

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 659 118**

51 Int. Cl.:

C12P 21/00 (2006.01)

C12N 9/10 (2006.01)

C12N 9/24 (2006.01)

C12N 1/13 (2006.01)

C07K 14/505 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.07.2011 PCT/EP2011/003756**

87 Fecha y número de publicación internacional: **02.02.2012 WO12013337**

96 Fecha de presentación y número de la solicitud europea: **27.07.2011 E 11741110 (8)**

97 Fecha y número de publicación de la concesión europea: **08.11.2017 EP 2598648**

54 Título: **N-glicosilación en *Phaeodactylum tricornutum* transformada**

30 Prioridad:

27.07.2010 EP 10007813

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

13.03.2018

73 Titular/es:

**UNIVERSITÉ DE ROUEN (50.0%)
1, rue Thomas Becket
76130 Mont-Saint-Aignan, FR y
INSTITUT FRANÇAIS DE RECHERCHE POUR
L'EXPLOITATION DE LA MER (IFREMER) (50.0%)**

72 Inventor/es:

**BARDOR, MURIEL;
LOUVET, ROMAIN;
SAINT-JEAN, BRUNO;
BUREL, CAROLE;
BAIET, BÉRANGÈRE;
CADORET, JEAN-PAUL;
LEROUGE, PATRICE;
MICHEL, RÉMY y
CARLIER, AUDE**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 659 118 T3

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

DESCRIPCIÓN

N-glicosilación en *Phaeodactylum tricornutum* transformada

Campo de la invención

5 La presente invención se refiere a métodos para mejorar el patrón de N-glicosilación de proteínas en *Phaeodactylum tricornutum* con el fin de producir en estas microalgas glicoproteínas que tengan perfiles de N-glicano similares a los de las glicoproteínas producidas por células de animales, especialmente células de seres humanos, que se puedan usar como agentes terapéuticos humanos o animales.

Antecedentes de la invención

10 Las proteínas farmacéuticas se producen como proteínas recombinantes por expresión en sistemas de expresión eucariotas. Después de la síntesis de la cadena principal de la proteína, la proteína recombinante se somete a un posterior proceso postraduccional incluyendo la unión de residuos de azúcares, un proceso conocido como glicosilación. Sin embargo, los organismos eucariotas presentan diferentes procesos de glicosilación que implican enzimas específicas (glicosiltransferasas y glicosidasas), y de modo que los patrones de glicosilación, incluso de la misma proteína, serán diferentes dependiendo de la célula eucariota en la que se está produciendo la proteína particular. Por tanto, el patrón de glicosilación de las proteínas farmacéuticas expresadas en células hospedantes eucariotas difiere sustancialmente del patrón de glicosilación de las proteínas naturales producidas en seres humanos y otros mamíferos.

N-glicosilación: una importante modificación postraduccional de proteínas segregadas

20 La N-glicosilación es una importante etapa de la modificación postraduccional en la síntesis de proteínas en células eucariotas. El proceso de N-glicanos en la vía secretora es esencial para las proteínas destinadas a ser segregadas o integradas en membranas. La N-glicosilación comienza cuando la proteína es traducida y translocada (movida) desde el ribosoma al lumen del retículo endoplásmico (RE). En este proceso, se ensambla inicialmente un precursor de oligosacáridos fosfato de dolicol ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-PP-dolicol}$) en la cara citoplasmática y termina en la cara luminal de la membrana del RE (BURDA AND AEBI, *Biochimica et Biophysica Acta*, vol.1426, pp. 239-257, 1999). Este precursor es usado por el complejo de multisubunidades de oligosacariltransferasa (OST) que cataliza su transferencia a los residuos de asparagina de las secuencias de consenso Asn-X-Ser/Thr, cuando X es diferente de prolina y ácido aspártico, de una proteína diana (BURDA AND AEBI, antes mencionadas, 1999). A continuación, el precursor se desglucosila/reglucosila para asegurar el control de calidad de la proteína neosintetizada por interacción con las proteínas chaperonas residentes en el RE, calreticulina y calnexina. Estos episodios en el RE son cruciales para el plegamiento y la oligomerización apropiados de las proteínas segregadas (HELENIUS AND AEBI, *Science*, vol. 291, pp. 2364-2369, 2001), altamente conservadas en las células eucariotas investigadas hasta ahora. Estas etapas conducen a la formación de un conjunto limitado de N-glicanos del tipo de alto contenido en manosa (Fig. 1). Por el contrario, la adaptación evolutiva del proceso de N-glicanos en el aparato de Golgi ha dado lugar a una gran variedad de estructuras complejas específicas de los organismos. Las manosidasas localizadas en este compartimento degradan en primer lugar el precursor de oligosacáridos en N-glicanos del tipo alto contenido en manosa que varían desde $\text{Man}_9\text{GlcNAc}_2$ (Man-9) hasta $\text{Man}_5\text{GlcNAc}_2$ (Man-5). A continuación se transfiere la N-acetilglucosaminiltransferasa I (GnT I) a un primer residuo de GlcNAc en el brazo de $\beta(1,3)$ -manosa de $\text{Man}_5\text{GlcNAc}_2$ y facilita la síntesis de múltiples N-glicanos de tipo complejo estructuralmente diferentes (Fig. 1). A continuación, las acciones de α -manosidasa II y GnT II permiten la síntesis del núcleo $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$. Los oligosacáridos de tipo complejo surgen de la transferencia en el aparato de Golgi de residuos de monosacáridos al núcleo $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ bajo la acción de glicosiltransferasas específicas de los organismos. Como consecuencia, las proteínas maduras que abandonan la vía secretora albergan múltiples N-glicanos complejos específicos del organismo que permiten a la proteína adquirir un conjunto de funciones biológicas mediadas por glicanos. Como ilustración, los N-glicanos en mamíferos son estructuras poli-antenarias, poli-sialiladas maduras que albergan un residuo de fucosa unido por $\alpha(1,6)$ en la N-acetilglucosamina próxima del núcleo (núcleo- $\alpha(1,6)$ -fucosa) (Fig. 1).

Remodelación en N-glicanos similares a los de seres humanos por estrategias de activación

50 Puesto que los perfiles de glicosilación difieren entre las células hospedantes de mamíferos y eucariotas, se han desarrollado estrategias para la remodelación *in vivo* de las estructuras de glicano unidas por N a proteínas. Estas estrategias incluyen la desactivación de genes endógenos que están implicados en la transferencia de algunos monómeros específicos, y metodologías de activación basadas en la expresión en las células hospedantes de enzimas de mamíferos. Los resultados del procedimiento de activación, complementando el repertorio enzimático de la célula hospedante, en la síntesis en el sistema de expresión recombinante de glicanos unidos a N son similares a los encontrados en células de mamíferos. Como ilustración, se ha conseguido la remodelación de N-glicanos de plantas en N-glicanos similares de mamíferos expresando una $\beta(1,4)$ -galactosiltransferasa humana en células vegetales. También se ha realizado la inserción dirigida de la $\beta(1,4)$ -galactosiltransferasa humana en *Physcomitrella patens* lo que conduce a la adición de $\beta(1,4)$ -galactosa terminal a N-glicanos endógenos. La N-acetilglucosaminiltransferasa III (GnT III) humana también se ha expresado con éxito en plantas con el fin de

modificar *in planta* N-glicanos endógenos. Esta transferasa es capaz de introducir el residuo $\beta(1,4)$ -GlcNAc en la β -manosa de los N-glicanos de mamífero del núcleo (biseccionando GlcNAc). El documento WO2009/101160 describe la modificación de *Phaeodactylum tricornutum* con diversas actividades enzimáticas. El documento WO2009/101160 describe que es necesaria la inactivación de ciertas enzimas en microalgas para humanizar el patrón de glicosilación. Las enzimas son alfa-manosidasa I, N-acetilglucosaminiltransferasa, alfa-1,3-fucosiltransferasa y beta-1,2-xilosiltransferasa.

Con la excepción de IgG, las proteínas séricas humanas requieren ácido siálico en las posiciones terminales de sus N-glicanos (Fig. 1). La mayoría de las células eucariotas que no son de mamífero, tales como plantas, no sintetizan ácidos N-acetil- y N-glicolil-neuramínicos (Neu₅Ac y Neu₅Gc), los dos principales ácidos siálicos de mamíferos. Como consecuencia, se ha desarrollado la manipulación genética para la síntesis *in planta* de proteínas sialiladas por expresión de enzimas capaces de sintetizar CMP-Neu₅Ac, su transportador de Golgi y las sialiltransferasas apropiadas (Paccalet et al., 2007, *Plant Biotechnology Journal*, vol.5, pp. 12-25; Castilho et al., 2008, *Plant Physiol.*, Vol. 147 (11), pp. 331-339; Castilho et al., 2010, *The Journal of Biological Chemistry*, vol. 285 (21), pp. 15923-30.

Breve descripción de los dibujos

La figura 1 muestra la biosíntesis de glicanos unidos a N en mamíferos.

La figura 2 muestra la detección por afinidad usando concanavalina A (Con A) y la inmunodetección usando anticuerpos producidos contra los epítomos de núcleo $\beta(1,2)$ -xilosa (anti-Xyl) y núcleo $\alpha(1,3)$ -fucosa (anti-Fuc) de proteínas aisladas de cebolleta usadas como control positivo (Pista 1) y de *Phaeodactylum tricornutum* (Pista 2).

La figura 3 muestra el vector de transformación BSJ-25 (SEQ ID NO 79). El casete de expresión está compuesto por el promotor 35S del virus del mosaico de la coliflor (CaMV) doble (SEQ ID NO 80), una secuencia Ω del virus del mosaico del tabaco como potenciador de la traducción y un péptido señal (PS) de quitinasa (SEQ ID NO 81) que está situado aguas arriba del sitio de clonación múltiple (MCS) y la secuencia del terminador *FcpA* de *Phaeodactylum tricornutum* en el extremo. El marcador de selección, el gen resistente a la bleomicina, *ShBle*, es activado por el promotor *FcpB*. La construcción también contiene el gen resistente a la ampicilina, *Amp*, y el origen de replicación de *Escherichia coli*.

La figura 4 muestra el casete de expresión de la GnT I de *Phaeodactylum tricornutum*. El primer casete comprende la secuencia que codifica la GnT I endógena situada bajo el control de secuencias reguladoras endógenas. En la segunda construcción, la GnT I está fusionada a la proteína verde fluorescente (GFP).

La figura 5 muestra el cribado de microalgas transformadas con la secuencia que codifica la GnT I de *Phaeodactylum tricornutum* (*Pt*-GnT I). *Pt*-GnT I: microalgas potencialmente transformadas; T-: control negativo de amplificación en agua; T+: amplificación por PCR realizada en el vector de transformación.

La figura 6 muestra amplificaciones por PCR del gen constitutivo H4 realizada en el cDNA obtenido de los linajes de *P. tricornutum* que fueron transformados con la secuencia que codifica la proteína GnT I (*Pt*-GnT I). *Pt*: cepa natural de *P. tricornutum*; T-: control negativo de amplificación en agua; RT+: amplificación por PCR realizada en el cDNA; RT-: control negativo de transcripción inversa.

La figura 7 muestra la amplificación por PCR de GnT I realizada en el cDNA obtenido a partir de *P. tricornutum* que fue transformada con la secuencia que codifica la GnT I endógena (*Pt*-GnT I). *Pt*: cepa natural de *P. tricornutum*; T-: control negativo de amplificación en agua; RT+: amplificación por PCR realizada en el cDNA; RT-: control negativo de transcripción inversa.

La figura 8 muestra la Q-PCR realizada en un linaje de *P. tricornutum* de tipo natural y en un linaje que expresa el transgén de GnT I. A: curva de fusión de los amplicones; B: perfiles de amplificación de los linajes de tipo natural y transformados de *P. tricornutum*. *Pt*-GnT I: curvas de amplificación obtenidas en 3 diluciones de cDNA a partir del linaje transformado con el gen GnT I. *Pt*-wt: curvas de amplificación obtenidas en 3 diluciones de cDNA a partir del linaje de tipo natural.

La figura 9 muestra la observación por microscopía confocal de microalgas *P. tricornutum* que fueron transformadas con los genes que codifican una proteína verde fluorescente citosólica (*Pt*-eGFP) y una proteína de fusión GnT I-eGFP (*Pt*-GnT I-eGFP).

La figura 10 muestra la fluorescencia observada en el cloroplasto y el retículo endoplásmico de microalgas *P. tricornutum* que expresan la proteína eGFP y la fluorescencia observada en el aparato de Golgi de microalgas que expresan GnT I-eGFP, después de un cultivo de dichas microalgas en brefeldin A (un agente capaz de dismantelar la estructura del aparato de Golgi bloqueando el transporte desde el retículo endoplásmico hasta el aparato de Golgi).

La figura 11 muestra la detección de la proteína de fusión GnT I-eGFP por transferencia de Western. La muestra de *Pt*-eGFP corresponde a un extracto de proteína de *P. tricornutum* que produce proteína eGFP citosólica. La muestra

negativa (*Pt*) corresponde al extracto de proteína de una microalga *P. tricornutum* no transformada. Por último, la *Pt*-GnT I-eGFP corresponde a extractos de proteínas de dos clones de *P. tricornutum* que fueron transformados con una secuencia que codifica la proteína de fusión GnT I-eGFP y que presenta fluorescencia en el aparato de Golgi.

La figura 12 muestra las construcciones usadas para la expresión de GnT I en *P. tricornutum*.

- 5 La figura 13 muestra N-glicanos del tipo con alto contenido en manosa que son los principales oligosacáridos unidos en N a proteínas de *P. tricornutum*. (A) Espectro de masas MALDI-TOF de glicanos unidos a N liberados por PNGasa A desde glicoproteínas de *P. tricornutum* y marcados con 2-aminobenzamida (2-AB). (B) Espectro de masas MALDI-TOF del agrupamiento de N-glicanos después de tratamiento con α -manosidasa de canavalia. (C) espectro de masas MALDI-TOF de glicanos unidos a N marcados con 2-AB liberados por PNGasa F desde glicoproteínas de *P. tricornutum*. Man-3 a Man-9: N-glicanos de tipo con alto contenido en manosa y de tipo con alto contenido en manosa Man₃GlcNAc₂ a Man₉GlcNAc₂. *: contaminantes. ■: aductos de potasio.

- 15 La figura 14 muestra la deficiencia de maduración de N-glicanos de los complementos GnT I de *P. tricornutum* en el mutante *Lec1* de CHO. Espectros de masas MALDI-TOF de glicanos unidos en N a proteínas extraídas de células de CHO. (A) tipo natural, (B) mutante *Lec1* de CHO y (C) transformante 4 del mutante *Lec1* de CHO complementado con el gen GnT I de *P. tricornutum*. Man-4 a Man-9: N-glicanos de tipo con alto contenido en manosa Man₄GlcNAc₂ a Man₉GlcNAc₂. Símbolos (70): cuadrado negro con contorno negro: GlcNAc, círculo gris con contorno negro: Man, círculo blanco con contorno negro: Gal, triángulo gris con contorno negro: Fucosa.

Sumario de la invención

- 20 Un primer aspecto de la invención se refiere a *Phaeodactylum tricornutum* transformada cuya vía de N-glicosilación ha sido modificada por la inactivación de β -N-acetilglucosaminidasas tanto de las secuencias de aminoácidos SEQ ID NO 9 como SEQ ID NO 11, no presentando dichas β -N-acetilglucosaminidasas inactivadas ninguna actividad enzimática.

- 25 En una realización preferida, dicha *P. tricornutum* comprende además otra secuencia de ácidos nucleicos unida operativamente a un promotor, codificando dicha otra secuencia de ácidos nucleicos un polipéptido que es expresado y glicosilado en la *P. tricornutum* transformada.

Un segundo aspecto de la invención se refiere a un método para producir un polipéptido glicosilado, comprendiendo dicho método las etapas de:

cultivar una *P. tricornutum* transformada como se ha descrito anteriormente,

purificar dicho polipéptido que es expresado y glicosilado en dicha *P. tricornutum* transformada.

- 30 En una realización preferida, dicho método comprende una etapa adicional (iii) de determinación del patrón de glicosilación de dicho polipéptido.

Un tercer aspecto de la invención se refiere a un uso de una *P. tricornutum* transformada como se ha descrito anteriormente para producir un polipéptido glicosilado.

Descripción detallada de la invención

- 35 Se ha predicho un gen que codifica una supuesta N-acetilglucosaminiltransferasa I (GnT I) en el genoma de *Phaeodactylum tricornutum* (*P. tricornutum*), pero los inventores establecieron que esta supuesta GnT I no mostraba ninguna actividad significativa en *Phaeodactylum tricornutum* en condiciones de cultivo estándares, puesto que las proteínas extraídas de *Phaeodactylum tricornutum* no presentaban ningún patrón de glicosilación GlcNAcMan₅GlcNAc₂ y llevaban alrededor del 95-97% de N-glicanos de tipo con alto contenido en manosa que variaba desde Man₉GlcNAc₂ (Man-9) hasta Man₅GlcNAc₂ (Man-5).

Sin embargo, los inventores encontraron sorprendentemente en otros experimentos, presentados en los siguientes ejemplos, que esta supuesta GnT I era capaz de restaurar la maduración de glicanos unidos a N en N-glicanos de tipo complejo en mutantes *Lec1* de CHO, defectuosos en su GnT I endógeno.

- 45 En consecuencia, los inventores muestran que esta supuesta GnT I tiene una actividad enzimática, cuya actividad enzimática puede restaurar la actividad defectuosa de GnT I de mamíferos.

Además, la N-acetilglucosaminiltransferasa I no es la única enzima identificada por los inventores en el genoma de *Phaeodactylum tricornutum*. Se ha identificado otro gen que codifica una alfa-manosidasa II (α -Man II) en el genoma de *Phaeodactylum tricornutum*, aunque no era detectable la actividad de α -Man II en *P. tricornutum* en condiciones de cultivos estándares.

GlcNAcMan₅GlcNAc₂, el producto de GnT I, se convierte sucesivamente en el aparato de Golgi en GlcNAcMan₄GlcNAc₂ y a continuación en GlcNAcMan₃GlcNAc₂ por la acción del α -Man II, seguido por la producción de GlcNAc₂Man₃GlcNAc₂ bajo la acción de GnT II.

5 No obstante, algunos organismos expresan β -N-acetilglucosaminidasas, que son enzimas responsables de la degradación de N-glicanos terminados en GlcNAc después de su biosíntesis en el aparato de Golgi con la acción de GnT I y Man II. La eliminación de GlcNAc terminal por β -N-acetilglucosaminidasas en el sistema secretor o en compartimentos en los que se acumulan proteínas pueden convertir entonces estos oligosacáridos en Man₄GlcNAc₂ y Man₃GlcNAc₂, reasociando así la vía de glicosilación.

10 Los inventores identificaron genes que codifican las supuestas β -N-acetilglucosaminidasas en el genoma de *Phaeodactylum tricornutum*: una primera β -N-acetilglucosaminidasa de secuencia de aminoácidos SEQ ID NO 9 es codificada por la secuencia de ácidos nucleicos SEQ ID NO 10 (número de registro 45073), mientras que una segunda β -N-acetilglucosaminidasa de secuencia de aminoácidos SEQ ID NO 11 es codificada por la secuencia de ácidos nucleicos SEQ ID NO 12 (número de registro 49563).

15 Estos supuestos genes pueden explicar la ausencia de la cantidad detectable de GlcNAcMan₅GlcNAc₂, el producto de GnT I, o GlcNAcMan₄GlcNAc₂ y GlcNAcMan₃GlcNAc₂, productos de GnT I y α -Man II, en proteínas de *P. tricornutum*.

20 Por tanto, un primer objeto de la invención es una *Phaeodactylum tricornutum* transformada cuya vía de N-glicosilación ha sido modificada por la inactivación de β -N-acetilglucosaminidasas tanto de secuencias de aminoácidos SEQ ID NO 9 como SEQ ID NO 11, no presentando dichas β -N-acetilglucosaminidasas inactivadas ninguna actividad enzimática.

En una primera realización, dicha *Phaeodactylum tricornutum* transformada comprende además una secuencia de ácidos nucleicos unida operativamente a un promotor, en donde:

- (i) dicha secuencia de ácidos nucleicos codifica una N-acetilglucosaminiltransferasa I que tiene la secuencia de aminoácidos SEQ ID NO 1, o una secuencia de aminoácidos que comprende al menos SEQ ID NO 4; y
- 25 (ii) dicho promotor unido operativamente tiene una identidad de secuencia de menos del 50% con la SEQ ID NO 3.

Phaeodactylum tricornutum es una microalga que pertenece a la clase Bacillariophyceae, al orden Naviculales, a la familia Phaeodactylaceae y al género Phaeodactylum.

30 El término "secuencia de ácidos nucleicos" utilizado en la presente memoria se refiere a secuencias de DNA (por ejemplo, cDNA o DNA genómico o sintético), así como a análogos de DNA que contienen análogos de nucleótidos no naturales, enlaces internucleosídicos no naturales, o ambos. Preferiblemente, dicha secuencia de ácidos nucleicos es una secuencia de DNA. Esta secuencia de ácidos nucleicos puede estar en cualquier conformación topológica, como lineal o circular.

La expresión promotor "unido operativamente" se refiere a una unión en la que el promotor es contiguo al gen de interés para controlar la expresión de dicho gen.

35 El término "*Phaeodactylum tricornutum* transformada" se refiere a una *P. tricornutum* en la que la secuencia de ácidos nucleicos unida operativamente a un promotor se ha introducido en dicha microalga por métodos convencionales de transformación, como se describe a continuación, para expresar dicha molécula de ácidos nucleicos en el núcleo de dicha *P. tricornutum*.

40 La transformación de *P. tricornutum* se puede realizar por métodos convencionales, tales como bombardeo de micropartículas, electroporación, perlas de vidrio, polietilenglicol (PEG), bigotes de carburo de silicio, o el uso de virus o agrobacterias. Dicho protocolo se describe en los ejemplos. La secuencia de ácidos nucleicos se puede introducir en *Phaeodactylum tricornutum* vía un plásmido, secuencias de virus, DNA monocatenario o bicatenario, DNA circular o lineal. En general es deseable incluir en cada secuencia o vector de ácidos nucleicos al menos un marcador seleccionable para permitir la selección de *Phaeodactylum tricornutum* que ha sido transformada establemente. Ejemplos de dichos marcadores son genes resistentes a antibióticos, tales como el gen *sh ble* que permite la resistencia a zeocina, los genes *nat* o *sat-1* que permiten la resistencia a nourseotricina y el gen *bar* que permite la resistencia a glufosinato.

45 La N-acetilglucosaminiltransferasa I, también conocida como GnT I o manosido-acetilglucosaminiltransferasa I (MGAT I) es una enzima de la vía de N-glicosilación, que es capaz de añadir un residuo de N-acetilglucosamina (GlcNAc) a Man₅GlcNAc₂ para producir un GlcNAcMan₅GlcNAc₂.

La N-acetilglucosaminiltransferasa I que tiene la secuencia de aminoácidos SEQ ID NO 1 corresponde a la "N-acetilglucosaminiltransferasa endógena" codificada por la secuencia de ácidos nucleicos SEQ ID NO 2 presente en el genoma de *Phaeodactylum tricornutum* de tipo natural.

La expresión "sitio catalítico" se refiere a la secuencia de aminoácidos de una enzima, siendo dicha secuencia de aminoácidos responsable de la actividad enzimática. Por tanto, el sitio catalítico de la N-acetilglucosaminiltransferasa I corresponde a la secuencia de aminoácidos responsable de la adición de un residuo de N-acetilglucosamina (GlcNAc) a $\text{Man}_5\text{GlcNAc}_2$ para producir un $\text{GlcNAcMan}_5\text{GlcNAc}_2$.

5 Como se usa en la presente memoria, el término "derivado de N-acetilglucosaminiltransferasa I" se refiere a una secuencia de aminoácidos capaz de añadir un residuo de N-acetilglucosamina (GlcNAc) a $\text{Man}_5\text{GlcNAc}_2$ para producir un $\text{GlcNAcMan}_5\text{GlcNAc}_2$ y que tiene una identidad mayor del 85% con la secuencia de aminoácidos SEQ ID NO 1 o uno de sus fragmentos, preferiblemente una identidad mayor del 90% con la secuencia de aminoácidos SEQ ID NO 1 o uno de sus fragmentos, y más preferiblemente una identidad mayor del 95% con la secuencia de aminoácidos SEQ ID NO 1 o uno de sus fragmentos.

10 Como se usa en la presente memoria, "porcentaje de identidad" entre dos secuencias de aminoácidos significa el porcentaje de aminoácidos idénticos, entre las dos secuencias que se van a comparar, obtenido con el mejor alineamiento de dichas secuencias, siendo este porcentaje puramente estadístico y distribuyéndose las diferencias entre estas dos secuencias aleatoriamente en las secuencias de aminoácidos. Como se usa en la presente memoria, "mejor alineamiento" o "alineamiento óptimo", significa el alineamiento para el que el porcentaje de identidad determinado (véase a continuación) es el mayor. La comparación de secuencias entre dos secuencias de aminoácidos se realiza generalmente comparando estas secuencias que se han alineado previamente de acuerdo con el mejor alineamiento; esta comparación se realiza en segmentos de comparación con el fin de identificar y comparar las regiones locales de similitud. El mejor alineamiento de secuencias para llevar a cabo la comparación se puede realizar usando programas informáticos que utilizan dichos algoritmos (GAP, BESTFIT, BLAST P, BLAST N, FASTA, TFASTA en el paquete de programas informáticos Wisconsin Genetics). Para obtener el mejor alineamiento local, se puede usar preferiblemente el programa BLAST, con la matriz BLOSUM 62 o la matriz PAM 30. El porcentaje de identidad entre dos secuencias de aminoácidos se determina comparando estas dos secuencias alineadas óptimamente, pudiendo comprender las secuencias de aminoácidos adiciones o deleciones respecto a la secuencia de referencia con el fin de obtener el alineamiento óptimo entre estas dos secuencias. El porcentaje de identidad se calcula determinando el número de posiciones idénticas entre estas dos secuencias, y dividiendo este número por el número total de posiciones comparadas, y multiplicando el resultado obtenido por 100 para obtener el porcentaje de identidad entre estas dos secuencias.

15 Muchos promotores diferentes que permiten la expresión de una secuencia de ácidos nucleicos en *Phaeodactylum tricornutum* son conocidos por los expertos en la técnica. Como ejemplo de dichos promotores, se pueden citar los promotores nucleares, tales como fcpA y fcpB de *Phaeodactylum tricornutum* descritos en ZAVLASKAÍA *et al.* (*J. Phycol.*, vol.36, pp. 379-386, 2000). Sin embargo, este promotor tiene una identidad de secuencias menor del 50% con la SEQ ID NO 3 correspondiente a la secuencia de 1047 pb de la secuencia de ácidos nucleicos aguas arriba del ATG de la secuencia de ácidos nucleicos SEQ ID NO 2 presente en el genoma de *Phaeodactylum tricornutum* de tipo natural, preferiblemente menor del 25% y lo más preferiblemente menor del 10%.

20 En otra realización preferida, dicha *Phaeodactylum tricornutum* transformada comprende además una secuencia de ácidos nucleicos unida operativamente a un promotor, en donde:

- (i) dicha secuencia de ácidos nucleicos codifica una alfa-manosidasa II que tiene la secuencia de aminoácidos SEQ ID NO 5; y
- 40 (ii) dicho promotor unido operativamente tiene una identidad de secuencias menor del 50% con la SEQ ID NO 7.

La α -manosidasa II, también conocida como α -Man II, es una enzima que cataliza la primera etapa comprometida en la biosíntesis de N-glicanos complejos. La α -Man II es capaz de hidrolizar los residuos terminales de alfa-D-manosa unidos por (1->3) y (1->6) en el manosil-oligosacárido $\text{GlcNAcMan}_5\text{GlcNAc}_2$. $\text{GlcNAcMan}_5\text{GlcNAc}_2$, el producto de GnT I, se convierte sucesivamente en el aparato de Golgi en $\text{GlcNAcMan}_4\text{GlcNAc}_2$ y luego en $\text{GlcNAcMan}_3\text{GlcNAc}_2$ por acción de la α -Man II.

La α -manosidasa II que tiene la secuencia de aminoácidos SEQ ID NO 5 corresponde a la " α -manosidasa II endógena" codificada por la secuencia de ácidos nucleicos SEQ ID NO 6 presente en el genoma de *Phaeodactylum tricornutum* de tipo natural.

50 La expresión "sitio catalítico" se refiere a la secuencia de aminoácidos de una enzima, siendo dicha secuencia de aminoácidos responsable de la actividad enzimática. Por tanto, el sitio catalítico de la α -manosidasa II corresponde a la secuencia de aminoácidos responsable de la conversión de $\text{GlcNAcMan}_5\text{GlcNAc}_2$, el producto de GnT I, en el aparato de Golgi en $\text{GlcNAcMan}_4\text{GlcNAc}_2$ y luego en $\text{GlcNAcMan}_3\text{GlcNAc}_2$.

55 Como se usa en la presente memoria, el término "derivado de α -Man II" se refiere a una secuencia de aminoácidos capaz de convertir $\text{GlcNAcMan}_5\text{GlcNAc}_2$, el producto de GnT I, en el aparato de Golgi en $\text{GlcNAcMan}_4\text{GlcNAc}_2$ y luego en $\text{GlcNAcMan}_3\text{GlcNAc}_2$ y que tiene una identidad mayor del 85% con la secuencia de aminoácidos SEQ ID NO 5 o uno de sus fragmentos, preferiblemente una identidad mayor del 90% con la secuencia de aminoácidos SEQ

ID NO 5 o uno de sus fragmentos, y más preferiblemente una identidad mayor del 95% con la secuencia de aminoácidos SEQ ID NO 5 o uno de sus fragmentos.

Muchos promotores diferentes que permiten la expresión de una secuencia de ácidos nucleicos en *Phaeodactylum tricornutum* son conocidos por los expertos en la técnica. Como ejemplo de dichos promotores, se pueden citar los promotores nucleares, tales como fcpA y fcpB, de *Phaeodactylum tricornutum* descrito en ZAVLASKAIA et al. (*J. Phycol.*, vol.36, pp. 379-386, 2000). Sin embargo, este promotor tiene una identidad de secuencias menor del 50% con la SEQ ID NO 7 correspondiente a la secuencia de 1000 pb de la secuencia de ácidos nucleicos aguas arriba del ATG de la secuencia de ácidos nucleicos SEQ ID NO 6 presente en el genoma de *Phaeodactylum tricornutum* de tipo natural, preferiblemente menor del 25% y lo más preferiblemente menor del 10%.

El término “inactivado” con referencia a β -N-acetilglucosaminidasas se refiere a β -N-acetilglucosaminidasas que no presentan ninguna actividad enzimática.

Este tipo de enzima es capaz de hidrolizar exclusivamente el residuo GlcNAc unido a la α -manosa unida por 1,3/1,6 del pentasacárido nuclear de N-glicanos. Cabe destacar que esta enzima no podría hidrolizar Man₅GlcNAc, sino que además actuaría solo “aguas abajo” en GlcNAcMan₅GlcNAc₂, GlcNAcMan₄GlcNAc₂, GlcNAcMan₃GlcNAc₂ o GlcNAcMan₃GlcNAc₂ sustituido por residuos de fucosa unidos al GlcNAc próximo.

Por actividades enzimáticas de β -N-acetilglucosaminidasa, los autores de la presente invención se refieren a la eliminación del residuo GlcNAc terminal unido a las manosas del núcleo de glicanos de tipo complejo unidos en N a proteínas.

Dicha inactivación se puede obtener de varias maneras conocidas por los expertos en la técnica. Por ejemplo, los métodos de inactivación comprenden silenciamiento de genes con interferencia de RNA (miRNA, siRNA) que se han usado en microalgas (Zhao T., et al., 2009, *The Plant Journal*, vol. 58, pp. 157-164; Molnar A., et al., 2009, *The Plant Journal* vol. 58, pp. 165-174), como en *P. tricornutum* (De Riso V., et al., 2009, *Nucleic Acids Research*, pp. 1-12). La inactivación de dichas β -N-acetilglucosaminidasas también se puede obtener por desactivación de los genes correspondientes, pero también por el uso de un método de inhibición de la actividad enzimática, por ejemplo, usando anticuerpo dirigido al sitio catalítico de dicha β -N-acetilglucosaminidasas.

La inactivación de β -N-acetilglucosaminidasas se puede confirmar analizando su actividad enzimática, que se puede medir por técnicas que son conocidas por los expertos en la técnica, por ejemplo, usando uno de los dos ensayos de actividad diferentes siguientes, descritos en Léonard R., et al., 2006, *The Journal of Biological Chemistry*, vol. 281, pp. 4867-4875.

De acuerdo con un primer ensayo, las N-acetilglucosaminidasas se incuban con los diferentes sustratos a 37°C durante 1-20 horas. Para experimentos con *p*-nitrofenil-GlcNAc, la concentración del sustrato es 5 mM en un volumen total de 0,04 mL de tampón de citrato/fosfato 0,1 M a pH 3-8. Las reacciones se terminan por la adición de 0,26 mL de tampón de glicina/NaOH 0,4 M a pH 10,4, y se mide la absorbancia a 405 nm con un lector de placas de microtitulación.

De acuerdo con un segundo ensayo, los oligosacáridos piridilaminados se usan a una concentración final de 0,1 mM en un volumen total de 0,02 mL de tampón de citrato/fosfato 0,1 M a pH 3-8. La incubación se termina por la adición de 0,18 mL de borato de sodio 20 mM enfriado con hielo. Se analizan partes alícuotas de 0,05 mL por HPLC en fase inversa como se ha descrito anteriormente.

En una realización preferida, dichas β -N-acetilglucosaminidasas se han inactivado con la técnica de interferencia de RNA.

Preferiblemente, todas las *P. tricornutum* transformadas como se ha descrito anteriormente comprenden además una secuencia de ácidos nucleicos adicional unida operativamente a un promotor, codificando dicha secuencia de ácidos nucleicos adicional un polipéptido que se expresa y glicosila en la *P. tricornutum* transformada.

El término “polipéptido”, como se usa en la presente memoria, se refiere a una secuencia de aminoácidos que comprende más de 50 aminoácidos que están unidos por enlaces peptídicos.

Después de la transformación de *P. tricornutum*, se seleccionan los transformantes que producen el polipéptido deseado. La selección se puede realizar por uno o más métodos convencionales que comprenden: ensayo por inmunoabsorción ligado a enzimas (ELISA), espectroscopia de masas, tal como MALDI-TOF-MS, cromatografía por ESI-MS, caracterización de células usando un clasificador de células activadas por fluorescencia, espectrofotometría, fluorimetría, inmunocitoquímica exponiendo las células a un anticuerpo que tenga una afinidad específica por la proteína deseada. Dichos métodos se detallan a continuación en los ejemplos.

Los polipéptidos glicosilados tienen al menos una estructura GlcNAcMan₅GlcNAc₂. Preferiblemente, dichos polipéptidos glicosilados tienen al menos una GlcNAcMan₄GlcNAc₂, GlcNAcMan₃GlcNAc₂ o GlcNAc₂Man₃GlcNAc₂.

Ventajosamente, el polipéptido expresado y glicosilado por la *Phaeodactylum tricornutum* transformada es un polipéptido de origen animal, preferiblemente de origen mamífero, y lo más preferiblemente de origen humano.

5 En una realización todavía preferida, el polipéptido expresado y glicosilado por la *P. tricornutum* transformada de la invención es un polipéptido que tiene un interés terapéutico. Preferiblemente, dicho polipéptido se selecciona del grupo que comprende eritropoyetina, citoquinas, tales como interferones, anticuerpos y sus fragmentos, factores de coagulación, hormonas, beta-glucocerebrosidasa, pentraxina-3, anti-TNF, α -glucosidasa ácida, α -L-iduronidasa y sus derivados.

10 Un anticuerpo es una molécula de inmunoglobulina correspondiente a un tetrámero que comprende cuatro cadenas polipeptídicas, dos cadenas pesadas (H) idénticas (aproximadamente 50-70 kDa con su longitud completa) y dos cadenas ligeras (L) idénticas (aproximadamente 25 kDa con su longitud completa) interconectadas por enlaces disulfuro. Las cadenas ligeras se clasifican como kappa y lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o épsilon, y definen el isotipo de anticuerpo como IgG, IgM, IgA, IgD e IgE, respectivamente. Cada cadena pesada está compuesta por una región variable de la cadena pesada en el extremo N (abreviada en la presente memoria como HCVR) y una región constante de la cadena pesada. La región constante de la cadena pesada está compuesta por tres dominios (CH1, CH2 y CH3) para IgG, IgD e IgA; y 4 dominios (CH1, CH2, CH3 y CH4) para IgM e IgE. Cada cadena ligera está compuesta por una región variable de la cadena ligera en el extremo N (abreviada en la presente memoria como LCVR) y una región constante de la cadena ligera. La región constante de la cadena ligera está compuesta por un dominio, CL. Las regiones HCVR y LCVR se pueden subdividir además en regiones de hipervariabilidad, denominadas regiones determinantes de la complementariedad (CDR), intercaladas con regiones que están más conservadas, denominadas regiones de entramado (FR). Cada HCVR y LCVR está compuesta de tres CDR y cuatro FR, dispuestas desde el extremo amino hasta el extremo carboxi en el siguiente orden: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. La asignación de aminoácidos a cada dominio está de acuerdo con los usos bien conocidos. La capacidad funcional del anticuerpo para unirse a un antígeno particular depende de las regiones variables de cada par de cadenas ligera/pesada, y está determinada en gran parte por las CDR.

20 El término "anticuerpo", como se usa en la presente memoria, se refiere a un anticuerpo monoclonal *per se*. Un anticuerpo monoclonal puede ser un anticuerpo humano, anticuerpo quimérico y/o anticuerpo humanizado.

30 El término "fragmentos de anticuerpos" como se usa en la presente memoria se refiere a fragmentos de anticuerpos que se unen a los antígenos particulares de dicho anticuerpo. Por ejemplo, están abarcados por la invención los fragmentos de anticuerpo capaces de unirse a antígenos particulares que incluyen Fab (por ejemplo, por digestión con papaína), Fab' (por ejemplo, por digestión con pepsina y reducción parcial) y F(ab')₂ (por ejemplo, por digestión con pepsina), facb (por ejemplo, por digestión con plasmina), pFc' (por ejemplo, por digestión con pepsina o plasmina), Fd (por ejemplo, por digestión con pepsina, reducción parcial y reagregación) y fragmentos Fv o ScFv (por ejemplo, por técnicas de biología molecular).

35 Dichos fragmentos se pueden producir por escisión enzimática, técnicas sintéticas o recombinantes, como se conoce en la técnica y/o como se describe en la presente memoria. Los anticuerpos también se pueden producir en una variedad de formas truncadas usando genes de anticuerpos en los que se han introducido uno o más codones de parada aguas arriba del sitio de parada natural. Por ejemplo, un gen de combinación que codifica una porción de cadena pesada F(ab')₂ se puede diseñar de forma que incluya secuencias de DNA que codifiquen el dominio CH₁ y/o la región bisagra de la cadena pesada. Las diversas porciones de anticuerpos se pueden unir entre sí químicamente por técnicas convencionales, o se pueden preparar como una proteína contigua usando técnicas de ingeniería genética.

45 El término "citoquinas" se refiere a proteínas de señalización que son liberadas por células específicas del sistema inmunitario para llevar una señal a otras células con el fin de alterar su función. Las citoquinas son agentes inmunomoduladores y se usan ampliamente en la comunicación celular. El término citoquinas abarca una amplia gama de reguladores polipeptídicos, tales como interferones, interleuquinas, quimioquinas o el factor de necrosis tumoral.

50 El término "factores de coagulación" se refiere a las proteínas plasmáticas que interactúan con las plaquetas en una cascada compleja de reacciones catalizadas por enzimas, conduciendo a la formación de fibrina para la iniciación de un coágulo de sangre en el proceso de coagulación de la sangre. Los factores de coagulación, en número de 13, son generalmente serina-proteasas, pero también comprenden glicoproteínas (Factores VIII y V) u otros tipos de enzimas, tales como transglutaminasa (Factor XIII).

55 El término "hormonas" se refiere a mensajeros químicos segregados por células específicas en el plasma o la linfa para producir sus efectos sobre otras células del organismo a una distancia de sus sitios de producción. La mayoría de las hormonas inician una respuesta celular combinándose inicialmente con una cualquiera de las proteínas receptoras asociada a la membrana intracelular o celular específica. Las hormonas conocidas comúnmente son, por ejemplo, la insulina para la regulación de la energía y la glucosa en el organismo, o la hormona del crecimiento que estimula el crecimiento y la reproducción y regeneración celular.

En otra realización preferida, la invención se refiere a una *P. tricornutum* transformada como se ha descrito antes, que comprende además otra secuencia de ácidos nucleicos unida operativamente a un promotor, en donde dicha secuencia de ácidos nucleicos codifica una N-acetilglucosaminiltransferasa II, uno de sus fragmentos o derivados.

5 La N-acetilglucosaminiltransferasa II, también conocida como GnT II o manosido-acetilglucosaminiltransferasa II (MGAT II) es una enzima de la vía de N-glicosilación, que es capaz de añadir un residuo de N-acetilglucosamina (GlcNAc) a GlcNAcMan₃GlcNAc₂, producto de α-Man II, para producir una GlcNAc₂Man₃GlcNAc₂.

10 Ejemplos de la GnT II, también conocida como manosil(alfa-1,6-)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa (MGAT 2), incluyen GnT II de *Mus musculus* (SEQ ID NO 13, número de registro NP_666147), de *Homo sapiens* (SEQ ID NO 14, número de registro NP_002399) o de *Phaeodactylum tricornutum* (SEQ ID NO 5).

Preferiblemente, la N-acetilglucosaminiltransferasa II tiene una secuencia de aminoácidos de SEQ ID NO 5, uno de sus fragmentos o derivados, y dicho promotor unido operativamente tiene una identidad de secuencias menor del 50% con SEQ ID NO 7 o uno de sus fragmentos.

15 Incluso en otra realización preferida, dicha N-acetilglucosaminiltransferasa II comprende la secuencia de aminoácidos SEQ ID NO 15 y dicho promotor unido operativamente tiene una identidad de secuencias menor del 50% con SEQ ID NO 7 o uno de sus fragmentos.

Los inventores han observado que la α-manosidasa II tiene un dominio de N-acetilglucosaminiltransferasa II de la secuencia de aminoácidos SEQ ID NO 15.

20 La expresión "sitio catalítico" se refiere a la secuencia de aminoácidos de una enzima, siendo dicha secuencia de aminoácidos responsable de la actividad enzimática. Por tanto, el sitio catalítico de la GnT II corresponde a la secuencia de aminoácidos responsable de la adición de un residuo de N-acetilglucosamina (GlcNAc) a GlcNAcMan₃GlcNAc₂ para producir una GlcNAc₂Man₃GlcNAc₂.

25 Como se usa en la presente memoria, el término "derivado de GnT II" se refiere a una secuencia de aminoácidos capaz de añadir un residuo de N-acetilglucosamina (GlcNAc) a GlcNAcMan₃GlcNAc₂, producto de α-Man II, para producir una GlcNAc₂Man₃GlcNAc₂ y que tiene una identidad mayor del 85% con la secuencia de aminoácidos SEQ ID NO 5, o uno de sus fragmentos, preferiblemente una identidad mayor del 90% con la secuencia de aminoácidos SEQ ID NO 5, o uno de sus fragmentos, y más preferiblemente una identidad mayor del 95% con la secuencia de aminoácidos SEQ ID NO 5 o uno de sus fragmentos.

30 En otra realización preferida, la invención se refiere a una *P. tricornutum* transformada como se ha descrito antes, que comprende además otra secuencia de ácidos nucleicos unida operativamente a un promotor, codificando dicha secuencia de ácidos nucleicos al menos una enzima de la vía de N-glicosilación humana, tal como N-acetilglucosaminiltransferasas III, IV, V, VI y glicosiltransferasas, tales como galactosiltransferasas, fucosiltransferasas o sialiltransferasas. Dichas enzimas se expresan en dicha *P. tricornutum* transformada y permiten la N-glicosilación de un polipéptido.

35 En otra realización preferida, dicha *P. tricornutum* transformada como se ha descrito antes, que comprende además otra secuencia de ácidos nucleicos unida operativamente a un promotor, comprende una secuencia de ácidos nucleicos que codifica las N-acetilglucosaminiltransferasas III, IV, V y VI.

40 Incluso en otra realización preferida, dicha *P. tricornutum* transformada como se ha descrito antes, que comprende además otra secuencia de ácidos nucleicos unida operativamente a un promotor, comprende una secuencia de ácidos nucleicos que codifica glicosiltransferasas que comprenden galactosiltransferasas, fucosiltransferasas y sialiltransferasas.

GnT III, GnT IV, GnT V, GnT VI, fucosiltransferasa, galactosiltransferasa (GalT) y sialiltransferasas (ST) son muy conocidas por los expertos en la técnica.

45 Ejemplos de GnT III, también conocida como manosil(beta-1,4-)-glicoproteína beta-1,4-N-acetilglucosaminiltransferasa (MGAT 3), incluyen GnT III de *Mus musculus* (SEQ ID NO 17, número de registro NP_034925) o de *Homo sapiens* (SEQ ID NO 18, número de registro NP_002400). Preferiblemente, dicha N-acetilglucosaminiltransferasa III (GnT III) corresponde a la SEQ ID NO 18 (número de registro NP_02400).

50 Ejemplos de GnT IV, también conocida como manosil(alfa-1,3-)-glicoproteína beta-1,4-N-acetilglucosaminiltransferasa (MGAT 4), incluyen la isoenzima A de GnT IV de *Mus musculus* (SEQ ID NO 19, número de registro NP_776295), la isoenzima B de *Mus musculus* (SEQ ID NO 20, número de registro NP_666038), la isoenzima C de *Mus musculus* (SEQ ID NO 21, número de registro NP_080519), la isoenzima A de GnT IV de *Homo sapiens* (SEQ ID NO 22, número de registro NP_036346), la isoenzima B de GnT IV de *Homo sapiens* (la isoforma 1, SEQ ID NO 23, número de registro NP_055090 o la isoforma 2, SEQ ID NO 24, número de registro NP_463459) o la isoenzima C de GnT IV de *Homo sapiens* (SEQ ID NO 25, número de registro NP_037376).

Ejemplos de GnT V incluyen GnT V de *Mus musculus* (SEQ ID NO 26, número de registro NP_660110), la isoenzima B de GnT V de *Mus musculus* (SEQ ID NO 27, número de registro NP_766536), la GnT V de *Homo sapiens* (SEQ ID NO 28, número de registro NP_002401), la isoenzima B de GnT V de *Homo sapiens* (la isoforma 1, SEQ ID NO 29, número de registro NP_653278 o la isoforma 2, SEQ ID NO 30, número de registro NP_945193).

5 Ejemplo de GnT VI incluye GnT VI de *Gallus gallus* (SEQ ID NO 31, número de registro NP_990012).

Las fucosiltransferasas son muy conocidas por los expertos en la técnica e incluyen, como ejemplo, alfa-(1,6)-fucosiltransferasa (fucosiltransferasa 8 (FUT8)), como FUT8 de *Mus musculus* (SEQ ID NO 32, número de registro NP_058589) o FUT8 de *Homo sapiens* (SEQ ID NO 33, número de registro Q9BYC5). Preferiblemente, dicha fucosiltransferasa corresponde a SEQ ID NO 33 (número de registro Q9BYC5).

