

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 368 267**

51 Int. Cl.:

C12N 9/10 (2006.01)

C12N 15/54 (2006.01)

C12N 15/62 (2006.01)

C12P 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07729499 .9**

96 Fecha de presentación: **24.05.2007**

97 Número de publicación de la solicitud: **2019864**

97 Fecha de publicación de la solicitud: **04.02.2009**

54 Título: **PREPARACIÓN Y UTILIZACIONES DE LAS SECUENCIAS GÉNICAS CODIFICANTES DE LAS GLUCOSILTRANSFERASAS QUIMÉRICAS CON ACTIVIDAD DE GLUCOSILACIÓN OPTIMIZADA.**

30 Prioridad:
24.05.2006 EP 06290843

45 Fecha de publicación de la mención BOPI:
15.11.2011

45 Fecha de la publicación del folleto de la patente:
15.11.2011

73 Titular/es:
**UNIVERSITE DE PROVENCE (AIX-MARSEILLE 1)
3, PLACE VICTOR HUGO
13331 MARSEILLE CEDEX 3, FR y
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE**

72 Inventor/es:
**RONIN, Catherine y
GUIRAUDIE-CAPRAZ, Gaëlle**

74 Agente: **Curell Aguila, Marcelino**

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).

ES 2 368 267 T3

DESCRIPCIÓN

Preparación y utilizaciones de las secuencias génicas codificantes de las glucosiltransferasas quiméricas con actividad de glucosilación optimizada.

La presente invención se refiere a la producción de secuencias génicas codificantes de las glucosiltransferasas quiméricas que presentan una actividad de glucosilación optimizada, y a sus utilizaciones en el contexto de la preparación de proteínas recombinantes de interés por parte de células transformadas con dichas secuencias y a secuencias codificantes de dichas proteínas recombinantes.

Glucosilación de proteínas y lípidos

Las proteínas y lípidos son componentes importantes de las membranas celulares. Los carbohidratos asociados a membrana se encuentran exclusivamente en forma de oligosacáridos unidos covalentemente a proteínas que forman glucoproteínas, y a lípidos que forman glucolípidos. Los glucoconjugados (incluyendo los glucolípidos y glucoproteínas) con frecuencia son moléculas de superficie celular claves que se considera que se encuentran implicadas en interacciones célula-célula y en la adhesión celular (Feizi, 1993, Crocker y Feizi, 1996).

Los monosacáridos predominantes presentes en las glucoproteínas eucarióticas son la glucosa (Glc), la galactosa (Gal), la manosa (Man), la fucosa (Fuc), la N-acetilgalactosamina (GalNAc), la N-acetilglucosamina (GlcNAc) y el ácido siálico (Sia), con frecuencia en forma de ácido neuramínico (NeuAc), que puede encontrarse N-acetilado o N-glucosilado en mamíferos, pero únicamente N-acetilado en seres humanos. Las cadenas de carbohidrato también denominadas glicanos se encuentran unidas al esqueleto polipeptídico mediante enlaces O-glucosídicos o N-glucosídicos. El enlace N-glucosídico se establece a través del grupo amida de la asparagina. El enlace O-glucosídico une el hidroxilo de la serina, treonina o hidroxilisina. En las glucoproteínas O-ligadas de tipo ser o de tipo thr, el monosacárido unido directamente a la proteína frecuentemente es GalNAc, mientras que en las glucoproteínas N-ligadas, únicamente se encuentra GlcNAc.

El enlace N-glucosídico se encuentra conservado en todo el reino eucariótico, incluyendo en levaduras, plantas, insectos, mamíferos y seres humanos. El sitio de glucosilación N-ligada se produce dentro de una secuencia de consenso de aminoácidos, Asn-X-Ser/Thr (N-X-S(T)), en la que X es cualquier aminoácido excepto prolina. En el caso de que se lleve a cabo un análisis de proteínas en las bases de datos públicas, puede demostrarse que aproximadamente 65% de todas las proteínas contienen por lo menos una de dichas secuencias de consenso. Todas las glucoproteínas N-ligadas contienen un pentasacárido común (invariante) unido al polipéptido. Este núcleo consta de tres residuos de Man y dos de GlcNAc. Unas antenas de secuencia variable completan el glicano y permiten la clasificación en tres subclases N-ligadas principales:

1. glicanos ricos en manosa que contienen únicamente manosa como azúcares terminales.
2. Glicanos híbridos que contienen por lo menos una antena GlcNAc-Gal.
3. Glicanos complejos que contienen entre 2 y 5 de dichas antenas terminadas en GlcNAc, Gal o ácido siálico.

En todas las células eucarióticas, las glucoproteínas N-ligadas son sintetizadas cotraduccionalmente a partir de polirribosomas unidos al retículo endoplasmático (RE). El tratamiento de los glicanos N-ligados se produce en etapas tempranas en la luz del RE y continua en el aparato de Golgi y la red trans-Golgi, en la que las glucoproteínas alcanzan su polimorfismo final y funcional. La unión de glicanos O-ligados se produce postraduccionalmente en el aparato de Golgi, en el que también se produce la glucosilación de los lípidos. Los azúcares utilizados para tanto la N-glucosilación como la O-glucosilación resultan activados mediante acoplamiento de nucleótidos específicos presentes en el citoplasma y se importan hacia el interior de los orgánulos mediante transportadores específicos.

La glucosilación es el conjunto más sofisticado de modificaciones postraduccionales que pueden producirse en las proteínas. Está destinada a: i) controlar la actividad biológica de las proteínas, ii) señalar las proteínas para la unión de receptores de tipo lectina y/o de sistemas de degradación, iii) enviar proteínas a la superficie celular (secreción), iv) dirigir proteínas a compartimentos celulares, v) definir una identidad inmunológica (grupos sanguíneos). En consecuencia, las proteínas glucosiladas existen en forma de una mezcla de glucoformas cuyas propiedades fisicoquímicas y biológicas difieren del producto directamente codificado por el gen relevante. Se admite ampliamente que las modificaciones postraduccionales de las proteínas permiten una mejor adaptación de la proteína a su función biológica (Helenius y Aebi, 2001).

Muchas glucoproteínas humanas son de elevada relevancia clínica. Por ejemplo, sobre las superficies celulares resultan importantes para la comunicación entre las células, para mantener la estructura celular y para el autorreconocimiento por parte del sistema inmunológico.

Las glucoproteínas son más abundantes en forma soluble en líquidos biológicos tales como la leche y la sangre. En este caso, los glicanos protegen a las proteínas frente a proteasas, incrementan su solubilidad y controlan su eliminación plasmática y los dirigen a órganos diana.

5 Enzimas de glucosilación

Las reacciones de glucosilación son de gran importancia biológica para tanto procariotas como eucariotas y requieren la acción coordinada de un gran número de enzimas denominados glucosiltransferasas (GT) (Breton y Imberty, 1999).

Las glucosiltransferasas (GT) constituyen una gran familia de enzimas implicados en la biosíntesis de polisacáridos y glucoconjugados en los reinos procariótico y eucariótico, respectivamente. Las secuencias de los enzimas procarióticos no son homólogas a las de las glucosiltransferasas de mamífero, mientras que los enzimas de seres humanos, mamíferos y de *Drosophila* con frecuencia comparten homología significativa. Hasta hoy, se conocen más de mil GT que median en un amplio abanico de funciones biológicas. El desarrollo de la biología molecular de estos enzimas ha revelado una inesperada diversidad, sugiriendo que el proceso de glucosilación probablemente requiere la participación de aproximadamente 250 genes en una sola célula humana viva (Breton *et al.*, 2001). Hasta hoy se han clonado aproximadamente 500 glucosiltransferasas, incluyendo 160 enzimas humanas. Está creciendo el número de enzimas clonados de origen humano y podría alcanzar los 200 en poco tiempo (Narimatsu, 2004). Las GT se clasifican según la estereoquímica de los sustratos y productos de reacción como enzimas conservadores o inversores (Sinnott, 1990). Se ha establecido una clasificación de las glucosiltransferasas utilizando nucleótido disfosfo-azúcar, nucleótido monofosfo-azúcares y azúcares fosfato y proteínas relacionadas en familias diferentes basadas en la secuencia. Demuestra que las GT humanas se reparten hasta el momento en 42 familias estructurales (<http://afmb.cnrs-mrs.fr/CAZY/Homo-sapiens.html>).

Las glucosiltransferasas catalizan la transferencia de residuos sacáridos de un nucleótido-azúcar (sustrato donador activado) a un aceptor (lípidos, proteína o cadena sacárida en crecimiento). Además de la clasificación oficial de los enzimas, pueden agruparse en familias funcionales basándose en sus similitudes de secuencia, lo que podría reflejar características enzimáticas tales como la especificidad de donador, la especificidad de aceptor y la especificidad de unión entre donador y aceptor. Basándose en el azúcar que transfieren, las GT se denominan N-acetilglucosaminiltransferasa (GlcNAcT), galactosiltransferasa (GalT), fucosiltransferasa (FucT) o sialiltransferasa (SiaT o ST). Debido a la acumulación de datos genéticos de glucosiltransferasas, se cree que éstas presentan una elevada especificidad para el tipo de enlace en los sustratos donador y aceptor, así como para la naturaleza del aceptor glucoconjugado (glucoproteínas o lípidos N-ligados/O-ligados).

Estructura de las glucosiltransferasas:

Todas las GT clonadas, hasta hoy en vertebrados, muestran la misma topología. Son proteínas membranales de tipo II (dominio citoplasmático N-terminal) compuestas de cuatro dominios principales: una cola citoplasmática corta (CT) en el extremo N-terminal, una región de anclaje membranal (TMD) de 10 a 20 aminoácidos, una región de tallo (SR) y un gran dominio catalítico C-terminal (CD) (figura 1) (Paulson *et al.*, 1987). La región de anclaje actúa como péptido de señal no escindible y también como dominio transmembranal (Wickner y Lodisch, 1985), orientando el dominio catalítico de las GT en la parte luminal del aparato de Golgi. La SR se considera ampliamente una región flexible que también permite la orientación del CD en la parte luminal del aparato de Golgi. De modo general, el CD se encuentra razonablemente conservado dentro de las GT de diversas especies, mientras que la SR constituye una parte hipervariable de la transferasa. Algunos de estos enzimas pueden ser cortados en su SR por una o más proteasas endógenas, y secretarse al exterior de la célula (Paulson y Colley, 1989), produciendo enzimas de la leche o la sangre. Está bien documentado que el corte proteolítico y la secreción de las glucosiltransferasas resultan afectadas por diversas condiciones patológicas tales como la transformación maligna y la inflamación, aunque los mecanismos moleculares subyacentes al corte y secreción no han sido clarificados todavía.

