enzima. Los resultados del experimento muestran que la α -1,2-manosidasa tiene un pH óptimo de aproximadamente 5,5 para su función (Figura 11).

En una realización preferida, se expresa un único gen de manosidasa clonado en el organismo huésped. Sin embargo, en algunos casos, puede ser deseable expresar varios genes diferentes de manosidasa, o varias copias de un gen particular, para lograr una producción adecuada de Man₅GlcNAc₂. En los casos en los que se usan múltiples genes, las manosidasas codificadas preferentemente todas tienen pH óptimos dentro del intervalo preferido de aproximadamente

5,1 a aproximadamente 8,0, o especialmente entre aproximadamente 5,9 y aproximadamente 7,5. Las actividades de manosidasa preferente incluyen α -1,2-manosidasas derivadas de ratón, ser humano, *Lepidoptera*, *Aspergillus nidulans*, o *Bacillus* sp., *C. elegans*, *D. melanogaster*, *P. citrinum*, *X. laevis* o *A. nidulans*.

5

Alteración in vivo de la glucosilación celular del huésped usando una adenoteca de combinaciones

Ciertos procedimientos de la invención preferentemente (pero no necesariamente) pueden realizarse usando una o más adenotecas de ácido nucleico. Una característica ejemplar de una adenoteca de combinaciones de ácido nucleico de la invención es que comprende secuencias que codifican péptidos y secuencias de señalización de direccionamiento celular que codifican proteínas a dirigir (por ejemplo, enzimas o sus dominios catalíticos, que incluyen pero sin limitación los que actúan de mediadores de la glucosilación).

Como se describe en el presente documento, una adenoteca de combinaciones de ácido nucleico comprende: (a) al menos dos secuencias de ácido nucleico que codifican diferentes péptidos de señalización de direccionamiento celular; y (b) al menos una secuencia de ácidos nucleicos que codifican un polipéptido a dirigir. Otra adenoteca de combinaciones de ácido nucleico que se describe en el presente documento comprende:

(a) al menos una secuencia de ácidos nucleicos que codifican un péptido de señalización de direccionamiento celular; y (b) al menos dos secuencias de ácido nucleico que codifican un polipéptido a dirigir en una célula huésped. Como se describe en más detalle más adelante, una secuencia de ácidos nucleicos derivada de (a)

y una secuencia de ácidos nucleicos derivada de (b) se ligan para producir una o más construcciones fusionadas que codifican un péptido de señalización de direccionamiento celular ligado funcionalmente a un dominio de un polipéptido de interés. Un ejemplo de un enlace funcional es cuando el péptido de señalización de direccionamiento celular se liga al dominio de interés del polipéptido en el mismo marco de lectura de traducción ("en marco").

Como se describe en el presente documento, una adenoteca de combinaciones expresa una o más proteínas de fusión que comprenden péptidos de señalización de direccionamiento celular ligados en marco a dominios de enzimas catalíticas. La proteína de fusión codificada preferentemente comprende un dominio catalítico de una enzima implicada en la modificación de *N*-glucanos de tipo mamífero o humanoide. Como se describe en el presente documento, el dominio catalítico se deriva de una enzima que se selecciona a partir del grupo constituido por manosidasas, glucosiltransferasas y otras glucosidasas que están ligadas en marco a uno o más péptidos de señalización de direccionamiento. El dominio enzimático puede ser exógeno y/o endógena para la célula huésped. Un péptido de señal particularmente preferente es uno normalmente asociado a una proteína que sufre transporte del RE al aparato de Golgi.

La adenoteca de combinaciones que se describe en el presente documento puede usarse para producir y localizar *in vivo* enzimas implicadas en la modificación de *N*-glucanos de tipo mamífero o humanoide. Las construcciones fusionadas de la adenoteca de combinaciones son genomanipuladas de forma que las enzimas codificadas estén localizadas en el RE, aparato de Golgi o la red del aparato de Golgi trans de las células huésped donde están implicadas en la producción de *N*-glucanos particulares sobre una glucoproteína de interés. La localización de enzimas que modifican *N*-glucanos de la presente invención se logra a través de un mecanismo de anclaje o a través de interacción entre proteínas donde el péptido de localización construido a partir de la adenoteca de combinaciones localiza un organelo deseado de la ruta de secreción como por ejemplo el RE, aparato de Golgi o la red del aparato de Golgi trans.

Un ejemplo de un *N*-glucano útil, que se produce eficazmente y en cantidades suficientes para la modificación posterior mediante reacciones de glucosilación humanoides (complejas) es $\text{Man}_5\text{GlcNAc}_2$. Es necesaria una cantidad suficiente de $\text{Man}_5\text{GlcNAc}_2$ en una glucoproteína de interés para el procesamiento humanoide posterior *in vivo* (por ejemplo, más de 30% molar). El intermedio $\text{Man}_5\text{GlcNAc}_2$ puede usarse como sustrato para la posterior modificación de *N*-glucanos para producir $\text{GlcNAcMan}_5\text{GlcNAc}_2$ (Figura 1 B; véase anteriormente). Por consiguiente, puede usarse la adenoteca de combinaciones que se describe en el presente documento para producir enzimas que posteriormente producen $\text{GlcNAcMan}_5\text{GlcNAc}_2$, u otros *N*-glucanos complejos deseados, en una cantidad útil.

Un aspecto adicional de las construcciones fusionadas producidas usando la adenoteca de combinaciones de la presente invención es que permiten una actividad de recorte intracelular de *N*-glucanos suficiente y a menudo casi completa en la célula huésped genomanipulada. Las construcciones fusionadas preferente producidas mediante la adenoteca de combinaciones de la invención codifican una enzima de glucosilación, por ejemplo, una manosidasa, que está localizada eficazmente en un compartimiento intracelular de la célula huésped y así muestra muy poca y preferentemente ninguna actividad extracelular. Se ha demostrado que las construcciones fusionadas preferentes de la presente invención que codifican una enzima manosidasa están localizadas donde los *N*-glucanos son modificados, en concreto, en el RE y el aparato de Golgi. Las enzimas de fusión de la presente invención se dirigen a los organelos particulares de la ruta secretora donde se localizan y actúan sobre *N*-glucanos como por ejemplo $\text{Man}_5\text{GlcNAc}_2$ para producir $\text{Man}_5\text{GlcNAc}_2$ sobre una glucoproteína de interés.

Las construcciones fusionadas de GnTIII generadas a partir de una adenoteca de combinaciones para producir glucanos diseccionados se ensayaron para determinar cualquier actividad extracelular. Un ejemplo de una construcción fusionada de GnTIII que muestra alteración *in vivo* de la glucosilación de las células huésped se denomina pVA53.

Después de transformar *P. pastoris* YSH-1 con la construcción fusionada pVA53, el sobrenadante se analizó para detectar cualquier actividad de GnTIII *ex vivo*. La Figura 33 no muestra cambios aparentes en el sustrato estándar GlcNAcMan₅GlcNAc₂ en condiciones que revelarían la actividad extracelular de GnTIII en el medio (Ejemplo 22). De forma similar, la Figura 34 no muestra actividad extracelular de GnTIII detectable en el medio en *P. pastoris* YSH-57 que reaccione con el sustrato GlcNAc₂Man₃GlcNAc₂ (Ejemplo 23).

Las enzimas producidas por la adenoteca de combinaciones de la presente invención puede modificar *N*-glucanos sobre una glucoproteína de interés como se ha demostrado para las proteínas K3 o IFN- β expresadas en *P. pastoris*, como se muestra en las Figuras 5, 6, y 25-34 (véanse también los Ejemplos 2, 4, y 18-23). Sin embargo, se entenderá que otros tipos de glucoproteínas, sin limitación, que incluyen eritropoyetina, citocinas como por ejemplo interferón- α , interferón- β , interferón- γ , interferón- ω , y CSF granulocítico, factores de coagulación como por ejemplo factor VIII, factor IX, y proteína C humana, cadena α del receptor soluble de IgE, IgG, fragmentos de IgG, IgM, interleucinas, urocinasa, quimasa, y inhibidor de tripsina de urea, proteína que se une a IGF, factor de crecimiento epidérmico, factor liberador de hormona del crecimiento, proteína de fusión de anexina V, angiostatina, factor de crecimiento endotelial vascular-2, factor inhibidor progenitor mieloide-1, osteoprotegerina, antitripsina α -1, ADNasa II, α -fetoproteínas, AAT, rhTBP-1 (oncept también conocida como proteína de unión a TNF 1), TACI-Ig (activador transmembrana y modulador del calcio e interacto con los ligandos de ciclofilina), FSH (hormona estimulante de folículos), GM-CSF, GLP-1 con o sin FC (proteína glucagonoide 1), agonista del receptor IL-1, sTNRr (enbrel, también conocida como fusión de TNF soluble y receptor Fc), ATIII, rhTrombina, glucocerebrosidasa y CTLA4-Ig (antígeno asociado a linfocitos T citotóxicos 4-Ig) pueden glucosilarse de este modo.

Construcción de una adenoteca de combinación de construcciones fusionadas:

Una adenoteca de combinaciones de construcciones fusionadas presenta uno o más péptidos de señalización de direccionamiento celular ("péptidos de direccionamiento") que se derivan de forma general a partir de dominios del extremo N de proteínas nativas (por ejemplo, realizando deleciones en el extremo C). Algunos péptidos directores, sin embargo, se derivan del extremo C de las proteínas nativas (por ejemplo *SEC12*). Preferentemente se usan las proteínas unidas a las membranas del RE o del aparato de Golgi como fuente para la secuencia de los péptidos directores. Estas proteínas tienen secuencias que codifican una cola citosólica (ct), un dominio transmembrana (tmd) y una región de tallo (sr) que son de longitud variable. Estas regiones pueden reconocerse mediante alineaciones de las secuencias de las proteínas y comparaciones con homólogos conocidos y/u otras proteínas localizadas (por ejemplo, comparando gráficas de capacidad hidrófoba).

Los péptidos directores se indican en el presente documento como corto (s), medio (m) y largos (l) con respecto a las partes de una membrana de tipo II. La secuencia del péptido director que se indica como corta (s) corresponde al dominio transmembrana (tmd) de la proteína unida a la membrana. La secuencia del péptido director que se indica como larga (l) corresponde al dominio transmembrana (tmd) y a la región del tallo (sr). La secuencia del péptido director que se indica como media (m) corresponde al dominio transmembrana (tmd) y aproximadamente la mitad de la longitud de la región del tallo (sr). Las regiones de los dominios catalíticos se indican en el presente documento mediante el número de deleción de nucleótidos con respecto a su enzima de glucosilación de tipo silvestre.

Sub-adenotecas

En algunos casos, puede compendiarse una adenoteca de combinaciones de ácido nucleico de la invención directamente a partir de genes existentes o de tipo silvestre. Preferentemente, la adenoteca se compone a partir de la fusión de dos o más subadenotecas. Mediante el ligado en marco de subadenotecas, es posible crear un gran número de construcciones genéticas novedosas que codifican dominios de proteínas dirigidas de utilidad como por ejemplo las que tienen actividades de glucosilación.

Subadenotecas de dominios catalíticos que codifican actividades de glucosilación

Una subadenoteca incluye secuencias de ADN que codifican enzimas como por ejemplo glucosidasas (por ejemplo, manosidasas), glucosiltransferasas (por ejemplo, fucosiltransferasas, galactosiltransferasas, glucosiltransferasas), GlcNAc transferasas y sialiltransferasas. Los dominios catalíticos pueden seleccionarse a partir del huésped a genomanipular, así como a partir de otros organismos relacionados o no. Las enzimas de mamífero, plantas, insectos, reptiles, algas u hongos son todas útiles y deberían elegirse para que representen un amplio espectro de propiedades bioquímicas con respecto a la temperatura y pH óptimos. Preferentemente, los genes son truncados para obtener fragmentos, algunos de los cuales codifican los dominios catalíticos de las enzimas. Al eliminar las secuencias de direccionamiento endógenas, las enzimas pueden redirigirse y expresarse después en otros locus celulares.

La elección de dichos dominios catalíticos puede guiarse por el conocimiento del entorno particular en el que el dominio catalítico estará activo posteriormente. Por ejemplo, si una enzima de glucosilación particular va a estar activa en el aparato de Golgi trans, y todas las enzimas conocidas del organismo huésped del aparato de Golgi trans tienen un cierto pH óptimo, o se sabe que el aparato de Golgi trans tiene un pH particular, entonces se elige un dominio catalítico que muestre una actividad adecuada, y preferentemente máxima, a ese pH, como se ha descrito anteriormente.

Subadenotecas de secuencias peptídicas de direccionamiento

Otra subadenoteca útil incluye secuencias de ácido nucleico que codifican péptidos de señalización de direccionamiento que provocan la localización de una proteína en un emplazamiento particular dentro de la red del RE, del aparato de Golgi, o del aparato de Golgi trans. Estos péptidos directores pueden seleccionarse a partir del organismo huésped a genomanipular, así como a partir de otros organismos relacionados o no. De forma general dichas secuencias entran en tres categorías: (1) secuencias del extremo N que codifican una cola citosólica (ct), un dominio transmembrana (tmd) y parte o toda de una región del tallo (sr), que juntos o individualmente anclan las proteínas a la membrana interna (del lumen) del aparato de Golgi; (2) señales de recuperación que de forma general se encuentran en el extremo C como por ejemplo el tetrapéptido HDEL (SEC. ID. N.º: 41) o KDEL (SEC. ID. N.º: 42); y (3) regiones que atraviesan la membrana de diversas proteínas, por ejemplo, transportadores de azúcares de nucleótidos, que se sabe que están localizados en el aparato de Golgi.

En el primer caso, cuando el péptido director está constituido por diversos elementos (ct, tmd y sr), la adenoteca se diseña de tal forma que estén representadas la ct, la tmd y diversas partes de la región del tallo. Por consiguiente, una subadenoteca de secuencias de péptidos directores preferentemente incluye secuencias de ct, tmd, y/o sr de proteínas unidas a la membrana del RE o aparato de Golgi. En algunos casos, podría ser deseable que la subadenoteca estuviera provista de longitudes variables de secuencias de sr. Esto puede lograrse mediante PCR usando cebadores que se unen al extremo 5' del ADN que codifica la región citosólica y empleando una serie de cebadores opuestos que se unen a diversas partes de la región del tallo.

Todavía otras fuentes de utilidad de secuencias de señalización de direccionamiento incluyen péptidos de señalización de recuperación, por ejemplo los tetrapéptidos HDEL o KDEL, que habitualmente se encuentran en el extremo C de las proteínas que se transportaban de vuelta al RE o al aparato de Golgi. Todavía otras fuentes de secuencias de señalización de direccionamiento incluyen (a) proteínas membrana de tipo II, (b) las enzimas que se enumeran en la Tabla 3, (c) transportadores de nucleótidos que atraviesan la membrana y azúcares que están localizados en el aparato de Golgi, y (d) secuencias a las que se hace referencia en la Tabla 5.

(Tabla pasa a página siguiente)

ES 2 336 790 T3

TABLA 5

Fuentes de secuencias de direccionamiento a compartimientos de utilidad

Gen o Secuencia	Organismo	Función	Locación de producto génico
<i>MNSI</i>	<i>A. nidulans</i>	α -1,2-manosidasa	RE
<i>MNSI</i>	<i>A. niger</i>	α -1,2-manosidasa	RE
<i>MNSI</i>	<i>S. cerevisiae</i>	α -1,2-manosidasa	RE
<i>GLSI</i>	<i>S. cerevisiae</i>	Glucosidasa	RE
<i>GLSI</i>	<i>A. niger</i>	Glucosidasa	RE
<i>GLSI</i>	<i>A. nidulans</i>	Glucosidasa	RE
HDEL en el extremo C	Universal en hongos	Señal de recuperación	RE
<i>SEC12</i>	<i>S. cerevisiae</i>	Proteína vesicular COPII	ER/aparato de Golgi
<i>SEC12</i>	<i>A. niger</i>	Proteína vesicular COPII	ER/aparato de Golgi
<i>OCH1</i>	<i>S. cerevisiae</i>	1,6-manosiltransferasa	Aparato de Golgi (cis)
<i>OCH1</i>	<i>P. pastoris</i>	1,6-manosiltransferasa	Aparato de Golgi (cis)
<i>MNN9</i>	<i>S. cerevisiae</i>	Complejo de 1,6-manosiltransferasa	Golgi
<i>MNN9</i>	<i>A. niger</i>	Sin determinar	Golgi
<i>VAN1</i>	<i>S. cerevisiae</i>	Sin determinar	Golgi
<i>VAN1</i>	<i>A. niger</i>	Sin determinar	Golgi
<i>ANP1</i>	<i>S. cerevisiae</i>	Sin determinar	Golgi
<i>HOCI</i>	<i>S. cerevisiae</i>	Sin determinar	Golgi
<i>MNN10</i>	<i>S. cerevisiae</i>	Sin determinar	Golgi
<i>MNN10</i>	<i>A. niger</i>	Sin determinar	Golgi
<i>MNN11</i>	<i>S. cerevisiae</i>	Sin determinar	Aparato de Golgi (cis)
<i>MNN11</i>	<i>A. niger</i>	Sin determinar	Aparato de Golgi (cis)
<i>MNT1</i>	<i>S. cerevisiae</i>	1,2-manosiltransferasa	Aparato de Golgi (cis, medio)
<i>KTR1</i>	<i>P. pastoris</i>	Sin determinar	Aparato de Golgi (medio)
<i>KRE2</i>	<i>P. pastoris</i>	Sin determinar	Aparato de Golgi (medio)
<i>KTR3</i>	<i>P. pastoris</i>	Sin determinar	Aparato de Golgi (medio)
<i>MNN2</i>	<i>S. cerevisiae</i>	1,2-manosiltransferasa	Aparato de Golgi

			(medio)
<i>KTR1</i>	<i>S. cerevisiae</i>	Sin determinar	Aparato de Golgi (medio)
<i>KTR2</i>	<i>S. cerevisiae</i>	Sin determinar	Aparato de Golgi (medio)
<i>MNN1</i>	<i>S. cerevisiae</i>	1,3-manosiltransferasa	Aparato de Golgi (trans)
<i>MNN6</i>	<i>S. cerevisiae</i>	Fosfomanosiltransferasa	Aparato de Golgi (trans)
2,6 ST	<i>H. sapiens</i>	2,6-sialiltransferasa	Red trans del aparato de Golgi
UDP-Gal T	<i>S. pombe</i>	Transportador de UDP-Gal	Golgi

En cualquier caso, es muy preferente que se seleccionen secuencias de los péptidos directores que sean apropiadas para la actividad o actividades enzimáticas particulares para que funcionen óptimamente dentro de la secuencia de reacciones de glucosilación deseadas. Por ejemplo, para desarrollar un microorganismo modificado con capacidad de sialilación terminal de *N*-glucanos en formación, un proceso que se produce en el aparato de Golgi trans en los seres humanos, es deseable utilizar una subadenoteca de secuencias de péptidos de direccionamiento derivadas de proteínas del aparato de Golgi trans. De forma similar, el recorte de Man₃GlcNAc₂ por una α -1,2-manosidasa para dar Man₅GlcNAc₂ es una etapa temprana en la formación de *N*-glucanos complejos en seres humanos (Figura 1B). Por lo tanto es deseable que esta reacción se produzca en el RE o en el aparato de Golgi cis de un microorganismo huésped genomanipulado. Se usa una subadenoteca que codifica señales de retención en el RE y en el aparato de Golgi cis.

Después se construye una serie de construcciones de proteínas de fusión (*es decir*, una adenoteca de combinaciones) ligando funcionalmente una o una serie de secuencias de péptidos directores a una o una serie de secuencias que codifican dominios catalíticos. Preferentemente, esto se logra ligando en marco una subadenoteca que comprende ADN que codifica secuencias de péptidos directores (anteriormente) con una subadenoteca que comprende ADN que codifica enzimas de glucosilación o sus fragmentos catalíticamente activos (véase más adelante).

La adenoteca resultante comprende genes sintéticos que codifican proteínas de fusión que contienen secuencias de péptidos directores. En algunos casos, es deseable proporcionar una secuencia de un péptido de direccionamiento en el extremo N de una proteína de fusión, o en otros casos en el extremo C. En algunos casos, las secuencias del péptido de direccionamiento pueden estar insertadas en el marco de lectura abierto de una enzima, con la condición de que no se altere la estructura de la proteína de los dominios plegados individualmente. Cada tipo de proteína de fusión se construye (o mediante un proceso dirigido por etapas o de forma semialeatoria) y pueden seleccionarse construcciones óptimas tras la transformación de células huésped y caracterización de patrones de glucosilación en células transformadas usando procedimientos de la invención.

Alteración de la glucosilación de la célula huésped usando construcciones fusionadas de adenotecas de combinaciones

La construcción de una adenoteca de combinaciones preferente se ilustra de forma esquemática en la Figura 2 y se describe en el Ejemplo 4. La construcción fusionada puede estar ligada operablemente a una multitud de vectores, como por ejemplo vectores de expresión notorios en la técnica. Se ensambló una amplia variedad de dichas construcciones fusionadas usando actividades representativas como se muestra en la Tabla 6. Las combinaciones de péptido director y dominios catalíticos puede ensamblarse para usar en el direccionamiento de las actividades de manosidasa, glucosiltransferasa y glucosidasa al RE, aparato de Golgi, y a la red del aparato de Golgi trans de acuerdo con la invención. Sorprendentemente, el mismo dominio catalítico puede tener desde ningún efecto hasta un efecto profundo sobre los patrones de *N*-glucosilación, dependiendo del tipo de péptido director que se use (véase, por ejemplo, la Tabla 7 y el Ejemplo 4).

Construcciones fusionadas de manosidasas

Un ejemplo representativo de una construcción fusionada de manosidasa derivada de una adenoteca de combinaciones de la invención es pFB8, que es un péptido director SEC12(m) de *Saccharomyces* (nucleótidos 988-1296 de SEC12 de SwissProt P11655) ligado en marco a una delección de 187 aminoácidos del extremo N de una α -manosidasa IA de ratón (Genbank AN 6678787). La nomenclatura que se usa en el presente documento, por lo tanto, se refiere al péptido director/región de dominio catalítico de una enzima de glucosilación como SEC12 (m) de *Saccharomyces*/manosidasa IA de ratón, Δ 187. La proteína de fusión codificada se localiza en el RE por medio de la secuencia del péptido director SEC12 a la vez que retiene su actividad de dominio catalítico de manosidasa y es capaz de producir *N*-glucanos *in vivo* que tienen una estructura Man₅GlcNAc₂ (Ejemplo 4; Figuras 6F y 7B).

La construcción fusionada pGC5, MNS1(m) de *Saccharomyces*/manosidasa IB de ratón Δ99, es otro ejemplo de una construcción fusionada que tiene actividad intracelular de recorte por manosidasas (Ejemplo 4; Figuras 50 y 8B). La construcción fusionada pBC18-5 (VAN1(s) de *Saccharomyces*/manosidasa IB de *C. elegans*, Δ80) es otro ejemplo más de una construcción fusionada eficaz capaz de producir *N*-glucanos que tienen una estructura Man₅GlcNAc₂ *in vivo*. Al crear una adenoteca de combinaciones de estas y otras construcciones fusionadas de manosidasas de ese tipo de acuerdo con la invención, un experto puede distinguir y seleccionar aquellas construcciones que tienen una actividad intracelular óptima de recorte de las que tienen una actividad relativamente baja o no la tienen. Los procedimientos que usan adenotecas de combinaciones que se describen en el presente documento son ventajosas porque sólo unas pocas construcciones fusionadas de manosidasas seleccionadas pueden producir un *N*-glucano particularmente deseado *in vivo*.

Además, la actividad de recorte de las manosidasas puede ser específica para una proteína de interés particular. Así, debe entenderse además que no todas las construcciones fusionadas de péptido director y dominio catalítico de las manosidasas puede funcionar igual de bien para producir la glucosilación adecuada sobre una glucoproteína de interés. Por consiguiente, una proteína de interés puede introducirse en una célula huésped transfectada con una adenoteca de combinaciones para identificar una o más construcciones fusionadas que expresan una actividad de manosidasa óptima para la proteína de interés. Un experto en la técnica será capaz de producir y seleccionar construcciones fusionadas óptimas usando la estrategia de adenoteca de combinaciones que se describe en el presente documento.

Es obvio, además, que otras de esas construcciones fusionadas que muestran dominios catalíticos de manosidasa activos localizados (o de forma más general, dominios de cualquier enzima) pueden prepararse usando técnicas como por ejemplo las que se ejemplifican en el Ejemplo 4 y que se describen en el presente documento. Será cuestión de experimentación rutinaria para un experto en la técnica preparar y usar la adenoteca de combinaciones de la presente invención para optimizar, por ejemplo, la producción de Man₅GlcNAc₂ a partir de una adenoteca de construcciones fusionadas en un vector de expresión particular introducido en una célula huésped particular.

Construcciones fusionadas de glucosiltransferasas

De forma similar, se preparó una adenoteca de combinaciones de glucosiltransferasas usando los procedimientos que se describen en el presente documento. Se ensambló una adenoteca de combinaciones de secuencias derivada de actividades de glucosiltransferasa I (GnTI) con péptidos directores y se cribó para determinar la producción eficaz en una célula huésped de un eucariota inferior de una estructura de *N*-glucano GlcNAcMan₅GlcNAc₂ sobre una glucoproteína marcadora. Se identificó una construcción fusionada que se demostró que producía GlcNAcMan₅GlcNAc₂ (pPB1 04), MNN9 (s) de *Saccharomyces*/GnTI humana, Δ38 (Ejemplo 8). Se ensambló una gran variedad de dichas construcciones fusionadas de GnTI (Ejemplo 8, Tabla 10). Pueden ensamblarse otras combinaciones de péptido director/dominios catalíticos de GnTI formando una adenoteca de combinaciones. También es obvio para un experto en la técnica que pueden prepararse otras construcciones fusionadas de ese tipo que muestran actividad de glucosiltransferasa como se demuestra en el Ejemplo 8. Será cuestión de experimentación rutinaria para un experto en la técnica usar el procedimiento de adenoteca de combinaciones que se describe en el presente documento para optimizar la producción de GlcNAcMan₅GlcNAc₂ usando una construcción fusionada seleccionada en un vector de expresión y línea de células huésped particular.

Como se ha indicado anteriormente, para las construcciones fusionadas de manosidasas no todas las construcciones fusionadas de péptido director y dominio catalítico de GnTI funcionarán igual de bien para producir la glucosilación adecuada sobre una glucoproteína de interés como se ha descrito. Sin embargo, un experto en la técnica será capaz de producir y seleccionar construcciones fusionadas óptimas usando una estrategia de adenoteca como se describe en el presente documento. El Ejemplo 8 ilustra una realización preferente de una adenoteca de combinaciones que comprende construcciones fusionadas de péptidos directores y dominios catalíticos de GnTI implicadas en la producción de glucoproteínas con una estructura predominantemente de GlcNAcMan₅GlcNAc₂.

Uso de múltiples construcciones fusionadas para alterar la glucosilación de la célula huésped

En otro ejemplo del uso de los procedimientos y adenotecas de la invención para alterar la glucosilación de las células huésped, una cepa de *P. pastoris* con una delección en *OCH1* que expresa una proteína testigo (K3) se transformó con múltiples construcciones fusionadas aisladas a partir de adenotecas de combinaciones de la invención para convertir *N*-glucanos con alto contenido en manosa en *N*-glucanos humanoides (Ejemplo 8). Primero, la construcción fusionada de manosidasa pFB8 (*SEC12* (m) de *Saccharomyces*/manosidasa IA de ratón, Δ187) se transformó en una cepa de *P. pastoris* que carecía de la actividad de manosiltransferasas de iniciación en 1,6 (es decir, delección en *och1*; Ejemplo 1). Segundo, se construyó pPB103 que comprendía un gen MNN2-2 de *K. lactis* (n.º de acceso de Genbank AF106080) que codificaba un transportador de UDP-GlcNAc para aumentar más la producción de GlcNAcMan₅GlcNAc₂. La adición del transportador de UDP-GlcNAc aumentó la producción de GlcNAcMan₅GlcNAc₂ significativamente en la cepa de *P. pastoris* como se ilustra en la Figura 10B. Tercero, se introdujo pPB104 que comprende MNN9 (s) de *Saccharomyces*/GnTI humana, Δ38 en la cepa. Esta cepa de *P. pastoris* se denominó “PBP-3”. (Véase la Figura 36).

Un experto en la técnica entenderá que las células huésped como por ejemplo las cepas de levadura que se describen más arriba pueden transformarse secuencialmente y/o co-transformarse con uno o más vectores de expresión. También debe entenderse que el orden de transformación no es particularmente relevante en la producción de la glucoproteína de interés. El experto reconoce las modificaciones rutinarias de los procedimientos que se describen en el presente documento pueden dar mejores resultados en la producción de la glucoproteína de interés.

La importancia del uso de una secuencia de péptido director particular con una secuencia de dominio catalítico particular se hace fácilmente obvia a partir de los experimentos que se describen en el presente documento. La adenoteca de combinaciones proporciona una herramienta para construir fusiones de enzimas que están implicadas en modificar los *N*-glucanos sobre una glucoproteína de interés, que es especialmente útil en la producción de glucoproteínas humanoides. (Sin embargo, puede seleccionarse cualquier fusión de enzimas usando las adenotecas y los procedimientos que se describen en el presente documento.) Los transformantes deseados que expresan α -1,2-manosidasa activa dirigida producen K3 con *N*-glucanos de la estructura Man₅GlcNAc₂ como se muestra en las Figuras 50 y 5E. Esto confiere una menor masa molecular al glucano escindido comparado con el K3 de la cepa con delección *OCH1*, según se detectó mediante espectrometría de masas MALDI-TOF en la Figura 5C.

De forma similar, se usó la misma estrategia para producir otra glucoproteína segregada: IFN- β que comprende predominantemente Man₅GlcNAc₂. El Man₅GlcNAc₂ se eliminó mediante digestión con PNGasa (Papac *et al.* (1998) Glycobiology 8:445-454) y se sometió a MALDI-TOF como se muestra en las Figuras 6A - 6F. Un único pico prominente en 1254 (*m/z*) confirma la producción de Man₅GlcNAc₂ sobre IFN- β en las Figuras 6E (pGC5) (MNS1(m) de *Saccharomyces*/manosidasa IB de ratón, Δ 99) y 6F (pFB8) (*SEC12* (m) de *Saccharomyces*/manosidasa IA de ratón, Δ 187). Además, en la cepa PBP-3 de *P. pastoris* que comprende pFB8 (*SEC12*(m) de *Saccharomyces*/manosidasa IA de ratón, Δ 187), pPB104 (*MNN9* (s) de *Saccharomyces*/GnTI humana, Δ 38) y pPB103 (gen *MNN2-2* de *K. lactis*), se detectó el *N*-glucano híbrido GlcNAcMan₅GlcNAc₂ [b] por MALDI-TOF (Figura 10).

Después de identificar a los transformantes con un elevado grado de recorte de manosas, se realizaron experimentos adicionales para confirmar que la actividad de manosidasa (recorte) se producía *in vivo* y no era predominantemente resultado de la actividad extracelular en el medio de cultivo (Ejemplo 6; Figuras 7-9).

Aunque la presente invención se ejemplifica usando un organismo huésped *P. pastoris*, los expertos en la técnica entenderán que pueden alterarse otras células huésped eucariotas, que incluyen otras especies de huéspedes de levaduras y fúngicos, como se describe en el presente documento para producir glucoproteínas humanoides. Las técnicas que se describen en el presente documento para la identificación y alteración de los genes de glucosilación indeseables de las células huésped, por ejemplo *OCH1*, se entiende que son aplicables a estos y/u otros genes homólogos o funcionalmente relacionados en otras células huésped eucariotas como por ejemplo otras cepas de levaduras y fúngicas. Como se describe en el Ejemplo 9, se borraron los genes *och1 mnn1* de *K. lactis* para diseñar una célula huésped que produjera *N*-glucanos que fueran convertidos completamente a Man₅GlcNAc₂ por 1,2-manosidasa (Figura 12C).

El gen *MNN1* se clonó a partir de *K. lactis* como se describe en el Ejemplo 9. Las secuencias de ácido nucleico y las de aminoácidos deducidos del gen *MNN1* de *K. lactis* se muestran en las SEC. ID. N.º: 43 y 44, respectivamente. Usando cebadores específicos de gen, se preparó una construcción para eliminar el gen *MNN1* del genoma de *K. lactis* (Ejemplo 9). Las células huésped sin actividades de *och1* y *mnn1* producen *N*-glucanos que tienen una estructura de carbohidrato Man₉GlcNAc₂ (véase, por ejemplo, Figura 12B). Dichas células huésped pueden genomanipularse además usando, por ejemplo, procedimientos de la invención y adenotecas que se describen en el presente documento, para producir glucoproteínas de tipo mamífero o humanoides.

Así, en el presente documento se describe una molécula de ácido nucleico aislada que tiene una secuencia de ácidos nucleicos que comprende o constituida por al menos 45, preferentemente al menos 50, más preferentemente al menos 60 y lo más preferentemente 75 o más restos nucleotídicos del gen *MNN1* de *K. lactis* (SEC. ID. N.º: 43), y sus homólogos, variantes y derivados. También se describen moléculas de ácido nucleico que se hibridan en condiciones rigurosas a las moléculas de ácido nucleico descritas anteriormente. De forma similar, en el presente documento se describen polipéptidos aislados (que incluyen muteínas, variantes de alelos, fragmentos, derivados, y análogos) codificados por las moléculas de ácido nucleico. También se describen vectores, que incluyen vectores de expresión, que comprenden dichas moléculas de ácido nucleico, como se describe adicionalmente en el presente documento. De forma similar, se describen células huésped transformadas con dichas moléculas de ácido nucleico o vectores de la invención.

Otro aspecto de la presente invención así se refiere a una cepa huésped de hongo unicelular o multicelular que expresa glucoproteínas que comprenden *N*-glucanos modificados que se parecen a los que forman las células humanas.

Preferentemente, se dirigen las enzimas de glucosilación o dominios catalíticos y similares a una localización subcelular de la ruta secretora de las células huésped donde pueden funcionar, y preferentemente, a donde han sido diseñadas o seleccionadas para funcionar del modo más eficaz.

También pueden genomanipularse células vegetales y de insectos para alterar la glucosilación de las proteínas expresadas usando la adenoteca de combinaciones y los procedimientos que se describen en el presente documento. Además, también puede modificarse la glucosilación en las células de mamífero, que incluyen células humanas, usando la adenoteca de combinaciones y procedimientos que se describen en el presente documento. Puede ser posible, por ejemplo, optimizar una actividad enzimática particular o modificar de otro modo las proporciones relativas de diversos *N*-glucanos formados en una célula huésped de mamífero usando la adenoteca de combinaciones y los procedimientos que se describen en el presente documento.

Los ejemplos de modificaciones a la glucosilación que pueden verse afectadas usando dicho procedimiento son: (1) diseñar una célula huésped eucariota para que recorte los restos de manosa de Man₈GlcNAc₂ proporcionando un *N*-glucano Man₅GlcNAc₂; (2) diseñar células huésped eucariotas para que añadan un resto *N*-acetilglucosamina (GlcNAc) a Man₅GlcNAc₂ mediante la acción de GlcNAc transferasa I; (3) diseñar una célula huésped eucariota para que exprese funcionalmente una enzima como por ejemplo una *N*-acetilglucosaminil transferasa (GnTI, GnTII, GnT-

TII, GnTIV, GnTV, GnTVI), manosidasa II, fucosiltransferasa (FT), galactosil transferasa (GaiT) o una sialiltransferasa (ST).

Al repetir el procedimiento, pueden genomanipularse rutas de glucosilación cada vez más complejas en un huésped diana, como por ejemplo un microorganismo eucariota inferior. En una realización preferida, el organismo huésped de la invención se transforma dos o más veces con adenotecas que incluyen secuencias que codifican actividades de glucosilación. La selección de los fenotipos deseados puede realizarse después de cada ronda de transformación o de forma alternativa después de que se hayan producido varias transformaciones. De este modo pueden genomanipularse rápidamente rutas de glucosilación complejas.

Reacciones de glucosilación secuenciales

En una realización preferente, dichas adenotecas de péptidos directores y dominios catalíticos se diseñan para que incorporen información existente sobre la naturaleza secuencial de las reacciones de glucosilación en los eucariotas superiores. Las reacciones que se sabe que se producen temprano en el transcurso del procesamiento de la glucoproteína requieren que se dirijan enzimas que catalizan dichas reacciones a una parte temprana del aparato de Golgi o del RE. Por ejemplo, el recorte de $\text{Man}_8\text{GlcNAc}_2$ a $\text{Man}_5\text{GlcNAc}_2$ por las manosidasas es una etapa temprana en la formación de *N*-glucanos complejos (Figuras 1B y 35A). Debido a que el procesamiento de las proteínas se inicia en el RE y después pasa a través del aparato de Golgi cis, medio y trans, es deseable que esta reacción se produzca en el RE o en el aparato de Golgi cis. Por lo tanto, cuando se diseña una adenoteca para la localización de manosidasa I, por ejemplo, se intentaría emparejar las señales de direccionamiento al RE y al aparato de Golgi cis con el dominio catalítico de manosidasa I.

Generación de diversidad adicional en las secuencias

El procedimiento que se describe en el presente documento es más eficaz cuando un ácido nucleico, por ejemplo, una adenoteca que transforma a un huésped contiene una gran diversidad de secuencias, aumentando así la probabilidad de que al menos un transformante muestre el fenotipo deseado. Las mutaciones de un único aminoácido, por ejemplo, pueden alterar drásticamente la actividad de las enzimas de procesamiento de las glucoproteínas (Romero *et al.* (2000) *J. Biol. Chem.* 275(15):11071-4). Por consiguiente, antes de la transformación, puede someterse una adenoteca o un constituyente de una subadenoteca a una o más técnicas para generar una diversidad adicional en las secuencias. Por ejemplo, puede realizarse una o más tandas de barajado génico, PCR con tendencia a errores, mutagénesis *in vitro* u otros procedimientos para generar diversidad en las secuencias, para obtener una mayor diversidad de secuencias dentro del conjunto de construcciones fusionadas.

Secuencias de control de la expresión

Además de las secuencias con marco de lectura abierto que se describen anteriormente, de forma general es preferible proporcionar a cada construcción de la adenoteca secuencias de control de la expresión, como por ejemplo promotores, terminadores de la transcripción, potenciadores, sitios de unión de ribosomas, y otras secuencias funcionales que puedan ser necesarios para asegurar una transcripción y traducción eficaces de las proteínas fusionadas al ser introducidas las construcciones fusionadas en el organismo huésped.

Los componentes adecuados para los vectores, por ejemplo, los marcadores seleccionables, las secuencias de control de la expresión (por ejemplo, promotor, potenciadores, terminadores y similares) y, opcionalmente secuencias necesarias para la replicación autónoma en una célula huésped, se seleccionan para en función de la célula huésped particular que se elija. Los criterios de selección para los componentes adecuados de los vectores para usar en un mamífero o una célula huésped particular de un eucariota inferior son rutinarios. Las células huésped eucariotas inferiores preferentes de la invención incluyen *Pichia pastoris*, *Pichia finlandica*, *Pichia trehalophila*, *Pichia koclamae*, *Pichia membranaefaciens*, *Pichia opuntiae*, *Pichia thermotolerans*, *Pichia salictaria*, *Pichia guercuum*, *Pichia pyperii*, *Pichia stiptis*, *Pichia methanolica*, *Pichia sp.*, *Saccharomyces cerevisiae*, *Saccharomyces sp.*, *Hansenula polymorpha*, *Kluyveromyces sp.*, *Kluyveromyces lactis*, *Candida albicans*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus otyzae*, *Trichoderma reesei*, *Chrysosporium lucknowense*, *Fusarium sp.*, *Fusarium gramineum*, *Fusarium venenatum* y *Neurospora crassa*. Cuando el huésped es *Pichia pastoris*, los promotores adecuados incluyen, por ejemplo, los promotores AOX1, AOX2, GAPDH, y P40.

Marcadores seleccionables

También es preferible dotar a cada construcción de al menos un marcador seleccionable, como por ejemplo un gen que confiera resistencia a fármacos o para complementar una carencia metabólica en el huésped. La presencia del marcador es de utilidad en la posterior selección de transformantes; por ejemplo, en las levaduras, pueden usarse los genes *URA3*, *HIS4*, *SUC2*, *G418*, *BLA*, o SH BLE. Se conoce una multitud de marcadores seleccionables y están disponibles para usar en células huésped de levaduras, hongos, plantas, insectos, mamífero y otras células huésped eucariotas.

Transformación

La adenoteca se transforma después en el organismo huésped. En las levaduras, puede usarse cualquier procedimiento conveniente de transferencia de ADN, como por ejemplo electroporación, el procedimiento con cloruro de litio, o el procedimiento con esferoplastos. En las células de hongos filamentosos y vegetales, los procedimientos convencionales incluyen bombardeo con partículas, electroporación y transformación mediada con agrobacterium. Para

producir una cepa estable adecuada para el cultivo a alta densidad (por ejemplo, fermentación en levaduras), es deseable integrar las construcciones de las adenotecas en el cromosoma del huésped. Preferentemente, la integración se produce mediante recombinación homóloga, usando técnicas notorias en la técnica. Por ejemplo, los elementos de la adenoteca se dotan de secuencias flanqueantes homólogas a secuencias del organismo huésped. De este modo, la integración se produce en un sitio definido en el genoma del huésped, sin alteración de genes deseables o esenciales.

Es especialmente preferente que el ADN de la adenoteca se integre en el sitio de un gen no deseado en un cromosoma de un huésped, logrando la alteración o delección del gen. Por ejemplo, la integración en los sitios de los genes *OCHI*, *MNN1*, o *MNN4* permite la expresión del ADN deseado de la adenoteca a la vez que se evita la expresión de enzimas implicadas en la hipermansilación de glucoproteínas en las levaduras. El ADN de la adenoteca también puede introducirse en el huésped mediante una molécula de ácido nucleico, plásmido, vector (por ejemplo, vector vírico o retrovírico), cromosoma, y puede introducirse como molécula de ácido nucleico autónoma o mediante integración homóloga o aleatoria en el genoma del huésped. En cualquier caso, generalmente es deseable incluir con cada construcción de ADN de la adenoteca al menos un gen marcador seleccionable para permitir una rápida selección de organismos huésped que han sido transformados de forma estable. Los genes marcadores reciclables, como por ejemplo *ura3*, que pueden seleccionarse para determinar su presencia o ausencia, son especialmente adecuados.

Procesos de cribado y selección

Después de la transformación de la cepa huésped con la adenoteca, se seleccionan los transformantes que presentan un fenotipo de glucosilación deseado. La selección puede realizarse en una única etapa o mediante una serie de etapas de enriquecimiento fenotípico y/o depleción usando cualquiera de una variedad de ensayos o procedimientos de detección. La caracterización fenotípica puede realizarse manualmente o usando equipo de cribado de alto rendimiento. Habitualmente un microorganismo huésped presenta *N*-glucanos de proteínas sobre la superficie celular, donde están localizadas diversas glucoproteínas.

Se pueden cribar las células que tengan la concentración más elevada del GlcNAc terminal sobre la superficie celular, por ejemplo, o las células que segreguen la proteína con el mayor contenido en GlcNAc terminal. Dicho cribado puede basarse en un procedimiento visual, como por ejemplo un procedimiento de tinción, en la capacidad de unirse a anticuerpos o lectinas que se unen a GlcNAc terminal conjugados a un marcador (dichas lectinas pueden obtenerse en E. Y. Laboratories Inc., San Mateo, CA), en la menor capacidad de lectinas específicas de unirse a restos terminales de manosa, en la capacidad de incorporar un azúcar marcado radiactivamente *in vitro*, en la unión alterada a tintes o a superficies cargadas, o puede lograrse usando un dispositivo de Selección celular asistida por fluorescencia (FACS) junto con una lectina o anticuerpo marcado con un fluoróforo (Guillen *et al.* (1998) Proc. Natl. Acad. Sci. USA 95(14):7888-7892).

Por consiguiente, pueden seleccionarse células intactas para determinar un fenotipo de glucosilación deseado exponiendo las células a una lectina o anticuerpo que se une de forma específica al *N*-glucano deseado. Hay disponible comercialmente una amplia variedad de lectinas específicas de oligosacáridos (por ejemplo en EY Laboratories, San Mateo, CA). De forma alternativa, los anticuerpos contra *N*-glucanos humanos o animales específicos están disponibles comercialmente o pueden producirse usando técnicas estándar. Una lectina o anticuerpo apropiado puede conjugarse con una molécula testigo, como por ejemplo un cromóforo, fluoróforo, radioisótopo, o una enzima que tiene un sustrato cromógeno (Guillen *et al.*, 1998. Proc. Natl. Acad. Sci. USA 95(14):7888-7892).

El cribado puede realizarse después usando procedimientos analíticos como por ejemplo espectrofotometría, fluorimetría, selección de células activada por fluorescencia o recuento de centelleo. En otros casos, puede ser necesario analizar glucoproteínas o *N*-glucanos aislados de células transformadas. El aislamiento de proteínas puede realizarse mediante técnicas conocidas en la técnica. Preferentemente, se segrega una proteína testigo al medio y se purifica mediante cromatografía por afinidad (por ejemplo afinidad a Ni, o cromatografía por afinidad de glutatona-S-transferasa). En los casos en los que se prefiere un *N*-glucano aislado, puede usarse una enzima como por ejemplo endo-p-*N*-acetilglucosaminidasa (Genzyme Co., Boston, MA; New England Biolabs, Beverly, MA) para escindir los *N*-glucanos de las glucoproteínas. Las proteínas o *N*-glucanos aislados pueden analizarse después mediante cromatografía de líquidos (por ejemplo HPLC), espectroscopia de masas u otros medios adecuados. La patente de Estados Unidos n.º 5.595.900 enseña varios procedimientos por los que pueden identificarse las células con las estructuras extracelulares de carbohidratos deseadas. Preferentemente, se usa espectrometría de masas MALDI-TOF para analizar los *N*-glucanos escindidos.

Antes de la selección de un transformante deseado, puede ser deseable eliminar de la población de células transformadas las que tienen fenotipos no deseados. Por ejemplo, cuando se usa el procedimiento para diseñar una actividad de manosidasa funcional en las células, los transformantes deseados tendrán niveles menores de manosa en la glucoproteína celular. La exposición de la población transformada a un radioisótopo letal de manosa en el medio elimina de la población los transformantes que tienen los fenotipos no deseados, es decir los que tienen niveles altos de manosa incorporada (Huffaker y Robbins (1983) Proc Natl Acad Sci USA. 80(24):7466-70). De forma alternativa, puede usarse una lectina o anticuerpo citotóxicos, dirigidos contra un *N*-glucano indeseable, para eliminar los fenotipos no deseados de la población transformada (por ejemplo, Stanley y Siminovitch (1977) Somatic Cell Genet 3(4):391-405). La patente de Estados Unidos n.º 5.595.900 enseña varios procedimientos por los que pueden identificarse las células con las estructuras extracelulares de carbohidratos deseadas. Realizar esta estrategia repetidamente permite diseñar secuencialmente glucanos más y más complejos en eucariotas inferiores.

Para detectar células huésped que tienen sobre su superficie un alto grado de *N*-glucanos intermedios humanos GlcNAcMan₅GlcNAc₂, por ejemplo, se pueden seleccionar los transformantes que permiten la transferencia más eficaz de GlcNAc mediante GlcNAc Transferasa de UDP-GlcNAc en un ensayo celular *in vitro*. Este cribado puede realizarse cultivando células que alojen la adenoteca transformada bajo presión selectiva en una placa de agar y transfiriendo las colonias individuales a una placa de microtitulación de 96 pocillos. Después de cultivar las células, éstas se centrifugan, se resuspenden en tampón, y después de la adición de UDP-GlcNAc y GnTII, se determina la liberación de UDP o por HPLC o por un ensayo ligado a enzimas para UDP. De forma alternativa, se puede usar UDP-GlcNAc y GnTII marcados radioactivamente, lavar las células y después buscar la liberación de GlcNAc radiactivo mediante *N*-actilglucosaminidasa. Todo esto puede realizarse de forma manual o automatizada usando equipo de cribado de alto rendimiento. Los transformantes que liberan más UDP, en el primer ensayo, o más GlcNAc marcado radiactivamente en el segundo ensayo, se espera que tenga un mayor grado de GlcNAcMan₅GlcNAc₂ en su superficie y así constituyen el fenotipo deseado. También pueden adaptarse ensayos similares para buscar los *N*-glucanos de las proteínas segregadas.

De forma alternativa, se puede usar cualquier otra selección adecuada como por ejemplo un ensayo de unión a lectina que es capaz de revelar patrones de glucosilación alterados sobre la superficie de las células transformadas. En este caso, la menor unión de las lectinas específicas a las manosas terminales puede ser una herramienta de selección adecuada. La lectina de *Galantus nivalis* se une específicamente a α -1,3 manosa terminal, que se espera que se reduzca si hay presente una actividad de manosidasa II suficiente en el aparato de Golgi. También puede enriquecerse en los transformantes deseados realizando una etapa de separación cromatográfica que permita eliminar las células que contienen un gran contenido en manosa terminal. Esta etapa de separación se realizaría con una columna de lectina que se una específicamente a las células con un alto contenido en manosa terminal (por ejemplo, lectina de *Galantus nivalis* unida a agarosa, Sigma, St. Louis, MO) y no a las que tienen un bajo contenido en manosa terminal.

Además, se pueden crear directamente dichas construcciones de proteínas de fusión, cuando haya información adicional sobre la localización de enzimas que modifican carbohidrato activo en diferentes huéspedes eucariotas inferiores en la bibliografía científica. Por ejemplo, es sabido que β 1,4-GaIT puede fusionarse al dominio de membrana de MNT, una manosiltransferasa de *S. cerevisiae*, y localizarse en el aparato de Golgi a la vez que mantiene su actividad catalítica (Schwientek *et al.* (1995) J. Biol. Chem. 270(10):5483-9). Si *S. cerevisiae* o un organismo relacionado es el huésped a genomanipular, se puede incorporar directamente dichos hallazgos en la estrategia global para obtener *N*-glucanos complejos a partir de dicho huésped. Se han identificado varias de tales fragmentos génicos en *P. pastoris* que están relacionados con las glucosiltransferasas en *S. cerevisiae* y así podrían usarse para el fin.

Sitos de integración

Dado que un objetivo final de esta tentativa de genomanipulación es una cepa con producción potente de proteínas que sea capaz de funcionar bien en un proceso de fermentación industrial, la integración de los múltiples genes en el cromosoma del huésped (por ejemplo, fúngico) preferentemente supone una planificación cuidadosa. La cepa genomanipulada podría tener que transformarse con un abanico de genes diferentes, y estos genes tendrán que ser transformados de forma estable para asegurar que se mantiene la actividad deseada durante todo el proceso de fermentación. Como se describe en el presente documento, la combinación de las diversas actividades enzimáticas deseadas debería genomanipularse en el huésped fúngico para la expresión de las proteínas, por ejemplo, sialiltransferasas, manosidasas, fucosiltransferasas, galactosiltransferasas, glucosiltransferasas, GlcNAc transferasas, transportadores específicos del RE y del aparato de Golgi (por ejemplo transportadores simport y antiport para UDP-galactosa y otros precursores), otras enzimas implicadas en el procesamiento de los oligosacáridos, y enzimas implicadas en la síntesis de precursores de oligosacáridos activados como por ejemplo UDP-galactosa, CMP-*N*-ácido acetilneuramínico. En la Tabla 6 se muestran ejemplos de procedimientos preferentes para modificar la glucosilación en una célula huésped de un eucariota inferior, como por ejemplo *Pichia pastoris*.

TABLA 6

Algunas realizaciones preferentes para modificar la glucosilación en un microorganismo eucariota inferior

Estructura deseada	Actividades catalíticas adecuadas	Fuentes adecuadas de secuencias de localización	Deleciones génicas adecuadas	Transportadores y/o fosfatasas adecuadas
Man ₅ GlcNAc ₂	α -1,2-manosidasa (murina, humana, <i>Bacillus sp.</i> , <i>A. nidulans</i>)	Mns1 (extremo N, <i>S. cerevisiae</i>) <i>Och1</i> (extremo N, <i>S. cerevisiae</i> , <i>P.</i>	<i>OCH1</i> MNN4 MNN6	Ninguno

ES 2 336 790 T3

5			<i>pastoris</i>) Ktr1 <i>Mnn9</i> Mnt1 (<i>S. cerevisiae</i>) KDEL, HDEL (extremo C)		
10			Och1 (extremo N, <i>S. cerevisiae</i> , <i>P. pastoris</i>) KTR1 (extremo N) KDEL, HDEL (extremo C) Mnn1 (extremo N, <i>S. cerevisiae</i>) Mnt1 (extremo N, <i>S. cerevisiae</i>) GDPasa (extremo N, <i>S. cerevisiae</i>)	<i>OCH1</i> MNN4 MNN6	Transportador de UDP-GlcNAc (humano, murino, <i>K. lactis</i>) UDPasa (humana)
15					
20		GlcNAc Transferasa I, (humana, murina, de rata, etc.)			
25	GlcNAcMan ₅ GlcNAc ₂				
30					
35					
40			Ktr1 Mnn1 (extremo N, <i>S. cerevisiae</i>) Mnt1 (extremo N, <i>S. cerevisiae</i>) Kre2/Mnt1 (<i>S. cerevisiae</i>) Kre2 (<i>P. pastoris</i>) Ktr1 (<i>S. cerevisiae</i>) Ktr1 (<i>P. pastoris</i>) Mnn1 (<i>S. cerevisiae</i>)	<i>OCH1</i> MNN4 MNN6	Transportador de UDP-GlcNAc (humano, murino, <i>K. lactis</i>) UDPasa (humana)
45					
50	GlcNAcMan ₃ GlcNAc ₂	Manosidasa II			
55					
60					

ES 2 336 790 T3

5			Mnn1 (extremo N, <i>S. cerevisiae</i>)		
10			Mnt1 (extremo N, <i>S. cerevisiae</i>)		
15	GlcNAc ₍₂₋₄₎ Man ₃ GlcNAc ₂	GlcNAc Transferasa II, III, IV, V (humana, murina)	Kre2/Mnt1 (<i>S. cerevisiae</i>) Kre2 (<i>P. pastoris</i>) Ktr1 (<i>S. cerevisiae</i>) Ktr1 (<i>S. pastoris</i>) Mnn1 (<i>S. cerevisiae</i>)	OCH1 MNN4 MNN6	Transportador de UDP-GlcNAc (humano, murino, <i>K. lactis</i>) UDPasa (humana)
20					
25					
30			Mnn1 (extremo N, <i>S. cerevisiae</i>)		
35			Mnt1 (extremo N, <i>S. cerevisiae</i>)		
40	Gal ₍₁₋₄₎ GlcNAc ₍₂₋₄₎ Man ₃ GlcNAc ₂	β-1,4-Galactosil transferasa (humana)	Kre2/Mnt1 (<i>S. cerevisiae</i>) Kre2 (<i>P. pastoris</i>) Ktr1 (<i>S. cerevisiae</i>) Ktr1 (<i>P. pastoris</i>) Mnn1 (<i>S. cerevisiae</i>)	OCH1 MNN4 MNN6	Transportador de UDP-Galactosa (humano, <i>S. pombe</i>)
45					
50					
55	NANA ₍₁₋₄₎ -Gal ₍₁₋₄₎ GlcNAc ₍₂₋₄₎ - Man ₃ GlcNAc ₂	α-2,6-sialiltransferasa (humano) (α 2,3-sialiltransferasa	KTR1 MNN1 (extremo N, <i>S. cerevisiae</i>) MNT1 (extremo N, <i>S. cerevisiae</i>) Kre2/Mnt1 (<i>S.</i>	OCH1 MNN4 MNN6	Transportador de CMP-ácido siálico (humana)
60					

		<i>cerevisiae</i>) Kre2 (<i>P.</i> <i>pastoris</i>) Ktr1 (<i>S. cerevisiae</i>) Ktr1 (<i>P.</i> <i>pastoris</i>) MNN1 (<i>S.</i> <i>cerevisiae</i>)		
--	--	---	--	--

Dado que cualquier estrategia para diseñar la formación de *N*-glucanos complejos en una célula huésped como por ejemplo un eucariota inferior supone tanto eliminar como añadir actividades de glucosiltransferasas particulares, un esquema global trataría de coordinar ambos requisitos. Los genes que codifican enzimas que son indeseables sirven como sitios de integración potenciales para los genes que son deseables. Por ejemplo, la actividad de 1,6 manosiltransferasa es una marca de glucosilación en muchos eucariotas inferiores conocidos. El gen que codifica alfa-1,6 manosiltransferasa (*OCHI*) ha sido clonado a partir de *S. cerevisiae* y las mutaciones en el gen dan lugar a un fenotipo viable con manosilación reducida. El locus génico que codifica la actividad de alfa-1,6 manosiltransferasa por lo tanto es una diana principal para la integración de genes que codifican la actividad de glucosiltransferasa. De forma similar, se puede elegir una variedad de otros sitios de integración cromosómica que, basándose en un evento de alteración génica en ese locus, pueda esperarse que: (1) mejoren la capacidad de glucosiltransferasa de glucosilar de un modo más de tipo humano, (2) mejorar la capacidad de las células de segregar proteínas, (3) reducir la proteólisis de proteínas exógenas y (4) mejorar otras características del procedimiento que faciliten el procedimiento de purificación o de fermentación en sí.