10 Las galactosiltransferasas son muy conocidas por los expertos en la técnica e incluyen, como ejemplo, una beta-(1,4)-galactosiltransferasa (B4GALT1), como B4GALT1 de *Homo sapiens* (SEQ ID NO 34, número de registro NP_001488) o B4GALT1 de *Mus musculus* (SEQ ID NO 35, número de registro CAM14782). Preferiblemente, dicha galactosiltransferasa corresponde a SEQ ID NO 34 (número de registro NP_001488).

15 Las sialiltransferasas son muy conocidas por los expertos en la técnica e incluyen, como ejemplo, alfa-2,6-sialiltransferasa (ST6 beta-galactosamida alfa-2,6-sialiltransferasa 1 (ST6GAL1) o beta galactosido alfa 2,6 sialiltransferasa 2 (ST6GAL2)), como ST6GAL2 de *Mus musculus* (SEQ ID NO 36, número de registro NP_766417) o ST6GAL1 de *Homo sapiens* (la isoforma a, SEQ ID NO 37, número de registro NP_775323 o la isoforma b, SEQ ID NO 38, número de registro NP_775324), o alfa 2,3 sialiltransferasa (ST3 beta-galactosido alfa-2,3-sialiltransferasa 6 (ST3GAL6), ST3 beta-galactosido alfa-2,3-sialiltransferasa 1 (ST3GAL1), ST3 beta-galactosido alfa-2,3-sialiltransferasa 2 (ST3GAL2), ST3 beta-galactosido alfa-2,3-sialiltransferasa 3 (ST3GAL3), como ST3GAL1 de *Mus musculus* (SEQ ID NO 39, número de registro NP_033203) o de *Homo sapiens* (SEQ ID NO 40, número de registro NP_003024), ST3GAL2 de *Homo sapiens* (SEQ ID NO 41, número de registro NP_008858), ST3GAL3 de *Homo sapiens* (la isoforma a, SEQ ID NO 42, número de registro NP_777623, la isoforma b, SEQ ID NO 43, número de registro NP_777624, la isoforma c, SEQ ID NO 44, número de registro NP_777625, la isoforma f, SEQ ID NO 45, número de registro NP_777628, la isoforma j, SEQ ID NO 46, número de registro NP_006270, la isoforma d, SEQ ID NO 47, número de registro NP_777626, la isoforma e, SEQ ID NO 48, número de registro NP_777627, la isoforma i, SEQ ID NO 49, número de registro NP_777631, la isoforma g, SEQ ID NO 50, número de registro NP_777629, la isoforma h, SEQ ID NO 51, número de registro NP_777630) o ST3GAL6 de *Homo sapiens* (SEQ ID NO 52, número de registro NP_006091).

30 Para que una glucosiltransferasa funcione satisfactoriamente en el aparato de Golgi, es necesario que la enzima sea proporcionada con concentraciones suficientes de un azúcar de nucleótidos apropiado, que es el donador de alta energía del resto de azúcar añadido a una glicoproteína naciente. En seres humanos, la gama completa de precursores del azúcar de nucleótidos generalmente es sintetizada en el citosol y transportada al aparato de Golgi, donde se une al oligosacárido del núcleo por las glicosiltransferasas. La firma solicitante observó en microalgas una concentración suficiente de GlcNAc, manosa, fucosa y galactosa, pero no de ácido siálico.

35 Por tanto, para que una sialiltransferasa funcione satisfactoriamente en el aparato de Golgi, es necesario expresar en las microalgas una o más enzimas necesarias para la síntesis de ácido siálico, su activación y su transporte dentro del aparato de Golgi entre UDP-GlcNAc 2-epimerasa, GlcNAc 2-epimerasa, GlcNAc-6P 2-epimerasa, NeuAc sintasa, NeuAc-9P sintasa, CMP-NeuAc sintasa y el transportador de CMP-ácido siálico (véanse, por ejemplo, los trabajos realizados en plantas: Misaki R et al., *Biochem Biophys Res Commun.* 2006 Jan 27; 339 (4): 1184-9; Paccalet T et al., *Plant Biotechnol J.* 2007 Jan; 5 (1): 16-25), Castihlo et al., 2008 y Castihlo et al., 2010 que se han descrito antes).

40 La UDP-GlcNAc 2-epimerasa, que es conocida también como glucosamina (UDP-N-acetil)-2-epimerasa/N-acetilmanosamina-quinasa (GNE), es muy conocida por los expertos en la técnica e incluye, por ejemplo, GNE de *Mus musculus* (SEQ ID NO 53, número de registro NP_056643) o GNE de *Homo sapiens* (SEQ ID NO 54, número de registro NP_005467). Preferiblemente, dicha GNE corresponde a la SEQ ID NO 54 (número de registro NP_005467).

La GlcNAc 2-epimerasa es muy conocida por los expertos en la técnica e incluye, como ejemplo, la proteína de unión a renina (RENBP) de *Homo sapiens* (SEQ ID NO 55, número de registro NP_002901).

50 La NeuAc-9-P-sintasa, también llamada ácido N-acetilneuramínico-sintasa (NANS), es muy conocida por los expertos en la técnica e incluye, como ejemplo, NANS de *Homo sapiens* (SEQ ID NO 56, número de registro NP_061819).

55 La CMP-NeuAc-sintasa, que es conocida también como citidina-monofosfo-ácido N-acetilneuramínico-sintetasa (CMAS), es muy conocida por los expertos en la técnica e incluye, como ejemplo, CMAS de *Mus musculus* (SEQ ID NO 57, número de registro NP_034038) o de *Homo sapiens* (SEQ ID NO 58, número de registro NP_061156). Preferiblemente, dicha CMAS corresponde a la SEQ ID NO 58 (número de registro NP_061156).

Los transportadores de CMP-ácido siálico son también muy conocidos por los expertos en la técnica e incluyen, como ejemplo, la familia de portadores de soluto 35 (transportador de CMP-ácido siálico), miembro AI (SLC35A1) de *Mus musculus* (SEQ ID NO 59, número de registro NP_036025) o de *Homo sapiens* (SEQ ID NO 60, número de registro NP_006407). Preferiblemente, dicho transportador de CMP-ácido siálico corresponde a SLC35A1 de *Homo sapiens* (SEQ ID NO 60, número de registro NP_006407).

La proteína transportadora añadida transporta un azúcar de nucleótido desde el citosol al aparato de Golgi, donde el azúcar de nucleótido puede reaccionar por medio de la glicosiltransferasa, por ejemplo, para prolongar un N-glicano. La reacción libera un nucleósido-difosfato o -monofosfato, por ejemplo, UDP, GDP o CMP. Como la acumulación de un nucleósido-difosfato inhibe la actividad adicional de una glicosiltransferasa, frecuentemente también es deseable proporcionar una copia expresada de un gen que codifique una nucleótido-difosfatasa. La difosfatasa (específica para UDP o GDP según proceda) hidroliza el difosfonucleósido proporcionando un nucleósido-monofosfato y un fosfato inorgánico. El nucleósido-monofosfato no inhibe la glicosiltransferasa y, en cualquier caso, es exportado desde el aparato de Golgi por un sistema celular endógeno.

Otro objeto de la invención es un método para producir un polipéptido glicosilado, comprendiendo dicho método las etapas de:

(i) cultivar una *P. tricornutum* transformada como se ha descrito antes;

(ii) purificar dicho polipéptido expresado y glicosilado en dicha *P. tricornutum* transformada.

En una realización preferida, dicho método para producir un polipéptido glicosilado comprende una etapa previa de transformación de una *Phaeodactylum tricornutum* de modo que se obtenga una *P. tricornutum* como se ha definido anteriormente.

Los métodos que se pueden emplear para la transformación de *P. tricornutum* son los descritos antes en la presente memoria. Dicha transformación, cultivo de *P. tricornutum* y purificación de los polipéptidos glicosilados se ilustran también a continuación.

Ventajosamente, el método de la invención comprende además una etapa (iii) para determinar el patrón de glicosilación de dicho polipéptido.

Este patrón de glicosilación se puede determinar por un método muy conocido por los expertos en la técnica. Como ejemplo, se pueden obtener informaciones preliminares sobre la N-glicosilación de la glicoproteína recombinante por análisis de afino- y de inmuno-transferencia usando sondas específicas, tales como lectinas (CON A; ECA; SNA; MAA ...) y anticuerpos de N-glicanos específicos (anti- β 1,2-xilosa; anti- α -1,3-fucosa; anti-Neu5Gc, anti-Lewis ...). Esto se realiza de acuerdo con FITCHETTE *et al.*, (*Methods Mol. Biol.*, vol. 355, pp. 317-342, 2007) y podría completarse por ensayos de desglicosilación.

Para investigar el perfil detallado de N-glicanos de la proteína recombinante, los oligosacáridos unidos en N se liberan a continuación de la proteína de una manera no específica usando digestión enzimática o tratamiento químico (FITCHETTE *et al.*, antes mencionado, 2007; SEVENO *et al.*, *Anal. Biochem.*, vol. 379 (1), pp. 66-72, 2008). La mezcla de oligosacáridos reductores resultante se puede perfilar por métodos de HPLC y/o espectrometría de masas (ESI-MS-MS y MALDI-TOF esencialmente) (BARDOR *et al.*, *Curr Opin Struct Biol.*, vol.16 (5), pp. 576-583, 2006; SEVENO *et al.*, antes mencionado, 2008). Estas estrategias, acopladas a la digestión con exoglicosidasa, permiten la identificación y cuantificación de N-glicanos (SEVENO *et al.*, antes mencionado, 2008).

Otra alternativa para estudiar el perfil de N-glicosilación de la proteína recombinante es trabajar directamente en sus glicopéptidos después de la digestión con proteasa de la proteína, purificación y análisis por espectrometría de masas de los glicopéptidos, como se ha sido descrito por BARDOR *et al.* (*Plant Biotechnol. J.*, vol.1 (6), pp. 451-462, 2003).

Otro objeto de la invención se refiere al uso de una *P. tricornutum* transformada como se ha definido anteriormente para producir un polipéptido glicosilado.

A continuación, se describe la invención con más detalle con referencia a métodos. Sin embargo, no se pretende ninguna limitación de la invención por los detalles de los ejemplos. Por el contrario, la invención se refiere a cualquier realización que comprenda detalles que no se mencionan explícitamente en los ejemplos de la presente memoria, pero que los expertos en la técnica encuentran sin esfuerzo excesivo.

EJEMPLOS

Ejemplo 1 No se detectó ninguna actividad significativa de GnT I en *Phaeodactylum tricornutum*

Se predijo un gen que codifica una supuesta GnT I en el genoma de *P. tricornutum* (Pt54844; <http://genome.jgi-psf.org/Phatr2/Phatr2.home.html>) (SEQ ID NO 2). En células eucariotas, esta enzima está implicada en la maduración de N-glicanos en N-glicanos de tipo complejo por transferencia de una GlcNAc terminal a Man-5 (Fig. 1). Se predice que esta supuesta secuencia de GnT I, SEQ ID NO 2, es una proteína de tipo II con una parte luminal que comparte altas homologías estructurales con el dominio catalítico de la GnT I de conejo.

Con el fin de determinar si este gen codifica una GnT I activa, se analizó el patrón de glicosilación de las proteínas de *P. tricornutum*.

Se cultivó *Phaeodactylum tricornutum* usando un método de cultivo por lotes estándar utilizando unas bombonas de vidrio que iban de 2 a 10 L en medios de Conway esterilizados (WALNE, L. *Fish Invest Serie II*, vol.25 (4), pp. 1-53, 1966) con agua de mar (salinidad = 3,3-3,4%), se filtró y aireó 1 μm con una mezcla de CO₂ al 2%/aire para mantener el pH en un intervalo de 7,5-8,1. El metasilicato de sodio se añadió a los medios a la concentración final de 40 mg/L.

Se cultivó *Phaeodactylum tricornutum* a 22-23°C bajo iluminación continua (280-350 μmol fotones m^2s^{-1}). El cultivo concentrado (aproximadamente $20 \cdot 10^6$ células/mL) se centrifugó en primer lugar a 5.000 g durante 20 minutos a 4°C y a continuación se liofilizó el sedimento.

Se trituraron 2 gramos de microalgas liofilizadas en presencia de arena en un mortero usando un tampón Tris-HCl 750 mM a pH 8 que contenía 15% (p/v) de sacarosa, 2% (v/v) de β -mercaptoetanol y fluoruro de fenilmetilsulfonio 1 mM y luego se centrifugó a 4°C durante 30 minutos a 11.500 g. Las proteínas del líquido sobrenadante se precipitaron luego con sulfato de amonio al 90% durante 2 horas a la temperatura ambiente. El sedimento se solubilizó en agua y luego se dializó frente a agua durante una noche a 4°C. Finalmente, el extracto de proteínas totales se ultracentrifugó a 100.000 g durante 1 hora a 4°C y se volvió a poner en suspensión en el volumen de agua más pequeño, antes de cuantificación de las proteínas y análisis posteriores. La cuantificación de las proteínas se realizó en los extractos de proteínas totales de *Phaeodactylum tricornutum* utilizando el kit de ensayo de proteínas BCA de PIERCE de acuerdo con las instrucciones del fabricante.

Se realizó luego el análisis estructural de glicanos unidos en N a las proteínas de *P. tricornutum* por análisis de transferencia de Western en un extracto de proteínas totales usando sondas específicas para los epítomos de glicano. Para este análisis, se separaron 50 μg de proteínas totales por SDS-PAGE. Se usaron como control proteínas de cebolla. Las proteínas separadas se transfirieron a membranas de nitrocelulosa y se tiñeron con Rojo Ponceau con el fin de controlar la eficacia de transferencia. Se realizó una detección por afinidad utilizando concanavalina A por incubación con la lectina a 25 $\mu\text{g} \cdot \text{mL}^{-1}$ durante 2 horas a temperatura ambiente en TBS-T, complementado con CaCl₂ 1 mM y MgCl₂ 1 mM. Después de lavar con TBS-T complementado con CaCl₂ y MgCl₂ (6 veces, 5 minutos), se detectó la unión de esta lectina usando peroxidasa de rábano silvestre diluida a 50 $\mu\text{g} \cdot \text{mL}^{-1}$, 1 hora a temperatura ambiente en TBS-T complementado con CaCl₂ 1 mM y MgCl₂ 1 mM. Después de lavar con el mismo TBS-T y luego con TBS, se realizó el desarrollo final de las transferencias usando 4-cloro-1-naftol como se ha descrito anteriormente (FITCHETTE *et al.*, *Methods in Molecular Biology*, publicado por Human Press, EE. UU. (Totowa, New Jersey), pp. 317-342, 2006).

También se realizó una inmunodetección usando anticuerpos específicos preparados en el laboratorio núcleo- β (1,2)-xilosa y núcleo- α (1,3)-fucosa (1:1000 en TBS que contenía 1% de gelatina, 2 horas a temperatura ambiente). Después de lavar con TBS-T (6 veces, 5 minutos), se detectó la unión de anticuerpos usando un anticuerpo anti-IgG de conejo de cabra conjugado con peroxidasa de rábano silvestre secundario diluido a 1:3000 en TBS que contenía gelatina al 1% durante 90 minutos a temperatura ambiente (Bio-Rad). El desarrollo final de las transferencias se realizó usando 4-cloro-1-naftol como se ha descrito anteriormente (FITCHETTE *et al.*, antes mencionado, 2006).

Los resultados se presentan en la Figura 2 que muestra detección por afinidad usando concanavalina A (Con A) e inmunodetección utilizando anticuerpos producidos contra los epítomos núcleo- β (1,2)-xilosa (anti-Xyl) y núcleo- α (1,3)-fucosa (anti-Fuc) de proteínas aisladas de cebolleta como control positivo (Pista 1) y de *Phaeodactylum tricornutum* cultivada en condiciones estándares (Pista 2).

Los resultados muestran que las proteínas expresadas en *Phaeodactylum tricornutum* no presentan ninguno de los epítomos núcleo- β (1,2)-xilosa (anti-Xyl) y núcleo- α (1,3)-fucosa (anti-Fuc). No obstante, las proteínas de *Phaeodactylum tricornutum* presentan secuencias de alto contenido en manosa como se reveló por la unión de Con A, una lectina específica para secuencias de alto contenido en manosa. La presencia de secuencias con alto contenido en manosa en las proteínas de *Phaeodactylum tricornutum* se confirmó también por ensayos de desglicosilación usando el péptido N-glicosidasa F (PNGasa F) y endoglicosidasa H (Endo H) (datos no mostrados).

Con el fin de determinar el patrón de glicosilación específico de las proteínas de *Phaeodactylum tricornutum*, dichas proteínas se sometieron a digestión antes del análisis por espectrometría de masas.

Las proteínas totales se sometieron a digestión por tratamientos sucesivos con pepsina y PNGasa A como han descrito anteriormente FITCHETTE *et al.*, (antes mencionado, 2006). Se sometieron a digestión 4 mg de proteínas con 6 mg de pepsina en 2 mL de HCl 10 mM, pH 2,2, a 37°C durante 48 horas. Después de neutralización con

hidróxido de amonio 1 M, la solución se calentó durante 5 minutos a 100°C y se liofilizó. A continuación, se desglicosilaron los glicopéptidos durante una noche a 37°C con PNGasa A (10 mU, BOEHRINGER MANNHEIM) en un tampón de acetato de sodio 100 mM, pH 5,0. Los N-glicanos se purificaron por elución sucesiva a través de una columna AG 50W-X2 (BIO-RAD) y un cartucho C18 (VARIAN) de acuerdo con FITCHETTE *et al.*, (antes mencionado, 2006). Los N-glicanos purificados se marcaron luego con 2-aminobenzamida (2-AB) usando el protocolo optimizado descrito por BIGGE *et al.*, (*Anal Biochem*, vol. 230, pp. 229-238, 1995). Luego, se detectaron los N-glicanos marcados con una luz ultravioleta y se eluyeron usando agua. Los N-glicanos marcados eluidos se liofilizaron finalmente antes de la digestión con exoglicosidasa y el análisis por espectrometría de masas MALDI-TOF. Para la digestión con exoglicosidasa, se desalaron 200 miliunidades de α -manosidasa de canavalia (SIGMA-ALDRICH) por ultrafiltración usando un Centricon y se incubaron durante una noche a 37°C con aproximadamente 50 pmoles de mezcla de N-glicanos marcados con 2-AB. Luego, los N-glicanos marcados con 2-AB se analizaron directamente por espectrometría de masas con ionización mediante desorción por láser asistida por matriz-en tiempo de vuelo (MALDI-TOF) en un instrumento Voyager DE-Pro MALDI-TOF (APPLIED BIOSYSTEMS) equipado con un láser de nitrógeno de 337 nm. Los espectros de masas se realizaron en el modo de extracción retardada del reflector usando como matriz ácido 2,5-dihidroxibenzoico (SIGMA-ALDRICH). La matriz, recién disuelta a 5 mg/mL en acetonitrilo/TFA al 0,1% 70:30 se mezcló con los oligosacáridos solubilizados en agua en una relación 1:1 (v/v). Estos espectros se registraron en un modo positivo, usando un voltaje de aceleración de 20.000 V con un tiempo de retardo de 100 ns. Se suavizaron una vez y se calibraron externamente usando mezclas comercialmente disponibles de péptidos y proteínas (APPLIED BIOSYSTEMS). En este estudio, los espectros se calibraron externamente usando des-Arg¹-bradiquinina (904,4681 Da), angiotensina I (1296,6853 Da), Glu¹-fibrinopéptido B (1570,6774 Da), ACTH₁₈₋₃₉ (2465,1989 Da) e insulina bovina (5730,6087 Da). Se acumularon 1000 disparos de láser para cada espectro y se acumularon varios espectros (entre 5 y 10 espectros) con el fin de obtener una buena relación señal a ruido.

Los resultados en la figura 13 han demostrado que los iones principales (95-97%) para la mezcla resultante de N-glicanos marcados corresponden a iones $(M+Na)^+$ de derivados 2-AB de Hexosa₅₋₉GlcNAc₂. Sin embargo, no se detectaron derivados GlcNAc-Hexosa₅GlcNAc₂.

En consecuencia, los inventores mostraron que *P. tricornutum* no presenta una actividad significativa de GnT I.

Ejemplo 2 Una supuesta GnT I detectada en células de CHO^{AGnT-I} del complemento de la GnT I de *P. tricornutum*

Incluso aunque no fuera identificada ninguna actividad de GnT I detectable en *Phaeodactylum tricornutum*, los inventores trataron de expresar la secuencia de ácidos nucleicos correspondiente a la supuesta secuencia de longitud completa de la GnT I de *P. tricornutum* (Pt54844) (SEQ ID NO 2) en el mutante *Lec1* de CHO, que está mutado en su GnT I endógena. Dicha secuencia de *P. tricornutum* se expresó en el mutante de CHO en fusión con un epítipo V5 para controlar la expresión de la proteína de fusión en los transformantes.

Basándose en la detección de este epítipo con anticuerpos anti-V5, los inventores mostraron que se encontró que la mayoría de los transformantes de CHO expresaban la proteína de fusión V5 (datos no mostrados). Se seleccionaron dos líneas celulares que expresaban la proteína de fusión V5 para el análisis de glicanos unidos en N.

Se aislaron las proteínas de estas líneas, así como de tipo natural, y las células *Lec1* CHO y se liberaron sus glicanos unidos en N por tratamiento con PNGasa F seguido de análisis por espectrometría de masas MALDI-TOF.

Los resultados mostraron que el mutante *Lec1* de CHO acumulaba N-glicanos del tipo de alto contenido en manosa, a diferencia de las células de CHO de tipo natural que presentaban N-glicanos tanto del tipo de alto contenido en manosa como del tipo complejo. Sorprendentemente, las proteínas del mutante *Lec1* de CHO que expresaban la proteína de fusión V5 llevaban tanto el tipo de alto contenido en manosa como un conjunto completo de N-glicanos complejos idénticos al observado en las células de CHO de tipo natural (Figura 14).

Finalmente, estos resultados muestran que incluso si ninguna actividad de GnT I era significativamente detectable en *P. tricornutum*, la expresión del gen Pt54844 era capaz de restaurar la biosíntesis de N-glicanos complejos en células de mamíferos.

En conclusión, el gen Pt54844 codifica una transferasa funcional diseñada como Pt GnT I.

Ejemplo 3 Expresión de GnT I en *Phaeodactylum tricornutum* y validación por RT-PCR

El gen Pt54844 descrito anteriormente (SEQ ID NO 2) se clonó bajo un promotor potenciado del virus del mosaico de la coliflor en un vector basado en pPha-T1 llamado vector BSJ-25. Como control, se construyeron también vectores que comprendían GnT I de *Arabidopsis thaliana* (AtGnT I, At4g38240) y GnT I de células humanas (hGnT I, MGAT1). El BSJ-25 (SEQ ID NO 79) procedía del vector pPha-T1 (Zaslavskaja and Lippmeier, 2000, *J. Phycol.* vol.36 (2), pp. 379-386) reemplazando el promotor FcpA del casete de expresión por el virus del mosaico de la coliflor potenciado doble (CaMV35S)(SEQ ID NO 80) fusionado al péptido señal de la planta. Para esto, pPha-T1 que contenía un casete de resistencia a la bleomicina guiado por el promotor FcpB se sometió a digestión con NdeI

y EcoR1 para eliminar el promotor FcpA. Esta construcción se denominó “pPha-T1-PfcpA delecionada”. El promotor 35S del virus del mosaico de la coliflor potenciado doble del casete de expresión se amplificó por PCR con el cebador directo, CaMV35Sfwd (5'-GAACATATGGTGGATTGATGTGATCTACTCC-3') (SEQ ID NO 61) y el cebador inverso, CaMV35Srev (5'-AATTCTCGAGGAATTCCGGCCGAGG-3') (SEQ ID NO 62) en la construcción PS1 (Kotzer *et al.*, 2004, *J Cell Sci*, vol.117 (Pt 26), pp. 6377-89). La construcción PS1 contenía un promotor 35S del virus del mosaico de la coliflor doble (SEQ ID NO 80), una secuencia Ω del virus del mosaico del tabaco como potenciador de la traducción fusionada al péptido señal de la quitinasa del tabaco de SEQ ID NO 81 (Haseloff *et al.*, 1997, *Proc Natl Acad Sci U S A*, vol. 94 (6), pp. 2122-7; Batoko *et al.*, 2000, *Plant Cell*, vol. 12 (11), pp. 2201-18). El producto de la PCR fue digerido por Nde1 y EcoR1 y luego clonado en “pPha-T1-PfcpA delecionada” para generar el vector BSJ-25.

Resumiendo, el casete de expresión del vector BSJ-25 contenía un promotor de CaMV doble, un potenciador de la traducción, un péptido señal, un sitio de clonación múltiple y el terminador FcpA (Fig 3). El gen GnT I de varios orígenes fue introducido en este vector usando la digestión enzimática desde los sitios de clonación múltiple y las inserciones fueron clonadas entre el péptido señal y la secuencia del terminador.

También se construyeron vectores específicos que expresaban los genes antes mencionados en fusión con la proteína verde fluorescente (GFP) para investigar la localización celular en el aparato de Golgi de las proteínas de fusión correspondientes.

Las construcciones se pueden ver en la Figura 12.

Dichos vectores se usaron para transformar *Phaeodactylum tricornutum*.

La transformación de *Phaeodactylum tricornutum* se realizó como ha sido descrito por Zaslavskaja *et al.*, (2000, *J. Phycol.* Vol.36 (2), pp. 379-386). Se cultivó *P. tricornutum* 1.8.6 (en matraz o placa de agar-agar al 10%) en agua de mar estéril (filtrada por un filtro de 0,22 μ m) enriquecida con 0,1% (v/v) del medio nutritivo (Conway) y 0,1% (v/v) de una solución de sílice (0,4 g/mL). Todos los cultivos se mantuvieron a 20°C bajo iluminación continua en condiciones estériles. Para la transformación genética, se contaron los cultivos de microalgas en fase exponencial de crecimiento y se concentraron por centrifugación, se diluyeron en agua de mar estéril y se inocularon aproximadamente 10^8 células en forma de una placa de 2,5 cm de diámetro sobre la superficie de la placa de agar-agar. Se recubrieron 500 microgramos de microportadores de oro (0,6 μ m de tamaño de partículas) con 1 μ g de DNA del vector en presencia de CaCl₂ 1,25 M y espermidina 20 μ M. La transformación se realizó con el sistema de suministro biolístico de partículas BIORAD PDS-1000/He para el bombardeo de partículas. Los experimentos se realizaron bajo una campana. El bombardeo se realizó a 63,274 (900 psi) y 77,335 kg/cm² (1100 psi) bajo una presión negativa de 27 mm de Hg con una distancia objetivo diferente (6 - 8 cm). Después del bombardeo, las células transformadas se pusieron en suspensión en 600 μ L de agua de mar nutritiva y se cultivaron 1 día bajo iluminación a 20°C. Las células transformadas se extendieron sobre una placa de agar-agar que contenía 100 μ g.mL⁻¹ de zeocina y se incubaron para su cultivo 3 o 4 semanas bajo luz continua.

A continuación, se controló la presencia de transcritos para la GnT I recombinante por PCR con transcripción inversa (RT-PCR). El sedimento de microalgas se puso en suspensión en 1,2 mL de Trizol (INVITROGEN) y se homogeneizó en 2 mL de Lysing Matrix E por 10 pasadas por Fastprep-24 (6,5 m/s, 60 segundos, MP BIOMEDICALS). Se añadieron 300 μ L de cloroformo al líquido sobrenadante, se homogeneizaron enérgicamente y se incubaron 3 minutos a la temperatura ambiente. Después de una centrifugación de 15 minutos a 4°C, la fase acuosa se mezcló con 1 volumen de etanol absoluto y el RNA se purificó con el kit RNEASY MINI (QIAGEN). El RNA purificado se eluyó en 50 μ L de agua exenta de RNasa, se valoró con NanoDrop (THERMO SCIENTIFIC) y se sometió a digestión con RQ1 DNasa (PROMEGA). Después de una segunda purificación con el kit RNEASY MINI y la cuantificación con NanoDrop, el RNA se conservó a -80°C.

La transcripción inversa (RT) se realizó en 1 μ g de RNA purificado. La primera cadena de cDNA fue sintetizada por 200 unidades de M-MLV RT RNasa H minus (PROMEGA) con cebadores oligodT. Se utilizaron 2 μ L de cDNA para la PCR con GoTaq polimerasa (PROMEGA) en 50 μ L. La reasociación se realizó con cebadores específicos de GnT I y secuencia de actina.

Los resultados confirmaron la expresión de los transcritos para la GnT recombinante en los recombinantes seleccionados (datos no mostrados).

El patrón de glicosilación de las proteínas expresadas en dichos transformantes se realizó luego como se ha descrito anteriormente. Simultáneamente, la localización celular de las proteínas de fusión se investigó observando la fluorescencia de la GFP de las proteínas de fusión GnT I-GFP por microscopía confocal.

Se realizó otra serie de ensayos con el fin de demostrar la expresión y localización de la GnT I endógena en *Phaeodactylum tricornutum* transformada.

Ejemplo 4 Expresión de GnT I endógena y validación por RT-PCR y Q-PCR

1) Construcciones que comprenden la secuencia de GnT I

Se realizaron diferentes construcciones que comprendían la secuencia que codifica la GnT I endógena de *Phaeodactylum tricorutum* con el vector de clonación pPHA-T1 construido por Zavlaskaia et al., (2000) para la transformación genética de *Phaeodactylum tricorutum*, incluyendo dicho vector secuencias de los promotores fcpA y fcpB (flucoxantina-clorofila a/c que se une a proteínas A y B) de *P. tricorutum* y el terminador de fcpA. Contiene un casete de selección con el gen *she ble* y un MCS que flanquea al promotor fcpA.

Dichas construcciones están esquematizadas en la figura 4.

El primer casete comprendía la secuencia que codifica la GnT I endógena situada bajo el control de secuencias reguladoras endógenas. En la segunda construcción, la GnT I estaba fusionada a la proteína verde fluorescente (GFP).

La secuencia de GnT I endógena de *Phaeodactylum tricorutum* se clonó en un vector de expresión con secuencias reguladoras de dichas microalgas. La GnT I se clonó sola o fusionada a la proteína verde fluorescente (eGFP). La expresión de dicha proteína de fusión permitía visualizar la expresión y localización de la GnT I en las microalgas.

Los vectores utilizados para la transformación de *Phaeodactylum tricorutum* comprendían también un casete de selección que comprendía un gen de resistencia a la zeocina. Dichos vectores permitían la transformación genética de la cepa Pt186 de *Phaeodactylum tricorutum*.

Los clones obtenidos se aislaron y cultivaron para ser analizados.

La transformación de *Phaeodactylum tricorutum* se realizó como ha sido descrita anteriormente.

2) Cribado de las microalgas transformadas

Los clones obtenidos después de la transformación de la cepa Pt186 de *Phaeodactylum tricorutum* se aislaron en medio de cultivo recién preparado. La inserción de la construcción que comprendía un gen que codifica la GnT I en el genoma de las microalgas fue verificada por amplificación por PCR con un conjunto de cebadores específicos para el transgén (SEQ ID NO 69 para el vector y SEQ ID NO 70 para la GnT I).

La reacción PCR se realizó en un volumen final de 50 µL que consistía en un tampón 1X PCR, 0,2 mM de cada dNTP, 5 µM de cada cebador, 20 ng del DNA molde y 1,25 U de *Taq* DNA-polimerasa (*Taq DNA polymerase*, ROCHE). Se realizaron treinta ciclos para la amplificación del DNA molde. La desnaturalización inicial se realizó a 94°C durante 3 minutos. Cada ciclo subsiguiente consistía en una etapa de fusión a 94°C (1 minuto), una etapa de reasociación a 55°C (1 minuto) y una etapa de extensión a 72°C (1 minuto). Las muestras obtenidas después de la reacción PCR se sometieron a electroforesis en gel de agarosa (1%) teñido con bromuro de etidio.

Los resultados se muestran en la figura 5 y presentan el cribado obtenido para 5 linajes diferentes de *Phaeodactylum tricorutum* potencialmente transformada. La amplificación de una secuencia específica de 120 pb del transgén obtenido para el control positivo (DNA plasmídico) se encontró también en los 5 clones analizados. Los clones de la microalga no integraban por consiguiente la construcción genética que comprendía el gen que codificaba la GnT I endógena.

3) Cribado de microalgas que expresan la GnT I

a) Análisis de la expresión de GnT I por RT-PCR

Se realizó un análisis transcriptómico de clones transformados con la secuencia que codifica la GnT I de *Phaeodactylum tricorutum*. Se purificaron los RNA totales de los diferentes clones. El cDNA correspondiente se sintetizó y analizó por amplificación por PCR con transcripción inversa de acuerdo con las instrucciones descritas en el Ejemplo 3 con cebadores específicos para el transgén (SEQ ID NO 71 y SEQ ID NO 72) y con el gen constitutivo H4 (SEQ ID NO 73 y SEQ ID NO 74).

Los resultados de las amplificaciones realizadas con los cebadores específicos del gen constitutivo H4 se presentan en la figura 6. La ausencia de amplificación en las muestras de RT permite validar que las muestras no contenían DNA genómico.

Las amplificaciones realizadas en las muestras RT+ permitieron obtener amplicones de 150 pb que eran homogéneos entre todos los clones *Pt*-GnT I. Este estudio permitió por tanto validar las condiciones experimentales.

A continuación, se analizaron los clones *Pt*-GnT I con cebadores que permitieran destacar la expresión de GnT I (véase la figura 7).

La figura 7 representa los resultados de la RT-PCR obtenidos de la expresión de GnT I en los clones transformados de *P. tricornutum*. La ausencia de amplificación en las muestras de RT valida la calidad de las muestras. Además, el fragmento de amplificación de 212 pb con una mayor intensidad que el de la cepa de tipo natural se obtiene a partir de los clones 1 y 3. Esta diferencia de señal corresponde a una sobreexpresión de GnT I. Por tanto, los inventores fueron capaces de identificar clones de microalgas que sobreexpresaban la GnT I. Se realizó un análisis más afinado de la expresión de GnT I por PCR cuantitativa (Q-PCR).

b) Análisis de la expresión de GnT I por Q-PCR

Se seleccionó por RT-PCR un linaje de *P. tricornutum* transformada con una construcción que comprendía el gen GnT I. A continuación, se analizó la expresión de GnT I por Q-PCR como ha descrito Siaut et al., 2007, *Molecular toolbox for studying diatom biology in Phaeodactylum tricornutum*, Gene 406 (1-2): 23-35, utilizando cebadores dirigidos al transgén (SEQ ID NO 75 y SEQ ID NO 76) y al gen constitutivo H4 (SEQ ID NO 77 y SEQ ID NO 78).

La figura 8A corresponde a una curva de fusión. Permite verificar que ha sido amplificado solo un producto de la PCR. La presencia de un pico muestra que se obtuvo solo un fragmento de amplificación. La temperatura de disociación de 81,2°C muestra una fuerte especificidad de la asociación amplicón-cebadores.

La figura 8B presenta el perfil de detección de la GnT I en las células transformadas (*Pt*-GnT I) y las células de tipo natural (*Pt*-wt). Para cada cepa, se analizaron 3 diluciones de cDNA. Los resultados se expresaron en número de ciclos en comparación con los valores relativos de detección del amplicón de GnT I.

Los 3 sigmoides correspondientes a los controles negativos (cDNA- y agua) muestran una detección tardía, que es indicativo de una amplificación no significativa de la GnT I.

El análisis del linaje transformado con GnT I mostró una detección que era significativamente más temprana que la de las células de tipo natural. Los datos obtenidos se normalizaron con un gen constitutivo H4. La normalización en comparación con el gen constitutivo usando el método Ct comparativo, también llamado $2^{-(\Delta\Delta Ct)}$ (método $\Delta\Delta Ct$) reveló una diferencia clara de perfil de expresión entre la cepa que expresaba el transgén de GnT I y las células de tipo natural ($\Delta\Delta Ct = 7$).

25 Ejemplo 5 Demostración de la expresión y localización de la proteína GnT I-eGFP

1) Por microscopía confocal

Los clones de las microalgas *P. tricornutum* que se transformaron con una secuencia que codificaba una proteína de fusión GnT I-eGFP se analizaron por microscopía confocal con parámetros estándares. La figura 9 presenta los resultados de dicho análisis.

Las observaciones realizadas muestran claramente una diferencia de localización de la fluorescencia entre las microalgas que expresan una eGFP citosólica y la proteína de fusión GnT I-eGFP. En ese caso, la fluorescencia está mucho más localizada y se presenta en forma de una concavidad o una judía. Este marcaje es típico del aparato de Golgi en microalgas *P. tricornutum*. Los inventores demostraron por tanto que la GnT I endógena se expresaba y estaba localizada en el compartimento celular adecuado.

2) Por análisis después de tratamiento del cultivo con brefeldin A y microscopía de epifluorescencia

Con el fin de validar los resultados, los inventores analizaron los clones que expresaban la proteína de fusión GnT I-eGFP después de un tratamiento del cultivo de las microalgas transformadas con brefeldin A en comparación con la cepa de microalgas de *P. tricornutum* que expresaba eGFP en el cloroplasto, en el retículo endoplásmico o en el aparato de Golgi después de un cultivo con brefeldin A (Figura 10).

La brefeldin A es una molécula que presenta la particularidad de dismantelar la estructura del aparato de Golgi en las células. Por tanto, es posible confirmar la localización de la proteína de fusión que se produce en las microalgas *P. tricornutum*. Las células se incubaron con brefeldin A 50 μ M y las observaciones se realizaron con una microscopía de epifluorescencia en un periodo de 48 horas.

La figura 10 presenta los resultados que se obtuvieron. No se observó ninguna variación de la fluorescencia en las microalgas *P. tricornutum* que expresaban eGFP en el cloroplasto o en el retículo endoplásmico después de un cultivo con brefeldin A. En las microalgas que expresaban la proteína de fusión GnT I-eGFP, aparecieron rápidamente pequeñas vesículas fluorescentes (después de 5 horas en la imagen), lo que es indicativo del efecto de la brefeldin A sobre el aparato de Golgi. Dicho efecto se hizo cada vez más visible después de 24 horas y 48 horas de cultivo con un marcaje que se extendía a toda la célula. El marcaje de Golgi al que estaba dirigido casi desapareció.

Este análisis permitió validar que la proteína de fusión GnT I-eGFP se expresa y está localizada en el aparato de Golgi de las microalgas *P. tricornutum* transformadas.

3) Análisis de inmunotransferencia por transferencia de Western

Se recogieron partes alícuotas de células de tipo natural y transformadas de un cultivo de *P. tricornutum* en fase exponencial de crecimiento y se separaron las células del medio de cultivo por centrifugación (10 minutos, 2150 g, 20°C). Los sedimentos celulares se volvieron a poner en suspensión en Tris-HCl 0,15 M a pH 8, sacarosa al 15%, SDS al 0,5%, PMSF 1 mM, coctel inhibidor de proteasas al 1% (SIGMA) y se trataron con ultrasonidos durante 30 minutos. Las suspensiones celulares obtenidas se centrifugaron (60 minutos, 15000 g, 4°C) para eliminar los residuos celulares y los líquidos sobrenadantes correspondían a la fracción intracelular.

Se separaron 10 µL de fracciones intracelulares de células de *Pt-eGFP*, *Pt-GnT I-eGFP* transformada y de tipo natural por SDS-PAGE usando un gel de poliacrilamida al 12%. Las proteínas separadas se transfirieron a una membrana de nitrocelulosa y se tiñeron con Rojo Ponceau con el fin de controlar la eficacia de la transferencia. La membrana de nitrocelulosa se bloqueó durante una noche en leche al 5% disuelta en TBS para inmunodetección. A continuación, se realizó la inmunodetección usando anti-GFP conjugado con peroxidasa de rábano silvestre (Santa Cruz, sc-9996-HRP) (1:2000 en TBS-T que contenía leche al 1% durante 2 horas a la temperatura ambiente). A continuación, se lavaron las membranas con TBS-T (6 veces, 5 minutos, temperatura ambiente) seguido por un lavado final con TBS (5 minutos, temperatura ambiente). El desarrollo final de las transferencias se realizó por un método de quimioluminiscencia.

Los extractos de proteínas totales de las microalgas *P. tricornutum* de tipo natural, de las microalgas de *P. tricornutum* que expresan la proteína eGFP (*Pt-eGFP*) y de las microalgas *P. tricornutum* que expresan la proteína de fusión GnT I-eGFP (*Pt-GnT I-eGFP*) se analizaron por transferencia de Western con anticuerpos anti-GFP que estaban acoplados a peroxidasa.

La figura 11 presenta los resultados obtenidos para la detección de la proteína de fusión GnT I-eGFP por transferencia de Western. La figura muestra una detección alrededor de 28 kDa para la eGFP citosólica y alrededor de 75 kDa para los clones que fueron transformados con una secuencia que codificaba la proteína de fusión GnT I-eGFP, cuyo tamaño teórico corresponde a 77 kDa.

Se demuestra por tanto la presencia de la proteína de fusión en las muestras de proteínas de microalgas transformadas. La ausencia de degradación es una prueba de que la proteína de fusión es estable.

Este análisis demuestra por tanto que es posible expresar la secuencia de GnT I endógena de *P. tricornutum* y que dicha proteína está dirigida al compartimento celular adecuado, es decir al aparato de Golgi.

Ejemplo 6 Patrón de glicosilación de proteínas expresadas en *P. tricornutum* transformada con GnT I endógena

Con el fin de investigar la presencia en proteínas de *P. tricornutum* de N-glicanos de tipo complejo que llevan GlcNAc terminal, los inventores trataron el extracto de proteínas con una β-1,4-galactosiltransferasa, una enzima capaz de transferir un residuo de galactosa a residuos de GlcNAc terminales, y luego analizaron la preparación mediante transferencia por afinidad de proteínas resultantes con RCA 120 o ECA, que son lectinas que se unen específicamente a secuencias de Galβ1-4GlcNAc de acuerdo con una estrategia previamente descrita (Bardor *et al.*, 2003, *Plant Biotech J*, vol.1, 451-462). Alternativamente, el patrón de glicosilación de proteínas expresadas en *P. tricornutum* transformada con GnT I endógena se estudió por métodos que han sido mencionados anteriormente.

Ejemplo 7 Expresión de α-Man II en *Phaeodactylum tricornutum*

Se predijo un gen que codificaría un supuesto α-Man II en el genoma de *P. tricornutum* (Pt52108).

El gen Pt52108 (SEQ ID NO 6) se clonó bajo un promotor potenciado del virus del mosaico de la coliflor en un vector basado en pPha-T1 denominado vector BSJ-25. Como control, también se construyeron vectores que comprendían α-Man II de *Arabidopsis thaliana* (NP_196999) y de célula humana (Q16706 o AAC50302). El BSJ-25 procedía del vector pPha-T1 (Zaslavskaja and Lippmeier, 2000, *J. Phycol.* vol. 36 (2), pp. 379-386) reemplazando el promotor FcpA del casete de expresión por el virus del mosaico de la coliflor mejorado doble (CaMV35S) fusionado al péptido señal de la planta. Para este fin, pPha-T1 que contenía el casete de resistencia a la bleomicina activado por el promotor *FcpB* se sometió a digestión con Nde1 y EcoR1 para eliminar el promotor FcpA. Esta construcción se denominó "pPha-T1-PfcpA deleccionada". El promotor del virus del mosaico de la coliflor 35S mejorado doble del casete de expresión se amplificó por PCR con el cebador directo, CaMV35Sfwd (5'-GAACATATGGTGGATTGATGTGATCTACTCC-3') (SEQ ID NO 61) y el cebador inverso, CaMV35Srev (5'-AATTCTCGAGGAATTCGCCGAGG-3') (SEQ ID NO 62) en la construcción PS1 (Kotzer *et al.*, 2004, *J. Cell Sci.* vol.117 (Pt 26), pp. 6377-89). La construcción PS1 contenía un promotor 35S del virus del mosaico de la coliflor doble, una secuencia Ω del virus del mosaico del tabaco como potenciador de la traducción fusionada al péptido señal de la quitinasa del tabaco (Haseloff *et al.*, 1997, *Proc Natl Acad Sci U S A*, vol.94 (6), pp. 2122-7; Batoko *et al.*, 2000, *Plant Cell*, vol.12 (11), pp. 2201-18). El producto de PCR se sometió a digestión con Nde1 y EcoR1 y luego se clonó en "pPha-T1-PfcpA deleccionada" para generar el vector BSJ-25.

Resumiendo, el casete de expresión del vector BSJ-25 contiene un promotor de CaMV doble, un potenciador de la traducción, un péptido señal, un sitio de clonación múltiple y el terminador *FcpA* (Fig 3).

5 También se construyeron vectores específicos que expresaban los genes antes mencionados en fusión con la proteína verde fluorescente (GFP) para investigar la localización celular en el aparato de Golgi de las proteínas de fusión correspondientes.

Dichos vectores se utilizan para transformar *Phaeodactylum tricorutum*.

La transformación de *Phaeodactylum tricorutum* se realiza como se ha descrito anteriormente en el ejemplo de expresión de GnT I en *Phaeodactylum tricorutum*.

10 La presencia de transcritos para el α -Man II recombinante se controló a continuación por PCR con transcripción inversa (RT-PCR). El sedimento de microalgas se volvió a poner en suspensión en 1,2 mL de Trizol (INVITROGEN) y se homogeneizó en 2 mL de Lysing Matrix E por 10 pasadas por Fastprep-24 (6,5 m/s, 60 segundos, MP BIOMEDICALS). Se añadieron al líquido sobrenadante 300 μ L de cloroformo, se homogeneizaron enérgicamente y se incubaron 3 minutos a la temperatura ambiente. Después de una centrifugación de 15 minutos a 4°C, la fase acuosa se mezcló con 1 volumen de etanol absoluto y los RNA se purificaron por el kit RNEASY MINI (QIAGEN).
15 Los RNA purificados se eluyeron en 50 μ L de agua exenta de RNasa, se valoraron con NanoDrop (Thermo SCIENTIFIC) y se sometieron a digestión con RQ1 DNasa (PROMEGA). Después de la segunda purificación por el kit RNEASY MINI y cuantificación por NanoDrop, los RNA se conservaron a -80°C.

20 La transcripción inversa (RT) se realizó en 1 μ g de RNA purificado. La primera cadena de cDNA se sintetizó con 200 unidades de M-MLV RT RNasa H minus (PROMEGA) con cebadores oligodT. Se utilizaron 2 μ L de cDNA para la PCR con GoTaq-polimerasa (PROMEGA) en 50 μ L. La reasociación se realizó con cebadores específicos de GnT I y secuencia de actina.

Los resultados confirmaron la expresión de transcritos para el α -Man II recombinante en los recombinantes seleccionados (datos no mostrados).

25 A continuación, se determinó el patrón de glicosilación de las proteínas expresadas en dichos transformantes como se ha descrito previamente. Simultáneamente, se investigó la localización celular de las proteínas de fusión observando la fluorescencia de GFP de las proteínas de fusión GnT I-GFP por microscopía confocal.

Ejemplo 8 Patrón de glicosilación de proteínas expresadas en *P. tricorutum* transformada con alfa-Man II endógena

30 Se estudió el patrón de glicosilación de proteínas expresadas en *P. tricorutum* transformada con alfa-Man II endógena por HPLC y espectrometría de masas como se ha mencionado antes.

Ejemplo 9 Inactivación de N-acetilglucosaminidasas en *Phaeodactylum tricorutum* con interferencia de RNA

Por el uso de vectores

35 Este ejemplo se basa en el uso de construcciones que contenían RNA antisentido o RNA de repetición invertida para la inactivación de la expresión de β -N-acetilglucosaminidasa como ha sido descrito por De Riso V. *et al.*, 2009, *Nucleic Acids Research*, pp. 1-12.