Las GT y las proteínas/lípidos neoglucosilados han sido localizados en el interior del ER y de los subcompartimentos del Golgi utilizando tanto el fraccionamiento subcelular de las membranas celulares como la microscopía inmunoelectrónica (Roth, 1987). Los primeros estudios sugerían que la compartimentalización de las GT podría reflejar la secuencia de modificación de la cadena oligosacárida (Kornfeld y Kornfeld, 1985). Se creía que esta localización estricta garantizaba una biosíntesis óptima de las cadenas de glicano al proporcionar una eficiente vecindad entre enzima, sustrato y donador de azúcar-nucleótido. Algunos estudios posteriores han demostrado que muchos enzimas de glucosilación se solapan en su localización, y han demostrado una subcompartimentalización en el Golgi específica de tipo celular (Colley, 1997). Las GT se encuentran extendidas en los dictiosomas del Golgi y esto puede diferir entre tipos celulares para una proteína dada (Roth, 1991).

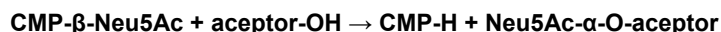
De manera general, se ha demostrado que el dominio transmembranal (TMD) de las proteínas es un determinante clave para la localización en el Golgi, esencial por lo menos para que las GalNAcT, GlcNAcT, GalT, FucT y SiaT encuentren sus aceptores. Se ha demostrado que la región flanqueante de TMD y/o la parte luminal, también contribuye a la localización (Munro, 1998; Yang *et al.*, 1996). Además, la CT y la SR de las GT especifican su sublocalización y estabilidad funcionales *in vivo* en el aparato de Golgi. La sustitución de cualquiera de los tres

dominios (CT, TMD y SR) no modifica la actividad catalítica del enzima, pero contribuye a alterar su distribución en los compartimentos del Golgi (Grabenhorst y Conradt, 1999). No se han encontrado secuencias específicas de direccionamiento claras durante la última década y sólo se han identificado regiones de las GT críticas para su compartimentalización (Opat *et al.*, 2001). Más recientemente, los inventores han encontrado que el dominio catalítico soluble de ST6Gal seguía una ruta subcelular diferente que el enzima de longitud completa (Ronin, Biochimie, 2003).

Entre el gran número de GT, las familias que son de mayor interés son aquellas responsables de la glucosilación terminal debido a que estos azúcares desempeñan un papel importante en fenómenos de reconocimiento y señalización en seres humanos. Entre ellos se incluyen esencialmente las sialiltransferasas (SiaTs), las fucosiltransferasas (FucT) y las galactosiltransferasas (GalT). La galactosa y la fucosa se encuentran implicadas en el reconocimiento de los antígenos de grupo sanguíneo (ABH/Lewis). Los oligosacáridos sialilados de las glucoproteínas y glucolípidos se encuentran implicados en muchos procesos biológicos tales como la adhesión celular y el reconocimiento de receptores en la inflamación y el cáncer (antígeno sialil-Lewis), así como en el crecimiento neuronal (antígeno polisialico) (Paulson, 1989). La diversidad estructural y la expresión regulada de los sialilglucoconjugados aparentemente se encuentran bien correlacionadas con las funciones de los mismos (Sasaki, 1996).

Sialiltransferasas:

La familia del ácido siálico está compuesta de más de cien derivados entre los que el ácido neuramínico es el más frecuente en mamíferos y seres humanos. Las sialiltransferasas (ST) catalizan la transferencia de un residuo de ácido siálico de su forma activada, el ácido citidil-monofosfo-N-acetil-neuramínico (CMP-Neu5-Ac), a una posición terminal no reductora en un aceptor glicano en glicoproteínas o en glicolípidos. La reacción catalizada es la siguiente:



Cada ST se clasifica según el tipo de enlace establecido entre el residuo de ácido siálico y la especificidad del sustrato aceptor (que puede ser una proteína o un lípido). De esta manera, se distinguen tres grupos: las α 2,3-sialiltransferasas (α 2,3-ST), las α 2,6-sialiltransferasas (α 2,6-ST) y las α 2,8-sialiltransferasas (α 2,8-ST). Las α 2,6-ST transfieren un residuo de ácido siálico en una posición alfa2,6 (ST6Gal) a un residuo de galactosa, o a una N-acetil-D-galactosamina (ST6GalNAc) o a una N-acetil-D-glucosamina (ST6GlcNAc). Sin embargo, todavía se desconoce cuál es el enzima implicado en la formación de este último tipo de enlace. Las α 2,3-ST transfieren un ácido siálico al carbono 3 de la galactosa (ST3Gal) y las α 2,8-ST, al carbono 8 de otro residuo de ácido siálico (ST8Sia).

Las ST también han sido identificadas en todos los animales (de aves a seres humanos), en células bacterianas y en virus (Sujino *et al.*, 2000). La familia de las ST humanas está compuesta de 20 genes diferentes clonados tal como se muestra en la Tabla 1 (Sasaki, 1996; Ronin, 2003). Sus especificidades de sustrato son las siguientes.

1) α 2,6-ST

• ST6Gal:

La ST6Gal I transfiere un residuo de ácido siálico al disacárido Gal β 1-4GlcNAc de los N-glicanos (van den Eijnden *et al.*, 1980; Weinstein *et al.*, 1982). Se expresa de modo ubicuo en los tejidos humanos, excepto en los testículos y en el cerebro, en donde se expresa a niveles más bajos (Tabla 2) (Kitagawa y Paulson, 1994a).

La ST6Gal II, identificada recientemente (Krzewinski-Recchi *et al.*, 2003; Takashima *et al.*, 2002), reconoce el disacárido Gal β 1-4GlcNAc como sustrato aceptor, particularmente en el caso de que se encuentre en un oligosacárido libre (proteína o lípido desconocido). El patrón de expresión de ST6Gal II se restringe al cerebro y muestra un nivel de expresión reducido en testículos, tiroides, ganglios linfáticos y en algunos tejidos fetales (Tabla 2).

• ST6GalNAc:

La segunda subfamilia de α 2,6-ST está representada por el grupo de ST6GalNAc, que transfiere un residuo de ácido siálico en un residuo de N-acetil-galactosamina. Se han identificado seis miembros en el ser humano (Tabla 2).

La ST6GalNAc I y la ST6GalNAc II presentan la especificidad de sustrato más amplia, debido a que pueden transferir un ácido siálico a las estructuras de O-glicano siguientes: Gal β 1-3GalNAc, GalNAc α -O-Ser/Thr, Sia α 2-3Gal β 1-3GalNAc α -O-Ser/Thr (Ikehara *et al.*, 1999; Kono *et al.*, 2000; Kurosawa *et al.*, 1994; Kurosawa *et al.*, 1996; Harduin-Lepers *et al.*, 2001). Su patrón de expresión es diferente: ST6GalNAc I se expresa en glándulas submaxilares y mamas, bazo y colon, mientras que ST6GalNAc II se expresa en muchos tejidos, tales como el testículo y las glándulas mamas lactantes (Kurosawa *et al.*, 1996; Kurosawa *et al.*, 2000).

Las ST6GalNAc III, IV, V y VI presentan una especificidad de sustrato reducida: reconocen únicamente el trisacárido Sia α 2,3Gal β 1-3GalNAc. Sin embargo, se ha establecido que ST6GalNAc III, V y VI catalizan preferentemente la formación del glucolípido GM1b (Sjoberg *et al.*, 1996; Lee *et al.*, 1999), mientras que ST6GalNAc IV cataliza la transferencia de ácido siálico a O-glicanos (Harduin-Lepers *et al.*, 2000; Ikehara *et al.*, 1999; Lee *et al.*, 1999; Okajima *et al.*, 2000). ST6GalNAc IV muestra una especificidad de sustrato restringida, al utilizar únicamente la secuencia de trisacárido Sia α 2,3Gal β 1-3GalNAc y no discriminar entre el GalNAc α -ligado y β -ligado. Se han descubierto dos isoformas diferentes. El análisis de transferencia northern detectó un transcrito de 2,2 kb en diversos tejidos adultos y niveles de expresión más bajos de un transcrito adicional en cerebro, corazón y músculo esquelético (Harduin-Lepers *et al.*, 2000).

2) α 2,3-ST

Todas son ST3Gal que transfieren ácido siálico de un residuo galactosa en un aceptor proteína o lípido.

ST3Gal I fue clonado a partir de una biblioteca de ADNc procedente de glándulas submaxilares y su expresión ubicua (Tabla 2) fue confirmada mediante transferencia northern (Chang *et al.*, 1995). Sintetiza Sia α 2,3Gal β 1-3GalNAc, una estructura común a muchos oligosacáridos O-ligados. Difiere de otras ST en su capacidad de utilizar sustratos aceptores glucolípido *in vitro* (Kitagawa y Paulson, 1994b).

ST3Gal II fue clonada por Kim *et al.* (1996) a partir de una biblioteca de ADNc procedente de hígado y el análisis de transferencia northern reveló niveles de expresión elevados en corazón, hígado y músculo esquelético, niveles intermedios en timo, nódulo linfático, apéndice y bazo, y niveles más bajos en linfocitos de sangre periférica pulmonar (Tabla 2). ST3Gal II puede transferir residuos de ácido siálico en Gal β 1-3-GalNAc o en los gangliósidos (lípidos) GM1 ó asialo-GM1 como sustratos aceptores (Giordanengo *et al.*, 1997).

ST3Gal III ha sido clonada mediante cribado de una biblioteca de ADNc procedente de placenta humana utilizando una sonda basada en un motivo sialilo, aislando el ADNc codificante de ST3Gal III (Kitagawa y Paulson, 1993). El transcrito se expresa abundantemente en el músculo esquelético y en tejido fetal y a un nivel de expresión más bajo en la placenta. Este enzima cataliza la transferencia de ácido siálico a los sustratos que contienen galactosa (Tabla 2).

ST3Gal IV se encuentra implicada en la sialilación de Gal β 1-3GalNAc O-ligada (Tabla 2) (Tetteroo *et al.*, 1987). Se expresan 5 ARNm diferentes, y codifican secuencias proteicas idénticas excepto en los extremos 5'. Estos transcritos se producen mediante una combinación de tratamiento alternativo. El análisis de transferencia *northern* ha demostrado que uno de ellos se expresa específicamente en placenta, testículos y ovarios, indicando que su expresión se encuentra regulada independientemente de la de los demás (Kitagawa *et al.*, 1996).

ST3Gal V ha sido aislada a partir de una biblioteca de ADNc procedente de células humanas y se ha denominado GM3 sintasa. Se expresa un transcrito importante de 2,4 kb en muchos tejidos, particularmente en cerebro, músculo esquelético, placenta y testículos. Se encuentra ampliamente distribuida en el cerebro humano, con un nivel de expresión ligeramente elevado en el córtex cerebral, lóbulo temporal y putamen. La especificidad de sustrato de esta enzima se encuentra muy restringida a la lacosilceramida como aceptor (Tabla 2) (Ishii *et al.*, 1998).