Glucoproteínas diana

Los procedimientos que se describen en el presente documento son de utilidad para producir glucoproteínas, especialmente glucoproteínas que se usan terapéuticamente en los seres humanos. Las glucoproteínas que tienen glucosilaciones específicas pueden ser de utilidad especial, por ejemplo en el direccionamiento de proteínas terapéuticas. Por ejemplo, se ha demostrado que manosa-6-fosfato dirige las proteínas al lisosoma, lo que puede ser esencial para el funcionamiento adecuado de varias enzimas relacionadas con trastornos de almacenamiento en los lisosomas como por ejemplo las enfermedades de Gaucher, Hunter, Hurler, Scheie, Fabry y Tay-Sachs, por mencionar sólo unas pocas. Del mismo modo, la adición de uno o más restos de ácido siálico a una cadena lateral de glucano puede aumentar la vida de una glucoproteína terapéutica *in vivo* después de la administración. Por consiguiente, las células huésped (por ejemplo, de eucariotas inferiores o de mamíferos) pueden diseñarse genéticamente para aumentar la cantidad de ácido siálico terminal en las glucoproteínas expresadas en las células. De forma alternativa, el ácido siálico puede conjugarse a la proteína de interés *in vitro* antes de la administración usando una ácido siálico transferasa y un sustrato apropiado. Los cambios en la composición del medio de cultivo puede emplearse además de la expresión de las actividades enzimáticas implicadas en la glucosilación de tipo humana para producir glucoproteínas que se parezcan más a las formas humanas (Weikert *et al.* (1999) *Nature Biotechnology* 17, 1116-1121; Werner *et al.* (1998) *Arzneimittelforschung* 48(8):870-880; Andersen y Goochee (1994) *Cur. Opin. Biotechnol.* 5:546-549; Yang y Butler (2000) *Biotechnol. Bioengin.* 68(4):370-380). Se ha demostrado que las modificaciones de glucanos específicos en los anticuerpos monoclonales (por ejemplo la adición de una GlcNAc diseccionadora) mejoran la citotoxicidad celular dependiente de anticuerpos (Umana *et al.* (1999) *Nat. Biotechnol.* 17(2):176-80), que puede ser deseable para la producción de anticuerpos u otras proteínas terapéuticas.

Las proteínas terapéuticas habitualmente se administran mediante inyección, por vía oral, pulmonar o por otros medios. Los ejemplos de glucoproteínas diana adecuadas que pueden producirse de acuerdo con la invención incluyen, sin limitación: eritropoyetina, citocinas como por ejemplo interferón- α , interferón- β , interferón- γ , interferón- ω , y CSF granulocítico, factores de coagulación como por ejemplo factor VIII, factor IX, y proteína C humana, cadena α del receptor soluble de IgE, IgG, fragmentos de IgG, IgM, interleucinas, urocinasa, quimasa, y inhibidor de tripsina de urea, proteína que se une a IGF, factor de crecimiento epidérmico, factor liberador de hormona del crecimiento, proteína de fusión de anexina V, angiostatina, factor de crecimiento endotelial vascular-2, factor inhibidor progenitor mieloide-1, osteoprotegerina, antitripsina α -1, ADNasa II, α -fetoproteínas, AAT, rhTBP-1 (onerecept también conocida como proteína de unión a TNF 1), TACI-Ig (activador transmembrana y modulador del calcio e interacto con los ligandos de ciclofilina), FSH (hormona estimulante de folículos), GM-CSF, GLP-1 con o sin FC (proteína glucagonoide 1), agonista del receptor IL-1, sTNRr (enbrel, también conocida como fusión de TNF soluble y receptor Fc), ATIII, rhTrombina, glucocerebrosidasa y CTLA4-Ig (antígeno asociado a linfocitos T citotóxicos 4-Ig).

Expresión de GnT-III para potenciar la funcionalidad de los anticuerpos

La adición de restos de *N*-acetilglucosamina a la estructura de $\text{GlcNAcMan}_3\text{GlcNAc}_2$ mediante las *N*-acetilglucosaminiltransferasas II y III proporciona lo que se denomina *N*-glucano $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ diseccionado (Figura 15). Esta estructura ha sido implicada en una mayor citotoxicidad celular dependiente de anticuerpos (ADCC). (1999) Nat. Biotechnol. 17(2):176-80). El rediseño de glucoformas de inmunoglobulinas expresadas por las células de mamífero es una tarea tediosa e incómoda. Especialmente en el caso de GnTIII, donde la hiperexpresión de esta enzima se ha implicado en la inhibición del crecimiento, y han tenido que emplearse procedimientos que suponen la expresión génica regulada (inducible) para producir inmunoglobulinas con *N*-glucanos diseccionados (Umana *et al.* (1999) Biotechnol Bloeng. 65(5):542-9; Umana *et al.* (1999) Nat. Biotechnol. 17(2):176-80); Umana *et al.* documento WO 03011878; patente de Estados Unidos n.º 6.602.684.

Por consiguiente, en otra realización, la invención proporciona sistemas y procedimientos para producir *N*-glucanos diseccionados que tienen *N*-acetilglucosamina (GlcNAc) diseccionadora en una estructura nuclear de trimanosa o pentammanosa. En una realización preferente, la invención proporciona un sistema y procedimiento para producir inmunoglobulinas que tienen *N*-glucanos diseccionados. Los sistemas y procedimientos que se describen en el presente documento no adolecerán de los problemas previos, por ejemplo, la citotoxicidad asociada a la hiperexpresión de GnTIII o ADCC, dado que las células huésped de la invención están genomanipuladas y se seleccionan para que sean células viables y preferentemente fuertes que produzcan *N*-glucanos que tienen glucoformas humanoides sustancialmente modificadas como por ejemplo $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$. Así, la adición de una *N*-acetilglucosamina diseccionadora a una célula huésped de la invención tendrá un efecto insignificante sobre el fenotipo de desarrollo o sobre la viabilidad de esas células huésped.

Además, las investigaciones de otros han demostrado que no existe correlación lineal entre los niveles de expresión de GnTIII y el grado de ADCC. Umana *et al.* (1999) Nature Biotechnol. 17:176-80. Así, encontrar el nivel de expresión óptimo en las células de mamífero y mantenerlo durante un procedimiento de fermentación autorizado por la FDA parece ser un reto. Sin embargo, en las células de la invención, como por ejemplo las células fúngicas, encontrar un promotor de potencia adecuada para establecer un nivel de expresión de GnTIII robusto, fiable y óptimo es una tarea comparativamente fácil para un experto en la técnica.

Una célula huésped como por ejemplo una cepa de levadura capaz de producir glucoproteínas con *N*-glucanos diseccionadores se genomanipula introduciendo en las células huésped de la invención una actividad de GnTIII (Ejemplo 12). Preferentemente, la célula huésped se transforma con un ácido nucleico que codifica GnTIII (véase, por ejemplo, la Figura 24) o un dominio de la misma que tiene actividad enzimática, opcionalmente fusionado a un péptido de señal director celular heterólogo (por ejemplo, usando las adenotecas y procedimientos asociados que se describen en el presente documento). Las células huésped genomanipuladas para que expresen GnTIII producirán mayores títulos de anticuerpos que los que pueden producir las células de mamífero. También producirán anticuerpos con mayor potencia que ADCC.

Los anticuerpos producidos por líneas celulares de mamífero transfectadas con GnTIII se ha demostrado que son eficaces como anticuerpos producidos por líneas celulares no transfectadas, pero a una concentración de 10-20 veces menor (Davies *et al.* (2001) Biotechnol. Bioeng. 74(4):288-94). Un aumento por un factor de veinte en la productividad del vehículo de producción de la invención comparado con los sistemas mamíferos, y un aumento de diez veces en la potencia conseguirá una mejora en la productividad neta de doscientos. La invención así proporciona un sistema y procedimiento para producir títulos elevados de un anticuerpo que tiene potencia elevada (por ejemplo, hasta varios órdenes de magnitud más potente de lo que puede producirse normalmente). El sistema y procedimiento es seguro y proporciona anticuerpos de gran potencia a bajo coste en periodos breves de tiempo. Las células huésped genomanipuladas para que expresen GnTIII como se describe en el presente documento producen inmunoglobulinas que tienen *N*-glucanos diseccionados en tasas de al menos 50 mg/litro/día a al menos 500 mg/litro/día. Además, cada molécula de inmunoglobulina (Ig) (que comprende GlcNAc diseccionadora) es más potente que la misma molécula de Ig producida sin GlcNAc diseccionadora.

Producción de estructuras multiantenarias para una funcionalidad de glucoproteína mejorada

Se ha encontrado que la síntesis de estructuras tetraantenarias es muy importante para la actividad biológica *in vivo* de una variedad de proteínas como por ejemplo EPO y α_1 -ácido glucoproteínas. Takeuchi *et al.*, Proc Natl Acad Sci USA, octubre de 1989; 86(20):7819-22; Boris *et al.*, Inflammation (1990) 14, 315-323. Los estudios farmacocinéticos han demostrado que la voluminosa estructura de la ramificación tetraantenaria evita que EPO se filtre a la orina. La modificación de proteínas, por ejemplo, con conjugados químicos (por ejemplo, polietilenglicol), se ha diseñado para que retrase el aclaramiento de glucoproteínas potencialmente terapéuticas. Por consiguiente, en una realización, la presente invención proporciona procedimientos para sintetizar glucoproteínas que comprenden múltiples estructuras antenarias en células huésped de hongos unicelulares o multicelulares (*P. pastoris*) que tienen mejor actividad biológica *in vivo* y se aclaran con menor rapidez que la misma glucoproteína que tiene menos antenas. En esencia, las glucoproteínas producidas de acuerdo con los procedimientos de la presente invención (véase también el documento WO 02/00879 y WO 03/056914 que se incorporan al presente documento por referencia) tienen una eficacia terapéutica mejorada.

Los siguientes son ejemplos que ilustran diversos aspectos de la presente invención. Estos ejemplos no deberían interpretarse como limitantes: los ejemplos se incluyen únicamente con fines ilustrativos.

Ejemplo 1

*Clonación y alteración del gen OCH1 en P. pastoris*5 *Generación de un mutante en OCH1 de P. pastoris*

Se amplificó un ORF del 1215 pb del gen *OCH1* de *P. pastoris* que codifica una α ,1,6 manosiltransferasa putativa a partir del ADN genómico de *P. pastoris* (cepa X-33, Invitrogen, Carlsbad, CA) usando los oligonucleótidos 5'-ATGGCGAAGGCAGATGGCAGT-3' (SEC. ID. N.º: 3) y 5'-TTAGTCCTTCCAACCTTCCTTC-3' (SEC. ID. N.º: 4) que se diseñaron basándose en la secuencia de *OCH1* de *P. pastoris* (Publicación de solicitud de patente japonesa n.º 8-336387). Posteriormente, se amplificaron 2685 pb aguas arriba y 1175 pb aguas abajo del ORF del gen *OCH1* a partir de una adenoteca genómica de *P. pastoris* (Boehm, T. *et al.* (1999) Yeast 15(7):563-72) usando los oligonucleótidos internos 5'-ACTGCCATCTGCCTTCGCCAT-3' (SEC. ID. N.º: 47) del gen *OCH1*, y los oligonucleótidos 5'-GTAATACGACTCACTATAGGGC-3' T7 (SEC. ID. N.º: 48) y 5'-AATTAACCCTCACTAAAGGG-3' T3 (SEC. ID. N.º: 49) del esqueleto de la adenoteca que porta el plásmido lambda ZAP II (Stratagene, La Jolla, CA). El fragmento de 5075 pb resultante se clonó en el vector pCR2.1-TOPO (Invitrogen, Carlsbad, CA) y se denominó pBK9.

Después de ensamblar una construcción de inactivación génica que sustituía el marco de lectura de *OCH1* por un gen de resistencia *HIS4*, *P. pastoris* se transformó y las colonias se cribaron para determinar la sensibilidad a la temperatura a 37°C. Los mutantes en *OCH1* de *S. cerevisiae* son sensibles a la temperatura y crecen despacio a temperaturas elevadas. Así se puede identificar a los homólogos funcionales de *OCH1* en *P. pastoris* complementando un mutante en *OCH1* de *S. cerevisiae* con un ADN o adenoteca de *P. pastoris*. Aproximadamente 20 cepas sensibles a la temperatura se sometieron después a cribado por PCR de las colonias para identificar las colonias con un gen *och1* eliminado. Se obtuvieron varias delecciones en *och1*.

El pBK9.1 linearizado, que tiene una secuencia de 2,1 kb aguas arriba y 1,5 kb aguas abajo del casete de la secuencia del gen *OCH1* que porta el gen *HIS4* de *Pichia*, se transformó en *P. pastoris* BK1 [GS115 (*his4* Invitrogen Corp., San Diego, CA) que porta el gen de IFN- β humano en el locus AOX1] una inactivación del gen *OCH1* de tipo silvestre. El cribado inicial de los transformantes se llevó a cabo usando medio de eliminación con histidina seguido de plaques replicados para seleccionar las colonias sensibles a la temperatura. Veinte de las doscientas colonias positivas para histidina mostraron un fenotipo sensible a la temperatura a 37°C. Para excluir la integración aleatoria de pBK9.1 en el genoma de *Pichia*, las 20 cepas aisladas sensibles a la temperatura se sometieron a PCR de colonias usando cebadores específicos para la secuencia aguas arriba del sitio de integración y al ORF de *HIS4*. Dos de las veinte colonias carecían de *och1* y se analizaron posteriormente usando una transferencia Southern y una transferencia Western que indicaba la alteración funcional de *och1* por la construcción de inactivación de *och1*. El ADN genómico se digirió usando dos enzimas de restricción distintos *Bgl*III y *Cla*I para confirmar la inactivación de *och1* y para confirmar la integración en el marco de lectura abierto. La transferencia Western mostró mutantes de *och1* que carecía de una banda marcada producida en la GS115 de tipo silvestre en 46,2 kDa.

Ejemplo 2

Genomanipulación de P. pastoris con α -1,2-manosidasa para producir precursores de IFN- β que contenían Man₅GlcNAc₂

Es necesaria una α -1,2-manosidasa para recortar Man₈GlcNAc₂ proporcionando Man₅GlcNAc₂, un intermedio esencial para la formación de *N*-glucanos complejos. Aunque la producción de un precursor de Man₅GlcNAc₂ es esencial, no es necesariamente suficiente para la producción de glucanos híbridos y complejos porque el isómero específico de Man₅GlcNAc₂ puede o no ser sustrato para GnTI. Se genomanipula un mutante en *och1* de *P. pastoris* para que exprese interferón- β controlado por un promotor *aox*. Se construye una adenoteca mediante el ligado en marco del dominio catalítico de manosidasa IB humana (una α -1,2-manosidasa) con una subadenoteca que incluye secuencias que codifican péptidos de localización en el aparato de Golgi cis y en el RE. El organismo huésped se transforma después con la adenoteca, lo que produce una población genéticamente mixta en la que los transformantes individuales expresan cada uno interferón- β así como un gen sintético de manosidasa de la adenoteca. Se cultivan colonias individuales de transformantes y se induce la producción de interferón mediante la adición de metanol. En estas condiciones, más del 90% de la proteína segregada es interferón- β glucosilado.

Los sobrenadantes se purifican para eliminar las sales y los contaminantes de bajo peso molecular mediante cromatografía de fase inversa sobre sílice en C18. Los transformantes deseados que expresan una α -1,2-manosidasa activa, dirigida de forma apropiada producen interferón- β que incluye *N*-glucanos de la estructura Man₅GlcNAc₂, que tiene una masa molecular menor que el interferón- β de la cepa parental. El interferón- β purificado se analiza mediante espectrometría de masas MALDI-TOF y se identifican las colonias que expresan la forma deseada de interferón- β .

Ejemplo 3

Generación de una cepa mutante para och1 que expresa una α -1,2-manosidasa, GnTI y GnTII para la producción de una glucoproteína humanoide

El marco de lectura abierto de 1215 pb del gen *OCH1* de *P. pastoris* así como 2685 pb aguas arriba y 1175 pb aguas abajo se amplificó por PCR (véase también el documento WO 02/00879), se clonó en el vector pCR2.1-TOPO (Invitrogen) y se denominó pBK9. Para crear una cepa con *och1* inactivado que contenía múltiples marcadores auxotróficos, se digirieron 100 μ g de pJN329, un plásmido que contiene un alelo mutante en *och1::URA3* flanqueado por sitios de restricción *SfiI* con *SfiI* y se usaron para transformar la cepa JC308 de *P. pastoris* (Cereghino *et al.* (2001) Gene 263:159-169) por electroporación. Tras la incubación en medio definido que carecía de uracilo durante 10 días a temperatura ambiente, se eligieron 1000 y se volvieron a plaquear. Los clones URA+ que no podían crecer a 37°C, pero crecieron a temperatura ambiente, se sometieron a PCR de colonias para analizar la integración correcta del alelo mutante de *och1::URA3*. Un clon que mostraba el patrón de PCR esperado se denominó YJN153. Se usó el dominio Kringle 3 de plasminógeno humano (K3) como proteína modelo. Un plásmido denominado Neo^R que contenía el gen K3 se transformó en la cepa Y JN153 y una cepa resultante, que expresaba K3, se denominó BK64-1.

El plásmido pPB103, que contiene el gen *MNN2-2* de *Kluyveromyces lactis* que codifica un transportador de UDP-N-acetilglucosamina al aparato de Golgi se construyó clonando un fragmento con extremos romos BgIII-HindIII del vector pDL02 (Abeijon *et al.* (1996) Proc. Natl. Acad. Sci. U.S.A. 93:5963-5968) en *BgRI* y se digirió con BamHI y pBLADE-SX con extremos romos que contenía el gen *ADE1* de *P. pastoris* (Cereghino *et al.* (2001) Gene 263:159-169). Este plásmido se linealizó con *EcoNI* y se transformó en la cepa BK64-1 mediante electroporación y una cepa que se confirmó que contenía el *MNN2-2* mediante análisis por PCR se denominó PBP1.

Se generó una adenoteca de construcciones de manosidasa, que comprendía las fusiones en marco de los dominios líder de varias proteínas de membrana de tipo I o de tipo II de *S. cerevisiae* y *P. pastoris* fusionados con los dominios catalíticos de varios genes de α -1,2-manosidasa de ser humano, ratón, mosca, gusano y levaduras (véase, por ejemplo, WO 02/00879). Esta adenoteca se creó en un vector de integración de *HIS4* de *P. pastoris* y se cribó linearizando con *Sall*, transformando mediante electroporación en la cepa PBP1, y analizando los glucanos liberados de la proteína testigo K3. Se eligió una construcción activa como quimera de los nucleótidos 988-1296 (extremo C) del gen *SEC12* de levadura fusionado con una delección del gen α -1,2-manosidasa IA de ratón (Figura 3), al que le faltaban los 187 nucleótidos. Una cepa de *P. pastoris* que expresaba esta construcción se denominó PBP2.

Se generó una adenoteca de construcciones de GnTII, que comprendía fusiones en marco de la misma adenoteca líder con los dominios catalíticos de los genes de GnTI de fuentes humana, de gusano, rana y mosca (WO 02/00879). Esta adenoteca se creó en un vector de integración de *ARG4* de *P. pastoris* y se cribó linearizando con *AatII*, transformando mediante electroporación en la cepa PBP2, y analizando los glucanos liberados de K3. Una construcción activa elegida era una quimera de las primeras 120 pb del gen *MNN9* de *S. cerevisiae* fusionado a una delección del gen GnTI humano, al que le faltaban los 154 pb primeros. Una cepa de *P. pastoris* que expresaba esta construcción se denominó PBP-3. (Véase también la Figura 36).

Se generó una adenoteca de construcciones de GnTI, que comprendía fusiones en marco de la misma adenoteca líder con los dominios catalíticos de los genes de GnTII de fuentes humana, y de rata (WO 02/00879). Esta adenoteca se creó en un vector de integración de *P. pastoris* que contenía el gen *NST^R* que confiere resistencia al fármaco nourseotricina. Los plásmidos de la adenoteca se linealizaron con *EcoRI*, transformado en la cepa RDP27 mediante electroporación, y las cepas resultantes se cribaron mediante análisis de los glucanos liberados de K3 purificado.

Materiales para las siguientes reacciones

Los reactivos MOPS, cacodilato sódico, cloruro de manganeso, UDP-galactosa y ácido CMP-N-acetilneuramínico eran de Sigma. El ácido trifluoroacético (TFA) era de Sigma/Aldrich, Saint Louis, MO. La α 2,6-sialiltransferasa de rata de *Spodoptera frugiperda* y β 1,4-galactosiltransferasa de leche bovina eran de Calbiochem (San Diego, CA). La N-glucosidasa F, las manosidasas, y los oligosacáridos eran de Glyko (San Rafael, CA). La resina ToyoPearl para DEAE era de TosoHaas. La resina "HisBind" quelante de metales era de Novagen (Madison, WI). Las placas de aclaramiento de lisado de 96 pocillos eran de Promega (Madison, WI). Las placas para la unión de proteínas de 96 pocillos eran de Millipore (Bedford, MA). Las sales y agentes tamponadores eran de Sigma (St. Louis, MO). Las matrices para MALDI eran de Aldrich (Milwaukee, WI).

Purificación de proteínas

Kringle 3 se purificó usando un formato de 96 pocillos con un robot Beckman BioMek 2000 para el manejo de las muestras (Beckman/Coulter Ranch Cucamonga, GA). La Kringle 3 se purificó a partir del medio de expresión usando un marcador de hexa-histidina en el extremo C. La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind. En resumen, se vierte un volumen de 150 μ l reposado de resina en los pocillos de una placa para la unión de lisados de 96 pocillos, se lava con 3 volúmenes de agua y se carga con 5 volúmenes de NiSO₄ 50 mM y se lava con 3 volúmenes de tampón de unión (imidazol 5 mM, NaCl 0,5M, Tris-HCl 20 mM a pH 7,9). El medio para la expresión de proteínas se diluye 3:2, medio/PBS (PO₄ 60 mM, KCl 16 mM, NaCl 822 mM a pH 7,4) y se carga en las columnas. Después de drenar, las columnas se lavan con 10 volúmenes de

tampón de unión y 6 volúmenes de tampón de lavado (imidazol 30 mM, NaCl 0,5 M, Tris-HCl 20 mM pH 7,9) y la proteína se eluye con 6 volúmenes de tampón de elución (imidazol 1 M, NaCl 0,5 M, Tris-HCl 20 mM a pH 7,9). Las glucoproteínas eluidas se evaporan a sequedad mediante liofilización.

5 Liberación de glucanos ligados en N

Los glucanos son liberados y separados de las glucoproteínas mediante una modificación de un procedimiento previamente reseñado (Papac, *et al.* A. J. S. (1998) *Glycobiology* 8, 445-454). Los pocillos de una placa MultiScreen IP (membrana ImmobiloN-P) de 96 pocillos (Millipore) se humedecen con 100 μ l de metanol, se lavan con 3X150 μ l de agua y 50 μ l de tampón RCM (urea 8 M, Tris 360 mM, EDTA 3,2 mM a pH8,6), drenando con vacío suave después de cada adición. Las muestras de proteína seca se disuelven en 30 μ l de tampón RCM y se transfieren a los pocillos que contenían 10 μ l de tampón RCM. Los pocillos se drenan y se lavan dos veces con tampón RCM. Las proteínas se reducen mediante la adición de 60 μ l de DTT 0,1 M en tampón RCM durante 1 h a 37°C. Los pocillos se lavan tres veces con 300 μ l de agua y se carboximetilan mediante la adición de 60 μ l de ácido yodoacético 0,1 M durante 30 min a oscuras a temperatura ambiente. Los pocillos se lavan de nuevo tres veces con agua y las membranas se bloquean mediante la adición de 100 μ l de 1% de PVP 360 en agua durante 1 h a temperatura ambiente. Los pocillos se drenan y se lavan tres veces con 300 μ l de agua y se desglucosilan mediante la adición de 30 μ l de NH_4HCO_3 10 mM a pH 8,3 que contiene una miliunidad de *N*-glucanasa (Glyko). Después de 16 horas a 37°C, la solución que contenía los glucanos se retira mediante centrifugación y se evapora a sequedad.

20 Espectrometría de masas de tiempo de vuelo con desorción ionización por láser asistida por matriz

Los pesos moleculares de los glucanos se determinaron usando un espectrómetro de masas para MALDI-TOF lineal Voyager DE PRO (Applied Biosciences) usando extracción retardada. Los glucanos secos de cada pocillo se disolvieron en 15 μ l de agua y 0,5 μ l se sembraron sobre placas de muestra de acero inoxidable y se mezclaron con 0,5 μ l de matriz S-DHB (9 mg/ml de ácido dihidroxibenzoico, 1 mg/ml de ácido 5-metoxisalicílico en agua/acetonitrilo 1:1 0,1% de TFA) y se dejó secar.

Se generaron iones por radiación con un láser de nitrógeno pulsado (337 nm) con un tiempo de pulso de 4 ns. El instrumento se hizo trabajar en modo de extracción retardado con un retardo de 125 ns y un voltaje de aceleración de 20 kV. El voltaje de los electrodos era de 93,00%, el voltaje del cable guía era de 0,10%, la presión interna era inferior a 5×10^{-7} torr, y la entrada de masa baja era 875 Da. Los espectros se generaron a partir de la suma de 100-200 pulsos de láser y se registraron en un digitalizador de 2 GHz. Se usó el oligosacárido $\text{Man}_5\text{GlcNAc}_2$ como patrón de peso molecular externo. Todos los espectros se generaron con el aparato en modo iónico positivo. La exactitud de la masa estimada de los espectros era 0,5%.

La masa de los *N*-glucanos eluida de la columna de forma general está asociada a un aducto iónico positivo, que aumenta la masa mediante el peso molecular del ión positivo. Los aductos más habituales son H^+ , Na^+ y K^+ .

40 Ejemplo 4

*Genomanipulación de P. pastoris para producir $\text{Man}_5\text{GlcNAc}_2$ como estructura de *N*-glucano predominante usando una adenoteca de combinaciones*

Se genomanipuló un mutante en *och1* de *P. pastoris* (véanse los Ejemplos 1 y 3) para que expresaran y segregaran proteínas como por ejemplo el dominio kringle 3 de plasminógeno humano (K3) controlado por el promotor inducible de *ADXI*. Se usó el dominio Kringle 3 de plasminógeno humano (K3) como proteína modelo. Se amplificó un fragmento de ADN que codificaba la K3 usando polimerasa Pfu turbo (Stratagene, La Jolla, CA) y se clonó en los sitios *EcoRI* y *XbaI* de pPICZ α A (Invitrogen, Carlsbad, CA), produciendo un marcador de 6- His en el extremo C. Para mejorar la eficiencia de glucosilación ligada en N de K3 (Hayes *et al.* 1975 J. Arch. Biochem. Biophys. 171, 651-655), se reemplazó Pro₄₆ por Ser₄₆ usando mutagénesis dirigida al sitio. El plásmido resultante se denominó pBK64. Se confirmó la secuencia correcta de la construcción por PCR mediante secuenciación del ADN.

Se construyó una adenoteca de combinaciones ligando en marco los dominios catalíticos de α -1,2-manosidasa IB murina (Genbank AN 6678787) e IA (Genbank AN 6754619) con una subadenoteca que incluía secuencias que codifican los péptidos de localización a la vesícula Cop II, RE, y aparato de Golgi cis de acuerdo con la Tabla 6. La adenoteca combinada se usó para generar construcciones fusionadas individuales, que después se transformaron en el organismo huésped que expresaba K3, produciendo una población genéticamente mixta en la que cada transformante individual expresa K3 así como un gen de fusión de señal de localización/manosidasa de la adenoteca. Se cultivaron los transformantes individuales y se indujo la producción de K3 mediante transferencia a un medio que contenía metanol. En estas condiciones, después de 24 horas de inducción, más del 90% de la proteína del medio era K3. La proteína testigo K3 se purificó del sobrenadante para eliminar las sales y los contaminantes de bajo peso molecular mediante cromatografía por afinidad a Ni. Tras la purificación por afinidad, la proteína se desaló mediante cromatografía de exclusión por tamaño en una resina a Sephadex G10 (Sigma, St. Louis, MO) y o bien se sometió directamente al análisis por MALDI-TOF que se describe más adelante o bien los *N*-glucanos se eliminaron mediante digestión con PNGasas como se describe más adelante (Liberación de *N*-glucanos) y se sometió a análisis por MALDI-TOF Miele *et al.* (1997) *Biotechnol. Appl. Biochem.* 25:151-157.

ES 2 336 790 T3

Siguiendo esta estrategia, se obtuvo un conjunto diverso de transformantes; algunos no mostraban modificación en los *N*-glucanos con respecto a la cepa con *och1* inactivado; y otros mostraban un grado elevado de recorte de manosas (Figuras 5D y 5E). Los transformantes deseados que expresaban α -1,2-manosidasa activa, dirigida de forma apropiada producía K3 con *N*-glucanos de la estructura Man₅GlcNAc₃. Esto confiere una masa molecular menor a la glucoproteína con respecto a la K3 de la cepa con la delección *och1* parental, una diferencia que se detectó fácilmente mediante espectrometría de masas MALDI-TOF (Figura 5). La Tabla 7 indica los niveles de producción relativos de Man₅GlcNAc₂.

TABLA 7

Una adenoteca de combinaciones representativa de secuencias de localización/dominios catalíticos que muestran los niveles relativos de producción de Man₅GlcNAc₂

		Secuencias peptídicas directoras				
		MNS1(s)	MNS1(m)	MNS1(l)	SEC12(s)	SEC12(m)
Dominios catalíticos	Manosidasa 1A de ratón Δ 187	FB4 ++	FB5 +	FB6 -	FB7 ++	FB8 ++++
	manosidasa 1B de ratón Δ 58	GB4 ++	GB5 +	GB6 +	GB7 ++	GB8 +
	manosidasa 1B de ratón Δ 99	GC4 -	GC5 +++	GC6 +	GC7 +	GC8 +
	manosidasa 1B de ratón Δ 170	GD4 -	GD5 -	GD6 -	GD7 +	GD8 +

TABLA 8

Otra adenoteca de combinaciones de secuencias de localización/dominios catalíticos que muestran los niveles relativos de producción de Man₅GlcNAc₂

			Secuencias peptídicas directoras				
			VAN1(s)	VAN1(m)	VAN1(l)	MNN10(s)	MNN10(l)
Dominios catalíticos	Manosidasa 1B de <i>C. elegans</i> Δ 80	BC18-5 +++++	BC19 ++++	BC20 +++	BC27 +++++	BC28 +++++	BC29 +++
	Manosidasa 1B de <i>C. elegans</i> Δ 31	BB18 +++++	BB19 +++++	BB20 ++++	BB18 +++++	BB19 +++++	BB20 ++++

Se seleccionaron péptidos directores a partir de *MNS I* (SwissProt P32906) en *S. cerevisiae* (largo, medio y corto) (véase anteriormente Adenotecas de ácidos nucleicos; Adenoteca de combinaciones de construcciones fusionadas) y *SEC12* (SwissProt P11655) en *S. cerevisiae* (988-1140 nucleótidos: corto) y (988-1296: medio). Aunque la mayoría de las secuencias de péptidos directores eran delecciones del extremo N, algunas secuencias de péptidos directores, como por ejemplo *SEC12* eran delecciones del extremo C. Los dominios catalíticos usados en este experimento se seleccionaron a partir de manosidasa 1A de ratón con una delección de 157 aminoácidos en el extremo N; y manosidasa 1B de ratón con una delección de 58, 99 y 170 aminoácidos. El número de (+), tal como se usa en el presente documento, indica los niveles relativos de producción de Man₅GlcNAc₂. La notación (-) indica sin producción aparente de Man₅GlcNAc₂. La notación (+) indica menos de 10% de producción de Man₅GlcNAc₂. La notación (++) indica aproximadamente 10-20% de producción de Man₅GlcNAc₂. La notación (+++) indica aproximadamente 20-40% de producción de Man₅GlcNAc₂. La notación (+++++) indica aproximadamente 50% de producción de Man₅GlcNAc₂. La notación (+++++) indica más de 50% de producción de Man₅GlcNAc₂.

La Tabla 9 muestra la cantidad relativa de Man₅GlcNAc₂ en la K3 segregada. Se generaron seiscientos ocho (608) cepas diferentes de *P. pastoris*, $\Delta och1$ transformándolas con una única construcción de una adenoteca genética de combinaciones que se generó fusionando diecinueve (19) dominios catalíticos de α -1,2 manosidasa a treinta y dos (32) líderes fúngicos de RE, y cis-aparato de Golgi.

TABLA 9

Cantidad de Man ₅ GlcNAc ₂ en la K3 segregada (% de total glucanos)	Número de construcciones (%)
S.D.*	19 (3,1)
0-10%	341 (56,1)
10-20%	50 (8,2)
20-40%	75 (12,3)
40-60%	72 (11,5)
Más de 60%	51 (8,4) †
Total	608 (100)
* varias construcciones fusionadas no se analizaron debido a que los plásmidos correspondientes no podían propagarse en <i>E.coli</i> antes de la transformación, en <i>P. pastoris</i>. † Los clones con el mayor grado de recorte de Man₅GlcNAc₂ (30/51) se analizaron posteriormente para determinar la actividad de las manosidasas en el sobrenadante del medio. La mayoría (28/30) presentaron una actividad detectable de manosidasas en el sobrenadante (por ejemplo Figura 4B). Sólo dos construcciones presentaron niveles altos de Man₅GlcNAc₂, careciendo de actividad de manosidasa en el medio (por ejemplo Figura 4C).	

La Tabla 7 muestra dos construcciones pFB8 y pGC5, entre otras, que presentan Man₅GlcNAc₂. La Tabla 8 muestra una construcción más preferente, pBC18-5, una secuencia del péptido director YAN1(s) de *S. cerevisiae* (de SwissProt 23642) ligada en marco a una delección en el extremo N de 80 aminoácidos (Van1(s) de *Saccharomyces*/manosidasa IB de *C.elegans* Δ 80) de la manosidasa IB de *C. elegans* (Genbank AN CAA9S114). Esta construcción fusionada también produce una estructura Man₅GlcNAc₂ predominante, como se muestra en la Figura 5E. Se demostró que esta construcción produce más del 50% de Man₅GlcNAc₂ (+++++).

Generación de una adenoteca de combinaciones de localización/manosidasas

La generación de una adenoteca de combinaciones de dominios catalíticos de α -1,2-manosidasa fusionados a péptidos directores requería la amplificación de dominios de manosidasa con longitudes variables de delecciones en

el extremo N de un número de organismos. Para conseguir este objetivo, se amplificó por PCR los marcos de lectura abiertos (ORF) de α -1,2-manosidasas de ADNc o de ADN genómico obtenido de las siguientes fuentes: *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Aspergillus nidulans* y *Penicillium citrinum*. En cada caso, el ADN se incubó en presencia de oligonucleótidos cebadores específicos para la secuencia de manosidasa deseada además de los reactivos necesarios para realizar la reacción por PCR. Por ejemplo, para amplificar el ORF de la α -1,2-manosidasa 1A de *M. musculus*, se incubaron el cebador 5' ATGCCCGTGGGGGGCCTGTTGCCGCTCTT CAGTAGC (SEC. ID. N.º: 52) y el cebador 3' TCATTCTCTTTGCCATCAATTCCTTCTTCTGTTTCACGG (SEC. ID. N.º: 53) en presencia de ADN polimerasa Pfu (Stratagene, La Jolla, CA) y se amplificaron en las condiciones que recomienda Stratagene usando los parámetros de ciclado: 94°C durante 1 min (1 ciclo); 94°C durante 30 s, 65°C durante 30 s, 72°C durante 3 min (30 ciclos). Tras la amplificación, la secuencia de ADN que codifica el ORF se incubó a 72°C durante 5 min con 1 U de ADN polimerasa Taq (Promega, Madison, WI) antes del ligado en pCR2.1- TOPO (Invitrogen, Carlsbad, CA) y la transformación en *E. coli*, TOP10 químicamente competentes como recomendaba Invitrogen. El producto de la PCR clonado se confirmó mediante secuenciación ABI usando los cebadores específicos para el ORF de la manosidasa.

Para generar los truncados deseados en el extremo N de cada manosidasa, se usó el ORF completo de cada manosidasa como plantilla en una tanda posterior de reacciones por PCR en las que la posición de hibridación del cebador 5' era específica para el extremo 5' del truncado deseado y el cebador 3' seguía siendo específico para el extremo 3' original del ORF. Para facilitar el subclonado del fragmento de manosidasa truncado en el vector de expresión de las levaduras, pJN347 (Figura 2C), se genomanipularon los sitios de restricción *AscI* y *PacI* en cada producto del truncado en los extremos 5'- y 3' respectivamente. El número y la posición de los truncados del extremo N generados para cada ORF de manosidasa dependía de la posición de la región transmembrana (TM) con respecto al dominio catalítico (CD). Por ejemplo, si la región del tallo localizada entre la TM y el CD tenía menos de 150 pb, entonces sólo se generaba un truncado para esa proteína. Si, sin embargo, la región del tallo tenía más de 150 pb, entonces se generaba una o más truncados dependiendo de la longitud de la región del tallo.

Un ejemplo de cómo se generaban los truncados para la manosidasa 1A de *M. musculus* (Genbank AN 6678787) se describe en el presente documento, y se usaba una estrategia similar para las otras manosidasas. La Figura 3 ilustra el ORF de la α -1,2-manosidasa 1A de *M. musculus*, donde los dominios transmembrana y catalíticos previstos se resaltan en negrita. Basándose en esta estructura, se diseñaron tres cebadores 5' (las posiciones de hibridación están subrayadas en la Figura 3) generando las delecciones del extremo N- Δ 65-, Δ 105- y Δ 187. Usando la delección de Δ 65 del extremo N como ejemplo, el cebador 5' usado fue 5'-GGCGCGCCGACTCCTCCMGCTGCTCAGCGGGGTCCTGTTCCAC-3' (SEC. ID. N.º: 54) (con el sitio de restricción *AscI* destacado en negrita) junto con el cebador 3' 5'-CCTTMTTMT CATTCTCTTTGCCATCMTTCTCTTCTGTTTCACGG-3' (SEC. ID. N.º: 55) (con el sitio de restricción *PacI* resaltado en negrita). Estos dos cebadores se usaron para amplificar un fragmento de 1561 pb en las condiciones que se describen anteriormente para la amplificación del ORF de la manosidasa 1A de longitud completa de *M. musculus*. Además, al igual que el producto obtenido para el ORF de longitud completa, el producto truncado también se incubó ADN polimerasa Taq, se ligó en pCR2.1-TOPO (Invitrogen, Carlsbad, CA), se transformó en TOP10 y se secuenció por ABI. Después de haber sido amplificado y confirmada la secuencia del fragmento de manosidasa truncado, el plásmido resultante, pCR2.1- Δ 65 mMannIA, se dirigió con *AscI* y *PacI* en tampón de New England Biolabs #4 (Beverly, MA) durante 16 h a 37°C. En paralelo, el pJN347 (Figura 2C) se dirigió con las mismas enzimas y se incubó como se describe anteriormente. Después de la digestión, tanto el esqueleto de pJN347 (Figura 2C) como el dominio catalítico truncado se extrajeron en gel y se ligaron usando el Quick Ligation Kit (New England Biolabs, Beverly, MA), como recomiendan los fabricantes, y se transformó en células DH5a químicamente competentes (Invitrogen, Carlsbad, CA). Se usó PCR de colonias para confirmar la generación de la construcción de pJN347 y manosidasa IM65 de ratón.

Habiendo generado una adenoteca de dominios catalíticos de α -1,2-manosidasa truncados en el vector de expresión de levadura pJN347 (Figura 2C) la etapa que quedaba para generar la adenoteca de péptido director/dominio catalítico fue clonar en marco las secuencias de péptidos directores (Figura 2). Tanto las construcciones de pJN347 y manosidasa (Figura 20) como las construcciones de pCR2.1TOPO-péptido director (Figura 28) se incubaron toda la noche a 37°C en tampón de New England Biolabs #4 en presencia de las enzimas de restricción *NotI* y *AscI*. Después de la digestión, tanto el esqueleto de la manosidasa en pJN347 como las regiones de péptidos directores se extrajeron en gel y se ligaron usando el Quick Ligation Kit (New England Biolabs, Beverly, MA), como recomiendan los fabricantes, y se transformó en células DH5a químicamente competentes (Invitrogen, Carlsbad, CA). Posteriormente, las construcciones pJN347 de péptido director/manosidasa se secuenciaron con ABI para confirmar que las fusiones generadas estaban en marco. El tamaño estimado de la adenoteca del péptido director final/alfa1,2-manosidasa contiene más de 1300 construcciones generadas mediante la estrategia que se describe anteriormente. La Figura 2 ilustra la construcción de la adenoteca de combinaciones.

60 Genomanipulación de una cepa de *P. pastoris* con inactivación en *OCH1* con múltiples marcadores auxotrófico

La primera etapa en la construcción de plásmidos suponía crear un conjunto de plásmidos universales que contenían regiones de ADN del gen *KEX1* de *P. pastoris* (Boehm *et al.* Yeast mayo de 1999; 15(7) :563-72) como espaciadores para las regiones 5' y 3' de los genes a inactivar. Los plásmidos también contenían el Ura-blaster de *S. cerevisiae* (AI ani *et al.* (1987) Genetics 116:541-545) como espaciador para los marcadores auxotróficos, y un casete de expresión con un sitio de clonación múltiple para la inserción de un gen exógeno. Un fragmento de 0,9 kb de la región 5' de *KEX1* de *P. pastoris* se amplificó mediante PCR usando cebadores GGCGAGCTCGGCC TACCCGGCCAAGGCTGAGATCATTTGTCCAGCTTCA GA (SEC. ID. N.º: 56) y GCCCACGTCGACGGATCCG

TTTAAACATCGATTGGAGAGGCTGACACCGCTACTA (SEC. ID. N.º: 57) y ADN genómico de *P. pastoris* como plantilla y se clonó en los sitios *SacI*, *Sall* de pUC19 (New England Biolabs, Beverly, MA). El plásmido resultante se cortó con *BamHI* y *Sall*, y se amplificó un fragmento de 0,8 kb de la región 3' de KEX1 que había sido amplificado usando cebadores CGGGATCCACTAGTATTAAATCATATGTGCGAGTGACAACCTCTTCCC ACATGG (SEC. ID. N.º: 58) y GGACGCGTCGACGGCCTACCCGGCCGTACGAGGAATTTCTCGGATGACTCTTTTC (SEC. ID. N.º: 59) se clonó en los sitios abiertos creando pJN262. Este plásmido se cortó con *BamHI* y el fragmento de 3,8 kb entre *BamHI*, *BglII* de pNKY51 (Alani *et al.* (1987) Genetics 116:541-545) se insertó en las dos orientaciones resultantes posibles en los plásmidos pJN263 (Figura 4A) y pJN284 (Figura 4B).

Se creó un casete de expresión con *NotI* y *PacI* como sitios de clonación. El promotor GAPDH de *P. pastoris* se amplificó usando cebadores CGGGATCCCTCGAGAGATCTTTTTTGTAGAAATGTCITGGTGCCT (SEC. ID. N.º: 60) y GGACATGCATGCACTAGTGCGGCCCGCCACGTGATAGTTGTTCAATTGATTGAAATAGGGACAA (SEC. ID. N.º: 61) y el plásmido pGAPZ-A (Invitrogen) como plantilla y se clonó en los sitios *BamHI*, *SphI* de pUC19 (New England Biolabs, Beverly, MA) (Figura 48). El plásmido resultante se cortó con *SpeI* y *SphI* y la región de terminación de la transcripción CYC1 ("TT") que había sido amplificado usando cebadores CCTTGCTAGCT TAATTAACCGCGGCACGTCCGACGGCGGCCCA CGGGTCCCA (SEC. ID. N.º: 62) y GGACATGCATGCG GATCCCTTAAGAGCCGGCAGCTTGCAAATTAAAGCCTTCGAGCGTCCC (SEC. ID. N.º: 63) y el plásmido pPICZ-A (Invitrogen) como plantilla se clonó en los sitios abiertos creando pJN261 (Figura 48).

Se creó un plásmido inactivado para el gen *OCH1* de *P. pastoris* digiriendo pJN263 con *Sall* y *SpeI* y un fragmento de 2,9 kb de ADN de la región 5' de *OCH1*, que había sido amplificado usando cebadores GAACCACGTCGACGGC CATTGCGGCCAAAACCTTTTTTCTATTCAAACACAAGGCATTGC (SEC. ID. N.º: 64) y CTCCAATACTAG TCGAAGATTATCTTCTACGGTGCCTGGACTC (SEC. ID. N.º: 65) y ADN genómico de *P. pastoris* como plantilla, se clonó en los sitios abiertos (Figura 4C). El plásmido resultante se cortó con *EcoRI* y *PmeI* y se insertó un fragmento de 1,0 kb de ADN de la región 3' de *OCH1* que se había generado usando los cebadores TGGAAGGTTTAAA CAAAGCTAGAGTAAAATAGATATAGCGAG ATTAGAGAATG (SEC. ID. N.º: 66) y AAGAATTCGGCTGGAAG GCCTTGACCTTGATGTAGTTCCCGTT TTCATC (SEC. ID. N.º: 67) para generar pJN298 (Figura 4C). Para permitir la posibilidad de usar de forma simultánea el plásmido para introducir un nuevo gen, se clonó el casete de expresión *BamHI* de pJN261 (Figura 48) en el único sitio *BamHI* de pJN298 (Figura 4C) para crear pJN299 (Figura 4E).

El casete Ura3-blaster de *P. pastoris* se construyó usando una estrategia similar como se describe en Lu *et al.* (1998) Appl. Microbiol. Biotechnol. 49:141-146. Se insertó un fragmento *PstI*, *SpeI* de 2,0 kb de URA3 de de *P. pastoris* en los sitios *PstI*, *XbaI* de pUC19 (New England Biolabs, Beverly, MA) para crear pJN306 (Figura 4D). Después se clonó un fragmento *SacI*, *PvuII* de 0,7 kb de ADN del marco de lectura abierto *lacZ* en los sitios *SacI*, *SmaI*, proporcionando pJN308 (Figura 40). Tras la digestión de pJN308 (Figura 4D) con *PstI*, y el tratamiento con ADN polimerasa de T4, se insertó el fragmento *SacI* - *PvuII* de *lacZ* que había sido cortado como con ADN polimerasa de T4 generando pJN315 (Figura 4D). El casete *lacZ*/URA3 se liberó mediante digestión con *SacI* y *SphI*, con extremos romos con ADN polimerasa de T4 y se clonó en el esqueleto de pJN299 que había sido digerido con *PmeI* y *AflIII* y se habían cortado los extremos romos con ADN polimerasa de T4. El plásmido resultante se denominó pJN329 (Figura 4E).

Se creó un plásmido de expresión HIS4 marcado cortando pJN261 (Figura 4F) con *EcoI* (Figura 4F). Se amplificó un fragmento de 2,7 kb del gen *HIS4* de *Pichia pastoris* que había sido amplificado usando los cebadores GCCCAAGCCCGCCTTAAGGGATCTCCTGATGACTGACTCACTGATAATAAAAATACGG (SEC. ID. N.º: 68) y GGGCGCGTATTTAAATACTAGTGGATCTATCGAATCTAAATGTAAGTTAAATCTCTAA (SEC. ID. N.º: 69) cortado con *NgoMIV* y *SwaI* y después se cortaron los extremos romos usando ADN polimerasa de T4, y después se ligó en el sitio abierto. Este plásmido se denominó pJN337 (Figura 4F). Para construir un plásmido con un sitio de clonación múltiple adecuado para la construcción de una adenoteca de fusiones, se cortó pJN337 con *NotI* y *PacI* y los dos oligonucleótidos GGCCGCTGCAGATTTAAATGAATTCGGCGCGCCTTAAT (SEC. ID. N.º: 70) y TAAGGCGCGCCGAATTCATTAAATCTGCAGGGC (SEC. ID. N.º: 71) que se habían hibridado *in vitro* se ligaron en los sitios abiertos, creando pJN347 (Figura 4F).

Para crear una cepa con *och1* inactivado que contuviera múltiples marcadores auxotróficos, se digirieron 100 µg de pJN329 con *SfiI* y se usó para transformar la cepa JC308 de *P. pastoris* (Cereghino *et al.* (2001) Gene 263:159-169) mediante electroporación. Tras la transformación, las placas de eliminación de URA se incubaron a temperatura ambiente durante 10 días. Se escogieron mil (1000) colonias y se volvieron a plaquear. Los 1000 clones después se plaquearon sobre 2 conjuntos de placas de eliminación de URA. Un conjunto se incubó a temperatura ambiente, mientras que el segundo conjunto se incubó a 37°C. Los clones que no podían crecer a 37°C, pero crecieron a temperatura ambiente, se sometieron a PCR de colonias para analizar la inactivación correcta de *OCH1*. Un clon que mostraba la señal de PCR esperada (aproximadamente 4,5 kb) se denominó YJN153.

Ejemplo 5

Caracterización de la adenoteca de combinaciones

Los transformantes positivos cribados mediante PCR de colonias que confirmaba la integración de la construcción de manosidasa en el genoma de *P. pastoris* posteriormente se cultivaron a temperatura ambiente en 50 ml de medio complejo de metanol tamponado con BMGY constituido por 1% de extracto de levadura, 2% de peptona, tampón de

fosfato potásico 100 mM, a pH 6,0, 1,34% de base nitrogenada de levadura, 4 X 10⁻⁵% de biotina y 1% de glicerol como medio de cultivo) hasta una DO_{600 nm} 2-6, momento en el cual se lavaron con 10ml de BMMY (medio complejo de metanol tamponado constituido por 1% de extracto de levadura, 2% de peptona, tampón de fosfato potásico 100 mM, a pH 6,0, 1,34% de base nitrogenada de levadura, 4 X 10⁻⁵% de biotina y 1% de glicerol como medio de cultivo) antes de la inducción de la proteína testigo durante 24 horas a temperatura ambiente en 5 ml de BMMY. Por consiguiente, la proteína testigo se aisló y se analizó como se describe en el Ejemplo 3 para caracterizar su estructura de glucanos. Usando los péptidos directores de la Tabla 6, los dominios catalíticos de manosidasas localizados en el RE o en el aparato de Golgi mostraron un nivel significativo de recorte de un glucano que contenía predominantemente MansGlcNAc₂ a un glucano que contenía predominantemente Man₅GlcNAc₂. Esto es evidente cuando la estructura del glucano de la glucoproteína testigo se compara entre la de la inactivación de *och1* en Pen las Figuras 5C y 6C y la misma cepa transformada con construcciones de manosidasas de *M. musculus* como se muestra en las Figuras 5D, 5E, 6D - 6F. Las Figuras 5 y 6 muestran la expresión de construcciones generadas en la adenoteca de combinaciones que muestra una actividad significativa de manosidasas en *P. pastoris*. La expresión de pGC5 (MNS1(m) de *Saccharomyces*/manosidasas IB de ratón Δ99) (Figuras 5D y 6E) produjo una proteína que tiene aproximadamente 30% de todos los glucanos recortados a Man₅GlcNAc₂, aunque la expresión de pFB8 (SEC12(m) de *Saccharomyces*/manosidasas IA de ratón Δ187) (Figura 6F) produjo aproximadamente 50% de Man₅GlcNAc₂ y la expresión de pBC18-5 (VANI(s) de *Saccharomyces*/ manosidasas IB de *C. elegans*, Δ80) (Figura 5E) produjo 70% de Man₅GlcNAc₂.

Ejemplo 6

Recorte *in vivo* mediante α-1,2-manosidasa

Para asegurar que las cepas genomanipuladas novedosas del Ejemplo 4 de hecho producían la estructura Man₅GlcNAc₂ deseada *in vivo*, se analizaron los sobrenadantes celulares para determinar la actividad de las manosidasas (véase las Figuras 7 - 9). Para cada construcción/cepa huésped que se describe más adelante, se realizó una HPLC a 30°C con una columna de 4,0 mm x 250 mm de Altech (Avondale, PA, USA) de resina Econosil/NH₂ (5 μm) a un caudal de 1,0 ml/min durante 40 min. En las Figuras 7 y 8, se demostró que se produce degradación del patrón Man₉GlcNAc₂ [b] produciendo un pico que muestra relación directa con ManBGlcNAc₂. En la Figura 7, el patrón Man₉GlcNAc₂ [b] eluyó a los 24,61 min y Man₅GlcNAc₂ [a] eluyó a los 18,59 min. En la Figura 8, Man₉GlcNAc₂ eluyó a los 21,37 min y Man₅GlcNAc₂ a los 15,67 min. En la Figura 9, se demostró que el patrón Man₈GlcNAc₂ [b] eluía a los 20,88 min.