40 Para la generación de un vector de β -N-acetilglucosaminidasa, se amplificaron dos fragmentos del cDNA de β -N-acetilglucosaminidasa con cebadores específicos, en donde uno de los fragmentos era más largo que el otro fragmento y también contenía la secuencia total del fragmento más corto. Los dos fragmentos obtenidos se sometieron luego a digestión con enzimas de restricción que también se pueden usar para la linealización del vector.

Para la construcción antisentido, uno de los dos fragmentos obtenidos se insertó en el vector en la orientación antisentido.

Para la construcción de repetición invertida, los fragmentos se ligaron en orientaciones en el mismo sentido y antisentido en el sitio de integración del vector.

45 Dichos vectores se usaron luego juntos o por separado eligiendo un vector para transformar *Phaeodactylum tricorutum*. La transformación de *Phaeodactylum tricorutum* se realizó como se ha descrito en el ejemplo 3.

b) Por el uso de siRNA

El silenciamiento de genes de las β -N-acetilglucosaminidasas también se puede obtener con la transformación de *Phaeodactylum tricorutum* con el uso de los siRNA específicos de al menos un gen que codifica una β -N-

acetilglucosaminidasa en dicha microalga (1 µg para cada siRNA). La transformación de *Phaeodactylum tricornutum* se realizó como se ha descrito en el ejemplo 3.

Los siRNA utilizados para la transformación de *Phaeodactylum tricornutum* tenían las siguientes secuencias:

- siRNA específicos de una primera β-N-acetilglucosaminidasa de la secuencia de aminoácidos SEQ ID NO 9 que está codificada por la secuencia de ácidos nucleicos SEQ ID NO 10

SEQ ID NO 63 siRNA1(β-N-acetilglucosaminidasa SEQ ID NO 10) GGCCAUUCGUUACUAGCAA

SEQ ID NO 64 siRNA2(β-N-acetilglucosaminidasa SEQ ID NO 10) GUGGUUCGUUGGGAAAUGA

SEQ ID NO 65 siRNA3(β-N-acetilglucosaminidasa SEQ ID NO 10) CCGUCUGUGUGAAAUUGGU

- siARN específicos de una segunda β-N-acetilglucosaminidasa de la secuencia de aminoácidos SEQ ID NO 11 que está codificada por la secuencia de ácidos nucleicos SEQ ID NO 12

SEQ ID NO 66 siRNA1(β-N-acetilglucosaminidasa SEQ ID NO 12) CGGUAGUAGUGCUUGUUGU

SEQ ID NO 67 siRNA2(β-N-acetilglucosaminidasa SEQ ID NO 12) CUGCCAUGGAUUAUUGUCA

SEQ ID NO 68 siRNA3(β-N-acetilglucosaminidasa SEQ ID NO 12) GCCUUUGGUCCUGAAGAAA

Ejemplo 10 Expresión y glicosilación de un polipéptido en *P. tricornutum* transformada

Como se ha explicado anteriormente, los polipéptidos de interés para ser expresados y glicosilados en *P. tricornutum* transformada de acuerdo con la invención son proteínas de interés terapéutico. Los inventores eligieron la eritropoyetina para ilustrar la expresión y glicosilación de dichas proteínas en *P. tricornutum* transformada de acuerdo con la invención. Sin embargo, la invención no se limita a dicha eritropoyetina y podría aplicarse a cualquier proteína de interés terapéutico que necesite ser glicosilada para presentar el patrón de glicosilación adecuado.

a) Construcción de expresión para la eritropoyetina

El vector utilizado para la expresión de eritropoyetina comprendía una secuencia de ácidos nucleicos unida operativamente a un promotor, codificando dicha secuencia de ácidos nucleicos la eritropoyetina. Dicho vector contenía preferiblemente un marcador seleccionable distinto de los marcadores seleccionables presentes en el vector BSJ-25 de acuerdo con el ejemplo 3.

b) Transformación de *P. tricornutum*

Las *P. tricornutum* transformadas del ejemplo 3 se transformaron también con un vector de acuerdo con la etapa a).

c) Análisis del patrón de glicosilación de la eritropoyetina expresada y glicosilada

El patrón de glicosilación de la eritropoyetina que se expresó y glicosiló en la *P. tricornutum* transformada de acuerdo con la invención se ha descrito en el ejemplo 1.

LISTA DE SECUENCIAS

<110> UNIVERSITE DE ROUEN
 IFREMER
 ALGENICS
 5 BARDOR, Muriel
 LOUVET, Romain
 SAINT-JEAN, Bruno
 BUREL, Carole
 BAIET, Bérange
 10 CARLIER, Aude
 CADORET, Jean-Paul
 LEROUGE, Patrice

<120> N-glicosilación en *Phaeodactylum tricornutum* transformada
 15
 <130> URO-B-0001 PCT1
 <160> 81
 20 <170> PatentIn versión 3.5

<210> 1
 <211> 444
 25 <212> PRT
 <213> *Phaeodactylum tricornutum*

<400> 1
 Met Arg Leu Trp Lys Arg Thr Thr Ser Gly Val Pro Pro Pro Arg Arg
 1 5 10 15
 Arg Arg Arg Gly Gly Arg Glu Glu Arg Leu Ala Ile Gly Leu Leu Leu
 20 25 30
 Thr Leu Ala Val Val Trp Gly Thr Val Leu Val Gly Thr Leu Val Ala
 35 40 45
 Leu Val Ala Ser Pro Thr Lys Ser Val Pro Thr Phe Pro Thr Val Pro
 50 55 60
 Thr Asp Ser Arg Pro Ser Ala Ala Phe Val Val Ser Asp Ala Pro Gly
 65 70 75 80
 Ala Tyr Glu Ser Pro Leu Leu Val Phe Thr Cys Arg Arg Asp Gln Tyr
 85 90 95
 Leu Arg Glu Thr Leu Arg Asp Ile Trp Asn Tyr Ile Pro Thr Asp Cys
 100 105 110
 Ser Val Gly Cys Pro Leu Val Ile Ser Gln Asp Gly Asn Asp Pro Ala
 115 120 125
 Val Arg Arg Val Val Arg Glu Phe Thr Asp Glu Phe Ala Thr Lys Asn
 130 135 140
 30 Val Pro Val Ile His Trp Thr His Thr Ser Ala Leu Arg Gly Ser Thr
 145 150 155 160

ES 2 659 118 T3

Asn Gly Tyr Gln Ala Leu Ala Ile His Tyr Gly Trp Ala Leu Arg Arg
165 170 175

Val Phe Asp Gly Gln Thr Leu Ser Gly Ser Val His Gly Ala Lys His
180 185 190

Gly Thr Pro Gln Arg Val Ile Ile Leu Glu Glu Asp Leu His Val Ala
195 200 205

Pro Asp Phe Phe Asp Tyr Phe Ala Ala Thr Ala Pro Leu Leu Asp His
210 215 220

Asp Ser Ser Leu Leu Ala Val Ser Ala Phe His Asp Asn Gly Phe Ala
225 230 235 240

His Asn Val Arg Asn Ala Ser Arg Ile Leu Arg Ser Asp Phe Phe Pro
245 250 255

Gly Leu Gly Trp Met Met Asn Arg Arg Leu Trp Val Asp Glu Leu Gln
260 265 270

Ser Lys Trp Pro Gly Gly Tyr Trp Asp Asp Trp Leu Arg Glu Pro Ala
275 280 285

Gln Arg Gln Asp Arg Ala Ile Leu Arg Pro Glu Ile Ser Arg Thr Tyr
290 295 300

His Phe Gly Thr Glu Gly Gly Thr Ser Ser Asn Gln Phe Gly Ser His
305 310 315 320

Leu Ser Lys Ile Leu Leu Asn Arg Glu Thr Val Asp Trp Ser Lys Ala
325 330 335

Ala Asp Leu Glu Ala Gln Leu Arg Pro Glu Val Tyr Asp Pro Ala Tyr
340 345 350

Trp Ala Met Val Gln Ala Ser Thr Leu Thr Tyr Thr Ile Pro Asp Ala
355 360 365

Leu Glu Gln Ala Lys Lys Ser Asn Ala Arg Leu Gln Tyr Thr Thr Ile
370 375 380

Glu Gln Phe Lys Tyr Leu Ala His Lys Leu Lys Leu Met Ala Asp Glu
385 390 395 400

Lys Ala Asn Val Pro Arg Thr Ala Tyr Lys Gly Ile Val Glu Thr Arg
405 410 415

Pro His Gly Ala Asp Tyr Phe Leu Phe Leu Thr Pro Pro Leu Ala Glu
420 425 430

Leu Gln Lys Glu Phe Asp Ile Pro Ser Pro Lys Arg
435 440

5 <210> 2
<211> 1332
<212> DNA
<213> Phaeodactylum tricornutum

10 <400> 2

ES 2 659 118 T3

atgcggttgt ggaaacgtac gacgtcgggt gtgccaccac cacgccgtag acgccgtggt 60
 ggacgggagg agcgtctagc gataggacta ctgctcacgc tggccgtcgt gtggggtacg 120
 gtactcgtcg gtaccctcgt ggcctcgtc gcgtccccc ccaaatccgt cccgacgttt 180
 cccacggtac cgaccgattc acgaccgtcc gccgccttcg tcgtctcggg cgcacccggg 240
 gcgtacgaat cgccttact cgtctttacc tgtcgacgag atcaatattt gcgcgaaaca 300
 ttgcgggata tttggaacta cattccgacg gattgttcgg taggttgtcc cttggtaatt 360
 tcacaagacg gcaacgaccc cgctgtccga cggtgtgtac gcgagtttac cgacgaattc 420
 gccaccaaga acgtccccgt aattcactgg acgcatacga gtgccttgcg cggcagcacg 480
 aacggctacc aagccttggc cattcactac ggttggcct tcgcaagagt ctttgacgga 540
 caaacgctgt ccgggagcgt tcacggtgcg aagcacggga cgccgcaacg ggtcattatc 600
 ctcaagaag atttgacgtg ggcaccggac ttttctgact actttgccgc caccgcgccc 660
 cttttggacc acgattcgtc cctcctggcc gtttcggcct ttcacgacaa cggctttgca 720
 cacaaagtgc ggaacgcatc cagaatactg cggctccgact ttttccccgg tctcggatgg 780
 atgatgaacc gacggttatg ggtggatgag ctccagtcca aatggcccgg cggatactgg 840
 gacgactggt tgcgggaacc ggcacaacgg caagatcgcg caattctaag acccgaaatc 900
 tcccgcacct accactttgg caccgagggg ggcacgagta gtaaccagtt tgggtcccat 960
 ttatcgaaaa tactgctcaa ccgtgaaaca gtggattgga gcaaggcagc agatctggaa 1020
 gcccaactgc gaccggaagt gtagcaccgg gcctactggg ccatggtgca agcctccacg 1080
 ttgacgtaca ccatcccaga cgcactcgag caggccaaaa aaagtaacgc gagattacag 1140
 tacaccacca ttgagcagtt caagtacctc gtcataaac tgaaactcat ggccgacgaa 1200
 aaggccaatg taccacggac ggcatacaag ggaattgtgg aaacgcgccc gcacggtgcc 1260
 gattactttc tattcctgac gccgccctc gcagagctgc agaaggaatt cgacattccg 1320
 tcaccgaaaa ga 1332

<210> 3

<211> 1047

5 <212> DNA

<213> Phaeodactylumtricornutum

<400> 3

cggtgtgaag caaacgccg tctgtgacga attcgatata gattacgact gttgggagga 60
 agcgacttgg gagtttctgc aggacgacgc gcaccccggt gtcctccagg ccgtgacgca 120

10

ES 2 659 118 T3

gttgcagaag ttctgggagc acgctacgga tgagcccgtḡ acgggatggc ggcccggaca 180
agccgccc aa gaggaagaag ccttgctcac ggccgaacgg gtcctcgaag tggcggcacg 240
gtactttgac gcgcacgctc aggcctttgc cgaactggtc caacagtacc aagcggacgg 300
tagggatacg accgaagccg cgacgcaaca agcactgcaa cgggactttg gcgcccggat 360
gagccaacag gatgccggag aagccgcgct cgagatggaa gacgtctcca tgaagtgttt 420
cgaaaaatcc gtcgccgctc atcagaacaa tcccgtgggt ggacgagcac tcggtatgat 480
gcagatgaaa cagcaacagg atatccaaaa cgccatggcc atggtggatt aatggacagt 540
aatactgtaa attgttaacc gccccaccgt tcctactgta aggattgtcg actcgcgttc 600
gtaccttate ggtgtgtgta actaacataa actgtacacc tcaacccaac aagagcgact 660
tgtcggcacc gtctcgggga ctaggttttc gagaggcggc ctcgttgccc tttgacccaa 720
ccctaccact gtttaggta tcacgagatt tttcgtgccg tcgaccactg cagtgcacta 780
acgtaatgtt agctctactg tagctagcgc gaatctcggc ctcgtccttg gggcacttca 840
ccatggcgca ttggtcggtc ggtaccgag gggcacattt ggttcggttg agtgtccaac 900
gtaaacacgt tcgtgctttc ttctgaggat acgccgtct ggggaagcaa cattcgcac 960
gttccaagca caacagcaca ccttaacacc acacaacaag ctccacaatc cacagcccta 1020
cggaacgaa tcttccggac ctctgcc 1047

<210> 4

<211> 363

5 <212> PRT

<213> Phaeodactylum tricornutum

<400> 4

Tyr Glu Ser Pro Leu Leu Val Phe Thr Cys Arg Arg Asp Gln Tyr Leu
1 5 10 15

Arg Glu Thr Leu Arg Asp Ile Trp Asn Tyr Ile Pro Thr Asp Cys Ser
20 25 30

Val Gly Cys Pro Leu Val Ile Ser Gln Asp Gly Asn Asp Pro Ala Val
35 40 45

Gln Arg Val Val Arg Glu Phe Thr Asp Glu Phe Ala Thr Lys Asn Val
50 55 60

Pro Val Val His Trp Thr His Thr Ser Ala Leu Arg Gly Gly Thr Asn
65 70 75 80

Gly Tyr Gln Ala Leu Ala Ile His Tyr Gly Trp Ala Leu Arg Arg Val
85 90 95

Phe Asp Gly Gln Thr Leu Ser Gly Ser Val His Gly Ala Lys His Gly
100 105 110

10

ES 2 659 118 T3

PRO-B-0001 P211 Histage Sequence DEFOI

Thr Pro Lys Arg Val Ile Ile Leu Glu Glu Asp Leu His Val Ala Pro
 115 120

Asp Phe Phe Asp Tyr Phe Ala Ala Thr Ala Pro Leu Leu Asp His Asp
 130 135 140

Ser Ser Leu Leu Ala Val Ser Ala Phe His Asp Asn Gly Phe Ala His
 145 150 155 160

Asn Val Arg Asn Ala Ser Arg Ile Leu Arg Ser Asp Phe Phe Pro Gly
 165 170 175

Leu Gly Trp Met Met Asn Arg Arg Leu Trp Val Asp Glu Leu Gln Ser
 180 185 190

Lys Trp Pro Gly Gly Tyr Trp Asp Asp Trp Leu Arg Glu Pro Ala Gln
 195 200 205

Arg Gln Asp Arg Ala Ile Leu Arg Pro Glu Ile Ser Arg Thr Tyr His
 210 215 220

Phe Gly Thr Glu Gly Gly Thr Ser Ser Asn Gln Phe Gly Ser His Leu
 225 230 240

Ser Lys Ile Leu Leu Asn Arg Glu Thr Val Asp Trp Ser Lys Ala Val
 245 250 255

Asp Leu Glu Ala Gln Leu Arg Pro Glu Val Tyr Asp Pro Ala Tyr Trp
 260 265 270

Ala Met Val Gln Ala Ser Thr Leu Thr Tyr Thr Ile Pro Asp Ala Leu
 275 280 285

Glu Gln Ala Lys Lys Ser Asn Ala Arg Leu Gln Tyr Thr Thr Ile Glu
 290 295 300

Gln Phe Lys Tyr Leu Ala His Lys Leu Lys Leu Met Ala Asp Glu Lys
 305 310 315 320

Ala Asn Val Pro Arg Thr Ala Tyr Lys Gly Ile Val Glu Thr Arg Pro
 325 330 335

His Gly Ala Asp Asn Phe Leu Phe Leu Thr Pro Pro Leu Ala Glu Leu
 340 345 350

Gln Lys Glu Phe Asp Ile Pro Ser Pro Lys Arg
 355 360

<210> 5

<211> 1256

Á

<212> PRT

<213> Phaeodactylum tricornutum

<400> 5

ES 2 659 118 T3

Met Arg Gln Leu Val Arg Thr Leu Ser Phe Asp Asn Pro Asp Gly Gly
1 5 10 15

Val Trp Lys Gln Gly Trp Asn Val Thr Ala Arg Asp Asp Phe Thr Asn
20 25 30

Ala Gln Pro Leu Gln Val Phe Val Val Pro His Ser His Cys Asp Pro
35 40 45

Gly Trp Ile Lys Thr Phe Asp Asp Tyr Phe Gln Ser Gln Thr Arg Gln
50 55 60

Ile Leu Thr Thr Val Val Gln Ala Leu Gln Arg Asp Gln Arg Arg Lys
65 70 75 80

Phe Ile Trp Ala Glu Ile Ser Tyr Phe Glu Trp Trp Tyr Arg Glu Gln
85 90 95

Lys Asp Asp Thr Arg Lys Val Val Gln Thr Leu Leu Asp Asn Lys Gln
100 105 110

Leu Gln Phe Val Thr Gly Gly Trp Val Gln Pro Asp Glu Ala Asn Ser
115 120 125

Glu Leu Tyr Ala Met Glu Ile Gln Leu Gln Glu Gly His Asp Trp Ile
130 135 140

Arg Gln Thr Phe Gly Asp Ala His Val Pro Lys His Gly Trp Ser Ile
145 150 155 160

Asp Pro Phe Gly Tyr Ser Pro Thr Met Ala Tyr Leu Leu Gln Lys Tyr
165 170 175

Gly Phe Lys Ala Met Leu Ile Gln Arg Val His Tyr Ala Val Lys Lys
180 185 190

Glu Leu Ala Gln Arg Arg His Leu Glu Phe Tyr Trp Arg Gln Thr Trp
195 200 205

Glu Asp Ala Thr Thr Ala Gly Thr His Asp Ile Phe Thr His Val Met
210 215 220

Pro Phe Phe Ser Tyr Asp Val Pro His Thr Cys Gly Pro Asp Pro Ser
225 230 235 240

Val Cys Cys Gln Phe Asp Phe Gln Arg Thr Thr Cys Pro Trp Leu Lys
245 250 255

Ala Pro Gln Thr Ile Thr Ser Lys Asn Val Ala Glu Arg Ala Met Leu

ES 2 659 118 T3

			805						810						815
Gln	Glu	Lys	Glu	Glu	Lys	Arg	Gly	Ala	Gly	Asp	Gln	Gln	Leu	Phe	Val
			820					825					830		
Asn	Asp	Lys	Pro	Lys	Glu	Ser	Ala	Leu	Lys	Gly	Phe	Trp	Asn	Thr	Gly
		835					840					845			
Ile	Ile	Thr	Ser	Asp	Phe	Phe	Gly	Phe	Phe	Arg	Lys	Ala	Ser	Glu	Thr
	850					855					860				
Leu	Pro	Lys	Asp	Gln	Gly	His	Glu	Asp	Gly	Lys	Gln	Leu	Gln	Lys	Met
865					870					875					880
Pro	Gly	Lys	Thr	Arg	Val	Ser	Lys	Gly	Asn	Arg	Arg	His	Arg	His	Pro
				885					890					895	
Pro	Val	Asp	Ser	Phe	Gln	Arg	Lys	Phe	Leu	Val	Asn	Asp	Ala	Asp	Trp
			900					905					910		
Met	Asn	Gly	Gly	Arg	Ser	Tyr	Ala	Ser	Arg	Lys	Thr	Arg	Glu	Ile	Lys
		915					920					925			
Asn	Arg	Arg	Phe	Thr	Met	Lys	Asn	Thr	Asp	Ile	Pro	Phe	Val	Leu	Met
	930					935					940				
Val	Tyr	Lys	Arg	Val	Asp	Tyr	Leu	Lys	Lys	Ala	Ile	Asp	Ser	Val	Arg
945					950					955					960
Arg	Ser	Asp	Phe	Pro	Lys	Ser	Arg	Val	Pro	Leu	Ile	Ile	Ser	His	Asp
				965					970					975	
Gly	Arg	Val	Pro	Glu	Val	Val	Glu	Tyr	Val	Glu	Ser	Leu	Lys	Asp	Glu
			980					985						990	
Phe	Lys	Ile	Ile	Gln	Leu	Ile	His	Pro	His	Ser	Cys	Tyr	Glu	His	Pro
		995					1000					1005			
Asn	Glu	Phe	Pro	Gly	Asp	Asp	Pro	Thr	Leu	Asn	Glu	Gly	Phe	Ala	
	1010					1015					1020				
Gly	Asp	Ser	Tyr	Gly	Asn	Pro	Arg	Ser	Ala	Trp	Ile	Thr	Cys	Cys	
	1025					1030					1035				
Lys	His	His	Phe	Thr	Trp	Met	Leu	His	Thr	Val	Phe	Arg	Arg	Asp	
	1040					1045					1050				
Phe	Thr	Asp	Pro	Ala	Val	Asp	Thr	Phe	Leu	Phe	Leu	Glu	Glu	Asp	
	1055					1060					1065				
Tyr	Ile	Val	Ala	Pro	Thr	Ile	Tyr	Ser	Ala	Val	Ile	Ala	Gly	Leu	

ES 2 659 118 T3

1070 1075 1080

Asn Val Met Glu Asp Met Asp Lys Glu Ile Pro Gly Gly Phe Phe
1085 1090

Gly Leu Gly Met Asp Pro Ser Met Ala Asn Thr Ala Phe Glu Pro
1100 1105 1110

Tyr Tyr Lys Lys Ala Thr Trp Tyr Val Glu Ala Phe Lys Ser Gly
1115 1120 1125

Pro Met Thr Met Asn Arg Asp Met Phe Lys Lys Leu Gln Gln His
1130 1135 1140

Ala Lys Glu Tyr Cys Thr Phe Asp Asp Tyr Asn Trp Asp Trp Ser
1145 1150 1155

Ile Val His Leu Gln Ser Lys Lys Leu Leu Pro Arg Thr Leu Leu
1160 1165 1170

Met Pro Ser Lys Val Leu Ala Lys His Ile Gly Val Lys Glu Gly
1175 1180 1185

Met His Thr Asn Lys Ser Phe Gly Lys Asp Phe Ser Asp Leu Phe
1190 1195 1200

Pro Asn Tyr Ala Pro Arg Arg Glu Gln Leu Thr Thr His Phe Ala
1205 1210 1215

Glu Tyr Arg Phe Thr Gly Asn Thr Ala Ala Ala Leu His Thr Glu
1220 1225 1230

Phe Asn Pro Gly Tyr Gly Gly Trp Gly His Pro Lys Asp His Glu
1235 1240 1245

His Cys Met Lys Val Leu Gln Ser
1250 1255

<210> 6

<211> 3771

5 <212> DNA

<213> Phaeodactylum tricornutum

<400> 6

atgcgccaat tggttcggac cttgagcttt gacaatcccc atggaggcgt ctggaaacag 60

ggatggaacg tcacggcacg ggacgacttt accaacgctc agccgctaca ggtattcgtc 120

gtcccgact cccactgtga tccgggatgg atcaagacct ttgacgatta cttccaatcg 180

cagactcgcc aaatccttac caccgtcgtc caggctctcc aacgagacca acgccgcaaa 240

ttcattctgg ccgaaatttc ctactttgaa tgggtgtacc gcgaacaaaa ggacgacact 300

cgaaaagtgg tccaacgct ccttgataat aagcagctgc aattcgtcac gggcggatgg 360

10

ES 2 659 118 T3

gtacaaccg acgaggccaa ttcggaatta tacgctatgg aaattcaact gcaggaaggg 420
 catgattgga ttcgacaaac gttcggggac gcccattgcc ccaaacacgg atggtccatt 480
 gatccctttg gctattcccc cactatggcc taccttttgc aaaaatacgg attcaaggcc 540
 atgctcattc aacgggtaca ctacgccgtc aagaaggaac tcgcgcaacg ccgtcacttg 600
 gaattttatt ggcggcaaac ctgggaggac gccaccacgg cgggcactca cgatattttc 660
 acccacgta tgccgttctt ttcctacgac gtaccgcata cgtgcggacc agatccctcg 720
 gtctgctgtc aatttgactt ccaacgcact acgtgtccgt ggctcaaggc gccgcaaacc 780
 atcacctcca aaaacgtggc cgaacgagca atgctcttgt tggatcagta tcgaaaaaag 840
 gcggctctct accggtccaa cgtagtctcg gcaccactgg gagatgactt tcgctacctg 900
 acggctcagg aagccgaggc ccagtacacc aattaccaag caatctttga ttacgtcaac 960
 gccaacctgc cgaatgttaa aatgcaattt ggaactttgt cggactatct tgatgccgtg 1020
 gtgggttcgt ttgacacacc catactacaa ggatccttct ttacctactc cgacattgat 1080
 caagactact ggtcggggta ttttacttcc cgtgtttccg acaaggcctt gggtcgttgg 1140
 ttggaacgcg tcttgactc tgcaacgcag atgggtgcat cgaacaaga cctccaggct 1200
 ccccgcggg ccctgtccct ctttcagcat catgacggcg tcaactgggac cgccaaaact 1260
 catgtacacg aagattacgc ccgccaaatg atggacgcca ttcataccac cgaggactgg 1320
 atgctgctgt cgattcatca acagtcagg actgaaactgc aacccttgtt gaccgcggat 1380
 actacggccg gcgcaatcca accgtgctgg gtcgcgcccg aaccacggac catgcccgaa 1440
 aatgcgtgtg aatcggagta cactgtagat tccgcagcgt cgtcgcgcc agcccttgta 1500
 gcggtctaca accgcgtgtc gacatcgcaa cactgtgta acgtggttgt gcccgccaa 1560
 aagctacgaa ctgccacact accctgtgaa ctccctggtc ccacgtcagt gtcacaaaca 1620
 aagtttgtct tccatcccga gacggggctt atgetagaac ccgtcaagga ggagtggaag 1680
 gtatggaag tcaagaagg aggagcatac ttgttcttcc ctggacaatt aaggtcttat 1740
 gagctaaca aacacgatgt gattatagag gacggtgat acgttgtgtc gactgagttc 1800
 tggaaagcga ctgtggtgga acgcgagata ccgacagatt ttggtatcag ctgcacggtg 1860
 atcgatttcg tttacgaaac taccttgatt gagggtaatc gcgaatggtt tgtgcgtttt 1920
 tccggtaatg tcgcgaaaca cggcattttc cacacggact tgaacggctt caacttcgat 1980
 acccattatt tccgagcggg tatgccatt caatcgcaag tttttccat gccgaccatg 2040
 tcggccatac aagacgacca aacacgcttg actgttttga gtgaacacgc gcaagggtgc 2100
 gctagcttgc aggatggcgc cattgatggt tggttggatc gtcgattaga tcaggacgat 2160
 gaccgaggtc taggccaggg tatatcgac aatagaccaa cccggacgcg acttcgagtt 2220
 gtggtggaac gagaaacgtt caatgtgcag aaggagttt acgtcacgcc actagtacgc 2280
 aggacctgg atgagcttca acatcctctt gtactttttg gaaagcatgt taaaaagtcg 2340
 atcgaccttg ttgccgatcc gtggaagcac cgatccgatt cagaagcccc acaagcacga 2400

ES 2 659 118 T3

agagaaaggc aacgtcaaca acgcgcgcga gagcggcaac ggcagctgca agaaaaagaa 2460
gaaaaagagg gcgctggtga ccagcaacta ttcgtgaatg ataaaccaa agagtcagct 2520
ttgaagggtt tttggaatac tggaatcatt acgtctgact tttttggctt tttccgcaa 2580
gcaagtgaaa ctcttcctaa agatcaaggt catgaagacg gcaaacaact gcaaaaaatg 2640
ccaggtaaaa cgagagtttc caaaggaaat cgacgccatc gacatccacc ggtcgattct 2700
tttcagcgca agttcttggc caacgacgct gactggatga atggggggag atcgtacgca 2760
tcgcgcaaaa cccgtgaaat caagaatcgt cgctttacaa tgaagaacac ggacatcccg 2820
ttcgtattga tggtttataa gcgcgtggat tttttgaaa aggcgattga ttcggtccgt 2880
cgatccgact ttccaaagtc tcgtgtgcct cttatcattt cgcacgatgg acgagtgccg 2940
gaagtcggtg agtatgtcga atcgcgtgaaa gacgagtcca aaattatcca actcattcat 3000
ccgcatctt gctacgagca tccaacgag tttcctgggg atgatccac actcaacgaa 3060
ggctttgctg gagatagcta tggtaatcca cggagtgcgt ggatcacctg ctgcaaacat 3120
cattttactt ggatgcttca cactgtcttt cgtcgggact ttacggacc agcagtggat 3180
acatttttgt ttctcgaaga agattacatc gtggctccta cgatttattc cgccgttatt 3240
gctgggttga atgttatgga agacatggac aaggagattc cgggcggctt cttcggtttg 3300
ggatggatc cgagtatggc gaatactgcc tttgagcctt actacaagaa ggcgacgtgg 3360
tatgtcgaag cattcaagtc aggtccgatg acgatgaacc gggacatggt caaaaagctc 3420
caacaacatg ctaaggagta ctgtacgttc gacgattaca attgggattg gtccattgtc 3480
catctacaaa gtaaaaagt actaccacgg actcttctca tgccaagcaa agtcctagca 3540
aaacatattg gtgtcaagga gggatgcac acgaacaaat cttttgaaa agacttttca 3600
gatttatttc ccaactatgc ccctcgtcgc gagcagctga ccacgcattt tgcggaatac 3660
cgttttaccg ggaacaccgc ggcagcttta cacaccgaat tcaatcccgg ttatggtggt 3720
tggggacatc cgaaagacca cgagcactgc atgaaagtgc tacagtcgta a 3771

<210> 7

<211> 1000

5 <212> DNA

<213> *Phaeodactylum tricornutum*

<400> 7

tacaattatc gagcgcactc tcgagagcgc agtagacgga ggggtgaccg tcactaccgg 60
aaattgtttt gggccccctg tggtagacc aagactcaag agaagaacag aaaacacagt 120
acgcagcata cgcgacactt tgtcgatgaa ctgcaactac aggctaccct cgcaagttat 180
ctctaaggc agcaatcctg tgacacttgc tcggttgttc ccgtctctca atccaatccg 240
cacggaagga tcgatatcgg attggatttc ggcagataag acgaccgtac agtgagtgc 300
gggtcgacac tgtccagcct tgttgttgtt ttccatccca ctcgttctct agcagattca 360
ctcttgatc tacgcaagca acggcttata acctaaatta ctactccaaa tgattcgag 420

10

ES 2 659 118 T3

```

tcgcaccggt gccaacggca ctaacggcgc ggtaccagcc cgtacgagtc cgaccaatc 480
ctcctccgcg ttgctcgaat ctccctccac cacaacgtcg tcagtgacgt ccaccaactaa 540
cctaaacagt aaccacaact ccgtttccgt cagccgattc tcccgttggc ttttggcacc 600
atcccgggcg tccccgtcgc cgccaacact ctcacctccg gcgcgctgg cacgcaatc 660
acgtctccga cggcgacccc tggcctcggc ttccgccgcc aaggcgtcca ccttgcgtca 720
ccatctcggg tgggttgcct cctcgcctct cgtattactc gcactcttta cgctcgtgct 780
ttttgtgagt atttatttac tcgaaaccac tgccaatgcc ggcgtctcta cggcggaggg 840
ctccaatgtg gtatcccccc gtcgtgcctt ttggcaacaa cacttgcgac ggctgcacc 900
ttccgacaca gccgtggtac caactgatca cgccatgcat cgacctccg gtgctaccat 960
cggaaacaaa atcagtagca aaccgaaga ctggattgat 1000

```

<210> 8

<211> 1220

5 <212> PRT

<213> Phaeodactylum tricornutum

<400> 8

Gln Val Phe Val Val Pro His Ser His Cys Asp Pro Gly Trp Ile Lys
1 5 10 15

Thr Phe Asp Asp Tyr Phe Gln Ser Gln Thr Arg Gln Ile Leu Thr Thr
20 25 30

Val Val Gln Ala Leu Gln Arg Asp Gln Arg Arg Lys Phe Ile Trp Ala
35 40 45

Glu Ile Ser Tyr Phe Glu Trp Trp Tyr Arg Glu Gln Lys Asp Asp Thr
50 55 60

Arg Lys Val Val Gln Thr Leu Leu Asp Asn Lys Gln Leu Gln Phe Val
65 70 75 80

Thr Gly Gly Trp Val Gln Pro Asp Glu Ala Asn Ser Glu Leu Tyr Ala
85 90 95

Met Glu Ile Gln Leu Gln Glu Gly His Asp Trp Ile Arg Gln Thr Phe
100 105 110

Gly Asp Ala His Val Pro Lys His Gly Trp Ser Ile Asp Pro Phe Gly
115 120 125

Tyr Ser Pro Thr Met Ala Tyr Leu Leu Gln Lys Tyr Gly Phe Lys Ala
130 135 140

Met Leu Ile Gln Arg Val His Tyr Ala Val Lys Lys Glu Leu Ala Gln
145 150 155 160

10

ES 2 659 118 T3

Arg Arg His Leu Glu Phe Tyr Trp Arg Gln Thr Trp Glu Asp Ala Thr
 165 170 175
 Thr Ala Gly Thr His Asp Ile Phe Thr His Val Met Pro Phe Phe Ser
 180 185 190
 Tyr Asp Val Pro His Thr Cys Gly Pro Asp Pro Ser Val Cys Cys Gln
 195 200 205
 Phe Asp Phe Gln Arg Thr Thr Cys Pro Trp Leu Lys Ala Pro Gln Thr
 210 215
 Ile Thr Ser Lys Asn Val Ala Glu Arg Ala Met Leu Leu Leu Asp Gln
 225 230 235 240
 Tyr Arg Lys Lys Ala Ala Leu Tyr Arg Ser Asn Val Val Leu Ala Pro
 245 250 255
 Leu Gly Asp Asp Phe Arg Tyr Leu Thr Ala Gln Glu Ala Glu Ala Gln
 260 265 270
 Tyr Thr Asn Tyr Gln Ala Ile Phe Asp Tyr Val Asn Ala Asn Leu Pro
 275 280 285
 Asn Val Lys Met Gln Phe Gly Thr Leu Ser Asp Tyr Phe Asp Ala Val
 290 295 300
 Val Gly Ser Phe Asp Thr Pro Ile Leu Gln Gly Ser Phe Phe Thr Tyr
 305 310 315 320
 Ser Asp Ile Asp Gln Asp Tyr Trp Ser Gly Tyr Phe Thr Ser Arg Val
 325 330 335
 Ser Asp Lys Ala Leu Gly Arg Trp Leu Glu Arg Val Leu Tyr Ser Ala
 340 345 350
 Thr Gln Met Gly Ala Ser Lys Gln Asp Leu Gln Ala Pro Arg Arg Ala
 355 360
 Leu Ser Leu Phe Gln His His Asp Gly Val Thr Gly Thr Ala Lys Thr
 370 375 380
 His Val His Glu Asp Tyr Ala Arg Gln Met Met Asp Ala Ile His Thr
 385 390 395 400
 Thr Glu Asp Trp Met Leu Arg Ala Ile His Gln Gln Tyr Gly Thr Glu
 405 410 415
 Leu Gln Pro Leu Leu Thr Ala Asp Thr Thr Ala Gly Ala Ile Gln Pro
 420 425 430

ES 2 659 118 T3

Cys Trp Val Ala Pro Glu Pro Arg Thr Met Pro Glu Asn Ala Cys Glu
 435 440 445
 Ser Glu Tyr Thr Val Asp Ser Ala Ala Ser Ser Pro Pro Ala Leu Val
 450 455 460
 Ala Val Tyr Asn Pro Leu Ser Thr Ser Gln His Cys Gly Asn Val Val
 465 470 475 480
 Val Pro Gly Gln Lys Leu Arg Thr Ala Thr Leu Pro Cys Glu Leu Pro
 485 490 495
 Gly Pro Thr Ser Val Ser Gln Thr Lys Phe Val Phe His Pro Glu Thr
 500 505 510
 Gly Leu Met Leu Glu Pro Val Lys Glu Glu Trp Lys Val Trp Lys Val
 515 520 525
 Lys Lys Gly Gly Ala Tyr Leu Phe Phe Pro Gly Gln Leu Arg Ser Tyr
 530 535
 Glu Leu Thr Lys His Asp Val Ile Ile Glu Asp Gly Gly Tyr Val Val
 545 550 555 560
 Ser Thr Glu Ser Trp Lys Arg Thr Val Val Glu Arg Glu Ile Pro Thr
 565 570 575
 Asp Phe Gly Ile Ser Ser Thr Val Ile Asp Phe Val Tyr Glu Thr Thr
 580 585 590
 Leu Ile Glu Gly Asn Arg Glu Trp Phe Val Arg Phe Ser Gly Asn Val
 595 600 605
 Ala Asn Asn Gly Ile Phe His Thr Asp Leu Asn Gly Phe Asn Phe Asp
 610 615 620
 Thr His Tyr Phe Arg Ala Asp Met Pro Ile Gln Ser Gln Val Phe Pro
 625 630 635 640
 Met Pro Thr Met Ser Ala Ile Gln Asp Asp Gln Thr Arg Leu Thr Val
 645 650 655
 Leu Ser Glu His Ala Gln Gly Ala Ala Ser Leu Gln Asp Gly Ala Ile
 660 665 670
 Asp Val Trp Leu Asp Arg Arg Leu Asp Gln Asp Asp Asp Arg Gly Leu
 675 680 685
 Gly Gln Gly Ile Ser Asp Asn Arg Pro Thr Arg Thr Arg Leu Arg Val
 690 695 700

ES 2 659 118 T3

Val Val Glu Arg Glu Thr Phe Asn Val Gln Lys Glu Phe Asp Val Thr
 705 710 715 720
 Pro Leu Val Arg Arg Thr Trp Asp Glu Leu Gln His Pro Leu Val Leu
 725 730 735
 Phe Gly Lys His Val Lys Lys Ser Ile Asp Leu Val Ala Asp Pro Trp
 740 745 750
 Lys His Arg Ser Asp Ser Glu Ala Arg Gln Ala Arg Arg Glu Arg Gln
 755 760 765
 Arg Gln Gln Arg Ala Arg Glu Arg Gln Arg Gln Leu Gln Glu Lys Glu
 770 775 780
 Glu Lys Arg Gly Ala Gly Asp Gln Gln Leu Phe Val Asn Asp Lys Pro
 785 790 795
 Lys Glu Ser Ala Leu Lys Gly Phe Trp Asn Thr Gly Ile Ile Thr Ser
 805 810 815
 Asp Phe Phe Gly Phe Phe Arg Lys Ala Ser Glu Thr Leu Pro Lys Asp
 820 825 830
 Gln Gly His Glu Asp Gly Lys Gln Leu Gln Lys Met Pro Gly Lys Thr
 835 840 845
 Arg Val Ser Lys Gly Asn Arg Arg His Arg His Pro Pro Val Asp Ser
 850 855 860
 Phe Gln Arg Lys Phe Leu Val Asn Asp Ala Asp Trp Met Asn Gly Gly
 865 870 875 880
 Arg Ser Tyr Ala Ser Arg Lys Thr Arg Glu Ile Lys Asn Arg Arg Phe
 885 890 895
 Thr Met Lys Asn Thr Asp Ile Pro Phe Val Leu Met Val Tyr Lys Arg
 900 905 910
 Val Asp Tyr Leu Lys Lys Ala Ile Asp Ser Val Arg Arg Ser Asp Phe
 915 920 925
 Pro Lys Ser Arg Val Pro Leu Ile Ile Ser His Asp Gly Arg Val Pro
 930 935 940
 Glu Val Val Glu Tyr Val Glu Ser Leu Lys Asp Glu Phe Lys Ile Ile
 945 950 955 960
 Gln Leu Ile His Pro His Ser Cys Tyr Glu His Pro Asn Glu Phe Pro
 965 970 975

ES 2 659 118 T3

Gly Asp Asp Pro Thr Leu Asn Glu Gly Phe Ala Gly Asp Ser Tyr Gly
 980 985 990

Asn Pro Arg Ser Ala Trp Ile Thr Cys Cys Lys His His Phe Thr Trp
 995 1000 1005

Met Leu His Thr Val Phe Arg Arg Asp Phe Thr Asp Pro Ala Val
 1010 1015 1020

Asp Thr Phe Leu Phe Leu Glu Glu Asp Tyr Ile Val Ala Pro Thr
 1025 1030 1035

Ile Tyr Ser Ala Val Ile Ala Gly Leu Asn Val Met Glu Asp Met
 1040 1045 1050

Asp Lys Glu Ile Pro Gly Gly Phe Phe Gly Leu Gly Met Asp Pro
 1055 1060 1065

Ser Met Ala Asn Thr Ala Phe Glu Pro Tyr Tyr Lys Lys Ala Thr
 1070 1075 1080

Trp Tyr Val Glu Ala Phe Lys Ser Gly Pro Met Thr Met Asn Arg
 1085 1090 1095

Asp Met Phe Lys Lys Leu Gln Gln His Ala Lys Glu Tyr Cys Thr
 1100 1105 1110

Phe Asp Asp Tyr Asn Trp Asp Trp Ser Ile Val His Leu Gln Ser
 1115 1120 1125

Lys Lys Leu Leu Pro Arg Thr Leu Leu Met Pro Ser Lys Val Leu
 1130 1135 1140

Ala Lys His Ile Gly Val Lys Glu Gly Met His Thr Asn Lys Ser
 1145 1150 1155

Phe Gly Lys Asp Phe Ser Asp Leu Phe Pro Asn Tyr Ala Pro Arg
 1160 1165 1170

Arg Glu Gln Leu Thr Thr His Phe Ala Glu Tyr Arg Phe Thr Gly
 1175 1180 1185

Asn Thr Ala Ala Ala Leu His Thr Glu Phe Asn Pro Gly Tyr Gly
 1190 1195 1200

Gly Trp Gly His Pro Lys Asp His Glu His Cys Met Lys Val Leu
 1205 1210 1215

Gln Ser
 1220

<210> 9

<211> 1085

<212> PRT

<213> Phaeodactylum tricornutum

<400> 9

Á

ES 2 659 118 T3

Met Gln Ser Ser Leu Arg Ala Arg Arg Gly Val Ala Pro Arg Arg Lys
 1 5 10 15

Ala Ala Val Gly Arg Tyr Gln Leu Val Leu Ala Ile Thr Leu Ser Val
 20 25 30

Ala Leu Thr Gly Leu Val Thr Thr Gln Phe Val Leu Ser Thr Ser Leu
 35 40 45

Tyr Ser Thr Gln Glu Lys Gln Gln Arg Gln Gln His Lys Ser His Ala
 50 55 60

Lys Gly Pro Asn Pro Ser Asn Leu Arg Thr Asp Ser Phe Pro Arg Glu
 65 70 75 80

Gln Leu His Val Val Val Pro Glu Glu Gln Lys Gln Ser Glu Arg Arg
 85 90 95

Leu Asn Glu Ser Ser Arg Glu His Glu Ser Arg Glu Asn Asp His Glu
 100 105 110

Asp Glu Gln Lys Thr Arg Pro Asp Ser Glu Lys Lys Asp Asp Leu Leu
 115 120 125

Thr Gln Gln Asp Gln Lys Thr Thr Lys Arg Asn Glu His Gly Pro Ala
 130 135 140

Gln Glu Gly Ile Arg Lys His Lys Thr Asp His Glu Ala Pro Glu Lys
 145 150 155 160

Val Glu Ser Pro Val Asp Glu Asp His Glu Val Gln Lys Ala His Arg
 165 170 175

Glu Thr Val Gln Lys Phe Val Asp Asp Arg Arg Val Arg Leu Lys His
 180 185 190

Arg Ile Ser Arg Pro Lys Val Ala Arg Ala Ala Ser Pro Pro Glu Val
 195 200 205

Glu Pro Gln Ile Glu Val Ala Pro Pro Pro Glu Lys Arg Pro Tyr Asp
 210 215 220

Ile Leu Asp Asp Pro Leu Gln Asn Pro Asp Phe Asn Lys Pro Ser Lys
 225 230 235 240

Pro Leu Asn Phe Thr Ala Ala Val Pro Tyr Leu Gly Val Leu Ile Asp

ES 2 659 118 T3

	515					520						525			
Thr	Trp	Trp	Tyr	Asp	Arg	Asn	Ile	Leu	Gly	Arg	Leu	Leu	Ala	Val	Ala
	530					535					540				
Leu	Gly	Ala	Arg	Asn	Glu	Thr	Leu	Pro	Lys	Thr	Met	Lys	Asp	Gln	Asp
545					550					555					560
His	Glu	Lys	Met	Val	Leu	Asp	Gln	Tyr	Gln	Val	Phe	Cys	Asp	Gln	Leu
				565					570					575	
Gly	Tyr	Ser	Gln	Ala	Ile	Cys	Glu	Thr	Asn	Gly	Gly	Pro	Ile	Ile	Pro
			580					585					590		
Thr	Pro	Glu	Tyr	Lys	Lys	Lys	Trp	Gly	Asp	Gly	Trp	Val	Val	Trp	Lys
		595					600					605			
Ala	His	Ile	Cys	Glu	Arg	Met	Thr	Thr	Thr	Glu	Val	Thr	Lys	Ala	Met
	610					615					620				
Arg	Pro	Arg	Ser	Ser	Asp	Arg	Val	Ala	Thr	Gln	Ala	Asn	Ser	Tyr	Phe
625					630					635					640
Trp	Asn	Val	Phe	Gly	Phe	Pro	Ala	His	Thr	His	Thr	Arg	Val	Gly	Gln
				645					650					655	
His	Pro	Thr	Leu	Pro	Asp	Asp	Leu	Gln	Ala	Leu	Gln	Arg	His	Leu	Ile
			660					665					670		
Pro	His	Cys	Gly	Val	Met	Leu	Asp	Thr	Thr	Arg	Ser	Leu	Val	Pro	Ala
		675					680					685			
Asp	Arg	Leu	Gly	Thr	Ile	Leu	Thr	Asp	Thr	Val	Ala	Lys	Leu	Gly	Phe
	690					695					700				
Asn	Ile	Ala	Gln	Leu	Arg	Leu	Val	Ser	Asn	Lys	Gly	Phe	Thr	Phe	Ala
705					710					715					720
Pro	Ser	Ser	Leu	Pro	His	Thr	Val	Gly	His	Ser	Leu	Leu	Ala	Thr	Lys
				725					730					735	
Glu	Ile	Lys	Val	Tyr	Thr	Arg	Ser	Asp	Leu	Met	Gly	Thr	Val	Ala	Lys
			740					745					750		
Ala	Ser	Ala	Val	Gly	Ile	Gln	Met	Ile	Pro	Glu	Ile	Ser	Met	Thr	Thr
		755					760					765			
Gly	Ser	Ala	Gly	Trp	Tyr	Glu	Ser	Gly	Tyr	Leu	Ala	Asn	Cys	Pro	Asn
	770					775					780				
Arg	Leu	Cys	Glu	Ile	Gly	Asp	Ala	Ser	Ile	Asp	Val	Thr	Asn	Pro	Phe