ST3Gal VI ha sido clonada a partir de una biblioteca de ADNc de melanoma humano. Esta ST muestra una especificidad de sustrato restringida: participa en la síntesis del sialil-paraglobósido, un precursor del determinante sialilo-Lewis X (Okajima *et al.*, 1999). Existen 2 formas de ARNm de la ST3Gal VI (denominadas tipo 1 y tipo 2), que difieren únicamente en la región 5' no traducida. Este enzima se expresa a niveles similares en la mayoría de tejidos (Tabla 2) (Taniguchi *et al.*, 2001).

3) α 2,8-ST

ST8Sia I también se denomina gangliósido GD3 sintasa y cataliza la transferencia de una molécula de ácido siálico al ácido siálico terminal de GM3 mediante un enlace α 2,8 (Sasaki *et al.*, 1994). Se ha demostrado que este enzima también puede utilizar GM1b, GD1a y GT1b como sustrato aceptor (Tabla 2) (Nakayama *et al.*, 1996; Nara *et al.*, 1996; Watanabe *et al.*, 1996).

ST8Sia II también se denomina STX. Scheidegger *et al.* (1995) utilizaron secuencias de STX de roedor para clonar el ADNc humano a partir de una biblioteca de corazón fetal. STX se expresa principalmente en tejidos embrionarios y a niveles modestos en corazón, cerebro y timo adultos (Angata *et al.*, 1997). Este enzima regula la formación de enlaces entre la molécula de adhesión de las células neurales (NCAM) y el ácido polisialílico (PSA) y de esta manera modula las propiedades adhesivas de NCAM. STX cataliza la transferencia del primer ácido siálico mediante un enlace α 2,8 a otro ácido siálico unido en enlace α 2,3 ó α 2,6 (Tabla 2) (Angata *et al.*, 1997) en un N-glicano, y después participa en el proceso de elongación, generando sucesivos enlaces α 2,8 con los residuos anteriores.

ST8Sia III transfiere un ácido siálico en las estructuras de Sia α 2-3Gal β 1-4GlcNAc dentro de N-glicanos y glucolípidos tales como GM3 (Tabla 2) (Lee *et al.*, 1998; Yoshida *et al.*, 1995a; Yoshida *et al.*, 1995b).

ST8Sia IV, también denominada PST (polisialiltransferasa) ha sido clonada por Nakayama *et al.* (1995). El análisis de transferencia northern ha revelado que PST se expresa en muchos tejidos fetales y en corazón, bazo y timo adultos, y a niveles más bajos en otros órganos y tejidos (Tabla 2). Este enzima también regula la formación de enlace entre la molécula de adhesión de las células neurales (NCAM) y el ácido polisialílico (PSA). Además, se ha demostrado que cataliza la misma reacción que STX (Angata *et al.*, 2000; Nakayama *et al.*, 1995).

ST8Sia V presenta una actividad de transferencia a los gangliósidos GM1b, GD1a, GT1b y GD3 (Tabla 2) (Kono *et al.*, 1996).

ST8Sia VI ha sido clonada bastante recientemente en el ser humano y se conoce poco de ella. La ST8Sia VI de ratón presenta una actividad de transferencia de ácido siálico a la estructura NeuAc_{2,3}(6)Gal₃ presente en el extremo no reducido de O-glicanos, N-glicanos y oligosacáridos libres tales como la sialil-lactosa (Tabla 2) (Takashima *et al.*, 2002).

4) Organización estructural necesaria para estimular la actividad de sialiltransferasa

Todos los miembros de la familia de las ST presentan tres regiones conservadas en su CD denominadas motivos sialilo: el motivo L de Large (grande), el motivo S de Small (pequeño) y el motivo VS de Very Small, (muy pequeño) (Datta y Paulson, 1995; Geremia *et al.*, 1997) basados en la comparación entre sus secuencias primarias. Recientemente se ha publicado un estudio de los genomas animales (Harduin-Lepers, 2005) que presenta el objetivo de encontrar secuencias desconocidas que presenten motivos sialilo y, de esta manera, potencialmente de encontrar nuevas ST.

El motivo sialilo L consiste de 44 ó 45 aminoácidos y contiene 5 residuos invariantes entre todos los enzimas humanos, tal como se muestra en la figura 2. Esta región se encuentra en el centro del CD. El segundo motivo sialilo S, en la parte COOH-terminal, consiste de 23 residuos aminoácidos, dos de los cuales son idénticos en todas las ST (Drickamer, 1993). El tercer motivo sialilo VS, en la parte terminal del enzima, consiste de 13 aminoácidos con dos residuos conservados (histidina y glutamato).

Se ha demostrado que el motivo sialilo L participa en la unión del sustrato donador CMP-ácido siálico (Datta y Paulson, 1995). Algunos aminoácidos de este motivo también pueden participar en la actividad catalítica de los enzimas (Sasaki, 1996). Se ha propuesto que el motivo sialilo S participa en la unión de tanto donador como aceptor (Datta *et al.*, 1998). El papel exacto del VS todavía no está claro pero algunos estudios recientes sobre STX y PST han estudiado el papel funcional de la His conservada en este motivo (Kitazume-Kawaguchi *et al.*, 2001). El cambio de la histidina por un residuo de lisina afecta a su actividad catalítica, demostrando que el motivo VS participa en la catálisis. El motivo sialilo VS resulta necesario para una eficiencia catalítica óptima y es parte del sitio activo, principalmente en el sitio aceptor o en proximidad a sustratos azúcar tanto donadores como aceptores (Jeanneau *et al.*, 2004). A partir de trabajos anteriores ahora resulta evidente que la parte C-terminal del CD (parte del motivo sialilo S, del motivo 3 y del motivo sialilo VS) se dedica principalmente al reconocimiento de los sustratos aceptores, mientras que la parte N-terminal (motivos sialilo L y parte del motivo S) participan principalmente en la unión del nucleótido-azúcar (Datta y Paulson, 1995; Datta *et al.*, 1998; Laroy *et al.*, 2001; Kitazume-Kawaguchi *et al.*, 2001; Jeanneau *et al.*, 2004). Sin embargo, no puede excluirse que el motivo sialilo F también participe en el reconocimiento de aceptores (Legaigneur *et al.*, 2001).

Tal como se ha indicado anteriormente, las secuencias de aminoácidos deducidas a partir del ADNc de sialiltransferasa humana clonada muestra la misma organización en cuatro dominios compartidos por muchas GT: i) una cola citoplasmática N-terminal corta (CT, de aproximadamente 10 aminoácidos), ii) un dominio transmembranal (TMD, de aproximadamente 20 aminoácidos), iii) una región de tallo (SR), de longitud muy variable, e iv) un dominio catalítico (CD, de aproximadamente 300 aminoácidos).

La SR se define como la región peptídica situada entre el TMD y el CD que puede eliminarse sin alterar la actividad enzimática (Paulson, 1989; Ronin, 2001; Vallejo-Ruiz *et al.*, 2001; Jeanneau *et al.*, 2004). Sin embargo, se ha demostrado que esta región hipervariable puede resultar esencial para definir un reconocimiento fino del aceptor glicano (tri>bi>tetraantenario) y propusieron que podría participar en un cambio conformacional que abriese el sitio catalítico (Ronin, 2001). Como resultado, se incrementó la eficiencia catalítica en un factor de 35 y se amplió el reconocimiento del glicano aceptor.

El CD resulta crucial para que las ST muestren actividad enzimática. Contiene por lo menos tres secuencias altamente conservadas que mantienen posiciones aminoácidas idénticas en todas las ST de mamífero clonadas. Estos dominios son motivos sialilo L, S y VS (Livingston y Paulson, 1993; Geremia *et al.*, 1997). También se ha aislado recientemente un dominio adicional (motivo 3) entre los motivos sialilo S y VS, que contiene cuatro residuos altamente conservados, con la secuencia de consenso siguiente: (H/y)Y(Y/F/W/h)(E/D/q/g) (las letras mayúsculas y minúsculas indican la aparición frecuente o rara del aminoácido, respectivamente) (Jeanneau *et al.*, 2004). Muchos estudios descritos en la literatura presentan el objetivo de comprender las relaciones estructura/función en la gran familia de ST y, de esta manera, definir particularmente un dominio catalítico mínimo dentro de la misma. De esta

manera, el CD mínimo ha sido definido experimentalmente mediante mutagénesis sitio-dirigida o, alternativamente, mediante alineación de secuencias y la comparación de unas cuantas ST (Vallejo-Ruiz *et al.*, 2001; Chen y Colley, 2000; Ronin, 2003).

A pesar de la amplia definición del CD y de la actividad catalítica, todavía resulta difícil delinear un CD mínimo de las ST. La ST6Gal I es el enzima más estudiado de entre las ST y la mayor parte de la investigación se ha centrado en definir tanto un dominio catalítico mínimo como la actividad catalítica misma. Esto se debe a la ausencia de ST6Gal en todas las células utilizadas para producir proteínas recombinantes humanas. Esta carencia es un cuello de botella tecnológico para sistemas heterólogos que utilizan células no animales (levaduras, insectos y plantas), debido a que no contienen nada de ácido siálico. En las células CHO, alternativamente, sólo se expresa una actividad de ST3.

La diferenciación entre el SR y el CD nunca ha sido bien definida, excepto experimentalmente. Basándose en la bioinformática, se ha diseñado una estrategia para identificar el extremo del SR para las 3 familias de ST (Ronin, 2003).

El CD generalmente coincide con el extremo de la región hipervariable en la mitad N-terminal de los enzimas (Ronin, 2003). Por lo tanto, se supone que el dominio catalítico se inicia aproximadamente 70 a 90 residuos cadena arriba del motivo sialilo L y aproximadamente 40 a 45 residuos en el caso de ST6GalNAc III a VI. El tamaño medio del CD se estima en aproximadamente 300 (± 20) aminoácidos, incluyendo los motivos sialilo (Jeanneau *et al.*, 2004) (Tabla 3). El CD mínimo ha sido definido para unas cuantas ST, mediante experimentos de truncado o la comparación entre secuencias (Legaigreur *et al.*, 2001; Vallejo-Ruiz *et al.*, 2001; Chen y Colley, 2000), tal como para hST3Gal I (CD mínimo consistente de los aminoácidos 57 a 350; Vallejo-Ruiz *et al.*, 2001) y hST8Sia IV (CD mínimo que consiste de los aminoácidos 62 a 359; Angata *et al.*, 2004).