Se cultivaron células de *P. pastoris* que comprendían el plásmido pFB8 (SEC12(m) de *Saccharomyces*/manosidasas IA de ratón Δ1887) a 30°C en BMGY hasta una DO₆₀₀ de aproximadamente 10. Las células se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. Después de 24 horas de inducción, las células se eliminaron por centrifugación proporcionando un sobrenadante esencialmente transparente. Se extrajo una alícuota del sobrenadante para realizar ensayos de manosidasas y el resto se usó para recuperar la K3 segregada soluble. Una única etapa de purificación usando cromatografía sobre CM-sefarosa y un gradiente de elución de NaAc 25 mM, a pH 5,0 a NaAc 25 mM, a pH 5,0, NaCl 1 M, consiguió la elución de K3 con una pureza del 95% con NaCl entre 300-500 mM. El análisis de los N-glucanos de glucanos derivados de K3 se muestra en la Figura 6F. La alícuota de sobrenadante extraída anteriormente se analizó además para determinar la presencia de actividad de manosidasas segregadas. Se añadió un patrón de oligomanosa 9 de tipo ligada en N marcada con 2-aminobenzamida (Man9-2-AB) disponible comercialmente (Glyko, Novato, CA) a: BMMY (Figura 7 A), el sobrenadante de la alícuota anterior (Figura 7 B), y BMMY que contenía 10 ng de α-1,2-manosidasa de *Trichoderma reesei* a 75 mU/ml (obtenida de Contreras *et al.*, documento WO 02/00856 A2) (Figura 7C). Después de incubar durante 24 horas a temperatura ambiente, se analizaron las muestras mediante HPLC sobre aminosílice para determinar el grado de recorte de las manosidasas.

Se cultivaron y ensayaron células de *P. pastoris* que comprendían el plásmido pGC5 (MNS1(m) de *Saccharomyces*/manosidasas IB de ratón Δ99) de forma similar. Las células se cultivaron a temperatura ambiente en BMGY hasta una DO₆₀₀ de aproximadamente 10. Las células se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 controlada por un promotor AOX1. Después de 24 horas de inducción, las células se eliminaron por centrifugación proporcionando un sobrenadante esencialmente transparente. Se extrajo una alícuota del sobrenadante para realizar ensayos de manosidasas y el resto se usó para recuperar la K3 segregada soluble. Una única etapa de purificación usando cromatografía sobre CM-sefarosa y un gradiente de elución de NaAc 25 mM, a pH 5,0 a NaAc 25 mM, a pH 5,0, NaCl 1 M, consiguió la elución de K3 con una pureza del 95% con NaCl entre 300-500 mM. El análisis de los N-glucanos de glucanos derivados de K3 se muestra en la Figura 5D. La alícuota de sobrenadante extraída anteriormente se analizó además para determinar la presencia de actividad de manosidasas segregadas como se muestra en la Figura 8B. Se añadió un patrón disponible comercialmente de Man9-2-AB (Glyko, Novato, CA) a: BMMY (Figura 8A), sobrenadante de la alícuota anterior (Figura 5B), y BMMY que contenía 10 ng de α-1,2-manosidasa de *Trichoderma reesei* a 75 mU/ml (obtenido de Contreras *et al.*, documento WO 02/00856 A2.) (Figura 8C). Después de incubar durante 24 horas a temperatura ambiente, se analizaron las muestras mediante HPLC sobre aminosílice para determinar el grado de recorte de las manosidasas.

Se usó Man9-2-AB como sustrato y es evidente que después de 24 horas de incubación, la actividad de las manosidasas era prácticamente nula en el sobrenadante de la digestión de la cepa pFB8 (SEC12(m) de *Saccharomyces*/manosidasas IA de ratón, Δ187) (Figura 7B) y la digestión de la cepa pGC5 (MNS1(m) de *Saccharomyces*/manosidasas de ratón IB Δ99) (Figura 8B) mientras que el control positivo (α-1,2-manosidasa de *T. reesei* pu-

rificada obtenida de Contreras) consigue la conversión completa de $\text{Man}_6\text{GlcNAc}_2$ a $\text{Man}_5\text{GlcNAc}_2$ en las mismas condiciones, como se muestra en las Figuras 7C y 8C. Estos datos son concluyentes para demostrar el recorte *in vivo* por la manosidasa en la cepa pGC5 de *P. pastoris*; y en la cepa pFB8, lo que claramente difiere de lo publicado hasta la fecha (Contreras *et al.*, documento WO 02/00856 A2.).

La Figura 9 es otro indicio más de la localización y la actividad de la enzima manosidasa. Se cultivó *P. pastoris* que comprendía p8C18-5 (*VAN1(s)*) de *Saccharomyces/manosidasa* IB de *C. elegans* $\Delta 80$) a temperatura ambiente en BMGY a una DO_{600} de aproximadamente 10. Las células se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 controlada por un promotor AOX1. Después de 24 horas de inducción, las células se eliminaron por centrifugación proporcionando un sobrenadante esencialmente transparente. Se extrajo una alícuota del sobrenadante para realizar ensayos de manosidasas y el resto se usó para recuperar la K3 segregada soluble. Una única etapa de purificación usando cromatografía sobre CM-sefarosa y un gradiente de elución de NaAc 25 mM, a pH 5,0 a NaAc 25 mM, a pH 5,0, NaCl 1 M, consiguió la elución de K3 con una pureza del 95% con NaCl entre 300-500 mM. El análisis de los *N*-glucanos de glucanos derivados de K3 se muestra en la Figura 5E. La alícuota de sobrenadante extraída anteriormente se analizó además para determinar la presencia de actividad de manosidasas segregadas como se muestra en la Figura 9B. Se añadió un patrón disponible comercialmente de $\text{Man}_8\text{-2-AB}$ (Glyko, Novato, CA) a: BMMY (Figura 9A), sobrenadante de la alícuota anterior p8C18-5 (*VAN1(s)*) de *Saccharomyces/manosidasa* IB de *C. elegans*, $\Delta 80$) (Figura 9B), y BMMY que contenía medios de una construcción fusionada diferente pDD28-3 (*MNN10(m)*) de *Saccharomyces* (de SwissProt 50108)/manosidasa IB de *H. sapiens*, $\Delta 99$) (Figura 9C). Después de incubar durante 24 horas a temperatura ambiente, se analizaron las muestras mediante HPLC sobre aminosílice para determinar el grado de recorte de las manosidasas. La Figura 9B demuestra la actividad de manosidasa intracelular comparada con una construcción fusionada pDD28-3 (*MNN10(m)*) de *Saccharomyces/manosidasa* IB de *H. sapiens* $\Delta 99$) que muestra un resultado negativo (Figura 9C).

Ejemplo 7

Ensayo a pH óptimo de α -1,2-manosidasa genommanipulada

Se cultivaron células de *P. pastoris* que comprendían el plásmido pBB27-2 (*MNN10(S)*) de *Saccharomyces* (de SwissProt 50108)/manosidasa IB de *C. elegans*, $\Delta 31$) a temperatura ambiente en BMGY a una DO_{600} de aproximadamente 17. Se inocularon aproximadamente 80 μl de estas células en 600 μl de BMGY y se cultivaron toda la noche. Posteriormente, las células se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. Después de 24 horas de inducción, las células se eliminaron por centrifugación proporcionando un sobrenadante esencialmente transparente (pH 6,43). Se extrajo el sobrenadante para realizar ensayos de pH óptimo para las manosidasas. Se añadió $\text{Man}_8\text{GlcNAc}_2$ marcado con fluorescencia (0,5 μg) a 20 μl de sobrenadante ajustado a diversos pH (Figura 11) y se incubó durante 8 horas a temperatura ambiente. Tras la incubación, la muestra se analizó mediante HPLC usando una columna de sílice aminoconjugado con Econosil NH2 de 4,6 X 250 mm, perlas de 5 micrómetros (Altech, Avondale, PA). El caudal era de 1,0 ml/min durante 40 min y la columna se mantuvo a 30°C. Después de eluir isocráticamente (68% de A:32% de B) durante 3 min, se empleó un gradiente lineal del disolvente (68% de A:32% de B a 40% de A:60% de B) durante 27 min para eluir los glucanos (18). El disolvente A (acetonitrilo) y el disolvente B (formiato de amonio, 50 mM, a pH 4,5). La columna se equilibró con disolvente (68% de A:32% de B) durante 20 min entre tandas.

Ejemplo 8

Genomanipulación de *P. pastoris* para producir *N*-glucanos con la estructura $\text{GlcNAcMan}_5\text{GlcNAc}_2$

Es necesaria la actividad de GlcNAc transferasa I para la maduración de *N*-glucanos complejos e híbridos (patente de Estados Unidos N.º 5.834.251). Los $\text{Man}_5\text{GlcNAc}_2$ s únicamente pueden ser recortados por manosidasa II, una etapa necesaria en la formación de glucoformas humanas, después de la adición de *N*-acetilglucosamina al resto de α -1,3 manosa terminal del tallo de trimanosas mediante GlcNAc transferasa I (Schachter, 1991 Glycobiology 1 (5):453-461). Por consiguiente, se preparó una adenoteca de combinaciones que incluía fragmentos de ADN que codificaban dominios catalíticos dirigidos adecuadamente de los genes de GlcNAc Transferasa I de *C. elegans* y de *Homo sapiens*; y secuencias de localización de GLS, MNS, SEC, *MNN9*, *VAN1*, *ANP1*, *HOC1*, *MNN10*, *MNN11*, *MNT1*, *KTR1*, *KTR2*, *MNN2*, *MNN5*, *YUR1*, *MNN1*, y *MNN6* de *S. cerevisiae* y α 1,2-manosiltransferasas putativas de *P. pastoris* basándose en la homología de *S. cerevisiae*: *D2*, *D9* y *D3*, que son homólogos de *KTR*. La Tabla 10 incluye pero sin limitación secuencias de péptidos directores como por ejemplo *SEC* y *OCH1*, de *P. pastoris* y GnTI de *K. lactis*, (Véase la Tabla 6 y Tabla 10).

ES 2 336 790 T3

TABLA 10

Una adenoteca representativa de combinaciones de secuencias de péptidos directores y dominio catalítico para UDP-N-Acetilglucosaminil Transferasa I (GnTI)

		Péptido director				
		OCH1(s)	OCH1(m)	OCH1(l)	MNN9(s)	MNN9(m)
5	Dominio catalítico					
10		Ser humano, GnTI, Δ38	PB105	PB106	PB107	PB104
15						N/A
20		Humana, GnTI, Δ86	NB12	NB13	NB14	NB15
25						NB
30		C. elegans, GnTI, Δ88	OA12	OA13	OA14	OA15
35						OA16
40		C. elegans, GnTI, Δ35	PA12	PA13	PA14	PA15
45						PA16
50		C. elegans, GnTI, Δ63	PB12	PB13	PB14	PB15
55						PB16
60		X. leavis, GnTI, Δ33	QA12	QA13	QA14	QA15
65						QA16
70		X. leavis, GnTI, Δ103	QB12	QB13	QB14	QB15
75						QB16

Se seleccionaron secuencias peptídicas de direccionamiento a partir de *och1* en *P. pastoris* (largo, medio y corto) (véase el Ejemplo 4) y *MNN9* (SwissProt P391 07) en *S. cerevisiae* (corto y medio). Los dominios catalíticos se seleccionaron a partir de GnTI humana con una delección de 38 y 86 aminoácidos en el extremo N, GnTI de *C. elegans* (gly-12) con una delección de 35 y 63 aminoácidos así como GnTI de *C. elegans* (gly-14) con una delección de 88 aminoácidos en el extremo N y *X. leavis* GnTI con una delección de 33 y 103 aminoácidos en el extremo N, respectivamente.

Una porción del gen que codificaba N-acetilglucosaminil transferasa I humana (MGATI, n.º de acceso NM002406), que carecía de las primeras 154 pb, se amplificó mediante PCR usando los oligonucleótidos 5' -TGGCAGGCGCGCC TCAGTCAGCGCTCTCG3' (SEC. ID. N.º: 72) y 5'-AGGTAAATTA AGTGCTAATTCCAGCTAGG-3' (SEC. ID. N.º: 73) y el vector pHG4.5 (n.º de ATCC 79003) como plantilla. El producto de la PCR resultante se clonó en pCR2.1-TOPO y se confirmó la secuencia correcta. Tras la digestión con *AscI* y *PacI*, la GnTI se insertó en el plásmido pJN346

para crear pNA. Después de la digestión de pJN271 con *NotI* y *AscI*, el inserto de 120 pb se ligó en pNA para generar una fusión en marco del dominio transmembrana de *MNN9* con la GnTI, para crear pNA15.

El organismo huésped es una cepa de *P. pastoris* que es deficiente en hipermanosilación (por ejemplo un mutante de *och1*), proporciona el sustrato UDP-GlcNAc en el aparato de Golgi y/o RE (*es decir*, contiene un transportador de UDP-GlcNAc funcional), y proporciona *N*-glucanos de la estructura $\text{Man}_5\text{GlcNAc}_2$ en el aparato de Golgi y/o RE (por ejemplo pFB8 de *P. pastoris* (*SEC12(m)* de *Saccharomyces/manosidasa IA* de ratón A187) anterior). Primero, pFB8 de *P. pastoris* se transformó con pPB103 que contenía el gen *MNN2-2* de *Kluyveromyces lactis* (Genbank AN AF1 06080) (que codifican transportador de UDP-GlcNAc) se clonó en el sitio *BamHI* y *BglII* del plásmido pBLADE-SX (Cereghino *et al.* (2001) Gene 263:159-169). Después, la anteriormente mencionada adenoteca de combinaciones que codifica una combinación de GnTI exógena o endógena/genes de localización se transformó y se seleccionaron las colonias y se analizaron para determinar la presencia de la construcción de GnTI mediante PCR de colonias. La eficacia de transformación e integración de los inventores de forma general estuvo por encima del 80% y puede omitirse el cribado por PCR una vez se han establecido parámetros de transformación sólidos.

Purificación de proteínas

K3 se purificó a partir del medio mediante cromatografía por afinidad de Ni utilizando un formato de 96 pocillos en un robot de laboratorio Beckman BioMek 2000. La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind. Puede realizarse otro procedimiento de cribado usando un anticuerpo que se une a un GlcNAc terminal específico, o a una lectina como por ejemplo la lectina GSII de *Griffonia simplicifolia*, que se une a GlcNAc terminal (EY Laboratories, San Mateo, CA). Estos cribados pueden automatizarse usando lectinas o anticuerpos que han sido modificados con marcadores fluorescentes como por ejemplo FITC o que se han analizado mediante MALDI-TOF.

La K3 segregada puede purificarse mediante cromatografía por afinidad de Ni, cuantificarse y pueden unirse cantidades iguales de proteína a una placa de 96 pocillos alta en proteína. Después de bloquear con BSA, las placas pueden sondarse con una lectina GSII-FACS y cribarse para determinar la máxima respuesta fluorescente. Un procedimiento preferente para detectar las anteriores proteínas glucosiladas supone el cribado mediante espectrometría de masas MALDI-TOF tras la purificación por afinidad de K3 segregada a partir del sobrenadante de los transformantes cultivados en 96 pocillos. Las colonias transformadas se escogieron y se cultivaron a una DO_{600} de 10 en una placa de 96 pocillos de 2 ml en BMGY a 30°C. Las células se recolectaron mediante centrifugación, se lavaron en BMMY y se resuspendieron en 250 μl de BMMY. A las 24 horas de la inducción, las células se retiraron por centrifugación, se recuperó el sobrenadante y K3 se purificó del sobrenadante mediante cromatografía por afinidad de Ni. Se liberaron los *N*-glucanos y se analizaron mediante espectrometría de masas MALDI-TOF con extracción retardada como se describe en el presente documento.

En resumen, los procedimientos de la invención producen cepas de *P. pastoris* que producen $\text{GlcNAcMan}_5\text{GlcNAc}_2$ en alto rendimiento, como se muestra en la Figura 10B. Al menos 60% de los *N*-glucanos son $\text{GlcNAcMan}_5\text{GlcNAc}_2$. Hasta la fecha, no existen informes que describan la formación de $\text{GlcNAcMan}_5\text{GlcNAc}_2$ en glucoproteínas segregadas solubles en ninguna levadura. Los resultados presentados en el presente documento muestran que la adición del transportador de UDP-GlcNAc junto con la actividad de GnTI producen una estructura $\text{GlcNAcMan}_5\text{GlcNAc}_2$ predominante, que se confirma por el pico en 1457 (m/z) (Figura 10B).

Construcción de la cepa PBP-3

La cepa de *P. pastoris* que expresa K3, ($\Delta\text{och } 1$, arg-, ade-, his-) se transformó sucesivamente con los siguientes vectores. Primero, pFB8 (*SEC12(m)* de *Saccharomyces/manosidasa IA* de ratón Δ 187) se transformó en la cepa de *P. pastoris* por electroporación. Segundo, pPB103 que contenía el gen *MNN2-2* de *Kluyveromyces lactis* (Genbank AN AF106(80) (que codifica el transportador de UDP-GlcNAc) se clonó en el plásmido pBLADE-SX (Cereghino *et al.* (2001) Gene 263:159-169) se digirió con las enzimas *BamHI* y *BglII*, se transformó en la cepa de *P. pastoris*. Tercero, pPB104 que contenía *MNN9(s)* de *Saccharomyces/GnTI humana* Δ 38 que codifica el gen clonado como el fragmento *NotI-PacI* en pJN336 se transformó en la cepa de *P. pastoris*.

Ejemplo 9

Genomanipulación de células de *K. lactis* para producir *N*-glucanos con la estructura $\text{Man}_5\text{GlcNAc}_2$

Identificación y alteración del gen *OCH1* de *K. lactis*

El gen *OCH1* de la levadura *S. cerevisiae* en desarrollo codifica una 1,6-manosiltransferasa que es responsable de la primera adición de manosa localizada en el aparato de Golgi a la estructura del *N*-glucano $\text{Man}_3\text{GlcNAc}_2$ sobre las proteínas segregadas (Nakanishi-Shindo *et al.* (1993) J. Biol. Chem.; 268(35):26338-45). Esta transferencia de manosa de forma general se reconoce como la etapa inicial clave en la polimanosilación fúngica específica de estructuras de *N*-glucano (Nakanishi-Shindo *et al.* (1993) J. Biol. Chem. 268(35):26338-26345; Nakayama *et al.* (1992) EMBO J. 11(7):2511-19; Mori-N-Ganet *et al.* (2000) Traffic 1(1):56-68). La delección de este gen en *S. cerevisiae* produce una estructura de *N*-glucano significativamente más corta que no incluye esta polimanosilación típica ni una falta de desarrollo a temperaturas elevadas (Nakayama *et al.* (1992) ENIBO J. 11(7):2511-19).

La secuencia de *Och1p* de *S. cerevisiae* se alineó con homólogos conocidos de *Candida albicans* (n.º de acceso de Genbank AAL49987), y *P. pastoris* junto con las proteínas *Hoc1* de *S. cerevisiae* (Neiman *et al* (1997) Genetics 145(3):637-45 y *K. lactis* (base de datos PENDANT EST) que son manosiltransferasas relacionadas pero diferentes. Se usaron las regiones con alta homología que tenían en común los homólogos de *Och1p* pero eran distintas de los homólogos de *Hoc1p* para diseñar pares de cebadores degenerados que se dirigieron contra el ADN genómico de la cepa MG1/2 de *K. lactis* (Bianchi *et al* (1987) Current Genetics 12:185-192). La amplificación por PCR con los cebadores RCD33 (CCAGMGAATTCAATTGTGTCARTGG) (SEC. ID. N.º: 74) y RCD34 (CAGTGAAAATACCTGGNCCNGTCCA) (SEC. ID. N.º: 75) consiguieron un producto de 302 pb que se clonó y se secuenció y se demostró que la traducción prevista tenía un gran grado de homología con las proteínas *Och 1* (>55% con *Och 1 p* de *S. cerevisiae*).

El producto de la PCR de 302 pb se usó para sondear una transferencia Southern de ADN genómico de la cepa (MG1/2) de *K. lactis* con condiciones muy rigurosas (Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Se observó hibridación con un patrón coherente con un único gen que indicaba que este segmento de 302 pb corresponde a una porción del genoma de *K. lactis* y *K. lactis* (*KIOCH1*) contiene una única copia del gen. Para clonar el gen *KIOCH1*, se usó la transferencia Southern para mapear el locus genómico. Por consiguiente, se clonó un fragmento *BamHI/PstI* de 5,2 kb digiriendo el ADN genómico y ligando esos fragmentos en el abanico de 5,2 kb en pUC19 (New England Biolabs, Beverly, MA) para crear una subadenoteca genómica de *K. lactis*. Esta adenoteca subgenómica se transformó en *E. coli* y se analizaron varios cientos de clones mediante PCR de colonias usando RCD 33/34. El clon de 5,2 kb que contenía el gen *KIOCH1* previsto se secuenció y un marco de lectura abierto de 1362 pb que codificaba una proteína prevista que es idéntica en un 46,5% al gen *OCH1* de *S. cerevisiae*. Se usó la secuencia de 5,2 kb para preparar los cebadores para la construcción de un alelo de delección *och1::KAN^R* usando un procedimiento de superposición de PCR (Davidson *et al.* (2002) Microbiol. 148(Pt 8):2607-15). Este alelo de delección se transformó en dos cepas de *K. lactis* y se seleccionaron colonias resistentes G418. Estas colonias se cribaron mediante PCR y en base a la sensibilidad a la temperatura obteniendo una cepa con el ORF de *OCH1* eliminado. Los resultados del experimento muestran cepas que revelan un patrón de PCR mutante, que se caracteriza por el análisis del desarrollo a diversas temperaturas y el análisis de carbohidratos *N*-glucanos segregados y proteínas de la pared celular tras la digestión con PNGasas. La mutación *och1* confería una sensibilidad a la temperatura que permitía que las cepas se desarrollaran a 30°C pero no a 35°C. La Figura 12A muestra un análisis por MALDI-TOF de una cepa de *K. lactis* de tipo silvestre que producía *N*-glucanos de Man₃GlcNAc₂ [c] y superiores.

Identificación, clonación, y alteración del gen *MNN1* de *K. lactis*

MNN1 de *S. cerevisiae* es el gen estructural para la α -1,3-manosiltransferasa del aparato de Golgi. El producto de *MNN1* es una proteína de membrana de tipo II de 762 aminoácidos (Yip *et al.* (1994) Proc Natl Acad Sci USA. 91 (7):2723-7). Tanto los oligosacáridos ligados en N y ligados en D aislados a partir de mutantes en *mnn1* carecen de enlaces α -1,3-manosa (Raschke *et al.* (1973) J Biol. Chem, 248(13):4660-66).

Se usó la secuencia de *Mnn1p* de *S. cerevisiae* para buscar las secuencias genómicas traducidas de *K. lactis* (PENDANT). Se identificó una secuencia de ADN de 405 pb que codificaba fragmento de proteína putativo de similitud considerable a *Mnn1p*. Posteriormente se amplificó por PCR un segmento de esta secuencia con los cebadores KMN1 (TGCCATCTTTTAGGTCCAGGCCCGTTC) (SEC. ID. N.º: 76) y KMN2 (GATCCCACGACGCATCGTATTTCTTTC), (SEC. ID. N.º: 77) y se usaron para sondear una transferencia Southern de ADN genómico de la cepa MG1/2 de *K. lactis*. Basándose en los datos de hibridación de la Southern se clonó un fragmento *BamHI-PstI* de 4,2 Kb generando una adenoteca seleccionada por el tamaño como se describe en el presente documento. Se identificó un único clon que contenía el gen *MNN1* de *K. lactis* mediante PCR de colonias entero usando los cebadores KMN1 y KMN2 y se secuenció. Dentro de este clon, se identificó un ORF de 2241 pb que codificaba una proteína predicha que era un 34% idénticas al gen *MNN1* de *S. cerevisiae*. Se diseñaron cebadores para la construcción de un alelo por delección de *mnn1::NAT^R* usando el procedimiento de superposición (Davidson *et al.* (2002) Microbiol. 148(Pt 8):2607-15).

Este alelo alterador se transformó en una cepa de *K. lactis* mediante electroporación y se seleccionaron los transformantes resistentes a nourseotricina y se amplificaron por PCR para determinar la inserción homóloga del alelo alterador. Las cepas que revelan un patrón mutante en la PCR pueden someterse a análisis de los carbohidratos *N*-glucanos de un gen testigo conocido.

La Figura 12B representa los *N*-glucanos de la cepa de *K. lactis* con la delección *och1 mnn1* observada tras la digestión con PNGasas. La MALDI-TOF como se describe en el presente documento. El pico predominante en 1908 (*m/z*) indicado con [d] es coherente con la masa de Man₉GlcNAc₂.

En la bibliografía se describen procedimientos y reactivos adicionales que pueden usarse en los procedimientos para modificar la glucosilación, como por ejemplo en la patente de Estados Unidos n.º 5.955.422, en la patente de Estados Unidos n.º 4.775.622, en la patente de Estados Unidos n.º 6.017.743, en la patente de Estados Unidos n.º 4.925.796, en la patente de Estados Unidos n.º 5.766.910, en la patente de Estados Unidos n.º 5.834.251, en la patente de Estados Unidos n.º 5.910.570, en la patente de Estados Unidos n.º 5.849.904, en la patente de Estados Unidos n.º 5.955.347, en la patente de Estados Unidos n.º 5.962.294, en la patente de Estados Unidos n.º 5.135.854, en la patente

de Estados Unidos n.º 4.935.349, en la patente de Estados Unidos n.º 5.707.828, y en la patente de Estados Unidos n.º 5.047.335. Los sistemas de expresión en levaduras apropiados pueden obtenerse de fuentes como por ejemplo la American Type Culture Collection, Rockville, MD. Los vectores están disponibles comercialmente de una variedad de fuentes.

Ejemplo 10

Identificación, clonación y delección del gen *ALG3* en *P. pastoris* y *X. lactis*

Se generaron cebadores degenerados basándose en una alineación de secuencias de la proteína Alg3 de *S. cerevisiae*, *H. sapiens*, y *D. melanogaster* y se usaron para amplificar un producto de 83 pb de ADN genómico de *P. pastoris*:

5'-GGTGTGTTTGTGTTTCTAGATCTITGCAYTAYCARTT-3' (SEC. ID. N.º: 78) y 5'-AGMTITGGTGGGTGGM TTCCARCACCAYTCRTG-3' (SEC. ID. N.º: 79). El producto resultante de la PCR se clonó en el vector pCR2.1 (Invitrogen, Carlsbad, CA) y el análisis de las secuencias reveló homología con los homólogos conocidos *ALG3/RHK1NOT56* (Genbank NC_001134.2, AF309689, NC_003424.1). Posteriormente, se amplificaron 1929 pb aguas arriba y 2738 pb aguas abajo del producto de la PCR inicial a partir de una adenoteca genómica de *P. pastoris* (Boehm (1999) Yeast 1S(7):S63-72) usando los oligonucleótidos internos 5'-CCTAAGCTGGTATGCGTTCTCTTTGCCA TATC-3' (SEC. ID. N.º: 80) y 5'-GCGGCATAAACAATAATAGATGCTATAAAG-3' (SEC. ID. N.º: 81) junto con T3 (5'-AATTAACCCTCACTAAAGGG-3') (SEC. ID. N.º: 49) y T7 (5'-GTAATACGACTCACTATAGGGC-3') (SEC. ID. N.º: 48) (Integrated ADN Technologies, Coralville, IA) en el esqueleto de la adenoteca que portaba el plásmido lambda ZAP II (Stratagene, La Jolla, CA). Los fragmentos resultantes se clonaron en el vector pCR2.1-TOPO (Invitrogen) y se secuenciaron. A partir de esta secuencia, se identificó un ORF de 1395 pb que codificaba una proteína con una identidad del 35% y una similitud del 53% con el gen *ALG3* de *S. cerevisiae* (usando programas BLAST). El gen se denominó *PpALG3*.

Se usó la secuencia de *PpALG3* para crear un conjunto de cebadores para generar una construcción por delección del gen *PpALG3* mediante superposición por PCR (Davidson *et al* (2002) Microbiol. 148(Pt 8):2607-15). Los cebadores siguientes se usaron para amplificar regiones de 1 kb en 5' y 3' del ORF de *PpALG3* y el gen *KAN^R*, respectivamente:

RCD142 (5'-CCACATCATCCGTGCTACATATAG-3') (SEC. ID. N.º: 82),

RCD144 (5'-ACGAGGCAAGCTAAACAGATCTCGAAGTATCGAGGGTTATCCAG-3') (SEC. ID. N.º: 83),

RCD145 (5'-CCATCCAGTGTCGAAAACGAGCCAATGGTTCATGTCTATAAATC-3') (SEC. ID. N.º: 84),

RCD147 (5'-AGCCTCAGCGCCAACAAGCGATGG-3') (SEC. ID. N.º: 85),

RCD143 (5'-CTGGATAACCCTCGATACTTCGAGATCTGTTTAGCTTGCCTCGT-3') (SEC. ID. N.º: 86), y

RCD146 (5'-GATTTATAGACATGAACCATTGGCTCGTTTTTCGACACTGGATGG-3') (SEC. ID. N.º: 87).

Posteriormente, se usaron los cebadores RCD142 y RCD147 para superponer los tres productos resultantes de la PCR en un único alelo de delección *Alg3::KAN^R* de 3,6 kb.

Identificación, clonación y delección del gen *ALG3* en *K. lactis*

Las secuencias de *ALG3p* de *S. cerevisiae*, *Drosophila melanogaster*, *Homo sapiens* *etc.* se alinearon con las secuencias de *K. lactis* (base de datos PENDANT EST). Se usaron las regiones con alta homología que tenían en común los homólogos en secuencia exacta pero que eran distintas de los homólogos para crear pares de cebadores degenerados que se dirigieron contra el ADN genómico de la cepa MG1/2 de *K. lactis* (Bianchi *et al* (1987). En el caso de *ALG3*, la amplificación por PCR con los cebadores KAL-1 (5'-ATCCTTTACCGATGCTGTAT-3') (SEC. ID. N.º: 88) y KAL-2 (5'-ATAACAGTATGTGTACACGCGTGTAG-3') (SEC. ID. N.º: 89) obtuvo un producto que se clonó y se secuenció y se demostró que la traducción prevista tenía un gran grado de homología con las proteínas *Alg3p* (>50% con *Alg3p* de *S. cerevisiae*).

El producto de la PCR se usó para sondar una transferencia Southern de ADN genómico de la cepa (MG1/2) de *K. lactis* con condiciones muy rigurosas (Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Se observó hibridación en un patrón coherente con un único gen. Esta transferencia Southern se usó para mapear los locus genómico. Se clonaron fragmentos genómicos digiriendo el ADN genómico y ligando esos fragmentos con el intervalo de tamaño apropiado en pUC19 para crear una subadenoteca genómica de *K. lactis*. Esta subadenoteca genómica se transformó en *E. coli* y se analizaron varios cientos de clones mediante PCR de colonias usando los cebadores KAL-1 y KAL-2. Los clones que contenían los genes *KIALG3* y *KIALG61* previstos se secuenciaron y se identificaron los marcos de lectura abiertos.

Se diseñaron cebadores para la construcción de un alelo de delección *Alg3::NAT^R*, usando un procedimiento de superposición (Davidson *et al.* (2002) Microbiol. 148(Pt 8):2607-15), y el alelo de delección resultante se transformó en dos cepas de *K. lactis* y se seleccionaron colonias resistentes a NAT. Esta colonias se cribaron mediante PCR y se obtuvieron transformantes en los que el ORF de *ALG3* fue sustituido por el alelo mutante *och1::NAT^R*.

Ejemplo 11

Generación de una cepa mutante para Alg3 que expresa una α -1,2-manosidasa, GnTI y GnTII para la producción de una glucoproteína humanoide

Se generó una construcción por delección de *Alg3::KAN^R* de *P. pastoris* como se describe en el Ejemplo 10. Aproximadamente 5 μ g del producto resultante de la PCR se transformó en la cepa PBP-3 (véase Ejemplo 3), y las colonias se seleccionaron sobre medio YPD que contenía 200 μ g/ml de G418. Se confirmó que una cepa de las 20 cribadas mediante PCR contenía la integración correcta del alelo mutante *Alg3::KAN^R* y carecía del alelo de tipo silvestre. Esta cepa se denominó RDP27 (Figura 36).

Después se generó una adenoteca de construcciones de GnTII, que comprendía fusiones en marco de la misma adenoteca líder con los dominios catalíticos de los genes de GnTII de fuentes humana, y de rata (WO 02/00879). Esta adenoteca se creó en un vector de integración de *P. pastoris* que contenía el gen NST^R que confiere resistencia al fármaco nourseotricina. Los plásmidos de la adenoteca se linearizaron con EcoRI, transformado en la cepa RDP27 mediante electroporación, y las cepas resultantes se cribaron mediante análisis de los glucanos liberados de K3 purificado. Una cepa de *P. pastoris* que expresaba el GnTII de rata fusionado en marco a la construcción *MNN9* (s) de *S. cerevisiae* se denominó PBP6-5 (Figura 36).

Generación de construcciones de expresión de GnTII

La construcción de un vector de expresión de GnTI (pNA 15) que contenía un gen de GnTI humano fusionado con la parte del extremo N del gen *MNN9* de *S. cerevisiae* se describe en Choi *et al.* (2003) Proc Natl Acad Sci USA. 100 (9):5022-27. De forma similar, se clonó el gen GnTII de rata. Se amplificó el gen GnTII de rata (número de acceso de GenBank U21662) por PCR usando la polimerasa Takara EX TaqTM (Panvera) de una adenoteca de ADNc de hígado de rata (Clontech) con los cebadores RAT1 (5'-TTCCTCACTGCAGTCTTCTATAACT-3') (SEC. ID. N.º: 90) y RAT2 (5'-TGGAGACCATGAGGTTCCGCATCTAC-3') (SEC. ID. N.º: 91). El producto de la PCR después se clonó en el vector pCR2.1-TOPO (Invitrogen) y se secuenció. Usando este vector como plantilla, se amplificó el fragmento *Ascl-PacI* de GnTII, que codifica los aminoácidos 88-443, con Pfu Turbopolimerasa (Stratagene) y los cebadores, RAT44 (5'-TTGGCGCGCCTCCCTAGTGTACCAGTTGAACCTTTG-3') (SEC. ID. N.º: 92) y RAT11 (5'-GATTAATTAACCTCACTGCAGTCTTCTATAACT-3') (SEC. ID. N.º: 93) respectivamente (los sitios de restricción *Ascl* y *PacI* introducidos están subrayados). Tras la confirmación mediante secuenciación, se clonó el dominio catalítico de GnTII de rata aguas abajo del promotor de *PMA 1* en forma de un fragmento *Ascl-PacI* en pBP124. En la etapa final, se clonó el fragmento de un gen que codificaba la señal de localización de *MNN2* de *S. cerevisiae* a partir de pJN281 en forma de fragmento *NotI-Ascl* para generar una fusión en marco con el dominio catalítico de GnTII, para generar el plásmido pTC53.

Ejemplo 12

Clonación y expresión de GnTIII para producir GlcNAc diseccionadora que potencia la función de los anticuerpos

La adición de una *N*-acetilglucosamina a la estructura de GlcNAc₂Man₃GlcNAc₂ mediante las *N*-acetilglucosaminiltransferasas III proporciona lo que se denomina *N*-glucano diseccionado (véase la Figura 15). Esta estructura ha sido implicada en una mayor citotoxicidad celular dependiente de anticuerpos (ADCC). (1999) Nat. Biotechnol. 17 (2):176-80).

Una célula huésped como por ejemplo una cepa de levadura capaz de producir glucoproteínas con *N*-glucanos diseccionados se genomanipula introduciendo en las células huésped de acuerdo con la invención una actividad de GnTIII. Preferentemente, la célula huésped se transforma con un ácido nucleico que codifica GnTIII (por ejemplo, una mamífera como por ejemplo la CnTIII murina que se muestra en la Figura 24) o un dominio de la misma que tiene actividad enzimática, opcionalmente fusionado a una péptido de señal director celular heterólogo (por ejemplo, usando las adenotecas y procedimientos asociados de la invención).

Las IgG están formadas por dos cadenas pesadas (V_H, C_{H1}, C_{H2} y C_{H3} en las Figura 22), interconectadas en la región de bisagra a través de tres puentes disulfuro, y dos cadenas ligeras (V_L, C_L en la Figura 22). Las cadenas ligeras (dominios V_L y C_L) están ligadas a otro puente disulfuro en la porción C_{H1} de la cadena pesada y, junto con los fragmentos C_{H1} y V_H, conforman la denominada región Fab. Los antígenos se unen a la porción terminal de la región Fab. La región Fc de las IgG está constituida por C_{H3}, C_{H2} y la región de bisagra y es responsable de ejercer las que se denominan funciones efectoras (véase más adelante). La función primaria de los anticuerpos es unirse a un antígeno. Sin embargo, a no ser que la unión al antígeno inactivo directamente el antígeno (como por ejemplo en el caso de las toxinas bacterianas), la mera unión carece de sentido a no ser que desencadene las denominadas funciones

efectoras. Los anticuerpos de las subclase de IgG ejercen dos funciones efectoras principales: la activación del sistema de complemento y la inducción de fagocitosis. El sistema de complemento consiste en un grupo complejo de proteínas serosas implicadas en controlar los eventos inflamatorios, en la activación de fagocitos y en la destrucción lítica de las membranas celulares. La activación de complemento se inicia con la unión del complejo C1 a la porción Fc de dos IgG cercanas. C1 está constituido por una molécula, C1q, y dos moléculas, C1r y C1s. La fagocitosis se inicia a través de una interacción entre el fragmento Fc de la IgG y los receptores de Fc-gamma (FcγR1, II y III en la Figura 22). Los receptores de Fc se expresan principalmente sobre la superficie de células efectoras del sistema inmunitario, en particular macrófagos, monocitos, células mieloides y células dendríticas.

La porción C_H2 alberga un sitio de *N*-glucosilación conservada en la asparragina 297 (Asn297). Los *N*-glucanos de Asn297 son muy heterogéneos y se sabe que afectan la unión al receptor de Fc y la activación de complemento. Sólo una minoría (*es decir*, aproximadamente 15-20%) de las IgG porta un *N*-glucano disialilado, y 3-10% tienen uno monosialilado (revisado en Jefferis (2001) *Biopharm.* 14:19-26). De forma interesante, la estructura de *N*-glucano mínima que se demuestra que es necesaria para anticuerpos totalmente funcionales capaces de activar complemento y de unión al receptor Fc es un pentasacárido con restos *N*-acetilglucosamina terminales (GlcNAc₂Man₃) (revisado en Jefferis, R., *Glycosilation of human IgG Antibodies*. BioPharm, 2001). Los anticuerpos con menos de un *N*-glucano GlcNAc₂Man₃ o sin *N*-glucosilación en Asn297 pueden ser capaces todavía de unirse a un antígeno pero, lo más probablemente es que no activen los eventos cruciales aguas abajo como por ejemplo fagocitosis y activación de complemento. Además, los anticuerpos con *N*-glucanos de tipo fúngico unidos en Asn297 con toda probabilidad desencadenarán una respuesta inmunitaria en un organismo mamífero que hará que el anticuerpo sea inútil como glucoproteína terapéutica.

Clonación y expresión de GnTIII

El fragmento de ADN que codifica parte de la proteína GnTIII de ratón que carece del dominio TM se amplificó por PCR a partir de ADN genómico murino (o de otro mamífero) usando los cebadores directo (5'-TCCTGGCGCGCCTT CCGAGAGAACTGGCCTCCCTC-3') (SEC. ID. N.º: 94) y inverso (5'-AATTAATTAACCCCTAGCCCTCCGCTG TATCCAACCTTG-3') (SEC. ID. N.º: 95). Esos cebadores incluyen los sitios de restricción *AscI* y *PacI* que pueden usarse para la clonación en el vector adecuado para la fusión con la adenoteca líder.

Las secuencias de ácido nucleico (SEC. ID. N.º: 45) y aminoácidos (SEC. ID. N.º: 46) de GnTIII murina se muestran en la Figura 24.

Clonación de secuencias que codifican inmunoglobulinas

Anteriormente se han publicado protocolos para la clonación de las regiones variables de anticuerpos, que incluyen secuencias cebadoras. Las fuentes de anticuerpos y los genes que codifican pueden ser, entre otras, linfocitos B humanos inmunizados *in vitro* (véase, por ejemplo, Borreback *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:3995-3999), linfocitos de sangre periférica o linfocitos B humanos únicos (véase, por ejemplo, Lagerkvist *et al.* (1995) *Biotechniques* 18:862-869; y Terness *et al.* (1997) *Hum. Immunol.* 56:17-27) y ratones transgénicos contenían locus de inmunoglobulina humana, que permite la creación de líneas celulares de hibridomas.

Usando técnicas de genomanipulación estándar, pueden clonarse secuencias de ácido nucleico que codifican anticuerpos. Las fuentes para la información genética que codifica inmunoglobulinas de interés son preparaciones de ARN total de las células de interés, como por ejemplo linfocitos sanguíneos o líneas celulares de hibridomas. Por ejemplo, empleando un protocolo con base de PCR con cebadores específicos, pueden clonarse regiones variables mediante transcripción inversa iniciada a partir de un cebador específico de secuencia que se hibrida al sitio del dominio C_H1 de la IgG y un segundo cebador que codifica los aminoácidos 111-118 de la región constante kappa murina. Los ADNc que codifican V_H y V_L después pueden amplificarse como se ha publicado anteriormente (véase, por ejemplo, Graziano *et al.* (1995) *J Immunol.* 155(10):4996-5002; Welschof *et al.* (1995) *J. Immunol. Methods* 179:203-214; y Orlandi *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 86:3833). También se han publicado procedimientos de clonación para inmunoglobulinas enteras (cadenas pesadas y ligeras) (véase, por ejemplo, Buckel *et al.* (1987) *Gene* 51:13-19; Recinos *et al.* (1994) *Gene* 149:385-386; Recinos *et al.* (1995) *Gene* 158:311-12). Se han descrito protocolos adicionales para la clonación y generación de construcciones de fragmentos de anticuerpos y para la expresión de anticuerpo en Antibody Engineering, Kontermann y Dübel (2001), Eds., Springer Verlag: Berlin Heidelberg Nueva York.

Se han descrito plásmidos de expresión fúngicos que codifican las cadenas pesadas y ligeras de las inmunoglobulinas (véase, por ejemplo, Abdel-Salam *et al.* (2001) *Appl. Microbiol. Biotechnol.* 56:157-164; y Ogunjimi *et al.* (1999) *Biotechnology Letters* 21:561-567). Así, se pueden generar plásmidos de expresión que albergan las regiones constantes de las inmunoglobulinas. Para facilitar la clonación de regiones variables en estos vectores de expresión, pueden situarse sitios de restricción adecuados cercanos a los extremos de las regiones variables. Las regiones constantes pueden construirse de tal manera que las regiones variables puedan fusionarse fácilmente en marco a ellas mediante un simple experimento de digestión con enzimas de restricción y ligado. La Figura 23 muestra un esquema resumen de una construcción de expresión así, diseñada de manera muy modular, que permite un intercambio fácil de promotores, terminadores de la transcripción, dominios de dirección de la integración e incluso marcadores de la selección.

Como se muestra en la Figura 23, los dominios V_L y V_H elegidos pueden clonarse fácilmente en marco con las regiones C_L y C_H, respectivamente. La integración inicial se dirige al locus *AOX* de *P. pastoris* (o a un locus homólogo

en otra célula fúngica) y el promotor *AOX* inducible con metanol dirigirá la expresión. De forma alternativa, puede usarse cualquier otro casete promotor constitutivo o inducible. Así, si se desea, las regiones 5' *AOX* y 3' *AOX* así como los fragmentos de terminación de la transcripción (TT) pueden sustituirse fácilmente con diferentes dominios promotor y de direccionamiento de la integración para optimizar la expresión. Inicialmente, se emplea la señal de secreción del factor alfa con el sitio de proteasa estándar KEX para facilitar la secreción de las cadenas pesadas y ligeras. Las propiedades del vector de expresión pueden refinarse aún más usando técnicas estándar.

Un vector de expresión de Ig como por ejemplo el que se describe anteriormente se introduce en una célula huésped de la invención que exprese GnTIII, preferentemente en el aparato de Golgi de las células huésped. Las moléculas de Ig expresadas en dicha célula huésped comprenden *N*-glucanos que tienen GlcNAc diseccionadora.

Ejemplo 13

Generación de la cepa de levadura YSH-1 (*Δoch1*, α 1,2-manosidasa, GnTI)

Se usó la cepa BK64 de *P. pastoris* previamente reseñada (Choi *et al.* (2003) Proc Natl Acad Sci USA. 100(9):5022-7), un auxótrofo triple (*ADE*, *ARG*, *HIS*) que posee la inactivación en *OCH1* y que expresa el dominio kringle 3 (K3) de plasminógeno humano, como cepa huésped. BK64 se transformó con el plásmido pPB103 linearizado con la enzima de restricción *EcoNI* para introducir el transportador de UDP-*N*-acetilglucosamina de *K. lactis* en las células huésped, creando así la cepa PBP-1. El MnsI de ratón se introdujo en esta cepa mediante transformación con el plásmido pFB8 linearizado con la enzima de restricción *EcoNI*, generando la cepa PBP-2. El análisis de los glucanos de K3 de proteínas aisladas a partir de la cepa PBP-2 demostró que la principal glucoforma presente era Man₅GlcNAc₂.

PBP-2 se transformó posteriormente con el plásmido de GnTI humano pNA15 linearizado con la enzima de restricción *AatII*, generando la cepa PBP-3. El análisis de las glucoformas de K3 producidas en la cepa PBP-3 demostró que el glucano híbrido GlcNAcMan₅GlcNAc₂ era la estructura predominante. Para recuperar el marcador URA3 de PBP-3, esta cepa se cultivó en YPD antes de la selección en medio mínimo que contenía 5-fluoroorótico (5-FOA, BioVectra) y uracilo (Boeke *et al.* (1984) Mol. Gen. Genet. 197:345-346). La cepa Ura- recuperada que producía glucoformas GlcNAcMan₅GlcNAc₂ se denominó YSH-1 (Figura 36). El perfil de *N*-glucano de la cepa YSH-1 se muestra en la Figura 25 (arriba) y presenta un pico predominante en 1465 m/z correspondiente a la masa de GlcNAcMan₅GlcNAc₂ [d].

Ejemplo 14

Generación de la Cepa de levadura YSH-37 (*P. pastoris* que expresa manosidasa II)

YSH-1 (Ejemplo 13) se transformó con el plásmido con manosidasa II de *D. melanogaster*/MNN2(s) de *S. cerevisiae* (pKD53) linearizado con la enzima de restricción *Apal*, generando la cepa YSH-37 (Figura 36). El análisis de las estructuras de glucanos de K3 producidos en la cepa YSH-37 (Figura 25 (abajo)) demostró que la glucoforma predominante en 1140 m/z corresponde a la masa de GlcNAcMan₅GlcNAc₂ [b] y otras glucoformas GlcNAcMan₄GlcNAc₂ [c] en 1303 m/z y GlcNAcMan₅GlcNAc₂ [d] en 1465 m/z.

Ejemplo 15

Generación de la cepa de levadura YSH-44

La cepa YSH-37 (Ejemplo 14) se transformó con un plásmido que codificaba un GnTII de rata/secuencia líder MNN2 (s), pTC53, linearizado con la enzima de restricción *EcoRI*. La cepa resultante, YSH-44 (Figura 36), producía un *N*-glucano de K3 que tiene una única glucoforma en 1356 m/z, correspondiente a la masa de GlcNAc₂Man₃IGlcNAc₂ [x], mediante espectrometría de masas MALDI-TOF en modo positivo (Figura 29).

Digestión con β -acetilhexosaminidasas

Los glucanos de YSH-44 fueron liberados y separados de las glucoproteínas mediante una modificación de un procedimiento previamente reseñado (Papac, *et al.* J. S. (1998) Glycobiology 8, 445-454). Después se redujeron las proteínas y se carboximetilaron y las membranas se bloquearon, los pocillos se lavaron tres veces con agua. La proteína se desglucosiló mediante la adición de 30 μ l de NH₄HCO₃ 10 mM a pH 8,3 que contenía una miliunidad de *N*-glucanasa (Glyko, Novato, CA). Después de 16 horas de digestión a 37°C, la solución que contenía los glucanos se retiró mediante centrifugación y se evaporó a sequedad. Los glucanos después se secaron en una secadora rápida a vacío SC210A (Thermo Savant, Halbrook, NY). Los glucanos secos se introdujeron en NH₄Ac 50 mM a pH 5,0 a 37°C toda la noche y se añadió 1 ml de hexos (Glyko, Novato, CA).

Ejemplo 16

Construcción del plásmido pJN 348

5 El plásmido pBLURA-SX (de Jim Cregg) se digirió con *Bam*HI y *Bgl*III para liberar el casete de expresión AOX. El fragmento BamHI que contenía el casete de expresión GAPDH/CYC1 de pJN261 (Figura 48) (Ejemplo 4) después se ligó en el esqueleto de pBLURA-SX para crear pJN338. El plásmido pJN338 se cortó con *Not*I y *Pac*I y los dos oligonucleótidos 5'-GGCCGCTGCAGATTTAAATGAATTCGGCGCGCCTTAAT-3' (SEC. ID. N.º: 96) y 5'-TAAGGCGCGCCGAATTCATTTAAATCTGCAGGGC-3' (SEC. ID. N.º: 97) que se habían hibridado *in vitro*, se ligaron en los sitios abiertos, para crear pJN348.

Ejemplo 17

Construcción de un plásmido de integración pRCD259

15 El vector de expresión GAPDH que contenía PpURA3, pJN348, se linearizó con *Xho*I y se cortó romo con ADN polimerasa de T4 y se trató con fosfatasa intestinal de ternero (CIP). El marcador de resistencia HYG se digirió a partir de pAG32 con *Bgl*III y *Sac*I y se cortó romo, después se ligó en pJN348 para crear pRCD259 que puede usarse como vector de expresión de HYG que integra el locus PpURA3.

Ejemplo 18

Generación de construcciones fusionadas de GnTIII

25 Se prepararon construcciones fusionadas entre GnTIII de mamífero y secuencias direccionadoras de levaduras usando el gen *Mgat3* de ratón (número de acceso de GenBank L39373, Bhaumik *et al.*, 1995). Se amplificaron por PCR tres fragmentos de ADN correspondientes a las delecciones Δ32, Δ86, y Δ212 del extremo N de gen GnTIII de ratón usando *Pfu Turbo* polimerasa (Stratagene) con los cebadores directos MG3-B (5'-TCCTGGCGCGCCTTCCCGAGA GAACTGGCCTCCCTC-3') (SEC. ID. N.º: 98), MG3-C (5'-CCGAGGCGCGCCACAGAGGAACTGCACCGGG TG-3') (SEC. ID. N.º: 99), MG3-D (5'-ACCGAGGCGCGCCATCAACGCCATCAACATCAACCAC-3') (SEC. ID. N.º: 100), e inversos MG3-A (5'-AATTAATTAACCCTAGCCCTCCGCTGTATCCAACCTTG-3') (SEC. ID. N.º: 101) cebadores. Los productos de la PCR después se clonaron en el vector pJN348 en forma de fragmentos *Asc*I-*Pac*I y se secuenciaron. Los vectores resultantes pVA (GnTIII Δ32), pVB (GnTIII Δ86), y pVC (GnTIII Δ12) se digirieron con enzimas *Not*I-*Asc*I y se usaron para ligar con la adenoteca de secuencias líder de levadura (líderes 20-67). Estos péptidos directores se fusionaron a los dominios catalíticos que se seleccionan a partir de GnTIII de ratón con delecciones de 32, 86 y 212 aminoácidos en el extremo N. Por ejemplo, el péptido director MNN2 de *S. cerevisiae* (largo, medio y corto) y GNTI de *K. lactis* (corto, y medio) (véase Ejemplo 11) se muestran en Tabla 11.

TABLA 11

Una adenoteca de combinaciones de secuencias de péptidos directores/dominios catalíticos representativa que muestra actividad de UDP-NAcetilglucosaminiltransferasa III (GnTIII) en *P. pastoris* YSH-1

Dominio catalítico	Péptido director				
		S. <i>cerevisiae</i> MNN2(s)	S. <i>cerevisiae</i> MNN2(m)	S. <i>cerevisiae</i> MNN2(l)	<i>K. lactis</i> GNT1(m)
GnTIII de ratón, Δ32		50% (pVA53)	30-40% (pVA54)	20-30% (pVA55)	0% (pVA51)
GnTIII de ratón, Δ86		20-30% (pVB53)	30-40% (pVB54)	20-30% (pVB55)	0% (pVB51)
GnTIII de ratón, Δ212		0% (pVC53)	0% (pVC54)	0% (pVC55)	0% (pVC51)

Ejemplo 19

Genomanipulación de P. pastoris para producir GlcNAc₂Man₅GlcNAc₂ diseccionada

La cepa de *P. pastoris* que produce GlcNAcMan₅GlcNAc₂ (PBP-3) (véase el Ejemplo 8) se contraseleccionó en 5-FOA, seleccionando así la pérdida del marcador URA3+ y un fenotipo ura3-. Esta cepa, se denominó YSH-1 (Figura 36), se transformó con los dominios catalíticos de la adenotoca de *N*-acetilglucosaminiltransferasa III (GnTIII) (vectores pYA, pVB, y pVC) y líderes. Los transformantes se cultivaron a 30°C en BMGY a una OD₆₀₀ de aproximadamente 10, se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. K3 se purificó a partir del medio mediante cromatografía por afinidad de Ni utilizando un formato de 96 pocillos en un robot de laboratorio Beckman BioMek 2000. La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind (Ejemplo 3). Los *N*-glucanos se liberaron por digestión con PNGasas (Ejemplo 3). Los *N*-glucanos se analizaron por EM MALDI-TOF (Ejemplo 3). Las actividades de GnTIII se muestran en la Tabla 11. El número de (+), tal como se usa en el presente documento, indica los niveles relativos de producción de *N*-glucanos diseccionados a partir del % de glucanos neutros. Las secuencias peptídicas de direccionamiento se seleccionaron a partir del grupo constituido por: GLS1 de *Saccharomyces*, MNS1 de *Saccharomyces*, SEC12 de *Saccharomyces*, SEC de *Pichia*, OCH1 de *Pichia*, MNN9 de *Saccharomyces*, VAN1 de *Saccharomyces*, ANP1 de *Saccharomyces*, HOC1 de *Saccharomyces*, MNN10 de *Saccharomyces*, MNN11 de *Saccharomyces*, MNT1 de *Saccharomyces*, D2 de *Pichia*, D9 de *Pichia*, J3 de *Pichia*, KTR1 de *Saccharomyces*, KTR2 de *Saccharomyces*, GnTI de *Kluyveromyces*, MNN2 de *Saccharomyces*, MNN5 de *Saccharomyces*, YUR1 de *Saccharomyces*, MNN1 de *Saccharomyces* y MNN6 de *Saccharomyces*. Los transformantes con pVA53 que muestran la GlcNAc diseccionadora (por ejemplo GlcNAc₂Man₅GlcNAc₂:1) se denominaron PBP26 (Figura 36).