ES 2 659 118 T3

<211> 3339
 <212> DNA
 <213> Phaeodactylum tricornutum

Á <400> 10
 atgcaatcct ccttgcgcgc acgacgtggg gtggctcccc ggagaaaagc tgccgtcggg 60
 cggtatcaac tggcctcgc tattacgttg tccgtcgcct tgaccgggtct cgtgacaact 120
 caatttgtgc tgcgacttc gttgtattcc actcaagaaa agcagcaacg acagcagcac 180
 aagagccacg caaagggacc taatccatcc aatctaagga ccgattcttt cccgagagag 240
 cagttgcacg tagttgtgcc ggaagaacag aaacaatcag agcggcgatt gaacgaaagc 300
 agtcgggaac acgaatctcg agagaatgac catgaagacg agcaaaagac cgggcccgat 360
 tcagagaaga aagatgatct tcttacgcag caggaccaa agacgacgaa acgaaacgaa 420
 catggtcctg cttaggaagg catccgcaag cacaagactg accacgaggc tcccgaaaag 480
 gtggaatcgc cagtgatga ggaccatgaa gtccaaaagg cacacagaga aacagtgcaa 540
 aaattcgtcg acgaccgaag agttcgcctc aagcacagaa tctctcgacc caaagtggcc 600
 agggctgcat caccgcctga ggtggaaccg caaattgagg ttgctccgcc acccgaaaag 660
 cgaccttatg acattctgga tgatcccttg caaatccag atttcaaca accctcgaag 720
 ccgttgaact tcacggccgc tgtaccatat ttgggtgtac tgattgacgg ggggcgcat 780
 ttctttcaa tggactggat gaaacgagcc gttgatcgcc tttctgattt gcgctataat 840
 ttgattcact tgcgtcttac ggacgaccaa gccttcaacg ttctattgga ttctcatccc 900
 gagctcgctt atcccccgcg cgtcaacaac ccacaccagc aagtttggac ggccagcgaa 960
 ttgctgact tgaccgcta cgcgaaatcc aaaggagtaa gcatcatgcc cgaagtcaac 1020
 gttcccggac acgccggtgc ctgggccggg attcctcacc tagtcgtgca ctgtcccgaa 1080
 ttcattctgcc aaagaggcta cggattgcca ctaaagttaa cccatcacga tctcaaacct 1140
 atcttgacga gtatcttgaa ggaagtcgtc gacatttttg atgatccgcc ctttctacat 1200
 ttgggtgggg atgaagtcaa cgtacgtttg ttttccgtgt ggaacttttc tgtgcgttgt 1260
 cggaaagtat aaggcgtttc taaacttttt tttctctacc agatggccgg cccctgtttt 1320
 aacgaagttc gcagtcccgt cttaattac acggctttcg aagttgttct taaagaaatc 1380
 attgccgatg taggctaccc cgaaaagcaa gtggttcggt gggaaatgac cgggcaggct 1440
 aatttgaac gcgctggcgg tgtggaacaa ttttgggagt cgtatccggg agaacggcac 1500
 aaggctgcgg gaccttttt catttcgaac cgtttgtatt ttgatacaaa ccaggatcaa 1560

ES 2 659 118 T3

aatgctacg aagtttggca gaataccgca cggttttatg taaatgatta ccagcccgag 1620
gcagttccaa ccgccattat tgccggcacc tttgagttgt cgacgacttg gtggtacgat 1680
cgcaatattt tgggacgttt gttggctggt gcggtgggtg cccgaaacga aactctacca 1740
aagacgatga aggaccaaga ccatgagaaa atggtgcttg atcaatacca agttttctgc 1800
gaccagttag gatatagcca ggcaatttgc gaaaccaacg gtggcccgat catccctacc 1860
ccggagtaca aaaagaaatg gggtagcggg tgggtagttt ggaaggcgca catctgtgaa 1920
cgcatgacga cgactgaggt aaccaaggca atgcgacccc gctctagcga ccgggtcgcc 1980
acgcaggcga acagctactt ttggaacgta tttggatttc ctgcgcacac acacacgcga 2040
gtgggccagc atccaacggt gccagacgat ctccaggccc tccagcgaca ttaattcct 2100
cattgtggcg ttatgttggg tacgaccaga tccctgggtc cagcggatcg gttgggaacg 2160
atthtgaccg acaccgttgc aaaattgggt ttcaacatag cccagctgcg tttggtcagc 2220
aacaagggct tcacgtttgc tccgagttag ctaccgcata ctgtaggcca ttcgttacta 2280
gcaacgaagg agatcaagggt atacactagg agtgacttga tgggtaccgt tgctaaagcg 2340
agtgcggtgg gaatcctaat gatccccgaa atcagcatga cgacaggaag tgctggttgg 2400
tacgagtcgg gctatctagc gaattgtcca aaccgtctgt gtgaaattgg tgacgcgtcg 2460
attgacgtga cgaaccggtt cttaccacc accgtgtact cgttgatcta cgagttgctg 2520
tccattttca gcagcagtec ctatattcat ctcggttcgg acgagcgtca agacgcggca 2580
gcttgctacc aagaagccaa tcccacgttc cacgcggacg tgggagcgtt cgagcgcaaa 2640
atggtcaaag tcttgagggc gagcgggatt gccaacgatt ctgtactgcg gtacgccaat 2700
tcgcaaggcg aggtgtacag cgaccggacg ggtggcgtca cacactacgg tccggaccat 2760
gcgacggaga ttcctgcgga cgcaccaata tttgtgagtg tggatttgtt gcgggacgat 2820
gggtggacat tgtaccagcg ggtgaaggaa ctcgtatcga aaaagcgtt gggcatcttg 2880
gcggaaatcc gtacgttgac ggctccccgt tgggaaggcc tggagattcc ggaacgtttg 2940
ctggtgtacg ccatggccgt atcggattg cccacgtacg cgaacgcggc ggcactgggc 3000
gagcggatcg gggagctttg ccgggcgtta tcggatcgat tgccgggatt gggctcgtcg 3060
cacgattcgg cgttgcggg tgctcgtctg gggaaaggta cgtttttggc cgacacgagt 3120
acctttgtgc agcagcagtg tcaaatggcg acgtatcccc tgacgcaaca ccacgccaag 3180
ttggtagcac cccggtacaa cgcgaccgag tgggagcaac tgcgcggggc gccccgcgtg 3240
tttccggcgg cgggtcggga tcccagcgg catcacccc tcgtcatcgg gcatggaaaa 3300
tcgggggacg agtccccggt cgagtccgta gctagctag 3339

- Á <210> 11
- <211> 973
- <212> PRT
- <213> *Phaeodactylum tricornutum*
- <400> 11

ES 2 659 118 T3

Met Glu Ser Ala Arg Ser Arg Asn Thr Arg Asp Gly Thr Arg Lys Gln
 1 5 10 15
 Gly Thr Ala Arg Arg Ser Gly Gly Thr Thr Ser Ser Asn Gly Asn Gln
 20 25 30
 Gly Phe Asp Glu Phe Gly Phe Gly Gln Pro Ala Phe Pro Asp Ser Ala
 35 40 45
 Phe Asp Asn His Gly Phe Glu Met Pro Gln Thr Arg Ile Gln Pro Thr
 50 55 60
 Lys Ile Arg Ser Arg Arg Arg Ala Ser Leu Ala Ala Ala Pro Asn Ile
 65 70 75 80
 Asp Val Val Ser Glu Asn Pro Ser Ile Gly Phe Thr Asn Gln Phe Gln
 85 90 95
 Ser Ser Gln Asp Glu Gln Val Ser Arg Gly Gly Ala Arg Leu Ala Lys
 100 105 110
 Ala Gly Arg Ser Ser Arg Ser Met Asp Gly Ile Glu Phe Pro Thr Ala
 115 120 125
 Arg Lys Asp Val Ser Ser Gln Asn Arg Pro Arg Arg Ser Gly Arg Arg
 130 135 140
 Ala Ser Met Ala Thr Ser Ser Asn His Ser Leu Ser Ala Ser Asn His
 145 150 155 160
 Thr Asn Pro Glu Leu Gly Tyr Gly Asp Ala Ile Pro Ser Val Ala Ala
 165 170 175
 Asn His Arg Lys Gly Asp Ser Asn Ser Gly Ile Leu Asp Phe Gly Phe
 180 185 190
 Gly Gly Gly Lys Asn Ala Gly Thr Ala Asn Ala Asp Tyr Gly Tyr Gly
 195 200 205
 Asp Thr Met Ser Ser Gly Phe Gly Asn Phe Glu Ser Met Pro Ser Ala
 210 215 220
 Pro Ser Thr Thr Pro Glu Ser Glu Arg Pro Arg Arg Ser Gly Arg Arg
 225 230 235 240
 Ser Ser Ile Ser Gly Gly Leu Glu Ser Leu Arg Ser Asp Leu Arg Gly
 245 250 255
 Gly Asp Leu Ser Gly Ala Pro Ser Ser Arg Val Leu Gly Gly Asn Ser
 260 265 270

ES 2 659 118 T3

Arg Ala Gln Asn Ile Val Leu Pro Met Ala Gly Pro Glu Lys Val Ala
 275 280 285

Gly Gly Asn Val Arg Arg Gly Arg Arg Gly Ser Leu Leu Gly Ser Val
 290 295 300

Gly Asn Ala Val Gly Ala Thr Met Gly Gly Phe Thr Gly Gly Asn Lys
 305 310 315 320

Asp Lys Glu Lys Leu Asp Asp Asp Thr Thr Lys Lys Ser Lys Ser Phe
 325 330 335

Leu Lys Asp Arg Lys Ala Glu Gly Arg Arg Gly Thr Thr Arg Gln Pro
 340 345 350

Ser Ala Asp Gly Asn Ile Ile Ser Ser Tyr Thr Gly Asp Arg Asp Arg
 355 360 365

Arg Arg Lys Pro Ala Ala Ser Ser Lys Thr Leu Gly Lys Glu Ser Asn
 370 375 380

Val Ser Tyr Ser Asp Arg Ile Leu Ala Gln Arg Asp Cys Ser Cys Gly
 385 390 400

Gly Glu Gly Ser Ser Ser Ser Gly Ser Gly Ser Ser Asp Gly Ser Ser
 405 410 415

Ala Cys Cys Arg Ser Lys Ile Trp Pro Ala Ala Arg Cys Glu Thr Tyr
 420 425 430

Arg Thr Leu Glu Ile Asp Ala Ser Ser Ser Met Thr Leu Arg Arg His
 435 440 445

Gly Leu Arg Gly Leu His Ile Ser Pro Thr Gln Ser Val Gly Asp Glu
 450 455 460

Ser Ser Phe Asp Val Asp Cys His Gly Tyr Cys Gln Asp Val Gln Ser
 465 470 475 480

Ile Leu Asp Ala Ala Tyr Val Arg Phe Leu Lys Ala Leu Arg Arg Ser
 485 490 495

Val Ser Ser Thr Pro Leu Ala His His Asp Arg Arg Glu Asn Glu Lys
 500 505 510

Val Ala Gln His Asp Asn Val Gln Ala Leu Leu Gly Ile His Ile Ser
 515 520 525

Ile Thr Thr Asn Glu Ser Ala Leu Val His Asp Ala Asp Glu Arg Tyr
 530 535 540

ES 2 659 118 T3

Gln Leu Asp Val Pro Gly Pro Thr Val Thr Glu Asn Asp Asp Asp Asp
 545 550 555 560

Asp Gly Ser Tyr Ile His Leu Thr Ala Pro Thr Val Tyr Gly Ile Leu
 565 570 575

His Ala Tyr Gln Ser Leu Leu Gln Leu Val Thr Phe Val Gly Arg Asp
 580 585 590

Ser Gln Thr Gly Ala Phe Val Phe Ala Met Pro Asp Thr Thr Leu Ile
 595 600 605

Arg Ile Arg Asp Gly Pro Val Tyr Pro Tyr Arg Gly Leu Met Ile Asp
 610 615 620

Thr Ala Arg His Phe Leu Pro Leu Pro Leu Ile Leu Gln Asn Leu Asp
 625 630 635 640

Ala Met Glu Ala Ser Lys Leu Asn Val Leu His Trp His Val Thr Asp
 645 650 655

Ser Gln Ser Trp Pro Tyr Val Ser Thr Ala Phe Pro Glu Leu Ser Ala
 660 665 670

Arg Gly Ala Phe Gly Pro Glu Glu Thr Tyr Thr Ala Thr Asp Ile Ala
 675 680 685

Leu Val Val Arg Glu Ala Ala Ala Arg Ala Ile Gly Arg Ser His Pro
 690 695 700

Glu Trp Leu Thr Pro Cys Gly Ser Lys Pro Arg Pro Gln Glu Pro Leu
 705 710 715 720

Asp Ala Thr Asn Pro Ala Val Tyr Glu Phe Val His Arg Leu Tyr Asp
 725 730 735

Glu Leu Ala Ile Leu Phe Ala His Glu Ser Phe Leu His Val Gly Gly
 740 745 750

Asp Glu Val Asn Leu Asp Cys Tyr His Asn Ser Thr Thr Val Gln Arg
 755 760 765

Trp Met Arg Lys His Asn Met Thr Gln Glu Leu Glu Val Leu Ser Tyr
 770 775 780

Phe Glu Arg Asp Leu Leu Ser Tyr Val Thr Ala Val Leu Asn Arg Arg
 785 790 795 800

Pro Ile Val Trp Gln Glu Leu Phe Asp Ser Gly Leu Gly Leu Pro Asn
 805 810 815

ES 2 659 118 T3

Gln Thr Ile Val Asp Val Trp Lys Ser Trp Glu Pro Ser Ser Arg Tyr
 820 825 830

Asn Ala Thr Leu Arg Gly His Glu Val Ile Leu Ser Ser Cys Trp Tyr
 835 840 845

Leu Asp His Leu Asn Glu Asp Trp Gln Ser Phe Tyr Ala Cys Asp Pro
 850 855 860

Arg Glu Phe Asn Gly Thr Lys Glu Gln Lys Asn Leu Ile Leu Gly Gly
 865 870 875 880

His Ala Ser Met Trp Gly Glu Arg Val Asp Ala Thr Asn Phe Leu Ser
 885 890 895

Arg Val Trp Pro Arg Ala Ser Ala Thr Ala Glu Lys Leu Trp Thr Gly
 900 905 910

Asn Leu Thr Ala Ala Ala Asp Ser Ala Ala Ser Arg Leu Ala Ala Phe
 915 920 925

Arg Cys His Leu Val Arg Arg Gly Ile Pro Ala Ser Pro Val Gly Pro
 930 935 940

Gly Ala Ser Cys Gly Arg Gln Pro Asn Gly Phe Pro Ala Val Ile Asp
 945 950 955 960

Ser Phe His Asp Glu Glu Leu Gln Glu Gly Lys Val Thr
 965 970

<210> 12

<211> 3341

5 <212> DNA

<213> Phaeodactylum tricornutum

<400> 12

atggaatctg ctcgctctcg aaataccg g gatggcacc gaaagcaagg tacggcacga	60
agaagtggcg gtactactag tagcaatggc aatcaaggct tcgacgagtt tggatttgg	120
caacctgcct ttccggattc tgcgtttgac aaccacggct ttgagatgcc gcaaactcga	180
attcagccaa cgaagattcg ttcccgccgt cgagcatctt tagctgcggc gccgaacatt	240
gacgttgtgt cgaaaaacc gtcaataggt ttaccaatc aatttcaatc ttcacaagac	300
gaacagggtg ctcgtgggtg agcccgcttg gcaaaggcgg gacgctcatc gcgctcaatg	360
gacggcatcg aattcccaac tgcacgcaag gacgtttcta gtcaaaatcg tcctcgtcgt	420
tcgggtcggc gggcttcaat ggctacttct tccaaccaca gcctttccgc ttccaatcac	480
accaaccggg aactcgggta cggagacgcc attccgtctg ttgctgctaa ccatagaaaa	540
ggagactcta atagcggaat tttggacttc ggctttgggt gtggtaagaa tgccggtaca	600
gccaatgccg actacggcta cggtgacaca atgtcgtcgg gttttggtaa tttcaggtct	660

10

ES 2 659 118 T3

atgccatccg cgccttccac cacacccgaa tctgaacgtc cgcgtcgcag cggacgacgc	720
tccagtatca gtggaggctc tgaagttcta cggctctgact tgcgcggagg cgacctgagt	780
ggtgctccgt ctagtccggg gttgggtgga aattctcgcg cccagaacat tgtgttgcct	840
atggccgggc cggaaaaagt ggccggtggc aatgttcgtc gtggacgtcg cggatcctta	900
ctgggtagtg ttggtaatgc agtcggagct accatggggg gattcactgg tggaaataag	960
gacaaggaaa aactcgcgca cgataccacc aaaaagtcta agtcctttct aaaggatcgc	1020
aaggctgaag gtcggcgagg cacgacacgt caaccatcgg ccgatggcaa tataatctct	1080
tcctataccg gcgatcgcga ccgacgccgc aagccggcag cgtcgtccaa gacctgggc	1140
aaagagagca acgtgtcgtc ctccgatcgt attttagcac agcggtaaga ggcaacataa	1200
aaacacagca attcaataat ttggcgtgtc acagactacc aatctaaatg tttaaagcct	1260
agcggtatag tccgctcggc caagatatag aagggcaggg aattgcagca aagtaaaggc	1320
atattagatt cagtgtacgt gatgtaccgg gaccagtgtg agagaataaa ggtccgaatg	1380
tgactcgcct gagatcgcgg aatcgcagaa aaaccgaga cactgtcaat ccgtttcttc	1440
gcgaaatcct ggccgctttc gcgcatttac attacatagt tcgcaccatg tgggggagac	1500
gaggactttg tcgtatttct cttctaactg tgetattatt actattcgtt tctaacagtg	1560
actgtagtgt tggaggtgaa gggagcagca gtagcggtag cggtagcagt gacggtagta	1620
gtgcttgttg tcgcagcaaa atttggccgg cggctcgtatg tgaaacctac cgaacactcg	1680
aaatcgtatc ttcctcctct atgacattgc gacggcacgg cttgcgagga ctgcatatct	1740
cccctactca aagcgtaggg gacgaaagta gcttcgatgt ggactgccat ggatattgtc	1800
aagacgttca atcgatactg gacgcccgct acgttcgctt tctcaaggcg ctccggcgaa	1860
gcgtgtcttc cacgcctttg gcgcattcac accgtcgcga aaatgaaaag gtcgctcaac	1920
atgacaacgt acaggctctg ttgggcattc acatttccat tactacgaat gagtctgcac	1980
tcgtacacga cgcggacgaa cgataccaac tggacgtccc agggcctacc gtcactgaaa	2040
acgacgacga cgacgatggc agctacattc atctcactgc acccaccgtc tacggcattc	2100
tgcacgccta ccaaagctta ctgcagctgg tgacgtttgt tggtagggac tctcaaacag	2160
gcgctttcgt attcgccatg ccggacacaa ccttcattcg aatccgtgat ggaccctgtg	2220
atccctaccg gggactcatg atcgacacgg cccgacattt tttgccacta ccgcttatct	2280
tgcaaaactt ggacgccatg gaggccagta aactgaacgt cttgcactgg cacgtgactg	2340
attcgcagtc gtggccctac gtcagtactg cttttccgga gcttagtgct cggggagcct	2400
ttggtcctga agaaacctac acggctacag atattgccct cgtcgtgcgg gaagccgccg	2460
cacgggttat tcgggtgatt cctgaattcg atttgcctgg acactcgtaa gcgattggac	2520
gctcacatcc ggaatggtta acaccctgtg ggtccaagcc acggccgcaa gaacctttgg	2580
atgcgaccaa tccggccgct tacgaattcg tacaccgct ctacgacgaa ttggcaatac	2640
tctttgcgca cgaatccttt ttacacgtcg gaggagacga agtcaattta gattgttacc	2700

ES 2 659 118 T3

acaatagcac gacgggtccaa agatggatgc gaaaacacaa tatgacacag gaacttgagg 2760
 ttctgagcta ttttgagcgt gatttgcttt cgtacgtcac cgctgtatta aatcgtcgtc 2820
 ccattgtgtg gcaggaactc ttcgattcgg gattgggtct tcccaatcag acaattgtcg 2880
 atgtctggaa atcgtgggaa ccttcgtcgc gatacaacgc cactttgcgg ggccacgaag 2940
 ttattttgtc ctcgtgctgg tatctcgatc atttgaacga agattggcaa agcttctacg 3000
 cctgtgatcc acgggagttc aacggtagca aagaacagaa gaacttgatt ctgggcggtc 3060
 acgcttccat gtggggggaa cgggtggatg cgaccaactt tctatctcgt gtttggeccc 3120
 gtgccagtgc tacggccgaa aagctgtgga caggcaactt aacagctgcg gcggattcgg 3180
 cggcttctcg attggccgcc tttcgtgtc atttggtccg cagaggaatt ccggccagtc 3240
 cggtcgggcc gggagcaagt tgcggcagac aaccaaagg tttccggct gtgatcgata 3300
 gctttcatga cgaggagtg caggaaggaa aggttacttg a 3341

<210> 13
 <211> 442
 <212> PRT
 <213> Mus musculus

<400> 13

5 <210> 13
 <211> 442
 <212> PRT
 <213> MUS musculus

10 <400> 13

Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Ser Asp Ala Leu Gly Pro Pro Leu Leu Asp Ala Glu Pro Val Arg
 35 40 45
 Gly Ala Gly His Leu Ala Val Ser Val Gly Ile Arg Arg Val Ser Asn
 50 55 60
 Glu Ser Ala Ala Pro Leu Val Pro Ala Val Pro Arg Pro Glu Val Asp
 65 70 75 80
 Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr Gln Leu Asn Phe Asp
 85 90 95
 Gln Met Leu Arg Asn Val Gly Asn Asp Gly Thr Trp Ser Pro Gly Glu
 100 105 110
 Leu Val Leu Val Val Gln Val His Asn Arg Pro Glu Tyr Leu Arg Leu
 115 120 125
 Leu Ile Asp Ser Leu Arg Lys Ala Gln Gly Ile Gln Glu Val Leu Val
 130 135 140
 Ile Phe Ser His Asp Phe Trp Ser Ala Glu Ile Asn Ser Leu Ile Ser

ES 2 659 118 T3

```

145                150                155                160
Arg Val Asp Phe Cys Pro Val Leu Gln Val Phe Phe Pro Phe Ser Ile
      165                170                175
Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp Pro Arg Asp Cys Pro
      180                185                190
Arg Asp Leu Lys Lys Asn Ala Ala Leu Lys Leu Gly Cys Ile Asn Ala
      195                200                205
Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu Ala Lys Phe Ser Gln
      210                215
Thr Lys His His Trp Trp Trp Lys Leu His Phe Val Trp Glu Arg Val
      225                230                235
Lys Val Leu Gln Asp Tyr Thr Gly Leu Ile Leu Phe Leu Glu Glu Asp
      245                250                255
His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe Lys Lys Met Trp Lys
      260                265                270
Leu Lys Gln Gln Glu Cys Pro Gly Cys Asp Val Leu Ser Leu Gly Thr
      275                280                285
Tyr Thr Thr Ile Arg Ser Phe Tyr Gly Ile Ala Asp Lys Val Asp Val
      290                295                300
Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly Leu Ala Leu Thr Arg
      305                310                315
Asp Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp Thr Phe Cys Thr Tyr
      325                330                335
Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr Leu Thr Leu Ala Cys
      340                345                350
Leu Pro Lys Ile Trp Lys Val Leu Val Pro Gln Ala Pro Arg Ile Phe
      355                360                365
His Ala Gly Asp Cys Gly Met His His Lys Lys Thr Cys Arg Pro Ser
      370                375                380
Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn Ser Asn Lys Gln Tyr
      385                390                395
Leu Phe Pro Glu Thr Leu Val Ile Gly Glu Lys Phe Pro Met Ala Ala
      405                410                415
Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly Asp Ile Arg Asp His
      420                425                430
Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
      435                440

```

Á
<210> 14
<211> 447
<212> PRT
<213> Homo sapiens

ES 2 659 118 T3

<400> 14

Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15

Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30

Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45

Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
 50 55 60

Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro
 65 70 75 80

Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
 85 90 95

Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
 100 105 110

Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
 115 120 125

Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
 130 135 140

Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
 145 150 155 160

Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
 165 170 175

Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190

Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
 195 200 205

Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220

ES 2 659 118 T3

Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240

Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
 245 250 255

Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270

Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
 275 280 285

Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
 290 295 300

Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320

Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335

Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350

Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365

Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380

Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 400

Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
 405 410 415

Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430

Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440 445

<210> 15

<211> 334

5 <212> PRT

<213> Phaeodactylum tricornutum

<400> 15

Arg Ser Tyr Ala Ser Arg Lys Thr Arg Glu Ile Lys Asn Arg Arg Phe
 1 5 10 15

10

ES 2 659 118 T3

Thr Met Lys Asn Thr Asp Ile Pro Phe Val Leu Met Val Tyr Lys Arg
 20 25 30
 Val Asp Tyr Leu Lys Lys Ala Ile Asp Ser Val Arg Arg Ser Asp Phe
 35 40 45
 Pro Lys Ser Arg Val Pro Leu Ile Ile Ser His Asp Gly Arg Val Pro
 50 55 60
 Glu Val Val Glu Tyr Val Glu Ser Leu Lys Asp Glu Phe Lys Ile Ile
 65 70 75 80
 Gln Leu Ile His Pro His Ser Cys Tyr Glu His Pro Asn Glu Phe Pro
 85 90 95
 Gly Asp Asp Pro Thr Leu Asn Glu Gly Phe Ala Gly Asp Ser Tyr Gly
 100 105 110
 Asn Pro Arg Ser Ala Trp Ile Thr Cys Cys Lys His His Phe Thr Trp
 115 120 125
 Met Leu His Thr Val Phe Arg Arg Asp Phe Thr Asp Pro Ala Val Asp
 130 135 140
 Thr Phe Leu Phe Leu Glu Glu Asp Tyr Ile Val Ala Pro Thr Ile Tyr
 145 150 155 160
 Ser Ala Val Ile Ala Gly Leu Asn Val Met Glu Asp Met Asp Lys Glu
 165 170 175
 Ile Pro Gly Gly Phe Phe Gly Leu Gly Met Asp Pro Ser Met Ala Asn
 180 185 190
 Thr Ala Phe Glu Pro Tyr Tyr Lys Lys Ala Thr Trp Tyr Val Glu Ala
 195 200 205
 Phe Lys Ser Gly Pro Met Thr Met Asn Arg Asp Met Phe Lys Lys Leu
 210 215 220
 Gln Gln His Ala Lys Glu Tyr Cys Thr Phe Asp Asp Tyr Asn Trp Asp
 225 230 235 240
 Trp Ser Ile Val His Leu Gln Ser Lys Lys Leu Leu Pro Arg Thr Leu
 245 250 255
 Leu Met Pro Ser Lys Val Leu Ala Lys His Ile Gly Val Lys Glu Gly
 260 265 270
 Met His Thr Asn Lys Ser Phe Gly Lys Asp Phe Ser Asp Leu Phe Pro
 275 280 285
 Asn Tyr Ala Pro Arg Arg Glu Gln Leu Thr Thr His Phe Ala Glu Tyr
 290 295 300
 Arg Phe Thr Gly Asn Thr Ala Ala Ala Leu His Thr Glu Phe Asn Pro
 305 310 315 320
 Gly Tyr Gly Gly Trp Gly His Pro Lys Asp His Glu His Cys
 325 330

Á <210> 16
 <211> 1002

ES 2 659 118 T3

<212> DNA

<213> Phaeodactylum tricornutum

<400> 16

```

agatcgtacg catcgcgcaa aaccctgaa atcaagaatc gtcgctttac aatgaagaac      60
acggacatcc cgttcgtatt gatggtttat aagcgcgtgg attatttgaa aaaggcgatt      120
gattcgggtcc gtcgatccga ctttccaaag tctcgtgtgc ctcttatcat ttcgcacgat      180
ggacgagtgc cggaaagtcgt tgagtatgtc gaatcgcctga aagacgagtt caaaattatc      240
caactcattc atccgcattc ttgctacgag catcccaacg agtttcctgg ggatgatccc      300
acactcaacg aaggctttgc tggagatagc tatggtaatc cacggagtgc gtggatcacc      360
tgctgcaaac atcattttac ttggatgctt cacactgtct ttcgctcgga ctttacggac      420
ccagcagtgg atacattttt gtttctcgaa gaagattaca tcgtggctcc tacgatttat      480
tccgccgtta ttgctgggtt gaatgttatg gaagacatgg acaaggagat tccgggcggc      540
ttcttcggtt tgggtatgga tccgagtatg gcgaatactg cctttgagcc ttactacaag      600
aaggcgacgt ggtatgtcga agcattcaag tcagggtccga tgacgatgaa ccgggacatg      660
ttcaaaaagc tccaacaaca tgctaaggag tactgtacgt tcgacgatta caattgggat      720
tgggccattg tccatctaca aagtaaaaag ttactaccac ggactcttct catgccaagc      780
aaagtcctag caaaacatat tgggtgcaag gagggtatgc acacgaacaa atcctttgga      840
aaagactttt cagatttatt tcccaactat gccctcgtc gcgagcagct gaccacgcat      900
tttgcggaat accgttttac cgggaacacc gcggcagctt tacacaccga attcaatccc      960
5  ggttatggtg gttggggaca tccgaaagac cgcgagcact gc                          1002

```

<210> 17

<211> 538

<212> PRT

10 <213> Mus musculus

<400> 17

```

Met Lys Met Arg Arg Tyr Lys Leu Phe Leu Met Phe Cys Met Ala Gly
1           5           10           15
Leu Cys Leu Ile Ser Phe Leu His Phe Phe Lys Thr Leu Ser Tyr Val
20           25           30

```

ES 2 659 118 T3

Thr Phe Pro Arg Glu Leu Ala Ser Leu Ser Pro Asn Leu Val Ser Ser
 35 40 45
 Phe Phe Trp Asn Asn Ala Pro Val Thr Pro Gln Ala Ser Pro Glu Pro
 50 55 60
 Gly Gly Pro Asp Leu Leu Arg Thr Pro Leu Tyr Ser His Ser Pro Leu
 65 70 75 80
 Leu Gln Pro Leu Ser Pro Ser Lys Ala Thr Glu Glu Leu His Arg Val
 85 90 95
 Asp Phe Val Leu Pro Glu Asp Thr Thr Glu Tyr Phe Val Arg Thr Lys
 100 105 110
 Ala Gly Gly Val Cys Phe Lys Pro Gly Thr Arg Met Leu Glu Lys Pro
 115 120 125
 Ser Pro Gly Arg Thr Glu Glu Lys Pro Glu Val Ser Glu Gly Ser Ser
 130 135 140
 Ala Arg Gly Pro Ala Arg Arg Pro Met Arg His Val Leu Ser Thr Arg
 145 150 155 160
 Glu Arg Leu Gly Ser Arg Gly Thr Arg Arg Lys Trp Val Glu Cys Val
 165 170 175
 Cys Leu Pro Gly Trp His Gly Pro Ser Cys Gly Val Pro Thr Val Val
 180 185 190
 Gln Tyr Ser Asn Leu Pro Thr Lys Glu Arg Leu Val Pro Arg Glu Val
 195 200 205
 Pro Arg Arg Val Ile Asn Ala Ile Asn Ile Asn His Glu Phe Asp Leu
 210 215 220
 Leu Asp Val Arg Phe His Glu Leu Gly Asp Val Val Asp Ala Phe Val
 225 230 235 240
 Val Cys Glu Ser Asn Phe Thr Ala Tyr Gly Glu Pro Arg Pro Leu Lys
 245 250 255
 Phe Arg Glu Met Leu Thr Asn Gly Thr Phe Glu Tyr Ile Arg His Lys
 260 265 270
 Val Leu Tyr Val Phe Leu Asp His Phe Pro Pro Gly Gly Arg Gln Asp
 275 280 285
 Gly Trp Ile Ala Asp Asp Tyr Leu Arg Thr Phe Leu Thr Gln Asp Gly
 290 295 300

ES 2 659 118 T3

Val Ser Arg Leu Arg Asn Leu Arg Pro Asp Asp Val Phe Ile Ile Asp
 305 310 315 320

Asp Ala Asp Glu Ile Pro Ala Arg Asp Gly Val Leu Phe Leu Lys Leu
 325 330 335

Tyr Asp Gly Trp Thr Glu Pro Phe Ala Phe His Met Arg Lys Ser Leu
 340 345 350

Tyr Gly Phe Phe Trp Lys Gln Pro Gly Thr Leu Glu Val Val Ser Gly
 355 360 365

Cys Thr Met Asp Met Leu Gln Ala Val Tyr Gly Leu Asp Gly Ile Arg
 370 375 380

Leu Arg Arg Arg Gln Tyr Tyr Thr Met Pro Asn Phe Arg Gln Tyr Glu
 385 390 395 400

Asn Arg Thr Gly His Ile Leu Val Gln Trp Ser Leu Gly Ser Pro Leu
 405 410 415

His Phe Ala Gly Trp His Cys Ser Trp Cys Phe Thr Pro Glu Gly Ile
 420 425 430

Tyr Phe Lys Leu Val Ser Ala Gln Asn Gly Asp Phe Pro Arg Trp Gly
 435 440 445

Asp Tyr Glu Asp Lys Arg Asp Leu Asn Tyr Ile Arg Ser Leu Ile Arg
 450 455 460

Thr Gly Gly Trp Phe Asp Gly Thr Gln Gln Glu Tyr Pro Pro Ala Asp
 465 470 475 480

Pro Ser Glu His Met Tyr Ala Pro Lys Tyr Leu Leu Lys Asn Tyr Asp
 485 490 495

Gln Phe Arg Tyr Leu Leu Glu Asn Pro Tyr Arg Glu Pro Lys Ser Thr
 500 505 510

Glu Glu Gly Gly Arg Arg Asn Gln Gly Ser Asp Gly Arg Pro Ser Ala
 515 520 525

Val Arg Gly Lys Leu Asp Thr Val Glu Gly
 530 535

<210> 18
 <211> 533
 <212> PRT
 <213> Homo sapiens

Á

<400> 18

ES 2 659 118 T3

Met Lys Met Arg Arg Tyr Lys Leu Phe Leu Met Phe Cys Met Ala Gly
 1 5 10 15

Leu Cys Leu Ile Ser Phe Leu His Phe Phe Lys Thr Leu Ser Tyr Val
 20 25 30

Thr Phe Pro Arg Glu Leu Ala Ser Leu Ser Pro Asn Leu Val Ser Ser
 35 40 45

Phe Phe Trp Asn Asn Ala Pro Val Thr Pro Gln Ala Ser Pro Glu Pro
 50 55 60

Gly Gly Pro Asp Leu Leu Arg Thr Pro Leu Tyr Ser His Ser Pro Leu
 65 70 75 80

Leu Gln Pro Leu Pro Pro Ser Lys Ala Ala Glu Glu Leu His Arg Val
 85 90 95

Asp Leu Val Leu Pro Glu Asp Thr Thr Glu Tyr Phe Val Arg Thr Lys
 100 105 110

Ala Gly Gly Val Cys Phe Lys Pro Gly Thr Lys Met Leu Glu Arg Pro
 115 120 125

Pro Pro Gly Arg Pro Glu Glu Lys Pro Glu Gly Ala Asn Gly Ser Ser
 130 135 140

Ala Arg Arg Pro Pro Arg Tyr Leu Leu Ser Ala Arg Glu Arg Thr Gly
 145 150 155 160

Gly Arg Gly Ala Arg Arg Lys Trp Val Glu Cys Val Cys Leu Pro Gly
 165 170 175

Trp His Gly Pro Ser Cys Gly Val Pro Thr Val Val Gln Tyr Ser Asn
 180 185 190

Leu Pro Thr Lys Glu Arg Leu Val Pro Arg Glu Val Pro Arg Arg Val
 195 200 205

Ile Asn Ala Ile Asn Val Asn His Glu Phe Asp Leu Leu Asp Val Arg
 210 215 220

Phe His Glu Leu Gly Asp Val Val Asp Ala Phe Val Val Cys Glu Ser
 225 230 235 240

Asn Phe Thr Ala Tyr Gly Glu Pro Arg Pro Leu Lys Phe Arg Glu Met
 245 250 255

Leu Thr Asn Gly Thr Phe Glu Tyr Ile Arg His Lys Val Leu Tyr Val
 260 265 270

ES 2 659 118 T3

Phe Leu Asp His Phe Pro Pro Gly Gly Arg Gln Asp Gly Trp Ile Ala
 275 280 285

Asp Asp Tyr Leu Arg Thr Phe Leu Thr Gln Asp Gly Val Ser Arg Leu
 290 295 300

Arg Asn Leu Arg Pro Asp Asp Val Phe Ile Ile Asp Asp Ala Asp Glu
 305 310 315 320

Ile Pro Ala Arg Asp Gly Val Leu Phe Leu Lys Leu Tyr Asp Gly Trp
 325 330 335

Thr Glu Pro Phe Ala Phe His Met Arg Lys Ser Leu Tyr Gly Phe Phe
 340 345 350

Trp Lys Gln Pro Gly Thr Leu Glu Val Val Ser Gly Cys Thr Val Asp
 355 360 365

Met Leu Gln Ala Val Tyr Gly Leu Asp Gly Ile Arg Leu Arg Arg Arg
 370 375 380

Gln Tyr Tyr Thr Met Pro Asn Phe Arg Gln Tyr Glu Asn Arg Thr Gly
 385 390 395 400

His Ile Leu Val Gln Trp Ser Leu Gly Ser Pro Leu His Phe Ala Gly
 405 410 415

Trp His Cys Ser Trp Cys Phe Thr Pro Glu Gly Ile Tyr Phe Lys Leu
 420 425 430

Val Ser Ala Gln Asn Gly Asp Phe Pro Arg Trp Gly Asp Tyr Glu Asp
 435 440 445

Lys Arg Asp Leu Asn Tyr Ile Arg Gly Leu Ile Arg Thr Gly Gly Trp
 450 455 460

Phe Asp Gly Thr Gln Gln Glu Tyr Pro Pro Ala Asp Pro Ser Glu His
 465 470 475 480

Met Tyr Ala Pro Lys Tyr Leu Leu Lys Asn Tyr Asp Arg Phe His Tyr
 485 490 495

Leu Leu Asp Asn Pro Tyr Gln Glu Pro Arg Ser Thr Ala Ala Gly Gly
 500 505 510

Trp Arg His Arg Gly Pro Glu Gly Arg Pro Pro Ala Arg Gly Lys Leu
 515 520 525

Asp Glu Ala Glu Val
 530

<210> 19
 <211> 535
 <212> PRT
 <213> Mus musculus
 <400> 19

Á

ES 2 659 118 T3

Met Arg Leu Arg Asn Gly Thr Val Ala Thr Ala Leu Val Phe Val Thr
1 5 10 15

Ser Phe Leu Thr Leu Ser Trp Tyr Thr Thr Trp Gln Asn Gly Lys Glu
20 25 30

Lys Leu Ile Ala Tyr Gln Arg Glu Phe Leu Ala Leu Lys Glu Arg Leu
35 40 45

Arg Val Ala Glu His Arg Ile Ser Gln Arg Ser Ser Glu Leu Asn Thr
50 55 60

Ile Val Gln Gln Phe Arg Arg Ala Gly Ala Glu Thr Asn Gly Ser Lys
65 70 75 80

Thr Ala Leu Ser Thr Ile Ser Asp Asn Thr Ile Lys Leu Leu Lys Glu
85 90 95

Leu Thr Ser Lys Lys Ser Leu Arg Val Pro Ser Ile Tyr Tyr His Leu
100 105 110

Pro His Leu Leu Gln Asn Glu Arg Ser Leu Gln Pro Ala Val Gln Ile
115 120 125

Gly Ser Gly Arg Thr Gly Val Ser Ile Val Met Gly Ile Pro Thr Val
130 135 140

Lys Arg Glu Val Lys Ser Tyr Leu Val Glu Thr Leu His Ser Leu Ile
145 150 155 160

Asp Asn Leu Tyr Pro Glu Glu Lys Leu Asp Cys Val Ile Val Val Phe
165 170 175

Ile Gly Glu Thr Asp Leu Asp Tyr Val His Ser Val Val Ala Asn Leu
180 185 190

Glu Lys Glu Phe Ser Arg Glu Ile Ser Ser Gly Leu Leu Glu Ile Ile
195 200 205

Ser Pro Pro Glu Ser Tyr Tyr Pro Asp Leu Thr Asn Leu Lys Glu Thr
210 215 220

Phe Gly Asp Ser Lys Glu Arg Val Arg Trp Arg Thr Lys Gln Asn Leu
225 230 235 240

Asp Tyr Cys Phe Leu Met Met Tyr Ala Gln Glu Lys Gly Ile Tyr Tyr

ES 2 659 118 T3

<400> 20

Met Arg Leu Arg Asn Gly Thr Phe Leu Thr Leu Leu Leu Phe Cys Leu
 1 5 10 15
 Cys Ala Phe Leu Ser Leu Ser Trp Tyr Ala Ala Leu Ser Gly Gln Lys
 20 25 30
 Gly Asp Val Val Asp Ile Tyr Gln Arg Glu Phe Leu Ala Leu Arg Asp
 35 40 45
 Arg Leu His Ala Ala Glu Gln Glu Ser Leu Lys Arg Ser Lys Glu Leu
 50 55 60
 Asn Leu Val Leu Glu Glu Ile Lys Arg Ala Val Ser Glu Arg Gln Ala
 65 70 75 80
 Leu Arg Asp Gly Glu Gly Asn Arg Thr Trp Gly Arg Leu Thr Glu Asp
 85 90 95
 Pro Arg Leu Lys Pro Trp Asn Val Ser His Arg His Val Leu His Leu
 100 105 110
 Pro Thr Val Phe His His Leu Pro His Leu Leu Ala Lys Glu Ser Ser
 115 120 125
 Leu Gln Pro Ala Val Arg Val Gly Gln Gly Arg Thr Gly Val Ser Val
 130 135 140
 Val Met Gly Ile Pro Ser Val Arg Arg Glu Val His Ser Tyr Leu Thr
 145 150 155 160
 Asp Thr Leu His Ser Leu Ile Ser Glu Leu Ser Pro Gln Glu Lys Glu
 165 170 175
 Asp Ser Val Ile Val Val Leu Ile Ala Glu Thr Asp Pro Gln Tyr Thr
 180 185 190
 Ser Ala Val Thr Glu Asn Ile Lys Ala Leu Phe Pro Thr Glu Ile His
 195 200 205
 Ser Gly Leu Leu Glu Val Ile Ser Pro Ser Pro His Phe Tyr Pro Asp
 210 215 220

ES 2 659 118 T3

Phe Ser Arg Leu Arg Glu Ser Phe Gly Asp Pro Lys Glu Arg Val Arg
 225 230 235 240

Trp Arg Thr Lys Gln Asn Leu Asp Tyr Cys Phe Leu Met Met Tyr Ala
 245 250 255

Gln Ser Lys Gly Ile Tyr Tyr Val Gln Leu Glu Asp Asp Ile Val Ala
 260 265 270

Lys Pro Asn Tyr Leu Ser Thr Met Lys Asn Phe Ala Leu Gln Gln Pro
 275 280 285

Ser Glu Asp Trp Met Ile Leu Glu Phe Ser Gln Leu Gly Phe Ile Gly
 290 295 300

Lys Met Phe Lys Ser Leu Asp Leu Ser Leu Ile Val Glu Phe Ile Leu
 305 310 315 320

Met Phe Tyr Arg Asp Lys Pro Ile Asp Trp Leu Leu Asp His Ile Leu
 325 330 335

Trp Val Lys Val Cys Asn Pro Glu Lys Asp Ala Lys His Cys Asp Arg
 340 345 350

Gln Lys Ala Asn Leu Arg Ile Arg Phe Lys Pro Ser Leu Phe Gln His
 355 360 365

Val Gly Thr His Ser Ser Leu Ala Gly Lys Ile Gln Lys Leu Lys Asp
 370 375 380

Lys Asp Phe Gly Lys His Ala Leu Arg Lys Glu His Val Asn Pro Pro
 385 390 395 400

Ala Glu Val Ser Thr Ser Leu Lys Thr Tyr Gln His Phe Thr Leu Glu
 405 410 415

Lys Ala Tyr Leu Arg Glu Asp Phe Phe Trp Ala Phe Thr Pro Ala Ala
 420 425 430

Gly Asp Phe Ile Arg Phe Arg Phe Phe Gln Pro Leu Arg Leu Glu Arg
 435 440 445

Phe Phe Phe Arg Ser Gly Asn Ile Glu His Pro Glu Asp Lys Leu Phe
 450 455 460

Asn Thr Ser Val Glu Val Leu Pro Phe Asp Asn Pro Gln Ser Glu Lys
 465 470 475 480

Glu Ala Leu Gln Glu Gly Arg Ser Ala Thr Leu Arg Tyr Pro Arg Ser
 485 490 495

Pro Asp Gly Tyr Leu Gln Ile Gly Ser Phe Tyr Lys Gly Val Ala Glu
 500 505 510

Gly Glu Val Asp Pro Ala Phe Gly Pro Leu Glu Ala Leu Arg Leu Ser
 515 520 525

Ile Gln Thr Asp Ser Pro Val Trp Val Ile Leu Ser Glu Ile Phe Leu
 530 535 540

Lys Lys Ala Asp
 545

ES 2 659 118 T3

<210> 21
 <211> 478
 <212> PRT
 <213> Mus musculus

Á

<400> 21
 Met Leu Lys Phe Tyr Gln Met Lys Tyr Ile Phe Gln Ile Leu Asp Lys
 1 5 10 15
 Met Arg Cys Leu Arg Lys Arg Ser Thr Val Ser Phe Leu Gly Val Leu
 20 25 30
 Val Val Phe Leu Leu Phe Met Asn Leu Tyr Ile Glu Asp Ser Tyr Val
 35 40 45
 Leu Glu Gly Asp Lys Gln Leu Ile Arg Glu Thr Ser Thr His Gln Leu
 50 55 60
 Asn Ser Glu Arg Tyr Val His Thr Phe Lys Asp Leu Ser Asn Phe Ser
 65 70 75 80
 Gly Thr Ile Asn Val Thr Tyr Arg Tyr Leu Ala Ala Thr Pro Leu Gln
 85 90 95
 Arg Lys Arg Tyr Leu Thr Ile Gly Leu Ser Ser Val Lys Arg Lys Lys
 100 105 110
 Gly Asn Tyr Leu Leu Asp Thr Ile Lys Ser Ile Phe Glu Gln Ser Ser
 115 120 125
 Tyr Glu Glu Leu Lys Glu Ile Ser Val Val Val His Leu Ala Asp Phe
 130 135 140
 Asn Ser Ser Trp Arg Asp Ala Met Val Gln Asp Ile Thr Gln Lys Phe
 145 150 155 160
 Ala His His Ile Ile Ala Gly Arg Leu Met Val Ile His Ala Pro Glu
 165 170 175