Tras la transfección en células CHO, se ha encontrado que el CD soluble de hST6Gal I (aminoácidos 90 a 406) muestra una especificidad incrementada para aceptores endógenos debido a que sigue la ruta secretoria intracelular dentro del aparato de Golgi. De esta manera, ha resultado posible delinear un CD mínimo para hST6Gal I que contiene una secuencia crítica capaz de desplazar el reconocimiento de aceptores de aceptores residentes intracelularmente a glicoconjugados de superficie celular (Donadio *et al.*, 2003). Además, el CD secretado soluble de hST6Gal I muestra una eficiencia de transferencia incrementada, con independencia del patrón de ramificación del aceptor glicano (Legaigreur *et al.*, 2001).

Glucosilación de proteínas (fármacos) producidos en sistemas de expresión heteróloga

Las proteínas de interés terapéutico se extrajeron por primera vez de fuentes naturales tales como sangre, placenta, o tejidos humanos o animales. Sin embargo, este enfoque se encuentra limitado por la fuente, cantidad y disponibilidad de los tejidos humanos y puede contener contaminantes potencialmente letales (priones, oncogenes, virus, etc.), así como alérgenos potenciales generados por proteínas de animales. Con el incremento de las proteínas-fármacos autorizados y de las necesidades clínicas, se han desarrollado nuevos enfoques para la producción de proteínas utilizando diferentes sistemas de expresión.

En la mayoría de casos, los mayores esfuerzos se están focalizando actualmente en la humanización del patrón de glucosilación de las proteínas recombinantes para aproximarse al máximo al patrón en las glucoproteínas naturales con el fin de mejorar la farmacocinética y reducir la inmunogenicidad del producto.

La maquinaria necesaria para la síntesis, la activación y la introducción de los residuos sialilo se expresa pobremente en los diversos sistemas de expresión de proteínas recombinantes existentes. Por lo tanto, las proteínas humanas recombinantes producidas con frecuencia se encuentran infrasialiladas o incluso no sialiladas en absoluto en comparación con sus contrapartidas nativas.

Uno de los sistemas más utilizados es la bacteria *Escherichia coli* (*E. coli*) (Swartz, 2001; Baneyx, 1999), pero su inconveniente principal es que este procariota no lleva a cabo las modificaciones postraduccionales humanas, en particular las glucosilaciones de proteínas, debido a que no se expresa ninguna glucosiltransferasa de este tipo en *E. coli*. Lo anterior puede llevar al rechazo de la proteína terapéutica de interés por parte del sistema inmunológico, a la reducción de la vida media en circulación y/o de su actividad biológica. La proteína producida finalmente puede plegarse incorrectamente y agregar formando cuerpos de inclusión.

Las levaduras y hongos filamentosos también son sistemas de expresión eucarióticos bien conocidos y presentan una maquinaria celular similar a la de las células humanas. Las levaduras producen proteínas complejas y pueden llevar a cabo varias modificaciones postraduccionales tales como glucosilaciones simples. Sin embargo, aunque el proceso de la N-glucosilación realizado por las levaduras y los hongos es el mismo que el proceso en los mamíferos en las etapas iniciales en el retículo endoplasmático, no se ha encontrado ningún oligosacárido complejo que contenga ácido siálico, galactosa, fucosa y N-acetilgalactosamina en las glucoproteínas producidas por estos organismos (Blanchard, 2004); tanto las levaduras como los hongos típicamente producen glicanos ricos en manosa mediante la adición de hasta 100 residuos de manosa (en el caso de las levaduras) en el núcleo pentasacárido

(Tanner y Lehele, 1987; Herscovics y Orlean, 1993) en el aparato de Golgi. Esta hipermanosilación promueve una respuesta inmunológica en el ser humano.

La introducción de la glucosilación en levaduras (Hamilton *et al.*, 2003; Roy *et al.*, 2000) en primer lugar ha permitido la reducción del número de residuos de manosa añadidos. Los enzimas que resultan necesarios para la N-acetilglucosaminilación y galactosilación han sido añadidos en estos sistemas (Maras *et al.*, 1999; Bretthauer, 2003; Vervecken *et al.*, 2004). Sin embargo, la adición del ácido siálico terminal todavía resulta difícil de conseguir debido al gran número de enzimas implicado en esta etapa terminal.

Las proteínas expresadas en las células de insecto se pliegan correctamente, pueden experimentar modificaciones postraduccionales y ser secretadas. La glucosilación N-ligada temprana realizada por estas células resulta similar a la realizada por las células de mamífero. Sin embargo, las estructuras de glicano obtenidas en este caso son del tipo paucimanosídico, es decir, truncadas debido a la presencia de una actividad de N-acetilhexosaminidasa no deseable que degrada las neoglucoproteínas expresadas durante la expresión baculovírica (Blanchard, 2004). Además, en algunas líneas de células de insecto pueden encontrarse residuos α 1,3-fucosa y estos residuos generalmente inducen una respuesta inmunológica no deseable en el ser humano. De esta manera, la utilización de este sistema se restringe a la producción de antígenos vacunales.

En las células de insecto, las GT que catalizan la transferencia de azúcares inmunogénicos han sido delecionadas y la sialilación ha podido conseguirse mediante la adición de 3 genes codificantes de N-acetilglucosamina 2-epimerasa, N-acetilneuraminil-liasa y CMP-Neu5Ac sintasa (Jarvis *et al.*, 1998; Aumiller *et al.*, 2003). Se ha creado una nueva línea celular de insecto (SfSWT-3) diseñada para sintetizar su propio CMP-ácido siálico. Las células resultantes expresan la totalidad de los 7 genes de mamífero, pueden producir CMP-ácido siálico y sialilato, una proteína heteróloga en el caso de que el cultivo se realice en un medio de cultivo sin suero (Aumiller *et al.*, 2003).

Al igual que las células eucarióticas, las plantas muestran una compleja y sofisticada maquinaria celular que puede utilizarse para producir proteínas terapéuticas. Las proteínas recombinantes presentan una calidad farmacológica muy buena debido a que las plantas presentan muchos de los enzimas necesarios para la maduración de las proteínas. Sin embargo, la ruta de glucosilación requiere ajustes importantes para no producir proteínas alergénicas. En efecto, las N-glucosilaciones en plantas (Lerouge *et al.*, 2000) son similares a las realizadas en el ser humano en el aspecto de la glucosilación nuclear. Sin embargo, los glicanos todavía no presentan antenas sialiladas y contienen β 1,2-xilosa y α 1,3-fucosa internas. Ambos residuos son altamente inmunogénicos en el ser humano y en la actualidad dificultan la aprobación de las plantas transgénicas como sistemas de expresión para uso terapéutico.

En las plantas, la primera estrategia utilizada estaba destinada a prevenir la adición de azúcares alergénicos al prevenir que las proteínas salgan del retículo endoplasmático. En consecuencia, los glicanos N-ligados no pueden madurar hasta el tipo complejo. Otra estrategia se basa en la inhibición de varias GT en el interior del aparato de Golgi. Esta inhibición no puede ser completa y/o puede competir con la maquinaria endógena para la maduración.

El sistema de expresión celular de mamífero, es decir las células CHO, en la actualidad es el único sistema farmacológico aprobado para producir terapéuticos recombinantes. Estas células muestran una ventaja importante debido a que pueden sintetizar glucoproteínas N-ligadas de elevado peso molecular y/o multiméricas. Las células de mamífero expresan naturalmente no sólo los enzimas participantes en la síntesis y transporte de los azúcares de nucleótido, sino también las glucosiltransferasas necesarias para conseguir la glucosilación compleja de las proteínas recombinantes con un contenido satisfactorio de ácido siálico 3-ligado. Sin embargo, existe una falta de algunos enzimas, tales como las α 1,3/4-fucosiltransferasas y las α 2,6-sialiltransferasas, que llevan a cabo la glucosilación O-ligada y N-ligada. Además, en las células de mamífero, la sialilación se produce a través del ácido N-glucolilneuramínico (NeuGc), que también difiere significativamente del derivado N-acetilado (NeuAc) presente en células humanas.

En el caso de las células de mamífero, se han llevado a cabo estudios utilizando la sobreexpresión de una α 2,3-St y una β 1,4GalT (Weikert *et al.*, 1999); ambos enzimas se encuentran presentes en el genoma, pero sus actividades son muy variables dependiendo de las condiciones del cultivo celular. Lo anterior ha conducido a una amplia variabilidad respecto a la presencia de la Gal y ácido siálico terminales y a una extensa microheterogeneidad en las proteínas glucosiladas. Los estudios también se han orientado hacia la optimización de la galactosilación y la sialilación (Granbenhorst *et al.*, 1999) mediante la introducción de una α 2,6-ST (Bragonzi *et al.*, 2000) en células de mamífero. Se ha establecido una línea celular CHO que expresa establemente α 2,6-ST de rata nativa. Las glucoproteínas producidas por estas células CHO muestran residuos de ácido siálico terminales tanto α 2,6-ligados como α 2,3-ligados, de manera similar al ser humano; la proporción observada entre los residuos presentes de ácido siálico terminales α 2,6-ligados y α 2,3-ligados era de 40,4% de residuos ácido α 2,6-siálico y 59,6% de residuos ácido α 2,3-siálico, lo que mejoró la farmacocinética en los estudios de eliminación (Bragonzi *et al.*, 2000). A pesar de estas mejoras en la humanización de las células, la proporción entre los residuos de ácido siálico terminal α 2,6-ligados y α 2,3-ligados no puede controlarse y no se ha encontrado que resulte favorable a la actividad en posición 6. De esta manera, aparentemente la primera sialilación es una etapa crítica para el control de las estructuras de glicano y, en segundo lugar, la dificultad reside también en expresar la ST heteróloga en un compartimento específico de la célula

(direccionamiento), en optimizar su actividad en el caso de que el sustrato donador para este enzima se encuentre presente (portador).

Por lo tanto, cabe indicar que la sialilación terminal de los fármacos glucoproteicos aprobados es la etapa más difícil de alcanzar en todos los sistemas de expresión disponibles hasta hoy.

Objetivo de la invención

El objetivo de la presente invención es proporcionar un procedimiento para generar un panel de nuevas secuencias génicas codificantes de enzimas membranales quiméricas con actividad de glucosiltransferasa, y en particular para diseñar secuencias génicas codificantes de sialiltransferasas quiméricas innovadoras que doten a células de una necesaria actividad sialilante y mejoren la calidad y rendimiento de proteínas glucosiladas recombinantes.