Ejemplo 20

Genomanipulación de P. pastoris YSH-44 para producir GlcNAc₃Man₅GlcNAc₂ diseccionada

Para la expresión de GnTIII en la cepa YSH-44 (Figura 36), se transfirieron construcciones de GnTIII desde los vectores pVA53, pVB53, pVA54, y pVB54 en forma de fragmentos *NotI-PacI* en pRCD259 para generar los vectores pPB135, pPB137, pPB136, y pPB 138. Los vectores contienen un marcador de resistencia HYG y el gen *URA3* de *P. pastoris* como secuencia de direccionamiento para la integración genómica. Los plásmidos se linearizan con *SanI*, transformado en la cepa YSH-44 mediante electroporación, se seleccionan sobre medio que contenía higromicina y la cepas resultantes se criban mediante análisis de los glucanos liberados a partir de K3 purificada. Los transformantes se cultivaron a 24°C en BMGY a una OD₆₀₀ de aproximadamente 10, se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. K3 se purificó a partir del medio mediante cromatografía por afinidad de Ni utilizando un formato de 96 pocillos en un robot de laboratorio Beckman BioMek 2000 (Ejemplo 3). La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind (Ejemplo 3). Los *N*-glucanos se liberaron por digestión con PNGasas. Los *N*-glucanos se analizaron por EM MALDI-TOF (Ejemplo 3). Los transformantes pPB135 que muestran la GlcNAc diseccionadora (por ejemplo GlcNAc₂Man₅GlcNAc₂) se denominaron YSH-57 (Figura 36). La Tabla 11 representa la actividad de GnTIII de ratón.

Ejemplo 21

Genomanipulación de P. pastoris PBP6-5 para producir GlcNAc₂Man₅GlcNAc₂ diseccionada

La cepa PBP6-5 de *P. pastoris* (Ejemplo 11) se transformó con el plásmido pPB135 (Tabla 11) que codifica un dominio catalítico de GnTIII de ratón (α 32) ligado en marco a un péptido director derivado de MNN2 de *S. cerevisiae*. Los transformantes se cultivaron a 30°C en BMGY a una OD₆₀₀ de aproximadamente 10, se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. K3 se purificó a partir del medio mediante cromatografía por afinidad de Ni utilizando un formato de 96 pocillos en un robot de laboratorio Beckman BioMek 2000. La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind (Ejemplo 3). Los *N*-glucanos se liberaron por digestión con PNGasas (Ejemplo 3). Los *N*-glucanos se analizaron por EM MALDI-TOF (Ejemplo 3). Los transformantes que muestran la GlcNAc diseccionadora (por ejemplo GlcNAc₂Man₅GlcNAc₂) se denominaron PBP-38 (Figura 36). La Tabla 11 representa la actividad de GnTIII de ratón.

Ejemplo 22

Actividad in vitro del ensayo GnTIII usando el sustrato GlcNAcMan₅GlcNAc₂ en la cepa YSH-57 genomanipulada de P. pastoris

Para analizar cualquier actividad potencial de GnTIII *ex vivo* en la cepa YSH-57 de *P. pastoris*, se analizaron los sobrenadantes del cultivo celular para determinar la actividad de GnTIII. Las células de *P. pastoris* YSH-57 se cul-

tivaron a 24°C en BMGY hasta una DO₆₀₀ de aproximadamente 10. Las células se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. Después de 24 horas de inducción, las células se eliminaron por centrifugación proporcionando un sobrenadante esencialmente transparente. Se extrajo una alícuota del sobrenadante para realizar ensayos de GnTIII y el resto se usó para recuperar la K3 segregada soluble. K3 se purificó a partir del medio mediante cromatografía por afinidad de Ni utilizando un formato de 96 pocillos en un robot de laboratorio Beckman BioMek 2000. La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind (Ejemplo 3). Los *N*-glucanos se liberaron por digestión con PNGasas (Ejemplo 3). La alícuota de sobrenadante extraída anteriormente se analizó además para determinar la presencia de actividad de GnTIII segregadas. El *N*-glucano GlcNAc₂Man₃GlcNAc₂ purificado a partir de K3 expresada en la cepa PBP-3 se añadió a: BMMY (A); UDP-GlcNAc 1 mM (Sigma Chemical Co., St. Louis, MO) en BMMY (B); el sobrenadante de YSH-44 transformado con pVA53 [YSH-57] (C); y el sobrenadante de YSH-57 + UDP-GlcNAc 1 mM (D). Después de incubar durante 24 horas a temperatura ambiente, se analizaron las muestras mediante HPLC sobre aminosílice para determinar el grado de actividad de GnTII.

Ejemplo 23

Actividad in vitro del ensayo GnTIII usando el sustrato GlcNAc₂Man₃GlcNAc₂ en la cepa YSH-57 genomanipulada de P. pastoris

Para analizar cualquier actividad potencial de GnTIII *ex vivo* en la cepa YSH-57 de *P. pastoris*, se analizaron los sobrenadantes del cultivo celular para determinar la actividad de GnTIII. Las células de *P. pastoris* YSH-57 se cultivaron a 24°C en BMGY hasta una DO₆₀₀ de aproximadamente 10. Las células se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. Después de 24 horas de inducción, las células se eliminaron por centrifugación proporcionando un sobrenadante esencialmente transparente. Se extrajo una alícuota del sobrenadante para realizar ensayos de GnTIII y el resto se usó para recuperar la K3 segregada soluble. K3 se purificó a partir del medio mediante cromatografía por afinidad de Ni utilizando un formato de 96 pocillos en un robot de laboratorio Beckman BioMek 2000. La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind (Ejemplo 3). Los *N*-glucanos se liberaron por digestión con PNGasas (Ejemplo 3). La alícuota de sobrenadante extraída anteriormente se analizó además para determinar la presencia de actividad de GnTIII segregadas. El *N*-glucano GlcNAc₂Man₃GlcNAc₂ purificado a partir de K3 expresada en la cepa YSH-44 se añadió a: BMMY (A); UDP-GlcNAc 1 mM (Sigma Chemical Co., St. Louis, MO) En BMMY (B); el sobrenadante de YSH-44 transformado con pVA53 [YSH-57] (C). Después de incubar durante 24 horas a temperatura ambiente, se analizaron las muestras mediante HPLC sobre aminosílice para determinar el grado de actividad de GnTIII.

Ejemplo 24

Clonación y expresión de GnTIV en P. pastoris

El fragmento de ADN que codifica parte de la proteína isoenzima A GnTIV humana (MGAT4A) que carece del dominio TM se amplificó por PCR a partir de ADNc humano usando el cebador directo HGIV-2 (5'-CTGATTGCTTATCAACGAGAATTCCTTG-3') y el cebador inverso HGIV-3 (5'-TGTTGGTTTCTCAGATGATCAGTTGGTG-3'), se clonó en el vector pCR2.1-TOPO (Invitrogen), y se secuenció. Para la clonación en un vector adecuado para la fusión con una adenoteca de secuencias líder, los fragmentos de ADN se amplificaron por PCR con cebadores que incluían los sitios de restricción *AscI* y *PacI*. Cebadores directos *AscI*: HGIV-ASC1 (5-TGGGCGCGCCAACTGATTGCTTATCAACGAGAA-3'), HGIV-ASC2 (5'-AGTGGGCGCGCCTTGAATAAGTTTTCAGATAATACC-3'), HGIV-ASC3 (5'-AAGGGCGCGCCCAAGTGCCAAGTATTTATTATC-3'). Cebador inverso *PacI* HGIV-PAC (5'-GTTTAATTAA GATCAGTTGGTGGCTTTTAAATATG-3'). El ácido nucleico y las secuencias de aminoácidos de GnTIVA humano se muestran en la Figura 41.

De forma similar, el fragmento de ADN que codifica parte de la proteína isoenzima B GnTIV humana (MGAT4B) que carece del dominio TM se amplificó por PCR a partir de ADNc humano usando el cebador directo HGIVB-2 (5'-AGCGGCCAGAAAGGCGACGTTGTGGAC-3') e inverso HGIVB-3 (5'-TACCCTCAGAAGCCCGCAGCTTAGTC-3'), se clonó en el vector pCR2.1- TOPO (Invitrogen), y se secuenció. Para la clonación en un vector adecuado para la fusión con una adenoteca de secuencias líder, los fragmentos de ADN se amplificaron por PCR con cebadores que incluían los sitios de restricción *AscI* y *PacI*. Cebadores directos *AscI*: HGIVB-A1 (5'-AGCGGGCGCGCCGGC GACGTTGTGGACGTTTAC-3'), HGIVB-A2 (5'-CCGTGGCGCGCCTCACACCGGCACGTGCTGCAC-3'). Cebador inverso *PacI* HGIVB-P (5'-TGTTAATTAAGCTTAGTCGGCCTTTTTCAGGAAG-3'). El ácido nucleico y las secuencias de aminoácidos de GnTIVB humano se muestran en la Figura 42.

ES 2 336 790 T3

TABLA 12

Plásmidos que contienen construcciones fusionadas de GnTV o GnTIV para la expresión de estructuras multiantenarias

Dominio catalítico/líder	Vector <i>HYG</i>	Vector <i>KAN</i>
GnTV de ratón		
$\Delta 95-53$		pPB146
$\Delta 145-53$	pPB125	pPB140
$\Delta 45-54$	pPB130	
$\Delta 209-53$	pPB126	
$\Delta 209-54$	pPB 131	
GnT IVA humana		
$\Delta 32-53$	pPB127	
$\Delta 82-53$	pPB128	pPB141
$\Delta 103-53$	pPB129	
$\Delta 32-54$	pPB132	
$\Delta 82-54$	pPB133	
$\Delta 103-54$	pPB134	
GnT IVB humana		
$\Delta 32-S3$	pPB143	
$\Delta 104-53$	pPB144	
GnTIX humana		
$\Delta 43-53$		pPB176

Ejemplo 25

*Cepa de *P. pastoris* que produce estructuras de glucanos triantenarios*

La cepa YSH-44 de *P. pastoris* (Ejemplo 15) que produce estructuras de glucanos complejos se transformó con el plásmido pPB144 (Tabla 12) que contenía un fragmento de un gen que codifica el dominio catalítico de GnTIVB humano ($\Delta 104$) ligado en marco a un péptido director de *MNN2(s)* de *S. cerevisiae* [nucleótidos 1-108]. El plásmido pPB144 también contiene un marcador de resistencia *HYG* y el gen *URA3* de *P. pastoris* como secuencia de direccionamiento para la integración genómica. Un μg de plásmido se linealizó con *San*, transformado en la cepa YSH-44 mediante electroporación, se seleccionaron sobre medio que contenía higromicina y la cepas resultantes se cribaron mediante análisis de los glucanos liberados a partir de K3 purificada. Los transformantes se cultivaron a 30°C en BMGY a una OD_{600} de aproximadamente 100, se recolectaron por centrifugación y se transfirieron a BMMY para inducir la producción de K3 (kringle 3 de plasminógeno humano) controlada por un promotor AOX1. K3 se purificó a partir del medio mediante cromatografía por afinidad de Ni utilizando un formato de 96 pocillos en un robot de laboratorio Beckman BioMek 2000. La purificación robótica es una adaptación del protocolo que proporciona Novagen para su resina HisBind (Ejemplo 3). Los *N*-glucanos se liberaron por digestión con PNGasas (Ejemplo 3). Los *N*-glucanos se analizaron por EM MALDI-TOF (Ejemplo 3). Los transformantes que mostraban la transferencia de restos GlcNAc sobre la ramificación $\text{Man}\alpha 1,3$ de la estructura de oligosacárido (por ejemplo $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$) se denominaron PBP43 (Figura 47). El análisis de los *N*-glucanos proporciona un pico predominante en 1543 m/z [y] es coherente con la masa del glucano $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$.

Ejemplo 26

Clonación y expresión de GnTV en P. pastoris

El fragmento de ADN que codifica parte de la proteína GnTV de ratón (MGAT45) que carece del dominio TM se amplificó por PCR a partir de ADNc murino usando los cebadores directo MGTV-2 (5'-AAATCAAGTGGATGAAG GACATGTGGC3') e inverso MGTV-3 (5'-AGCGATGCTATAGGCAGTCTTTGCAGAG-3'), se clonó en el vector pCR2.1-TOPO (Invitrogen), y se secuenció. Para la clonación en el vector adecuado para la fusión con una adenoteca de secuencias líder, los fragmentos de ADN se amplificaron por PCR con cebadores que incluían los sitios de restricción *AscI* y *PacI*. Cebadores directos *AscI*: MGTV-ASC1 (5'-TATGGGCGCGCCGATCATAACTCATTGGCGGAAATC-3'), MGTV-ASC2 (5'-GAAGGGCGCGCCTTGCCTCCTATGGATGGCTACCCCCAC3'), MGTV-ASC3 (5'-TGGGG CGCGCCGGAAGCTCGAGTCAAAGGTGGACAAT-3'). Cebador inverso *PacI* MGTV-PAC (5'-AGTTAATTAAT GCTATAGGCAGTCTTTGCAGAG-3'). El ácido nucleico y las secuencias de aminoácidos de GnTV de ratón se muestran en la Figura 43.

Ejemplo 27

Cepa de P. pastoris que expresa GnTV

La cepa YSH-44 de *P. pastoris* (Ejemplo 15) que produce estructuras de glucanos complejos se transformó con el plásmido pPB140 (Tabla 12) que contenía un fragmento de un gen que codifica el dominio catalítico de GnTV de ratón ($\Delta 45$) ligado en marco a un péptido director derivado de *MNN2(s)* de *S. cerevisiae*. Las condiciones de cultivo eran idénticas a las del Ejemplo 25. Se analizó la proteína testigo K3 de dos transformantes usando MALDI-TOF. Un pico en 1559 m/z [y] es coherente con la masa del glucano GlcNAc₃Man₃GlcNAc₂ (Figura 48). Un pico secundario en 1355 m/z [u] es coherente con la masa de GlcNAc₂Man₅GlcNAc₂.

Ejemplo 28

Cepa de P. pastoris que produce estructuras tetraantenarias sobre las glucoproteínas

La cepa PBP43 de *P. pastoris* (Ejemplo 25) que produce estructuras de glucanos triantenarios (por ejemplo, GlcNAcMan₃GlcNAc₂) se transformó con el plásmido pPB140 (Figura 40B) que codifica GnTV de ratón (Ejemplo 27). El vector pPB140 contiene el marcador de resistencia KAN y el gen *HIS3* de *P. pastoris* como secuencia de direccionamiento para la integración genómica. Un μ g de plásmido se linearizó con *KpnI*, transformado en la cepa PBP43 mediante electroporación, se seleccionaron sobre medio que contenía kanomicina y la cepas resultantes se cribaron mediante análisis de los glucanos liberados a partir de K3 purificada. Las condiciones de cultivo eran idénticas a las del Ejemplo 25. El análisis de la proteína testigo K3 mediante MALDI-TOF mostró un pico predominante en 1747 m/z [z], que es coherente con la masa del glucano tetraantenario GlcNAc₄Man₅GlcNAc₂ (Figura 49). La digestión con hexosaminidasa (Véase el Ejemplo 15) de los glucanos resultantes mostró una masa del pico correspondiente a Man₃GlcNAc₂ (no se muestran los datos).

En un segundo experimento, YSH-44 de *P. pastoris* se transformó con pPB128 y pPB140 (Tabla 12). El análisis de los transformantes que producen la proteína testigo K3 mediante MALDI-TOF mostró un pico predominante en 1743 m/z [z], que es coherente con la masa del glucano tetraantenario GlcNAc₄Man₃GlcNAc₂ (Figura 50).

Ejemplo 29

Clonación y expresión de GnTIX en P. pastoris

El ácido nucleico y las secuencias de aminoácidos de GnTIX humano (AB109185.1) se muestran en la Figura 45. El fragmento de ADN con codones optimizados que codifican parte del GnTIX humano que carece del dominio TM ($\Delta 43$) se sintetizó a partir de oligonucleótidos usando PCR (Figura 46). El fragmento de ADN que codifica el dominio catalítico de GnTIX se ligó en marco a un péptido director derivado de *MNN2(s)* de *S. cerevisiae*. El plásmido resultante pPB176 (Figuras 40C) se linearizó con *KpnI* y se transformó en la cepa YSH-44 de *P. pastoris* (Ejemplo 15) produciendo estructuras complejas de glucanos tetraantenarios. Las condiciones de cultivo eran idénticas a las del Ejemplo 25. Se analizó la proteína testigo K3 de un transformante usando EM MALDI-TOF.

REIVINDICACIONES

1. Una célula huésped de hongo unicelular o multicelular que es capaz de producir glucoproteínas que comprenden una estructura nuclear de *N*-glucano triantenario, en la que las células huésped son genomanipuladas para que produzcan glucoproteínas que tienen un *N*-glucano $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ y en la que las células huésped además incluyen un ácido nucleico que codifica el dominio catalítico de la *N*-acetilglucosaminiltransferasa IV exógena GnT IVB ($\Delta 104-53$) humana fusionado con los nucleótidos 1-108 del gen *MNN2* de *S. cerevisiae* que actúa como péptido de señalización de direccionamiento celular no asociado normalmente con el dominio catalítico, que dirige dicho dominio catalítico al RE o al aparato de Golgi de la célula huésped.
2. La célula huésped de la reivindicación 1, en la que la estructura triantenaria producida en dicha célula huésped comprende la ramificación del de oligosacárido $\text{GlcNAc } \beta 1,2 - \text{Man} \alpha 1,6 (\text{GlcNAc } \beta 1,4 \text{ GlcNAc } \beta 1,2 \text{ Man} \alpha 1,3) \text{ Man } \beta 1,4 - \text{GlcNAc } \beta 1,4 - \text{GlcNAc } \beta 1,4 - \text{Asn}$.
3. La célula huésped de hongo unicelular o multicelular de la reivindicación 1 ó 2, en la que dicha célula huésped además incluye un ácido nucleico que codifica un dominio catalítico de *N*-acetilglucosaminiltransferasa exógena para producir una glucoproteína que tiene una estructura nuclear de *N*-glucano tetraantenario fusionado a un péptido de señalización de direccionamiento celular no asociado normalmente al dominio catalítico, que dirige el dominio catalítico al RE o al aparato de Golgi de la célula huésped.
4. La célula huésped de hongo unicelular o multicelular de cualquiera de las reivindicaciones 1 a 3, en la que la célula huésped incluye además un ácido nucleico que codifica un dominio catalítico de *N*-acetilglucosaminiltransferasa IX exógeno fusionado a un péptido de señalización de direccionamiento celular no asociado normalmente al dominio catalítico, que dirige el dominio catalítico al RE o al aparato de Golgi de la célula huésped para producir la glucoproteína que tiene la estructura nuclear del *N*-glucano tetraantenario.
5. La célula huésped de la reivindicación 3 ó 4, en la que dichas estructuras tetraantennarias producidas por dicha célula huésped comprenden estructuras de $\text{GlcNAc}_4\text{Nan}_3\text{GlcNAc}_2$ que pueden sufrir reacciones adicionales por GnT VI.
6. La célula huésped de cualquiera de las reivindicaciones 1 a 5, en la que dicha célula huésped además comprende un ácido nucleico que codifica una enzima que tiene actividad de *N*-acetilglucosaminiltransferasa III.
7. La célula huésped de cualquiera de las reivindicaciones 1 a 6, en la que dicha célula huésped es deficiente en una actividad de manosiltransferasa *OCH1* y/o en una actividad de manosiltransferasa Dol-P-Man:Man₅GlcNAc₂-PP-Dol.
8. La célula huésped de cualquiera de reivindicaciones 1 a 7, en la que dicha célula huésped está genomanipulada además para que exprese niveles más elevados de actividad del transportador de UDP-GlcNAc para aumentar los niveles de UDP-GlcNAc en la célula.
9. La célula huésped de cualquiera de las reivindicaciones 1 a 8, en la que la célula huésped se selecciona a partir del grupo constituido por *Pichia pastoris*, *Pichia finlandica*, *Pichia trehalophila*, *Pichia koclamae*, *Pichia membranaefaciens*, *Pichia minutia*, *Ogataea minuta*, *Pichia lindneri*, *Pichia opuntiae*, *Pichia thermotolerans*, *Pichia salictaria*, *Pichia guercuum*, *Pichia pijperi*, *Pichia stiptis*, *Pichia methanolica*, *Pichia sp.*, *Saccharomyces cerevisiae*, *Saccharomyces sp.*, *Hansenula polymorpha*, *Kluyveromyces sp.*, *Kluyveromyces laetis*, *Candida albicans*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei*, *Chrysosporium lucknowense*, *Fusarium sp.*, *Fusarium gramineum*, *Fusarium venenatum*, y *Neurospora crassa*.
10. La célula huésped de la reivindicación 9, en la que la célula huésped es *Pichia pastoris*.
11. La célula huésped de cualquiera de las reivindicaciones 1 a 10, en la que la glucoproteína es una proteína terapéutica.
12. La célula huésped de la reivindicación 11, en la que la proteína terapéutica se selecciona a partir del grupo constituido por dominios kringle de plasminógeno humano, eritropoyetina, citocinas, factores de coagulación, cadena a del receptor de IgE soluble, IgG, fragmentos de IgG, IgM, interleucinas, urocinasa, quimasa, inhibidor de urea tripsina, proteína que se une a IGF, factor de crecimiento epidérmico, factor liberador de hormona del crecimiento, proteína de fusión anexina V, angiostatina, factor de crecimiento endotelial vascular-2, factor inhibidor progenitor mielóide-1, osteoprotegerina, antitripsina α -1, ADNasa II, α -fetoproteínas, AAT, rhTBP-1 (proteína de unión a TNF 1), TACI-Ig (activador transmembrana y modulador del calcio e interactor con los ligandos de ciclofilina), FSH (hormona estimulante de folículos), GM-CSF, GLP-1 con o sin FC (proteína glucagonoide 1), agonista del receptor IL-1, sTNFR (fusión de TNF soluble y receptor Fc), ATIII, rhTrombina, glucocerebrosidasa y CTLA4-Ig (antígeno asociado a linfocitos T citotóxicos 4-Ig).
13. Un procedimiento para producir una glucoproteína que comprende expresar un ácido nucleico que codifica la glucoproteína en una célula huésped de una cualquiera de las reivindicaciones 1 a 12.

ES 2 336 790 T3

14. El procedimiento de la reivindicación 13, en el que dicha glucoproteína comprende *N*-glucanos que tienen estructuras triantenarias.

15. El procedimiento de la reivindicación 13, en el que dicha glucoproteína comprende *N*-glucanos que tienen estructuras tetraantenarias.

16. El procedimiento de cualquiera de reivindicaciones 13 a 15, en el que dicha glucoproteína además comprende un resto GlcNAc diseccionado.

17. El procedimiento de cualquiera de reivindicaciones 13 a 16, en el que dicha glucoproteína carece de fucosa.

18. El procedimiento de una cualquiera de las reivindicaciones 13 a 17, en el que dicha glucoproteína se selecciona a partir del grupo constituido por dominios kringle de plasminógeno humano, eritropoyetina, citocinas, factores de coagulación, cadena α del receptor de IgE soluble, IgG, fragmentos de IgG, IgM, interleucinas, urocinasa, quimasa, inhibidor de urea tripsina, proteína que se une a IGF, factor de crecimiento epidérmico, factor liberador de hormona del crecimiento, proteína de fusión anexina V, angiostatina, factor de crecimiento endotelial vascular-2, factor inhibidor progenitor mieloide-1, osteoprotegerina, antitripsina α -1, ADNasa II, α -fetoproteínas, AAT, rhTBP-1 (proteína de unión a TNF 1), TACI-Ig (activador transmembrana y modulador del calcio e interacto con los ligandos de ciclofilina), FSH (hormona estimulante de folículos), GM-CSF, GLP-1 con o sin FC (proteína glucagonoide 1), agonista del receptor IL-1, sTNRr (fusión de TNF soluble y receptor Fc), ATIII, rhTrombina, glucocerebrosidasa y CTLA4-Ig (antígeno asociado a linfocitos T citotóxicos 4-Ig).

19. El procedimiento de cualquiera de reivindicaciones 13 a 18, que además comprende la etapa de aislar la glucoproteína del huésped.

20. Una composición farmacéutica que comprende una composición de glucoproteína producida por una cualquiera de las células huésped de las reivindicaciones 1 a 11, en la que más del 50% molar de las estructuras nucleares de los *N*-glucanos de las glucoproteínas de dicha composición tienen estructuras triantenarias.

21. La composición farmacéutica de la reivindicación 20, en la que dichas estructuras triantenarias comprenden al menos tres GlcNAc sobre un oligosacárido $\text{Man}_3\text{GlcNAc}_2$.

22. La composición farmacéutica de la reivindicación 20 ó 21, en la que algunas o todas las estructuras triantenarias comprenden la ramificación del oligosacárido $\text{GlcNAc } \beta 1,2 - \text{Man} \alpha 1,6 (\text{GlcNAc } \beta 1,4 \text{ GlcNAc } \beta 1,2 \text{ Man} \alpha 1,3) \text{ Man } \beta 1,4 - \text{GlcNAc } \beta 1,4 \text{ GlcNAc } \beta 1,4 - \text{Asn}$.

23. La composición farmacéutica de cualquiera de las reivindicaciones 20 a 22, en la que más del 50% molar o más de 75% molar de las estructuras nucleares de los *N*-glucanos están modificadas además por GnTV.

24. La composición farmacéutica de la reivindicación 23, en la que algunas o todas las estructuras modificadas con GnTV comprenden la ramificación del oligosacárido $\text{GlcNAc } \beta 1,4 \text{ GlcNAc } \beta 1,2 - \text{Man} \alpha 1,6 (\text{GlcNAc } \beta 1,4 \text{ GlcNAc } \beta 1,2 \text{ Man} \alpha 1,3) \text{ Man } \beta 1,4 \text{ GlcNAc } \beta 1,4 - \text{GlcNAc } \beta 1,4 - \text{Asn}$.

25. Las composiciones farmacéuticas de cualquiera de las reivindicaciones 20 a 24, en las que dichas glucoproteínas de dicha composición carecen de fucosa.

26. Un vector capaz de expresar en una célula huésped de hongo unicelular o multicelular una enzima que tiene actividad de *N*-acetilglucosaminiltransferasa IV (GnT IV) que es sustancialmente intracelular, en el que dicha enzima comprende el dominio catalítico de GnT IVB humana ($\Delta 104-53$) fusionado al péptido de señalización de direccionamiento celular *MNN2* de *S. cerevisiae* (1-36) no asociado normalmente con el dominio catalítico, que dirige el dominio catalítico al RE o al aparato de Golgi de dicha célula huésped.

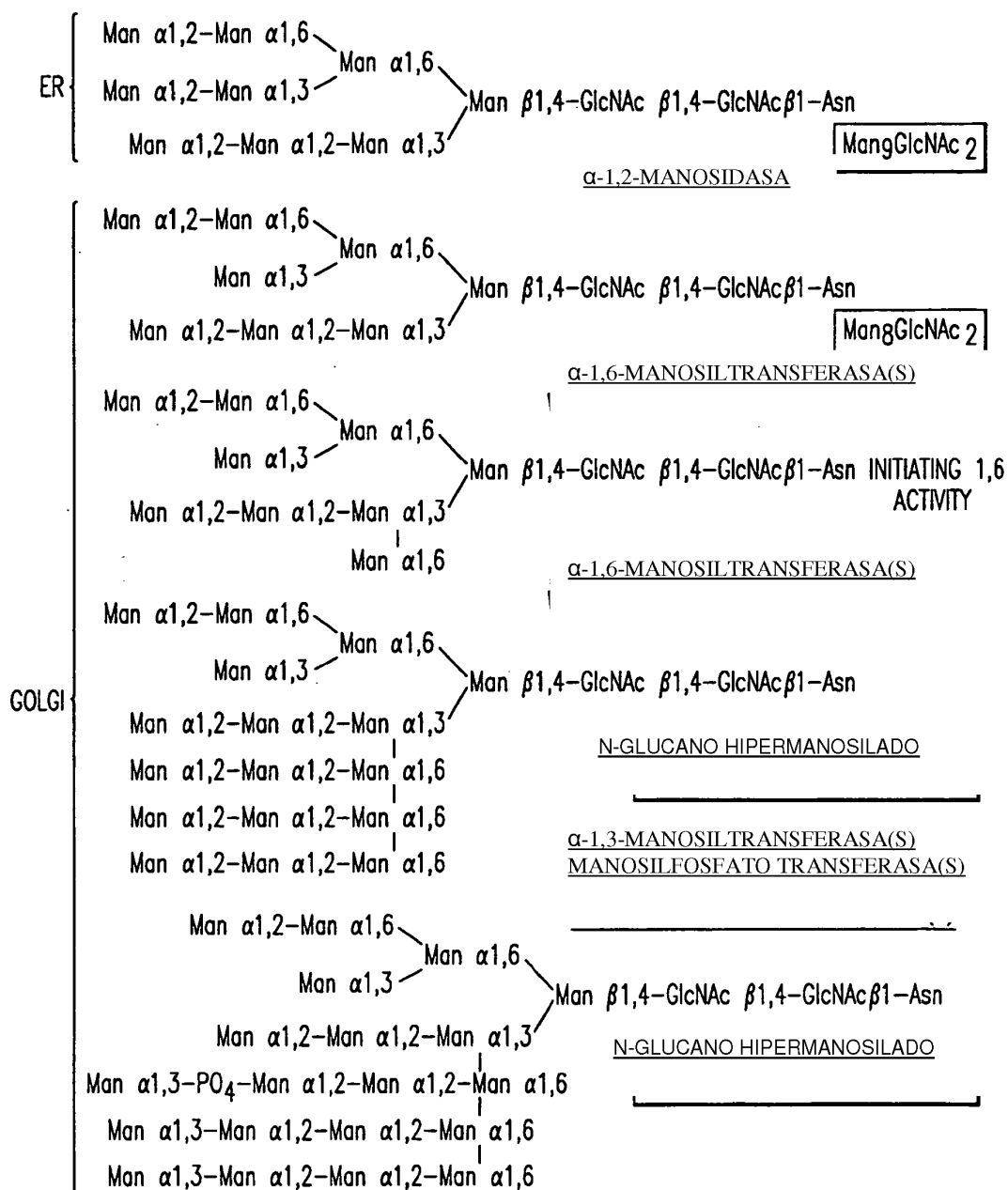


FIG. 1A

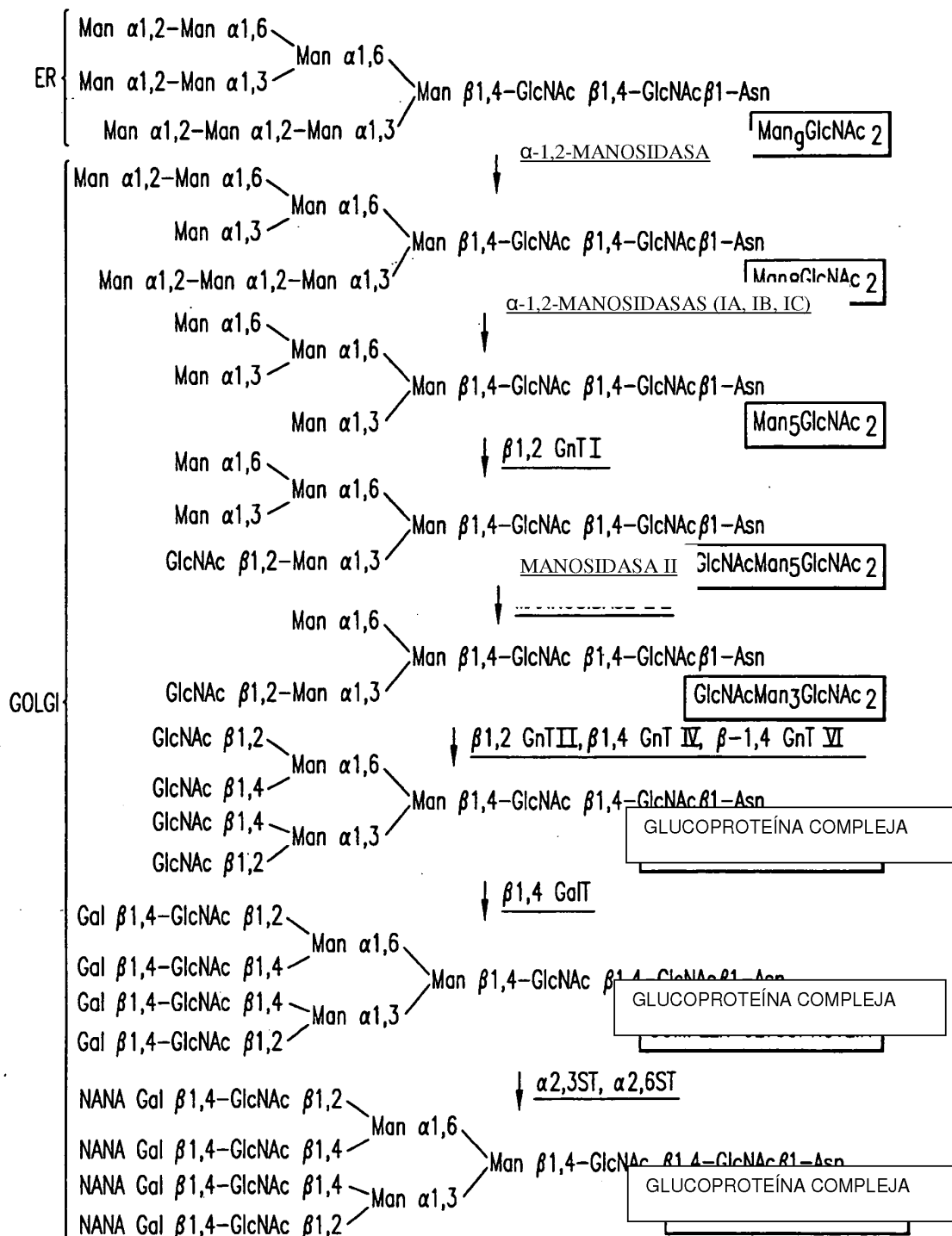


FIG.1 B

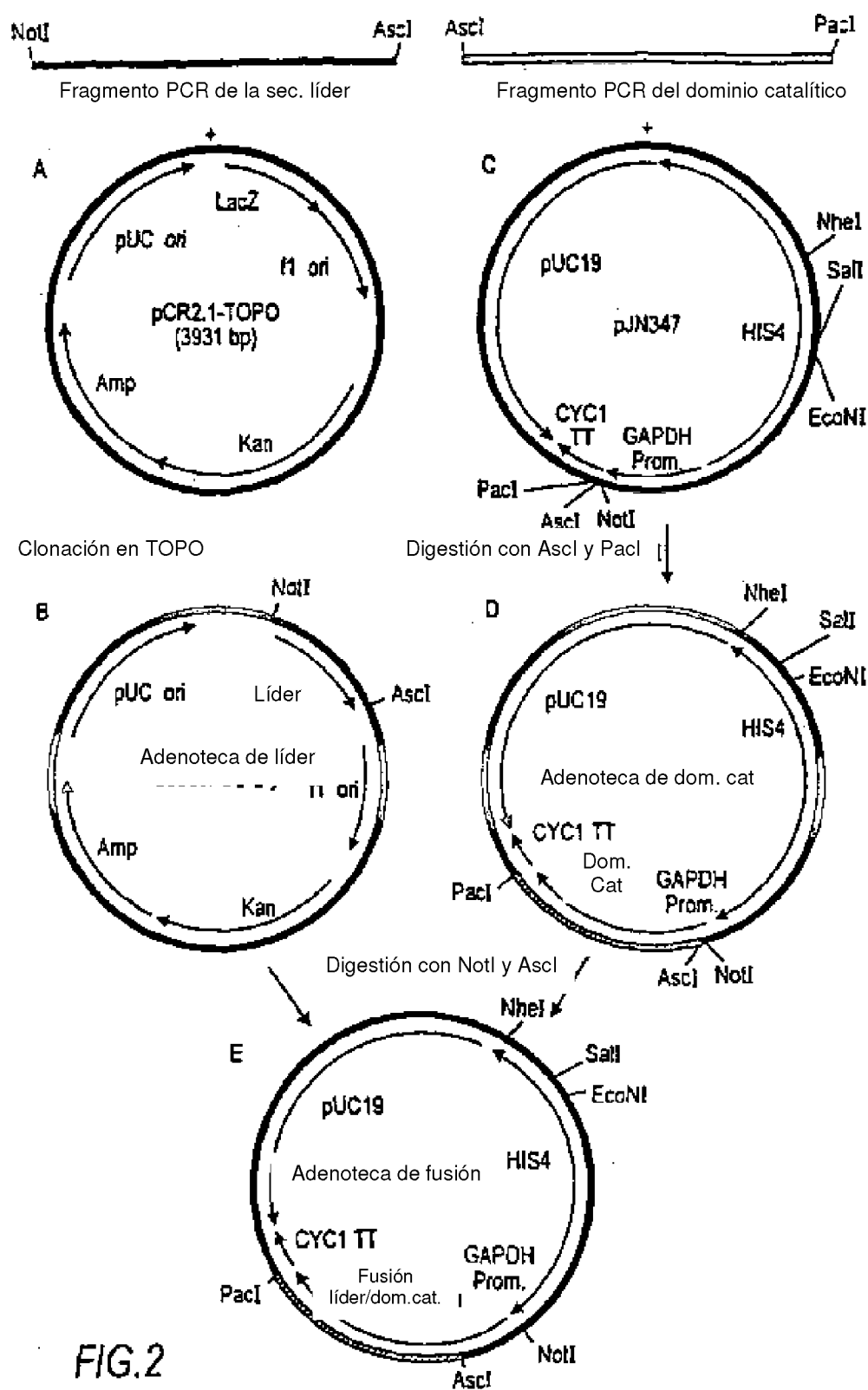


FIG.2

Marco de lectura abierto de alfa-1,2-manosidasa IA de *M. musculus*. Los dominios transmembrana y catalítico están destadados en negrita respectivamente. La secuencia de los cebadores que se usaron para generar los truncados del extremon N están destacados mediante subrayado y el inicio de cada fragmento proteínico respectivo está indicado mediante una flecha.

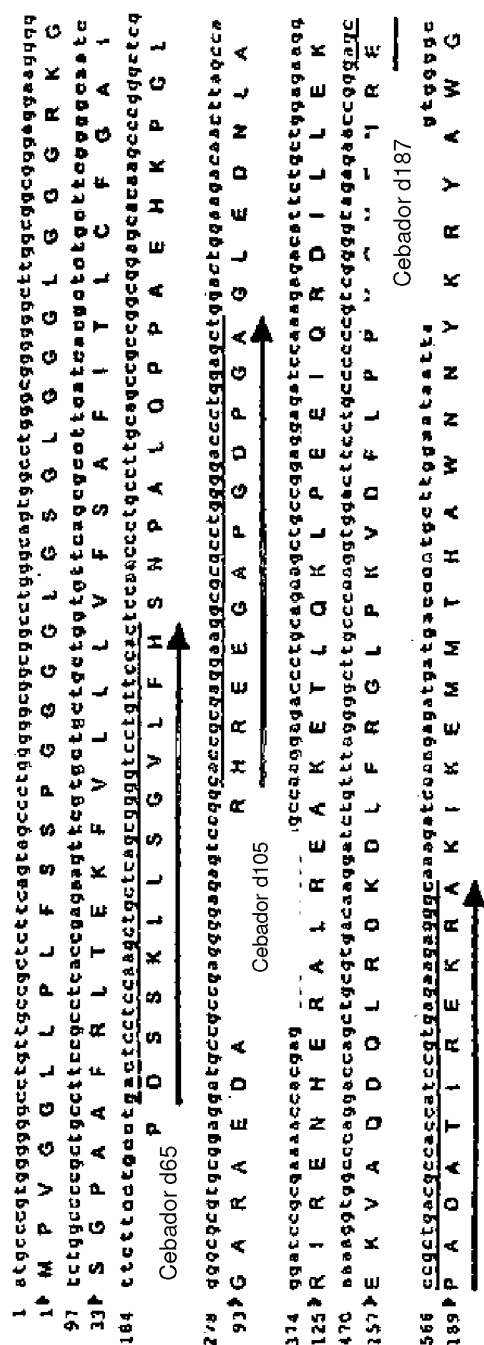


FIG. 3

[illegible]