ES 2 659 118 T3

Glu Tyr Tyr Pro Val Leu Asp Gly Leu Lys Arg Asn Tyr Asn Asp Pro
 180 185 190
 Glu Asp Arg Val Arg Phe Arg Ser Lys Gln Asn Val Asp Tyr Ala Phe
 195 200 205
 Leu Leu Asn Phe Cys Ala Asn Thr Ser Asp Tyr Tyr Val Met Leu Glu
 210 215 220
 Asp Asp Val Arg Cys Ser Arg Asn Phe Leu Thr Ala Ile Lys Lys Val
 225 230 235 240
 Ile Ala Ser Leu Glu Gly Thr Tyr Trp Val Thr Leu Glu Phe Ser Lys
 245 250 255
 Leu Gly Tyr Ile Gly Lys Leu Tyr His Ser His Asp Leu Pro Arg Leu
 260 265 270
 Ala His Phe Leu Leu Met Phe Tyr Gln Glu Met Pro Cys Asp Trp Leu
 275 280 285
 Leu Thr His Phe Arg Gly Leu Leu Ala Gln Lys Asn Val Ile Arg Phe
 290 295 300
 Lys Pro Ser Leu Phe Gln His Met Gly Tyr Tyr Ser Ser Tyr Lys Gly
 305 310 315 320
 Thr Glu Asn Lys Leu Lys Asp Asp Asp Phe Glu Glu Glu Ser Phe Asp
 325 330 335
 Ile Pro Asp Asn Pro Pro Ala Ser Phe Tyr Thr Asn Met Asn Val Phe
 340 345 350
 Glu Asn Tyr Glu Ala Ser Lys Ala Tyr Ser Ser Val Asp Glu Tyr Phe
 355 360 365
 Trp Gly Lys Ser Pro Ser Met Gly Asp Thr Phe Val Ile Val Phe Glu
 370 375 380
 Asn Pro Ile Thr Ile Lys Lys Ile Lys Val Asn Thr Gly Thr Glu Asp
 385 390 395 400
 Arg Gln Asn Asp Ile Leu Gln His Gly Ala Leu Asp Val Gly Glu Lys
 405 410 415
 Leu Ile Phe Ser Lys Gln Ile Arg Gln Cys Asp Thr Tyr Leu Arg Leu
 420 425 430
 Gly Glu Phe Lys Asn Gly Tyr Phe Glu Met Ser Asp Val Asn Gln Lys
 435 440 445
 Ile Pro Phe Asp Ile His Cys Met Arg Ile Cys Val Thr Lys Thr Gln
 450 455 460
 Lys Glu Trp Leu Ile Ile Arg Ser Ile Ser Ile Trp Thr Ser
 465 470 475

5 <210> 22
 <211> 535
 <212> PRT
 <213> Homo sapiens

10 <400> 22

ES 2 659 118 T3

Met Arg Leu Arg Asn Gly Thr Val Ala Thr Ala Leu Ala Phe Ile Thr
 1 5 10 15
 Ser Phe Leu Thr Leu Ser Trp Tyr Thr Thr Trp Gln Asn Gly Lys Glu
 20 25 30
 Lys Leu Ile Ala Tyr Gln Arg Glu Phe Leu Ala Leu Lys Glu Arg Leu
 35 40 45
 Arg Ile Ala Glu His Arg Ile Ser Gln Arg Ser Ser Glu Leu Asn Thr
 50 55 60
 Ile Val Gln Gln Phe Lys Arg Val Gly Ala Glu Thr Asn Gly Ser Lys
 65 70 75 80
 Asp Ala Leu Asn Lys Phe Ser Asp Asn Thr Leu Lys Leu Leu Lys Glu
 85 90 95
 Leu Thr Ser Lys Lys Ser Leu Gln Val Pro Ser Ile Tyr Tyr His Leu
 100 105 110
 Pro His Leu Leu Lys Asn Glu Gly Ser Leu Gln Pro Ala Val Gln Ile
 115 120 125
 Gly Asn Gly Arg Thr Gly Val Ser Ile Val Met Gly Ile Pro Thr Val
 130 135 140
 Lys Arg Glu Val Lys Ser Tyr Leu Ile Glu Thr Leu His Ser Leu Ile
 145 150 155 160
 Asp Asn Leu Tyr Pro Glu Glu Lys Leu Asp Cys Val Ile Val Val Phe
 165 170 175
 Ile Gly Glu Thr Asp Ile Asp Tyr Val His Gly Val Val Ala Asn Leu
 180 185 190
 Glu Lys Glu Phe Ser Lys Glu Ile Ser Ser Gly Leu Val Glu Val Ile
 195 200 205

ES 2 659 118 T3

Ser Pro Pro Glu Ser Tyr Tyr Pro Asp Leu Thr Asn Leu Lys Glu Thr
 210 215 220

Phe Gly Asp Ser Lys Glu Arg Val Arg Trp Arg Thr Lys Gln Asn Leu
 225 230 235 240

Asp Tyr Cys Phe Leu Met Met Tyr Ala Gln Glu Lys Gly Ile Tyr Tyr
 245 250 255

Ile Gln Leu Glu Asp Asp Ile Ile Val Lys Gln Asn Tyr Phe Asn Thr
 260 265 270

Ile Lys Asn Phe Ala Leu Gln Leu Ser Ser Glu Glu Trp Met Ile Leu
 275 280 285

Glu Phe Ser Gln Leu Gly Phe Ile Gly Lys Met Phe Gln Ala Pro Asp
 290 295 300

Leu Thr Leu Ile Val Glu Phe Ile Phe Met Phe Tyr Lys Gln Lys Pro
 305 310 315 320

Ile Asp Trp Leu Leu Asp His Ile Leu Trp Val Lys Val Cys Asn Pro
 325 330 335

Glu Lys Asp Ala Lys His Cys Asp Arg Gln Lys Ala Asn Leu Arg Ile
 340 345 350

Arg Phe Arg Pro Ser Leu Phe Gln His Val Gly Leu His Ser Ser Leu
 355 360 365

Ser Gly Lys Ile Gln Lys Leu Thr Asp Lys Asp Tyr Met Lys Pro Leu
 370 375 380

Leu Leu Lys Ile His Val Asn Pro Pro Ala Glu Val Ser Thr Ser Leu
 385 390 395 400

Lys Val Tyr Gln Gly His Thr Leu Glu Lys Thr Tyr Met Gly Glu Asp
 405 410 415

Phe Phe Trp Ala Ile Thr Pro Ile Ala Gly Asp Tyr Ile Leu Phe Lys
 420 425 430

Phe Asp Lys Pro Val Asn Val Glu Ser Tyr Leu Phe His Ser Gly Asn
 435 440 445

Gln Glu His Pro Gly Asp Ile Leu Leu Asn Thr Thr Val Glu Val Leu
 450 455 460

Pro Phe Lys Ser Glu Gly Leu Glu Ile Ser Lys Glu Thr Lys Asp Lys
 465 470 475 480

Arg Leu Glu Asp Gly Tyr Phe Arg Ile Gly Lys Phe Glu Asn Gly Val
 485 490 495

Ala Glu Gly Met Val Asp Pro Ser Leu Asn Pro Ile Ser Ala Phe Arg
 500 505 510

Leu Ser Val Ile Gln Asn Ser Ala Val Trp Ala Ile Leu Asn Glu Ile
 515 520 525

His Ile Lys Lys Ala Thr Asn
 530 535

ES 2 659 118 T3

<210> 23
 <211> 548
 <212> PRT
 <213> Homo sapiens

Á

<400> 23
 Met Arg Leu Arg Asn Gly Thr Phe Leu Thr Leu Leu Leu Phe Cys Leu
 1 5 10 15
 Cys Ala Phe Leu Ser Leu Ser Trp Tyr Ala Ala Leu Ser Gly Gln Lys
 20 25 30
 Gly Asp Val Val Asp Val Tyr Gln Arg Glu Phe Leu Ala Leu Arg Asp
 35 40 45
 Arg Leu His Ala Ala Glu Gln Glu Ser Leu Lys Arg Ser Lys Glu Leu
 50 55 60
 Asn Leu Val Leu Asp Glu Ile Lys Arg Ala Val Ser Glu Arg Gln Ala
 65 70 75 80
 Leu Arg Asp Gly Asp Gly Asn Arg Thr Trp Gly Arg Leu Thr Glu Asp
 85 90 95
 Pro Arg Leu Lys Pro Trp Asn Gly Ser His Arg His Val Leu His Leu
 100 105 110
 Pro Thr Val Phe His His Leu Pro His Leu Leu Ala Lys Glu Ser Ser
 115 120 125
 Leu Gln Pro Ala Val Arg Val Gly Gln Gly Arg Thr Gly Val Ser Val
 130 135 140
 Val Met Gly Ile Pro Ser Val Arg Arg Glu Val His Ser Tyr Leu Thr
 145 150 155 160
 Asp Thr Leu His Ser Leu Ile Ser Glu Leu Ser Pro Gln Glu Lys Glu
 165 170 175
 Asp Ser Val Ile Val Val Leu Ile Ala Glu Thr Asp Ser Gln Tyr Thr

ES 2 659 118 T3

450 455 460

Asn Thr Ser Val Glu Val Leu Pro Phe Asp Asn Pro Gln Ser Asp Lys
 465 470 475 480

Glu Ala Leu Gln Glu Gly Arg Thr Ala Thr Leu Arg Tyr Pro Arg Ser
 485 490 495

Pro Asp Gly Tyr Leu Gln Ile Gly Ser Phe Tyr Lys Gly Val Ala Glu
 500 505 510

Gly Glu Val Asp Pro Ala Phe Gly Pro Leu Glu Ala Leu Arg Leu Ser
 515 520 525

Ile Gln Thr Asp Ser Pro Val Trp Val Ile Leu Ser Glu Ile Phe Leu
 530 535 540

Lys Lys Ala Asp
 545

<210> 24
 <211> 563
 <212> PRT
 <213> Homo sapiens

5

<400> 24
 Met Ser Arg Val Ala Gly Thr Arg Thr Asp Val Asn Glu Leu Leu Gln
 1 5 10 15

Arg Trp Thr Pro Arg Cys Val Arg Trp His Thr Gly Gly Ala Arg Arg
 20 25 30

Val Ala Leu Asp Arg Pro Leu Val Thr Ala Cys Leu Pro Pro Ala Gly
 35 40 45

Asp Val Val Asp Val Tyr Gln Arg Glu Phe Leu Ala Leu Arg Asp Arg
 50 55 60

Leu His Ala Ala Glu Gln Glu Ser Leu Lys Arg Ser Lys Glu Leu Asn
 65 70 75 80

Leu Val Leu Asp Glu Ile Lys Arg Ala Val Ser Glu Arg Gln Ala Leu
 85 90 95

Arg Asp Gly Asp Gly Asn Arg Thr Trp Gly Arg Leu Thr Glu Asp Pro
 100 105 110

Arg Leu Lys Pro Trp Asn Gly Ser His Arg His Val Leu His Leu Pro
 115 120 125

Thr Val Phe His His Leu Pro His Leu Leu Ala Lys Glu Ser Ser Leu
 130 135 140

10

ES 2 659 118 T3

Gln Pro Ala Val Arg Val Gly Gln Gly Arg Thr Gly Val Ser Val Val
 145 150 155 160
 Met Gly Ile Pro Ser Val Arg Arg Glu Val His Ser Tyr Leu Thr Asp
 165 170 175
 Thr Leu His Ser Leu Ile Ser Glu Leu Ser Pro Gln Glu Lys Glu Asp
 180 185 190
 Ser Val Ile Val Val Leu Ile Ala Glu Thr Asp Ser Gln Tyr Thr Ser
 195 200 205
 Ala Val Thr Glu Asn Ile Lys Ala Leu Phe Pro Thr Glu Ile His Ser
 210 215 220
 Gly Leu Leu Glu Val Ile Ser Pro Ser Pro His Phe Tyr Pro Asp Phe
 225 230 235 240
 Ser Arg Leu Arg Glu Ser Phe Gly Asp Pro Lys Glu Arg Val Arg Trp
 245 250 255
 Arg Thr Lys Gln Asn Leu Asp Tyr Cys Phe Leu Met Met Tyr Ala Gln
 260 265 270
 Ser Lys Gly Ile Tyr Tyr Val Gln Leu Glu Asp Asp Ile Val Ala Lys
 275 280 285
 Pro Asn Tyr Leu Ser Thr Met Lys Asn Phe Ala Leu Gln Gln Pro Ser
 290 295 300
 Glu Asp Trp Met Ile Leu Glu Phe Ser Gln Leu Gly Phe Ile Gly Lys
 305 310 315 320
 Met Phe Lys Ser Leu Asp Leu Ser Leu Ile Val Glu Phe Ile Leu Met
 325 330 335
 Phe Tyr Arg Asp Lys Pro Ile Asp Trp Leu Leu Asp His Ile Leu Trp
 340 345 350
 Val Lys Val Cys Asn Pro Glu Lys Asp Ala Lys His Cys Asp Arg Gln
 355 360 365
 Lys Ala Asn Leu Arg Ile Arg Phe Lys Pro Ser Leu Phe Gln His Val
 370 375 380
 Gly Thr His Ser Ser Leu Ala Gly Lys Ile Gln Lys Leu Lys Asp Lys
 385 390 395 400
 Asp Phe Gly Lys Gln Ala Leu Arg Lys Glu His Val Asn Pro Pro Ala
 405 410 415

ES 2 659 118 T3

Glu Val Ser Thr Ser Leu Lys Thr Tyr Gln His Phe Thr Leu Glu Lys
 420 425 430

Ala Tyr Leu Arg Glu Asp Phe Phe Trp Ala Phe Thr Pro Ala Ala Gly
 435 440 445

Asp Phe Ile Arg Phe Arg Phe Phe Gln Pro Leu Arg Leu Glu Arg Phe
 450 455 460

Phe Phe Arg Ser Gly Asn Ile Glu His Pro Glu Asp Lys Leu Phe Asn
 465 470 475 480

Thr Ser Val Glu Val Leu Pro Phe Asp Asn Pro Gln Ser Asp Lys Glu
 485 490 495

Ala Leu Gln Glu Gly Arg Thr Ala Thr Leu Arg Tyr Pro Arg Ser Pro
 500 505 510

Asp Gly Tyr Leu Gln Ile Gly Ser Phe Tyr Lys Gly Val Ala Glu Gly
 515 520 525

Glu Val Asp Pro Ala Phe Gly Pro Leu Glu Ala Leu Arg Leu Ser Ile
 530 535 540

Gln Thr Asp Ser Pro Val Trp Val Ile Leu Ser Glu Ile Phe Leu Lys
 545 550 555 560

Lys Ala Asp

<210> 25

<211> 478

5 <212> PRT

<213> Homo sapiens

<400> 25

Met Phe Lys Phe His Gln Met Lys His Ile Phe Glu Ile Leu Asp Lys
 1 5 10 15

Met Arg Cys Leu Arg Lys Arg Ser Thr Val Ser Phe Leu Gly Val Leu
 20 25 30

Val Ile Phe Leu Leu Phe Met Asn Leu Tyr Ile Glu Asp Ser Tyr Val
 35 40 45

Leu Glu Gly Asp Lys Gln Leu Ile Arg Glu Thr Ser Thr His Gln Leu
 50 55 60

Asn Ser Glu Arg Tyr Val His Thr Phe Lys Asp Leu Ser Asn Phe Ser
 65 70 75 80

10

ES 2 659 118 T3

Gly Ala Ile Asn Val Thr Tyr Arg Tyr Leu Ala Ala Thr Pro Leu Gln
85 90 95

Arg Lys Arg Tyr Leu Thr Ile Gly Leu Ser Ser Val Lys Arg Lys Lys
100 105 110

Gly Asn Tyr Leu Leu Glu Thr Ile Lys Ser Ile Phe Glu Gln Ser Ser
115 120 125

Tyr Glu Glu Leu Lys Glu Ile Ser Val Val Val His Leu Ala Asp Phe
130 135 140

Asn Ser Ser Trp Arg Asp Ala Met Val Gln Asp Ile Thr Gln Lys Phe
145 150 155 160

Ala His His Ile Ile Ala Gly Arg Leu Met Val Ile His Ala Pro Glu
165 170 175

Glu Tyr Tyr Pro Ile Leu Asp Gly Leu Lys Arg Asn Tyr Asn Asp Pro
180 185 190

Glu Asp Arg Val Lys Phe Arg Ser Lys Gln Asn Val Asp Tyr Ala Phe
195 200 205

Leu Leu Asn Phe Cys Ala Asn Thr Ser Asp Tyr Tyr Val Met Leu Glu
210 215 220

Asp Asp Val Arg Cys Ser Lys Asn Phe Leu Thr Ala Ile Lys Lys Val
225 230 235 240

Ile Ala Ser Leu Glu Gly Thr Tyr Trp Val Thr Leu Glu Phe Ser Lys
245 250 255

Leu Gly Tyr Ile Gly Lys Leu Tyr His Ser His Asp Leu Pro Arg Leu
260 265 270

Ala His Phe Leu Leu Met Phe Tyr Gln Glu Met Pro Cys Asp Trp Leu
275 280 285

Leu Thr His Phe Arg Gly Leu Leu Ala Gln Lys Asn Val Ile Arg Phe
290 295 300

Lys Pro Ser Leu Phe Gln His Met Gly Tyr Tyr Ser Ser Tyr Lys Gly
305 310 315 320

Thr Glu Asn Lys Leu Lys Asp Asp Asp Phe Glu Glu Glu Ser Phe Asp
325 330 335

Ile Pro Asp Asn Pro Pro Ala Ser Leu Tyr Thr Asn Met Asn Val Phe
340 345 350

ES 2 659 118 T3

Glu Asn Tyr Glu Ala Ser Lys Ala Tyr Ser Ser Val Asp Glu Tyr Phe
355 360 365

Trp Gly Lys Pro Pro Ser Thr Gly Asp Val Phe Val Ile Val Phe Glu
370 375 380

Asn Pro Ile Ile Ile Lys Lys Ile Lys Val Asn Thr Gly Thr Glu Asp
385 390 395 400

Arg Gln Asn Asp Ile Leu His His Gly Ala Leu Asp Val Gly Glu Asn
405 410 415

Val Met Pro Ser Lys Gln Arg Arg Gln Cys Ser Thr Tyr Leu Arg Leu
420 425 430

Gly Glu Phe Lys Asn Gly Asn Phe Glu Met Ser Gly Val Asn Gln Lys
435 440 445

Ile Pro Phe Asp Ile His Cys Met Arg Ile Tyr Val Thr Lys Thr Gln
450 455 460

Lys Glu Trp Leu Ile Ile Arg Ser Ile Ser Ile Trp Thr Ser
465 470 475

<210> 26

<211> 740

<212> PRT

<213> Mus musculus

5

<400> 26

Met Ala Phe Phe Ser Pro Trp Lys Leu Ser Ser Gln Lys Leu Gly Phe
1 5 10 15

Phe Leu Val Thr Phe Gly Phe Ile Trp Gly Met Met Leu Leu His Phe
20 25 30

Thr Ile Gln Gln Arg Thr Gln Pro Glu Ser Ser Ser Met Leu Arg Glu
35 40 45

Gln Ile Leu Asp Leu Ser Lys Arg Tyr Ile Lys Ala Leu Ala Glu Glu
50 55 60

Asn Arg Asp Val Val Asp Gly Pro Tyr Ala Gly Val Met Thr Ala Tyr
65 70 75 80

Asp Leu Lys Lys Thr Leu Ala Val Leu Leu Asp Asn Ile Leu Gln Arg
85 90 95

Ile Gly Lys Leu Glu Ser Lys Val Asp Asn Leu Val Asn Gly Thr Gly
100 105 110

10

ES 2 659 118 T3

Ala Asn Ser Thr Asn Ser Thr Thr Ala Val Pro Ser Leu Val Ser Leu
115 120 125

Glu Lys Ile Asn Val Ala Asp Ile Ile Asn Gly Val Gln Glu Lys Cys
130 135 140

Val Leu Pro Pro Met Asp Gly Tyr Pro His Cys Glu Gly Lys Ile Lys
145 150 155 160

Trp Met Lys Asp Met Trp Arg Ser Asp Pro Cys Tyr Ala Asp Tyr Gly
165 170 175

Val Asp Gly Thr Ser Cys Ser Phe Phe Ile Tyr Leu Ser Glu Val Glu
180 185 190

Asn Trp Cys Pro Arg Leu Pro Trp Arg Ala Lys Asn Pro Tyr Glu Glu
195 200 205

Ala Asp His Asn Ser Leu Ala Glu Ile Arg Thr Asp Phe Asn Ile Leu
210 215 220

Tyr Gly Met Met Lys Lys His Glu Glu Phe Arg Trp Met Arg Leu Arg
225 230 235 240

Ile Arg Arg Met Ala Asp Ala Trp Ile Gln Ala Ile Lys Ser Leu Ala
245 250 255

Glu Lys Gln Asn Leu Glu Lys Arg Lys Arg Lys Lys Ile Leu Val His
260 265 270

Leu Gly Leu Leu Thr Lys Glu Ser Gly Phe Lys Ile Ala Glu Thr Ala
275 280 285

Phe Ser Gly Gly Pro Leu Gly Glu Leu Val Gln Trp Ser Asp Leu Ile
290 295 300

Thr Ser Leu Tyr Leu Leu Gly His Asp Ile Arg Ile Ser Ala Ser Leu
305 310 315 320

Ala Glu Leu Lys Glu Ile Met Lys Lys Val Val Gly Asn Arg Ser Gly
325 330 335

Cys Pro Thr Val Gly Asp Arg Ile Val Glu Leu Ile Tyr Ile Asp Ile
340 345 350

Val Gly Leu Ala Gln Phe Lys Lys Thr Leu Gly Pro Ser Trp Val His
355 360 365

Tyr Gln Cys Met Leu Arg Val Leu Asp Ser Phe Gly Thr Glu Pro Glu
370 375 380

ES 2 659 118 T3

Phe Asn His Ala Ser Tyr Ala Gln Ser Lys Gly His Lys Thr Pro Trp
 385 390 395 400
 Gly Lys Trp Asn Leu Asn Pro Gln Gln Phe Tyr Thr Met Phe Pro His
 405 410 415
 Thr Pro Asp Asn Ser Phe Leu Gly Phe Val Val Glu Gln His Leu Asn
 420 425 430
 Ser Ser Asp Ile His His Ile Asn Glu Ile Lys Arg Gln Asn Gln Ser
 435 440
 Leu Val Tyr Gly Lys Val Asp Ser Phe Trp Lys Asn Lys Lys Ile Tyr
 450 455 460
 Leu Asp Ile Ile His Thr Tyr Met Glu Val His Ala Thr Val Tyr Gly
 465 470 475 480
 Ser Ser Thr Lys Asn Ile Pro Ser Tyr Val Lys Asn His Gly Ile Leu
 485 490 495
 Ser Gly Arg Asp Leu Gln Phe Leu Leu Arg Glu Thr Lys Leu Phe Val
 500 505 510
 Gly Leu Gly Phe Pro Tyr Glu Gly Pro Ala Pro Leu Glu Ala Ile Ala
 515 520 525
 Asn Gly Cys Ala Phe Leu Asn Pro Lys Phe Asn Pro Pro Lys Ser Ser
 530 535 540
 Lys Asn Thr Asp Phe Phe Ile Gly Lys Pro Thr Leu Arg Glu Leu Thr
 545 550 555 560
 Ser Gln His Pro Tyr Ala Glu Val Phe Ile Gly Arg Pro His Val Trp
 565 570 575
 Thr Val Asp Leu Asn Asn Arg Glu Glu Val Glu Asp Ala Val Lys Ala
 580 585 590
 Ile Leu Asn Gln Lys Ile Glu Pro Tyr Met Pro Tyr Glu Phe Thr Cys
 595 600 605
 Glu Gly Met Leu Gln Arg Ile Asn Ala Phe Ile Glu Lys Gln Asp Phe
 610 615 620
 Cys His Gly Gln Val Met Trp Pro Pro Leu Ser Ala Leu Gln Val Lys
 625 630 635 640
 Leu Ala Glu Pro Gly Gln Ser Cys Lys Gln Val Cys Gln Glu Ser Gln
 645 650 655

ES 2 659 118 T3

Leu Ile Cys Glu Pro Ser Phe Phe Gln His Leu Asn Lys Glu Lys Asp
660 665 670

Leu Leu Lys Tyr Lys Val Thr Cys Gln Ser Ser Glu Leu Tyr Lys Asp
675 680 685

Ile Leu Val Pro Ser Phe Tyr Pro Lys Ser Lys His Cys Val Phe Gln
690 695 700

Gly Asp Leu Leu Leu Phe Ser Cys Ala Gly Ala His Pro Thr His Gln
705 710 715 720

Arg Ile Cys Pro Cys Arg Asp Phe Ile Lys Gly Gln Val Ala Leu Cys
725 730 735

Lys Asp Cys Leu
740

<210> 27

<211> 792

5 <212> PRT

<213> Mus musculus

<400> 27

Met Ile Thr Val Asn Pro Asp Gly Lys Ile Met Val Arg Arg Cys Leu
1 5 10 15

Val Thr Leu Arg Pro Phe Arg Leu Phe Val Leu Gly Ile Gly Phe Phe
20 25 30

Thr Leu Cys Phe Leu Met Thr Ser Leu Gly Gly Gln Phe Ser Ala Arg
35 40 45

Arg Leu Gly Asp Ser Pro Phe Thr Ile Arg Thr Glu Val Pro Gly Ser
50 55 60

Pro Glu Ser Arg Gly Ala Leu Arg Lys Met Ser Asp Leu Leu Glu Leu
65 70 75 80

Met Val Lys Arg Met Asp Met Leu Ala Arg Leu Glu Asn Ser Ser Glu
85 90 95

Leu His Arg Thr Ala Ser Val Ala His Leu Ala Ala Asp Arg Leu Thr
100 105 110

Pro Gly Ala Ser Leu Ile Glu Arg Ile Gln Ala Ile Ala Gln Asn Val
115 120 125

Ser Asp Ile Ala Val Lys Val Asp Gln Ile Leu Arg His Ser Leu Ile
130 135 140

Leu His Ser Lys Val Ser Glu Gly Arg Arg Asp Gln Cys Glu Ala Pro

10

ES 2 659 118 T3

145 150 155 160
 Ser Asp Pro Lys Phe Pro Asp Cys Ser Gly Lys Val Glu Trp Met Arg
 165 170 175
 Ala Arg Trp Thr Ser Asp Pro Cys Tyr Ala Phe Phe Gly Val Asp Gly
 180 185 190
 Thr Glu Cys Ser Phe Leu Ile Tyr Leu Ser Glu Val Glu Trp Phe Cys
 195 200 205
 Pro Pro Leu Pro Trp Arg Asn Gln Thr Ala Ala Arg Thr Ala Pro Lys
 210 215 220
 Ser Leu Pro Arg Val Gln Ala Val Phe Arg Ser Asn Leu Ser His Leu
 225 230 235 240
 Leu Glu Leu Met Gly Ser Gly Lys Glu Ser Leu Ile Phe Met Lys Lys
 245 250 255
 Arg Thr Arg Arg Phe Thr Ala Gln Trp Thr Lys Ala Ala Lys Tyr Leu
 260 265 270
 Ala Gln Lys Leu Gly Asp Ile Arg Arg Asp Gln Lys Gln Ile Leu Val
 275 280 285
 His Ile Gly Phe Leu Thr Glu Glu Ser Gly Asp Val Phe Ser Pro Arg
 290 295 300
 Val Leu Lys Gly Gly Pro Leu Gly Glu Met Val Gln Trp Ala Asp Ile
 305 310 315 320
 Leu Ala Ala Leu Tyr Val Leu Gly His Ser Leu Arg Ile Thr Val Ser
 325 330 335
 Leu Lys Glu Leu Gln Ser Asn Leu Gly Val Pro Pro Gly Arg Gly Asn
 340 345 350
 Cys Pro Leu Thr Val Pro Leu Pro Phe Asp Leu Ile Tyr Thr Asp Tyr
 355 360 365
 His Gly Leu Gln Gln Met Lys Gln His Met Gly Leu Ser Phe Lys Lys
 370 375 380
 Tyr Arg Cys Arg Ile Arg Val Ile Asp Thr Phe Gly Thr Glu Pro Ala
 385 390 395 400
 Tyr Asn His Glu Glu Tyr Ala Thr Leu His Gly Tyr Arg Thr Asn Trp
 405 410 415
 Gly Tyr Trp Asn Leu Asn Pro Lys Gln Phe Met Thr Met Phe Pro His

ES 2 659 118 T3

420 425 430
 Thr Pro Asp Asn Ser Phe Met Gly Phe Val Ser Glu Glu Leu Asn Glu
 435 440 445
 Thr Glu Lys Gln Leu Ile Lys Asp Gly Lys Ala Ser Asn Met Ala Val
 450 455 460
 Val Tyr Gly Lys Glu Ala Ser Ile Trp Lys Leu Gln Gly Lys Glu Lys
 465 470 475 480
 Phe Leu Ala Val Leu Asn Lys Tyr Met Glu Ile His Gly Thr Val Tyr
 485 490
 Tyr Glu Ser Gln Arg Pro Pro Glu Val Pro Ala Phe Val Lys Asn His
 500 505 510
 Gly Leu Leu Pro Gln Pro Glu Phe Gln Gln Leu Leu Arg Lys Ala Lys
 515 520 525
 Leu Phe Ile Gly Phe Gly Phe Pro Tyr Glu Gly Pro Ala Pro Leu Glu
 530 535 540
 Ala Ile Ala Asn Gly Cys Ile Phe Leu Gln Ser Arg Phe Ser Pro Pro
 545 550 555 560
 His Ser Ser Leu Asn His Glu Phe Phe Arg Gly Lys Pro Thr Ser Arg
 565 570 575
 Glu Val Phe Ser Gln His Pro Tyr Ala Glu Asn Phe Ile Gly Lys Pro
 580 585 590
 His Val Trp Thr Val Asp Tyr Asn Asn Ser Asp Glu Phe Glu Thr Ala
 595 600
 Ile Lys Ala Ile Met Asn Thr Gln Val Asp Pro Tyr Leu Pro Tyr Glu
 610 615 620
 Tyr Thr Cys Ala Gly Met Leu Glu Arg Ile Asn Ala Tyr Ile Gln His
 625 630 635 640
 Gln Asp Phe Cys Val Gly Pro Ser Pro Leu Pro Pro Gly Ala Ser Thr
 645
 Ala Gln Ser Pro Phe Val Leu Ala Pro Asn Ala Thr His Leu Glu Trp
 660 665 670
 Ala Gln Asn Ile Ser Ser Val Pro Gly Ala Trp Pro Pro Thr His Ser
 675 680 685
 Leu Arg Ala Trp Leu Ala Ala Pro Gly Arg Ala Cys Thr Asp Ala Cys

ES 2 659 118 T3

690 695 ~ 700 ~

Leu Asp His Gly Leu Ile Cys Glu Pro Ser Phe Phe Pro Phe Leu Asn
705 710 715 720

Ser Gln Asn Ser Phe Leu Lys Leu Gln Val Pro Cys Asp Ser Thr Glu
725 730 735

Trp Glu Met His His Leu Tyr Pro Ala Phe Ala Gln Pro Gly Gln Glu
740 745 750

Cys Tyr Leu Gln Lys Glu Pro Leu Leu Phe Ser Cys Ala Gly Ala Ser
755 760 765

Thr Lys Tyr Gln Arg Leu Cys Pro Cys Arg Asp Phe Arg Lys Gly Gln
770 775 780

Val Ala Leu Cys Gln Gly Cys Leu
785 790

<210> 28

<211> 741

5 <212> PRT

<213> Homo sapiens

<400> 28

Met Ala Leu Phe Thr Pro Trp Lys Leu Ser Ser Gln Lys Leu Gly Phe
1 5 10 15

Phe Leu Val Thr Phe Gly Phe Ile Trp Gly Met Met Leu Leu His Phe
20 25 30

Thr Ile Gln Gln Arg Thr Gln Pro Glu Ser Ser Ser Met Leu Arg Glu
35 40 45

Gln Ile Leu Asp Leu Ser Lys Arg Tyr Ile Lys Ala Leu Ala Glu Glu
50 55 60

Asn Arg Asn Val Val Asp Gly Pro Tyr Ala Gly Val Met Thr Ala Tyr
65 70 75 80

Asp Leu Lys Lys Thr Leu Ala Val Leu Leu Asp Asn Ile Leu Gln Arg
85 90 95

Ile Gly Lys Leu Glu Ser Lys Val Asp Asn Leu Val Val Asn Gly Thr
100 105 110

Gly Thr Asn Ser Thr Asn Ser Thr Thr Ala Val Pro Ser Leu Val Ala
115 120 125

Leu Glu Lys Ile Asn Val Ala Asp Ile Ile Asn Gly Ala Gln Glu Lys
130 135 140

10

ES 2 659 118 T3

Cys Val Leu Pro Pro Met Asp Gly Tyr Pro His Cys Glu Gly Lys Ile
 145 150 155 160
 Lys Trp Met Lys Asp Met Trp Arg Ser Asp Pro Cys Tyr Ala Asp Tyr
 165 170 175
 Gly Val Asp Gly Ser Thr Cys Ser Phe Phe Ile Tyr Leu Ser Glu Val
 180 185 190
 Glu Asn Trp Cys Pro His Leu Pro Trp Arg Ala Lys Asn Pro Tyr Glu
 195 200 205
 Glu Ala Asp His Asn Ser Leu Ala Glu Ile Arg Thr Asp Phe Asn Ile
 210 215 220
 Leu Tyr Ser Met Met Lys Lys His Glu Glu Phe Arg Trp Met Arg Leu
 225 230 235 240
 Arg Ile Arg Arg Met Ala Asp Ala Trp Ile Gln Ala Ile Lys Ser Leu
 245 250 255
 Ala Glu Lys Gln Asn Leu Glu Lys Arg Lys Arg Lys Lys Val Leu Val
 260 265 270
 His Leu Gly Leu Leu Thr Lys Glu Ser Gly Phe Lys Ile Ala Glu Thr
 275 280 285
 Ala Phe Ser Gly Gly Pro Leu Gly Glu Leu Val Gln Trp Ser Asp Leu
 290 295 300
 Ile Thr Ser Leu Tyr Leu Leu Gly His Asp Ile Arg Ile Ser Ala Ser
 305 310 315 320
 Leu Ala Glu Leu Lys Glu Ile Met Lys Lys Val Val Gly Asn Arg Ser
 325 330 335
 Gly Cys Pro Thr Val Gly Asp Arg Ile Val Glu Leu Ile Tyr Ile Asp
 340 345 350
 Ile Val Gly Leu Ala Gln Phe Lys Lys Thr Leu Gly Pro Ser Trp Val
 355 360 365
 His Tyr Gln Cys Met Leu Arg Val Leu Asp Ser Phe Gly Thr Glu Pro
 370 375 380
 Glu Phe Asn His Ala Asn Tyr Ala Gln Ser Lys Gly His Lys Thr Pro
 385 390 395 400
 Trp Gly Lys Trp Asn Leu Asn Pro Gln Gln Phe Tyr Thr Met Phe Pro
 405 410 415

ES 2 659 118 T3

His Thr Pro Asp 420 Asn Ser Phe Leu Gly 425 Phe Val Val Glu Gln His Leu 430
 Asn Ser Ser 435 Asp Ile His His Ile 440 Asn Glu Ile Lys Arg 445 Gln Asn Gln
 Ser Leu Val Tyr Gly Lys Val 455 Asp Ser Phe Trp Lys 460 Asn Lys Lys Ile
 Tyr 465 Leu Asp Ile Ile His 470 Thr Tyr Met Glu Val 475 His Ala Thr Val Tyr 480
 Gly Ser Ser Thr Lys 485 Asn Ile Pro Ser Tyr Val 490 Lys Asn His Gly Ile 495
 Leu Ser Gly Arg 500 Asp Leu Gln Phe Leu 505 Leu Arg Glu Thr Lys 510 Leu Phe
 Val Gly Leu 515 Gly Phe Pro Tyr Glu 520 Gly Pro Ala Pro Leu 525 Glu Ala Ile
 Ala Asn Gly Cys Ala Phe Leu 535 Asn Pro Lys Phe Asn 540 Pro Pro Lys Ser
 Ser 545 Lys Asn Thr Asp Phe 550 Phe Ile Gly Lys Pro 555 Thr Leu Arg Glu Leu 560
 Thr Ser Gln His Pro 565 Tyr Ala Glu Val Phe 570 Ile Gly Arg Pro His Val 575
 Trp Thr Val Asp 580 Leu Asn Asn Gln Glu 585 Glu Val Glu Asp Ala Val Lys 590
 Ala Ile Leu 595 Asn Gln Lys Ile Glu 600 Pro Tyr Met Pro Tyr 605 Glu Phe Thr
 Cys Glu Gly Met Leu Gln Arg Ile 615 Asn Ala Phe Ile 620 Glu Lys Gln Asp
 Phe Cys His Gly Gln Val 630 Met Trp Pro Pro Leu Ser Ala Leu Gln Val 640
 Lys Leu Ala Glu Pro 645 Gly Gln Ser Cys Lys 650 Gln Val Cys Gln Glu Ser 655
 Gln Leu Ile Cys 660 Glu Pro Ser Phe Phe 665 Gln His Leu Asn Lys Asp Lys 670
 Asp Met Leu 675 Lys Tyr Lys Val Thr 680 Cys Gln Ser Ser Glu 685 Leu Ala Lys
 Asp Ile Leu Val Pro Ser Phe 695 Asp Pro Lys Asn Lys 700 His Cys Val Phe
 Gln Gly Asp Leu Leu 710 Phe Ser Cys Ala Gly 715 Ala His Pro Arg His 720
 Gln Arg Val Cys Pro 725 Cys Arg Asp Phe Ile 730 Lys Gly Gln Val Ala Leu 735
 Cys Lys Asp Cys 740 Leu

ES 2 659 118 T3

<210> 29
 <211> 790
 <212> PRT
 <213> Homo sapiens

Á

<400> 29
 Met Ile Thr Val Asn Pro Asp Gly Lys Ile Met Val Arg Arg Cys Leu
 1 5 10 15
 Val Thr Leu Arg Pro Phe Arg Leu Phe Val Leu Gly Ile Gly Phe Phe
 20 25 30
 Thr Leu Cys Phe Leu Met Thr Ser Leu Gly Gly Gln Phe Ser Ala Arg
 35 40 45
 Arg Leu Gly Asp Ser Pro Phe Thr Ile Arg Thr Glu Val Met Gly Gly
 50 55 60
 Pro Glu Ser Arg Gly Val Leu Arg Lys Met Ser Asp Leu Leu Glu Leu
 65 70 75 80
 Met Val Lys Arg Met Asp Ala Leu Ala Arg Leu Glu Asn Ser Ser Glu
 85 90 95
 Leu His Arg Ala Gly Gly Asp Leu His Phe Pro Ala Asp Arg Met Pro
 100 105 110
 Pro Gly Ala Gly Leu Met Glu Arg Ile Gln Ala Ile Ala Gln Asn Val
 115 120 125
 Ser Asp Ile Ala Val Lys Val Asp Gln Ile Leu Arg His Ser Leu Leu
 130 135 140
 Leu His Ser Lys Val Ser Glu Gly Arg Arg Asp Gln Cys Glu Ala Pro
 145 150 155 160
 Ser Asp Pro Lys Phe Pro Asp Cys Ser Gly Lys Val Glu Trp Met Arg
 165 170 175

ES 2 659 118 T3

Ala Arg Trp Thr Ser Asp Pro Cys Tyr Ala Phe Phe Gly Val Asp Gly
 180 185 190

Thr Glu Cys Ser Phe Leu Ile Tyr Leu Ser Glu Val Glu Trp Phe Cys
 195 200 205

Pro Pro Leu Pro Trp Arg Asn Gln Thr Ala Ala Gln Arg Ala Pro Lys
 210 215 220

Pro Leu Pro Lys Val Gln Ala Val Phe Arg Ser Asn Leu Ser His Leu
 225 230 235 240

Leu Asp Leu Met Gly Ser Gly Lys Glu Ser Leu Ile Phe Met Lys Lys
 245 250 255

Arg Thr Lys Arg Leu Thr Ala Gln Trp Ala Leu Ala Ala Gln Arg Leu
 260 265 270

Ala Gln Lys Leu Gly Ala Thr Gln Arg Asp Gln Lys Gln Ile Leu Val
 275 280 285

His Ile Gly Phe Leu Thr Glu Glu Ser Gly Asp Val Phe Ser Pro Arg
 290 295 300

Val Leu Lys Gly Gly Pro Leu Gly Glu Met Val Gln Trp Ala Asp Ile
 305 310 315 320

Leu Thr Ala Leu Tyr Val Leu Gly His Gly Leu Arg Val Thr Val Ser
 325 330 335

Leu Lys Glu Leu Gln Ser Asn Leu Gly Val Pro Pro Gly Arg Gly Ser
 340 345 350

Cys Pro Leu Thr Met Pro Leu Pro Phe Asp Leu Ile Tyr Thr Asp Tyr
 355 360 365

His Gly Leu Gln Gln Met Lys Arg His Met Gly Leu Ser Phe Lys Lys
 370 375 380

Tyr Arg Cys Arg Ile Arg Val Ile Asp Thr Phe Gly Thr Glu Pro Ala
 385 390 395 400

Tyr Asn His Glu Glu Tyr Ala Thr Leu His Gly Tyr Arg Thr Asn Trp
 405 410 415

Gly Tyr Trp Asn Leu Asn Pro Lys Gln Phe Met Thr Met Phe Pro His
 420 425 430

Thr Pro Asp Asn Ser Phe Met Gly Phe Val Ser Glu Glu Leu Asn Glu
 435 440 445

ES 2 659 118 T3

Thr Glu Lys Arg Leu Ile Lys Gly Gly Lys Ala Ser Asn Met Ala Val
 450 455 460
 Val Tyr Gly Lys Glu Ala Ser Ile Trp Lys Gly Lys Glu Lys Phe Leu
 465 470 475 480
 Gly Ile Leu Asn Lys Tyr Met Glu Ile His Gly Thr Val Tyr Tyr Glu
 485 490 495
 Ser Gln Arg Pro Pro Glu Val Pro Ala Phe Val Lys Asn His Gly Leu
 500 505 510
 Leu Pro Gln Pro Glu Phe Gln Gln Leu Leu Arg Lys Ala Lys Leu Phe
 515 520 525
 Ile Gly Phe Gly Phe Pro Tyr Glu Gly Pro Ala Pro Leu Glu Ala Ile
 530 535 540
 Ala Asn Gly Cys Ile Phe Leu Gln Ser Arg Phe Ser Pro Pro His Ser
 545 550 555 560
 Ser Leu Asn His Glu Phe Phe Arg Gly Lys Pro Thr Ser Arg Glu Val
 565 570 575
 Phe Ser Gln His Pro Tyr Ala Glu Asn Phe Ile Gly Lys Pro His Val
 580 585 590
 Trp Thr Val Asp Tyr Asn Asn Ser Glu Glu Phe Glu Ala Ala Ile Lys
 595 600 605
 Ala Ile Met Arg Thr Gln Val Asp Pro Tyr Leu Pro Tyr Glu Tyr Thr
 610 615 620
 Cys Glu Gly Met Leu Glu Arg Ile His Ala Tyr Ile Gln His Gln Asp
 625 630 635 640
 Phe Cys Arg Ala Pro Asp Pro Ala Leu Pro Glu Ala His Ala Pro Gln
 645 650 655
 Ser Pro Phe Val Leu Ala Pro Asn Ala Thr His Leu Glu Trp Ala Arg
 660 665 670
 Asn Thr Ser Leu Ala Pro Gly Ala Trp Pro Pro Ala His Ala Leu Arg
 675 680 685
 Ala Trp Leu Ala Val Pro Gly Arg Ala Cys Thr Asp Thr Cys Leu Asp
 690 695 700
 His Gly Leu Ile Cys Glu Pro Ser Phe Phe Pro Phe Leu Asn Ser Gln
 705 710 715 720

ES 2 659 118 T3

Asp Ala Phe Leu Lys Leu Gln Val Pro Cys Asp Ser Thr Glu Ser Glu
725 730 735

Met Asn His Leu Tyr Pro Ala Phe Ala Gln Pro Gly Gln Glu Cys Tyr
740 745 750

Leu Gln Lys Glu Pro Leu Leu Phe Ser Cys Ala Gly Ser Asn Thr Lys
755 760 765

Tyr Arg Arg Leu Cys Pro Cys Arg Asp Phe Arg Lys Gly Gln Val Ala
770 775 780

Leu Cys Gln Gly Cys Leu
785 790

<210> 30

<211> 801

<212> PRT

<213> Homo sapiens

5

<400> 30

Met His Ser Phe Val Lys His Leu Cys Ser Arg Tyr Val Val Glu Arg
1 5 10 15

Gln Gly Thr Met Ala Leu Pro Ala Leu Leu Thr Arg Leu Leu Pro Leu
20 25 30

Arg Arg Leu Phe Val Leu Gly Ile Gly Phe Phe Thr Leu Cys Phe Leu
35 40 45

Met Thr Ser Leu Gly Gly Gln Phe Ser Ala Arg Arg Leu Gly Asp Ser
50 55 60

Pro Phe Thr Ile Arg Thr Glu Val Met Gly Gly Pro Glu Ser Arg Gly
65 70 75 80

Val Leu Arg Lys Met Ser Asp Leu Leu Glu Leu Met Val Lys Arg Met
85 90 95

Asp Ala Leu Ala Arg Leu Glu Asn Ser Ser Glu Leu His Arg Ala Gly
100 105 110

Gly Asp Leu His Phe Pro Ala Asp Arg Met Pro Pro Gly Ala Gly Leu
115 120 125

Met Glu Arg Ile Gln Ala Ile Ala Gln Asn Val Ser Asp Ile Ala Val
130 135 140

Lys Val Asp Gln Ile Leu Arg His Ser Leu Leu Leu His Ser Lys Val
145 150 155 160

10

ES 2 659 118 T3

Ser Glu Gly Arg Arg Asp Gln Cys Glu Ala Pro Ser Asp Pro Lys Phe
 165 170 175

Pro Asp Cys Ser Gly Lys Val Glu Trp Met Arg Ala Arg Trp Thr Ser
 180 185 190

Asp Pro Cys Tyr Ala Phe Phe Gly Val Asp Gly Thr Glu Cys Ser Phe
 195 200 205

Leu Ile Tyr Leu Ser Glu Val Glu Trp Phe Cys Pro Pro Leu Pro Trp
 210 215 220

Arg Asn Gln Thr Ala Ala Gln Arg Ala Pro Lys Pro Leu Pro Lys Val
 225 230 240

Gln Ala Val Phe Arg Ser Asn Leu Ser His Leu Leu Asp Leu Met Gly
 245 250 255

Ser Gly Lys Glu Ser Leu Ile Phe Met Lys Lys Arg Thr Lys Arg Leu
 260 265 270

Thr Ala Gln Trp Ala Leu Ala Ala Gln Arg Leu Ala Gln Lys Leu Gly
 275 280 285

Ala Thr Gln Arg Asp Gln Lys Gln Ile Leu Val His Ile Gly Phe Leu
 290 295 300

Thr Glu Glu Ser Gly Asp Val Phe Ser Pro Arg Val Leu Lys Gly Gly
 305 310 315 320

Pro Leu Gly Glu Met Val Gln Trp Ala Asp Ile Leu Thr Ala Leu Tyr
 325 330 335

Val Leu Gly His Gly Leu Arg Val Thr Val Ser Leu Lys Glu Leu Gln
 340 345 350

Ser Asn Leu Gly Val Pro Pro Gly Arg Gly Ser Cys Pro Leu Thr Met
 355 360 365

Pro Leu Pro Phe Asp Leu Ile Tyr Thr Asp Tyr His Gly Leu Gln Gln
 370 375 380

Met Lys Arg His Met Gly Leu Ser Phe Lys Lys Tyr Arg Cys Arg Ile
 385 390 395 400

Arg Val Ile Asp Thr Phe Gly Thr Glu Pro Ala Tyr Asn His Glu Glu
 405 410 415

Tyr Ala Thr Leu His Gly Tyr Arg Thr Asn Trp Gly Tyr Trp Asn Leu
 420 425 430

ES 2 659 118 T3

Asn Pro Lys Gln Phe Met Thr Met Phe Pro His Thr Pro Asp Asn Ser
 435 440 445
 Phe Met Gly Phe Val Ser Glu Glu Leu Asn Glu Thr Glu Lys Arg Leu
 450 455 460
 Ile Lys Gly Gly Lys Ala Ser Asn Met Ala Val Val Tyr Gly Lys Glu
 465 470 475 480
 Ala Ser Ile Trp Lys Gly Lys Glu Lys Phe Leu Gly Ile Leu Asn Lys
 485 490 495
 Tyr Met Glu Ile His Gly Thr Val Tyr Tyr Glu Ser Gln Arg Pro Pro
 500 505 510
 Glu Val Pro Ala Phe Val Lys Asn His Gly Leu Leu Pro Gln Pro Glu
 515 520 525
 Phe Gln Gln Leu Leu Arg Lys Ala Lys Leu Phe Ile Gly Phe Gly Phe
 530 535 540
 Pro Tyr Glu Gly Pro Ala Pro Leu Glu Ala Ile Ala Asn Gly Cys Ile
 545 550 555 560
 Phe Leu Gln Ser Arg Phe Ser Pro Pro His Ser Ser Leu Asn His Glu
 565 570 575
 Phe Phe Arg Gly Lys Pro Thr Ser Arg Glu Val Phe Ser Gln His Pro
 580 585 590
 Tyr Ala Glu Asn Phe Ile Gly Lys Pro His Val Trp Thr Val Asp Tyr
 595 600 605
 Asn Asn Ser Glu Glu Phe Glu Ala Ala Ile Lys Ala Ile Met Arg Thr
 610 615 620
 Gln Val Asp Pro Tyr Leu Pro Tyr Glu Tyr Thr Cys Glu Gly Met Leu
 625 630 635 640
 Glu Arg Ile His Ala Tyr Ile Gln His Gln Asp Phe Cys Arg Ala Pro
 645 650 655
 Asp Pro Ala Leu Pro Glu Ala His Ala Pro Gln Ser Pro Phe Val Leu
 660 665 670
 Ala Pro Asn Ala Thr His Leu Glu Trp Ala Arg Asn Thr Ser Leu Ala
 675 680 685
 Pro Gly Ala Trp Pro Pro Ala His Ala Leu Arg Ala Trp Leu Ala Val
 690 695 700