La invención se refiere a un procedimiento para producir secuencias génicas codificantes de glucosiltransferasas membranales quiméricas que presentan una actividad de glucosilación optimizada en células transformadas con dichas secuencias, en comparación con la actividad de glucosilación de las glucosiltransferasas nativas correspondientes, es decir, una glucosilación menos selectiva y/o más eficiente (hasta por lo menos 30 veces más elevada que la actividad inicial de las glucosiltransferasas de longitud completa nativas) hacia el sustrato aceptor, comprendiendo dicho procedimiento la fusión:

- de una primera secuencia de ácidos nucleicos codificante de un fragmento mínimo C-terminal del dominio catalítico (CD) de la glucosiltransferasa de longitud completa nativa, mostrando dicho fragmento CD mínimo C-terminal una actividad de transferasa (que puede incrementarse por lo menos 30 veces respecto a la actividad inicial del CD nativo), obteniendo dicha primera secuencia de ácidos nucleicos mediante la eliminación de los nucleótidos codificantes de uno o varios aminoácidos contiguos que se extienden desde el primer aminoácido del extremo N-terminal de dicho CD, y la selección del ácido nucleico codificante de dicho fragmento CD mínimo C-terminal, el cual, en el caso de que n represente el número de aminoácidos contiguos tal como se ha indicado anteriormente que han sido delecionados, el fragmento obtenido al delecionar por lo menos n+1 aminoácidos contiguos tal como se ha definido anteriormente, no presenta actividad de transferasa sustancial,

- con un segundo ácido nucleico de secuencia variable codificante de una cadena peptídica transmembranal que especifica un anclaje de las glucosiltransferasas en comportamientos intracelulares, y que comprende (o que consiste) en su región N-terminal, una región de cola citoplasmática (CT) situada cadena arriba de un dominio transmembranal (TMB), éste localizado cadena arriba de una región tallo (SR) o de un fragmento de por lo menos 3 aminoácidos contiguos de la SR, estando unida dicha SR o parte de la misma a dicho dominio catalítico, opcionalmente mediante un conector o péptido de conexión de por lo menos 2 aminoácidos codificados por un sitio de restricción que no existe en la secuencia de nucleótidos codificante del CD nativo indicado anteriormente,

con la condición de que por lo menos uno de dichos péptidos CT, TMD y SR sea diferente de la estructura primaria de los péptidos nativos correspondientes en la glucosiltransferasa nativa de la que se deriva el fragmento de CD con actividad de glucosiltransferasa óptima tal como se ha definido anteriormente,

llevando a cabo dicha fusión de manera que el primer ácido nucleico se localice cadena abajo del segundo ácido nucleico y proporcione un producto proteico en el que el CD se encuentre en la mitad C-terminal.

Debe apreciarse la referencia a las glucosiltransferasas membranales quiméricas que presentan una actividad de glucosilación optimizada en células transformadas con dichas secuencias, en comparación con la actividad de glucosilación de las glucosiltransferasas nativas correspondientes, como que dichas glucosiltransferasas membranales quiméricas:

- presentan una actividad de transferencia de azúcares menos selectiva hacia sustratos aceptores que las glucosiltransferasas nativas correspondientes en ensayo *in vitro* con glucoproteínas aceptoras exógenas/disponibles comercialmente que presentan glicanos biantenarios, triantenarios o tetraantenarios conocidos, es decir, que presentan una actividad de glucosilación *in vivo* en células hacia todas las glucoproteínas aceptoras celulares, siendo medible dicha actividad de glucosilación preferentemente en compartimentos intracelulares y de superficie celular según el procedimiento general siguiente utilizando la unión de lectina SNA.

- y/o presentan una actividad de transferencia de azúcares más eficiente hacia sus sustratos aceptores que las glucosiltransferasas nativas correspondientes, es decir, muestran una actividad de glucosilación por lo menos 30 veces superior que la actividad inicial de las glucosiltransferasas de longitud completa nativas, siendo medible dicha actividad de glucosilación según el procedimiento *in vitro* general con glucoproteínas aceptoras exógenas/disponibles comercialmente que presentan glicanos biantenarios, triantenarios o tetraantenarios.

La expresión "glucosiltransferasas quiméricas" utilizada anteriormente corresponde a una glucosiltransferasa cuya secuencia de longitud completa no es el producto directo de un gen nativo o un transcrito, sino que ha sido diseñada utilizando secuencias de otras glucosiltransferasas exclusivamente.

La expresión "glucosiltransferasas nativas" utilizada anteriormente corresponde a la secuencia del enzima de longitud completa según se encuentra representada por una secuencia codificante natural.

- 5 La expresión "actividad de glucosilación" utilizada anteriormente corresponde a la reacción catalítica de transferir un azúcar de un nucleótido donador a un sustrato aceptor.

10 La expresión "dominio catalítico mínimo (CD)" utilizada anteriormente corresponde al dominio peptídico C-terminal de una secuencia de glucosiltransferasa que no puede deletionarse adicionalmente sin pérdida de actividad de transferencia.

La expresión "dominio transmembranal (TMD)" utilizada anteriormente corresponde a una parte peptídica compuesta de un tramo de 17 a 24 aminoácidos esencialmente hidrofóbicos.

- 15 La expresión "cola citoplasmática (CT)" utilizada anteriormente corresponde al péptido N-terminal de una glucosiltransferasa que puede comprender por lo menos más de 3 aminoácidos cadena arriba del TMD.

20 La expresión "región tallo (SR)" utilizada anteriormente corresponde a un tramo de, como máximo, 246 aminoácidos cadena abajo del TMD/cadena arriba del CD.

25 La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el primer y el segundo ácidos nucleicos se derivan de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en glucosiltransferasas de origen eucariótico, preferentemente mamíferos y seres humanos, participando dichas glucosiltransferasas en:

- la O-glucosilación de las proteínas en las células, tales como las transferasas N-acetilgalactosaminilo, N-acetilglucosaminilo, glucosilo, fucosilo, galactosilo o sialilo,
- la N-glucosilación de las proteínas en células, tales como las transferasas N-acetilglucosaminilo, galactosilo, fucosilo o sialiltransferasas,
- la glucosilación de lípidos tales como las transferasas N-acetilgalactosaminilo, N-acetilglucosaminilo, fucosilo, galactosilo o sialilo.

- 35 La expresión "O-glucosilación" utilizada anteriormente corresponde a la ruta biosintética que elabora monosacáridos u oligosacáridos unidos covalentemente a aminoácidos que no son residuos asparagina, pero que preferentemente pueden ser hidroxiaminoácidos tales como los residuos serina, treonina o hidroxilisina.

40 La expresión "N-glucosilación" utilizada anteriormente corresponde a la ruta biosintética que elabora oligosacáridos unidos a los residuos asparagina dentro del tripéptido de consenso Asn-X-Ser/Thr (X no siendo prolina).

La expresión "galactosiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que transfiere un residuo galactosa de UDP-Gal a una proteína o glucolípido aceptor N-ligado/O-ligado.

- 45 La expresión "sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que transfiere un residuo ácido siálico, preferentemente un derivado del ácido neuramínico, de CMP-NeuAc a una proteína o glucolípido aceptor N-ligado/O-ligado.

50 La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque los primer y segundo ácidos nucleicos se derivan de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente en N-acetilgalactosaminiltransferasas, N-acetilglucosaminiltransferasas, glucosiltransferasas, galactosiltransferasas, fucosiltransferasas o sialiltransferasas.

55 La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el primer y segundo ácidos nucleicos se derivan de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente en alfa-6-fucosiltransferasas (glucosilación del núcleo), beta-2/4/6-N-acetilglucosaminiltransferasas (ramificación), beta-4-galactosiltransferasas y alfa-3/6/8-sialiltransferasas (glucosilación terminal).

- 60 Ventajosamente, el primer y segundo ácidos nucleicos se derivan de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, de enzimas implicados en la ruta de N-glucosilación.

65 La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el primer y segundo ácidos nucleicos se derivan de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en sialiltransferasas.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el primer y segundo ácidos nucleicos se derivan de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en α 2,6-sialiltransferasas, α 2,3-sialiltransferasas o α 2,8-sialiltransferasas.

Ventajosamente, las secuencias de nucleótidos codificantes de CT, TMD y/o SR o fragmento de SR mencionados anteriormente son preferentemente de origen sintético.

La expresión " α 2,6-sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que puede transferir un residuo ácido siálico, preferentemente un derivado de ácido neuramínico, de CMP-NeuAc a la posición 6 de un carbohidrato aceptor del tipo proteína o glucolípido N-ligado/O-ligado.

La expresión " α 2,3-sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que puede transferir un residuo ácido siálico, preferentemente un derivado del ácido neuramínico, de CMP-NeuAc a la posición 3 un carbohidrato aceptor del tipo proteína o glucolípido N-ligado/O-ligado.

La expresión " α 2,8-sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa capaz de transferir un residuo ácido siálico, preferentemente un derivado del ácido neuramínico, de CMP-NeuAc a la posición 8 de un aceptor sialilado del tipo proteína o glucolípido N-ligado/O-ligado.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque los primer y segundo ácidos nucleicos se derivan de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en:

- α 2,6-sialiltransferasas seleccionadas de entre:

* las β 1,4-galactósido α 2,6-sialiltransferasas I y II humanas (hST6Gal I y hST6Gal II) representadas por las secuencias SEC ID nº 2 y SEC ID nº 4, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 1 y SEC ID nº 3, respectivamente, o

* las N-acetilgalactosaminido- α 2,6-sialiltransferasas I a VI humanas (hST6GalNAc I a VI), representadas por las secuencias SEC ID nº 6, SEC ID nº 8, SEC ID nº 10, SEC ID nº 12, SEC ID nº 14 y SEC ID nº 16, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 5, SEC ID nº 7, SEC ID nº 9, SEC ID nº 11, SEC ID nº 13 y SEC ID nº 15, respectivamente,

- α 2,3-sialiltransferasas seleccionadas de entre las galactósido- α 2,3-sialiltransferasas I a VI humanas (hST3Gal I a VI) representadas por las secuencias SEC ID nº 18, SEC ID nº 20, SEC ID nº 22, SEC ID nº 24, SEC ID nº 26 y SEC ID nº 28, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 17, SEC ID nº 19, SEC ID nº 21, SEC ID nº 23, SEC ID nº 25 y SEC ID nº 27, respectivamente, o las galactósido- α 2,3-sialiltransferasas I a VI de rata (rST3Gal I a VI), tales como la rST3Gal III representada por la secuencia SEC ID nº 30 codificada por la secuencia de nucleótidos SEC ID nº 29, o cualquier ST de otro origen animal con la condición de que comparta una homología de por lo menos 85% con el enzima humano.

- α 2,8-sialiltransferasas seleccionadas de entre las ácido siálico- α 2,8-sialiltransferasas I a VI humanas (hST8Sia I a VI) representadas por las secuencias SEC ID nº 32, SEC ID nº 34, SEC ID nº 36, SEC ID nº 38, SEC ID nº 40 y SEC ID nº 42, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 31, SEC ID nº 33, SEC ID nº 35, SEC ID nº 37, SEC ID nº 39 y SEC ID nº 41, respectivamente.