FIG. 3 CONT

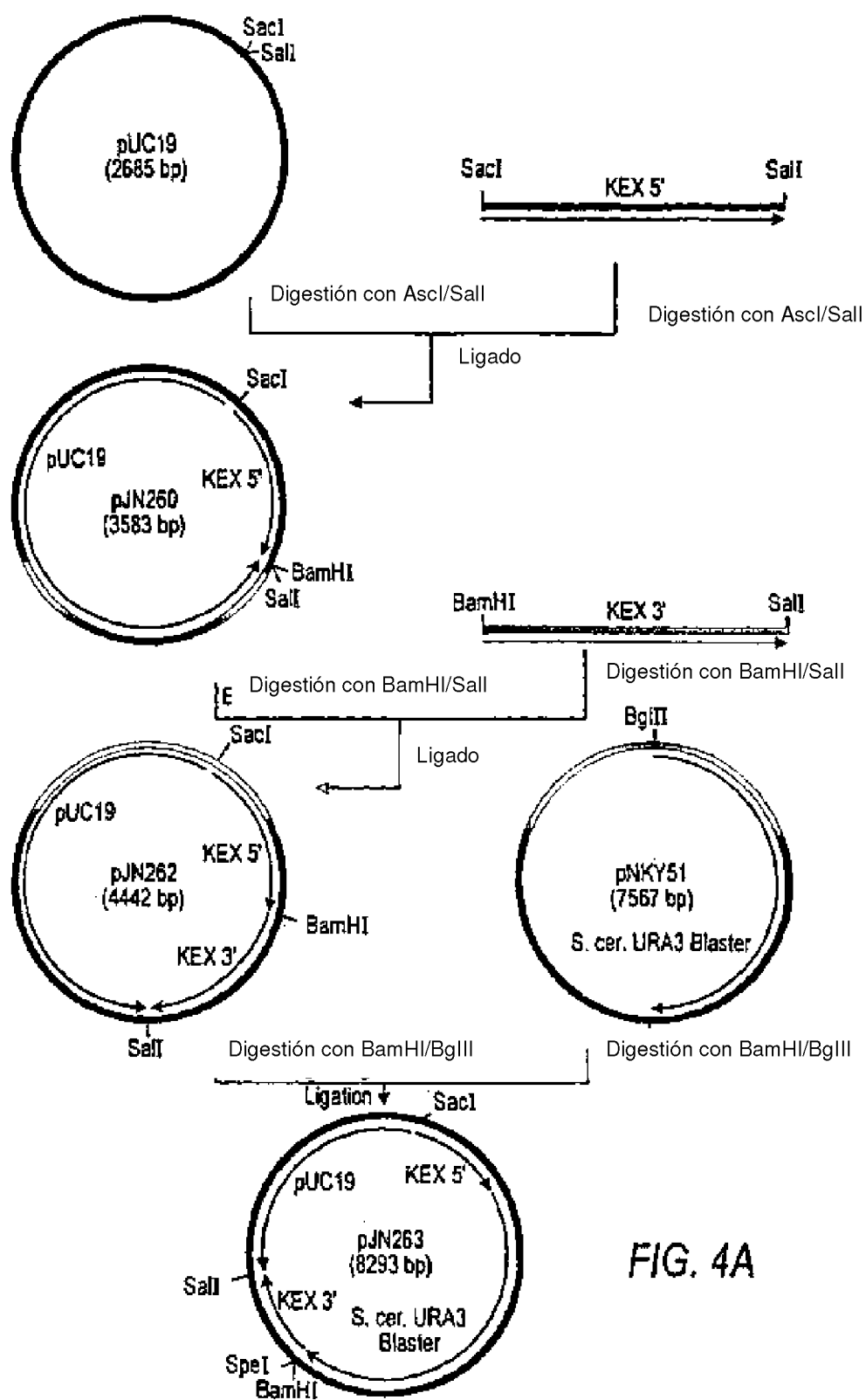


FIG. 4A

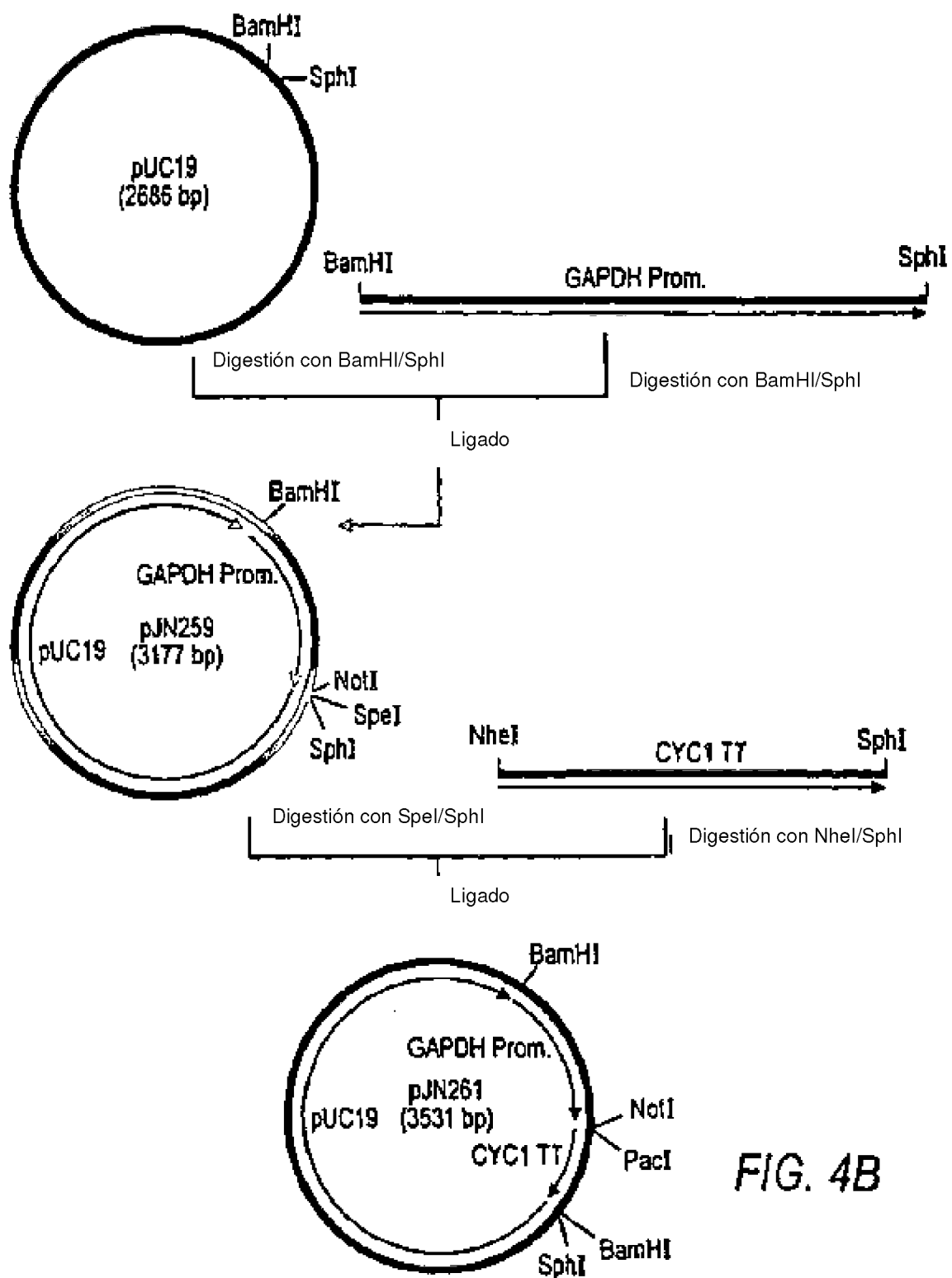


FIG. 4B

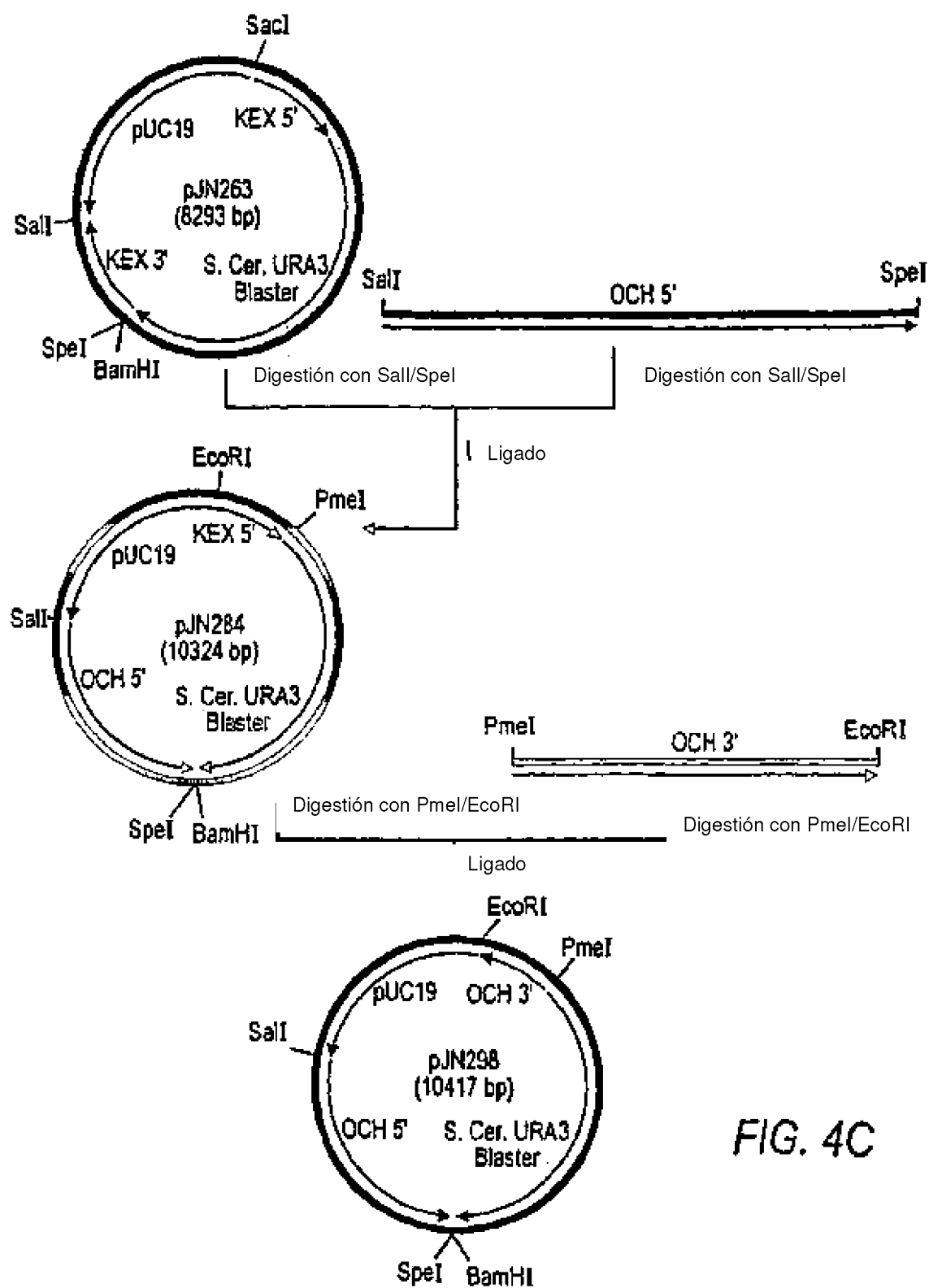


FIG. 4C

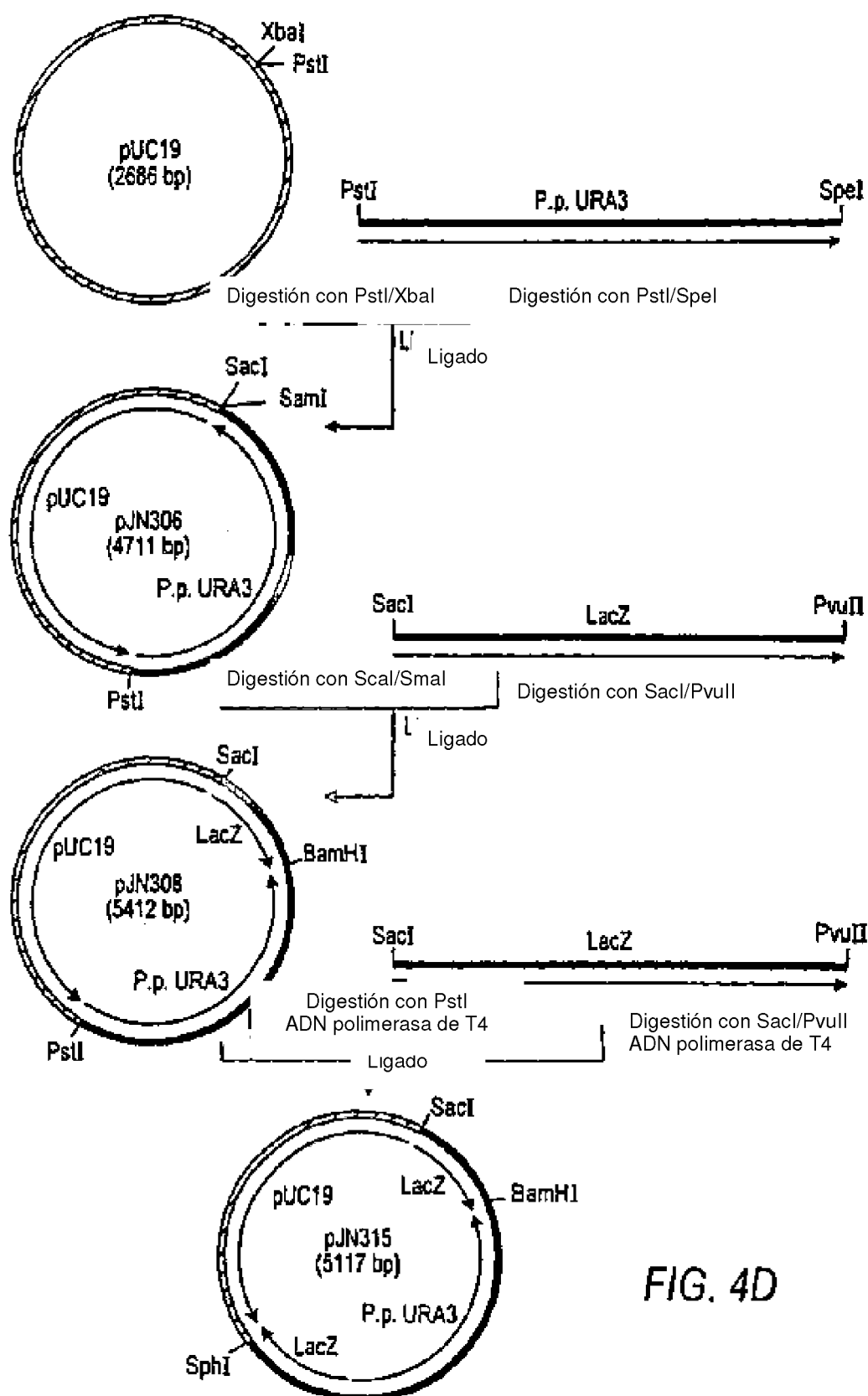


FIG. 4D

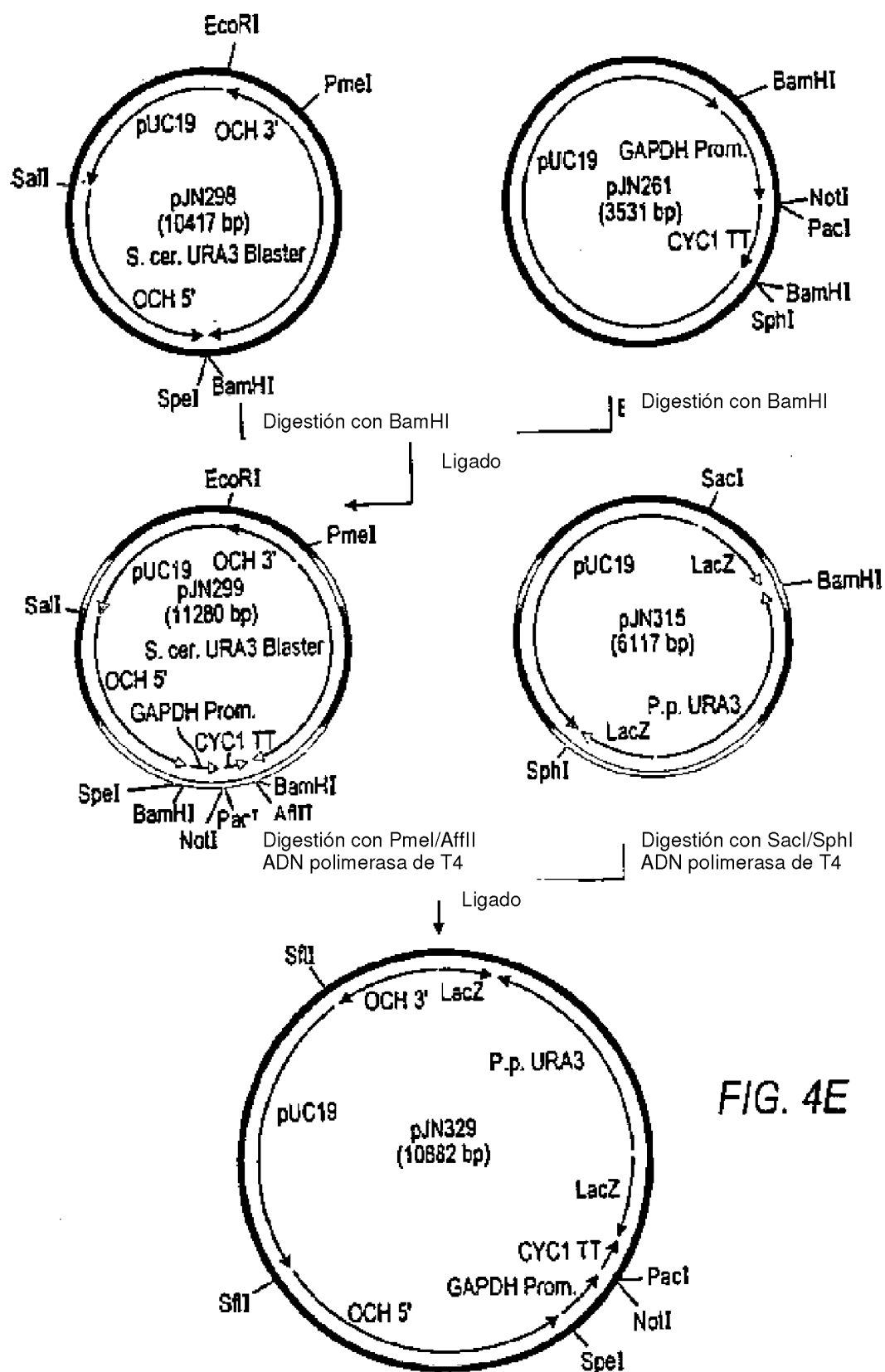
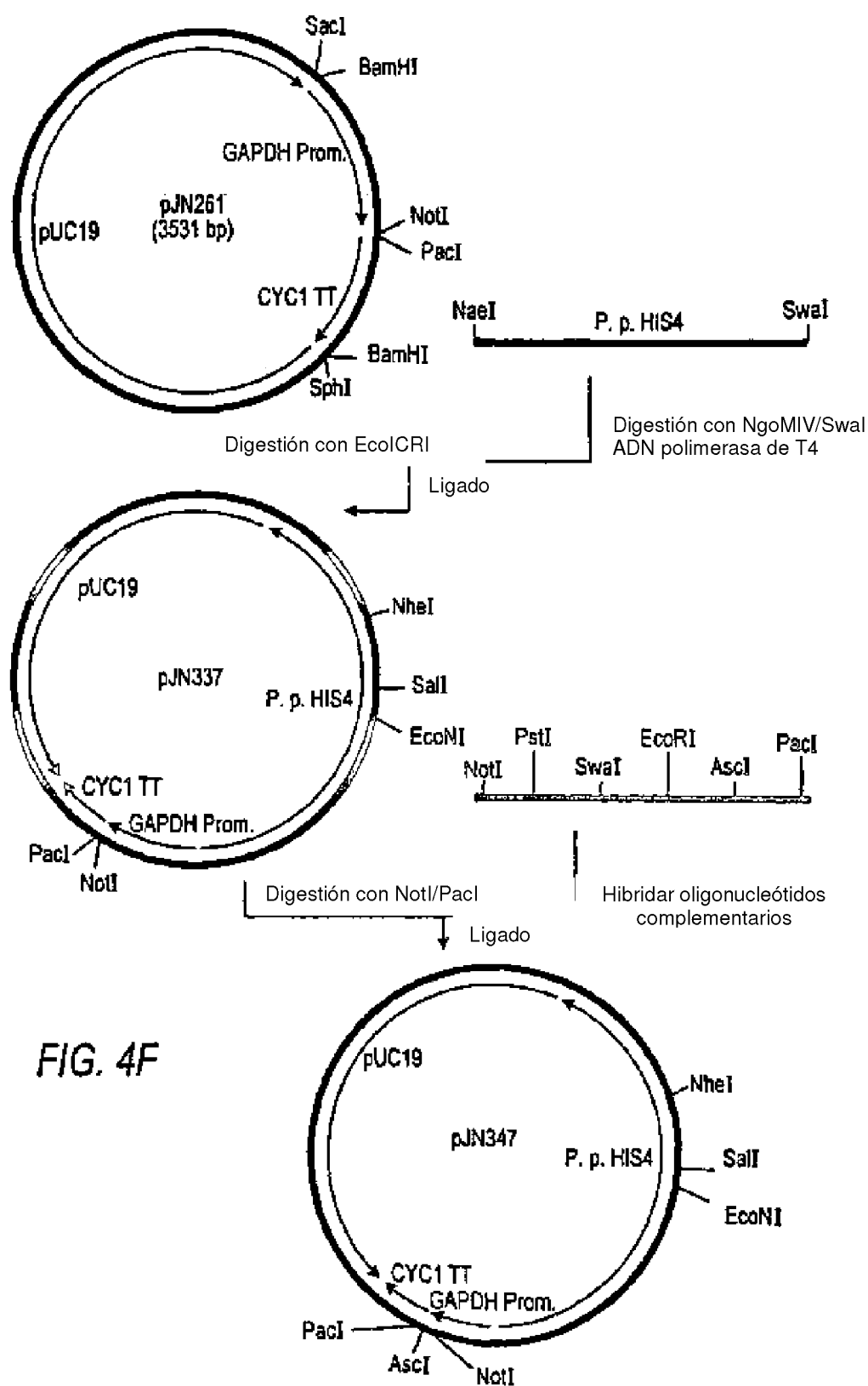


FIG. 4E



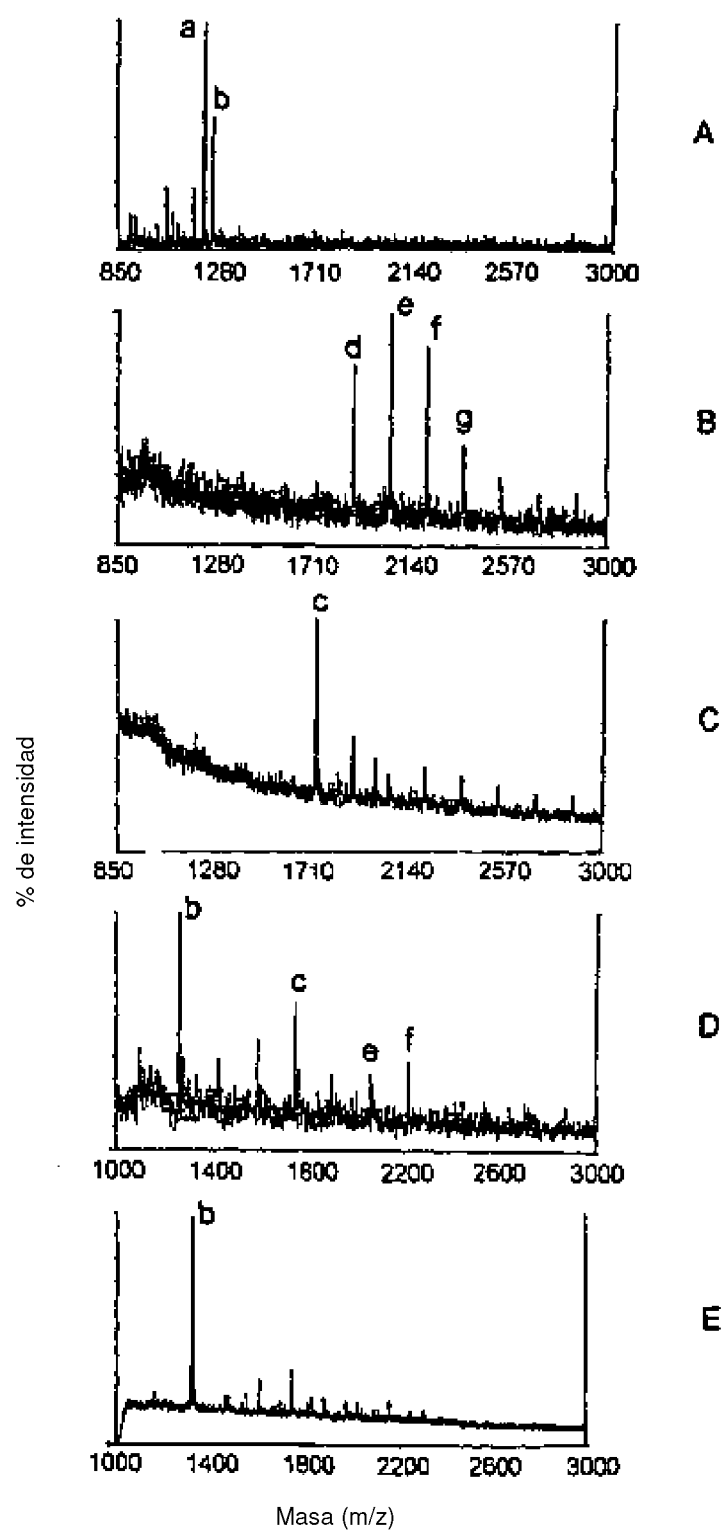
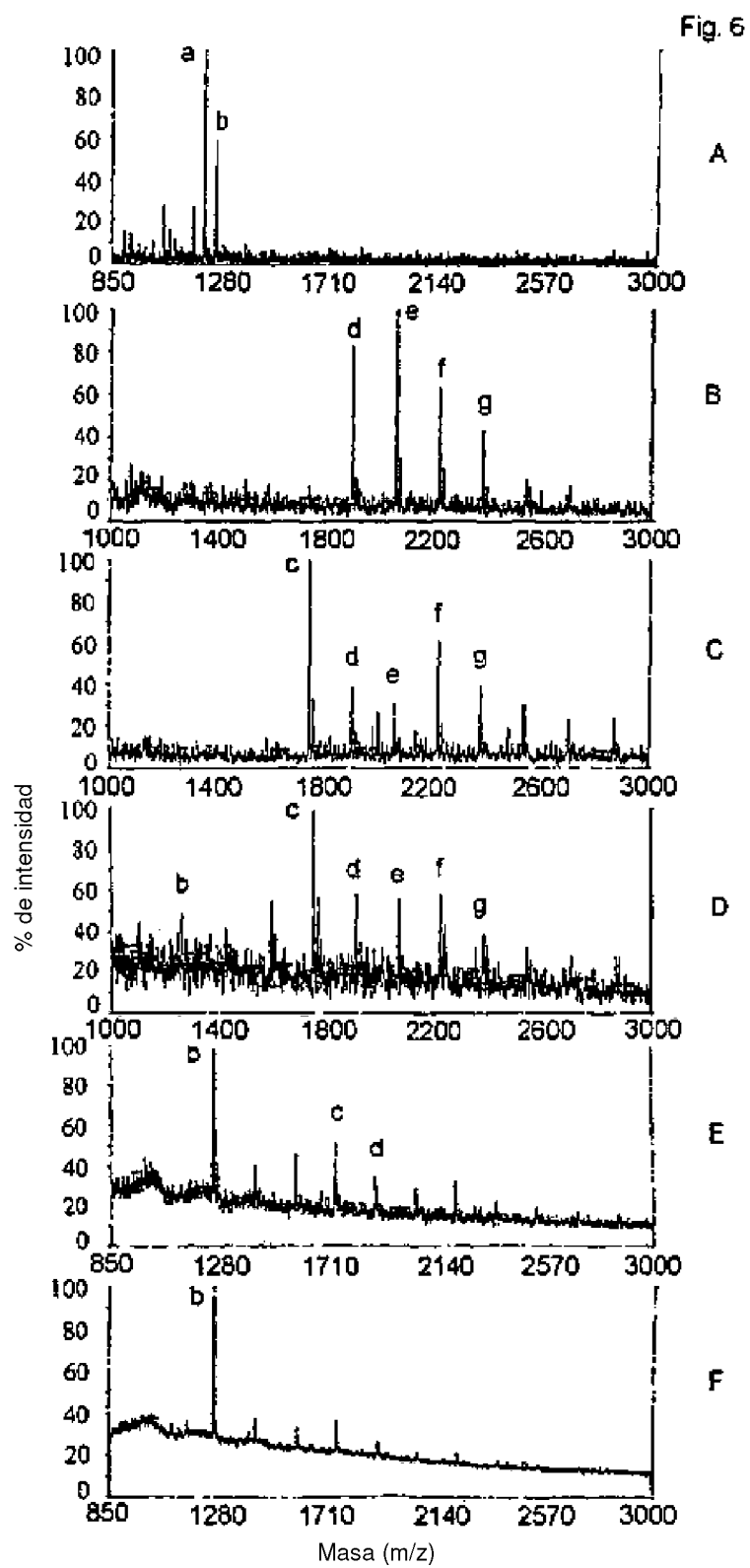
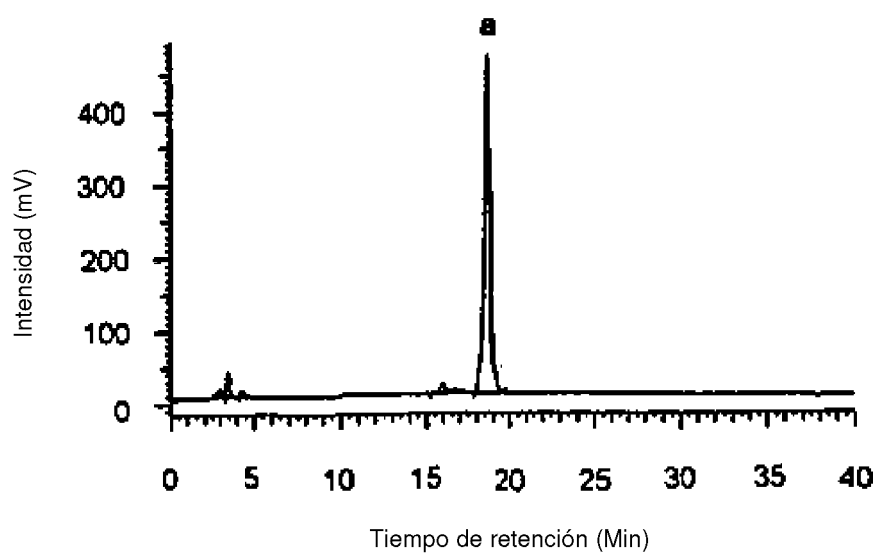
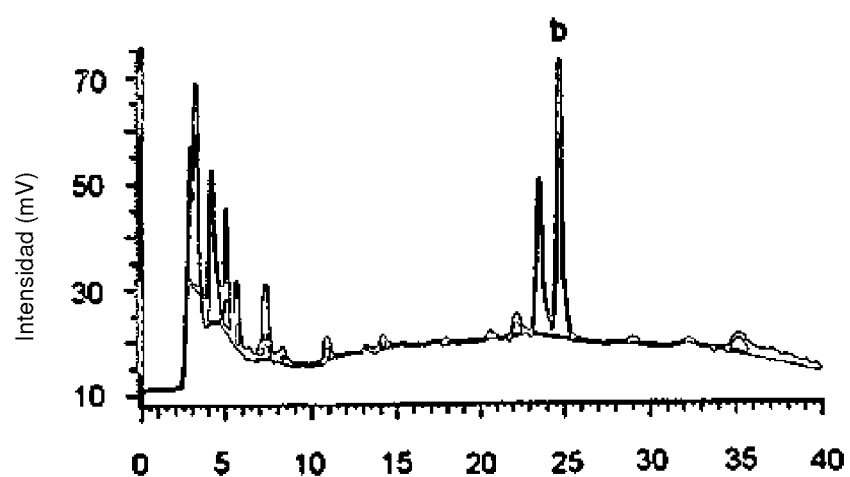
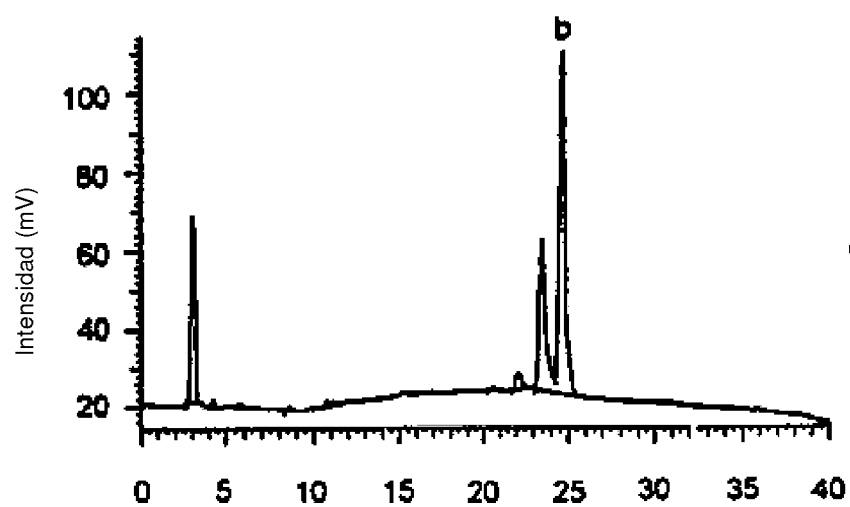
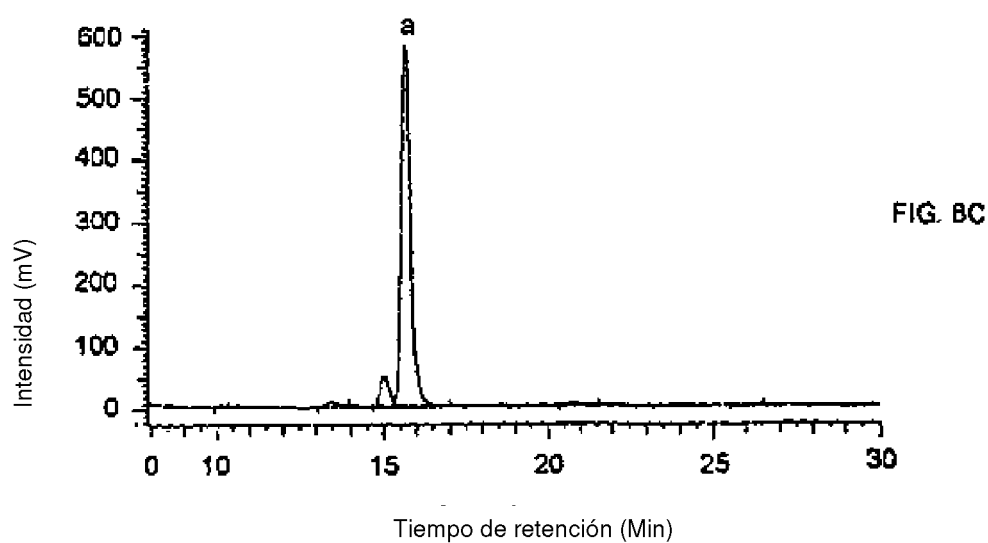
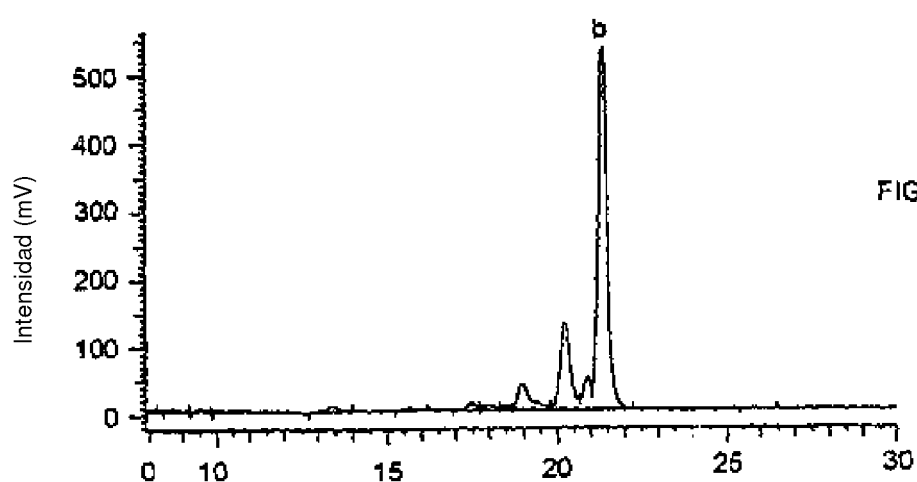
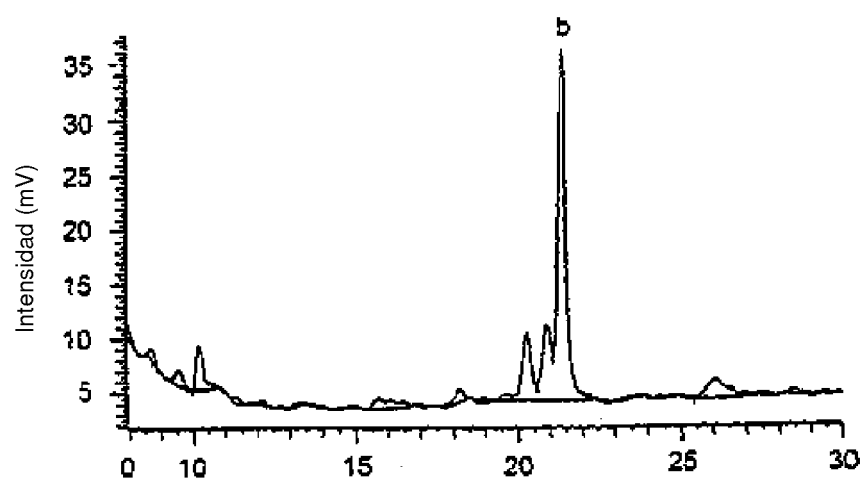
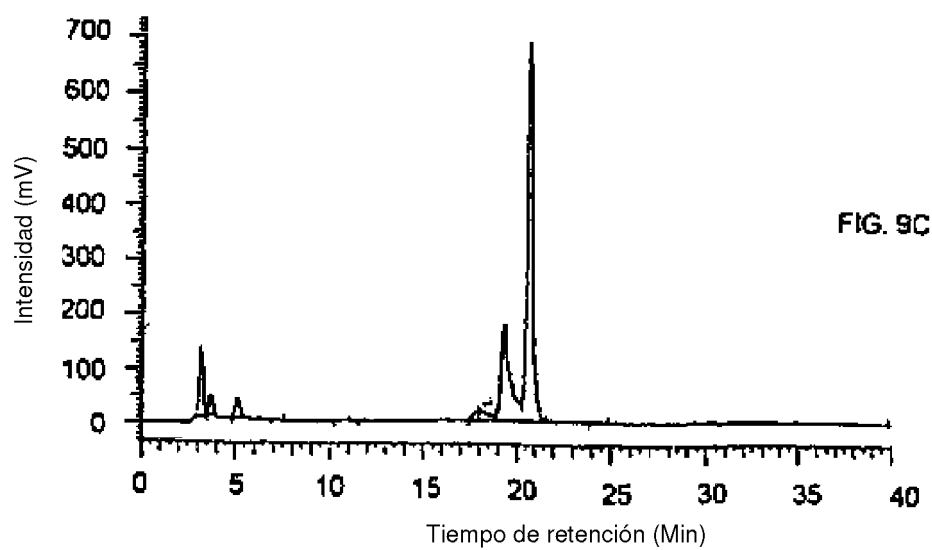
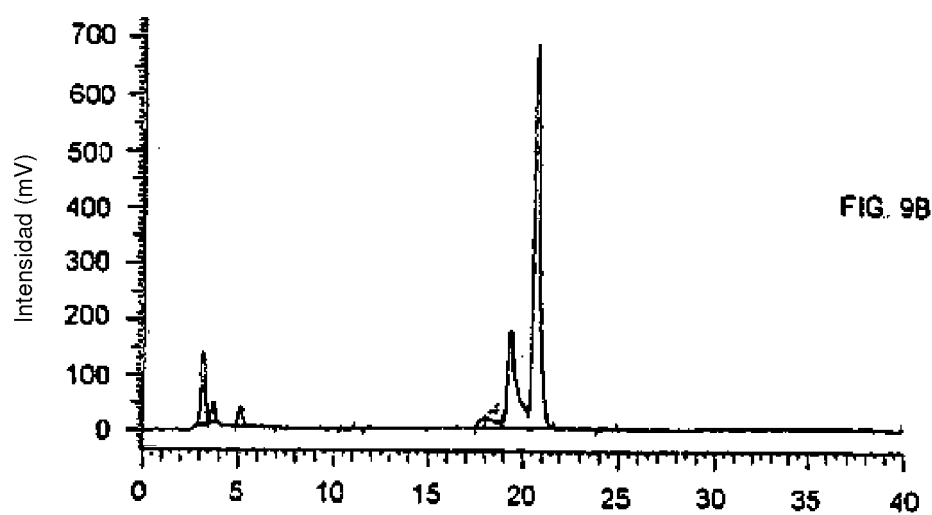
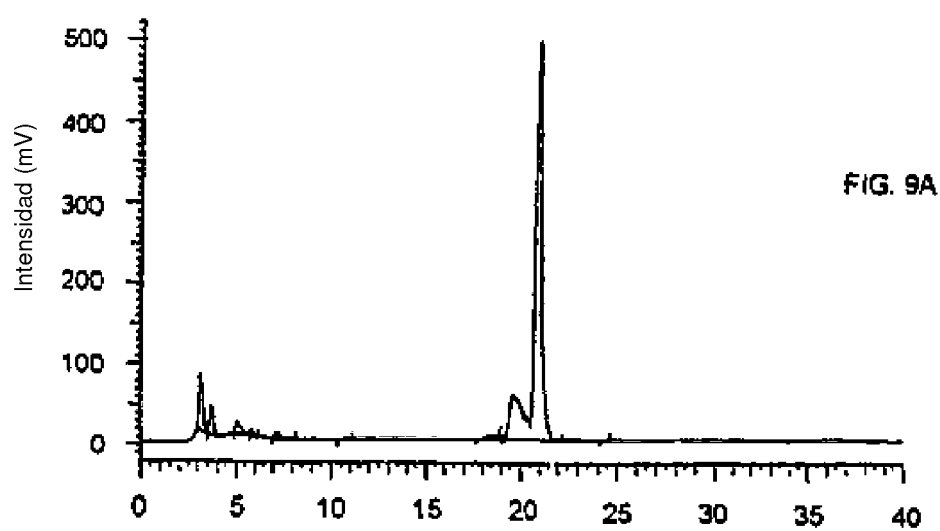


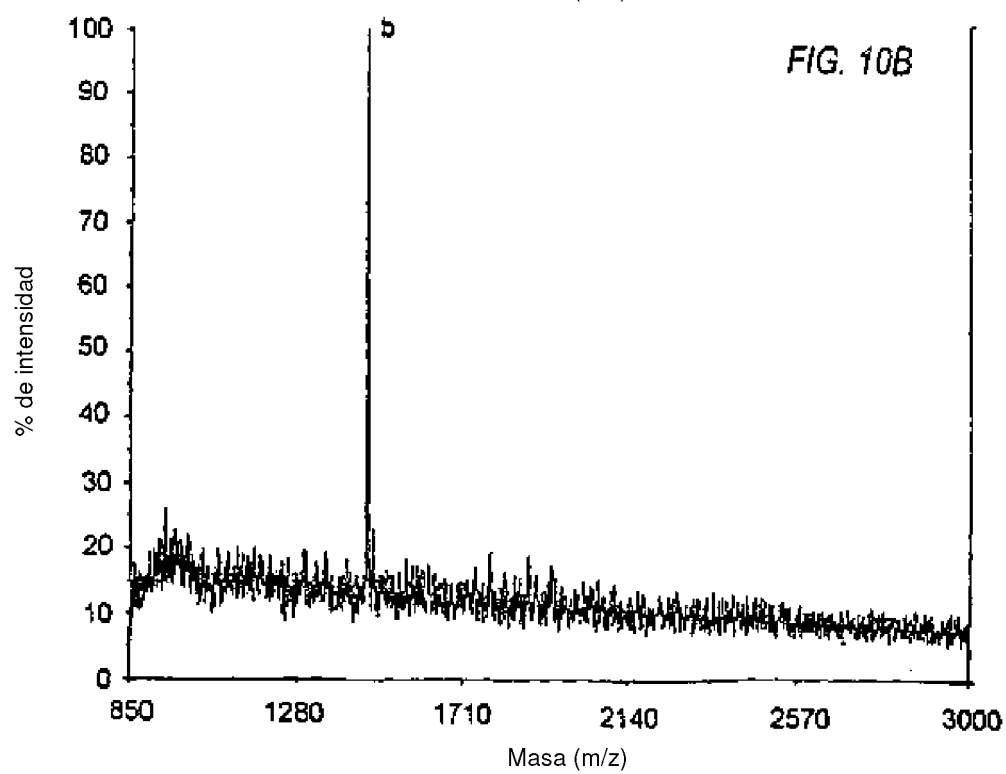
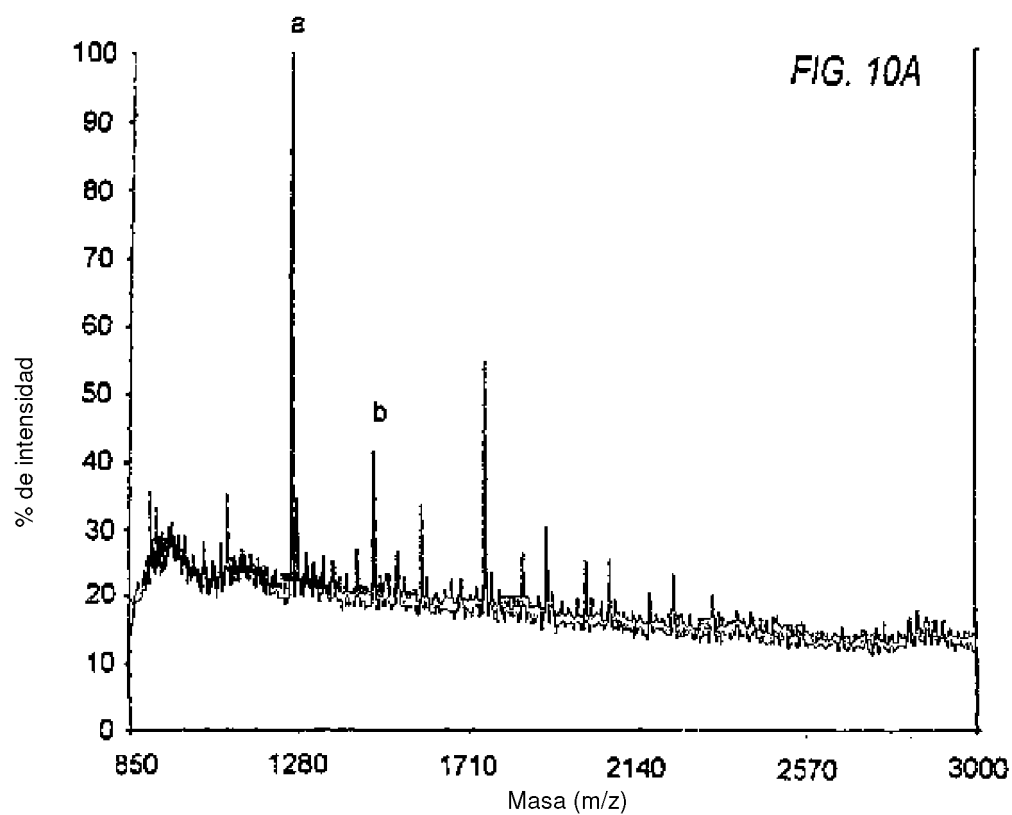
Fig. 5











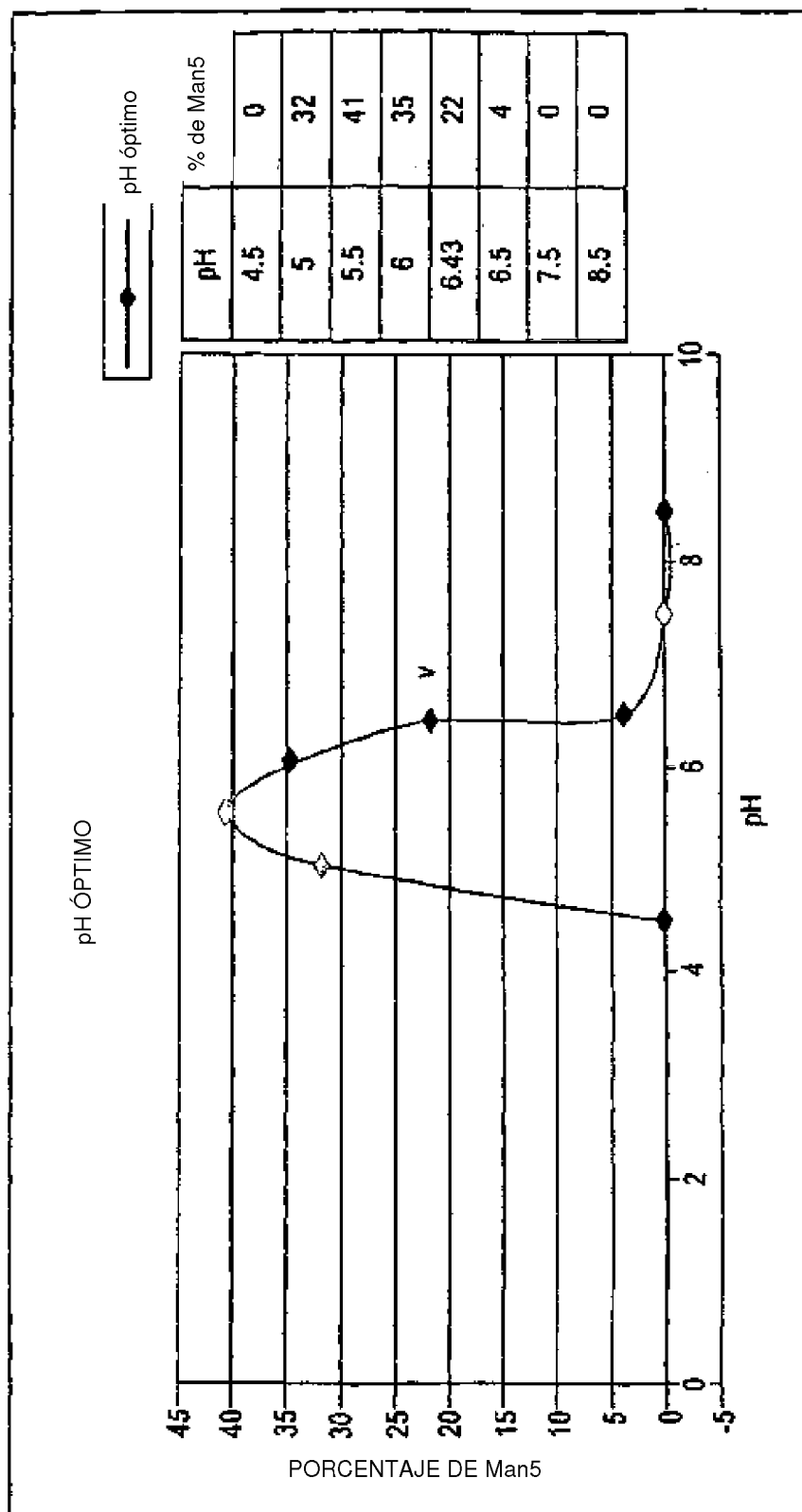
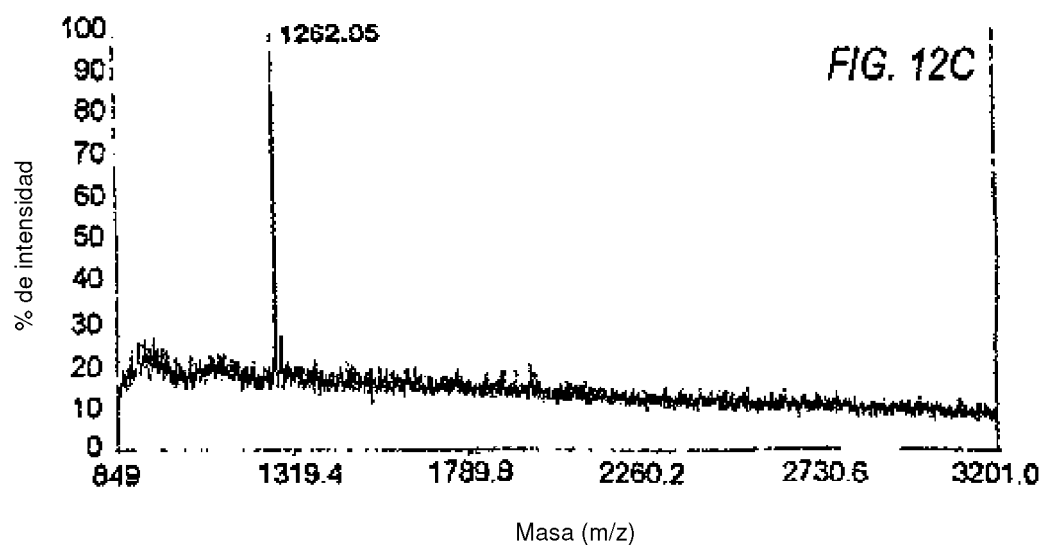
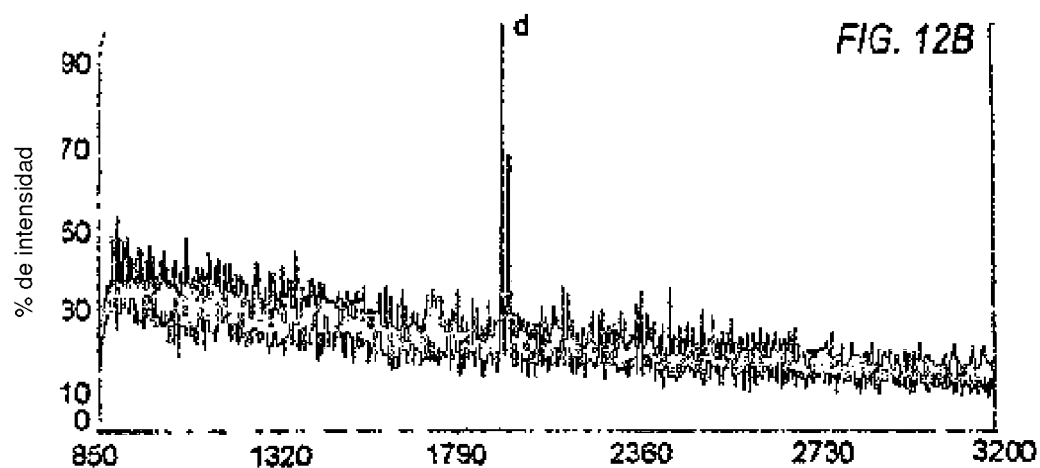
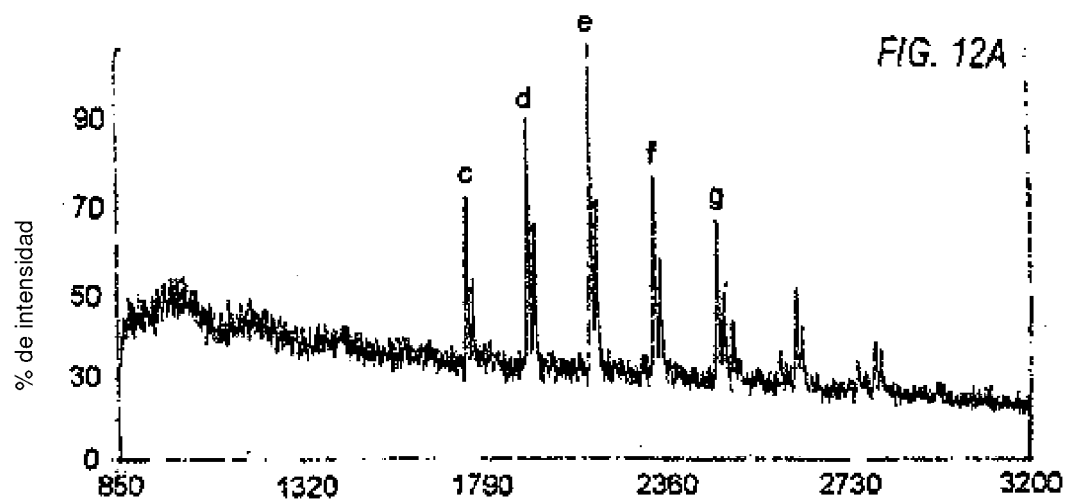


FIG. 11



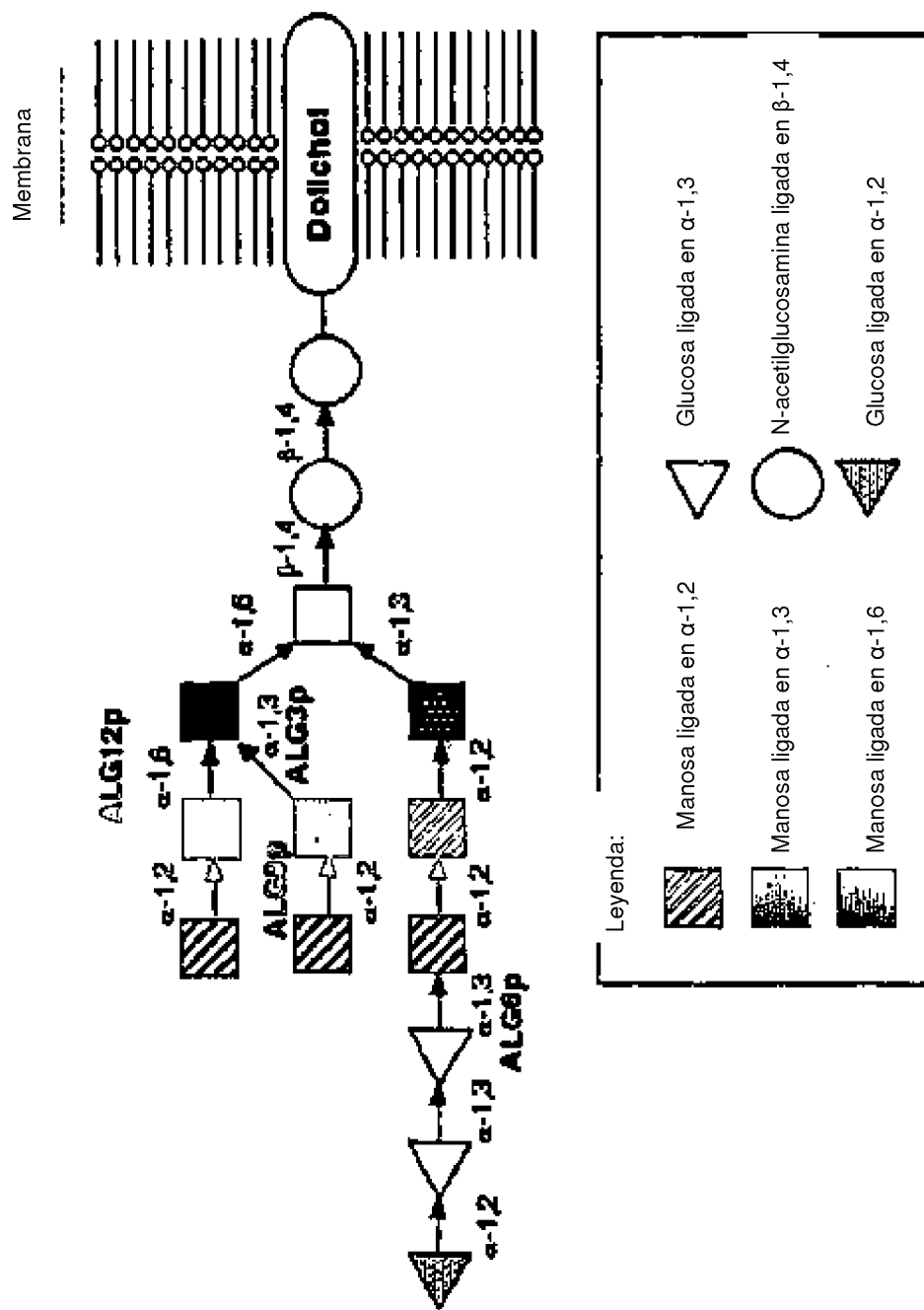


FIG. 13

N-glucanos ligados con lípidos

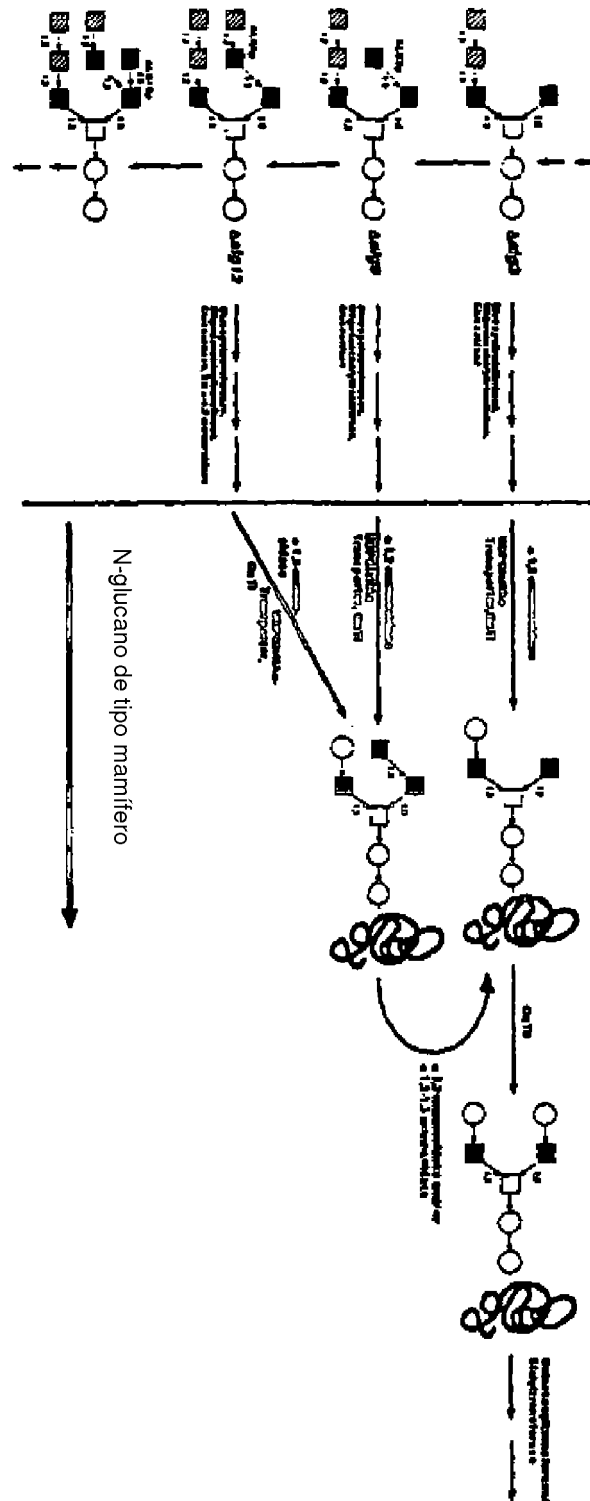


FIG. 14

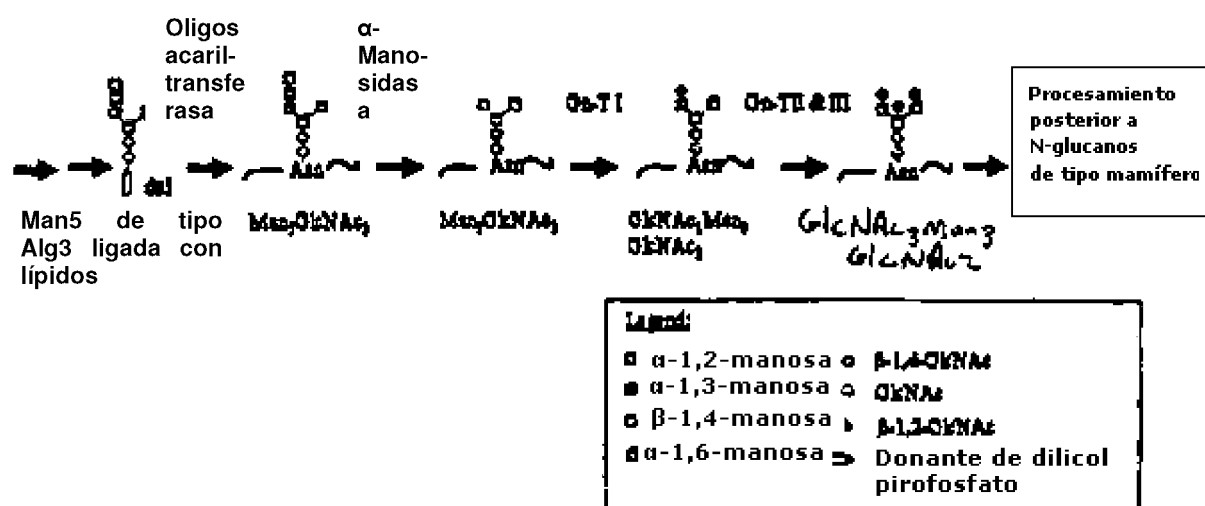


FIG. 15

ALG3 Blast 05-22-01

Secuencias que producían alineaciones significativas:

bits valor

```

gi|586444|sp|P38179|ALG3_YEAST DOLICHYL-P-MAN:MAN(5)GLCNAC(...797 0.0
gi|3024226|sp|Q92685|ALG3_HUMAN DOLICHYL-P-MAN:MAN(5)GLCNAC...173 7e-43
gi|3024221|sp|Q24332|NT56_DROVI LETHAL(2)NEIGHBOUR OF TID P...145 3e-34
gi|3024222|sp|Q27333|NT56_DROME LETHAL(2)NEIGHBOUR OF TID P...121 3e-27
gi|10720253|sp|P82149|NT53_DROME LETHAL(2)NEIGHBOUR OF TID ...121 5e-27
gi|1707982|sp|P40989|GLS2_YEAST 1,3-BETA-GLUCAN SYNTHASE CO... 32 2.8
gi|1346146|sp|P19631|GLS1_YEAST 1,3-BETA-GLUCAN SYNTHASE CO... 31 6.6

```

Alineación

Levadura:

```

>gi | 586444 | sp | P38179 | ALG3_DOLICHYL-P-MAN:MAN(5)GLCNAC(2).PP.DOLICHYL
MANOSILTRANSFERASA DE LEVADURA
ALFA(1-3)-MANOSILTRANSFERASA DEPENDIENTE DE DOL-P-MAN
PROTEÍNA DE RESISTENCIA A TOXINA INACTIVADORA HM-2
LONGITUD = 458