ES 2 659 118 T3

Pro Gly Arg Ala Cys Thr Asp Thr Cys Leu Asp His Gly Leu Ile Cys
705 710 715 720

Glu Pro Ser Phe Phe Pro Phe Leu Asn Ser Gln Asp Ala Phe Leu Lys
725 730 735

Leu Gln Val Pro Cys Asp Ser Thr Glu Ser Glu Met Asn His Leu Tyr
740 745 750

Pro Ala Phe Ala Gln Pro Gly Gln Glu Cys Tyr Leu Gln Lys Glu Pro
755 760 765

Leu Leu Phe Ser Cys Ala Gly Ser Asn Thr Lys Tyr Arg Arg Leu Cys
770 775 780

Pro Cys Arg Asp Phe Arg Lys Gly Gln Val Ala Leu Cys Gln Gly Cys
785 790 795 800

Leu

<210> 31

<211> 464

<212> PRT

<213> Gallus gallus

<400> 31

Met Arg Cys Ser Pro Lys Arg Ser Leu Thr Ala Val Ile Ala Ala Ser
1 5 10 15

Phe Leu Leu Leu Leu Leu Leu Leu Leu Leu His Arg Gly Ser Trp Gln
20 25 30

Asp Pro Gln Glu Val Gln Phe Arg Asp Leu Pro Ser Asp Ala Val Leu
35 40 45

Lys Ile Leu Lys Gln Gly Ser Leu His Ile Leu Gln Asp Thr Asp Asn
50 55 60

Leu Cys Ala Leu His Asn Ile Ser Tyr His Leu Leu Ala Gly Ser Pro
65 70 75 80

Leu Pro His Lys Lys Phe Leu Ala Val Gly Leu Ser Ser Val Arg Arg
85 90 95

Pro Arg Gly Tyr Tyr Leu Pro Asp Thr Leu Gln Ser Leu Phe Lys Gln
100 105 110

Ser Ser Glu Glu Glu Leu Gln Glu Met Val Val Val Val His Leu Ala
115 120 125

Asp Ala Asp Pro Ile Trp Asn Ala Gln Val Ala Ala Asp Ile Ser His

5

10

ES 2 659 118 T3

130		135		140	
Arg Phe Ala His His Ile Leu Leu Gly Arg Leu Val Leu Ile His Thr	145	150	155	160	
Pro His Glu Phe Tyr Pro Thr Leu Glu Gly Leu Lys Arg Asn Tyr Asn	165	170	175		
Asp Pro Glu Glu Arg Val Lys Phe Arg Ser Lys Gln Asn Val Asp Tyr	180	185	190		
Ala Phe Leu Phe Thr Phe Ala Ala Asn Leu Ser Ser Tyr Tyr Leu Met	195	200	205		
Ile Glu Asp Asp Val Trp Ser Ala Lys Ser Phe Phe Thr Ala Ile Arg	210	215	220		
Lys Ala Val Ala Ser Gln Glu Gly Ser Asn Trp Ala Thr Leu Glu Phe	225	230	235		
Ser Lys Leu Gly Tyr Ile Gly Lys Leu Tyr Arg Ser Ser Asp Leu Pro	245	250	255		
Arg Leu Ala Arg Phe Leu Leu Leu Phe Tyr Gln Glu Met Pro Cys Asp	260	265	270		
Trp Leu Leu Thr His Phe Arg Leu Leu Leu Thr Gln Lys Asp Val Ile	275	280	285		
Arg Phe Lys Pro Ser Leu Phe Gln His Met Gly Leu Tyr Ser Ser Phe	290	295	300		
Gln Gly Thr Val Asn Arg Leu Glu Asp Asp Glu Phe Gln Ala Asp Ala	305	310	315	320	
Met Asp Leu Pro Asp Asn Pro Pro Ala Ala Leu Phe Thr Asn Met Val	325	330	335		
Val Phe Glu Asn Tyr Glu Pro Ser Lys Ala Tyr Ser Thr Ala Arg Gly	340	345	350		
Tyr Phe Trp Gly Lys Asn Pro Ala Val Gly Ser Ile Phe Ser Ile Val	355	360	365		
Phe His Gln Pro Ala Arg Val Thr Arg Val Arg Val Gln Thr Gly Ser	370	375	380		
Ser Glu Arg Pro Gly Asp Phe Leu His Ala Gly Val Leu Glu Leu Gly	385	390	395	400	
Arg Gly Arg Arg Ala Asp Gly Arg Asp Cys Ser Val Tyr Thr Thr Val	405	410	415		
Gly Thr Phe Glu Lys Gly Asn Leu Glu Trp Arg Gly Leu Glu Lys Gly	420	425	430		
Met Pro Asn Pro Val Glu Cys Val Arg Ile Arg Val Thr Gln Ser Gln	435	440	445		
Ser Glu Trp Leu Ile Ile Gln Ser Ile Gly Ile Trp Thr Ala Gly Thr	450	455	460		

ES 2 659 118 T3

<210> 32
 <211> 575
 <212> PRT
 <213> Mus musculus

Á

<400> 32
 Met Arg Ala Trp Thr Gly Ser Trp Arg Trp Ile Met Leu Ile Leu Phe
 1 5 10 15
 Ala Trp Gly Thr Leu Leu Phe Tyr Ile Gly Gly His Leu Val Arg Asp
 20 25 30
 Asn Asp His Pro Asp His Ser Ser Arg Glu Leu Ser Lys Ile Leu Ala
 35 40 45
 Lys Leu Glu Arg Leu Lys Gln Gln Asn Glu Asp Leu Arg Arg Met Ala
 50 55 60
 Glu Ser Leu Arg Ile Pro Glu Gly Pro Ile Asp Gln Gly Thr Ala Thr
 65 70 75 80
 Gly Arg Val Arg Val Leu Glu Glu Gln Leu Val Lys Ala Lys Glu Gln
 85 90 95
 Ile Glu Asn Tyr Lys Lys Gln Ala Arg Asn Gly Leu Gly Lys Asp His
 100 105 110
 Glu Ile Leu Arg Arg Arg Ile Glu Asn Gly Ala Lys Glu Leu Trp Phe
 115 120 125
 Phe Leu Gln Ser Glu Leu Lys Lys Leu Lys His Leu Glu Gly Asn Glu
 130 135 140
 Leu Gln Arg His Ala Asp Glu Ile Leu Leu Asp Leu Gly His His Glu
 145 150 155 160
 Arg Ser Ile Met Thr Asp Leu Tyr Tyr Leu Ser Gln Thr Asp Gly Ala
 165 170 175
 Gly Asp Trp Arg Glu Lys Glu Ala Lys Asp Leu Thr Glu Leu Val Gln
 180 185 190

ES 2 659 118 T3

Arg Arg Ile Thr Tyr Leu Gln Asn Pro Lys Asp Cys Ser Lys Ala Arg
 195 200 205

Lys Leu Val Cys Asn Ile Asn Lys Gly Cys Gly Tyr Gly Cys Gln Leu
 210 215 220

His His Val Val Tyr Cys Phe Met Ile Ala Tyr Gly Thr Gln Arg Thr
 225 230 240

Leu Ile Leu Glu Ser Gln Asn Trp Arg Tyr Ala Thr Gly Gly Trp Glu
 245 250 255

Thr Val Phe Arg Pro Val Ser Glu Thr Cys Thr Asp Arg Ser Gly Leu
 260 265 270

Ser Thr Gly His Trp Ser Gly Glu Val Asn Asp Lys Asn Ile Gln Val
 275 280 285

Val Glu Leu Pro Ile Val Asp Ser Leu His Pro Arg Pro Pro Tyr Leu
 290 295 300

Pro Leu Ala Val Pro Glu Asp Leu Ala Asp Arg Leu Leu Arg Val His
 305 310 315 320

Gly Asp Pro Ala Val Trp Trp Val Ser Gln Phe Val Lys Tyr Leu Ile
 325 330 335

Arg Pro Gln Pro Trp Leu Glu Lys Glu Ile Glu Glu Ala Thr Lys Lys
 340 345 350

Leu Gly Phe Lys His Pro Val Ile Gly Val His Val Arg Arg Thr Asp
 355 360 365

Lys Val Gly Thr Glu Ala Ala Phe His Pro Ile Glu Glu Tyr Met Val
 370 375 380

His Val Glu Glu His Phe Gln Leu Leu Ala Arg Arg Met Gln Val Asp
 385 390 395 400

Lys Lys Arg Val Tyr Leu Ala Thr Asp Asp Pro Thr Leu Leu Lys Glu
 405 410 415

Ala Lys Thr Lys Tyr Ser Asn Tyr Glu Phe Ile Ser Asp Asn Ser Ile
 420 425 430

Ser Trp Ser Ala Gly Leu His Asn Arg Tyr Thr Glu Asn Ser Leu Arg
 435 440 445

Gly Val Ile Leu Asp Ile His Phe Leu Ser Gln Ala Asp Phe Leu Val
 450 455 460

ES 2 659 118 T3

Cys Thr Phe Ser Ser Gln Val Cys Arg Val Ala Tyr Glu Ile Met Gln
465 470 475 480

Thr Leu His Pro Asp Ala Ser Ala Asn Phe His Ser Leu Asp Asp Ile
485 490 495

Tyr Tyr Phe Gly Gly Gln Asn Ala His Asn Gln Ile Ala Val Tyr Pro
500 505 510

His Lys Pro Arg Thr Glu Glu Glu Ile Pro Met Glu Pro Gly Asp Ile
515 520 525

Ile Gly Val Ala Gly Asn His Trp Asp Gly Tyr Ser Lys Gly Ile Asn
530 535 540

Arg Lys Leu Gly Lys Thr Gly Leu Tyr Pro Ser Tyr Lys Val Arg Glu
545 550 555 560

Lys Ile Glu Thr Val Lys Tyr Pro Thr Tyr Pro Glu Ala Glu Lys
565 570 575

<210> 33

<211> 575

<212> PRT

<213> Homo sapiens

5

<400> 33

Met Arg Pro Trp Thr Gly Ser Trp Arg Trp Ile Met Leu Ile Leu Phe
1 5 10 15

Ala Trp Gly Thr Leu Leu Phe Tyr Ile Gly Gly His Leu Val Arg Asp
20 25 30

Asn Asp His Pro Asp His Ser Ser Arg Glu Leu Ser Lys Ile Leu Ala
35 40 45

Lys Leu Glu Arg Leu Lys Gln Gln Asn Glu Asp Leu Arg Arg Met Ala
50 55 60

Glu Ser Leu Arg Ile Pro Glu Gly Pro Ile Asp Gln Gly Pro Ala Ile
65 70 75 80

Gly Arg Val Arg Val Leu Glu Glu Gln Leu Val Lys Ala Lys Glu Gln
85 90 95

Ile Glu Asn Tyr Lys Lys Gln Thr Arg Asn Gly Leu Gly Lys Asp His
100 105 110

Glu Ile Leu Arg Arg Arg Ile Glu Asn Gly Ala Lys Glu Leu Trp Phe
115 120 125

10

ES 2 659 118 T3

Phe Leu Gln Ser Glu Leu Lys Lys Leu Lys Asn Leu Glu Gly Asn Glu
 130 135 140
 Leu Gln Arg His Ala Asp Glu Phe Leu Leu Asp Leu Gly His His Glu
 145 150 155 160
 Arg Ser Ile Met Thr Asp Leu Tyr Tyr Leu Ser Gln Thr Asp Gly Ala
 165 170 175
 Gly Asp Trp Arg Glu Lys Glu Ala Lys Asp Leu Thr Glu Leu Val Gln
 180 185 190
 Arg Arg Ile Thr Tyr Leu Gln Asn Pro Lys Asp Cys Ser Lys Ala Lys
 195 200 205
 Lys Leu Val Cys Asn Ile Asn Lys Gly Cys Gly Tyr Gly Cys Gln Leu
 210 215 220
 His His Val Val Tyr Cys Phe Met Ile Ala Tyr Gly Thr Gln Arg Thr
 225 230 235 240
 Leu Ile Leu Glu Ser Gln Asn Trp Arg Tyr Ala Thr Gly Gly Trp Glu
 245 250 255
 Thr Val Phe Arg Pro Val Ser Glu Thr Cys Thr Asp Arg Ser Gly Ile
 260 265 270
 Ser Thr Gly His Trp Ser Gly Glu Val Lys Asp Lys Asn Val Gln Val
 275 280 285
 Val Glu Leu Pro Ile Val Asp Ser Leu His Pro Arg Pro Pro Tyr Leu
 290 295 300
 Pro Leu Ala Val Pro Glu Asp Leu Ala Asp Arg Leu Val Arg Val His
 305 310 315 320
 Gly Asp Pro Ala Val Trp Trp Val Ser Gln Phe Val Lys Tyr Leu Ile
 325 330 335
 Arg Pro Gln Pro Trp Leu Glu Lys Glu Ile Glu Glu Ala Thr Lys Lys
 340 345 350
 Leu Gly Phe Lys His Pro Val Ile Gly Val His Val Arg Arg Thr Asp
 355 360 365
 Lys Val Gly Thr Glu Ala Ala Phe His Pro Ile Glu Glu Tyr Met Val
 370 375 380
 His Val Glu Glu His Phe Gln Leu Leu Ala Arg Arg Met Gln Val Asp
 385 390 395 400

ES 2 659 118 T3

Lys Lys Arg Val Tyr Leu Ala Thr Asp Asp Pro Ser Leu Leu Lys Glu
 405 410 415
 Ala Lys Thr Lys Tyr Pro Asn Tyr Glu Phe Ile Ser Asp Asn Ser Ile
 420 425 430
 Ser Trp Ser Ala Gly Leu His Asn Arg Tyr Thr Glu Asn Ser Leu Arg
 435 440 445
 Gly Val Ile Leu Asp Ile His Phe Leu Ser Gln Ala Asp Phe Leu Val
 450 455 460
 Cys Thr Phe Ser Ser Gln Val Cys Arg Val Ala Tyr Glu Ile Met Gln
 465 470 475 480
 Thr Leu His Pro Asp Ala Ser Ala Asn Phe His Ser Leu Asp Asp Ile
 485 490 495
 Tyr Tyr Phe Gly Gly Gln Asn Ala His Asn Gln Ile Ala Ile Tyr Ala
 500 505 510
 His Gln Pro Arg Thr Ala Asp Glu Ile Pro Met Glu Pro Gly Asp Ile
 515 520 525
 Ile Gly Val Ala Gly Asn His Trp Asp Gly Tyr Ser Lys Gly Val Asn
 530 535 540
 Arg Lys Leu Gly Arg Thr Gly Leu Tyr Pro Ser Tyr Lys Val Arg Glu
 545 550 555 560
 Lys Ile Glu Thr Val Lys Tyr Pro Thr Tyr Pro Glu Ala Glu Lys
 565 570 575

<210> 34
 <211> 398
 <212> PRT
 <213> Homo sapiens

5

<400> 34
 Met Arg Leu Arg Glu Pro Leu Leu Ser Gly Ser Ala Ala Met Pro Gly
 1 5 10 15
 Ala Ser Leu Gln Arg Ala Cys Arg Leu Leu Val Ala Val Cys Ala Leu
 20 25 30
 His Leu Gly Val Thr Leu Val Tyr Tyr Leu Ala Gly Arg Asp Leu Ser
 35 40 45
 Arg Leu Pro Gln Leu Val Gly Val Ser Thr Pro Leu Gln Gly Gly Ser
 50 55 60

10

ES 2 659 118 T3

Asn Ser Ala Ala Ala Ile Gly Gln Ser Ser Gly Glu Leu Arg Thr Gly
 65 70 75 80
 Gly Ala Arg Pro Pro Pro Pro Leu Gly Ala Ser Ser Gln Pro Arg Pro
 85 90 95
 Gly Gly Asp Ser Ser Pro Val Val Asp Ser Gly Pro Gly Pro Ala Ser
 100 105 110
 Asn Leu Thr Ser Val Pro Val Pro His Thr Thr Ala Leu Ser Leu Pro
 115 120 125
 Ala Cys Pro Glu Glu Ser Pro Leu Leu Val Gly Pro Met Leu Ile Glu
 130 135 140
 Phe Asn Met Pro Val Asp Leu Glu Leu Val Ala Lys Gln Asn Pro Asn
 145 150 155 160
 Val Lys Met Gly Gly Arg Tyr Ala Pro Arg Asp Cys Val Ser Pro His
 165 170 175
 Lys Val Ala Ile Ile Ile Pro Phe Arg Asn Arg Gln Glu His Leu Lys
 180 185 190
 Tyr Trp Leu Tyr Tyr Leu His Pro Val Leu Gln Arg Gln Gln Leu Asp
 195 200 205
 Tyr Gly Ile Tyr Val Ile Asn Gln Ala Gly Asp Thr Ile Phe Asn Arg
 210 215 220
 Ala Lys Leu Leu Asn Val Gly Phe Gln Glu Ala Leu Lys Asp Tyr Asp
 225 230 235 240
 Tyr Thr Cys Phe Val Phe Ser Asp Val Asp Leu Ile Pro Met Asn Asp
 245 250 255
 His Asn Ala Tyr Arg Cys Phe Ser Gln Pro Arg His Ile Ser Val Ala
 260 265 270
 Met Asp Lys Phe Gly Phe Ser Leu Pro Tyr Val Gln Tyr Phe Gly Gly
 275 280 285
 Val Ser Ala Leu Ser Lys Gln Gln Phe Leu Thr Ile Asn Gly Phe Pro
 290 295 300
 Asn Asn Tyr Trp Gly Trp Gly Gly Glu Asp Asp Asp Ile Phe Asn Arg
 305 310 315 320
 Leu Val Phe Arg Gly Met Ser Ile Ser Arg Pro Asn Ala Val Val Gly
 325 330 335
 Arg Cys Arg Met Ile Arg His Ser Arg Asp Lys Lys Asn Glu Pro Asn
 340 345 350
 Pro Gln Arg Phe Asp Arg Ile Ala His Thr Lys Glu Thr Met Leu Ser
 355 360 365
 Asp Gly Leu Asn Ser Leu Thr Tyr Gln Val Leu Asp Val Gln Arg Tyr
 370 375 380
 Pro Leu Tyr Thr Gln Ile Thr Val Asp Ile Gly Thr Pro Ser
 385 390 395

ES 2 659 118 T3

<210> 35
 <211> 346
 <212> PRT
 <213> Mus musculus

Á

<400> 35
 Met Arg Phe Arg Glu Gln Phe Leu Gly Gly Ser Ala Ala Met Pro Gly
 1 5 10 15
 Ala Thr Leu Gln Arg Ala Cys Arg Leu Leu Val Ala Val Cys Ala Leu
 20 25 30
 His Leu Gly Val Thr Leu Val Tyr Tyr Leu Ser Gly Arg Asp Leu Ser
 35 40 45
 Arg Leu Pro Gln Leu Val Gly Val Ser Ser Thr Leu Gln Gly Gly Thr
 50 55 60
 Asn Gly Ala Ala Ala Ser Lys Gln Pro Pro Gly Glu Gln Arg Pro Arg
 65 70 75 80
 Gly Ala Arg Pro Pro Pro Pro Leu Gly Val Ser Pro Lys Pro Arg Pro
 85 90 95
 Gly Leu Asp Ser Ser Pro Gly Ala Ala Ser Gly Pro Gly Leu Lys Ser
 100 105 110
 Asn Leu Ser Ser Leu Pro Val Pro Thr Thr Thr Gly Leu Leu Ser Leu
 115 120 125
 Pro Ala Cys Pro Glu Glu Ser Pro Leu Leu Val Gly Pro Met Leu Ile
 130 135 140
 Asp Phe Asn Ile Ala Val Asp Leu Glu Leu Leu Ala Lys Lys Asn Pro
 145 150 155 160
 Glu Ile Lys Thr Gly Gly Arg Tyr Ser Pro Lys Asp Cys Val Ser Pro
 165 170 175
 His Lys Val Ala Ile Ile Ile Pro Phe Arg Asn Arg Gln Glu His Leu

ES 2 659 118 T3

180 185 190

Lys Tyr Trp₁₉₅ Leu Tyr Tyr Leu His₂₀₀ Pro Ile Leu Gln Arg₂₀₅ Gln Gln Leu

Asp Tyr Gly₂₁₀ Ile Tyr Val Ile₂₁₅ Asn Gln Ala Gly₂₂₀ Asp Thr Met Phe Asn

Arg₂₂₅ Ala Lys Leu Leu Asn₂₃₀ Ile Gly Phe Gln Glu₂₃₅ Ala Leu Lys Asp Tyr₂₄₀

Asp Tyr Asn Cys Phe₂₄₅ Val Phe Ser Asp₂₅₀ Val Asp Leu Ile Pro Met₂₅₅ Asp

Asp Arg Asn Ala₂₆₀ Tyr Arg Cys Phe₂₆₅ Ser Gln Pro Arg His Ile₂₇₀ Ser Val

Ala Met Asp₂₇₅ Lys Phe Gly Phe₂₈₀ Ser Leu Pro Tyr Val Gln Tyr Phe Gly

Gly Val₂₉₀ Ser Ala Leu Ser Lys₂₉₅ Gln Gln Phe Leu Ala₃₀₀ Ile Asn Gly Phe

Pro Asn Asn Tyr Trp Gly₃₁₀ Trp Gly Gly Glu₃₁₅ Asp Asp Ile Phe Asn₃₂₀

Arg Ser Lys Pro Lys₃₂₅ Ala Ser Ala Glu₃₃₀ Thr Gly Gly Ser Leu Gly₃₃₅

Lys Ala Leu Ser₃₄₀ Pro Ala Ser Thr Arg₃₄₅ Ala

<210> 36
 <211> 438
 <212> PRT
 <213> Mus musculus

5

<400> 36

Met Lys Pro His₅ Leu Lys Gln Trp Arg₁₀ Gln Arg Met Leu Phe Gly₁₅ Ile

Phe Val Trp Gly₂₀ Leu Leu Phe Leu Ala₂₅ Ile Phe Ile Tyr Phe₃₀ Thr Asn

Ser Asn Pro₃₅ Ala Ala Pro Met₄₀ Pro Ser Ser Phe Ser₄₅ Phe Leu Glu Arg

Arg Gly₅₀ Leu Leu Pro Leu Gln₅₅ Gly Lys Gln Arg Val₆₀ Ile Met Gly Ala

Leu Gln Glu Pro Ser₇₀ Leu Pro Arg Ser Leu Asp₇₅ Ala Ser Lys Val Leu₈₀

10

ES 2 659 118 T3

Leu Asp Ser His Pro Gln Asn Pro Phe His Pro Trp Pro Gly Asp Pro
 85 90 95
 Gln Lys Trp Asp Gln Ala Pro Asn Gly Phe Asp Asn Gly Asp Glu Phe
 100 105 110
 Phe Thr Ser Gln Val Gly Arg Lys Ser Gln Ser Ala Phe Tyr Pro Glu
 115 120 125
 Glu Asp Ser Tyr Phe Phe Val Ala Asp Gln Pro Glu Leu Tyr His His
 130 135 140
 Arg Gln Gly Ala Leu Glu Leu Pro Ser Pro Gly Glu Thr Ser Trp Arg
 145 150 155 160
 Ser Gly Pro Val Gln Pro Lys Gln Lys Leu Leu His Pro Arg Arg Gly
 165 170 175
 Ser Leu Pro Glu Glu Ala Tyr Asp Ser Asp Met Leu Ser Ala Ser Met
 180 185 190
 Ser Arg Ala Phe Leu Tyr Arg Leu Trp Lys Gly Ala Val Ser Ser Lys
 195 200 205
 Met Leu Asn Pro Arg Leu Gln Lys Ala Met Arg Tyr Tyr Met Ser Phe
 210 215 220
 Asn Lys His Gly Val Arg Phe Arg Arg Arg Gly Arg Arg Glu Ala Thr
 225 230 235 240
 Arg Thr Gly Pro Glu Leu Leu Cys Glu Met Arg Arg Arg Val Arg Val
 245 250 255
 Arg Thr Leu Asp Gly Arg Glu Ala Pro Phe Ser Gly Leu Gly Trp Arg
 260 265 270
 Pro Leu Val Pro Gly Val Pro Leu Ser Gln Leu His Pro Arg Gly Leu
 275 280 285
 Ser Ser Cys Ala Val Val Met Ser Ala Gly Ala Ile Leu Asn Ser Ser
 290 295 300
 Leu Gly Glu Glu Ile Asp Ser His Asp Ala Val Leu Arg Phe Asn Ser
 305 310 315 320
 Ala Pro Thr Arg Gly Tyr Glu Lys Asp Val Gly Asn Lys Thr Thr Val
 325 330 335
 Arg Ile Ile Asn Ser Gln Ile Leu Ala Asn Pro Ser His His Phe Ile
 340 345 350

ES 2 659 118 T3

Asp Ser Ala Leu Tyr Lys Asp Val Ile Leu Val Ala Trp Asp Pro Ala
 355 360 365

Pro Tyr Ser Ala Asn Leu Asn Leu Trp Tyr Lys Lys Pro Asp Tyr Asn
 370 375 380

Leu Phe Thr Pro Tyr Ile Gln His Arg Arg Lys Tyr Pro Thr Gln Pro
 385 390 395 400

Phe Tyr Ile Leu His Pro Lys Phe Ile Trp Gln Leu Trp Asp Ile Ile
 405 410 415

Gln Glu Asn Thr Arg Glu Lys Ile Gln Pro Asn Pro Pro Ser Ser Gly
 420 425 430

Phe Ile Gly Thr Cys Val
 435

<210> 37

<211> 406

5 <212> PRT

<213> Homo sapiens

<400> 37

Met Ile His Thr Asn Leu Lys Lys Lys Phe Ser Cys Cys Val Leu Val
 1 5 10 15

Phe Leu Leu Phe Ala Val Ile Cys Val Trp Lys Glu Lys Lys Lys Gly
 20 25 30

Ser Tyr Tyr Asp Ser Phe Lys Leu Gln Thr Lys Glu Phe Gln Val Leu
 35 40 45

Lys Ser Leu Gly Lys Leu Ala Met Gly Ser Asp Ser Gln Ser Val Ser
 50 55 60

Ser Ser Ser Thr Gln Asp Pro His Arg Gly Arg Gln Thr Leu Gly Ser
 65 70 75 80

Leu Arg Gly Leu Ala Lys Ala Lys Pro Glu Ala Ser Phe Gln Val Trp
 85 90 95

Asn Lys Asp Ser Ser Ser Lys Asn Leu Ile Pro Arg Leu Gln Lys Ile
 100 105 110

Trp Lys Asn Tyr Leu Ser Met Asn Lys Tyr Lys Val Ser Tyr Lys Gly
 115 120 125

Pro Gly Pro Gly Ile Lys Phe Ser Ala Glu Ala Leu Arg Cys His Leu
 130 135 140

10

ES 2 659 118 T3

Arg Asp His Val Asn Val Ser Met Val Glu Val Thr Asp Phe Pro Phe
 145 150 155 160
 Asn Thr Ser Glu Trp Glu Gly Tyr Leu Pro Lys Glu Ser Ile Arg Thr
 165 170 175
 Lys Ala Gly Pro Trp Gly Arg Cys Ala Val Val Ser Ser Ala Gly Ser
 180 185 190
 Leu Lys Ser Ser Gln Leu Gly Arg Glu Ile Asp Asp His Asp Ala Val
 195 200 205
 Leu Arg Phe Asn Gly Ala Pro Thr Ala Asn Phe Gln Gln Asp Val Gly
 210 215 220
 Thr Lys Thr Thr Ile Arg Leu Met Asn Ser Gln Leu Val Thr Thr Glu
 225 230 235 240
 Lys Arg Phe Leu Lys Asp Ser Leu Tyr Asn Glu Gly Ile Leu Ile Val
 245 250 255
 Trp Asp Pro Ser Val Tyr His Ser Asp Ile Pro Lys Trp Tyr Gln Asn
 260 265 270
 Pro Asp Tyr Asn Phe Phe Asn Asn Tyr Lys Thr Tyr Arg Lys Leu His
 275 280 285
 Pro Asn Gln Pro Phe Tyr Ile Leu Lys Pro Gln Met Pro Trp Glu Leu
 290 295 300
 Trp Asp Ile Leu Gln Glu Ile Ser Pro Glu Glu Ile Gln Pro Asn Pro
 305 310 315 320
 Pro Ser Ser Gly Met Leu Gly Ile Ile Ile Met Met Thr Leu Cys Asp
 325 330 335
 Gln Val Asp Ile Tyr Glu Phe Leu Pro Ser Lys Arg Lys Thr Asp Val
 340 345 350
 Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp Ser Ala Cys Thr Met Gly Ala
 355 360 365
 Tyr His Pro Leu Leu Tyr Glu Lys Asn Leu Val Lys His Leu Asn Gln
 370 375 380
 Gly Thr Asp Glu Asp Ile Tyr Leu Leu Gly Lys Ala Thr Leu Pro Gly
 385 390 395 400
 Phe Arg Thr Ile His Cys
 405

<210> 38

<211> 175

Á <212> PRT

<213> Homo sapiens

<400> 38

ES 2 659 118 T3

Met Asn Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser
 1 5 10 15
 Leu Tyr Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val Tyr His
 20 25 30
 Ser Asp Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn
 35 40 45
 Asn Tyr Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe Tyr Ile
 50 55 60
 Leu Lys Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile
 65 70 75 80
 Ser Pro Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met Leu Gly
 85 90 95
 Ile Ile Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe
 100 105 110
 Leu Pro Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe
 115 120 125
 Phe Asp Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu Tyr Glu
 130 135 140
 Lys Asn Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr
 145 150 155 160
 Leu Leu Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys
 165 170 175

<210> 39
 <211> 337
 <212> PRT
 <213> Mus musculus

5

<400> 39
 Met Arg Arg Lys Thr Leu Lys Tyr Leu Thr Phe Phe Leu Leu Phe Ile
 1 5 10 15

Phe Leu Thr Ser Phe Val Leu Asn Tyr Ser Asn Thr Gly Val Pro Ser
 20 25 30

Ala Trp Phe Pro Lys Gln Met Leu Leu Glu Leu Ser Glu Asn Phe Arg

10

ES 2 659 118 T3

	35		40		45																
Arg	Phe	Ile	Lys	Ser	Gln	Pro	Cys	Thr	Cys	Arg	His	Cys	Ile	Ser	Gln						
	50					55					60										
Asp	Lys	Val	Ser	Tyr	Trp	Phe	Asp	Gln	Arg	Phe	Asn	Lys	Thr	Met	Gln						
	65				70					75					80						
Pro	Leu	Leu	Thr	Val	His	Asn	Ala	Leu	Met	Glu	Glu	Asp	Thr	Tyr	Arg						
				85					90					95							
Trp	Trp	Leu	Arg	Leu	Gln	Arg	Glu	Arg	Lys	Pro	Asn	Asn	Leu	Ser	Asp						
			100					105					110								
Thr	Val	Lys	Glu	Leu	Phe	Arg	Leu	Val	Pro	Gly	Asn	Val	Asp	Pro	Met						
		115					120					125									
Leu	Asn	Lys	Arg	Leu	Val	Gly	Cys	Arg	Arg	Cys	Ala	Val	Val	Gly	Asn						
	130					135					140										
Ser	Gly	Asn	Leu	Lys	Asp	Ser	Ser	Tyr	Gly	Pro	Glu	Ile	Asp	Ser	His						
	145				150					155					160						
Asp	Phe	Val	Leu	Arg	Met	Asn	Lys	Ala	Pro	Thr	Val	Gly	Phe	Glu	Ala						
				165					170					175							
Asp	Val	Gly	Ser	Arg	Thr	Thr	His	His	Leu	Val	Tyr	Pro	Glu	Ser	Phe						
			180					185					190								
Arg	Glu	Leu	Gly	Glu	Asn	Val	Asn	Met	Val	Leu	Val	Pro	Phe	Lys	Thr						
		195					200					205									
Thr	Asp	Leu	Gln	Trp	Val	Ile	Ser	Ala	Thr	Thr	Thr	Gly	Thr	Ile	Thr						
	210					215						220									
His	Thr	Tyr	Val	Pro	Val	Pro	Pro	Lys	Ile	Lys	Val	Lys	Gln	Glu	Lys						
	225				230					235					240						
Ile	Leu	Ile	Tyr	His	Pro	Ala	Phe	Ile	Lys	Tyr	Val	Phe	Asp	Asn	Trp						
				245					250					255							
Leu	Gln	Gly	His	Gly	Arg	Tyr	Pro	Ser	Thr	Gly	Ile	Leu	Ser	Ile	Ile						
			260					265					270								
Phe	Ser	Ile	His	Ile	Cys	Asp	Glu	Val	Asp	Leu	Tyr	Gly	Phe	Gly	Ala						
		275					280					285									
Asp	Ser	Lys	Gly	Asn	Trp	His	His	Tyr	Trp	Glu	Asn	Asn	Pro	Ser	Ala						
	290					295					300										
Gly	Ala	Phe	Arg	Lys	Thr	Gly	Val	His	Asp	Gly	Asp	Phe	Glu	Tyr	Asn						
	305						310					315			320						
Ile	Thr	Thr	Thr	Leu	Ala	Ala	Ile	Asn	Lys	Ile	Arg	Ile	Phe	Lys	Gly						
				325					330					335							

Arg

Á <210> 40
<211> 340

ES 2 659 118 T3

<212> PRT
 <213> Homo sapiens

<400> 40

Met Val Thr Leu Arg Lys Arg Thr Leu Lys Val Leu Thr Phe Leu Val
 1 5 10 15
 Leu Phe Ile Phe Leu Thr Ser Phe Phe Leu Asn Tyr Ser His Thr Met
 20 25
 Val Ala Thr Thr Trp Phe Pro Lys Gln Met Val Leu Glu Leu Ser Glu
 35 40 45
 Asn Leu Lys Arg Leu Ile Lys His Arg Pro Cys Thr Cys Thr His Cys
 50 55 60
 Ile Gly Gln Arg Lys Leu Ser Ala Trp Phe Asp Glu Arg Phe Asn Gln
 65 70 75 80
 Thr Met Gln Pro Leu Leu Thr Ala Gln Asn Ala Leu Leu Glu Asp Asp
 85 90 95
 Thr Tyr Arg Trp Trp Leu Arg Leu Gln Arg Glu Lys Lys Pro Asn Asn
 100 105 110
 Leu Asn Asp Thr Ile Lys Glu Leu Phe Arg Val Val Pro Gly Asn Val
 115 120 125
 Asp Pro Met Leu Glu Lys Arg Ser Val Gly Cys Arg Arg Cys Ala Val
 130 135 140
 Val Gly Asn Ser Gly Asn Leu Arg Glu Ser Ser Tyr Gly Pro Glu Ile
 145 150 155 160
 Asp Ser His Asp Phe Val Leu Arg Met Asn Lys Ala Pro Thr Ala Gly
 165 170 175
 Phe Glu Ala Asp Val Gly Thr Lys Thr Thr His His Leu Val Tyr Pro
 180 185 190
 Glu Ser Phe Arg Glu Leu Gly Asp Asn Val Ser Met Ile Leu Val Pro
 195 200 205

Á

ES 2 659 118 T3

Phe Lys Thr Ile Asp Leu Glu Trp Val Val Ser Ala Ile Thr Thr Gly
 210 215 220

Thr Ile Ser His Thr Tyr Ile Pro Val Pro Ala Lys Ile Arg Val Lys
 225 230 235 240

Gln Asp Lys Ile Leu Ile Tyr His Pro Ala Phe Ile Lys Tyr Val Phe
 245 250 255

Asp Asn Trp Leu Gln Gly His Gly Arg Tyr Pro Ser Thr Gly Ile Leu
 260 265 270

Ser Val Ile Phe Ser Met His Val Cys Asp Glu Val Asp Leu Tyr Gly
 275 280 285

Phe Gly Ala Asp Ser Lys Gly Asn Trp His His Tyr Trp Glu Asn Asn
 290 295 300

Pro Ser Ala Gly Ala Phe Arg Lys Thr Gly Val His Asp Ala Asp Phe
 305 310 315 320

Glu Ser Asn Val Thr Ala Thr Leu Ala Ser Ile Asn Lys Ile Arg Ile
 325 330 335

Phe Lys Gly Arg
 340

<210> 41

<211> 350

<212> PRT

<213> Homo sapiens

5

<400> 41

Met Lys Cys Ser Leu Arg Val Trp Phe Leu Ser Val Ala Phe Leu Leu
 1 5 10 15

Val Phe Ile Met Ser Leu Leu Phe Thr Tyr Ser His His Ser Met Ala
 20 25 30

Thr Leu Pro Tyr Leu Asp Ser Gly Ala Leu Asp Gly Thr His Arg Val
 35 40 45

Lys Leu Val Pro Gly Tyr Ala Gly Leu Gln Arg Leu Ser Lys Glu Arg
 50 55 60

Leu Ser Gly Lys Ser Cys Ala Cys Arg Arg Cys Met Gly Asp Ala Gly
 65 70 75 80

Ala Ser Asp Trp Phe Asp Ser His Phe Asp Gly Asn Ile Ser Pro Val
 85 90 95

10

ES 2 659 118 T3

Trp Thr Arg Glu Asn Met Asp Leu Pro Pro Asp Val Gln Arg Trp Trp
100 105 110

Met Met Leu Gln Pro Gln Phe Lys Ser His Asn Thr Asn Glu Val Leu
115 120

Glu Lys Leu Phe Gln Ile Val Pro Gly Glu Asn Pro Tyr Arg Phe Arg
130 135 140

Asp Pro His Gln Cys Arg Arg Cys Ala Val Val Gly Asn Ser Gly Asn
145 150 155 160

Leu Arg Gly Ser Gly Tyr Gly Gln Asp Val Asp Gly His Asn Phe Ile
165 170 175

Met Arg Met Asn Gln Ala Pro Thr Val Gly Phe Glu Gln Asp Val Gly
180 185 190

Ser Arg Thr Thr His His Phe Met Tyr Pro Glu Ser Ala Lys Asn Leu
195 200 205

Pro Ala Asn Val Ser Phe Val Leu Val Pro Phe Lys Val Leu Asp Leu
210 215 220

Leu Trp Ile Ala Ser Ala Leu Ser Thr Gly Gln Ile Arg Phe Thr Tyr
225 230 235 240

Ala Pro Val Lys Ser Phe Leu Arg Val Asp Lys Glu Lys Val Gln Ile
245 250 255

Tyr Asn Pro Ala Phe Phe Lys Tyr Ile His Asp Arg Trp Thr Glu His
260 265 270

His Gly Arg Tyr Pro Ser Thr Gly Met Leu Val Leu Phe Phe Ala Leu
275 280 285

His Val Cys Asp Glu Val Asn Val Tyr Gly Phe Gly Ala Asp Ser Arg
290 295 300

Gly Asn Trp His His Tyr Trp Glu Asn Asn Arg Tyr Ala Gly Glu Phe
305 310 315 320

Arg Lys Thr Gly Val His Asp Ala Asp Phe Glu Ala His Ile Ile Asp
325 330 335

Met Leu Ala Lys Ala Ser Lys Ile Glu Val Tyr Arg Gly Asn
340 345 350

<210> 42
<211> 444
Á <212> PRT
<213> Homo sapiens
<400> 42

ES 2 659 118 T3

Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
1 5 10 15

Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
20 25 30

Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His Ser Ser Ser Pro
35 40 45

Gln Glu Lys Pro Val Ala Asp Ser Val Val Leu Ser Phe Asp Ser Ala
50 55 60

Gly Gln Thr Leu Gly Ser Glu Tyr Asp Arg Leu Gly Phe Leu Leu Asn
65 70 75 80

Leu Asp Ser Lys Leu Ser Pro Arg Thr Leu Cys Thr Val Val Phe Gly
85 90 95

Leu Asp Cys Ile Leu Glu Ser Pro Gly Glu Pro Lys Lys Leu Leu Met
100 105 110

Pro Ala Ser His Pro Leu Glu Ile Leu Lys Ser Leu Ser Glu Asp Thr
115 120 125

Ala Phe Ala Leu Gly Phe Leu Lys Leu Pro Arg Pro Ala Glu Leu Ala
130 135 140

Thr Lys Tyr Ala Asn Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala
145 150 155 160

Ser Ala Leu Met Thr Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro
165 170 175

Met Phe Leu Asp Asp Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe
180 185 190

Val Pro Pro Phe Gly Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala Ile
195 200 205

Leu Ser Val Thr Lys Glu Tyr Arg Leu Thr Pro Ala Leu Asp Ser Leu
210 215 220

Arg Cys Arg Arg Cys Ile Ile Val Gly Asn Gly Gly Val Leu Ala Asn
225 230 235 240

Lys Ser Leu Gly Ser Arg Ile Asp Asp Tyr Asp Ile Val Val Arg Leu
245 250 255

ES 2 659 118 T3

Asn Ser Ala Pro Val Lys Gly Phe Glu Lys Asp Val Gly Ser Lys Thr
 260 265 270
 Thr Leu Arg Ile Thr Tyr Pro Glu Gly Ala Met Gln Arg Pro Glu Gln
 275 280 285
 Tyr Glu Arg Asp Ser Leu Phe Val Leu Ala Gly Phe Lys Trp Gln Asp
 290 295 300
 Phe Lys Trp Leu Lys Tyr Ile Val Tyr Lys Glu Arg Val Ser Ala Ser
 305 310 315
 Asp Gly Phe Trp Lys Ser Val Ala Thr Arg Val Pro Lys Glu Pro Pro
 325 330 335
 Glu Ile Arg Ile Leu Asn Pro Tyr Phe Ile Gln Glu Ala Ala Phe Thr
 340 345 350
 Leu Ile Gly Leu Pro Phe Asn Asn Gly Leu Met Gly Arg Gly Asn Ile
 355 360 365
 Pro Thr Leu Gly Ser Val Ala Val Thr Met Ala Leu His Gly Cys Asp
 370 375 380
 Glu Val Ala Val Ala Gly Phe Gly Tyr Asp Met Ser Thr Pro Asn Ala
 385 390 395 400
 Pro Leu His Tyr Tyr Glu Thr Val Arg Met Ala Ala Ile Lys Glu Ser
 405 410 415
 Trp Thr His Asn Ile Gln Arg Glu Lys Glu Phe Leu Arg Lys Leu Val
 420 425 430
 Lys Ala Arg Val Ile Thr Asp Leu Ser Ser Gly Ile
 435 440

<210> 43
 <211> 390
 <212> PRT
 <213> Homo sapiens

5

<400> 43
 Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15
 Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30
 Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His Ser Ser Ser Pro
 35 40 45
 Gln Glu Lys Pro Val Ala Asp Ser Val Val Leu Ser Phe Asp Ser Ala

10

ES 2 659 118 T3

325 330 335

Phe Gly Tyr Asp Met Ser Thr Pro Asn Ala Pro Leu His Tyr Tyr Glu
340 345 350

Thr Val Arg Met Ala Ala Ile Lys Glu Ser Trp Thr His Asn Ile Gln
355 360 365

Arg Glu Lys Glu Phe Leu Arg Lys Leu Val Lys Ala Arg Val Ile Thr
370 375 380

Asp Leu Ser Ser Gly Ile
385 390

<210> 44

<211> 170

5 <212> PRT

<213> Homo sapiens

<400> 44

Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
1 5 10 15

Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
20 25 30

Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His Ser Ser Ser Pro
35 40 45

Gln Glu Lys Pro Val Ala Asp Ser Val Val Leu Ser Phe Asp Ser Ala
50 55 60

Gly Gln Thr Leu Gly Ser Glu Tyr Asp Arg Leu Gly Phe Leu Leu Asn
65 70 75 80

Leu Asp Ser Lys Leu Pro Ala Glu Leu Ala Thr Lys Tyr Ala Asn Phe
85 90 95

Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala Ser Ala Leu Met Thr Ala
100 105 110

Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe Leu Asp Asp Ser
115 120 125

Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe Gly Ile
130 135 140

Lys Gly Gln Val Leu Asp Ala Gln Tyr Pro Ala Arg Glu Arg Val Ser
145 150 155 160

Ala Glu Ala Gly Glu Ser Ser Arg His His
165 170

10

<210> 45

<211> 429

<212> PRT

<213> Homo sapiens

15

<400> 45

ES 2 659 118 T3

Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
1 5 10 15

Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
20 25 30

Leu Gln Trp Glu Glu Asp Ser Asn Ser Val Val Leu Ser Phe Asp Ser
35 40 45

Ala Gly Gln Thr Leu Gly Ser Glu Tyr Asp Arg Leu Gly Phe Leu Leu
50 55 60

Asn Leu Asp Ser Lys Leu Ser Pro Arg Thr Leu Cys Thr Val Val Phe
65 70 75 80

Gly Leu Asp Cys Ile Leu Glu Ser Pro Gly Glu Pro Lys Lys Leu Leu
85 90 95

Met Pro Ala Ser His Pro Leu Glu Ile Leu Lys Ser Leu Ser Glu Asp
100 105 110

Thr Ala Phe Ala Leu Gly Phe Leu Lys Leu Pro Arg Pro Ala Glu Leu
115 120 125

Ala Thr Lys Tyr Ala Asn Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr
130 135 140

Ala Ser Ala Leu Met Thr Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala
145 150 155 160

Pro Met Phe Leu Asp Asp Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu
165 170 175

Phe Val Pro Pro Phe Gly Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala
180 185 190