La expresión "galactósido- α 2,6-sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que transfiere un residuo ácido siálico, preferentemente un derivado del ácido neuramínico, de CMP-NeuAc a la posición 6 de un receptor galactosilado del tipo proteína o lípido N-ligado/O-ligado.

La expresión "N-acetilgalactosaminido- α 2,6-sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que transfiere un residuo ácido siálico, preferentemente un derivado del ácido neuramínico de CMP-NeuAc a la posición 6 de un residuo N-acetilgalactosaminilo de un aceptor proteína o glucolípido N-ligado/O-ligado.

La expresión "galactósido- α 2,3-sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que transfiere un residuo ácido siálico, preferentemente un derivado del ácido neuramínico, de CMP-NeuAc a la posición 3 de un aceptor proteína o lípido N-ligado/O-ligado galactosilado.

La expresión "ácido siálico- α 2,8-sialiltransferasas" utilizada anteriormente corresponde a una glucosiltransferasa que transfiere un residuo ácido siálico, preferentemente un derivado del ácido neuramínico, de CMP-NeuAc a un aceptor proteína o glucolípido N-ligado/O-ligado sialilado.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque la secuencia de nucleótidos codificante del dominio CD comprendido en el primer ácido nucleico se selecciona de entre las secuencias constituidas por, o que comprenden:

- 5 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 268 y 330, y en su extremo 3' por el nucleótido situado en la posición 1.218 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6Gal I delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 90 y 110 y en su extremo C-terminal por el aminoácido situado en la posición 406 de la secuencia SEC ID nº 2,
- 10 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 307 y 369, y en su extremo 3' por el nucleótido situado en la posición 1.587 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6Gal II delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 103 y 123, y en su extremo C-terminal por el aminoácido situado en la posición 529 de la secuencia SEC ID nº 4,
- 15 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 814 y 876, y en su extremo 3' por el nucleótido situado en la posición 1.800 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac I delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 272 y 292 y en su extremo C-terminal por el aminoácido situado en la posición 600 de la secuencia SEC ID nº 6,
- 20 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 172 y 234, y en su extremo 3' por el nucleótido situado en la posición 1.122 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac II delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 58 y 78 y en su extremo C-terminal por el aminoácido situado en la posición 374 de la secuencia SEC ID nº 8,
- 25 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 73 y 135, y en su extremo 3' por el nucleótido situado en la posición 915 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac III delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 25 y 45, y en su extremo C-terminal por el aminoácido situado en la posición 305 de la secuencia SEC ID nº 10,
- 30 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 61 y 123, y en su extremo 3' por el nucleótido situado en la posición 906 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac IV delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 21 y 41, y en su extremo C-terminal por el aminoácido situado en la posición 302 de la secuencia SEC ID nº 12,
- 35 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 121 y 183, y en su extremo 3' por el nucleótido situado en la posición 1.008 de la secuencia SEC ID nº 13, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac V delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 41 y 61, y en su extremo C-terminal por el aminoácido situado en la posición 336 de la secuencia SEC ID nº 14,
- 40 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 154 y 216, y en su extremo 3' por el nucleótido situado en la posición 999 de la secuencia SEC ID nº 15, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac VI delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 52 y 72 y en su extremo C-terminal por el aminoácido situado en la posición 333 de la secuencia SEC ID nº 16,
- 45 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 145 y 207, y en su extremo 3' por el nucleótido situado en la posición 1.020 de la secuencia SEC ID nº 17, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal I delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 49 y 69, y en su extremo C-terminal por el aminoácido situado en la posición 340 de la secuencia SEC ID nº 18,
- 50 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 175 y 237, y en su extremo 3' por el nucleótido situado en la posición 1.050 de la secuencia SEC ID nº 19, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal II delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 59 y 79, y en su extremo C-terminal por el aminoácido situado en la posición 350 de la secuencia SEC ID nº 20,
- 55 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 199 y 261, y en su extremo 3' por el nucleótido situado en la posición 1.332 de la secuencia SEC ID nº 21, codificando dicha
- 60
- 65

secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal III delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 67 y 87, y en su extremo C-terminal por el aminoácido situado en la posición 444 de la secuencia SEC ID nº 22,

5 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 79 y 141, y en su extremo 3' por el nucleótido situado en la posición 987 de la secuencia SEC ID nº 23, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal IV delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 27 y 47, y en su extremo C-terminal por el aminoácido situado en la posición 329 de la secuencia SEC ID nº 24,

10 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 136 y 198, y en su extremo 3' por el nucleótido situado en la posición 1.086 de la secuencia SEC ID nº 25, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal V delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 46 y 66, y en su extremo C-terminal por el aminoácido situado en la posición 362 de la secuencia SEC ID nº 26,

15 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 73 a 135, y en su extremo 3' por el nucleótido situado en la posición 993 de la secuencia SEC ID nº 27, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal VI delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 25 y 45, y en su extremo C-terminal por el aminoácido situado en la posición 331 de la secuencia SEC ID nº 28,

20 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 103 y 165, y en su extremo 3' por el nucleótido situado en la posición 1.122 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de ST3Gal III de rata delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 35 y 55, y en su extremo C-terminal por el aminoácido situado en la posición 374 de la secuencia SEC ID nº 30,

25 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 133 a 195, y en su extremo 3' por el nucleótido situado en la posición 1.068 de la secuencia SEC ID nº 31, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia I delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 45 y 65, y en su extremo C-terminal por el aminoácido situado en la posición 356 de la secuencia SEC ID nº 32,

30 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 190 a 252, y en su extremo 3' por el nucleótido situado en la posición 1.125 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia II delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 64 y 84, y en su extremo C-terminal por el aminoácido situado en la posición 375 de la secuencia SEC ID nº 34,

35 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 205 a 267, y en su extremo 3' por el nucleótido situado en la posición 1.140 de la secuencia SEC ID nº 35, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia III delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 69 y 89, y en su extremo C-terminal por el aminoácido situado en la posición 380 de la secuencia SEC ID nº 36,

40 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 145 a 207, y en su extremo 3' por el nucleótido situado en la posición 1.077 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia IV delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 49 y 69, y en su extremo C-terminal por el aminoácido situado en la posición 359 de la secuencia SEC ID nº 38,

45 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 211 a 273, y en su extremo 3' por el nucleótido situado en la posición 1.128 de la secuencia SEC ID nº 39, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia V delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 71 y 91, y en su extremo C-terminal por el aminoácido situado en la posición 376 de la secuencia SEC ID nº 40,

50 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 277 a 339, y en su extremo 3' por el nucleótido situado en la posición 1.194 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia VI delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 93 y 113, y en su extremo C-terminal por el aminoácido situado en la posición 398 de la secuencia SEC ID nº 42.

55 La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el primer ácido nucleico es la secuencia de nucleótidos SEC ID nº 43, correspondiente a la secuencia

delimitada por los nucleótidos situados en las posiciones 268 y 1.218 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos SEC ID nº 43 la secuencia polipeptídica SEC ID nº 44, correspondiente al fragmento mínimo C-terminal del dominio CD de hST6Gal I delimitado por los aminoácidos situados entre las posiciones 90 y 406 de la secuencia SEC ID nº 2.

5 La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque:

10 - la secuencia de nucleótidos codificante de la región CT comprendida en el segundo ácido nucleico se selecciona de entre:

15 * la secuencia de nucleótidos SEC ID nº 45 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 27 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos SEC ID nº 45 la secuencia polipeptídica SEC ID nº 46 correspondiente a la región CT de hST6Gal I delimitada por los aminoácidos situados entre las posiciones 1 y 9 de la secuencia SEC ID nº 2,

20 * la secuencia de nucleótidos SEC ID nº 47 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 30 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos SEC ID nº 47 la secuencia polipeptídica SEC ID nº 48 correspondiente a la región CT de hST6Gal II delimitada por los aminoácidos situados entre las posiciones 1 y 10 de la secuencia SEC ID nº 4,

25 * la secuencia de nucleótidos SEC ID nº 49 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 42 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos SEC ID nº 50 la secuencia polipeptídica SEC ID nº 49 correspondiente a la región CT de hST6GalINAc I delimitada por los aminoácidos situados entre las posiciones 1 y 14 de la secuencia SEC ID nº 6,

30 * la secuencia de nucleótidos SEC ID nº 51 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 21 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos SEC ID nº 51 la secuencia polipeptídica SEC ID nº 52 correspondiente a la región CT de hST6GalINAc II delimitada por los aminoácidos situados entre las posiciones 1 y 7 de la secuencia SEC ID nº 8,

35 * la secuencia de nucleótidos SEC ID nº 53 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos SEC ID nº 53 la secuencia polipeptídica SEC ID nº 54 correspondiente a la región CT de hST6GalINAc III delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 10,

40 * la secuencia de nucleótidos SEC ID nº 55 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 18 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos SEC ID nº 55 la secuencia polipeptídica SEC ID nº 56 correspondiente a la región CT de hST6GalINAc IV delimitada por los aminoácidos situados entre las posiciones 1 y 6 de la secuencia SEC ID nº 12,

45 * la secuencia de nucleótidos SEC ID nº 57 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 13, codificando dicha secuencia de nucleótidos SEC ID nº 57 la secuencia polipeptídica SEC ID nº 58 correspondiente a la región CT de hST6GalINAc V delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 14,

50 * la secuencia de nucleótidos SEC ID nº 59 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 129 de la secuencia SEC ID nº 15, codificando dicha secuencia de nucleótidos SEC ID nº 59 la secuencia polipeptídica SEC ID nº 60 correspondiente a la región CT de hST6GalINAc VI delimitada por los aminoácidos situados entre las posiciones 1 y 43 de la secuencia SEC ID nº 16,

55 * la secuencia de nucleótidos SEC ID nº 61 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 39 de la secuencia SEC ID nº 17, codificando dicha secuencia de nucleótidos SEC ID nº 61 la secuencia polipeptídica SEC ID nº 62 correspondiente a la región CT de hST3Gal I delimitada por los aminoácidos situados entre las posiciones 1 y 13 de la secuencia SEC ID nº 18,

60 * la secuencia de nucleótidos SEC ID nº 63 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 18 de la secuencia SEC ID nº 19, codificando dicha secuencia de nucleótidos SEC ID nº 63 la secuencia polipeptídica SEC ID nº 64 correspondiente a la región CT de hST3Gal II delimitada por los aminoácidos situados entre las posiciones 1 y 6 de la secuencia SEC ID nº 20,

65 * la secuencia de nucleótidos SEC ID nº 65 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos SEC ID nº 65 la secuencia polipeptídica SEC ID nº 66 correspondiente a la región CT de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 22,