```

Puntuación = 797 bits (2059), Esperada = 0,0

Identidades = 422/458 (92%), Positivos = 422/458 (92%)

```

Query: 1  MEGEQSPQGEKSLQKQFVRPFLDLWQDLKDGVRVYVIFDCRANLIVMPLLLIFESMLCKI 60
        MEGEQSPQGEKSLQKQFVRPFLDLWQDLKDGVRVYVIFDCRANLIVMPLLLIFESMLCKI
Sbjct: 1  MEGEQSPQGEKSLQKQFVRPFLDLWQDLKDGVRVYVIFDCRANLIVMPLLLIFESMLCKI 60

Query: 61  IIKKVAYTEIDYKAYMEQIEMIQLDGFLDYSQVSGGTGSLVYPAGHVLIYKQMYWLTEGM 120
        IIKKVAYTEIDYKAYMEQIEMIQLDGFLDYSQVSGGTGSLVYPAGHVLIYKQMYWLTEGM
Sbjct: 61  IIKKVAYTEIDYKAYMEQIEMIQLDGFLDYSQVSGGTGSLVYPAGHVLIYKQMYWLTEGM 120

Query: 121 DHVERGQVFFRYLYLLTLALQMACYLLHLPPWCVVLACLSKRLHSIYVLRLEFNDCTTL 180
        DHVERGQVFFRYLYLLTLALQMACYLLHLPPWCVVLACLSKRLHSIYVLRLEFNDCTTL
Sbjct: 121 DHVERGQVFFRYLYLLTLALQMACYLLHLPPWCVVLACLSKRLHSIYVLRLEFNDCTTL 180

Query: 181 FMVVTVLGAIVASRCHQRPKLKSLALVISATYSMAVSIKMNALLYFPAMMISLFIILNDA 240
        FMVVTVLGAIVASRCHQRPKLKSLALVISATYSMAVSIKMNALLYFPAMMISLFIILNDA
Sbjct: 181 FMVVTVLGAIVASRCHQRPKLKSLALVISATYSMAVSIKMNALLYFPAMMISLFIILNDA 240

Query: 241 NVILTLLDLVAMIANQVAVAVPFLRSFPQQYLHCAFNFGRKFMYQWSINWQMDEEAFND 300
        NVILTLLDLVAMIANQVAVAVPFLRSFPQQYLHCAFNFGRKFMYQWSINWQMDEEAFND
Sbjct: 241 NVILTLLDLVAMIANQVAVAVPFLRSFPQQYLHCAFNFGRKFMYQWSINWQMDEEAFND 300

Query: 301 KRFXXXXXXXXXXXXXXXXXXKVTRYPRILPDLWSSLCHPLRKNVAVLNANPAKTIFFVLIASN 360
        KRF FVTRYPRILPDLWSSLCHPLRKNVAVLNANPAKTIFFVLIASN
Sbjct: 301 KRFRLALLISHLIALTTLFVTRYPRILPDLWSSLCHPLRKNVAVLNANPAKTIFFVLIASN 360

Query: 361 FIGVLFERSLHYQFLSWYHNTLPILIFWSGMPFFVGPDIWYVLHENCWNSYPFNSQXXXQX 420
        FIGVLFERSLHYQFLSWYHNTLPILIFWSGMPFFVGPDIWYVLHENCWNSYPFNSQ
Sbjct: 361 FIGVLFERSLHYQFLSWYHNTLPILIFWSGMPFFVGPDIWYVLHENCWNSYPFNSQASTLL 420

Query: 421 XXXXXXXXXXXXXXXXXXXXSGSVALAKSHLRITSSMEKKLN 458
        SGSVALAKSHLRITSSMEKKLN
Sbjct: 421 LALNTVLLLLLALTQLSGSVALAKSHLRITSSMEKKLN 458

```

FIG. 16-1

ES 2 336 790 T3

H. sapiens

>gi | 30242261 | sp | Q92685 | ALG3_DOLIQUIL-P-MAN:MAN(5)GLCNAC(2).PP.DOLIQUIL
MANOSILTRANSFERASA DE LEVADURA
ALFA(1-3)-MANOSILTRANSFERASA DEPENDIENTE DE DOL-P-MAN
PROTEÍNA DE TIPO NOT56
LONGITUD = 438

Puntuación = 173 bits (2059), Esperada = 7e-43

Identidades = 133/396 (33%), Positivos = 195/396 (48%), Gaps = 28/396 (7%)

```
Query: 26 WQDLKDCVRVIFDCRANLIVMFLILFESMLCKIIKKVAYTEIDYKAYMEQIEMIQLD 85
      WQ+   R ++ + R L+V L L E + +I -VAYTEID+KAYM ++E + ++
Sbjct: 29 WQER----RLLRPRYTLVLAACLCLAEVGITFWVLRVAYTEIDWKAYMAEVEGV- IN 83

Query: 86 GMLDYSQVSGGTGPLVYPAGHVLIIYKMYWLTGMDHVERGQVFRYLYLLTLALQMACY 145
      G DY+Q+ G TGPLVYPAG V I+ +Y+ T + Q F LYL TL L Y
Sbjct: 84 GTDYDTQLQGDGTGPLVYPAGVYVIFMGLYYATSRGTDIRMAQNIFAVLVLAATLLLVFLIY 143

Query: 146 Y+LLHLPPHC-VVLACL SKRLHSIYVLRLFNDCTTLFMVVTVLGAIVASRCHORPKLKK 203
      + +PP+ + C S R+HSI+VLRLFND + + +L + OR
Sbjct: 144 HQICKVPPFFVFFMCCASYRVHSIFVLRLFNDP-----VAMVLLFLSINLLLAQRNGWG- 197

Query: 204 SLALVISATYSMAVSIMKNALLYFPAMHISLFILNDANVILTLLDLVAMIANQVAVAVPF 263
      +S+AVS+KQW LL+ P ++ L L L + A + CV + +PF
Sbjct: 198 -----CCFFSLAVSVKQWVLLFAPGLLFLLLLTQPGFRGALFKLGICAGL--QVVLGLPF 249

Query: 264 LREFFQQLYHCAFNFGRKPFMYQNSINWQFMDEEAFNDKRFXXXXXXXXXXXXXXXXXVUTRY 323
      L P YL +F+ GR+P++ W++NW+ + E P + F + E+
Sbjct: 250 LLENFSGYLSRSFDLGRQFLPHWTVWVRFLPEALFLNRAFLALLTANLTLLLLFALCRW 309

Query: 324 PRILFDLWSSLCHPLRIQAVLNANPAKTITFVLIASNFIOVLFPSRLHYQFLSWYHTLP 383
      R + S L P ++ I L SNFIQ+ FSRSLHYQF WY TLP
Sbjct: 310 HRTGESILSLLRDPENRKKVFPQPLTPNQIVSTLFTSNFIGICFSRSLHYQFYVWYHTLP 369

Query: 386 ILIF----NSGMFFVGPINVTLHEHCNMSYPFNS 414
      L++ H + + + E NN+YF S
Sbjct: 370 YLLNAMPARNLTMLLALLVLGLI--ELSMNTYPSTS 403
```

Drosophila Vi

>gi | 30242211 | sp | Q24332 | NT56_COLINDANTE DE DROVI LETAL (2) DE LA PROTEÍNA
TID (NOT58) LONGITUD = 526

Puntuación = 145 bits (366), Esperada = 3e-34

Identidades = 103/273 (37%), Positivos = 157/273 (56%), Gaps = 17/273 (6%)

```
Query: 33 VRYVIFDCRANLIVMFLILFESMLCKIIKKVAYTEIDYKAYMEQIEMIQLDGMLDYSQ 92
      ++Y+ F+ A IV L++L E+++ ++I++V YTEID+KAYM++ E L+G +YS
Sbjct: 34 IKYLAFEPAALPIVSVLIVLAEAVINVLVIQRPVYTEIDWKAYMQECTGF-LNGITNYSL 92

Query: 93 VSGGTGPLVYPAGHVLIIYKMYWLTGMDHVERGQVFRYLYLLTLALQMACYLLH-LP 151
      + G TGPLVYFA V IY +Y+LT +V Q F +YLL + L + Y +P
Sbjct: 93 LRGDIGPLVYPAAFVYIYSGLYYLTGQGTNVRLAQYIFACIYLLQMLVRLDYTKSRKVP 152

Query: 152 FWCVVLACL-SKRLHSIYVLRLFNDCTTLFMVVTVLGAIVASRCHORPKLKKSLALVIS 210
      P+ +VL+ S R+HSIYVLRLFND L +L A + OR L S
Sbjct: 153 FVYVLSAFTSYRIHSIYVLRLFNDPVAIL-----LLYAALNLFDRWTLG-----S 200

Query: 211 ATYSMAVSIMKNALLYFPAMHISLFILNDANVILTLLDLVAMIANQVAVAVPFELRSFQ 270
      YS+AV +KQW + A + LP L + V+ TL+ L Q+ + PFLR+ P +
Sbjct: 201 ICYSLAVGVKQW--ILLFAPALLFLYLANLGVLRITLVQLTICAVLQLFICAPFLRTNPME 258
```

FIG. 16-2

ES 2 336 790 T3

Query: 271 YLHCAFNFGRKFMYQWSINWQMDEEAFNDKRF 303
 YL +F+ GR F ++W++N++ + +E F + F
 Sbjct: 259 YLRGSFDLGRIFENKWTVNYRFLSKELFEQREF 291

Score = 53.3 bits (177), Expect = 1e-06
 Identities = 31/62 (50%), Positives = 41/62 (66%), Gaps = 6/62 (9%)

Query: 352 IPFVLIASNFIVGLFSRSLHYQFLSWYHWTLPIILIFWSGMFFVGPINWYVLH--EWQWNS 409
 +PF L NFIGV +RSLHYQF WY +LP L+ WS P+ +G + +L E+CWN+
 Sbjct: 412 LPFFL--CNFIGVACARSLHYQFYINWYFNSLPYLV-WS-TPYSLGVRYLILGITEYCWN 467

Query: 410 YP 411
 YP
 Sbjct: 468 YP 469

Drosophila melanogaster

>gi | 3024222| sp | Q27333 | NT56_COLINDANTE DE DROVI LETAL (2) DE LA PROTEÍNA
 TID (NOT56) (NOT45)
 LONGITUD = 510

Puntuación = 121 bits (305), Esperada = 3e-27
 Identidades = 96/272 (35%), Positivos = 154/272 (56%), Gaps = 17/272 (6%)

Query: 34 RYVIFDCRAMLIIVMPLILFESHICKI IIKCVAYTEIDYKAYMEQIENIQDGMIDYSQV 93
 +Y++ + A IV ++L E ++ ++I++V YTEID+ AYM++ E L+G +YS +
 Sbjct: 36 KYLLLEPAALPIVGLFVLLAELVINVVVIQRPVTEIDWVAYMQECEGF-LNGTINYSLL 94

Query: 94 SEGTSPLVYPAGHVLIYKMMYWLTEGMORVERGOVFFRYLYLLTLALOMACTYLLH-LPP 152
 G TGPLVYPA V IY +Y++T +V Q F +YLL LAL + Y +PP
 Sbjct: 95 RGDTPGLVYPAAFVYIYSALYFVTSHGTVRLAQYIFAGIYLLQLALVRLYSKSRKVP 154

Query: 153 MCVVLAEL-SKRLHSIYVLRLEFNDCTILEMVTTLGAIVASRCHQRPFLKKSLALVISA 211
 + +VL+ S R+HSIYVLRLEFND + V +L A + +R L S
 Sbjct: 155 YVLVLSAFTSYRINSIYVLRLEFNDP-----VAVLLLYAALNLFLDRHWTLG-----ST 202

Query: 212 TYSMAVSIGNALLYFFPMISLFIENANVILTLDLVAMIANQVAVAVFFLRSPFQQY 272
 +S+AV +KMN + A + LF L + ++ T+L L Q+ + PFL + P +Y
 Sbjct: 203 FFSLAVGVIOGN--ILLFAPALLLFYLANLGLLRTILQLAVCGVIQLLLGAPFLLTHPVEY 260

Query: 272 LHCAFNFGRKFMYQWSINWQMDEEAFNDKRF 303
 L +F+ GR F ++W++N++ + + F ++ F
 Sbjct: 261 LRGSFDLGRIFENKWTVNYRFLSRDVFENRTF 292

Score = 45.4 bits (117), Expect = 2e-05
 Identities = 27/60 (45%), Positives = 35/60 (58%), Gaps = 2/60 (3%)

Query: 352 IPFVLIASNFIVGLFSRSLHYQFLSWYHWTLPIILIFWSGMFFVGPINWYVLH--EWQWNSYP 411
 +PF L N +GV SRSLHYQF WY +LP L + + V + L E+CWN+YP
 Sbjct: 407 LPFFL--CNLVGVACSRSLHYQFYVWYFHSPLAWSTPYSLGVRCILGLIEYCWN 464

FIG. 16-3

ES 2 336 790 T3

Matriz : BLOSUM62

Penalización por gap: Existencia: 11, Extensión: 1

Número de aciertos con BD: 28883317

Número de secuencias: 96469

Número de extensiones: 1107545

Número de extensiones con éxito: 2870

Número de secuencias superiores a 10,0: 16

Número de HSP superiores a 10,0 sin gaps: 5

Número de HSP con gaps exitosas en la prueba preliminar: 11

Número de HSP que se intentó regularizar en la prueba preliminar: 2839

Número de HSP regularizadas (no preliminar): 23

Longitud de la interrogación: 458

Longitud de la base de datos: 35.174.128

Longitud de HSP efectiva: 45

Longitud de la interrogación efectiva: 413

Longitud de la base de datos efectiva: 30.833.023

Espacio de búsqueda efectivo: 12734038499

Espacio de búsqueda efectivo usado: 12734036499

T: 11

A: 4D

X1: 15 (7,1 bits)

X2: 38 (14,6 bits)

X3: 64 (24,7 bits)

S1: 40 (21,8 bits)

S2: 67 (30,4 bits)

FIG. 16-4

S. cerevisiae ALG3

ATGGAAGGTGAACAGTCTCCGCAAGGTGAAAAGTCTCTGCAAAGGAAGC
 AATTTGTCAGACCTCCGCTGGATCTGTGGCAGGATCTCAAGGACGGTGTG
 CGCTACGTGATCTTCGATTGTAGGGCCAATCTTATCGTTATGCCCCTTTTG
 ATTTTGTTCGAAAGCATGCTGTGCAAGATTATCATTAAAGAAGGTAGCTTAC
 ACAGAGATCGATTACAAGGCGTACATGGAGCAGATCGAGATGATTCAGCT
 CGATGGCATGCTGGACTACTCTCAGGTGAGTGGTGGAAACGGGGCCCGCTGG
 TGTATCCAGCAGGCCACGTCTTGATCTACAAGATGATGTACTGGCTAACA
 GAGGGAATGGACCACGTTGAGCGCGGGCAAGTGTTTTTCAGATACTTGTA
 TCTCCTTACACTGGCGTTACAAATGGCGTGTTACTACCTTTTACATCTACC
 ACCGTGGTGTGTGGTCTTGGCGTGCCCTCTCTAAAAGATTGCACTCTATTTA
 CGTGCTACGGTTATTCAATGATTGCTTCACTACITTTGTTTATGGTCGTCACG
 GTTTTGGGGGCTATCGTGGCCAGCAGGTGCCATCAGCGCCCCAAATTA
 GAAGTCCCTTGCGCTGGTGATCTCCGCAACATACAGTATGGCTGTGAGCA
 TTAAGATGAATGCGCTGTTGTATTTCCCTGCAATGATGATTTCTCTATTTCAT
 CCTTAATGACGCGAACGTAATCCTTACTTTGTTGGATCTCGTTGCGATGAT
 TGCATGGCAAGTCGCAGTTGCAGTGCCCTTCCCTGCGCAGCTTTCCGCAACA
 GTACCTGCATTGCGCTTTTAAATTTCCGCGAGGAAGTTTATGTACCAATGGAG
 TATCAATTGGCAAATGATGGATGAAGAGGCTTTCAATGATAAGAGGTTCC
 ACTTGGCCCTTTTAAATCAGCCACCTGATAGCGCTCACCACACTGTTGCTCA
 CAAGATACCCTCGCATCCTGCCCGATTATGTTCTTCCCTGTGCCATCCGC
 TGAGGAAAAATGCAGTGCTCAATGCCAATCCCGCCAAGACTATTCCATTC
 GTTCTAATCGCATCCAACCTTCATCGGCGTCCTATTTTCAAGGTCCCTCCAC
 TACCAGTTTCTATCCTGGTATCACTGGACTTTGCCTATACTGATCTTTTGGT
 CGGGAATGCCCTTCTTCGTTGGTCCCATTTGGTACGTCTTGACAGAGTGGT
 GCTGGAATTCTATCCACCAAACCTCACAAGCAAGCACGCTATTGTTGGCA
 TTGAATACTGTTCTGTTGCTTCTATTGGCCTTGACGCAGCTATCTGGTTCCG
 TCGCCCTCGCCAAAAGCCATCTTCGTACCACCAGCTCTATGGAAAAAAG
 CTCAACTGA

S. cerevisiae Alg3p

MEGEQSPQGEKSLQRKQFVRPPLDLWQDLKDGVRVYVIFDCRANLIVMPLLI
 FESMLCKIIKKVAYTEIDYKAYMEQIEMIQLDGMLDYSQVSGGTGPLVYPAG
 HVLIYKMMYWLTEGMDHVERGQVFFRYLYLLTLALQMACYLLHLPPWCV
 VLACLSKRLHSIYVLRLEFNDCTTTFMVVTVLGAIVASRCHQRPKLKSLALV
 ISATYSMAVSIMNALLYFPAMMISLFILNDANVILTLLDLVAMIAWQVAVA
 VPFLRSFPQQYLHCAFNFGKRFMYQWSINWQMMDEEAFNDKRFHLALLISHL
 LALTTLFVTRYPRILPDLWSSLCHPLRKNVNLNANPAKTIPFVLIASNFIGVLFS
 RSLHYQFLSWYHWTLPLIFWSGMPFFVGPIWYVLHEWCWNSYPPNSQASTL
 LLALNTVLLLLLALTQLSGSVALAKSHLRRTSSMEKKLN

FIG. 17

P. pastoris ALG3

ATGCCTCCGATAGAGCCAGCTGAAAGGCCAAAGCTTACGCTGAAAAATGT
 TATCGGTGATCTAGTGGCTCTTATTCAAACGTTTTATTAAACCCAGATTTT
 AGTGTCTTCGTTGCACCTCTTTTATGGTTAGCTGATTCCATTGTTATCAAGG
 TGATCATIGGCACTGTTTCCTACACAGATATTGATTTTTCTTCATATATGCA
 ACAAATCTTTAAAATTTCGACAAGGAGAATTAGATTATAGCAACATATTTG
 GTGACACCGGTCCATTGGTTTACCCAGCCGGCCATGTTTCATGCTTACTCAG
 TACTTTTCGTGGTACAGTGATGGTGGAGAAGACGTCAGTTTCGTTCAACAA
 GCATTTGGTTGGTTATACCTAGGTTGCTTGTTACTATCCATCAGCTCCTACT
 TTTTCTCTGGCTTAGGGAAAATACCTCCGGTTTATTTTGTTTTGTTGGTAGC
 GTCCAAGAGACTGCATTCAATATTTGTATTGAGACTCTTCAATGACTGTTT
 AACAACATTTTTGATGTTGGCAACTATAATCATCCTTCAACAAGCAAGTAG
 CTGGAGGAAAGATGGCACAACCTATCCATTATCTGTCCCTGATGCTGCAG
 ATACGTACAGTTTAGCCATCTCTGTAAAGATGAATGCGCTGCTATACCTCC
 CAGCATTCCTACTACTCATATATCTCATTGTGACGAAAATTTGATTAAAG
 CCTTGGCACCTGTTCTAGTTTTGATATTGGTGCAAGTAGGAGTCGGTTATT
 CGTTCATTTTACCGTTGCACTATGATGATCAGGCAAATGAAATTCGTTCTG
 CCTACTTTAGACAGGCTTTTGACTTTAGTCGCCAATTTCTTTATAAGTGGA
 CGGTTAATTGGCGCTTTTTGAGCCAAGAACTTTCAACAATGTCCATTTTC
 ACCAGCTCCTGTTTGCTCTCCATATTATTACGTTAGTCTTGTTTCATCCTCAA
 GTTCTCTCTCTCTAAAAACATTGGAAAACCGCTTGGTAGATTTGTGTTGGA
 CATTTTCAAATTTTGGAAGCCAACCTTATCTCCAACCAATATTATCAACGA
 CCCAGAAAGAAGCCCAGATTTTGTTTACACCGTCATGGCTACTACCAACTT
 AATAGGGGTGCTTTTTTGCAAGATCTTTACACTACCAGTTCCTAAGCTGGTA
 TGCGTTCTCTTTGCCATATCTCCTTTACAAGGCTCGTCTGAACTTTATAGCA
 TCTATTATTGTTTATGCCGCTCACGAGTATTGCTGGTTGGTTTTCCAGCTA
 CAGAACAAAGTTCGCGGTTGTTGGTATCTATCTTACTACTTATCCTGATTC
 TCATTTTTACCAACGAACAGTTATTTCTTCTCAATCGGTCCCTGCAGAAA
 AAAAGAATACATAA

P. pastoris Alg3p

MPPIEPAERP KLT LKNVIGDLV ALIQNVLFNPDFS VFVAPLLWLADSI VIKVIIG
 TVSYTDIDFSSYMQQIFKIRQGELDYSNIFGDTGPLVYPAGHVHAYSVLSWYS
 DGGEDVSFVQQA FGWLYLGCLLSISSYFFSGLGKIPVYFVLLVASKRLHSIF
 VLRLFNDCLTTF LMLATHILQQASSWRKDGTTPLSVPDAADTYS LAISVKMN
 ALLYLPAFLLLIYLICDENLIKALAPVLVLILVQVGVGYSFILPLHYDDQANEIR
 SAYFRQAFDFSRQFLYKWTNVNWRFLSQETFN NVHFHQLLFALHITLVLFILKF
 LSPKNIGKPLGRFVLDFKFWKPTLSPTNINDPERSPDFVYTVMATTNLIGVLF
 ARSLHYQFLSWYAFSLPYLLYKARLNFIASIVYAAHEYCWLVPATEQSSAL
 LVSILLILILIFTNEQLFPSQSVPAEKKNT

FIG. 18

ES 2 336 790 T3

Neurospora crassa

Puntuación = 212 bits (540), Esperada = 8e-54

Identidades = 140/400 (35%), Positivos = 212/400 (53%), Gaps = 29/400 (7%)

```

Query: 35 SVFVAFLLWLADSIIVIRVIIGTVSYTDIDFSSYMOQIFKIROGELDYSNIFGDTGPLVYP 34
S + P L+L D+++ +II V YT-ID+++YM+Q+ +I GE DY+ + G TGPLVYP
Sbjct: 33 SKLIPPALFLVDALLCGLI INKVPYTEIDWAAYMEQVSQILSGERDYTEVRGGTGPLVYP 92

Query: 95 AGHVHAYSIVLSWYSDGGEDVSFVQQAFCWLYLGCILLSSISSYFSGLGKIPPVYFVLLWA 154
A HV+ Y+ L +D Q ++ QQ F LY+ L + + Y+ K PP F LL
Sbjct: 93 AAKVYIYTGLYHLTDEGENTLLAQQLFAGLYMVLAVVMGCYH---QAKAPPYLFPLLTL 149

Query: 155 SKRLHSIFVLRLEFNDCLITFLMLATITIILOQASSWRKDGTTIFLSVPDAADTYSLAISVK 214
SKRLHSIFVLR FNDCLITFLMLATITIILOQASSWRKDGTTIFLSVPDAADTYSLAISVK
Sbjct: 150 SKRLHSIFVLRLEFNDCLITFLMLATITIILOQASSWRKDGTTIFLSVPDAADTYSLAISVK 214

Query: 215 MXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 274
M + + L F+ HY + Y
Sbjct: 198 MTLLESLPAVGIVLFLGSG-SFVTTLQVLATMGLVQILIGVDFL---AHYPTL-----Y 247

Query: 275 FROAFDFSRQFLYKWTVMWRFLSQETFMNVHFMQLLALHITLVLFI-LKPLSPKNIK 331
+AF+ SRQF +KWTVMWRFLSQETFMNVHFMQLLALHITLVLFI-LKPLSPKNIK
Sbjct: 248 LSRAPFELSRQFFKWTVMWRFLSQETFMNVHFMQLLALHITLVLFI-LKPLSPKNIK 305

Query: 334 ELGRFVLDIFKFWKPTLS-PTNIINDPERSPDFVYTMATTNLIGVLFARSLHYQFLSWY 392
L + + + KP L+ P + + +P + + T + + M +G-LFARSLHYQF +
Sbjct: 306 SLVQLISFVLLAGRPPLTVPEKTAARDVTPRYIMTTILSANAVGLLFARSLHYQFYAYV 369

Query: 393 AFSLPYLLYKARLNFIAIIVYAAWEYCHLVFPATEQSSA 432
A+S P+LL+A L+ + +A HE+ W VFP+T SSA
Sbjct: 366 AWSTPFLWRAGLRPVLYVLLNAVHEWAWKVPFSTPASSA 405

```

Schizosaccharomyces pombe

Puntuación = 176 bits (405), Esperada = 8e-43

Identidades = 132/390 (33%), Positivos = 154/390 (49%), Gaps = 35/390 (8%)

```

Query: 42 LWLADSIIVIRVIIGTVSYTDIDFSSYMOQIFKIROGELDYSNIFGDTGPLVYPAGHVHAY 101
L L + + II V YT-ID+ +YM+Q+ GE DY ++ G TGPLVYP GHV Y
Sbjct: 30 LLLLEIPFVFAIISKVPYTEIDWIAAYMEQVNSFLGGERDYKSLVGCTGPLVYPQGHVFLY 89

Query: 102 SVLSWYSDGGEDVSFVQQAFCWLYLGCILLSSISSYFSGLGKIPPVYFVLLVASKRLHSI 161
++L + +DGG ++ Q F ++Y + +I Y F + + P +VLL+ SKRLHSI
Sbjct: 90 TLLYTLTDGCTNIIVRAQYIFAFVYH--ITTAIVGYLFK-IVRAPFYIYVLLILSKRLHSI 146

Query: 162 FVLRLEFNDCLITFLMLATITIILOQASSWRKDGTTIFLSVPDAADTYSLAISVKMXXXXX 221
F+LRLFNDCLITFLMLATITIILOQASSWRKDGTTIFLSVPDAADTYSLAISVKMXXXXX
Sbjct: 147 FVLRLEFNDCLITFLMLATITIILOQASSWRKDGTTIFLSVPDAADTYSLAISVKMXXXXX 221

Query: 222 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 281
L++ L P + + + +Y+ QAFDF
Sbjct: 195 PAYLVL-----LLQILGPKXTWGHIFVITVQILFSIPF----LAYFWSYWTQAFDF 242

Query: 282 SRQFLYKWTVMWRFLSQETFMNVHFMQLLALHITLVLFI-LKPLSPKNIKPLGRFVLD 341
R F YKWTVMWRFLSQETFMNVHFMQLLALHITLVLFI-LKPLSPKNIKPLGRFVLD
Sbjct: 241 GRAFDYKWTVMWRFLSQETFMNVHFMQLLALHITLVLFI-LKPLSPKNIKPLGRFVLD 295

Query: 342 IFKFWKPTLSPTNIINDPERSPDFVYTMATTNLIGVLFARSLHYQFLSWYAFSLPYLLY 401
F L+ + + +P+P+T +AT+NLIG+L ARSLHYQF +W+A+ PYL Y

```

FIG. 19-2

Sbjct: 296 -FAMVNSMLTLKPLPKLQLATPNFIFTALATSNLIGILCARSLHYQFYAWFAWYSPYLCY 354

Query: 402 KARLNFIASIIIVYAAHETCWLVPATEQSS 431

+A I ++ EY W VFP+T+ SS

Sbjct: 355 QASFPAPIVIGLWMLQEYAWNVPSTKLS 384

Arabidopsis thaliana

Puntuación = 164 bits (415), Esperada = 2e-39

Identidades = 131/391 (33%), Positivos = 194/391 (49%), Gaps = 29/391 (7%)

Query: 42 LWLADSIIVIKVIIGTVSYTDIDFSSYMQQIFKIROGELDYSNIFGDTGPLVYPAGHVHAY 101

L LAD+I++ +II V YT ID+ +YM Q+ GE DY E+ GDTGPLVYPAG ++ Y

Sbjct: 39 LILADAILVALLIAYVPYTKIDWDAYMSQVSGFLGGERDYGNLKEGDTGPLVYPAGFLVYVY 98

Query: 102 SVLSWYSDGGEDVSVFQQAFGWLYLGCLLLSISSTFFSGLGKI9PVYFVLLVASKRLHSI 161

S + + G +V Q FG LY+ L + + Y + + +P LL SKR+HSI

Sbjct: 99 SAVQNLITGG--EVYPAQILPGVLYIVNLGIVLITYVKTGV--VPWALSLCLSKRHSI 154

Query: 162 FVLRLFNDCLTTFMLATIILQOASSWRNDGTTIFLSVPDAADTYSLAISVQDQXXXKX 221

FVLRLFNDC L+ A++ + +RK + + +S A+SVQDQ

Sbjct: 195 FVLRLFNDCFAMTLHASTALFL---YRCHLGLLV-----FSGAVSVQDQVLLYA 302

Query: 222 XXXXXXXXXCDENLIKALAPXXXXXXXCYSFILPLHYDDQANEIRSAYTROAFDF 281

N+I ++ F++ +Y AFD

Sbjct: 203 PTLRLLLKAM--NIIGVVSALAGAALAQILVGLPFLITYPV-----SYIANAFDL 251

Query: 282 SRQFLYKWTYNWRFLSOETFNNVHFHQLLFALHIIITLVLFILKFLSPKNIGKPLGRFVLD 341

R F++ W+VN++F+ + F + F L H+ LV F + K+ G +G

Sbjct: 252 GRVFIRFHSVVFKEFVPERVFSKEFAVCLLIAHLFLLVAFANRKNCKHEGGIIGFMRSR 310

Query: 342 IFKFWKP-TLSPTMIINDPERSDFVTTVMATTNLIGVLFARSLHYQFLSWYAFSLPYLL 400

F F +LS +++ + + V T N N IG++FARSLHYQF SWY +SLPYLL

Sbjct: 311 HFFLTLPSSLSFSDVSASRIITKEHVVTAMFVGNFIGIVFARSLHYQFYSWYFYSLPYLL 370

Query: 401 YKARLNFIASIIIVYAAHEYCWLVPATEQSS 431

++ +I++ E CN V+P+T SS

Sbjct: 371 WRTFFPTMLRLIMFLGIELCWNVYPSTPSSS 401

FIG. 19-3

K. lactis **ALG3**

TTTGTTTACAAGCTGATACCAACGAACATGAATACACCGGCAGGTTTACT
GAAGATTGGCAAAGCTAACCTTTTACATCCTTTTACCGATGCTGTATTCAG
TGCGATGAGAGTAAACGCAGAACAAATTGCATACATTTTACTTGTTACCA
ATTACATTGGAGTACTATTGCTCGATCATTACACTACCAATTCCTATCTT
GGTACCATTGGACGTTACCAGTACTATTGAATTGGGCCAATGTTCCGTATC
CGCTATGTGTGCTATGGTACCTAACACATGAGTGGTGCTGGAACAGCTAT
CCGCCAAACGCTACTGCATCCACACTGCTACACGCGTGTAACACATACTG
TTATTGGCTGTATTCTTAAGAGGACCCGCAAACCTCGAAAAGTGGTGATAA
CGAAACAACACACGAGAAAGCTGAG

K. lactis **Alg3p**

FVYKLIPTNMNTPAGLLKIGKANLLHPFTDAVFSAMRVNAEQIAYILLVTNYI
GVLFARSLHYQFLSWYHWTLPVLLNWANVPYPPLCVLWYLTHEWCWNSYPP
NATASTLLHACNTYCYWLYSZEDPQTRKVVTTKQHTRKL

FIG. 20

***K. lactis* ALG3 BLAST**

Puntuación

Secuencias que producían alineaciones significativas:

bits

valor

<u>gi 586444 sp P38179 ALG3 YEAST Dolichyl-P-Man:Man(5)GlcNAc(...125</u>	<u>1e-28</u>
<u>gi 984725 gb AAA75352.1 ORF 1</u>	<u>4e-19</u>
<u>gi 16226531 gb AAL16193.1 AF428424.1 At2g47760/F17A22.13 {A... 72</u>	<u>1e-12</u>
<u>gi 25367230 pir B84919 Not56-like protein [imported] - Ara... 72</u>	<u>1e-12</u>
<u>gi 21292031 gb EAA04176.1 agCP3388 [Anopheles gambiae str.... 69</u>	<u>2e-11</u>
<u>gi 20892051 ref XP_148657.1 similar to Lethal(2)neighbour ... 65</u>	<u>2e-10</u>

Alineaciones

S. cerevisiae

Puntuación = 125 bits (314), Esperada = 1e-28

Identidades = 60/120 (50%), Positivos = 83/120 (69%), Gaps = 1/120 (0%)

Marco = + 3

Query: 66 ANLLKFFT-DAVFSAMRVNABQIAYILLVTNYIGVLFARSLHYQFLSWYHWTLPVLLNWA 262
 ++L HP +AV +A A+ I ++L+ +N+IGVLF+RSLHYQFLSWYHWTLP+L+ W+

Sbjct: 332 SSLCHPLRKNVLRNRP--AKTIPFVLIASNFIGVLFPSRSLHYQFLSWYHWTLPIILFWS 389

Query: 243 NVYPYPLCVLWYLTWENWCWNSYPFNATASTLLHACNTYCYWLYS*EDPQTRKVVITKQHT 422
 +P+ + +WY+ HENCWNSYPFN+ ASTLL A NT .L + + V + K H R

Sbjct: 390 GHPFFVGPINIVLHENCWNSYPFNASQASTLLALNTVILLLLA-LTQLSGSVALAKSHLR 448

A. thaliana

Puntuación = 72,0 bits (175), Esperada = 1e-12

Identidades = 42/107 (39%), Positivos = 57/107 (53%), Gaps = 3/107 (2%)

Marco = + 3

Query: 84 FIDAVFSAMRVNABQIAYILLVTNYIGVLFARSLHYQFLSWYHWTLPVLLNWANVFPYFLC 263
 F+D S + + E + + V N+IC++FARSLHYQF SWY ++LP LL P L

Sbjct: 322 PSDVBSRI-ITKENVVTAMFVGNFIGIVFARSLHYQFYSNYFYSLPVLLNWRTPFPPTWLR 380

Query: 264 VLWYLTWENWCWNSYPFNATASTL---LHACNTYCYWLYS*EDPQTRK 395
 ++ +L E CWN YP ++S L LH WL DP K

Sbjct: 381 LIMFLGIELCWNVYPSTPSSSGLLLCLHLIILVGLWLAPSVDFYQLK 427

FIG. 21

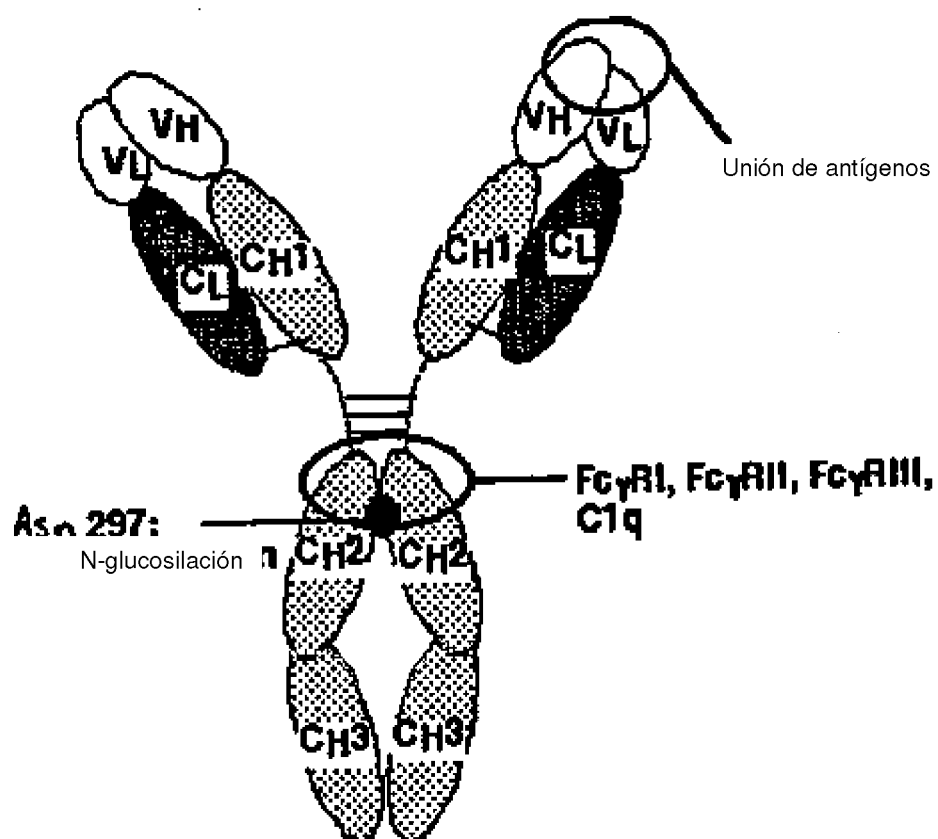


FIG. 22

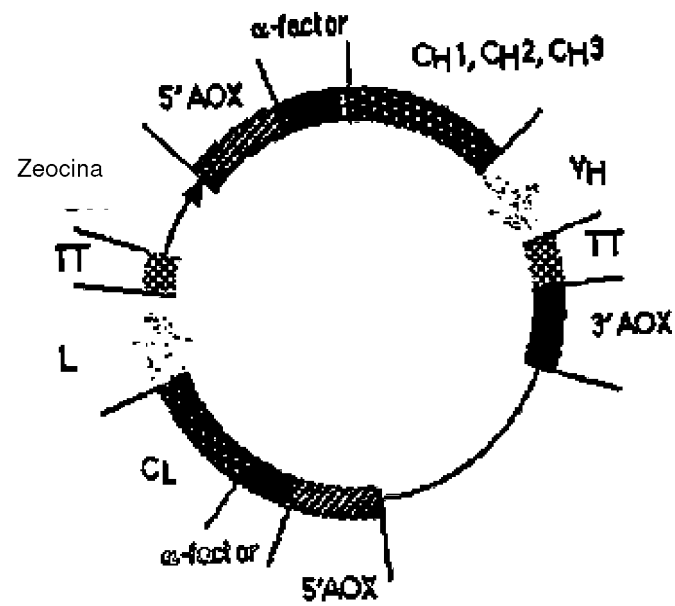


FIG. 23

>gi|6754685|ref|NM_010793.1|

Manosida acetil glucosaminiltransferasa 3 (mgat3) ARNm de *Mus musculus*

ATGAAGATGAGACGGCTACAAGCTCTTTCTCATGTTCTGTATGGCTGGCCTGTGCTCATATCCTTCCTGC
 ACTTCTTTAAGACCTTATCCTATGTACCTTCCCGAGAGAACTGGCCTCCCTCAGCCCTAACCTCGTATC
 CAGCTTCTTCTGGAACAATGCCCTGTCACTCCCCAGGCCAGTCCGGAGCCCGGTGGCCCCGACCTATTG
 CGGACACCCCTCTACTCCCACTCTCCCTGCTCCAGCCACTGTCCCCGAGCAAGGCCACAGAGGAACTGC
 ACCGGGTGGACTTCTGTGTTGCCGGAGGACACCACGGAGTATTTTGTGCGCACCAAGCTGGTGGTGTGTG
 CTTCAAACCAGGTACCAAGATGCTGGAGAAACCTTCGCCAGGGCGGACAGAGGAGAAGCCCCGAAGTGTCT
 GAGGGCTCCTCAGCCCGGGGACCTGCTCGGAGGCCCATGAGGCACGTGTTGAGTACCGCGGAGCGCCTGG
 GCAGCCGGGGCACTAGGCGCAAGTGGGTTGAGTGTGTGTGCTGCCAGGCTGGCACGGGCCCAGTTGCGG
 GGTGCCACCGGTGGTGCAGTATTCGAACCTGCCCAACCAAGGAACGCCCTGGTACCCAGGAGGTACCGGAG
 CGGGTTATCAACGCCATCAACATCAACCAAGAGTTCGACCTGCTGGATGTGCGCTTCCATGAGCTGGGAG
 ATGTTGTGGACCCCTTCGTGGTCTGTGAATCTAATTCACCGCCTACGGGGAGCCTCGGCCGCTCAASTT
 CCGAGAGATGCTGACCAATGGCACCTTCAGTACATCCGCCACAAGGTGCTCTATGTCTTCCTGGACCAT
 TTCCACCTGGTGGCCGTCAGGACGGCTGGATTGCGGATGACTACCTGCGCACCTTCCTCACCAGGATG
 GCGTCTCCCGCCTGCGCAACCTGCGGCCCGATGACGTCTTTATCATCGACGATGCGGACGAGATCCCTGC
 GCGTATGGTGTGCTGTTCCTCAAACCTTACGATGCTGGACAGAGCCCTTCGCTTCCACATGCGGAAG
 TCCCTGTATGGTTTCTTCTGGAAGCAGCCCGGACACTGGAGGTGGTGTGAGGCTGCACCATGGACATGC
 TGCAGGCCGTGTATGGGCTGGATGGCATCGGCTGCGCCGCGCCAGTACTACACCATGCCCAACTTCCG
 GCAGTATGAGAACCGCACCGCCACATCCTAGTGCAGTGGTCTCTCGGCAGCCCCCTGCACTTCGCGGGC
 TGGCATTGTCTCTGCTCTTCCACACCCGAGGCTCTACTTTAACTCGTGTGAGCCAGAAATGGCGACT
 TCCCCCGCTGGGCTGACTATGAGGACAGAGCGGACCTCAATTACATCCGCACTTGTATCCGCACTGGGG
 ATGOTTGAGCGGAACCGCAGCAGGAGTACCTCCTGCGGACCCAGTGAAGCACATGTATGCTCCTAAATAC
 CTGCTCAAGAACTATGACCAAGTTCGGCTACTTGTGGAATCCCTACCGGGAGCCCAAGAGCACTGTAG
 AGGGTGGCCGCCAGAACAGGGCTCAGATGGAAGGCCATCTGCTGTGAGGGGCAAGTTGGATACAGTGGAA
 GGGCTAG

>gi|2117717|pir|JC4362 Beta-1,4-manosil-glucoproteína 4-beta-N-acetilglucosaminiltransferasa (EC
 2.4.1.144) III - Ratón

ERRYKLFLEFCNAGLCCLISFLHFFKTL SYVTFFRELASLSPNLISSFFWNNAPVTPQASPEPGDPDLLET
 PLYSHSPILQPLSPSKATEELHRVDFVLPEDTTEYFVRTKAGGVCFKPGTRMLEKPSPGRTTEKTEVSEG
 SSARGFARRPMRHVLSRERLGRGTRRKWVECVCLPGWKGPSCGVPTVVOYSNLPTEKRLVPREVPRRV
 INAININEEFDLDDVRFHELGDVVDVAVVCDNSNFTAYGEPRLKFRMLNLTNGTFEYIRHKVLYVFLDHFP
 PGGRQDGIADDDYLRFTLTQDGVSRRLRLRPDDVFIIDDADEIPARDQVLPLKLYDGWTEPPAFHMRKSL
 YGFFWKQPGTLEVVSGCTMDMLQAVYGLDGIRLRQRYYTMTNFRQYENRTGHILVQNSLGSPLHFAGWH
 CSWCFTPEGIYFKLVSAQNGDFRWGDIYEDKRDLYIRSLIRTGWFDTGTQOEYPPADPSEHMYAPKYL
 KNYDQFRYLLNENPYREPSTVEGGRQNGSDGRSSAVRGKLDTAEG

FIG. 24

Fig. 25

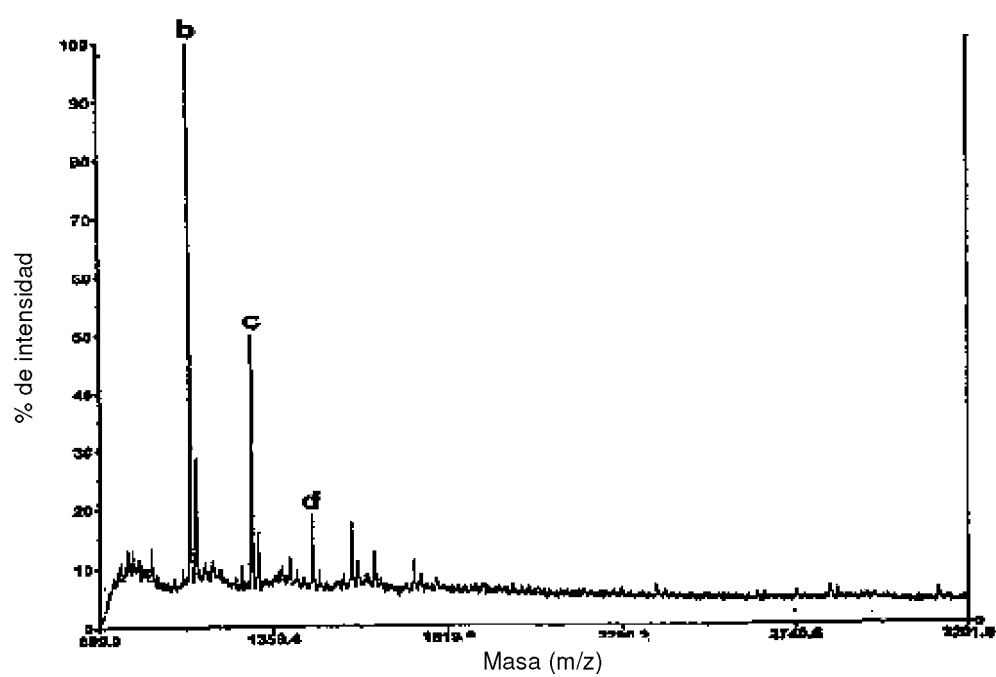
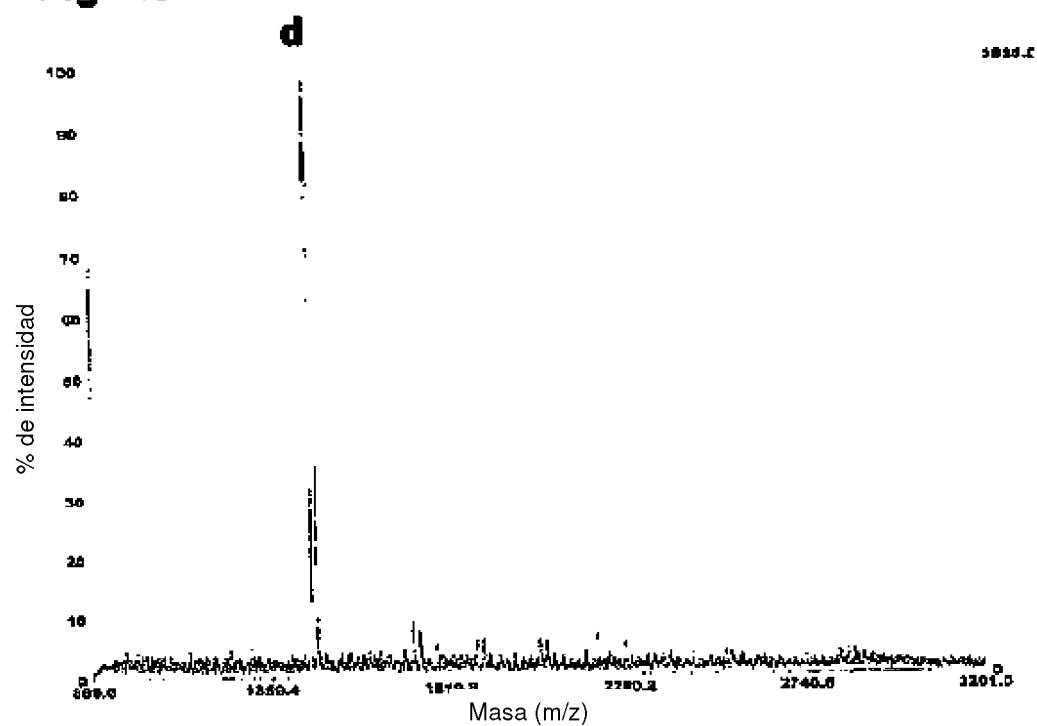


Fig. 26

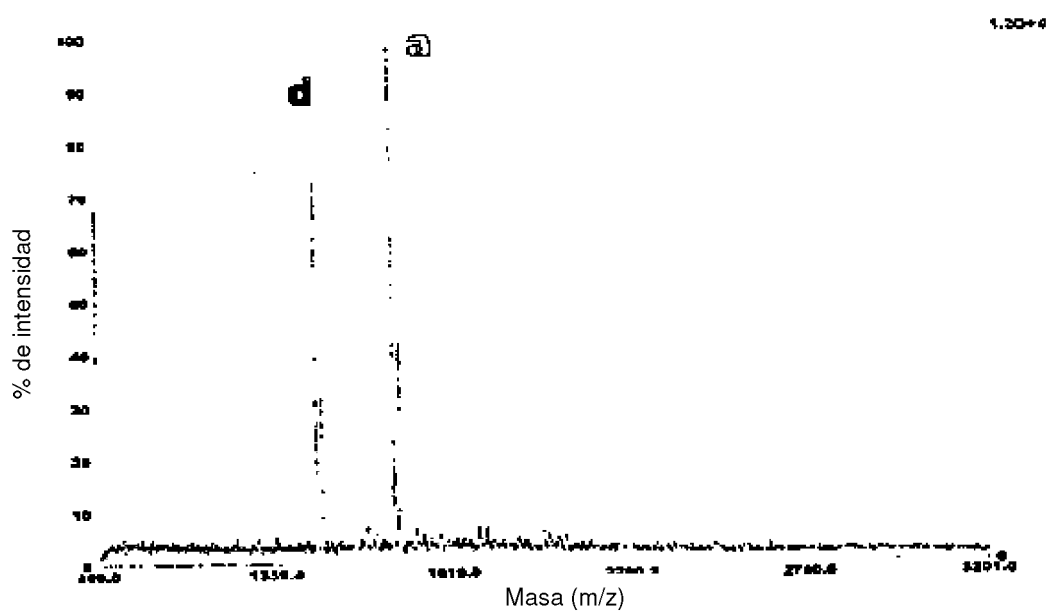
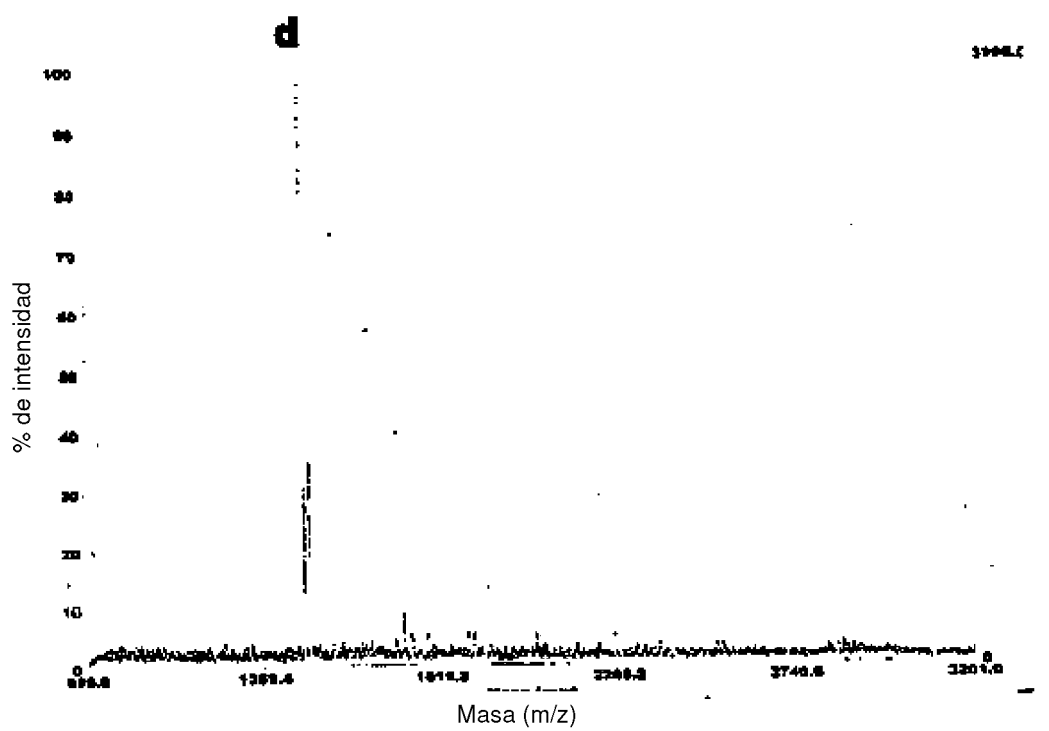


Fig. 27

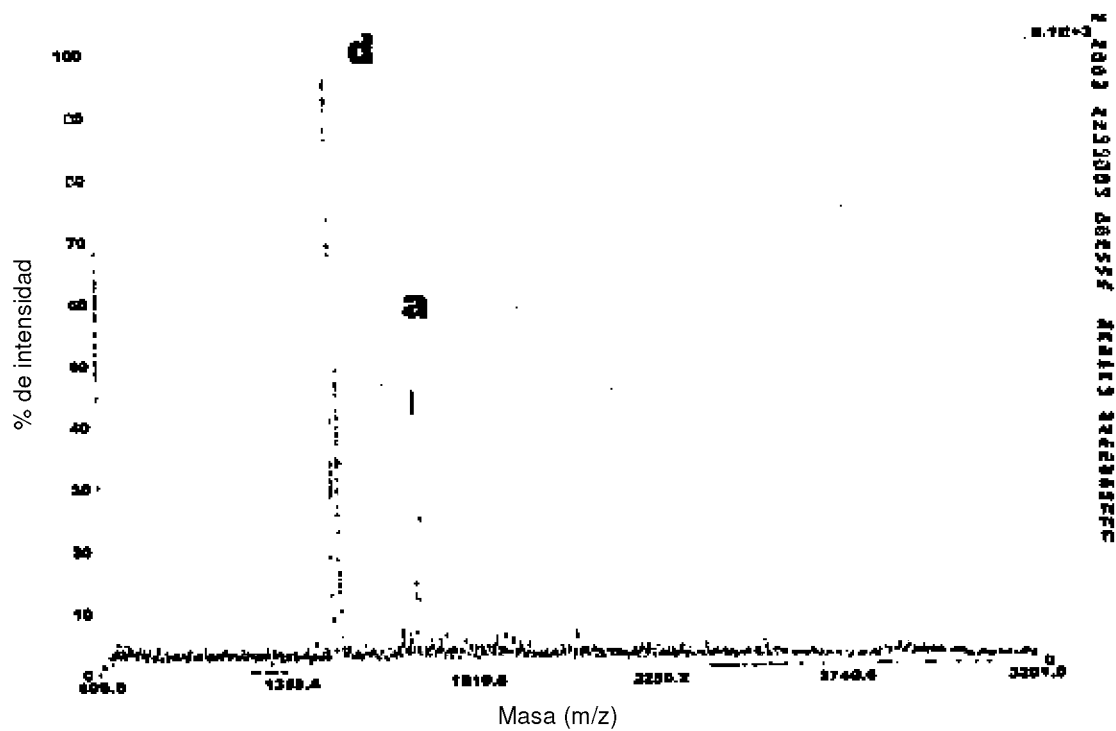
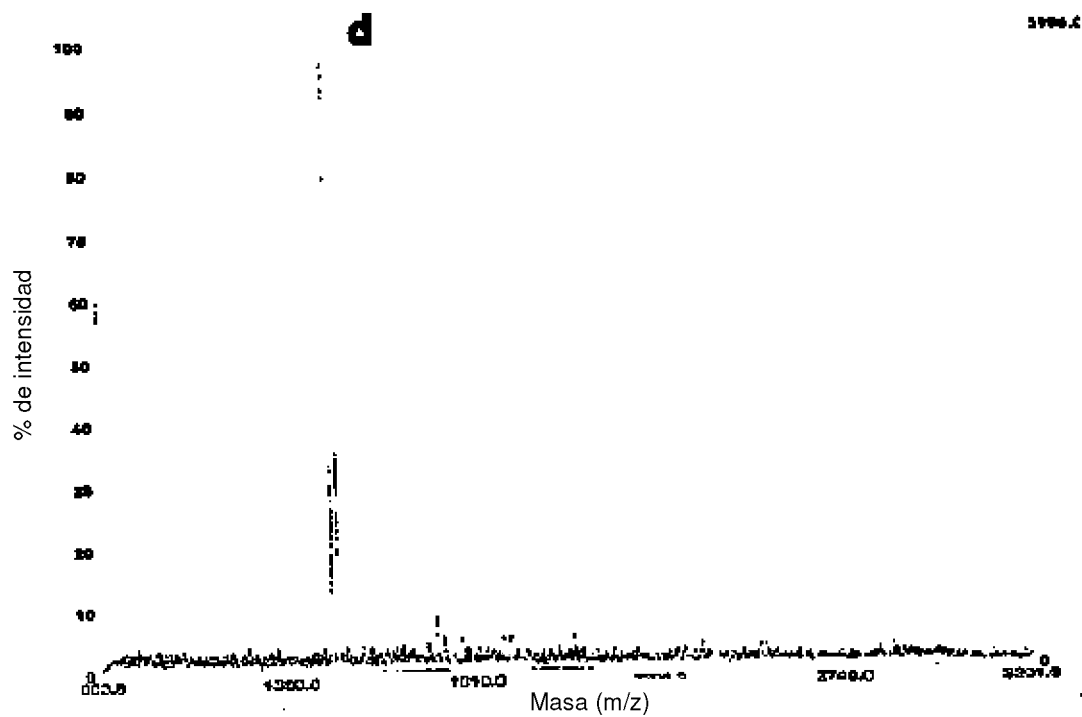