Ile Leu Ser Val Thr Lys Glu Tyr Arg Leu Thr Pro Ala Leu Asp Ser
195 200 205

Leu Arg Cys Arg Arg Cys Ile Ile Val Gly Asn Gly Gly Val Leu Ala
210 215 220

Asn Lys Ser Leu Gly Ser Arg Ile Asp Asp Tyr Asp Ile Val Val Arg
225 230 235 240

ES 2 659 118 T3

Leu Asn Ser Ala Pro Val Lys Gly Phe Glu Lys Asp Val Gly Ser Lys
 245 250 255
 Thr Thr Leu Arg Ile Thr Tyr Pro Glu Gly Ala Met Gln Arg Pro Glu
 260 265 270
 Gln Tyr Glu Arg Asp Ser Leu Phe Val Leu Ala Gly Phe Lys Trp Gln
 275 280 285
 Asp Phe Lys Trp Leu Lys Tyr Ile Val Tyr Lys Glu Arg Val Ser Ala
 290 295 300
 Ser Asp Gly Phe Trp Lys Ser Val Ala Thr Arg Val Pro Lys Glu Pro
 305 310 315 320
 Pro Glu Ile Arg Ile Leu Asn Pro Tyr Phe Ile Gln Glu Ala Ala Phe
 325 330 335
 Thr Leu Ile Gly Leu Pro Phe Asn Asn Gly Leu Met Gly Arg Gly Asn
 340 345 350
 Ile Pro Thr Leu Gly Ser Val Ala Val Thr Met Ala Leu His Gly Cys
 355 360 365
 Asp Glu Val Ala Val Ala Gly Phe Gly Tyr Asp Met Ser Thr Pro Asn
 370 375 380
 Ala Pro Leu His Tyr Tyr Glu Thr Val Arg Met Ala Ala Ile Lys Glu
 385 390 395 400
 Ser Trp Thr His Asn Ile Gln Arg Glu Lys Glu Phe Leu Arg Lys Leu
 405 410 415
 Val Lys Ala Arg Val Ile Thr Asp Leu Ser Ser Gly Ile
 420 425

<210> 46
 <211> 375
 <212> PRT
 <213> Homo sapiens

5

<400> 46
 Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15
 Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30
 Leu Gln Trp Glu Glu Asp Ser Asn Ser Val Val Leu Ser Phe Asp Ser
 35 40 45

10

ES 2 659 118 T3

Ala Gly Gln Thr Leu Gly Ser Glu Tyr Asp Arg Leu Gly Phe Leu Leu
50 55 60

Asn Leu Asp Ser Lys Leu Pro Ala Glu Leu Ala Thr Lys Tyr Ala Asn
65 70 75 80

Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala Ser Ala Leu Met Thr
85 90 95

Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe Leu Asp Asp
100 105 110

Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe Gly
115 120 125

Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala Ile Leu Ser Val Thr Lys
130 135 140

Glu Tyr Arg Leu Thr Pro Ala Leu Asp Ser Leu Arg Cys Arg Arg Cys
145 150 155 160

Ile Ile Val Gly Asn Gly Gly Val Leu Ala Asn Lys Ser Leu Gly Ser
165 170 175

Arg Ile Asp Asp Tyr Asp Ile Val Val Arg Leu Asn Ser Ala Pro Val
180 185 190

Lys Gly Phe Glu Lys Asp Val Gly Ser Lys Thr Thr Leu Arg Ile Thr
195 200 205

Tyr Pro Glu Gly Ala Met Gln Arg Pro Glu Gln Tyr Glu Arg Asp Ser
210 215 220

Leu Phe Val Leu Ala Gly Phe Lys Trp Gln Asp Phe Lys Trp Leu Lys
225 230 235 240

Tyr Ile Val Tyr Lys Glu Arg Val Ser Ala Ser Asp Gly Phe Trp Lys
245 250 255

Ser Val Ala Thr Arg Val Pro Lys Glu Pro Pro Glu Ile Arg Ile Leu
260 265 270

Asn Pro Tyr Phe Ile Gln Glu Ala Ala Phe Thr Leu Ile Gly Leu Pro
275 280 285

Phe Asn Asn Gly Leu Met Gly Arg Gly Asn Ile Pro Thr Leu Gly Ser
290 295 300

Val Ala Val Thr Met Ala Leu His Gly Cys Asp Glu Val Ala Val Ala
305 310 315 320

Gly Phe Gly Tyr Asp Met Ser Thr Pro Asn Ala Pro Leu His Tyr Tyr
325 330 335

Glu Thr Val Arg Met Ala Ala Ile Lys Glu Ser Trp Thr His Asn Ile
340 345 350

Gln Arg Glu Lys Glu Phe Leu Arg Lys Leu Val Lys Ala Arg Val Ile
355 360 365

Thr Asp Leu Ser Ser Gly Ile
370 375

Á <210> 47

ES 2 659 118 T3

<211> 277
 <212> PRT
 <213> Homo sapiens

5 <400> 47
 Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15
 Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30
 Leu Gln Trp Glu Glu Asp Ser Asn Ser Val Val Leu Ser Phe Asp Ser
 35 40 45
 Ala Gly Gln Thr Leu Gly Ser Glu Tyr Asp Arg Leu Gly Phe Leu Leu
 50 55 60
 Asn Leu Asp Ser Lys Leu Pro Ala Glu Leu Ala Thr Lys Tyr Ala Asn
 65 70 75 80
 Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala Ser Ala Leu Met Thr
 85 90 95
 Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe Leu Asp Asp
 100 105 110
 Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe Gly
 115 120 125
 Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala Ile Leu Ser Val Thr Lys
 130 135 140
 Glu Tyr Arg Leu Thr Pro Ala Leu Asp Ser Leu Arg Cys Arg Arg Cys
 145 150 155 160
 Ile Ile Val Gly Asn Gly Gly Val Leu Ala Asn Lys Ser Leu Gly Ser
 165 170 175
 Arg Ile Asp Asp Tyr Asp Ile Val Val Arg Leu Asn Ser Ala Pro Val
 180 185 190
 Lys Gly Phe Glu Lys Asp Val Gly Ser Lys Thr Thr Leu Arg Ile Thr
 195 200 205
 Tyr Pro Glu Gly Ala Met Gln Arg Pro Glu Gln Tyr Glu Arg Asp Ser
 210 215 220
 Leu Phe Val Leu Ala Gly Phe Lys Trp Gln Asp Phe Lys Trp Leu Lys
 225 230 235 240
 Tyr Ile Val Tyr Lys Glu Arg Val Ser Trp Thr His Asn Ile Gln Arg
 245 250 255
 Glu Lys Glu Phe Leu Arg Lys Leu Val Lys Ala Arg Val Ile Thr Asp
 260 265 270
 Leu Ser Ser Gly Ile
 275

10 <210> 48
 <211> 155
 <212> PRT
 <213> Homo sapiens

ES 2 659 118 T3

<400> 48

Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15
 Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30
 Leu Gln Trp Glu Glu Asp Ser Asn Ser Val Val Leu Ser Phe Asp Ser
 35 40 45
 Ala Gly Gln Thr Leu Gly Ser Glu Tyr Asp Arg Leu Gly Phe Leu Leu
 50 55 60
 Asn Leu Asp Ser Lys Leu Pro Ala Glu Leu Ala Thr Lys Tyr Ala Asn
 65 70 75 80
 Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala Ser Ala Leu Met Thr
 85 90 95
 Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe Leu Asp Asp
 100 105 110
 Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe Gly
 115 120 125
 Ile Lys Gly Gln Val Leu Asp Ala Gln Tyr Pro Ala Arg Glu Arg Val
 130 135 140
 Ser Ala Glu Ala Gly Glu Ser Ser Arg His His
 145 150 155

5

<210> 49

<211> 413

<212> PRT

10 <213> Homo sapiens

<400> 49

ES 2 659 118 T3

Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15

Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30

Leu Gln Trp Gln Glu Asp Ser Lys Tyr Asp Arg Leu Gly Phe Leu Leu
 35 40 45

Asn Leu Asp Ser Lys Leu Ser Pro Arg Thr Leu Cys Thr Val Val Phe
 50 55 60

Gly Leu Asp Cys Ile Leu Glu Ser Pro Gly Glu Pro Lys Lys Leu Leu
 65 70 75 80

Met Pro Ala Ser His Pro Leu Glu Ile Leu Lys Ser Leu Ser Glu Asp
 85 90 95

Thr Ala Phe Ala Leu Gly Phe Leu Lys Leu Pro Arg Pro Ala Glu Leu
 100 105 110

Ala Thr Lys Tyr Ala Asn Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr
 115 120 125

Ala Ser Ala Leu Met Thr Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala
 130 135 140

Pro Met Phe Leu Asp Asp Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu
 145 150 155 160

Phe Val Pro Pro Phe Gly Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala
 165 170 175

Ile Leu Ser Val Thr Lys Glu Tyr Arg Leu Thr Pro Ala Leu Asp Ser
 180 185 190

Leu Arg Cys Arg Arg Cys Ile Ile Val Gly Asn Gly Gly Val Leu Ala
 195 200 205

Asn Lys Ser Leu Gly Ser Arg Ile Asp Asp Tyr Asp Ile Val Val Arg
 210 215 220

ES 2 659 118 T3

Leu Asn Ser Ala Pro Val Lys Gly Phe Glu Lys Asp Val Gly Ser Lys
 225 230 235 240
 Thr Thr Leu Arg Ile Thr Tyr Pro Glu Gly Ala Met Gln Arg Pro Glu
 245 250 255
 Gln Tyr Glu Arg Asp Ser Leu Phe Val Leu Ala Gly Phe Lys Trp Gln
 260 265 270
 Asp Phe Lys Trp Leu Lys Tyr Ile Val Tyr Lys Glu Arg Val Ser Ala
 275 280 285
 Ser Asp Gly Phe Trp Lys Ser Val Ala Thr Arg Val Pro Lys Glu Pro
 290 295 300
 Pro Glu Ile Arg Ile Leu Asn Pro Tyr Phe Ile Gln Glu Ala Ala Phe
 305 310 315 320
 Thr Leu Ile Gly Leu Pro Phe Asn Asn Gly Leu Met Gly Arg Gly Asn
 325 330 335
 Ile Pro Thr Leu Gly Ser Val Ala Val Thr Met Ala Leu His Gly Cys
 340 345 350
 Asp Glu Val Ala Val Ala Gly Phe Gly Tyr Asp Met Ser Thr Pro Asn
 355 360 365
 Ala Pro Leu His Tyr Tyr Glu Thr Val Arg Met Ala Ala Ile Lys Glu
 370 375 380
 Ser Trp Thr His Asn Ile Gln Arg Glu Lys Glu Phe Leu Arg Lys Leu
 385 390 395 400
 Val Lys Ala Arg Val Ile Thr Asp Leu Ser Ser Gly Ile
 405 410

<210> 50
 <211> 359
 <212> PRT
 <213> Homo sapiens

5

<400> 50
 Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15
 Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30
 Leu Gln Trp Glu Glu Asp Ser Lys Tyr Asp Arg Leu Gly Phe Leu Leu
 35 40 45

10

ES 2 659 118 T3

Asn Leu Asp Ser Lys Leu Pro Ala Glu Leu Ala Thr Lys Tyr Ala Asn
 50 55 60
 Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala Ser Ala Leu Met Thr
 65 70 75 80
 Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe Leu Asp Asp
 85 90 95
 Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe Gly
 100 105 110
 Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala Ile Leu Ser Val Thr Lys
 115 120 125
 Glu Tyr Arg Leu Thr Pro Ala Leu Asp Ser Leu Arg Cys Arg Arg Cys
 130 135 140
 Ile Ile Val Gly Asn Gly Gly Val Leu Ala Asn Lys Ser Leu Gly Ser
 145 150 155 160
 Arg Ile Asp Asp Tyr Asp Ile Val Val Arg Leu Asn Ser Ala Pro Val
 165 170 175
 Lys Gly Phe Glu Lys Asp Val Gly Ser Lys Thr Thr Leu Arg Ile Thr
 180 185 190
 Tyr Pro Glu Gly Ala Met Gln Arg Pro Glu Gln Tyr Glu Arg Asp Ser
 195 200 205
 Leu Phe Val Leu Ala Gly Phe Lys Trp Gln Asp Phe Lys Trp Leu Lys
 210 215 220
 Tyr Ile Val Tyr Lys Glu Arg Val Ser Ala Ser Asp Gly Phe Trp Lys
 225 230 235 240
 Ser Val Ala Thr Arg Val Pro Lys Glu Pro Pro Glu Ile Arg Ile Leu
 245 250 255
 Asn Pro Tyr Phe Ile Gln Glu Ala Ala Phe Thr Leu Ile Gly Leu Pro
 260 265 270
 Phe Asn Asn Gly Leu Met Gly Arg Gly Asn Ile Pro Thr Leu Gly Ser
 275 280 285
 Val Ala Val Thr Met Ala Leu His Gly Cys Asp Glu Val Ala Val Ala
 290 295 300
 Gly Phe Gly Tyr Asp Met Ser Thr Pro Asn Ala Pro Leu His Tyr Tyr
 305 310 315 320
 Glu Thr Val Arg Met Ala Ala Ile Lys Glu Ser Trp Thr His Asn Ile
 325 330 335
 Gln Arg Glu Lys Glu Phe Leu Arg Lys Leu Val Lys Ala Arg Val Ile
 340 345 350
 Thr Asp Leu Ser Ser Gly Ile
 355

Á <210> 51
 <211> 139
 <212> PRT

ES 2 659 118 T3

<213> Homo sapiens

<400> 51

Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15

Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30

Leu Gln Trp Glu Gln Asp Ser Lys Tyr Asp Arg Leu Gly Phe Leu Leu
 35 40 45

Asn Leu Asp Ser Lys Leu Pro Ala Glu Leu Ala Thr Lys Tyr Ala Asn
 50 55 60

Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala Ser Ala Leu Met Thr
 65 70 75 80

Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe Leu Asp Asp
 85 90 95

Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe Gly
 100 105 110

Ile Lys Gly Gln Val Leu Asp Ala Gln Tyr Pro Ala Arg Glu Arg Val
 115 120 125

Ser Ala Glu Ala Gly Glu Ser Ser Arg His His
 130 135

5

<210> 52

<211> 331

<212> PRT

<213> Homo sapiens

10

<400> 52

Met Arg Gly Tyr Leu Val Ala Ile Phe Leu Ser Ala Val Phe Leu Tyr
 1 5 10 15

Tyr Val Leu His Cys Ile Leu Trp Gly Thr Asn Val Tyr Trp Val Ala
 20 25 30

ES 2 659 118 T3

Pro Val Glu Met Lys Arg Arg Asn Lys Ile Gln Pro Cys Leu Ser Lys
35 40 45

Pro Ala Phe Ala Ser Leu Leu Arg Phe His Gln Phe His Pro Phe Leu
50 55 60

Cys Ala Ala Asp Phe Arg Lys Ile Ala Ser Leu Tyr Gly Ser Asp Lys
65 70 75 80

Phe Asp Leu Pro Tyr Gly Met Arg Thr Ser Ala Glu Tyr Phe Arg Leu
85 90 95

Ala Leu Ser Lys Leu Gln Ser Cys Asp Leu Phe Asp Glu Phe Asp Asn
100 105 110

Ile Pro Cys Lys Lys Cys Val Val Val Gly Asn Gly Gly Val Leu Lys
115 120 125

Asn Lys Thr Leu Gly Glu Lys Ile Asp Ser Tyr Asp Val Ile Ile Arg
130 135 140

Met Asn Asn Gly Pro Val Leu Gly His Glu Glu Glu Val Gly Arg Arg
145 150 155 160

Thr Thr Phe Arg Leu Phe Tyr Pro Glu Ser Val Phe Ser Asp Pro Ile
165 170 175

His Asn Asp Pro Asn Thr Thr Val Ile Leu Thr Ala Phe Lys Pro His
180 185 190

Asp Leu Arg Trp Leu Leu Glu Leu Leu Met Gly Asp Lys Ile Asn Thr
195 200 205

Asn Gly Phe Trp Lys Lys Pro Ala Leu Asn Leu Ile Tyr Lys Pro Tyr
210 215 220

Gln Ile Arg Ile Leu Asp Pro Phe Ile Ile Arg Thr Ala Ala Tyr Glu
225 230 235 240

Leu Leu His Phe Pro Lys Val Phe Pro Lys Asn Gln Lys Pro Lys His
245 250 255

Pro Thr Thr Gly Ile Ile Ala Ile Thr Leu Ala Phe Tyr Ile Cys His
260 265 270

Glu Val His Leu Ala Gly Phe Lys Tyr Asn Phe Ser Asp Leu Lys Ser
275 280 285

Pro Leu His Tyr Tyr Gly Asn Ala Thr Met Ser Leu Met Asn Lys Asn
290 295 300

Ala Tyr His Asn Val Thr Ala Glu Gln Leu Phe Leu Lys Asp Ile Ile
305 310 315 320

Glu Lys Asn Leu Val Ile Asn Leu Thr Gln Asp
325 330

5 <210> 53
<211> 753
<212> PRT
<213> Mus musculus

10 <400> 53

ES 2 659 118 T3

Met Glu Thr His Ala His Leu His Arg Glu Gln Ser Tyr Ala Gly Pro
 1 5 10 15
 His Glu Leu Tyr Phe Lys Lys Leu Ser Ser Lys Lys Lys Gln Val Met
 20 25 30
 Glu Lys Asn Gly Asn Asn Arg Lys Leu Arg Val Cys Val Ala Thr Cys
 35 40 45
 Asn Arg Ala Asp Tyr Ser Lys Leu Ala Pro Ile Met Phe Gly Ile Lys
 50 55 60
 Thr Glu Pro Ala Phe Phe Glu Leu Asp Val Val Val Leu Gly Ser His
 65 70 75 80
 Leu Ile Asp Asp Tyr Gly Asn Thr Tyr Arg Met Ile Glu Gln Asp Asp
 85 90 95
 Phe Asp Ile Asn Thr Arg Leu His Thr Ile Val Arg Gly Glu Asp Glu
 100 105 110
 Ala Ala Met Val Glu Ser Val Gly Leu Ala Leu Val Lys Leu Pro Asp
 115 120 125
 Val Leu Asn Arg Leu Lys Pro Asp Ile Met Ile Val His Gly Asp Arg
 130 135 140
 Phe Asp Ala Leu Ala Leu Ala Thr Ser Ala Ala Leu Met Asn Ile Arg
 145 150 155 160
 Ile Leu His Ile Glu Gly Gly Glu Val Ser Gly Thr Ile Asp Asp Ser
 165 170 175
 Ile Arg His Ala Ile Thr Lys Leu Ala His Tyr His Val Cys Cys Thr
 180 185 190
 Arg Ser Ala Glu Gln His Leu Ile Ser Met Cys Glu Asp His Asp Arg
 195 200 205

ES 2 659 118 T3

Ile Leu Leu Ala Gly Cys Pro Ser Tyr Asp Lys Leu Leu Ser Ala Lys
 210 215 220

Asn Lys Asp Tyr Met Ser Ile Ile Arg Met Trp Leu Gly Asp Asp Val
 225 230 235 240

Lys Cys Lys Asp Tyr Ile Val Ala Leu Gln His Pro Val Thr Thr Asp
 245 250 255

Ile Lys His Ser Ile Lys Met Phe Glu Leu Thr Leu Asp Ala Leu Ile
 260 265 270

Ser Phe Asn Lys Arg Thr Leu Val Leu Phe Pro Asn Ile Asp Ala Gly
 275 280 285

Ser Lys Glu Met Val Arg Val Met Arg Lys Lys Gly Ile Glu His His
 290 295 300

Pro Asn Phe Arg Ala Val Lys His Val Pro Phe Asp Gln Phe Ile Gln
 305 310 315 320

Leu Val Ala His Ala Gly Cys Met Ile Gly Asn Ser Ser Cys Gly Val
 325 330 335

Arg Glu Val Gly Ala Phe Gly Thr Pro Val Ile Asn Leu Gly Thr Arg
 340 345 350

Gln Ile Gly Arg Glu Thr Gly Glu Asn Val Leu His Val Arg Asp Ala
 355 360 365

Asp Thr Gln Asp Lys Ile Leu Gln Ala Leu His Leu Gln Phe Gly Lys
 370 375 380

Gln Tyr Pro Cys Ser Lys Ile Tyr Gly Asp Gly Asn Ala Val Pro Arg
 385 390 395 400

Ile Leu Lys Phe Leu Lys Ser Ile Asp Leu Gln Glu Pro Leu Gln Lys
 405 410 415

Lys Phe Cys Phe Pro Pro Val Lys Glu Asn Ile Ser Gln Asp Ile Asp
 420 425 430

His Ile Leu Glu Thr Leu Ser Ala Leu Ala Val Asp Leu Gly Gly Thr
 435 440 445

Asn Leu Arg Val Ala Ile Val Ser Met Lys Gly Glu Ile Val Lys Lys
 450 455 460

Tyr Thr Gln Phe Asn Pro Lys Thr Tyr Glu Glu Arg Ile Ser Leu Ile
 465 470 475 480

ES 2 659 118 T3

Leu Gln Met Cys Val Glu Ala Ala Ala Glu Ala Val Lys Leu Asn Cys
 485 490 495

Arg Ile Leu Gly Val Gly Ile Ser Thr Gly Gly Arg Val Asn Pro Gln
 500 505

Glu Gly Val Val Leu His Ser Thr Lys Leu Ile Gln Glu Trp Asn Ser
 515 520

Val Asp Leu Arg Thr Pro Leu Ser Asp Thr Leu His Leu Pro Val Trp
 530 535 540

Val Asp Asn Asp Gly Asn Cys Ala Ala Met Ala Glu Arg Lys Phe Gly
 545 550 555 560

Gln Gly Lys Gly Gln Glu Asn Phe Val Thr Leu Ile Thr Gly Thr Gly
 565 570 575

Ile Gly Gly Gly Ile Ile His Gln His Glu Leu Ile His Gly Ser Ser
 580 585 590

Phe Cys Ala Ala Glu Leu Gly His Leu Val Val Ser Leu Asp Gly Pro
 595 600 605

Asp Cys Ser Cys Gly Ser His Gly Cys Ile Glu Ala Tyr Ala Ser Gly
 610 615 620

Met Ala Leu Gln Arg Glu Ala Lys Lys Leu His Asp Glu Asp Leu Leu
 625 630 635 640

Leu Val Glu Gly Met Ser Val Pro Lys Asp Glu Ala Val Gly Ala Leu
 645 650 655

His Leu Ile Gln Ala Ala Lys Leu Gly Asn Val Lys Ala Gln Ser Ile
 660 665 670

Leu Arg Thr Ala Gly Thr Ala Leu Gly Leu Gly Val Val Asn Ile Leu
 675 680 685

His Thr Met Asn Pro Ser Leu Val Ile Leu Ser Gly Val Leu Ala Ser
 690 695 700

His Tyr Ile His Ile Val Lys Asp Val Ile Arg Gln Gln Ala Leu Ser
 705 710 715 720

Ser Val Gln Asp Val Asp Val Val Val Ser Asp Leu Val Asp Pro Ala
 725 730 735

Leu Leu Gly Ala Ala Ser Met Val Leu Asp Tyr Thr Thr Arg Arg Ile
 740 745 750

His

- 5 <210> 54
- <211> 722
- <212> PRT
- <213> Homo sapiens

- 10 <400> 54

ES 2 659 118 T3

Met Glu Lys Asn Gly Asn Asn Arg Lys Leu Arg Val Cys Val Ala Thr
 1 5 10 15
 Cys Asn Arg Ala Asp Tyr Ser Lys Leu Ala Pro Ile Met Phe Gly Ile
 20 25 30
 Lys Thr Glu Pro Glu Phe Phe Glu Leu Asp Val Val Val Leu Gly Ser
 35 40 45
 His Leu Ile Asp Asp Tyr Gly Asn Thr Tyr Arg Met Ile Glu Gln Asp
 50 55 60
 Asp Phe Asp Ile Asn Thr Arg Leu His Thr Ile Val Arg Gly Glu Asp
 65 70 75 80
 Glu Ala Ala Met Val Glu Ser Val Gly Leu Ala Leu Val Lys Leu Pro
 85 90 95
 Asp Val Leu Asn Arg Leu Lys Pro Asp Ile Met Ile Val His Gly Asp
 100 105 110
 Arg Phe Asp Ala Leu Ala Leu Ala Thr Ser Ala Ala Leu Met Asn Ile
 115 120 125
 Arg Ile Leu His Ile Glu Gly Gly Glu Val Ser Gly Thr Ile Asp Asp
 130 135 140
 Ser Ile Arg His Ala Ile Thr Lys Leu Ala His Tyr His Val Cys Cys
 145 150 155 160
 Thr Arg Ser Ala Glu Gln His Leu Ile Ser Met Cys Glu Asp His Asp
 165 170 175
 Arg Ile Leu Leu Ala Gly Cys Pro Ser Tyr Asp Lys Leu Leu Ser Ala
 180 185 190
 Lys Asn Lys Asp Tyr Met Ser Ile Ile Arg Met Trp Leu Gly Asp Asp
 195 200 205
 Val Lys Ser Lys Asp Tyr Ile Val Ala Leu Gln His Pro Val Thr Thr
 210 215 220

ES 2 659 118 T3

Asp Ile Lys His Ser Ile Lys Met Phe Glu Leu Thr Leu Asp Ala Leu
 225 230 235 240
 Ile Ser Phe Asn Lys Arg Thr Leu Val Leu Phe Pro Asn Ile Asp Ala
 245 250 255
 Gly Ser Lys Glu Met Val Arg Val Met Arg Lys Lys Gly Ile Glu His
 260 265 270
 His Pro Asn Phe Arg Ala Val Lys His Val Pro Phe Asp Gln Phe Ile
 275 280 285
 Gln Leu Val Ala His Ala Gly Cys Met Ile Gly Asn Ser Ser Cys Gly
 290 295
 Val Arg Glu Val Gly Ala Phe Gly Thr Pro Val Ile Asn Leu Gly Thr
 305 310 315 320
 Arg Gln Ile Gly Arg Glu Thr Gly Glu Asn Val Leu His Val Arg Asp
 325 330 335
 Ala Asp Thr Gln Asp Lys Ile Leu Gln Ala Leu His Leu Gln Phe Gly
 340 345 350
 Lys Gln Tyr Pro Cys Ser Lys Ile Tyr Gly Asp Gly Asn Ala Val Pro
 355 360 365
 Arg Ile Leu Lys Phe Leu Lys Ser Ile Asp Leu Gln Glu Pro Leu Gln
 370 375 380
 Lys Lys Phe Cys Phe Pro Pro Val Lys Glu Asn Ile Ser Gln Asp Ile
 385 390 395 400
 Asp His Ile Leu Glu Thr Leu Ser Ala Leu Ala Val Asp Leu Gly Gly
 405 410 415
 Thr Asn Leu Arg Val Ala Ile Val Ser Met Lys Gly Glu Ile Val Lys
 420 425 430
 Lys Tyr Thr Gln Phe Asn Pro Lys Thr Tyr Glu Glu Arg Ile Asn Leu
 435 440 445
 Ile Leu Gln Met Cys Val Glu Ala Ala Ala Glu Ala Val Lys Leu Asn
 450 455 460
 Cys Arg Ile Leu Gly Val Gly Ile Ser Thr Gly Gly Arg Val Asn Pro
 465 470 475 480
 Arg Glu Gly Ile Val Leu His Ser Thr Lys Leu Ile Gln Glu Trp Asn
 485 490 495

ES 2 659 118 T3

Ser Val Asp Leu Arg Thr Pro Leu Ser Asp Thr Leu His Leu Pro Val
 500 505 510

Trp Val Asp Asn Asp Gly Asn Cys Ala Ala Leu Ala Glu Arg Lys Phe
 515 520 525

Gly Gln Gly Lys Gly Leu Glu Asn Phe Val Thr Leu Ile Thr Gly Thr
 530 535 540

Gly Ile Gly Gly Gly Ile Ile His Gln His Glu Leu Ile His Gly Ser
 545 550 555 560

Ser Phe Cys Ala Ala Glu Leu Gly His Leu Val Val Ser Leu Asp Gly
 565 570 575

Pro Asp Cys Ser Cys Gly Ser His Gly Cys Ile Glu Ala Tyr Ala Ser
 580 585 590

Gly Met Ala Leu Gln Arg Glu Ala Lys Lys Leu His Asp Glu Asp Leu
 595 600 605

Leu Leu Val Glu Gly Met Ser Val Pro Lys Asp Glu Ala Val Gly Ala
 610 615 620

Leu His Leu Ile Gln Ala Ala Lys Leu Gly Asn Ala Lys Ala Gln Ser
 625 630 635 640

Ile Leu Arg Thr Ala Gly Thr Ala Leu Gly Leu Gly Val Val Asn Ile
 645 650 655

Leu His Thr Met Asn Pro Ser Leu Val Ile Leu Ser Gly Val Leu Ala
 660 665 670

Ser His Tyr Ile His Ile Val Lys Asp Val Ile Arg Gln Gln Ala Leu
 675 680 685

Ser Ser Val Gln Asp Val Asp Val Val Val Ser Asp Leu Val Asp Pro
 690 695 700

Ala Leu Leu Gly Ala Ala Ser Met Val Leu Asp Tyr Thr Thr Arg Arg
 705 710 715 720

Ile Tyr

<210> 55

<211> 427

5 <212> PRT

<213> Homo sapiens

<400> 55

Met Ser Lys Gly Leu Pro Ala Arg Gln Asp Met Glu Lys Glu Arg Glu

10

ES 2 659 118 T3

Glu Leu Gln Arg Tyr Ala Glu Glu Val Gly Ile Phe Phe Thr Ala Ser
 100 105 110
 Gly Met Asp Glu Met Ala Val Glu Phe Leu His Glu Leu Asn Val Pro
 115 120 125
 Phe Phe Lys Val Gly Ser Gly Asp Thr Asn Asn Phe Pro Tyr Leu Glu
 130 135 140
 Lys Thr Ala Lys Lys Gly Arg Pro Met Val Ile Ser Ser Gly Met Gln
 145 150 155
 Ser Met Asp Thr Met Lys Gln Val Tyr Gln Ile Val Lys Pro Leu Asn
 165 170 175
 Pro Asn Phe Cys Phe Leu Gln Cys Thr Ser Ala Tyr Pro Leu Gln Pro
 180 185 190
 Glu Asp Val Asn Leu Arg Val Ile Ser Glu Tyr Gln Lys Leu Phe Pro
 195 200 205
 Asp Ile Pro Ile Gly Tyr Ser Gly His Glu Thr Gly Ile Ala Ile Ser
 210 215 220
 Val Ala Ala Val Ala Leu Gly Ala Lys Val Leu Glu Arg His Ile Thr
 225 230 235 240
 Leu Asp Lys Thr Trp Lys Gly Ser Asp His Ser Ala Ser Leu Glu Pro
 245 250 255
 Gly Glu Leu Ala Glu Leu Val Arg Ser Val Arg Leu Val Glu Arg Ala
 260 265 270
 Leu Gly Ser Pro Thr Lys Gln Leu Leu Pro Cys Glu Met Ala Cys Asn
 275 280 285
 Glu Lys Leu Gly Lys Ser Val Val Ala Lys Val Lys Ile Pro Glu Gly
 290 295 300
 Thr Ile Leu Thr Met Asp Met Leu Thr Val Lys Val Gly Glu Pro Lys
 305 310 315 320
 Gly Tyr Pro Pro Glu Asp Ile Phe Asn Leu Val Gly Lys Lys Val Leu
 325 330 335
 Val Thr Val Glu Glu Asp Asp Thr Ile Met Glu Glu Leu Val Asp Asn
 340 345 350
 His Gly Lys Lys Ile Lys Ser
 355

<210> 57

<211> 432

Á <212> PRT

<213> Mus musculus

<400> 57

ES 2 659 118 T3

Met Asp Ala Leu Glu Lys Gly Ala Ala Thr Ser Gly Pro Ala Pro Arg
 1 5 10 15

Gly Arg Pro Ser Arg Gly Arg Pro Pro Lys Leu Gln Arg Ser Arg Gly
 20 25 30

Ala Gly Arg Gly Leu Glu Lys Pro Pro His Leu Ala Ala Leu Val Leu
 35 40 45

Ala Arg Gly Gly Ser Lys Gly Ile Pro Leu Lys Asn Ile Lys Arg Leu
 50 55 60

Ala Gly Val Pro Leu Ile Gly Trp Val Leu Arg Ala Ala Leu Asp Ala
 65 70 75 80

Gly Val Phe Gln Ser Val Trp Val Ser Thr Asp His Asp Glu Ile Glu
 85 90 95

Asn Val Ala Lys Gln Phe Gly Ala Gln Val His Arg Arg Ser Ser Glu
 100 105 110

Thr Ser Lys Asp Ser Ser Thr Ser Leu Asp Ala Ile Val Glu Phe Leu
 115 120 125

Asn Tyr His Asn Glu Val Asp Ile Val Gly Asn Ile Gln Ala Thr Ser
 130 135 140

Pro Cys Leu His Pro Thr Asp Leu Gln Lys Val Ala Glu Met Ile Arg
 145 150 155 160

Glu Glu Gly Tyr Asp Ser Val Phe Ser Val Val Arg Arg His Gln Phe
 165 170 175

Arg Trp Ser Glu Ile Gln Lys Gly Val Arg Glu Val Thr Glu Pro Leu
 180 185 190

Asn Leu Asn Pro Ala Lys Arg Pro Arg Arg Gln Asp Trp Asp Gly Glu
 195 200 205

Leu Tyr Glu Asn Gly Ser Phe Tyr Phe Ala Lys Arg His Leu Ile Glu
 210 215 220

Met Gly Tyr Leu Gln Gly Gly Lys Met Ala Tyr Tyr Glu Met Arg Ala
 225 230 235 240

ES 2 659 118 T3

Glu His Ser Val Asp Ile Asp Val Asp Ile Asp Trp Pro Ile Ala Glu
 245 250 255
 Gln Arg Val Leu Arg Phe Gly Tyr Phe Gly Lys Glu Lys Leu Lys Glu
 260 265 270
 Ile Lys Leu Leu Val Cys Asn Ile Asp Gly Cys Leu Thr Asn Gly His
 275 280 285
 Ile Tyr Val Ser Gly Asp Gln Lys Glu Ile Ile Ser Tyr Asp Val Lys
 290 295 300
 Asp Ala Ile Gly Ile Ser Leu Leu Lys Lys Ser Gly Ile Glu Val Arg
 305 310 315 320
 Leu Ile Ser Glu Arg Ala Cys Ser Lys Gln Thr Leu Ser Ala Leu Lys
 325 330 335
 Leu Asp Cys Lys Thr Glu Val Ser Val Ser Asp Lys Leu Ala Thr Val
 340 345 350
 Asp Glu Trp Arg Lys Glu Met Gly Leu Cys Trp Lys Glu Val Ala Tyr
 355 360 365
 Leu Gly Asn Glu Val Ser Asp Glu Glu Cys Leu Lys Arg Val Gly Leu
 370 375 380
 Ser Ala Val Pro Ala Asp Ala Cys Ser Gly Ala Gln Lys Ala Val Gly
 385 390 400
 Tyr Ile Cys Lys Cys Ser Gly Gly Arg Gly Ala Ile Arg Glu Phe Ala
 405 410 415
 Glu His Ile Phe Leu Leu Ile Glu Lys Val Asn Asn Ser Cys Gln Lys
 420 425 430

<210> 58
 <211> 434
 <212> PRT
 <213> Homo sapiens

5

<400> 58
 Met Asp Ser Val Glu Lys Gly Ala Ala Thr Ser Val Ser Asn Pro Arg
 1 5 10 15
 Gly Arg Pro Ser Arg Gly Arg Pro Pro Lys Leu Gln Arg Asn Ser Arg
 20 25 30
 Gly Gly Gln Gly Arg Gly Val Glu Lys Pro Pro His Leu Ala Ala Leu
 35 40 45

10

ES 2 659 118 T3

Ile Leu Ala Arg Gly Gly Ser Lys Gly Ile Pro Leu Lys Asn Ile Lys
50 55 60

His Leu Ala Gly Val Pro Leu Ile Gly Trp Val Leu Arg Ala Ala Leu
65 70 75 80

Asp Ser Gly Ala Phe Gln Ser Val Trp Val Ser Thr Asp His Asp Glu
85 90 95

Ile Glu Asn Val Ala Lys Gln Phe Gly Ala Gln Val His Arg Arg Ser
100 105 110

Ser Glu Val Ser Lys Asp Ser Ser Thr Ser Leu Asp Ala Ile Ile Glu
115 120 125

Phe Leu Asn Tyr His Asn Glu Val Asp Ile Val Gly Asn Ile Gln Ala
130 135 140

Thr Ser Pro Cys Leu His Pro Thr Asp Leu Gln Lys Val Ala Glu Met
145 150 155 160

Ile Arg Glu Glu Gly Tyr Asp Ser Val Phe Ser Val Val Arg Arg His
165 170 175

Gln Phe Arg Trp Ser Glu Ile Gln Lys Gly Val Arg Glu Val Thr Glu
180 185 190

Pro Leu Asn Leu Asn Pro Ala Lys Arg Pro Arg Arg Gln Asp Trp Asp
195 200 205

Gly Glu Leu Tyr Glu Asn Gly Ser Phe Tyr Phe Ala Lys Arg His Leu
210 215 220

Ile Glu Met Gly Tyr Leu Gln Gly Gly Lys Met Ala Tyr Tyr Glu Met
225 230 235 240

Arg Ala Glu His Ser Val Asp Ile Asp Val Asp Ile Asp Trp Pro Ile
245 250 255

Ala Glu Gln Arg Val Leu Arg Tyr Gly Tyr Phe Gly Lys Glu Lys Leu
260 265 270

Lys Glu Ile Lys Leu Leu Val Cys Asn Ile Asp Gly Cys Leu Thr Asn
275 280 285

Gly His Ile Tyr Val Ser Gly Asp Gln Lys Glu Ile Ile Ser Tyr Asp
290 295 300

Val Lys Asp Ala Ile Gly Ile Ser Leu Leu Lys Lys Ser Gly Ile Glu
305 310 315 320

ES 2 659 118 T3

Val Arg Leu Ile Ser Glu Arg Ala Cys Ser Lys Gln Thr Leu Ser Ser
 325 330 335

Leu Lys Leu Asp Cys Lys Met Glu Val Ser Val Ser Asp Lys Leu Ala
 340 345 350

Val Val Asp Glu Trp Arg Lys Glu Met Gly Leu Cys Trp Lys Glu Val
 355 360 365

Ala Tyr Leu Gly Asn Glu Val Ser Asp Glu Glu Cys Leu Lys Arg Val
 370 375 380

Gly Leu Ser Gly Ala Pro Ala Asp Ala Cys Ser Thr Ala Gln Lys Ala
 385 390 395 400

Val Gly Tyr Ile Cys Lys Cys Asn Gly Gly Arg Gly Ala Ile Arg Glu
 405 410 415

Phe Ala Glu His Ile Cys Leu Leu Met Glu Lys Val Asn Asn Ser Cys
 420 425 430

Gln Lys

<210> 59

<211> 336

<212> PRT

<213> Mus musculus

5

<400> 59

Met Ala Pro Ala Arg Glu Asn Val Ser Leu Phe Phe Lys Leu Tyr Cys
 1 5 10 15

Leu Thr Val Met Thr Leu Val Ala Ala Ala Tyr Thr Val Ala Leu Arg
 20 25 30

Tyr Thr Arg Thr Thr Ala Glu Glu Leu Tyr Phe Ser Thr Thr Ala Val
 35 40 45

Cys Ile Thr Glu Val Ile Lys Leu Leu Ile Ser Val Gly Leu Leu Ala
 50 55 60

Lys Glu Thr Gly Ser Leu Gly Arg Phe Lys Ala Ser Leu Ser Glu Asn
 65 70 75 80

Val Leu Gly Ser Pro Lys Glu Leu Ala Lys Leu Ser Val Pro Ser Leu
 85 90 95

Val Tyr Ala Val Gln Asn Asn Met Ala Phe Leu Ala Leu Ser Asn Leu
 100 105 110

Asp Ala Ala Val Tyr Gln Val Thr Tyr Gln Leu Lys Ile Pro Cys Thr

10

ES 2 659 118 T3

115 120 125

Ala Leu Cys Thr Val Leu Met Leu Asn Arg Thr Leu Ser Lys Leu Gln
 130 135

Trp Ile Ser Val Phe Met Leu Cys Gly Gly Val Thr Leu Val Gln Trp
 145 150 155 160

Lys Pro Ala Gln Ala Thr Lys Val Val Val Ala Gln Asn Pro Leu Leu
 165 170 175

Gly Phe Gly Ala Ile Ala Ile Ala Val Leu Cys Ser Gly Phe Ala Gly
 180 185

Val Tyr Phe Glu Lys Val Leu Lys Ser Ser Asp Thr Ser Leu Trp Val
 195 200 205

Arg Asn Ile Gln Met Tyr Leu Ser Gly Ile Val Val Thr Leu Ala Gly
 210 215 220

Thr Tyr Leu Ser Asp Gly Ala Glu Ile Gln Glu Lys Gly Phe Phe Tyr
 225 230 235 240

Gly Tyr Thr Tyr Tyr Val Trp Phe Val Ile Phe Leu Ala Ser Val Gly
 245 250 255

Gly Leu Tyr Thr Ser Val Val Val Lys Tyr Thr Asp Asn Ile Met Lys
 260 265 270

Gly Phe Ser Ala Ala Ala Ala Ile Val Leu Ser Thr Ile Ala Ser Val
 275 280 285

Leu Leu Phe Gly Leu Gln Ile Thr Leu Ser Phe Ala Leu Gly Ala Leu
 290 295 300

Leu Val Cys Val Ser Ile Tyr Leu Tyr Gly Leu Pro Arg Gln Asp Thr
 305 310 315 320

Thr Ser Ile Gln Gln Glu Ala Thr Ser Lys Glu Arg Ile Ile Gly Val
 325 330 335

<210> 60
 <211> 337
 <212> PRT
 <213> Homo sapiens

<400> 60
 Met Ala Ala Pro Arg Asp Asn Val Thr Leu Leu Phe Lys Leu Tyr Cys
 1 5 10 15

Leu Ala Val Met Thr Leu Met Ala Ala Val Tyr Thr Ile Ala Leu Arg
 20 25 30

5

10

ES 2 659 118 T3

Tyr Thr Arg Thr Ser Asp Lys Glu Leu Tyr Phe Ser Thr Thr Ala Val
 35 40 45
 Cys Ile Thr Glu Val Ile Lys Leu Leu Leu Ser Val Gly Ile Leu Ala
 50 55 60
 Lys Glu Thr Gly Ser Leu Gly Arg Phe Lys Ala Ser Leu Arg Glu Asn
 65 70 75 80
 Val Leu Gly Ser Pro Lys Glu Leu Leu Lys Leu Ser Val Pro Ser Leu
 85 90 95
 Val Tyr Ala Val Gln Asn Asn Met Ala Phe Leu Ala Leu Ser Asn Leu
 100 105 110
 Asp Ala Ala Val Tyr Gln Val Thr Tyr Gln Leu Lys Ile Pro Cys Thr
 115 120 125
 Ala Leu Cys Thr Val Leu Met Leu Asn Arg Thr Leu Ser Lys Leu Gln
 130 135 140
 Trp Val Ser Val Phe Met Leu Cys Ala Gly Val Thr Leu Val Gln Trp
 145 150 155 160
 Lys Pro Ala Gln Ala Thr Lys Val Val Val Glu Gln Asn Pro Leu Leu
 165 170 175
 Gly Phe Gly Ala Ile Ala Ile Ala Val Leu Cys Ser Gly Phe Ala Gly
 180 185 190
 Val Tyr Phe Glu Lys Val Leu Lys Ser Ser Asp Thr Ser Leu Trp Val
 195 200 205
 Arg Asn Ile Gln Met Tyr Leu Ser Gly Ile Ile Val Thr Leu Ala Gly
 210 215 220
 Val Tyr Leu Ser Asp Gly Ala Glu Ile Lys Glu Lys Gly Phe Phe Tyr
 225 230 235 240
 Gly Tyr Thr Tyr Tyr Val Trp Phe Val Ile Phe Leu Ala Ser Val Gly
 245 250 255
 Gly Leu Tyr Thr Ser Val Val Val Lys Tyr Thr Asp Asn Ile Met Lys
 260 265 270
 Gly Phe Ser Ala Ala Ala Ala Ile Val Leu Ser Thr Ile Ala Ser Val
 275 280 285
 Met Leu Phe Gly Leu Gln Ile Thr Leu Thr Phe Ala Leu Gly Thr Leu
 290 295 300
 Leu Val Cys Val Ser Ile Tyr Leu Tyr Gly Leu Pro Arg Gln Asp Thr
 305 310 315 320
 Thr Ser Ile Gln Gln Gly Glu Thr Ala Ser Lys Glu Arg Val Ile Gly
 325 330 335

val

Á <210> 61
 <211> 31
 <212> DNA

<213> Secuencia Artificial
 <220>
 <223> CaMV35Sfwd
 5 <400> 61
 gaacatatgg tggattgatg tgatctactc c 31
 <210> 62
 10 <211> 24
 <212> DNA
 <213> Secuencia Artificial
 <220>
 15 <223> CaMV35Srev
 <400> 62
 aattctcgag gaattcgcc gagg 24
 20 <210> 63
 <211> 19
 <212> RNA
 <213> Secuencia Artificial
 25 <220>
 <223> siRNA
 <400> 63
 30 ggccauucgu uacuagcaa 19
 <210> 64
 <211> 19
 <212> RNA
 <213> Secuencia Artificial
 35 <220>
 <223> siRNA
 <400> 64
 40 gugguucguu gggaaauga 19
 <210> 65
 <211> 19
 <212> RNA
 45 <213> Secuencia Artificial
 <220>
 <223> siRNA
 50 <400> 65
 ccgucugugu gaaauuggu 19
 <210> 66
 <211> 19
 55 <212> RNA
 <213> Secuencia Artificial
 <220>
 <223> siRNA
 60 <400> 66
 cgguaguagu gcuuguugu 19
 <210> 67
 65 <211> 19

<212> RNA
 <213> Secuencia Artificial

 <220>
 5 <223> siRNA

 <400> 67
 cugccaugga uauugucaa 19

 10 <210> 68
 <211> 19
 <212> RNA
 <213> Secuencia Artificial

 15 <220>
 <223> siRNA

 <400> 68
 gccuuugguc cugaagaaa 19
 20
 <210> 69
 <211> 20
 <212> DNA
 <213> Secuencia Artificial
 25
 <220>
 <223>Cebador

 <400> 69
 30 agcaagagtg ctcgtgtgc 20

 <210> 70
 <211> 20
 <212> DNA
 35 <213> Secuencia Artificial

 <220>
 <223>Cebador

 <400> 70
 40 acccgacgtc gtacgttcc 20

 <210> 71
 <211> 20
 <212> DNA
 45 <213> Secuencia Artificial

 <220>
 <223>Cebador
 50
 <400> 71
 atccgtcgga atgtagtcc 20

 <210> 72
 <211> 19
 <212> DNA
 55 <213> Secuencia Artificial

 <220>
 <223>Cebador
 60
 <400> 72
 tactcgtcgg taccctcgt 19

 65 <210> 73

<211> 20
 <212> DNA
 <213> Secuencia Artificial

5 <220>
 <223>Cebador

<400> 73
 aggtccttcg cgacaatac 20

10 <210> 74
 <211> 20
 <212> DNA
 <213> Secuencia Artificial

15 <220>
 <223>Cebador

<400> 74
 acggaatcac gaatgacgtt 20

<210> 75
 <211> 21
 <212> DNA
 25 <213> Secuencia Artificial

<220>
 <223>Cebador

30 <400> 75
 accgtgaaac agtggattgg a 21

<210> 76
 <211> 22
 35 <212> DNA
 <213> Secuencia Artificial

<220>
 <223>Cebador

40 <400> 76
 ctgggatggt gtacgtcaat gt 22

<210> 77
 <211> 20
 <212> DNA
 <213> Secuencia Artificial

<220>
 <223>Cebador

50 <400> 77
 aggtccttcg cgacaatac 20

55 <210> 78
 <211> 20
 <212> DNA
 <213> Secuencia Artificial

<220>
 <223>Cebador

<400> 78
 acggaatcac gaatgacgtt 20

65

ES 2 659 118 T3

<210> 79
 <211> 4433
 <212> DNA
 <213> Secuencia Artificial

5

<220>
 <223> vector BSJ-25

<400> 79
 gatatcatcg actaattcga gctcgggtacc cggggatcct ctagagtcga cctgcaggca 60
 tgcaagcttc agaagcgtgc tctcgaactc aaccaggagc gtgcggcaca aatgggcatc 120
 cttgctctca tggcgcacga acagttggga gtctctatcc ttccttaaaa atttaatttt 180
 cattagttgc agtcaactcg ctttggtttc acagtcagga ataacactag ctcgtcttca 240
 ccattggatgc caatctcgcc tattcatggt gtataaaagt tcaacatcca aagctagaac 300
 ttttggaaaag agaaagaata tccgaatagg gcacggcgtg ccgattggt ggagtggact 360
 agcagaaaagt gaggaaggca caggatgagt tttctcgaga cataccttca gcgtcgtctt 420
 cactgtcaca gtcaactgac agtaatcgtt gatccggaga gattcaaat tcaatctggt 480
 tggacctgga taagacaca gagcgcacac ctgacatgaa cgccgtaaac agcaaatcct 540
 ggttgaacac gtatcctttt gggggcctcc gctacgacgc tcgctccagc tggggcttcc 600
 ttactataca cagcgcgcat atttcacggt tgccagatgt caagatggcc aagttgacca 660
 gtgccgttcc ggtgctcacc gcgcgcgacg tcgccggagc ggtcaggttc tggaccgacc 720
 ggctcgggtt ctcccgggac ttcgtggagg acgacttcgc cgggtgtggtc cgggacgacg 780
 tgaccctggt catcagcgcg gtccaggacc aggtggtgcc ggacaacacc ctggcctggg 840
 tgtgggtgcg cggcctggac gagctgtacg ccgagtggtc ggaggtcgtg tccacgaact 900
 tccgggacgc ctcccgggac gccatgaccg agatcggcga gcagccgtgg gggcgggagt 960
 tcgccctgcg cgaccgggcc ggcaactgcg tgcacttcgt ggccgaggag caggactgaa 1020

10

ES 2 659 118 T3

ccttccttaa aaatttaatt ttcattagtt gcagtcactc cgctttggtt tcacagtcag 1080
 gaataaact agctcgtctt caccatggat gccaatctcg cctattcatg gtgtataaaa 1140
 gttcaacatc caaagctaga acttttggaa agagaaagaa tatccgaata gggcacggcg 1200
 tgccgtattg ttggagtgga ctagcagaaa gtgaggaagg cacaggatga gttttctcga 1260
 ggccggcttc cctatagtga gtcgtattaa ttctgataag ccaggttaac ctgcattaat 1320
 gaatcggcca acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc gcttcctcgc 1380
 tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 1440
 cggtaatcag gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 1500
 gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgcg ggcgtttttc cataggtccc 1560
 gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagac 1620
 gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 1680
 ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaaagcgtg gcgctttctc 1740
 aatgctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 1800
 tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 1860
 ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 1920
 gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 1980
 ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttacctc ggaaaagag 2040
 ttggtagctc ttgatccggc aacaaacca ccgctggtag cggtggtttt tttgtttgca 2100
 agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 2160
 ggtctgacgc tcagtggaac gaaaactcac gtaagggat tttggtcatg agattatcaa 2220
 aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 2280
 tatatgagta aacttggctt gacagttacc aatgcttaat cagtgaggca cctatctcag 2340
 cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga 2400
 tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac 2460
 cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc 2520
 ctgcaacttt atccgcctcc atccagtcta ttaattggtg ccgggaagct agagtaagta 2580
 gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac 2640
 gctcgtcgtt tggatggct tcattcagct ccggttccca acgatcaagg cgagttacat 2700
 gatccccat gttgtgcaaa aaagcggtta gctcctcgg tcctccgac gttgtcagaa 2760
 gtaagtggc cgcagtgtta tcaactatgg ttatggcagc actgcataat tctcttactg 2820
 tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag 2880
 aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc 2940
 cacatagcag aactttaaaa gtgctcatca ttggaaacg ttcttcgggg cgaaaactct 3000
 caaggatctt accgctggtg agatccagtt cgatgtaacc cactcgtgca cccaactgat 3060

ES 2 659 118 T3

cttcagcatc tttfactttc accagcgttt ctgggtgagc aaaaacagga aggcaaatg 3120
 ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc ttcctttttc 3180
 aatattattg aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta 3240
 tttagaaaaa taaacaaata ggggttccgc gcacatttcc cggaaaagtg ccacctgacg 3300
 tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct 3360
 ttcgtctcgc gcgtttcggg gatgacggtg aaaacctctg acacatgcag ctcccggaga 3420
 cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag 3480
 cgggtgttgg cgggtgtcgg ggctggctta actatgcggc atcagagcag attgtactga 3540
 gagtgcacca tatggtggat tgatgtgatc tactccaaaa atatcaaaga tacagtctca 3600
 gaagaccaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcggg 3660
 ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 3720
 tacaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 3780
 ggtcccaaa atggaccccc acccagcagg agcatcgtgg aaaaagaaga cgttccaacc 3840
 acgtcttcaa agcaagtggg ttgatgtgat ctactccaaa aatatcaaag atacagtctc 3900
 agaagaccaa agggcaattg agacttttca acaaagggtg atatccggaa acctcctcgg 3960
 attccattgc ccagctatct gtcactttat tgtgaagata gtggaaaagg aaggtggctc 4020
 ctacaaatgc catcattgcg ataaaggaaa ggccatcgtt gaagatgcct ctgcccagag 4080
 tggccccaaa gatggacccc caccacagag gagcatcgtg gaaaaagaag acgttccaac 4140
 cacgtcttca aagcaagtgg attgatgtga tatctccact gacgtaaggg atgacgcaca 4200
 atcccactat ccttcgcaag accttctctc tatataagga agttcatttc attcggagag 4260
 gtacgtatct ttacaacaat taccaacaac aacaacaaac aacaacaaca ttacatttta 4320
 cattctacaa ctacatctag aactagtggg tccaaggaga tataacaatg aagactaatc 4380
 tttttctctt tctcatcttt tcacttctcc tatcattatc ctcggccgaa ttc 4433

- <210> 80
- <211> 819
- <212> DNA
- <213> Virus del mosaico de la coliflor

5

<400> 80
 catatggtgg attgatgtga tctactccaa aaatatcaaa gatacagtct cagaagacca 60
 aagggcaatt gagacttttc aacaagggtt aatatccgga aacctcctcg gattccattg 120
 cccagctatc tgtcacttta ttgtgaagat agtggaagaa gaaggtggct cctacaaatg 180
 ccattcattgc gataaaggaa aggccatcgt tgaagatgcc tctgcccaga gtgggtccaa 240
 agatggaccc ccaccacga ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc 300
 aaagcaagtg gattgatgtg atctactcca aaaatatcaa agatacagtc tcagaagacc 360
 aaagggcaat tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt 420

10

gccagctat ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat 480
 gccatcattg cgataaagga aagccatcgt ttgaagatgc ctctgcccag agtgggtcca 540
 aagatggacc cccaccacag aggagcatcg tggaaaaaga agacgttcca accacgtctt 600
 caaagcaagt ggattgatgt gatattctca ctgacgtaag ggatgacgca caatcccact 660
 atccttcgca agacccttcc tctatataag gaagttcatt tcattcggag aggtacgtat 720
 ttttacaaca attaccaaca acaacaacaa acaacaacaa cattacattt tacattctac 780
 aactacatct agaactagtg gatccaagga gatataaca 819

ES 2 659 118 T3

<210> 81
<211> 66
<212> DNA
<213> Virus del mosaico del tabaco

5

<400> 81
atgaagacta atctttttct ctttctcatc ttttcacttc tcctatcatt atcctcggcc 60
gaattc 66

10

REIVINDICACIONES

1. Una *Phaeodactylum tricornutum* transformada cuya vía de N-glicosilación ha sido modificada por la inactivación de β -N-acetilglucosaminidasas tanto de las secuencias de aminoácidos SEQ ID NO 9 como SEQ ID NO 11, no presentando dicha β -N-acetilglucosaminidasa inactivada ninguna actividad enzimática.
- 5 2. Una *Phaeodactylum tricornutum* transformada de acuerdo con la reivindicación 1, en donde dicha *Phaeodactylum tricornutum* transformada cuya vía de N-glicosilación ha sido modificada comprende una secuencia de ácidos nucleicos unida operativamente a un promotor,
 - 10 (i) dicha secuencia de ácidos nucleicos codifica una N-acetilglucosaminiltransferasa I que tiene la secuencia de aminoácidos SEQ ID NO 1, o una secuencia de aminoácidos que comprende al menos la SEQ ID NO 4, y
 - (ii) dicho promotor unido operativamente tiene una identidad de secuencias inferior al 50% con la SEQ ID NO 3.
- 15 3. La *P. tricornutum* transformada de acuerdo con la reivindicación 1, en donde dicha *P. tricornutum* transformada cuya vía de N-glicosilación ha sido modificada comprende una secuencia de ácidos nucleicos unida operativamente a un promotor,
 - (i) dicha secuencia de ácidos nucleicos codifica una α -manosidasa II que tiene la secuencia de aminoácidos SEQ ID NO 5, y
 - (ii) dicho promotor unido operativamente tiene una identidad de secuencias inferior al 50% con la SEQ ID NO 7.
- 20 4. La *P. tricornutum* transformada de acuerdo con una cualquiera de las reivindicaciones 1 a 3, que comprende además una secuencia de ácidos nucleicos unida operativamente a un promotor, en donde dicha secuencia de ácidos nucleicos codifica una N-acetilglucosaminiltransferasa II.
5. La *P. tricornutum* transformada de acuerdo con la reivindicación 4, en donde:
 - 25 (i) dicha N-acetilglucosaminiltransferasa II comprende la secuencia de aminoácidos SEQ ID NO 15, y
 - (ii) dicho promotor unido operativamente tiene una identidad de secuencias inferior al 50% con la SEQ ID NO 7.
6. La *P. tricornutum* transformada de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en donde dicha *P. tricornutum* transformada comprende además otra secuencia de ácidos nucleicos unida operativamente a un promotor, codificando dicha secuencia de ácidos nucleicos un polipéptido que se expresa y glicosila en la *P. tricornutum* transformada.
- 30 7. La *P. tricornutum* transformada de acuerdo con la reivindicación 6, en donde dicho polipéptido glicosilado tiene al menos una estructura $\text{GlcNAcMan}_5\text{GlcNAc}_2$, preferiblemente una estructura $\text{GlcNAcMan}_4\text{GlcNAc}_2$, $\text{GlcNAcMan}_3\text{GlcNAc}_2$ y/o $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$.
- 35 8. La *P. tricornutum* transformada de acuerdo con una cualquiera de las reivindicaciones 6 o 7, en donde dicho polipéptido se selecciona del grupo que comprende eritropoyetina, citoquinas, anticuerpos y sus fragmentos, factores de coagulación, hormonas, beta-glucocerebrosidasa, pentraxina-3, anticuerpos anti-TNF, α -glucosidasa ácida, α -L-iduronidasa y sus derivados.
- 40 9. La *P. tricornutum* transformada de acuerdo con una cualquiera de las reivindicaciones 1 a 8, en donde dicha *P. tricornutum* transformada comprende además todavía otra secuencia de ácidos nucleicos unida operativamente a un promotor, codificando dicha secuencia de ácidos nucleicos al menos una enzima seleccionada entre N-acetilglucosaminiltransferasa III, IV, V y VI.
- 45 10. La *P. tricornutum* transformada de acuerdo con la reivindicación 9, en donde dicha *P. tricornutum* comprende además otra secuencia de ácidos nucleicos unida operativamente a un promotor, codificando dicha secuencia de ácidos nucleicos al menos una enzima glicosiltransferasa seleccionada del grupo que comprende galactosiltransferasas, fucosiltransferasas y sialiltransferasas.
11. Un método para producir un polipéptido glicosilado, comprendiendo dicho método las etapas de:
 - (i) cultivar una *P. tricornutum* transformada como se ha definido en una cualquiera de las reivindicaciones 6 a 10, y
 - (ii) purificar dicho polipéptido que se expresa y glicosila en dicha *P. tricornutum* transformada.

12. El método de acuerdo con la reivindicación 11, en donde dicho método comprende una etapa (iii) para determinar el patrón de glicosilación de dicho polipéptido.
13. Un uso de una *P. tricornutum* transformada como se define en una cualquiera de las reivindicaciones 6 a 10, para producir un polipéptido glicosilado.

Á

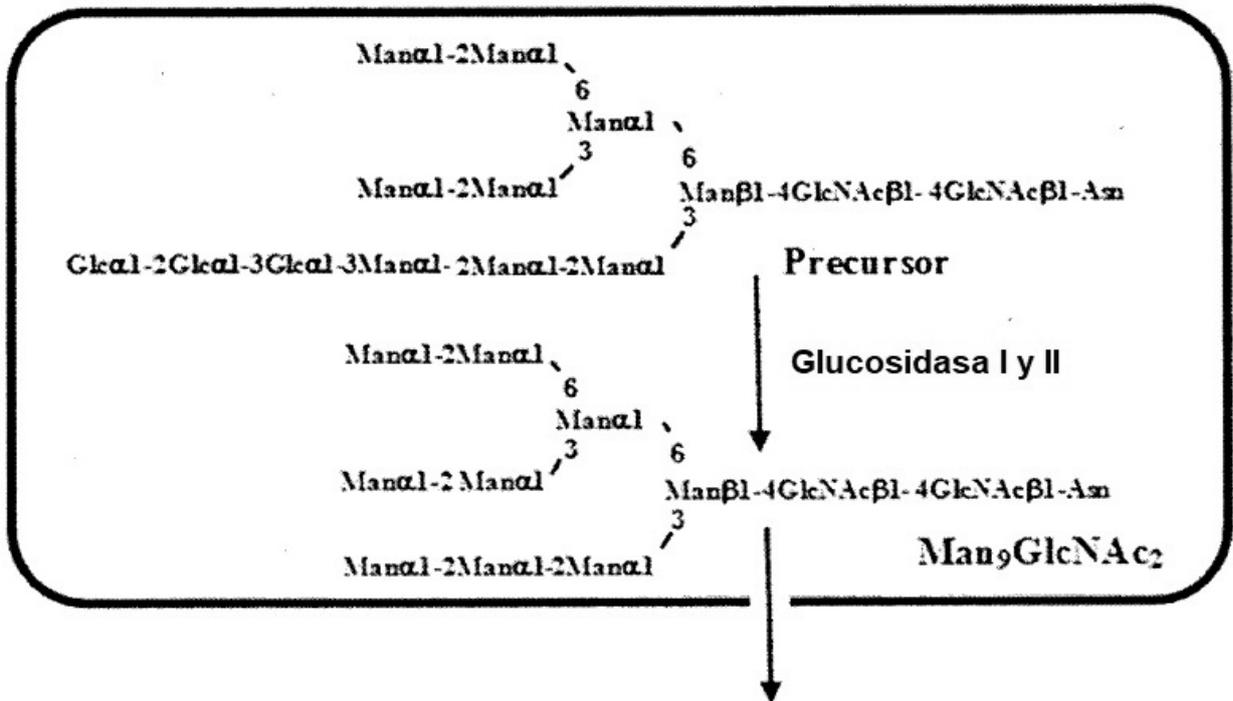


Figura 1 (principio)

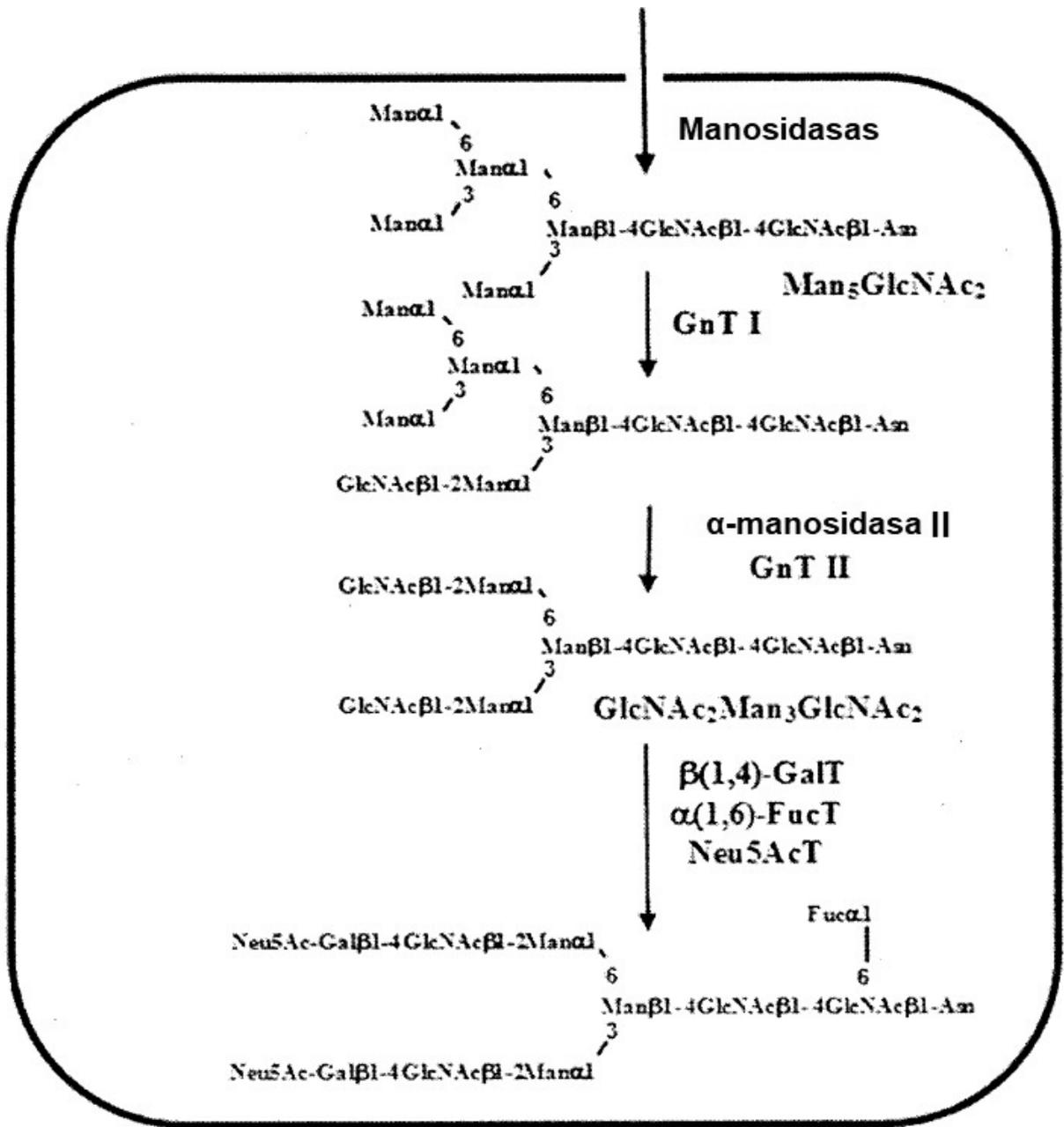


Figura 1 (final)

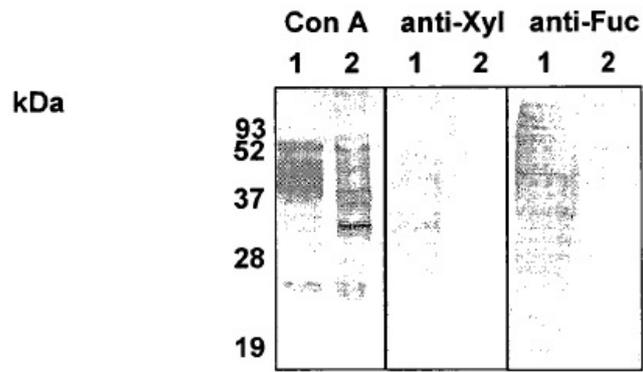


Figura 2

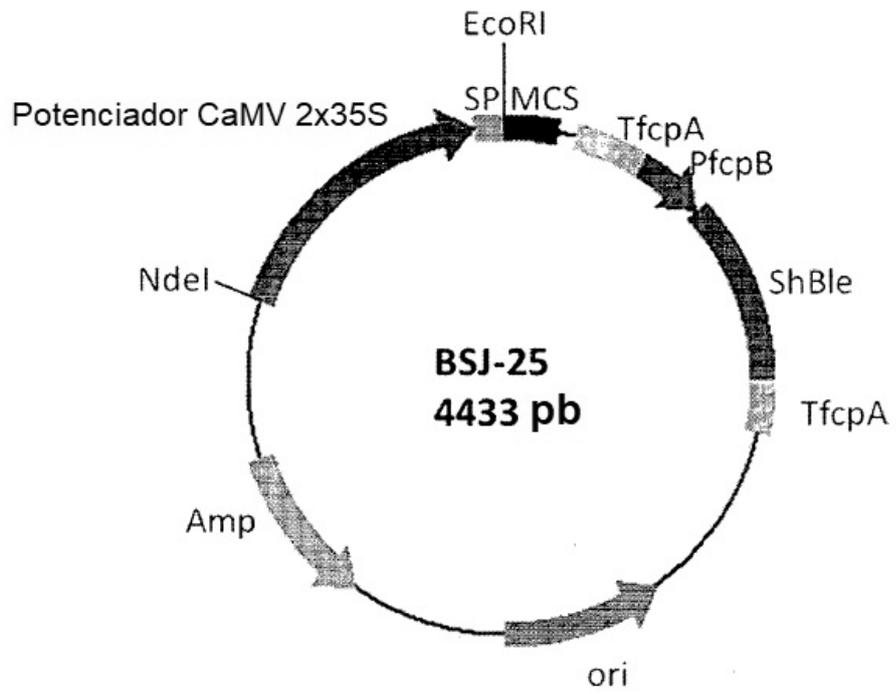


Figura 3

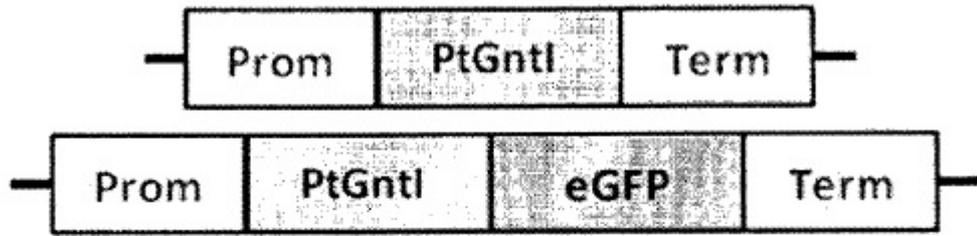


Figura 4

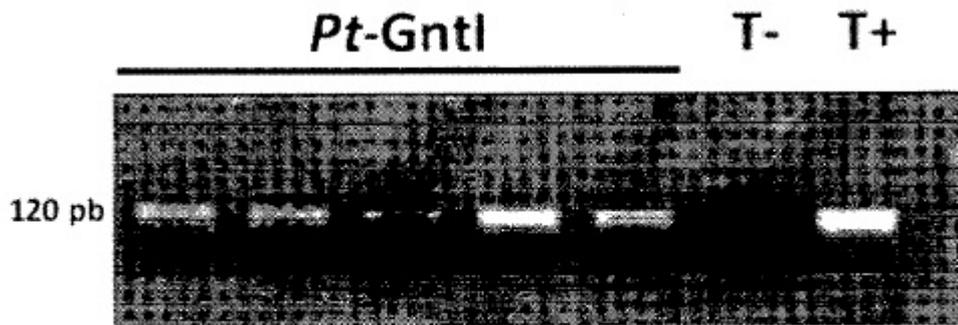


Figura 5

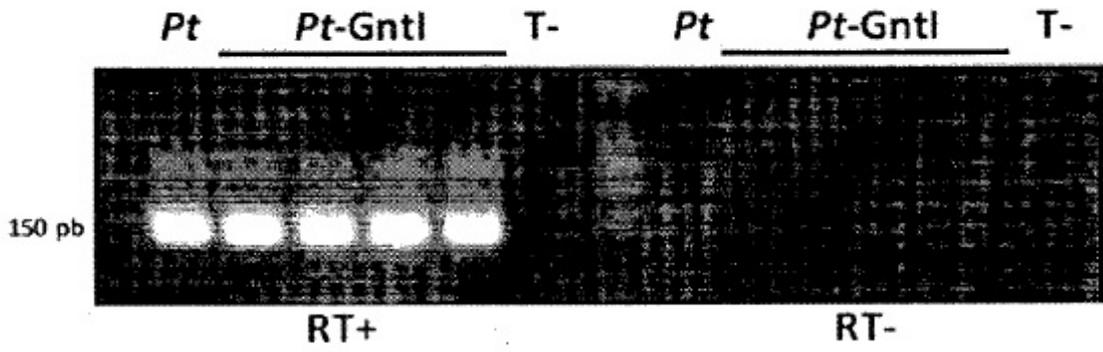


Figura 6

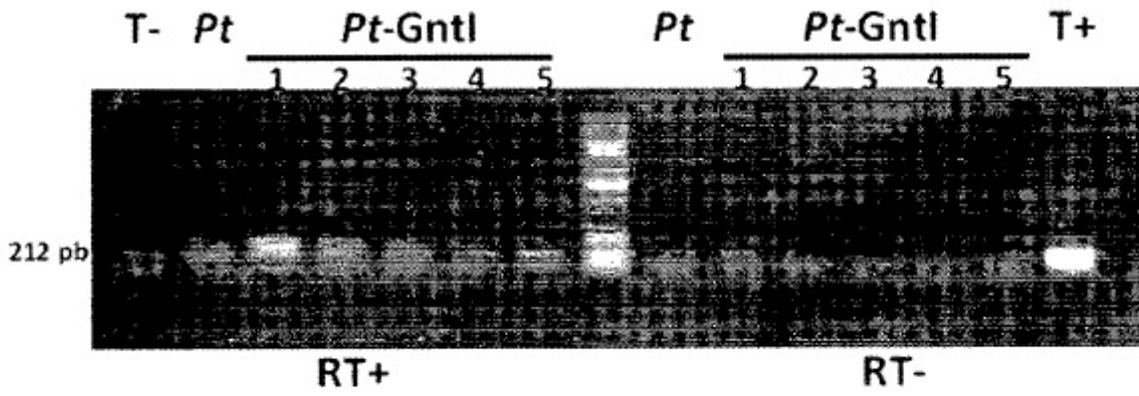


Figura 7

Curva de fusión

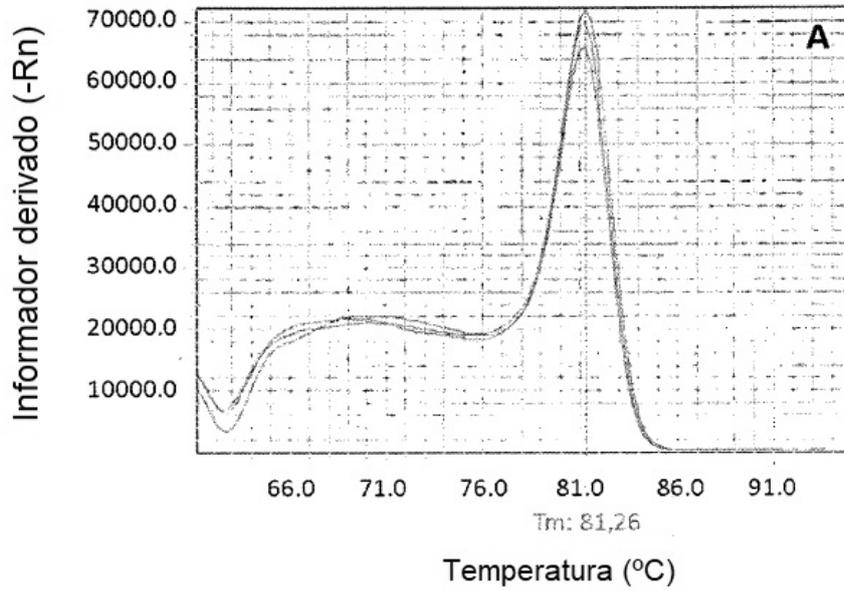


Figura 8A

Gráfico de amplificación

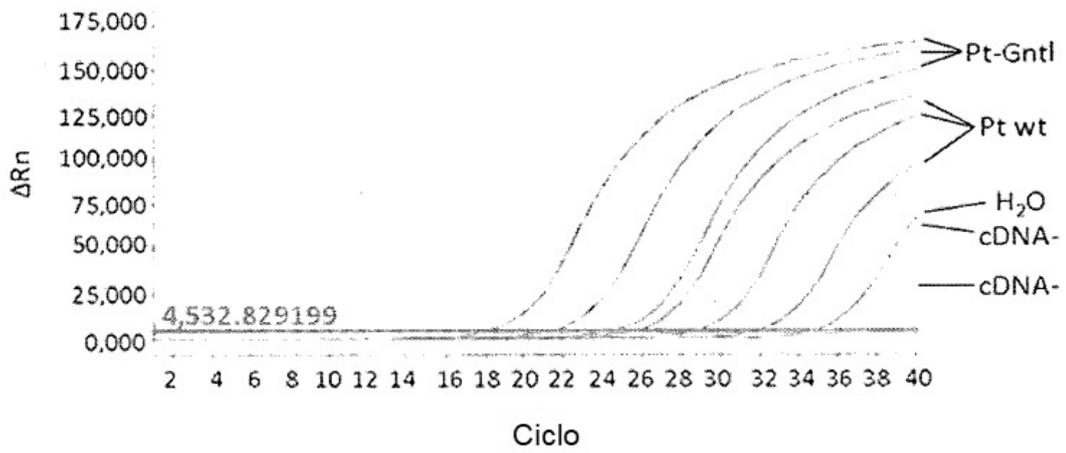


Figura 8B

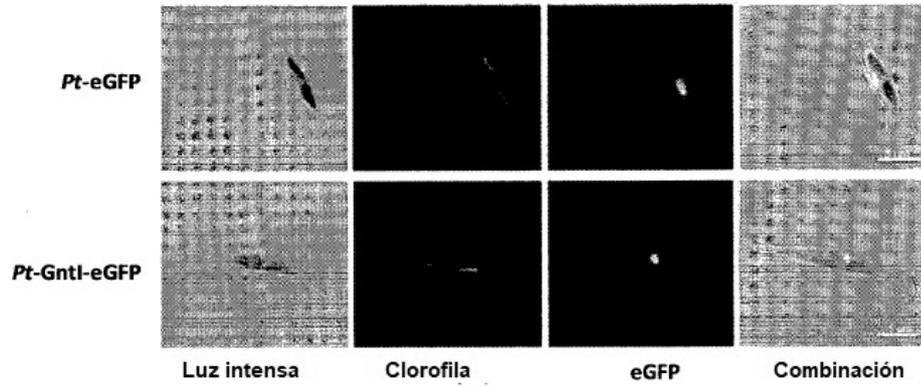


Figura 9

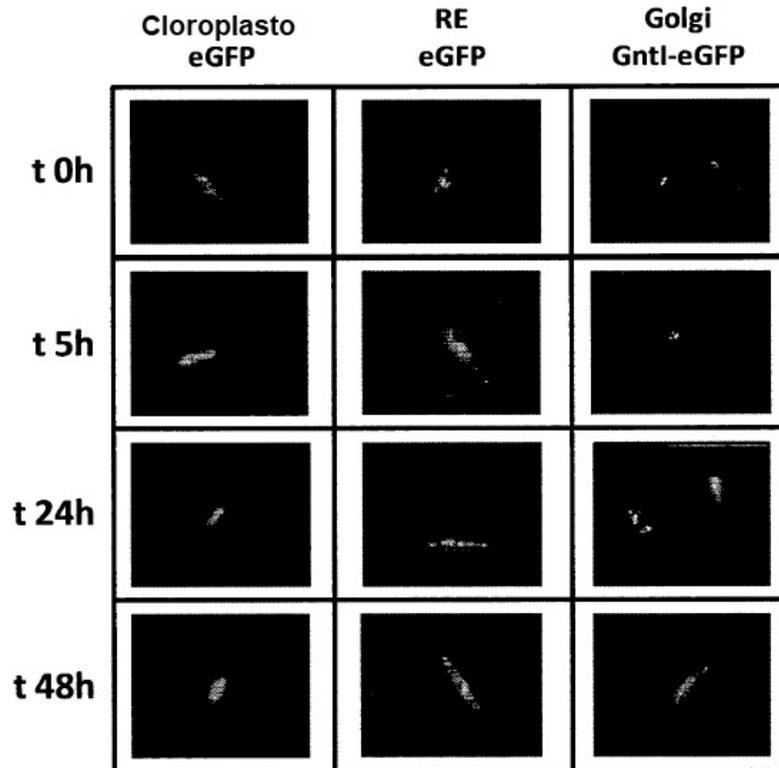
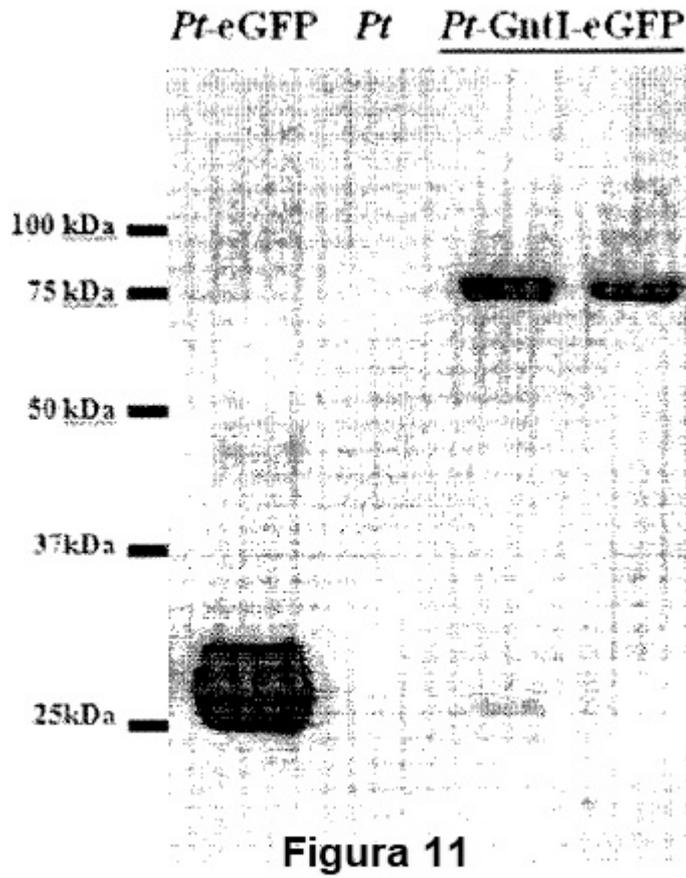


Figura 10



Secuencia de GnT1 de Arabidopsis expresada en *Phaeodactylum tricornutum*

2x35S	GnT1	eGFP	Terminador
2x35S	GnT1	Terminador	

Secuencia de GnT1 humana expresada en *Phaeodactylum tricornutum*

2x35S	GnT1	eGFP	Terminador
2x35S	GnT1	Terminador	
<i>pfcpA</i>	GnT1	eGFP	Terminador

Secuencia de GnT1 endógena expresada en *Phaeodactylum tricornutum*

2x35S	GnT1	eGFP	Terminador	
2x35S	GnT1	Terminador		
2x35S	PS	GnT1	Terminador	
2x35S	PS	GnT1	eGFP	Terminador
<i>pfcpA</i>	GnT1	eGFP	Terminador	
<i>pfcpA</i>	GnT1	Terminador		

Figura 12

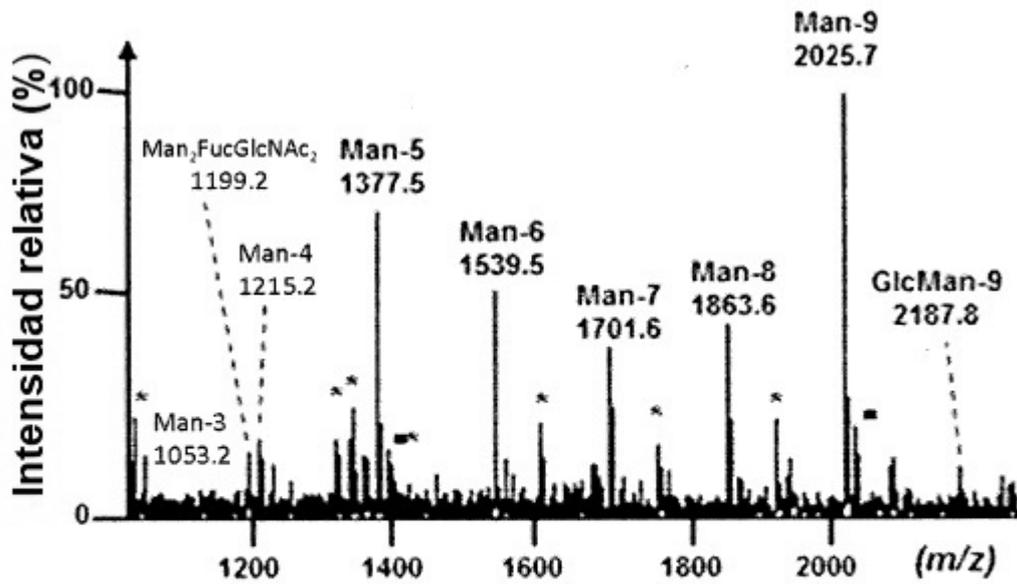


Figura 13A

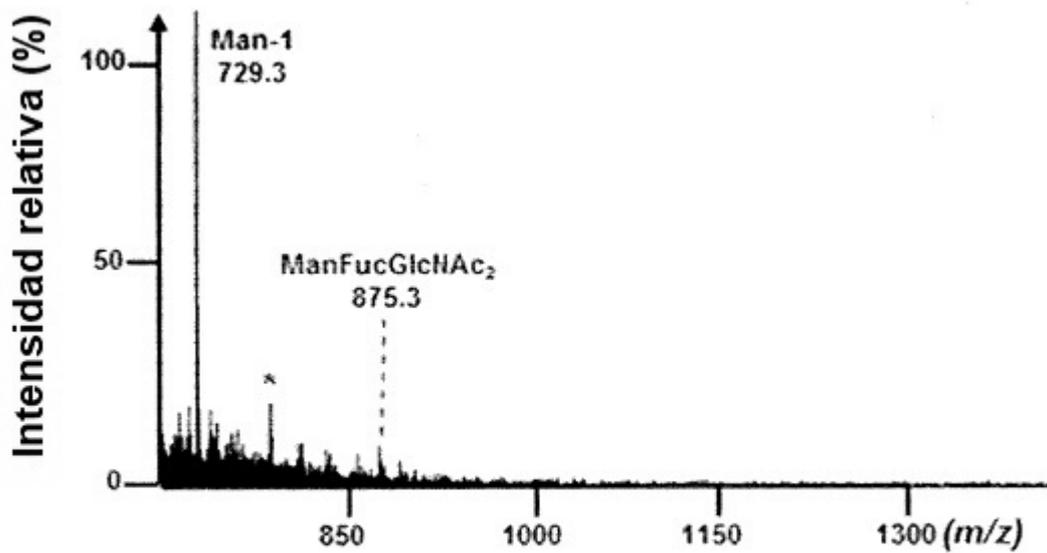


Figura 13B

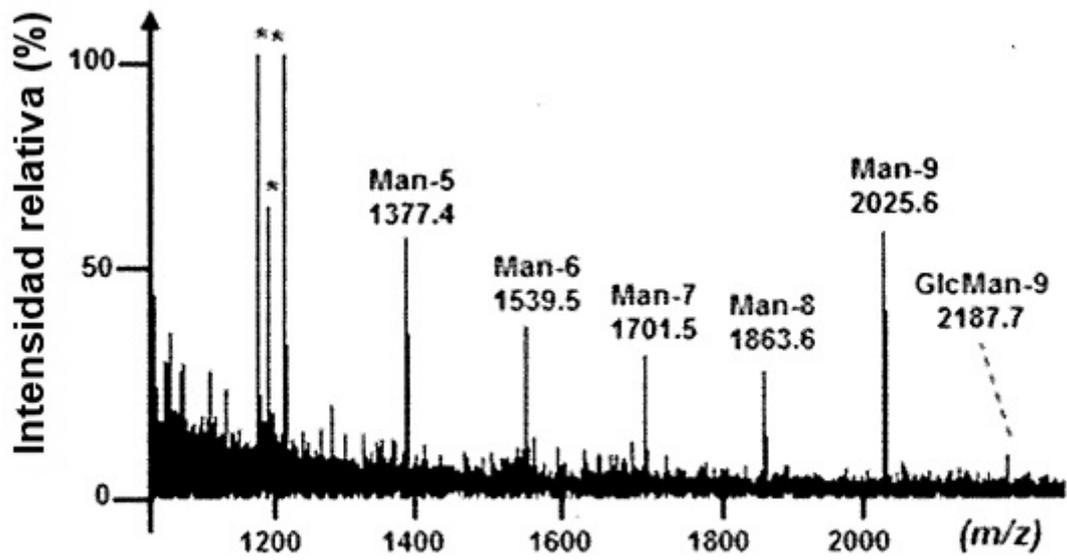


Figura 13C

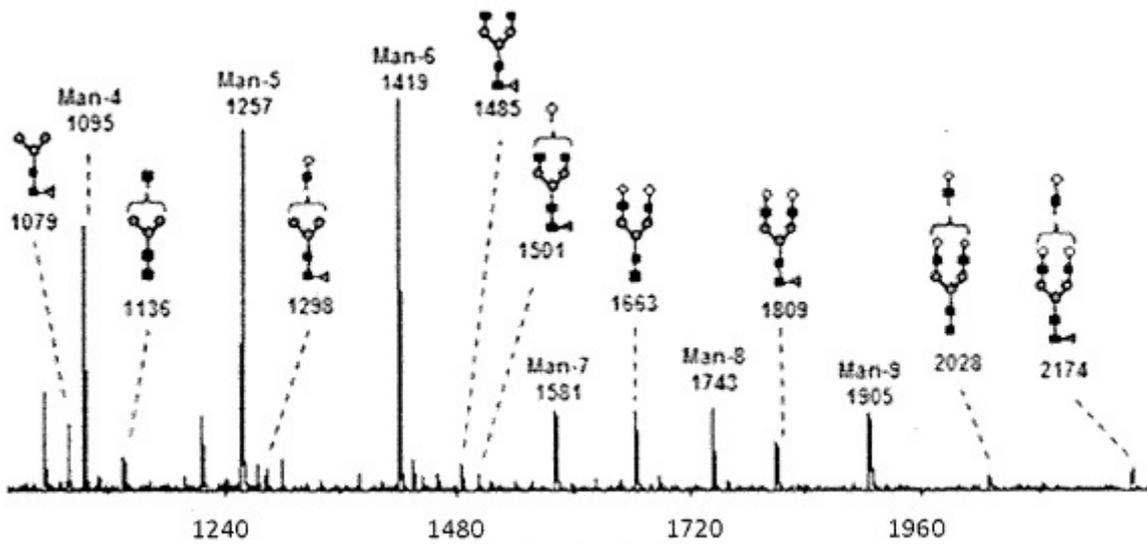


Figura 14A

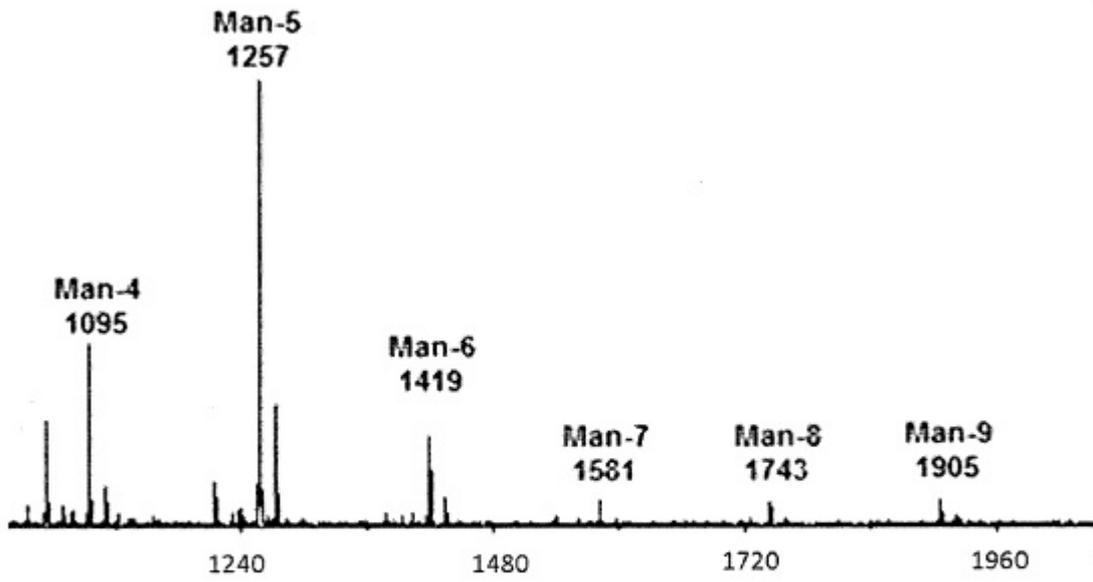
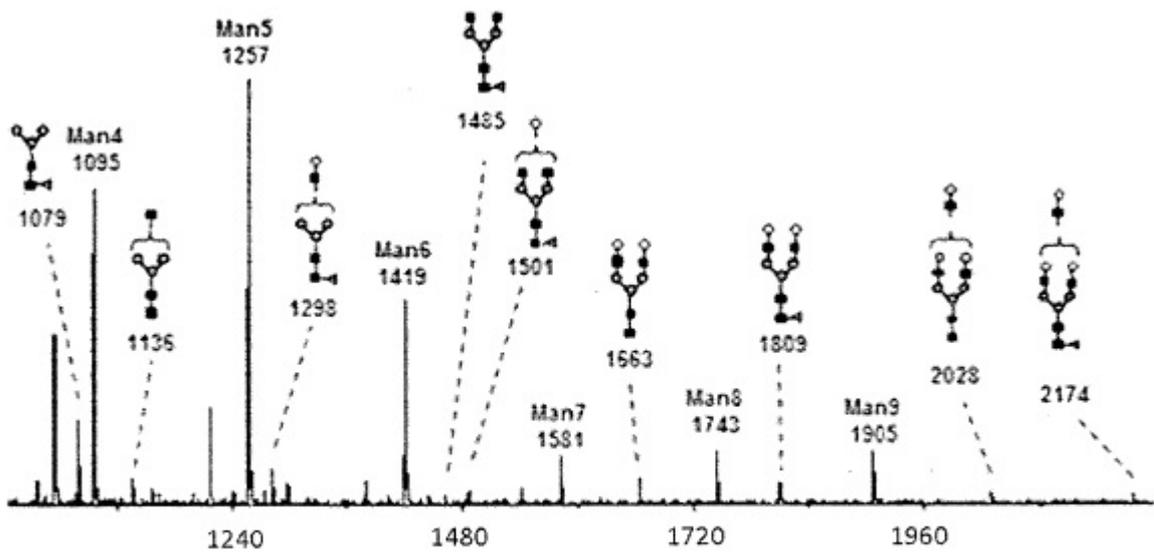


Figura 14B



Masa (m/z)

Figura 14C