Fig. 28

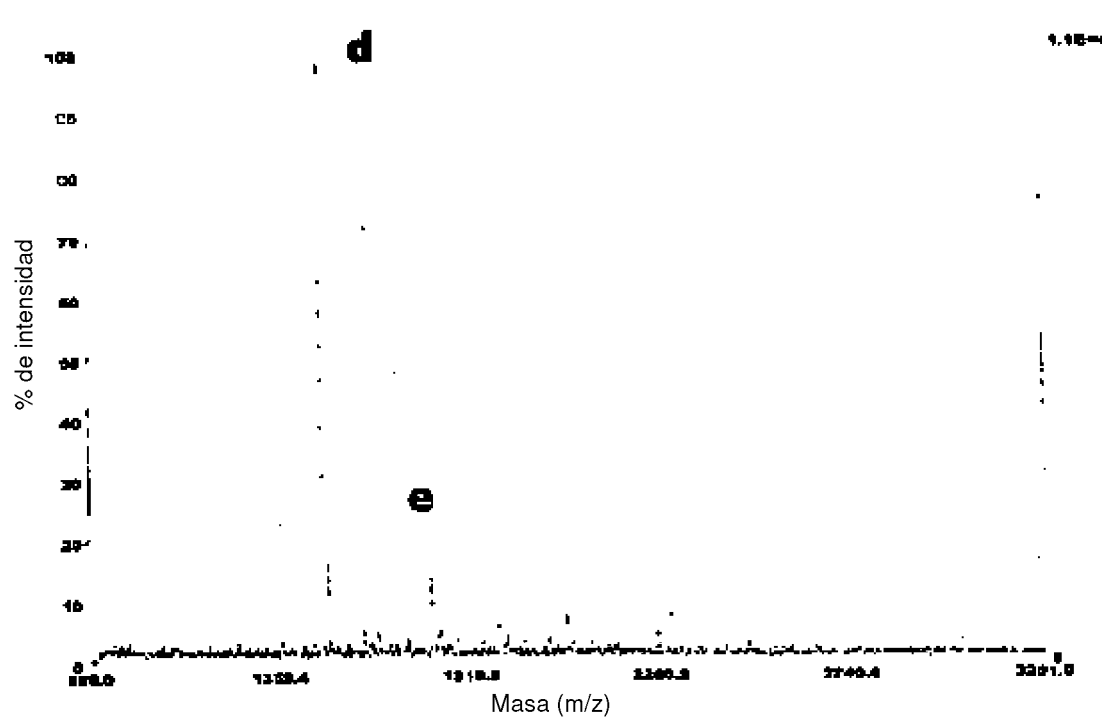
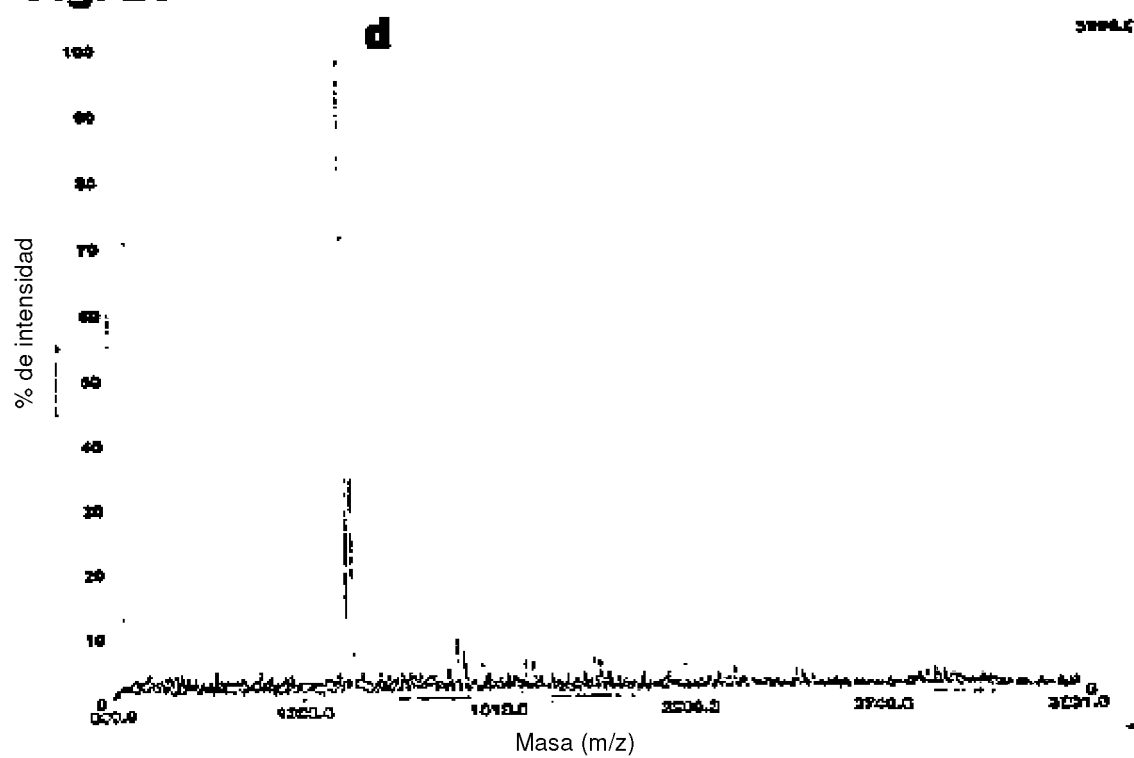


Fig. 29

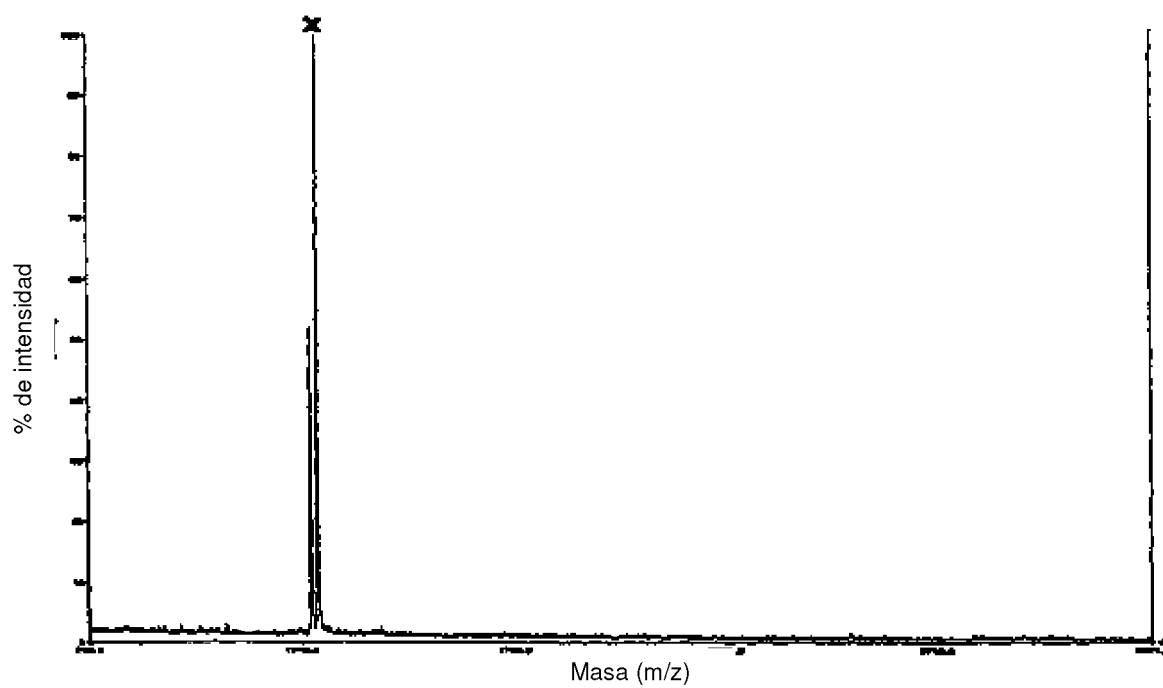


Fig. 30

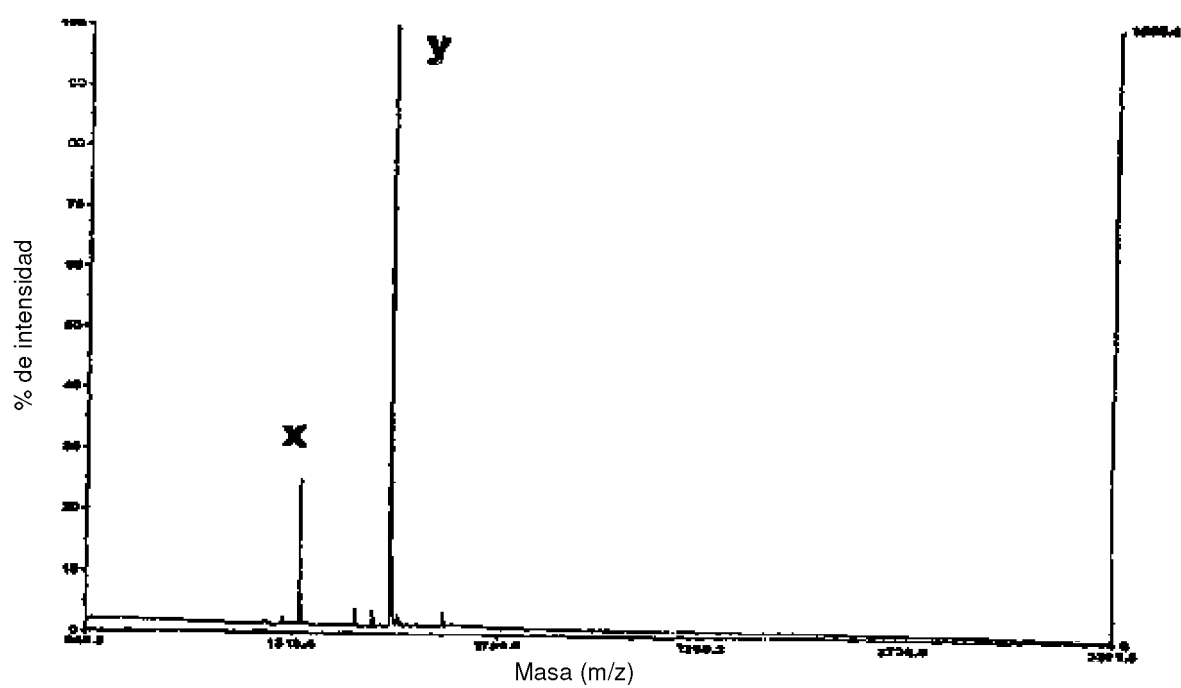


Fig. 31

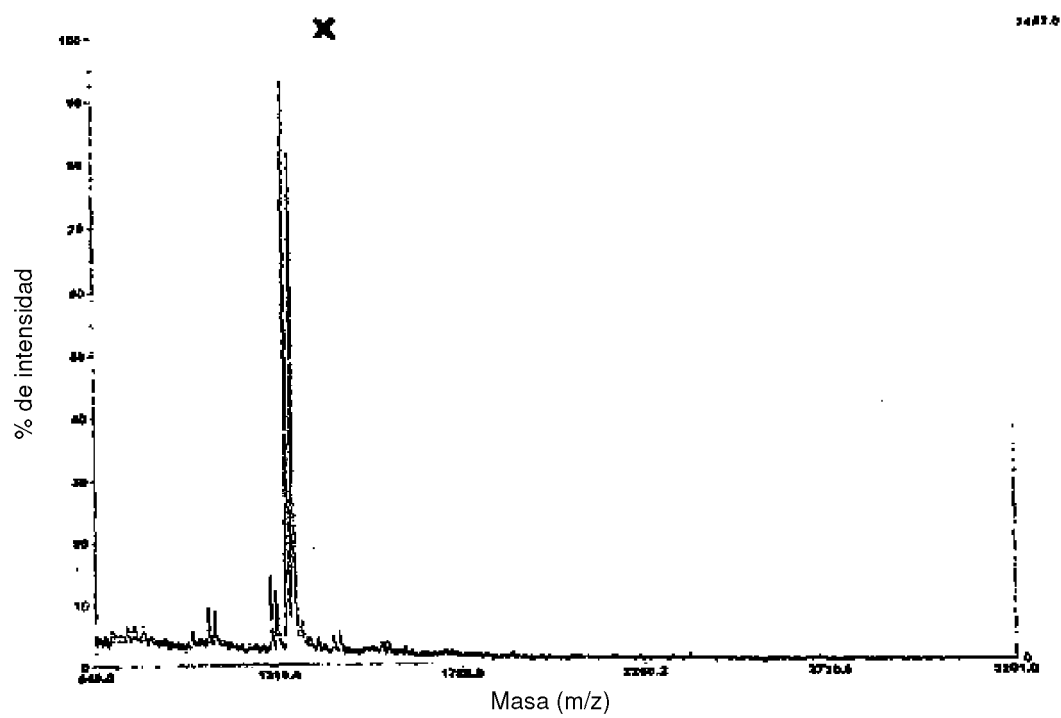
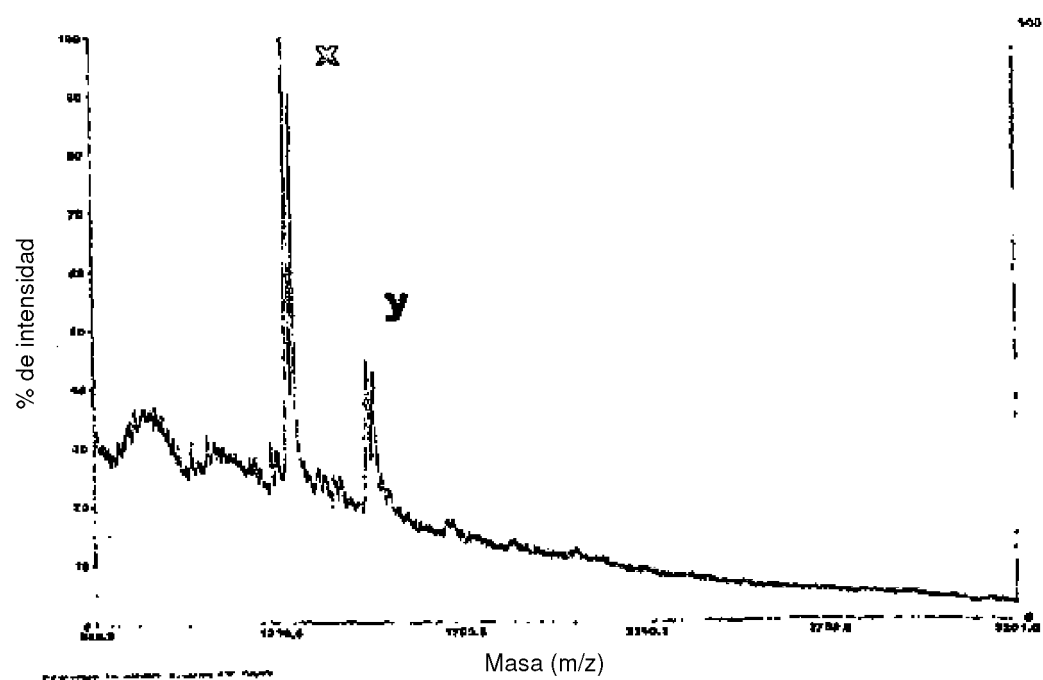


Fig. 32



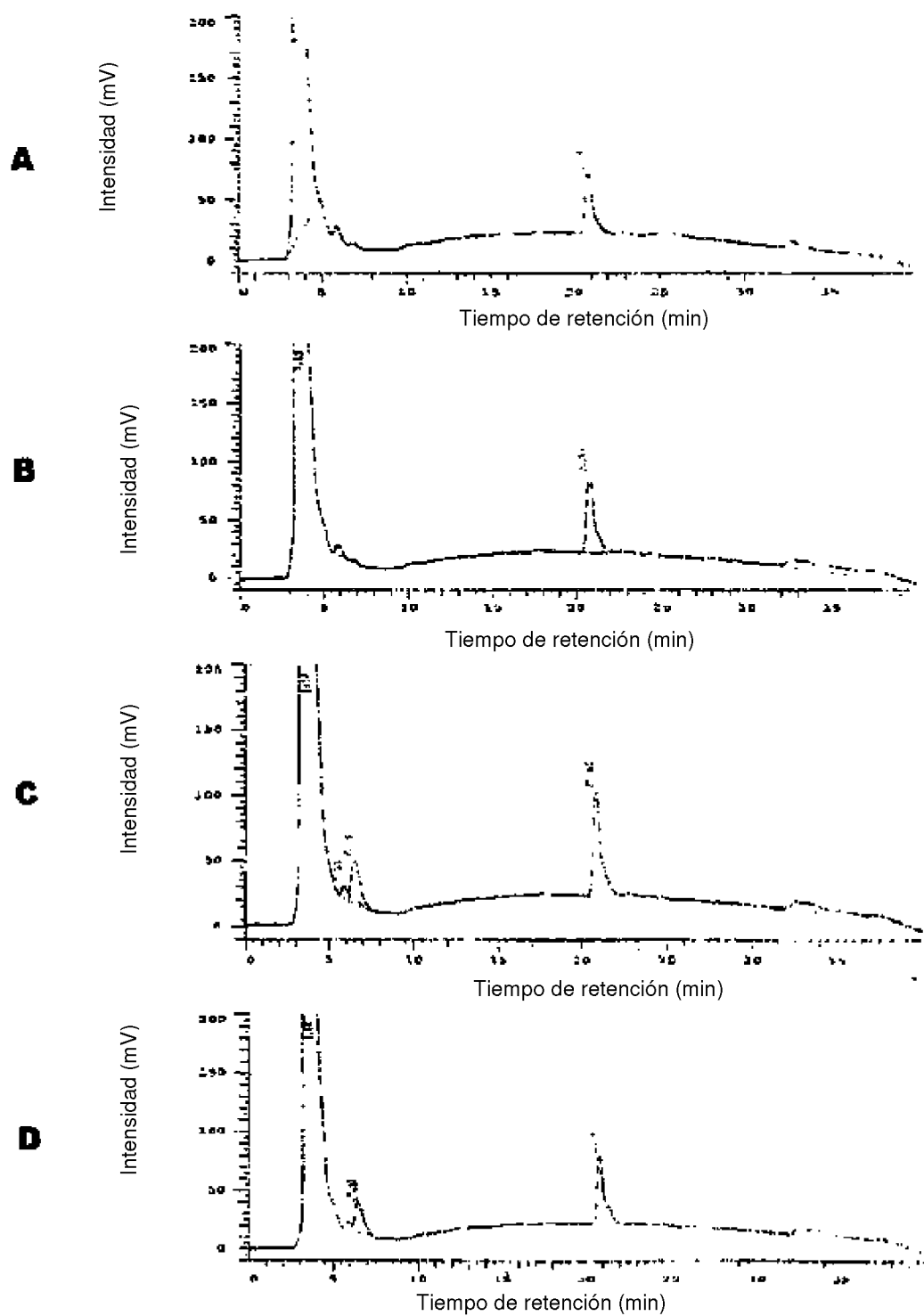


Fig. 33

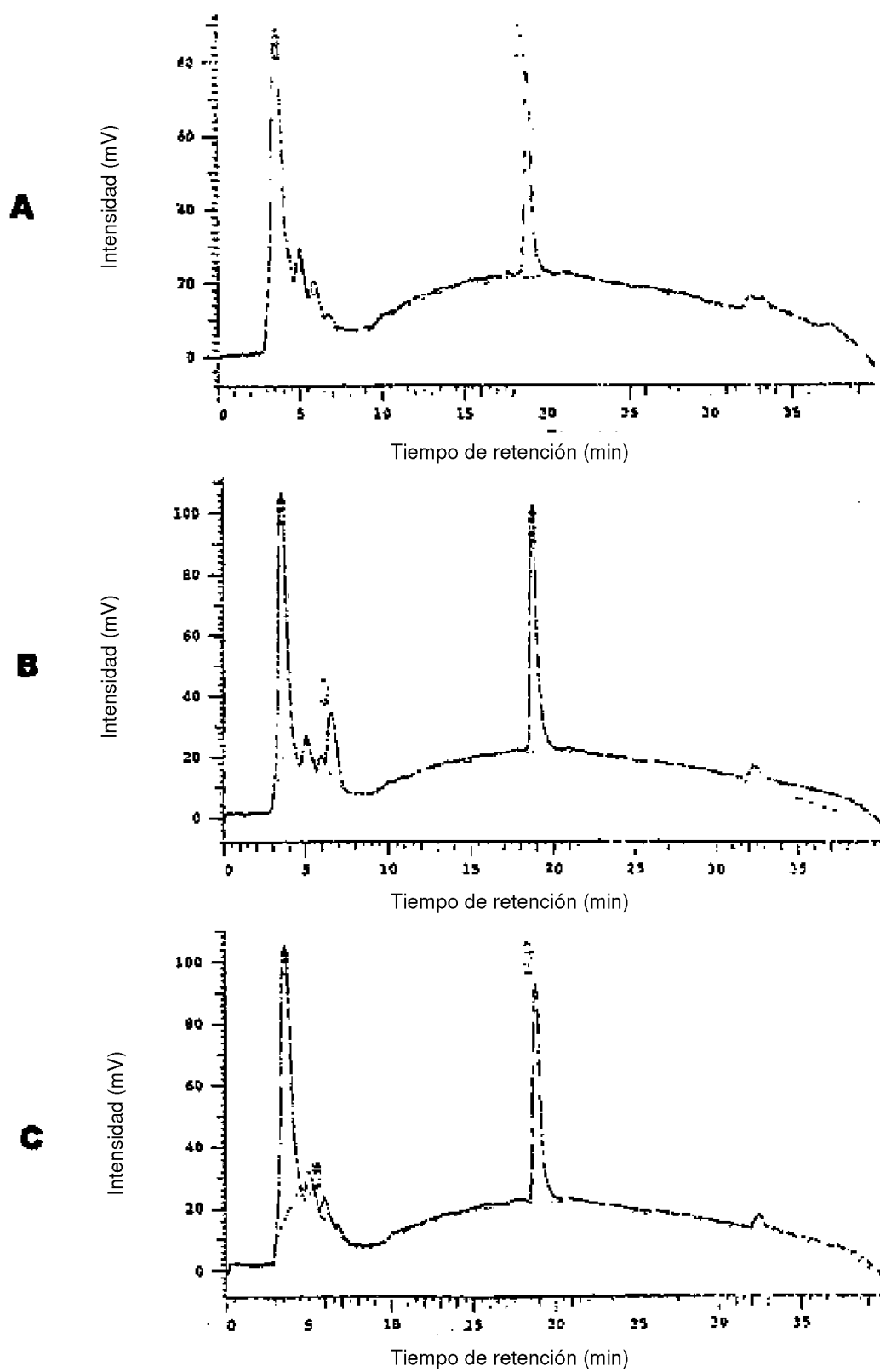


Fig. 34

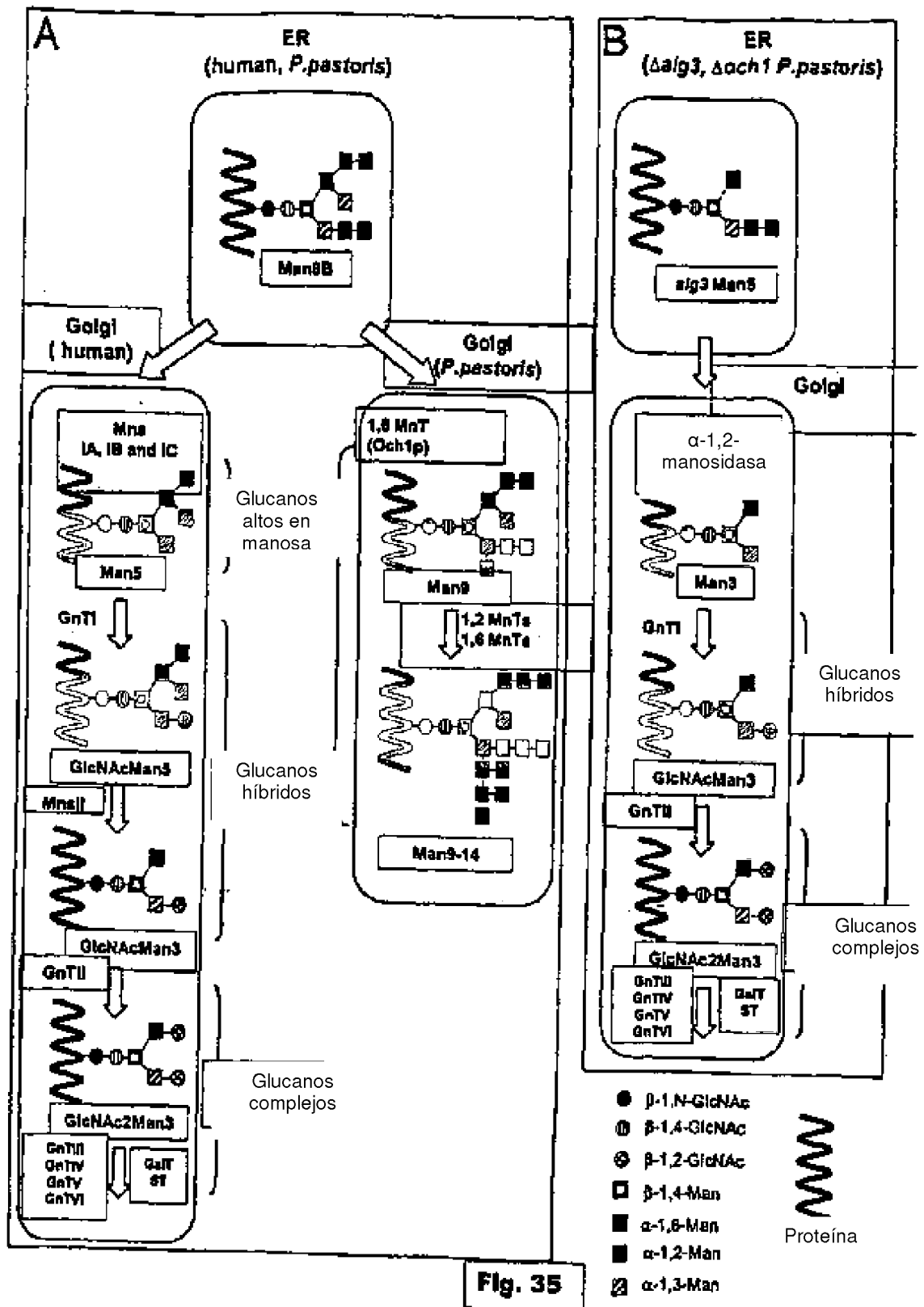


Fig. 36

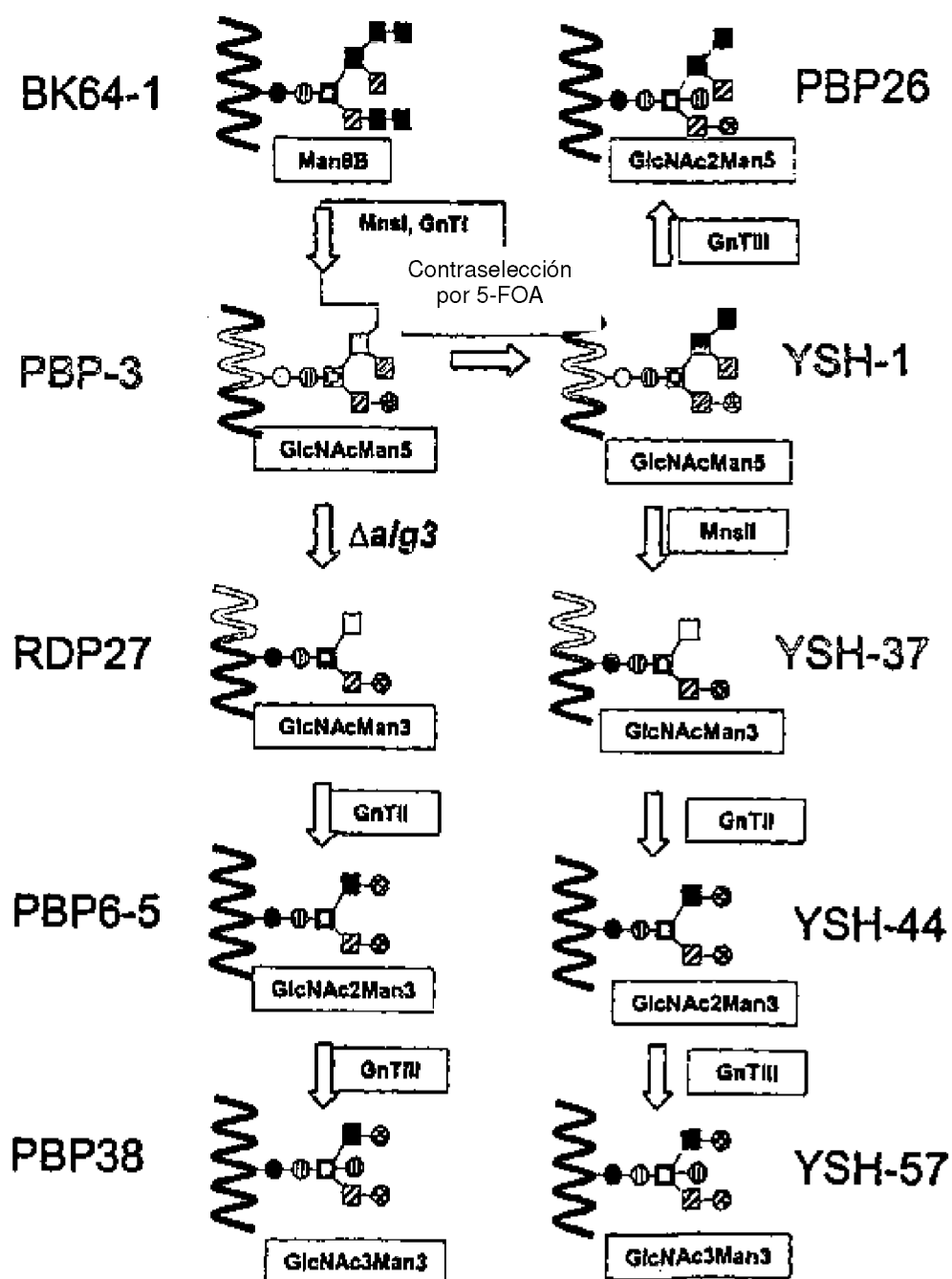


Fig. 37

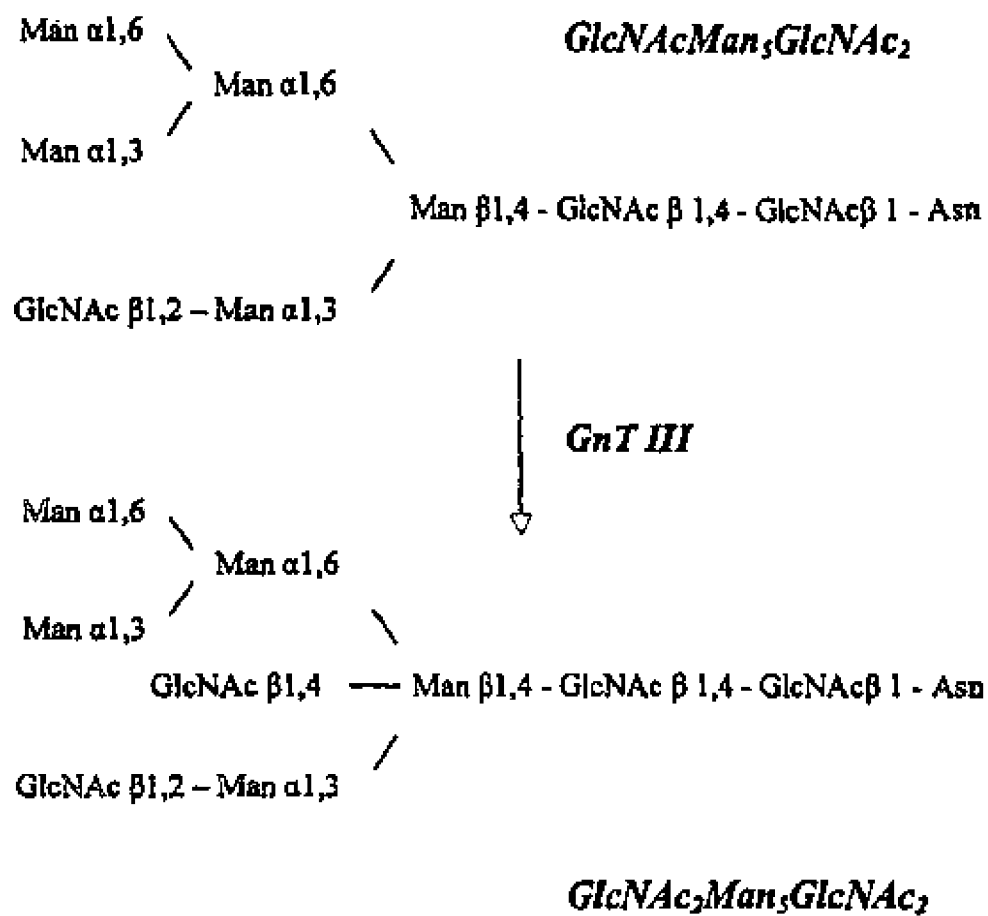
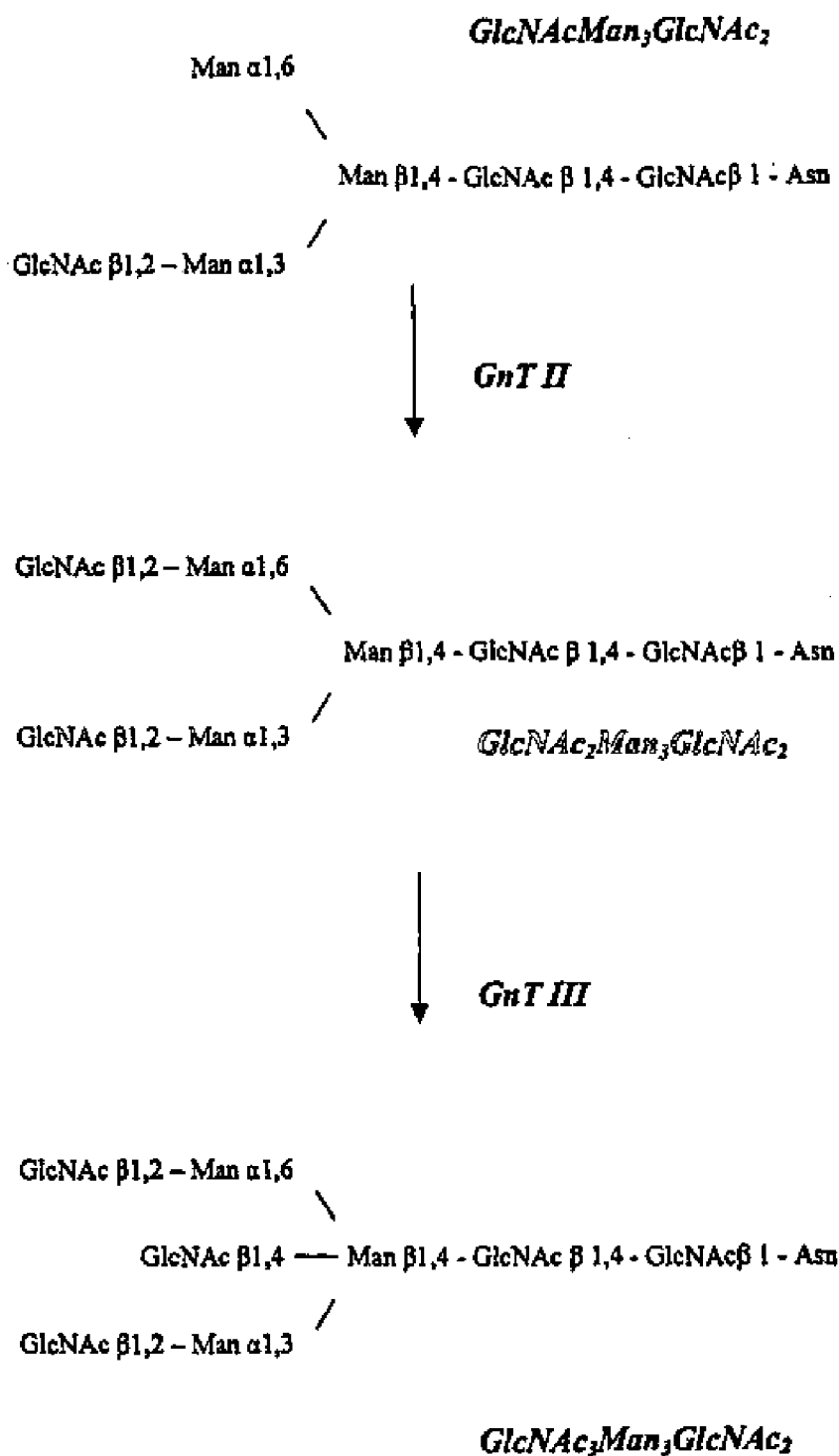


Fig. 38



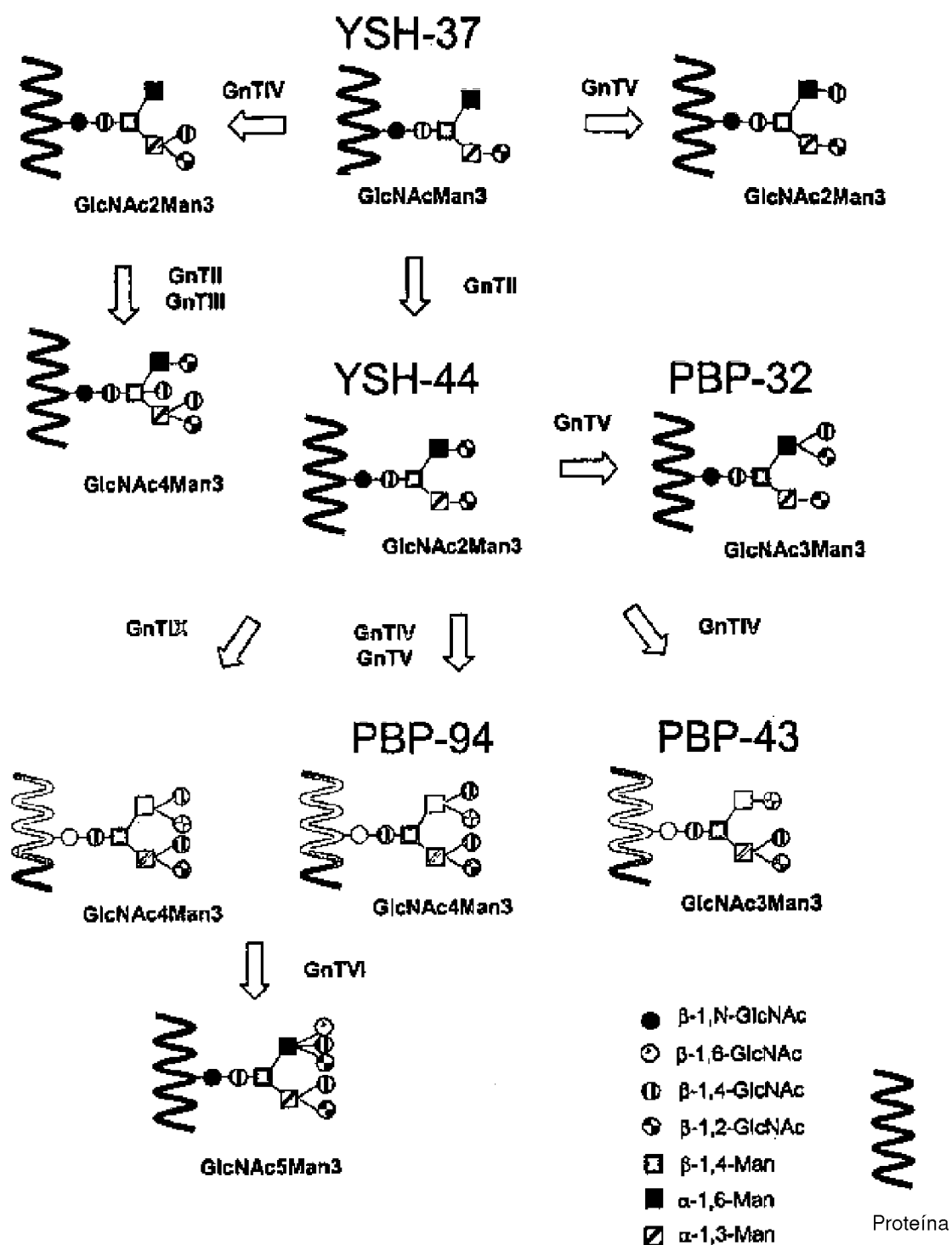


Fig. 39

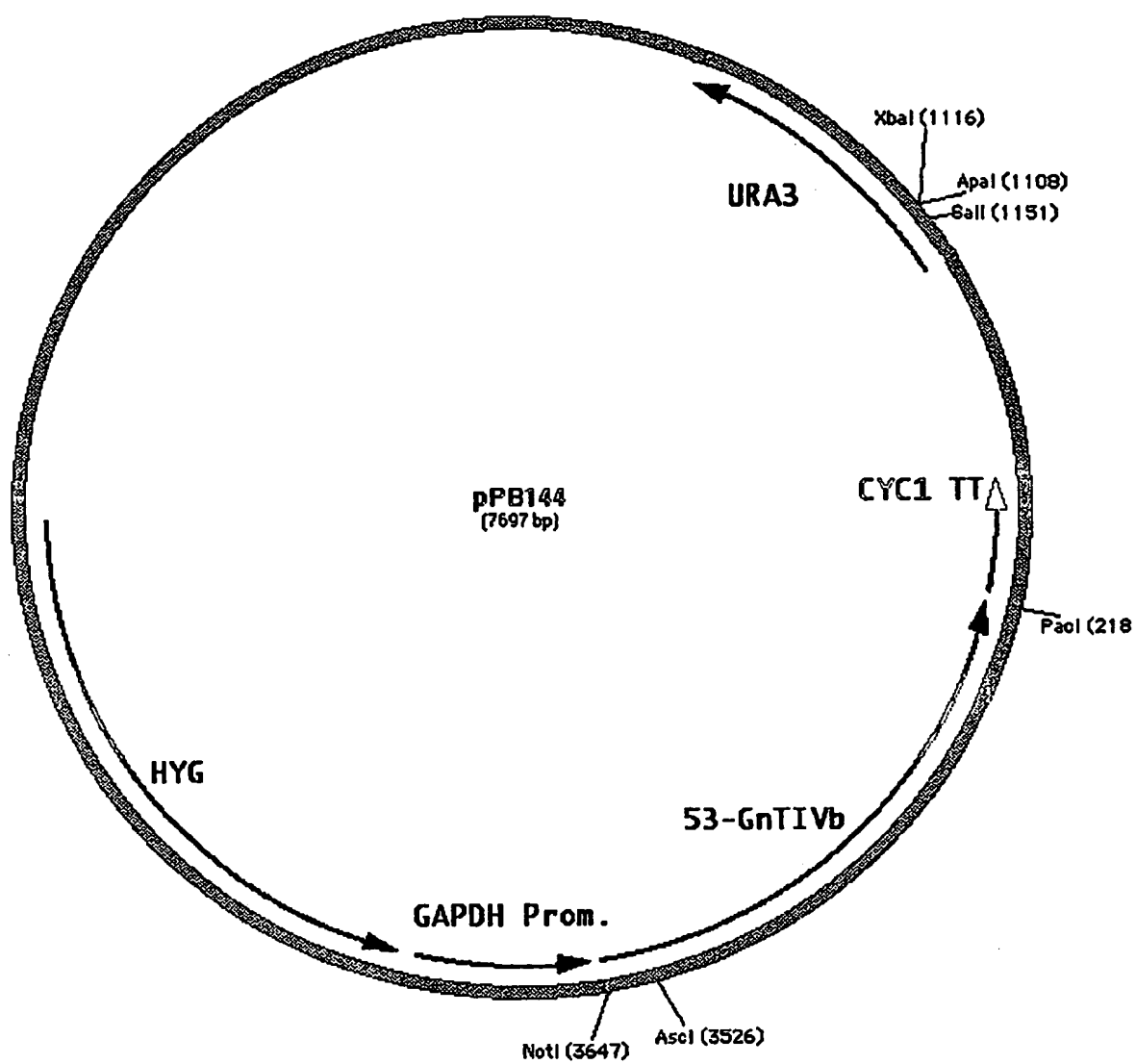


Fig. 40A

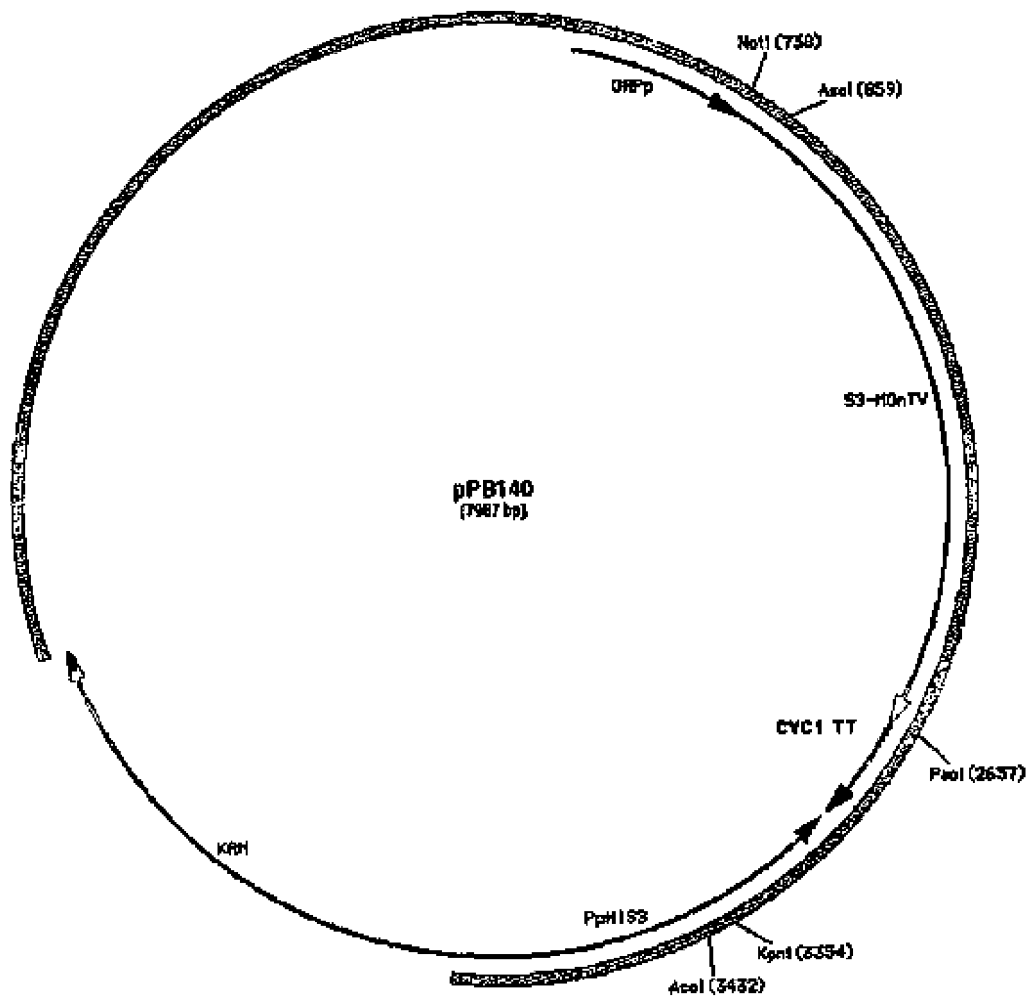


Fig. 40B

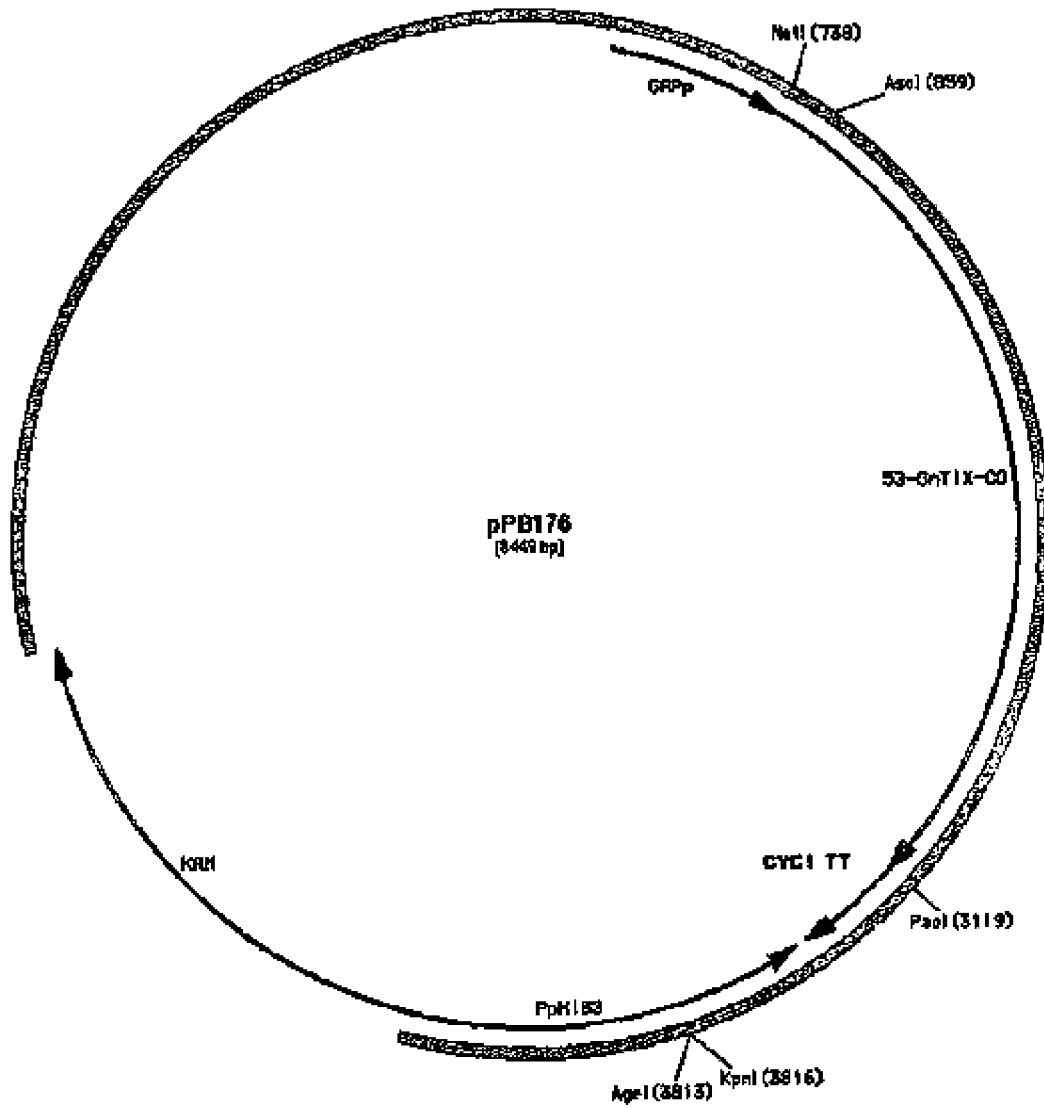


Fig. 40C

Figura 41

Manosil (alfa-1,3)-glucoproteína beta-1,4-N-acetilglucosaminiltransferasa, isoenzima A (MGAT4A) de *Homo sapiens*
 Número de acceso NM_012214

MRLRNGTVATALAFITSEFLTSLWYTTWQNGKEKLIAYQREFLALKERLRIAEHRISQRS
 SELNTIVQQFKRVGAETNGSKDALNKFSDNTLKLKELTSKKSQVPSIYYHLPHELLKN
 EGSLOPAVQIGNGRTGVSIVMGIPTVKREVKSYLIETLHSLIDNLYPEEKLD CVIVVFI
 GETDIDYVHGVVANLEKEFSKEISSGLVEVISPPESYYPDLTNLKETFGDSKERVRWRT
 KQNLDYCFLLMYAQEKGIYYIQLEDDIIVKQNYFNTIKNFALQLSSEBWMILEPSQLGF
 IGKMFQAPDLTLIVEFIFMFYKEKPIDWLLDHILWVKVCNPEKDAKHCDROKANLRIRF
 RPSLFQHVGLHSSLSGKIQKLTDKDYMKPLLLKIHVNPPEAVSTSLKVYQGHTEKTYM
 GEDFFWAIITPIAGDYILFKFDKPVNVESYLFHSGNQEHFGDILLNTTVEVLFPFKSEGLE
 ISKETKDKRLEDGYFRIGKPFENGVAEGMVDPSLNPI SAFRLSVIQNSAVWAILNEIHK
 KATN

Manosil (alfa-1,3)-glucoproteína beta-1,4-N-acetilglucosaminiltransferasa, isoenzima A
 (MGAT4A) de *Homo sapiens*, ARNm

GAAATGAACCTCTCTTATTGATTTTTATTGGCCTAGAGCCAGGAGTACTGCATTTCAGTT
 GACTTTCAGGGTAAAAAGAAAACAGTCCTGGTTGTTGTCATCATAAACATATGGACCAG
 TGTGATGGTGAATGAGATGAGGCTCCGCAATGGAAGTGTAGCCACTGCTTTAGCATTT
 ATCACTTCCCTTCTTACTTTGTCTTGGTATACTACATGGCAAAATGGGAAAGAAAACT
 GATTGCTTATCAACGAGAATTCCTTGCCTTGAAGAAGCTCTTCGAATAGCTGAACACA
 GAATCTCACAGCGCTCTTCTGAATTAAATACGATTGTGCAACAGTTCAAGCGTGTAGGA
 GCAGAAACAAATGGAAGTAAGGATGCGTTGAATAAGTTTTTCAGATAATACCTTAAAGCT
 GTTAAAGGAGTTAAACAAGCAAAAAATCTCTTCAAGTGCCAAGTATTTATTATCATTTCG
 CTCATTTATTGAAAAATGAAGGAAGTCTTCAACCTGCTGTACAGATTGGCAACGGAAGA
 ACAGGAGTTTCAATAGTCATGGGCATTCCACAGTGAAGAGAGAAGTTAAATCTTACCT
 CATAGAACTCTTCATTCCCTTATTGATAACCTGTATCCTGAAGAGAAGTTGGACTGTG
 TTATAGTAGTCTTCATAGGAGAGACAGATAATTGATTATGTACATGGTGTGTAGCCAAC
 CTGGAGAAAGAATTTTCTAAAGAAATCAGTTCTGGCTTGGTGGAGTCAATCACCCCC
 TGAAAGCTATTATCCTGACTTGACAAACCTAAAGGAGACATTTGGAGACTCCAAAGAAA
 GAGTAAGATGGAGAACAAAGCAAAACCTAGATTACTGTTTTCTAATGATGTATGCTCAA
 GAAAAGGGCATATATTACATTTCAGCTTGAAGATGATATTATGTCAAACAAAATTATTT
 TAATACCATAAAAAATTTTGCACTTCAACTTTCTTCTGAGGAATGGATGATTCTAGAGT
 TTTCCAGCTGGGCTTCATTGGTAAATGTTTCAAGCGCCGATCTTACTCTGATTGTA
 GAATTCATATTCATGTTTACAAGGAGAAACCCATTGATTGGCTCCTGGACCATATTC
 TCTGGGTGAAAGTCTGCAACCCCTGAAAAAGATGCAAAACATTGTGATAGACAGAAAGCA
 AATCTGCGAATTCGCTTCAGACCTTCCCTTTTCCAACATGTTGGTCTGCACTCATCACT
 ATCAGGAAAAATCCAAAACTCACGSATAAAGATTATATGAAACCATTACTTCTTAAAA
 TCCATGTAAACCCACCTGCGGAGGTATCTACTTCTTGAAGGTCTACCAAGGGCATACG
 CTGGAGAAAACTTACATGGGAGAGGATTTCTTCTGGGCTATCACACCGATAGCTGGAGA
 CTACATCTTGTTTAAATTTGATAAACCAGTCAATGTAGAAAGTTATTTGTTCCATAGCG
 GCAACCAAGAACATCCTGGAGATATTCTGCTAAACACAACCTGTGGAAGTTTTGCCTTTT
 AAGAGTGAAGGTTTTGCAAATAAGCAAAGAAACCAAGACAAACGATTAGAAGATGCCTA
 TTTCAGAATAGGAAAATTTGAGAATGGTGTTCGAGAAGGAATGGTGGATCCAAGTCTCA

Figura 41 (cont.)

ATCCCATTTTCAGCCTTTTCGACTTTCAGTTATTCAGAATTCTGCTGTTTGGGCCATTCTT
AATGAGATTCATATTAAAAAGCCACCAACTGATCATCTGAGAAACCAACACATTTTTT
CCTGTGAATTTGTTAATTAAAGATAGTTAAGCATGTATCTTTTTTTTATTTCTACTTGA
ACACTACCTCTTGTGAAGTCTACTGTAGATAAGACGATTGTCATTTCCACTTGGAAAGT
GAATCTCCATAATAATTGTATTTGTTTGAAACTAAGCTGTCCTCAGATTTTAACTTGA
CTCAAACATTTTTCAATTATGACAGCCTGTTAATATGACTTGTACTATTTTGGTATTAT
ACTAATACATAAGAGTTGTACATATTGTTACATTCTTTAAATTTGAGAAAACTAATGT
TACATACATTTTATGAAGGGGGTACTTTTGAGGTTCACTTATTTTACTATT

Figura 42

Manosil (alfa-1,3)-glucoproteína beta-1,4-N-acetilglucosaminiltransferasa, isoenzima B (MGAT4B) de *Homo sapiens*

Número de acceso NM_014275

MRLRNGTFLTLLEFCLCAFLSLSWYAALSQKGDVVDVYQRHFLALRDRHLHAAEQESLK
RSKELNLVLDDEIKRAVSEKQALRDGDGNRTWGRLTEDPRLKPWNGSHRHVLHLPTVFHH
LPHLLAKBSSLQPAVRVGGQRTGVSVMGIPSVRREVHSYLTDTLHSLISELSPQEKED
SVIVVLLAETDSQYTSAVTENIKALFPTBIHSGLLLEVISPSPHFYPDFSRLRESFGDPK
ERVRWRWKQNLDYCFLLMYAQSKGIYYVQLEDDIVAKPNYLSTMKNFALQQPSDWMLL
BFSQLGFIGKMFKSLDLSLIVEFILMFYRDKPIDWLLDHILWVKVCNPEKDAKHCDROK
ANLRIRFKPSLFQHVGTSSLAGKIQKLKDKDFGKQALRKEHVNPPAEVSTSLKTYQHF
TLEKAYLREDFFWAFTPAAGDFIRFRFPQPLRLERFFFRSGNIEHPEDKLFNTSVEVLP
FDNPFQSDKEALQEGRTATLRYPRSPDGYLQIGSPYKGVAEGEVDPAFGPLEALRLSIQT
DSPVWVILSEIFLKKAD

Manosil (alfa-1,3)-glucoproteína beta-1,4-N-acetilglucosaminiltransferasa, isoenzima B (MGAT4B) de *Homo sapiens*, ARNm

TGCAGCCTCGGCCCGCGGGCGCCCGCCGCGCACCCGAGGAGATGAGG
CTCCGCAATGGCACCTTCTGACGCTGCTGCTCTTCTGCCTGTGCGCCTT
CCTCTCGCTGTCTGTTACGCGGCACTCAGCGGCCAGAAAGGCGACGTT
GTGGACGTTTACCAGCGGGAGTTCTGCGGCTGCGCGATCGGTTGACAG
CAGCTGAGCAGGAGAGCCTCAAGCGCTCCAAGGAGCTCAACCTGGTGCT
GGACGAGATCAAGAGGGCCGTGTGAGAAAGGCAGGCGCTGCGAGACGG
AGACGGCAATCGCACCTGGGGCCGCCTAACAGAGGACCCCGATTGAAG
CCGTGGAACGGCTCACACCGGCACGTGCTGCACCTGCCACCGTCTTCC
ATCACCTGCCACACCTGCTGGCCAAGGAGAGCAGTCTGCAGCCCGCGGTGCGC
GTGGGCCAGGGCCGACCCGAGTGTGCGGTGGTGATGGGCATCCCGAGCGTGCGGCGCGA
GGTGCACTCGTACCTGACTGACACTCTGCACTCGCTCATCTCCGAGCTGAGCCCGCAGG
AGAAGGAGGACTCGGTTCATCGTGGTGCTGATCGCCGAGACTGACTCACAGTACACTTCG
GCAGTGACAGAGAACATCAAGGCCTTGTTCCCCACGGAGATCCATTCTGGGCTCCTGGA
GGTCATCTCACCTCCCCCACTTCTACCTGACTTCTCCCGCCTCCGAGAGTCTTTTG
GGGACCCCAAGGAGAGAGTCAAGTGAGGACCAACAGAACCTCGATTACTGCTTCCTC
ATGATGTACGCGCAGTCCAAAGGCATCTACTACGTGCAGCTGGAGGATGACATCGTGGC
CAAGCCCAACTACCTGAGCACCATGAAGAACTTTGCACTGCAGCAGCCTTCAGAGGACT
GGATGATCCTGGAGTTCTCCAGCTGGGCTTCATTGGTAAGATGTTCAAGTCGCTGGAC
CTGAGCCTGATTGTAGAGTTTATTCTCATGTTCTACCGGGACAAGCCCATCGACTGGCT
CCTGGACCATATTCTGTGGGTGAAAGTCTGCAACCCCGAGAAGGATGCGAAGCACTGTG
ACCGGCAGAAAGCCAACCTGCGGATCC
GCTTCAAACCGTCCCTCTTCCAGCACGTGGGCACTCACTCCTCGCTGGCTGGCAAGATC
CAGAACTGAAGGACAAAGACTTTGGAAGCAGGCGCTGCGGAAGGAGCATGTGAACCC
GCCAGCAGAGGTGAGCACGAGCCTGAAGACATAACGACACTTCACCTGGAGAAAGCCT
ACCTGCGCGAGGACTTCTTCTGGGCTTCACCCCTGCCGCGGGGACTTCATCCGCTTC

Figura 42 (cont.)

CGCTTCTTCCAACCTCTAAGACTGGAGCGGTTCTTCTTCCGCAGTGGGAACATCGAGCA
 CCCGGAGGACAAGCTCTTCAACACGTCTGTGGAGGTGCTGCCCTTCGACAACCCTCAGT
 CAGACAAGGAGGCCCTGCAGGAGGGCCGCACCGCCACCCTCCGGTACCCTCGGAGCCCC
 GACGGCTACCTCCAGATCGGCTCCTTCTACAAGGGAGTGGCAGAGGGAGAGGTGGACCC
 AGCCTTTCGGCCCTCTGGAAGCACTGCGCCTCTCGATCCAGACGGACTCCCCTGTGTGGG
 TGATTCTGAGCGAGATCTTCCTGAAAAAGGCCGACTAAGCTGCGGGCTTCTGAGGGTAC
 CCTGTGGCCAGCCCTGAAGCCACATTTCTGGGGGTGTGCTCACTGCCGTCCCCGGAGG
 GCCAGATACGGCCCCGCCCAAAGGGTTCTGCCTGGCGTCGGGCTTGGGCCGGCCTGGGG
 TCCGCCGCTGGCCCGGAGGCCCTAGGAGCTGGTGCTGCCCCCGCCCGCCGGGCCGCGGA
 GGAGGCAGGCGGCCCCCCACACTGTGCCTGAGGCCCGGAACCGTTCCGACCCCGGCCTGCC
 CCAGTCAGGCCGTTTTAGAAAGAGCTTTTACTTGGGCGCCCGCCGTCTCTGGCGCGAACA
 CTGGAATGCATATACTACTTTATGTGCTGTGTTTTTTTATTCTTGGATACATTTGATTTT
 TTCACGTAAGTCCACATATACTTCTATAAGAGCGTGACTTGTAATAAAGGGTTAATGAA
 G

Figura 43

N-acetilglucosaminiltransferasa V de *Mus musculus* (Mgat 5)
Número de acceso AF474154

MAFFSPWKLSSQKLGFFLVTFGFIWGMMLLHFTIQQRTQPESSSMLREQI
LDLSKRYIKALAEENRDVVDGPPYAGVMTAYDLKKTAVLLDNILQRIGKL
ESKVDNLVNGTGANSTNSTTAVPSLVSLEKINVADIINGVQEKCVLPMD
GYPHCEGKIKWMKDMWRS DPCYADYGVDGTSCSFFIYLSEVENWCPRLPWRAKNPYEE
ADHNSLABIRTDENILYGMKKHEEPRMRLRIRRMADAWIQAIKSLAEKONLEKRRK
KILVHLGLLTKEGFKIAETAFSGGPLGELVQWSDLITSLYLLGHDIRISASLAELKEI
MKKVGNRSGCPTVGDRIVELIYIDIVGLAQFKKTLGPSWVHYQCMLRVLD SFGTEPEF
NHASYAQSKGHKTPWGKWNLPQQFYTMFPHTPDNSFLGFVVEQHLNSSDIHHINEIKR
QNQSLVYGKVD SFWKNKKIYLDI IHTYMEVHATVYGSSTKNIPSYVKNHGILSGRDLQF
LLRETKLFVGLGFPYEGPAPLEAIANGCAFLNPKFNPPKSSKNTDFFIGKPTLRELT SQ
HPYAEVFIGRPHVWTVDLNNREEVEDAVKAILNQKIEPYMPYEFTCEGMLQRINAFIEK
QDFCHGQVMWPPLSALQVKLAEPGQSCQVCQESQLICEPSFFQHLNKEKDLLKYKVT C
QSSELYKDILVPSFYPKSKHCVFQGDLLLFSCAGAHPTHQRICPCRDFIKGQVALCKDC
L

N-acetilglucosaminiltransferasa V de *Mus musculus* (Mgat 5), ARNm

ATTGCTAGAGAGAGATGGCTTTCTTTTCTCCCTGGAAGTTGTCCTCTCAG
AAGCTGGGCTTTTTCCTGCTGACTTTCGGCTTCATCTGGGGCATGATGCT
TCTGCACTTCACCATCCAGCAGCGGACTCAGCCCGAGAGCAGCTCCATG
TTACGGGAGCAGATCCTTGACCTCAGCAAGAGGTACATTAAGGCACTGG
CAGAGGAGAAACAGGGACGTGGTGGATGGCCCCCTACGCTGGTGT CATGACAGCCTATG
ATCTGAAGAAAACGCTCGCCGCTCTTGCTGGATAACATCCTGCAGCGCATTTGGCAAGCTC
GAGTCAAAGGTGGACAATCTGGTCAACGGCACAGGAGCGAACTCCACCAACTCCACCAC
GGCTGTCCCCAGCTTGGTGTGCTTGAGAAAATTAATGTGGCAGATATCATTAAATGGAG
TTCAGGAAAAATGTGTATTGCCTCCTATGGATGGCTACCCCCACTGCGAGGGGAAAAATC
AAGTGGATGAAGGACATGTGGCGCTCGGACCCCTGCTACGCAGACTATGGAGTGGACGG
GACCTCCTGCTCCTTTTTTATTTACCTCAGTGAGGTTGAAATTGGTGTCTCTGTTTAC
CTTGGAGAGCAAAAAATCCCTATGAAGAAGCTGATCATAACTCATTTGGCGGAAATCCGT
ACGGATTTTAAATTCTCTACGGCATGATGAAGAAGCACGAGGAGTTCCGTTGGATGAG
GCTTCGGATCCGCGCAATGCCTGACGCGTGGATCCAAGCTATCAAGTCTCTGGCGGAGA
AACAAAACCTTGAGAAGAGGAAACGGAAGAAAATCCTTGTTACCTGGGGCTCCTGACC
AAGGAATCGGGCTTCAAGATTGCGGAGACAGCATT CAGCGGTGGCCCTCTGGGTGAACT
CGTTCAGTGGAGTGACTTAATCACATCTCTGTACCTGCTGGGCCATGACATCCGGATCT
CGGCCTCACTGGCTGAGCTCAAGGAGATAATGAAGAAGCTTGTTGGAAACCGGTCTGGC
TGTCCAACCTGTAGGAGACAGAATCGTTGAGCTGATTTATATCGATATTGTGGGACTTGC
TCAATTTAAGAAAACACTAGGGCCATCCTGGGTTTCAATTA
CCAGTGCATGCTCCGGGTGCTAGACTCCTTTGGAACAGAACCTGAGTTCAATCATGCCA
GCTATGCCCAGTCAAAAGGCCACAAGACCCCTGGGGAAAGTGGAATCTGAACCCGCAG
CAGTTTTACACCATGTTCCCTCATACCCAGACAACAGCTTTCTGGGCTTCGTGGTGG A
GCAGCACTGAACCTCCAGCGACATTCACCACATCAACGAGATCAAAGGCAGAACCAAGT
CCCTTGTTGATGGCAAAGTGGATAGTTTCTGGAAGAATAAGAAAATCTACCTGGATATC

Figura 42 (cont.)

ATTCACACGTACATGGAAGTGCACGCCACTGTTTATGGCTCCAGTACCAAGAACATTCC
 CAGTTACGTGAAAAACCATGGCATTCTCAGTGGACGTGACCTGCAGTTTCTTCTCCGGG
 AAACCAAGCTGTTCTGTTGGGCTCGGATTCCCTTATGAAGGCCCAGCTCCCCTGGAGGCC
 ATCGCGAATGGATGTGCTTTCCTGAACCCCAAGTTCAACCCTCCCAAAGCAGCAAAAA
 CACAGACTTCTTCATTGGCAAGCCAACACTGAGAGAGCTGACATCCAGCATCCTTACG
 CAGAAGTCTTCATCGGCCCGGCCACACGTCTGGACTGTGGATCTCAATAACCGAGAGGAA
 GTAGAAGATGCAGTAAAAGCCATCTTAAACCAGAAGATTGAGCCGTATATGCCATATGA
 GTTCACATGTGAAGGCATGCTGCAGAGAATCAACGCTTTCATTGAAAAACAGGACTTCT
 GCCATGGCCAAGTGATGTGCGCGCCCTCAGCGCCCTGCAGGTTAAGCTGGCTGAGCCA
 GGGCAGTCTTGCAAACAGGTGTGCCAGGAGAGCCAGCTCATCTGCGAGCCATCCTTCTT
 TCAACACCTCAACAAGGAAAAGGACCTGCTGAAGTATAAGGTGACCTGCCAAAGCTCAG
 AACTGTACAAGGACATCCTGGTGCCCTCCTTCTACCCCAAGAGCAAGCACTGTGTGTTC
 CAAGGGGACCTCCTGCTCTTCAGTTGTGCCGGAGCCCATCCACACACCAGCGGATCTG
 CCCCTGCCGGGACTTCATCAAGGGCCAAGTGGCCCTCTGCAAAGACTGCCTATAGCAT
 CGCTGCCCTGAATTAACCTCAGACGGGAAAGACGTGGCTCCACTGGGCAGGGCCAAGGGG
 CACAAAGACATTCAGGGACTCTGACCAGAGCCTGAGATCTTTGGTCCAGGGCTTGAGTT
 TAGTACCGCTCCAGCCACAGCCAGTGCATCCAGTTTACACCAAAACCACAAGGGGAACA
 GGTTAGAACAGGAACCTGGGTTCTCCTCAGTGTAAGGAATGTCTCTCTGTCTGGGAGA
 TCGAGCGACTGTAGGGAAAGGATCCAGGCAGTTGCTCCCGGAATTTTTTTTTTTTTTTT
 TTTTAAAGAAGGGATAAAAGTCCGGAGACTCATTCAAACCTGAAAACAAAACAGGAAGA
 GGGAAATTGAGCCAATTGGGAAGGACTTTGGGGCCGATCCTAAACCAATTAATTTATTTA
 TTTGGGAGGATGGGGGCGGGCTCGGGAGGGAGGAGAGGGGTTGAACAGTTTCCTTTTGT
 TCCTCACTGTTAATTCGCCCACCTTCGGGCCCTTCTTGTTCTGCAGCGCCAAGCAGGGT
 GCAGAGGGGCTGTGGCTTGCTTGAGGGGGCCACTGTGGGGCTTCACTCCTGGTCACAGGT
 GGCAGCAGAGAAAAGAGATGTCTATAAGCAGGGGGATGTAGCTCAGTTTGTAGAATGCT
 TGCATAGCATAAATGAAGTCCTGGGTTCATCCCCAGCACCACATAAATGCAGGTAAGA
 AACAGAGTCAGGAGGACCAAGCATTCTCCTTGCTACATAACAAAAGCAAGGCCTTTGT
 CCCCATGTCTTGGCTACAAGAGACCTATCTCAGAAAATTGTGGGGGGGAGGGGGGGGG
 AAATGGCCTTGAAAACACAGCCAGTCACTGTCACTGCATTGCCAGAACTGGTGGATCCC
 AGGTGTGCTTGGCAGATAACAGCTAAAAGGCACATAACCTTGGTGGGGAATAAATGCC
 TGTGGTGTCTGAGGGCCCCACCAAGTTCCAAAAA

Figura 44

N-acetilglucosaminiltransferasa VI de *Gallus gallus*, GnT.VI
Número de acceso AB040608

MRCSPKRSITAVIAASFLLLLLLLLLHGRGSWQDPQEVQFRDLPSDAVLKILKQGSLLHIL
QDNDNLCAHNI SYHLLAGSPLPHKKFLAVGLSSVRRPRGYLPLDTLQSLFKQSSEEL
QEMVVVHLADADPIWNAQVAADISHRFAHHILLGRLVLIHTPHEFYPTLEGLKRNYND
PEERVKFRSKQNVDAFLFTFAANLSSYYLMIEDDVWSAKSFFTAIRKAVASQEGSNWA
TLEFSKLG YIGLYRSSLPLRLARFLLLFYQEMPCDWLLTHFRLLLTOKDVIRFKPSLF
QHMGLYSSFQGTVNRLEDDFQADAMDLPDNPAAALFTNMVVFENYEPSKAYSTARGYF
WGKNPAVGSI PSIVFHQPARVTRVRVQTGSSERPGDFLHAGVLELGRGRRADGRDCSVY
TTVGTFEKGNLEWRGLEKMPNPFVECVRIRVTQSQSEWLI IQSIGIWTAGT

ARNm de Gnt-VI de *Gallus gallus* para *N*-acetilglucosaminiltransferasa VI, dc completas

GGCAGCAGGCAGAGCCAGCAGGGACGCTGCCAGGGCCCAACGCCCCACGAGCCCACAG
CTGCCCACGCACCCCGGCAGACCCAGTTGCATTAATAATGGAAAAATCCCAGGTAAACTC
AGGAGCCTTTGAGCCTCATGCCATGCCTCTGCTGCGAGGAGAAACGTCTCAAGGGCACT
AATGTCTCCATGAAGGGCTCTCAGCAGGCCCGCCTGGTGCTATGCGGTGCTCCCCGAAA
CGCTCCCTCAGCGCTGTGATTGCAGCCTCCTTCCTCCTCCTCCTCCTCCTCCTCCTCCT
GCACAGGGGCAGCTGGCAGGACCCCCAGGAGGTGCAGTTTAGGGATCTACCTTCAGATG
CGGTCTGAAGATACTGAAGCAAGGATCCCTGCACATCCTCCAGGACACGACAACCTC
TGTGCACTCCACAACATCTCCTACCACCTCCTTGCTGGTTCCCCATTACCCACAAAAA
GTTCTTGCGGTGGGGCTGTCATCAGTGCGACGACCAACGTGGATATTACCTCCAGACA
CGCTGCAGTCCCTCTTCAAGCAGTCATCAGAGGAGGAGCTGCAGGAGATGGTGGTGGTG
GTGCACCTGGCAGATGCAGACCCCATCTGGAATGCCAGGTGGCCGCGGACATCAGCCA
TAGGTTGCGTCAACCATCTCCTGGGCGCGCTCGTGCTTATCCATACTCCCATGAGT
TTTACCCAAACCCTGGAAGGCCTCAAGAGAACTACAACGACCCAGAGGAGCGGGTGAAG
TTCAGGTCCAAGCAGAACGTGGATTACGCCTTCCTCTTCACCTTTGCTGCCAACCTTTC
CTCCTACTACTTGATGATTGAGGATGACGTGTGGTCTGCCAAGTCCTTCTTCACTGCCA
TCCGCAAAGCTGTGGCCTCCCAGGAAGGCTCCAACCTGGGCCACCTTGAGTTCTCCAAG
CTGGGCTACATCGGTAAGCTCTACCGCTCCAGTGACCTTCCTCGCTTGGCTCGCTTCCT
CCTCCTCTTCTACCAGGAGATGCCCTGTGACTGGCTGCTGACCCACTTCCGCCTCCTGC
TCACCCAGAAGGATGTGATCCGCTTCAAGCCCTCCCTCTTCCAGCACATGGGCCTCTA
CTCCTCCTTCCAAGGTACCGTCAACCGACTGGAGGATGACGAGTTTCAGGCTGATGCCA
TGGACCTTCCAGACAACCCGCCAGCAGCCCTGTTCAACACATGGTTGTCTTTGAGAAC
TATGAGCCCTCCAAGGCTTACAGCACAGCAAGGGGTATTTCTGGGGGAAAAACCCAGC
AGTTGGCAGCATTTTCTCCATTGTCTTCCACCAACCAGCCCGTGTACCCCGCGTCCGGG
TGCAGACGGGATCCAGTGAGCGCCCTGGGGACTTCCTGCATGCAGGGGTTCTGAGCTG
GGCCGGGGGGCGGCGGGCTGATGGCCGAGACTGCTCTGTGTACACCACTGTGGGCACCTT
TGAGAAAGGGAACCTTAGAGTGGCGGGGGCTGGAGAAGGGAATGCCCAACCTTGIGGAGT
GCGTGAGGATCCGGGTGACCCAGAGCCAGAGTGAAGTGGCTCATCATCCAGAGCATTGGT
ATTTGGAAGTGCAGGCACCTGACCAGGGCTGTGATGGGTCAACCACTGTGGTTGGATTTTG
CTCTAAGAAGAGCTTTATTTTCTCAGTCCCTTTTTTTCGATGGGGAATTAAATTATTCA
GTCAAACCGGCTCCTGCTGTGAACGTAGAGGGGTGGCAGGGCAGCTGCGGGGTCTGCT
TCCTGCACGGAGGTGGACGGGGTTGGCTGTAGGGCCCACTGTGCTGCACCAGACTGGGG
GATGCTGCAGAAAGCAGTGCCAGCCCCAGGCTGCAGCCCTACGGCCCATCAGTATGGG

Figura 44 (cont.)

GAAAGTGATGGACAGGCAGCTCTGCATACGCTTTGTGTCCTGATGGAGTGCCAGTTTTC
GTGCTCCAAGCAGAGTCCTGCTTCCTTTGTACCCAGTGCCCTTCTTGATGCTTCCTTA
TGCCCTGACTCAGCTAATTAATTAAAAACGCTGAGTCAATT

Figura 45

N-acetilglucosaminiltransferasa IX GnT-IX de *Homo sapiens* (GnT IX)

MITVNPDGKIMVRRCLVTLRPFRFLFVLGIGFFTLCEFLMTSLGGQFSARRLGDSFPTIRT
 EVMGGPESRGVLRKMSDLLELMVKRMDALARLENSSELHRAGGDLHFPADRMPPGAGLM
 ERIQAIQNVSDIAVKVDQILRHSLLLHSKVSEGRDQCEAPSDPKFPDCSGKVEWMRA
 RWTSDPCYAFFGVGTECSFLIYLSEVEWFCPLPWRNQTAAQRAPKPLPKVQAVFRSN
 LSHLLDLMSGKESLIFMKKRTKRITAQWALAAQRLAQKLGATORDOKQILVHIGFLTE
 ESGDVFSPRVLKGGPLGEMVQWADILTALYVLGHGLRVTVSLKELOSNLGVPPGRGSCP
 LTMPLPFDLIYTDYHGLQOMKRHMGLSFKKYRCRIRVIDTFGTEPAYNHBEYATLHGYR
 TNWGYWNLNPKQFMIMFPHTPDNSFMGFVSEELNETEKRLIKGGKASNMVAVYGKEASI
 WKLQKKEKFLGILNKYMEIHGTVVYESQRPPEVPAFVNHLGLLPQPEFQQLLRKAKLFI
 GFGFPYEGPAPLEAIANGCIFLQSRFSPHSSLNHEFFRGKPTSREVFSSQHPYAENFIG
 KPHVWTVDYNNSEEFEEAAIKAIMRTQVDPYLPYEYTCGMLERIHAYIQHDFCRAPDP
 ALPEAHAPQSPFVLAPNATHLEWARNTSLAPGAWPPAHALRAWLAVPGRACDTDTCLDHG
 LICEPSFFPFLNSQDAFLKLQVPCDSTESEMNHLYPAFAQPGQECYLQKEPLLFSAGS
 NTKYRRLCPCRDFRKGQVALCQGCL

ARNm de GnT-IX de *Homo sapiens* para *N*-acetilglucosaminiltransferasa, cd completas;
 Número de acceso AB109185.1

CTGCTCGCACCAACAAGTTTGAACAATGATCACCGTCAACCCCGATGGGAAG
 ATAATGGTCAGAAGATGCCTGGTCACCGTGAGACCCCTTTCGGCTTTTGTCTCT
 GGGCATCGGCTTCTTCACTCTCTGCTTCCTGATGACGTCTCTGGGAGGCCAGT
 ICTCGGCCCGGCGCCTGGGGGACTCGCCATTCACCATCCGCACAGAAAGTGAT
 GGGGGGCCCCGAGTCCCGCGGCGTCTGCGCAAGATGAGCGACCTGCTGGAG
 CTGATGGTGAAGCGCATGGACGCACTGGCCAGGCTGGAGAACAGCAGTGAG
 CTGCACCGGGCCGGCGGCGACCTGCACTTCCCGCAGACAGGATGCCCCCTG
 GGGCCGGCCTCATGGAGCGGATCCAGGCTATTGCCCAGAACGTCTCCGACAT
 CGCTGTGAAGGTGGACCAGATCCTGCGCCACAGTCTGCTCCTGCACAGCAAG
 GTGTCAAGAGCCGGCGGGACCAGTGTGAGGCACCCAGTGACCCCAAGTTCC
 CTGACTGCTCAGGGAAGGTGGAGTGGATGCGTGCCCGCTGGACCTCTGACCC
 CTGCTACGCCTTCTTTGGGGTGGACGGCACCGAGTGCTCCTTCTCATCTACC
 TCAGTGAGGTGAGTGTTCTGCCCCCGCTGCCCTGGAGGAACCAGACGGC
 TGCCCAAGAGGGCACCAAGCCCTCCCAAGTCCAGGCAGTTTTCCGAAGC
 AACCTGTCCACCTTCTGGACCTGATGGGCAGCGGGAAGG

AGTCCCTGATCTTCATGAAGAAGCGGACCAAGAGGCTCACAGCCAGTGGGCGCT
 GGCTGCCAGCGCCTGGCACAGAAGCTGGGGGCCACCCAGAGGGACCAGAAGCAG
 ATCCTGGTCCACATCGGCTTCTGACGGAGGAGTCCGGGGACGTGTTTCAGCCCTC
 GGGTCTGAAGGGCGGGCCCTAGGGGAGATGGTGAGTGCGGCGACATTCTGAC
 TGCACCTCTATGTCTTGGGCCATGGCCTGCGGGTCACAGTCTCCCTGAAGGAGCTG
 CAGAGTAACCTTAGGGGTACCGCCAGGCGGGGAAGCTGCCCGCTCACCATGCCCC
 TGCCCTTCGACCTCATCTACACCGACTACCACGGCCTGCAGCAGATGAAGCGGCA
 CATGGGACTCTCCTTCAAGAAGTACCGGTGCCGAATCAGGGTCATCGACACCTTC

Figura 45 (cont.)

GGGACGGAACTGCGTACAACCACGAGGAGTACGCCACGCTGCACGGCTACCGGA
 CCAACTGGGGCTACTGGAACCTCAACCCCAAGCAGTTCATGACCATGTTTCCTCA
 TACCCCCGACAACTCCTTCATGGGCTTCGTGTCCGAGGAGCTCAACGAGACGGAG
 AAGCGGCTCATCAAAGGCGGCAAGGCCAGCAACATGGCCGTGGTGTACGGCAAGG
 AGGCGAGCATCTGGAAGCTCCAGGGGAAGGAGAAGTTCCTGGGCATCCTGAACAA
 ATACATGGAGATCCATGGCACCGTGTACTACGAGAGCCAGCGGCCCCCGAGGTG
 CCAGCCTTTGTGAAGAACCACGGCCTCTTACCGCAGCCTGAGTTTCAGCAGCTGC
 TGC GCAAGGCCAAACTCTTCATCGGGTTTGGCTTCCCCTACGAGGGCCCCGCCCC
 CCTGGAGGCCATCGCCAATGGTTGCATCTTCCTGCAGTCCCGCTTCAGCCCGCCC
 CACAGCTCCCTCAACCACGAGTTCCTCCGAGGCAAGCCCACCTCCAGAGAGGTGT
 TCTCCAGCATCCCTACGCGGAGAACTTCATCGGCAAGCCCCAAGTGTGGACAGT
 CGACTACAACAACCTCAGAGGAGTTTGAAGCAGCCATCAAGGCCATTATGAGAACT
 CAGGTAGACCCCTACCTACCCTATGAGTACACCTGCGAGGGGATGCTGGAGCGGA
 TCCACGCCTACATCCAGCACCCAGGACTTCTGCAGAGCTCCAGACCCTGCCCTACC
 AGAGGGCCACGCCCCGCAGAGCCCCTTTGTCTCTGGCCCCCAATGCCACCCACCTC
 GAGTGGGCTCGGAACACCAGCTTGGCTCCTGGGGCCTGGCCCCCGCGCACGCCC
 TGCGGGCCTGGCTGGCCGTGCCTGGGAGGGCCTGCACCGACACCTGCCTGGACCA
 CGGGCTAATCTGTGAGCCCTCCTTCTTCCCCTTCTGAAACAGCCAGGACGCCTTC
 CTCAAGCTGCAGGTGCCCTGTGACAGCACCGAGTCGGAGATGAACCACCTGTACC
 CGGCGTTTCGCCAGCCTGGCCAGGAGTGCTACCTGCAGAAGGAGCCTCTGCTCTT
 CAGCTGCGCCGGCTCCAACACCAAGTACCGCCGGCTCTGCCCCCTGCCGCGACTTC
 CGCAAGGGCCAGGTGGCCTTGTGCCAGGGCTGTCTGTGAATCCGCCTCTGCCGCC
 CTGCCTGGCACCCACGCTGGCTCTCTCCTGCCGCGGGAGAAAGCACCAGCAGGTT
 C

Figura 46

ggcgsgccCAGTTCTCTGCTAGAAGATTGGGTGACTCTCCATTCACTATTAGAACCGAGGTTATGGGTGGA
 C
 CAGAGTCC
 AGAGGTGTCTTGAGAAAGATGTCTGACTTGTGGAGCTGATGGTTAAGAGAATGGACGCTTTGGCTAGATT
 G
 GAGAACTC
 CTCTGAGTTGCACAGAGCTGGTGGTGACTTACACTTTCTGCTGACAGAATGCCACCAGGTGCAGGTTTGA
 T
 GGAGAGAA
 TCCAAGCTATCGCACAGAACGTTTCTGACATTGCTGTTAAGGTTGACCAAATCTTGAGACACTCTTTGCTG
 T
 TGCACICT
 AAGGTTTTCTGAGGTTAGAAGAGATCAATGTGAGGCTCCTTCTGATCCAAAGTTCCTGACTGTTCTGTGTA
 G
 GTTGAATG
 GATGAGAGCAAGATGGACTTCTGACCCATGTTACGCTTTCTTCGGTGTGACGGTACTGAGTGTCTTTCT
 T
 GATTTACT
 TGTCTGAAGTTGAGTGGTTCTGTCCACCTTTGCCTTGGAGAAACCAAACCGCTGCTCAAAGAGCACCAAAG
 C
 CTTTGCCA
 AAGGTTCAAGCTGTCTTCAGATCCAACCTGTCTCACTTGTGGACTTGATGGGTTCTGGTAAGGAGTCTTT
 G
 ATCTTCAT
 GAAGAAGAGAACTAAGAGATTGACTGCTCAATGGGCTTTGGCTGCACAGAGATTGGCACAGAAGTTGGGTG
 C
 TACTCAA
 GAGACCAGAAGCAAATCTTGGTTTACATCGGTTTCTTGACTGAGGAGTCTGGTGACGTCTTCTCTCCAAGA
 G
 TCTTGAAG
 GGTGCTCCATTGGGAGAGATGGTCCAATGGGCTGACATCTTGACTGCTTTGTACGTCTTGGGTACCGTTT
 G
 AGAGTTAC
 TGTTTCTTTGAAGGAGTTGCAATCTAACTTGGGTGTTCCCTCCAGGTAGAGGTTCTGTCCACTTACTATGC
 C
 TTTGCCAT
 TCGACTTGATTTACACTGATTACCACGGTTTGCAACAGATGAAGAGACACATGGGTTTATCTTTCAAGAAG
 T
 ACAGATGT
 AGAATTAGAGTCATTGACACTTTCCGTACTGAACCTGCTTACAACCCAGAGGAGTACGCTACTTTGCACGG
 T
 TACAGAAC
 TAAGTGGGGATACTGGAACTTGAACCCAAAGCAATTCATGACTATGTTCCCTCACACCCAGACAACTCTT
 T
 CATGGGT
 TCGTTTCTGAGGAGTTGAACGAGACTGAGAAGAGATTGATTAAAGGTGGAAGGCTTCTAACATGGCTGTT
 G
 TCTACGGT
 AAGGAGGCTTCTATTTGGAAGTTGCAAGGTAAAGGAGAAGTTCTTGGGAATCTTGAACAAGTACATGGAGAT
 T
 CACGGTAC
 TGTTTACTACGAGTCTCAAAGACCACCTGAAGTTCCAGCATTTGTCAAGAACCACGGTTTGTGCCACAAC
 C

Figura 46 (Cont.)

TGAGTTCC
 AACAGTTGTTAAGAAAGGCTAAGTTGTTTCATCGGTTTCGGTTTCCCATACGAGGGACCTGCTCCTTTGGAG
 G
 CTATTGCT
 AATGGTTGTATCTTCTTGCACTCCAGATTCTCTCCACCTCACTCTTCTTTGAACCACGAGTTCTTCAGAGG
 T
 AAGCCAAC
 TTCCAGAGAGGTCTTCTCTCAACACCCATACGCTGAGAACTTCATTGGTAAGCCACACCTTTGGACTGTTG
 A
 CTACAACA
 ACTCTGAGGAGTTTGAAGCTGCTATTAAAGGCTATTATGAGAACTCAAGTTGACCCATACTTGCCCTTACGAG
 T
 AACTTTGT
 GAGGGTATGTTGGAGAGAATCCACGCATACATTC AACACCAAGACTTCTGTAGAGCACCAGACCCTGCATT
 G
 CCAGAGGC
 TCACGCTCCACAATCTCCATTGCTCTTGGCTCCTAATGCCAACTCACTTGGAGTGGGCTAGAAACACTTCTT
 T
 GGCTCCTG
 GTGCATGGCCACCTGCTCAGCTTTGAGAGCTTGGTTGGCTGTTCCCTGGTAGAGCTTGTACTGACACTTGT
 T
 TGGACCAC
 GGTTTGATTGTGAGCCATCTTTCTTCCCATTCCTTGAACCTCCAAGACGCTTTCTTGAAGTTGCAAGTTCC
 A
 TGTGACTC
 TACTGAGTCTGAGATGAACCACTTGTACCCAGCTTTCGCACAACCTGGTCAAGAGTGTACTTGCAGAAGG
 A
 GCCTTTGT
 TGTTCTCTTGTGCTGCTTCTAACACCAAGTACAGAAGATTGTGTCCATGTAGAGACTTCAGAAAGGGTCAA
 G
 TCGCTTTG
 TGTCAAGGTTGTTTGTAAGcttaattaa

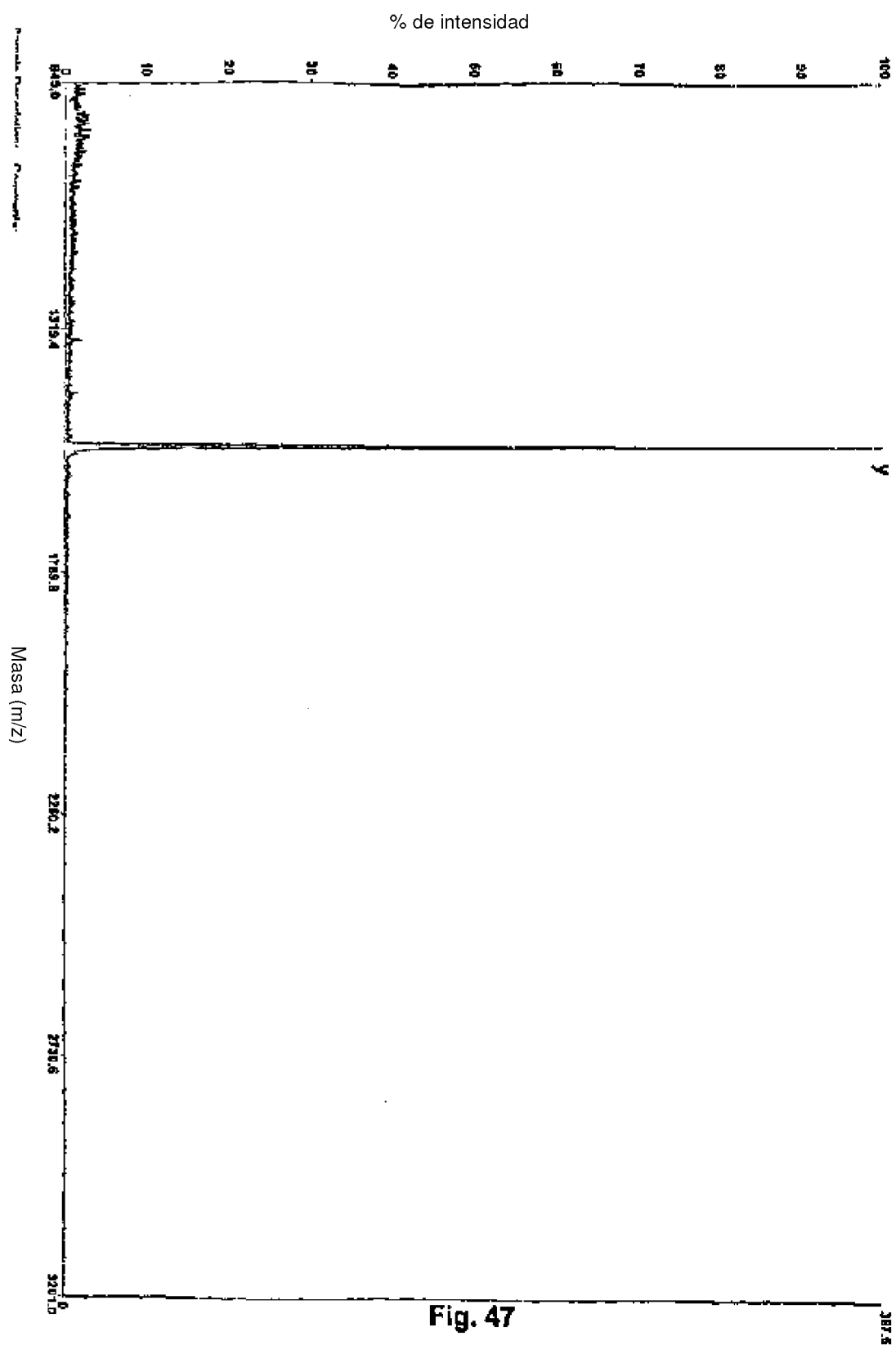
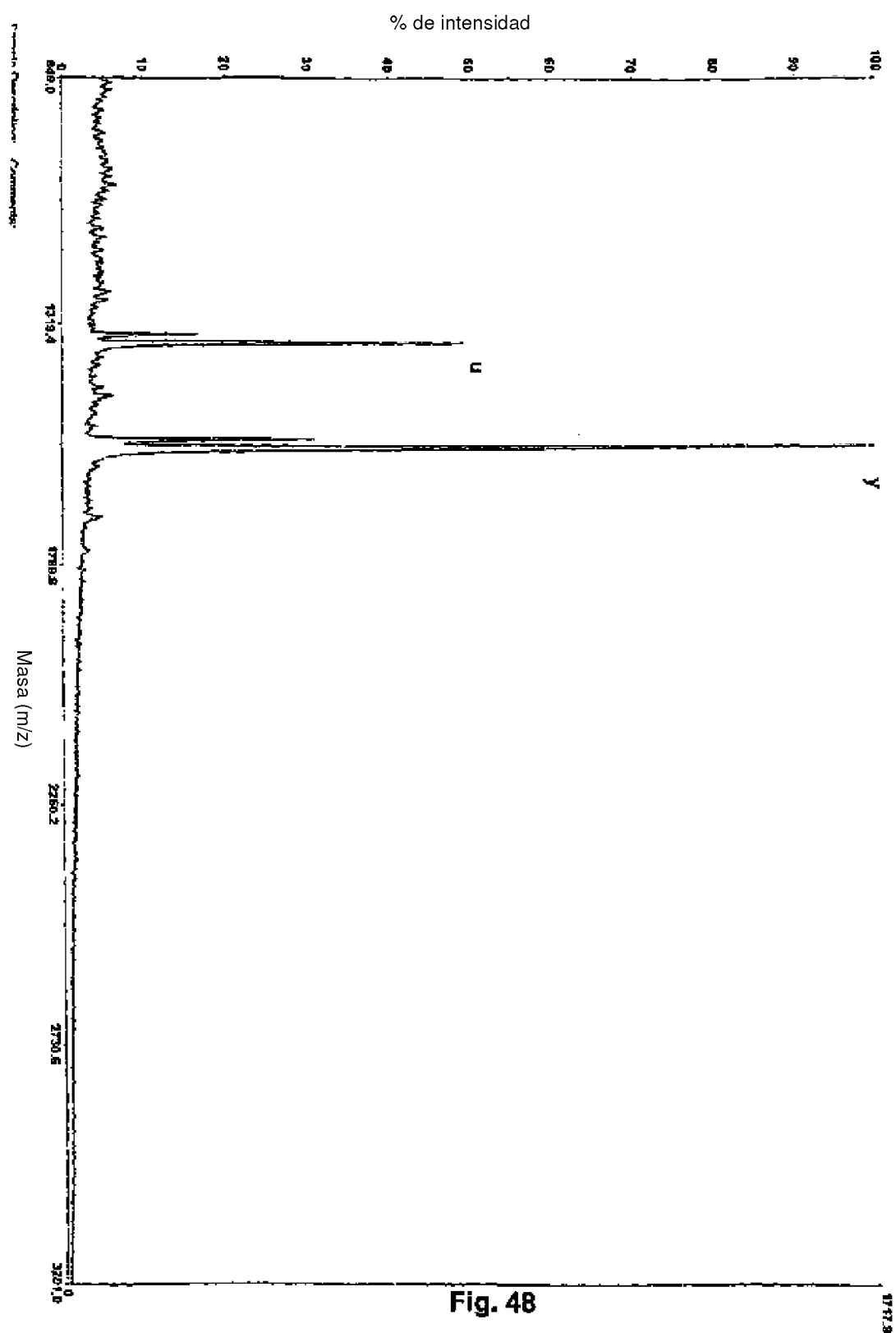


Fig. 47



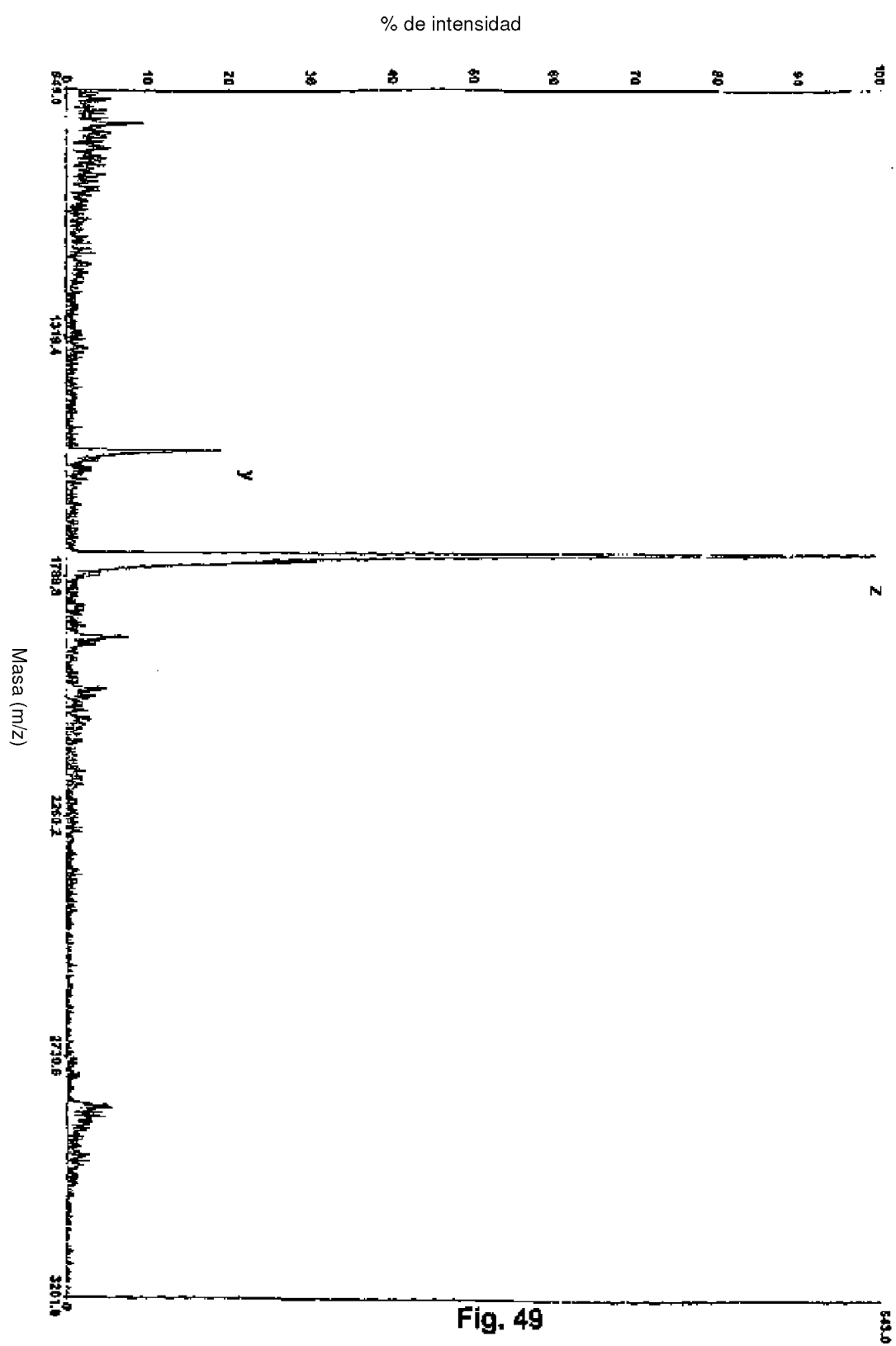


Fig. 49

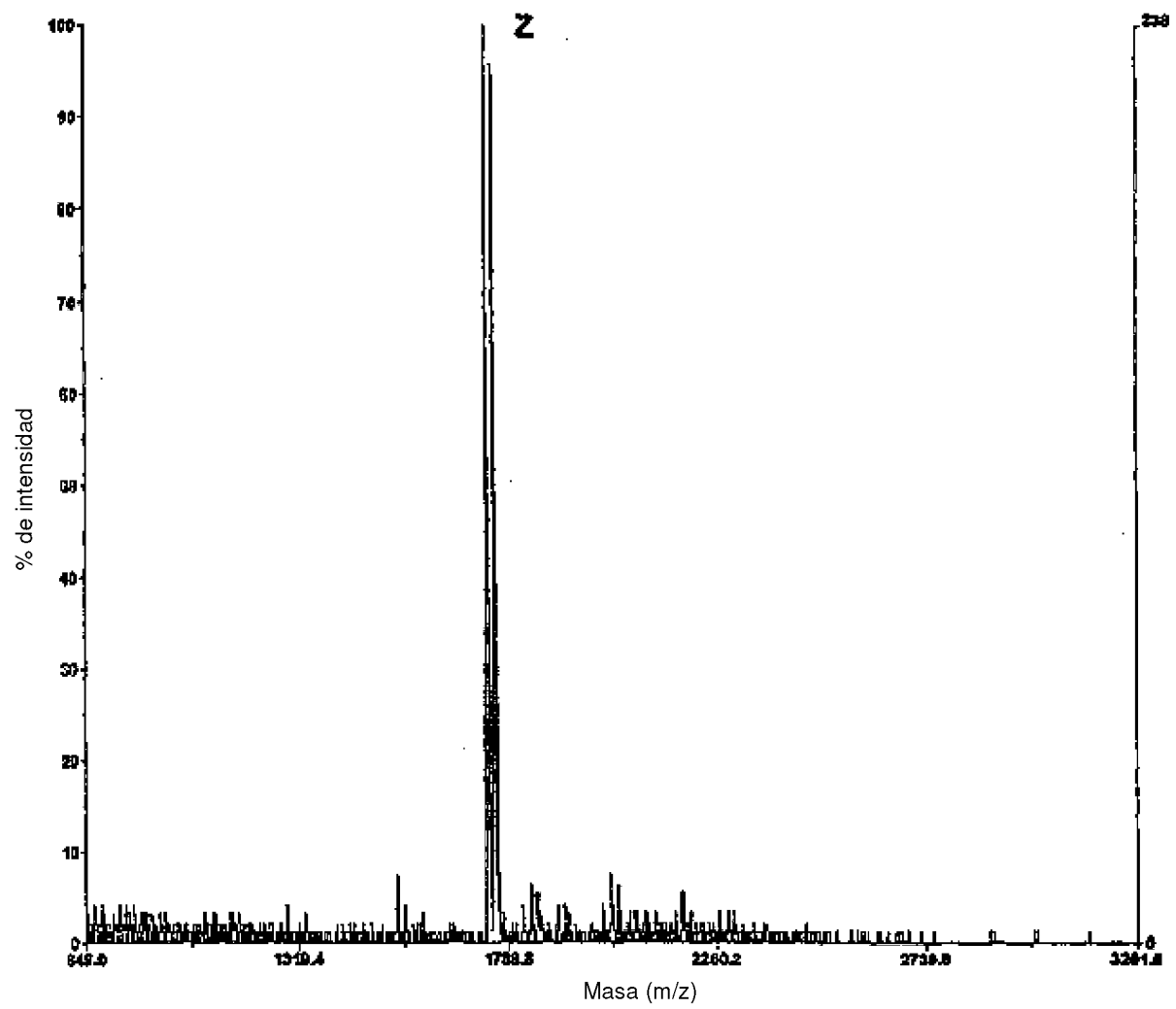


Fig. 50