* la secuencia de nucleótidos SEC ID nº 67 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 23, codificando dicha secuencia de nucleótidos SEC ID nº 67 la secuencia polipeptídica SEC ID nº 68 correspondiente a la región CT de hST3Gal IV delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 24,

* la secuencia de nucleótidos SEC ID nº 69 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 15 de la secuencia SEC ID nº 25, codificando dicha secuencia de nucleótidos SEC ID nº 69 la secuencia polipeptídica SEC ID nº 70 correspondiente a la región CT de hST3Gal V delimitada por los aminoácidos situados entre las posiciones 1 y 5 de la secuencia SEC ID nº 26,

* la secuencia de nucleótidos SEC ID nº 71 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 12 de la secuencia SEC ID nº 27, codificando dicha secuencia de nucleótidos SEC ID nº 71 la secuencia polipeptídica SEC ID nº 72 correspondiente a la región CT de hST3Gal VI delimitada por los aminoácidos situados entre las posiciones 1 y 4 de la secuencia SEC ID nº 28,

* la secuencia de nucleótidos SEC ID nº 73 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos SEC ID nº 73 la secuencia polipeptídica SEC ID nº 74 correspondiente a la región CT de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 30,

* la secuencia de nucleótidos SEC ID nº 75 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 87 de la secuencia SEC ID nº 31, codificando dicha secuencia de nucleótidos SEC ID nº 75 la secuencia polipeptídica SEC ID nº 76 correspondiente a la región CT de hST8Sia I delimitada por los aminoácidos situados entre las posiciones 1 y 29 de la secuencia SEC ID nº 32,

* la secuencia de nucleótidos SEC ID nº 77 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 18 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos SEC ID nº 77 la secuencia polipeptídica SEC ID nº 78 correspondiente a la región CT de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 1 y 6 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos SEC ID nº 79 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 27 de la secuencia SEC ID nº 35, codificando dicha secuencia de nucleótidos SEC ID nº 79 la secuencia polipeptídica SEC ID nº 80 correspondiente a la región CT de hST8Sia III delimitada por los aminoácidos situados entre las posiciones 1 y 9 de la secuencia SEC ID nº 36,

* la secuencia de nucleótidos SEC ID nº 81 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 21 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos SEC ID nº 81 la secuencia polipeptídica SEC ID nº 82 correspondiente a la región CT de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 1 y 7 de la secuencia SEC ID nº 38,

* la secuencia de nucleótidos SEC ID nº 83 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 51 de la secuencia SEC ID nº 39, codificando dicha secuencia de nucleótidos SEC ID nº 83 la secuencia polipeptídica SEC ID nº 84 correspondiente a la región CT de hST8Sia V delimitada por los aminoácidos situados entre las posiciones 1 y 17 de la secuencia SEC ID nº 40,

* la secuencia de nucleótidos SEC ID nº 85 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 9 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos SEC ID nº 85 la secuencia polipeptídica SEC ID nº 86 correspondiente a la región CT de hST8Sia VI delimitada por los aminoácidos situados entre las posiciones 1 y 3 de la secuencia SEC ID nº 42,

- la secuencia de nucleótidos codificante de la región TMD comprendida en el segundo ácido nucleico se selecciona de entre:

* la secuencia de nucleótidos SEC ID nº 87 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 28 y 78 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos SEC ID nº 87 la secuencia polipeptídica SEC ID nº 88 correspondiente a la región TMD de hST6Gal I delimitada por los aminoácidos situados entre las posiciones 10 y 26 de la secuencia SEC ID nº 2,

* la secuencia de nucleótidos SEC ID nº 89 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 31 y 90 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos SEC ID nº 89 la secuencia polipeptídica SEC ID nº 90 correspondiente a la región TMD de hST6Gal II delimitada por los aminoácidos situados entre las posiciones 11 y 30 de la secuencia SEC ID nº 4,

* la secuencia de nucleótidos SEC ID nº 91 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 43 y 105 de la secuencia SEC ID nº 5, codificando dicha secuencia de

nucleótidos SEC ID nº 91 la secuencia polipeptídica SEC ID nº 92 correspondiente a la región TMD de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 15 y 35 de la secuencia SEC ID nº 6,

- 5 * la secuencia de nucleótidos SEC ID nº 93 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 22 y 84 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos SEC ID nº 93 la secuencia polipeptídica SEC ID nº 94 correspondiente a la región TMD de hST6GalNAc II delimitada por los aminoácidos situados entre las posiciones 8 y 28 de la secuencia SEC ID nº 8,
- 10 * la secuencia de nucleótidos SEC ID nº 95 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos SEC ID nº 95 la secuencia polipeptídica SEC ID nº 96 correspondiente a la región TMD de hST6GalNAc III delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 10,
- 15 * la secuencia de nucleótidos SEC ID nº 97 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 19 y 81 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos SEC ID nº 97 la secuencia polipeptídica SEC ID nº 98 correspondiente a la región TMD de hST6GalNAc IV delimitada por los aminoácidos situados entre las posiciones 7 y 27 de la secuencia SEC ID nº 12,
- 20 * la secuencia de nucleótidos SEC ID nº 99 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 87 de la secuencia SEC ID nº 13, codificando dicha secuencia de nucleótidos SEC ID nº 99 la secuencia polipeptídica SEC ID nº 100 correspondiente a la región TMD de hST6GalNAc V delimitada por los aminoácidos situados entre las posiciones 9 y 29 de la secuencia SEC ID nº 14,
- 25 * la secuencia de nucleótidos SEC ID nº 101 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 130 y 177 de la secuencia SEC ID nº 15, codificando dicha secuencia de nucleótidos SEC ID nº 101 la secuencia polipeptídica SEC ID nº 102 correspondiente a la región TMD de hST6GalNAc VI delimitada por los aminoácidos situados entre las posiciones 44 y 59 de la secuencia SEC ID nº 16,
- 30 * la secuencia de nucleótidos SEC ID nº 103 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 40 y 102 de la secuencia SEC ID nº 17, codificando dicha secuencia de nucleótidos SEC ID nº 103 la secuencia polipeptídica SEC ID nº 104 correspondiente a la región TMD de hST3Gal I delimitada por los aminoácidos situados entre las posiciones 14 y 34 de la secuencia SEC ID nº 18,
- 35 * la secuencia de nucleótidos SEC ID nº 105 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 19 y 81 de la secuencia SEC ID nº 19, codificando dicha secuencia de nucleótidos SEC ID nº 105 la secuencia polipeptídica SEC ID nº 106 correspondiente a la región TMD de hST3Gal II delimitada por los aminoácidos situados entre las posiciones 7 y 27 de la secuencia SEC ID nº 20,
- 40 * la secuencia de nucleótidos SEC ID nº 107 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos SEC ID nº 107 la secuencia polipeptídica SEC ID nº 108 correspondiente a la región TMD de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 22,
- 45 * la secuencia de nucleótidos SEC ID nº 109 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 78 de la secuencia SEC ID nº 23, codificando dicha secuencia de nucleótidos SEC ID nº 109 la secuencia polipeptídica SEC ID nº 110 correspondiente a la región TMD de hST3Gal IV delimitada por los aminoácidos situados entre las posiciones 9 y 26 de la secuencia SEC ID nº 24,
- 50 * la secuencia de nucleótidos SEC ID nº 111 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 16 y 78 de la secuencia SEC ID nº 25, codificando dicha secuencia de nucleótidos SEC ID nº 111 la secuencia polipeptídica SEC ID nº 112 correspondiente a la región TMD de hST3Gal V delimitada por los aminoácidos situados entre las posiciones 6 y 26 de la secuencia SEC ID nº 26,
- 55 * la secuencia de nucleótidos SEC ID nº 113 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 13 y 75 de la secuencia SEC ID nº 27, codificando dicha secuencia de nucleótidos SEC ID nº 113 la secuencia polipeptídica SEC ID nº 114 correspondiente a la región TMD de hST3Gal VI delimitada por los aminoácidos situados entre las posiciones 5 y 25 de la secuencia SEC ID nº 28,
- 60 * la secuencia de nucleótidos SEC ID nº 115 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos SEC ID nº 115 la secuencia polipeptídica SEC ID nº 116 correspondiente a la región TMD de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 30,
- 65 * la secuencia de nucleótidos SEC ID nº 117 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 88 y 144 de la secuencia SEC ID nº 31, codificando dicha secuencia de

nucleótidos SEC ID nº 117 la secuencia polipeptídica SEC ID nº 118 correspondiente a la región TMD de hST8Sia I delimitada por los aminoácidos situados entre las posiciones 30 y 48 de la secuencia SEC ID nº 32,

* la secuencia de nucleótidos SEC ID nº 119 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 19 y 69 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos SEC ID nº 119 la secuencia polipeptídica SEC ID nº 120 correspondiente a la región TMD de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 7 y 23 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos SEC ID nº 121 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 28 y 99 de la secuencia SEC ID nº 35, codificando dicha secuencia de nucleótidos SEC ID nº 121 la secuencia polipeptídica SEC ID nº 122 correspondiente a la región TMD de hST8Sia III delimitada por los aminoácidos situados entre las posiciones 10 y 33 de la secuencia SEC ID nº 36,

* la secuencia de nucleótidos SEC ID nº 123 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 22 y 60 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos SEC ID nº 123 la secuencia polipeptídica SEC ID nº 124 correspondiente a la región TMD de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 8 y 20 de la secuencia SEC ID nº 38,

* la secuencia de nucleótidos SEC ID nº 125 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 52 y 114 de la secuencia SEC ID nº 39, codificando dicha secuencia de nucleótidos SEC ID nº 125 la secuencia polipeptídica SEC ID nº 126 correspondiente a la región TMD de hST8Sia V delimitada por los aminoácidos situados entre las posiciones 18 y 38 de la secuencia SEC ID nº 40,

* la secuencia de nucleótidos SEC ID nº 127 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 10 y 72 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos SEC ID nº 127 la secuencia polipeptídica SEC ID nº 128 correspondiente a la región TMD de hST8Sia VI delimitada por los aminoácidos situados entre las posiciones 4 y 24 de la secuencia SEC ID nº 42,

- la secuencia de nucleótidos codificante de la región SR comprendida en el segundo ácido nucleico, o codificante de un fragmento de por lo menos 2 aminoácidos del mismo, se selecciona de entre las secuencias constituidas o que comprenden:

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 79, y en su extremo 3' por el nucleótido situado entre las posiciones 267 y 327 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal I delimitada en su extremo N-terminal por el aminoácido situado en la posición 27 y en su extremo C-terminal por el aminoácido situado entre las posiciones 89 y 109 de la secuencia SEC ID nº 2,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 91, y en su extremo 3' por el nucleótido situado entre las posiciones 306 y 336 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal II delimitada en su extremo N-terminal por el aminoácido situado en la posición 31 y en su extremo C-terminal por el aminoácido situado entre las posiciones 102 y 112 de la secuencia SEC ID nº 4,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 106, y en su extremo 3' por el nucleótido situado entre las posiciones 813 y 873 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc I delimitada en su extremo N-terminal por el aminoácido situado en la posición 36 y en su extremo C-terminal por el aminoácido situado entre las posiciones 271 y 291 de la secuencia SEC ID nº 6,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 171 y 231 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc II delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 57 y 77 de la secuencia SEC ID nº 8,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 102 y 132 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc III delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 34 y 44 de la secuencia SEC ID nº 10,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 82, y en su extremo 3' por el nucleótido situado entre las posiciones 90 y 120 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc IV delimitada

en su extremo N-terminal por el aminoácido situado en la posición 28 y en su extremo C-terminal por el aminoácido situado entre las posiciones 30 y 40 de la secuencia SEC ID nº 12,

- 5 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 88, y en su extremo 3' por el nucleótido situado entre las posiciones 120 y 180 de la secuencia SEC ID nº 13, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc V delimitada en su extremo N-terminal por el aminoácido situado en la posición 30 y en su extremo C-terminal por el aminoácido situado entre las posiciones 40 y 60 de la secuencia SEC ID nº 14,
- 10 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 178, y en su extremo 3' por el nucleótido situado en la posición 183 de la secuencia SEC ID nº 15, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc VI delimitada en su extremo N-terminal por el aminoácido situado en la posición 60 y en su extremo C-terminal por el aminoácido situado en la posición 61 de la secuencia SEC ID nº 16,
- 15 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 103, y en su extremo 3' por el nucleótido situado entre las posiciones 144 y 204 de la secuencia SEC ID nº 17, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal I delimitada en su extremo N-terminal por el aminoácido situado en la posición 35 y en su extremo C-terminal por el aminoácido situado entre las posiciones 48 y 68 de la secuencia SEC ID nº 18,
- 20 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 82, y en su extremo 3' por el nucleótido situado entre las posiciones 174 y 234 de la secuencia SEC ID nº 19, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal II delimitada en su extremo N-terminal por el aminoácido situado en la posición 28 y en su extremo C-terminal por el aminoácido situado entre las posiciones 58 y 78 de la secuencia SEC ID nº 20,
- 25 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 198 y 258 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal III delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 66 y 86 de la secuencia SEC ID nº 22,
- 30 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 79, y en su extremo 3' por el nucleótido situado entre las posiciones 108 y 138 de la secuencia SEC ID nº 23, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal IV delimitada en su extremo N-terminal por el aminoácido situado en la posición 27 y en su extremo C-terminal por el aminoácido situado entre las posiciones 36 y 46 de la secuencia SEC ID nº 24,
- 35 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 79, y en su extremo 3' por el nucleótido situado entre las posiciones 135 y 195 de la secuencia SEC ID nº 25, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal V delimitada en su extremo N-terminal por el aminoácido situado en la posición 27 y en su extremo C-terminal por el aminoácido situado entre las posiciones 45 y 65 de la secuencia SEC ID nº 26,
- 40 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 76, y en su extremo 3' por el nucleótido situado entre las posiciones 103 y 132 de la secuencia SEC ID nº 27, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal VI delimitada en su extremo N-terminal por el aminoácido situado en la posición 26 y en su extremo C-terminal por el aminoácido situado entre las posiciones 34 y 44 de la secuencia SEC ID nº 28,
- 45 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 105 y 165 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de rST3Gal III delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 35 y 55 de la secuencia SEC ID nº 30,
- 50 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 145, y en su extremo 3' por el nucleótido situado entre las posiciones 162 y 192 de la secuencia SEC ID nº 31, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia I delimitada en su extremo N-terminal por el aminoácido situado en la posición 49 y en su extremo C-terminal por el aminoácido situado entre las posiciones 54 y 64 de la secuencia SEC ID nº 32,
- 55 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 70, y en su extremo 3' por el nucleótido situado entre las posiciones 189 y 249 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia II delimitada
- 60
- 65

en su extremo N-terminal por el aminoácido situado en la posición 24 y en su extremo C-terminal por el aminoácido situado entre las posiciones 63 y 83 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 100, y en su extremo 3' por el nucleótido situado entre las posiciones 204 y 264 de la secuencia SEC ID nº 35, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia III delimitada en su extremo N-terminal por el aminoácido situado en la posición 34 y en su extremo C-terminal por el aminoácido situado entre las posiciones 68 y 88 de la secuencia SEC ID nº 36,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 61, y en su extremo 3' por el nucleótido situado entre las posiciones 144 y 204 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia IV delimitada en su extremo N-terminal por el aminoácido situado en la posición 21 y en su extremo C-terminal por el aminoácido situado entre las posiciones 48 y 68 de la secuencia SEC ID nº 38,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 115, y en su extremo 3' por el nucleótido situado entre las posiciones 210 y 270 de la secuencia SEC ID nº 39, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia V delimitada en su extremo N-terminal por el aminoácido situado en la posición 39 y en su extremo C-terminal por el aminoácido situado entre las posiciones 70 y 90 de la secuencia SEC ID nº 40,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 73, y en su extremo 3' por el nucleótido situado entre las posiciones 276 y 336 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia VI delimitada en su extremo N-terminal por el aminoácido situado en la posición 25 y en su extremo C-terminal por el aminoácido situado entre las posiciones 92 y 112 de la secuencia SEC ID nº 42,

* o cualquier fragmento de por lo menos 6 nucleótidos de las secuencias de nucleótidos codificantes secuencias polipeptídicas correspondientes a las regiones SR indicadas anteriormente, y codificantes de por lo menos 2 aminoácidos contiguos de dichas regiones SR, tales como:

** el fragmento SEC ID nº 129 delimitado por los nucleótidos situados entre las posiciones 106 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 130 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 74 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 131 delimitado por los nucleótidos situados entre las posiciones 109 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 132 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 37 y 74 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 133 delimitado por los nucleótidos situados entre las posiciones 106 y 420 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 134 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 140 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 135 delimitado por los nucleótidos situados entre las posiciones 106 y 774 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 136 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 258 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 137 delimitado por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 21, codificante de la secuencia polipeptídica SEC ID nº 138 correspondiente al fragmento de la región SR de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 22,

** el fragmento SEC ID nº 139 delimitado por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 29, codificante de la secuencia polipeptídica SEC ID nº 140 correspondiente al fragmento de la región SR de ST3GalII de rata delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 30,

** el fragmento SEC ID nº 141 delimitado por los nucleótidos situados entre las posiciones 70 y 237 de la secuencia SEC ID nº 33, codificante de la secuencia polipeptídica SEC ID nº 142 correspondiente al fragmento de la región SR de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 24 y 79 de la secuencia SEC ID nº 34,

** el fragmento SEC ID nº 143 delimitado por los nucleótidos situados entre las posiciones 61 y 201 de la secuencia SEC ID nº 37, codificante de la secuencia polipeptídica SEC ID nº 144 correspondiente al fragmento de la región SR de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 21 y 67 de la secuencia SEC ID nº 38.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque:

- la secuencia de nucleótidos codificante de la región CT comprendida en el segundo ácido nucleico se selecciona de entre:

* la secuencia de nucleótidos SEC ID nº 49 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 1 y 42 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos SEC ID nº 49 la secuencia polipeptídica SEC ID nº 50 correspondiente a la región CT de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 1 y 14 de la secuencia SEC ID nº 6,

* la secuencia de nucleótidos SEC ID nº 65 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 1 y 24 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos SEC ID nº 65 la secuencia polipeptídica SEC ID nº 66 correspondiente a la región CT de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 22,

* la secuencia de nucleótidos SEC ID nº 73 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 1 y 24 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos SEC ID nº 73 la secuencia polipeptídica SEC ID nº 74 correspondiente a la región CT de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 30,

* la secuencia de nucleótidos SEC ID nº 77 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 1 y 18 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos SEC ID nº 77 la secuencia polipeptídica SEC ID nº 78 correspondiente a la región CT de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 1 y 6 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos SEC ID nº 81 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 1 y 21 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos SEC ID nº 81 la secuencia polipeptídica SEC ID nº 82 correspondiente a la región CT de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 1 y 7 de la secuencia SEC ID nº 38,

-la secuencia de nucleótidos codificante de la región TMD comprendida en el segundo ácido nucleico se selecciona de entre:

* la secuencia de nucleótidos SEC ID nº 91 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 43 y 105 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos SEC ID nº 91 la secuencia polipeptídica SEC ID nº 92 correspondiente a la región TMD de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 15 y 35 de la secuencia SEC ID nº 6,

* la secuencia de nucleótidos SEC ID nº 107 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 25 y 84 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos SEC ID nº 107 la secuencia polipeptídica SEC ID nº 108 correspondiente a la región TMD de hST3GalIII delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 22,

* la secuencia de nucleótidos SEC ID nº 115 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 25 y 84 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos SEC ID nº 115 la secuencia polipeptídica SEC ID nº 116 correspondiente a la región TMD de rST3Gal III delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 30,

* la secuencia de nucleótidos SEC ID nº 119 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 19 y 69 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos SEC ID nº 119 la secuencia polipeptídica SEC ID nº 120 correspondiente a la región TMD de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 7 y 23 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos SEC ID nº 123 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados en las posiciones 22 y 60 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos SEC ID nº 123 la secuencia polipeptídica SEC ID nº 124 correspondiente a la región TMD de hST8 Sia IV delimitada por los aminoácidos situados entre las posiciones 8 y 20 de la secuencia SEC ID nº 38,

- la secuencia de nucleótidos codificante de la región SR o un fragmento de la misma comprendido en el segundo ácido nucleico se selecciona de entre:

* la secuencia SEC ID nº 129 delimitada por los nucleótidos situados entre las posiciones 106 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 130 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 74 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 131 delimitada por los nucleótidos situados entre las posiciones 109 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 132 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 37 y 74 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 133 delimitada por los nucleótidos situados entre las posiciones 106 y 420 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 134 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 140 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 135 delimitada por los nucleótidos situados entre las posiciones 106 y 774 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 136 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 258 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 137 delimitada por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 21, codificante de la secuencia polipeptídica SEC ID nº 138 correspondiente al fragmento de la región SR de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 22,

* la secuencia SEC ID nº 139 delimitada por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 29, codificante de la secuencia polipeptídica SEC ID nº 140 correspondiente al fragmento de la región SR de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 30,

* la secuencia SEC ID nº 141 delimitada por los nucleótidos situados entre las posiciones 70 y 237 de la secuencia SEC ID nº 33, codificante de la secuencia polipeptídica SEC ID nº 142 correspondiente al fragmento de la región SR de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 24 y 79 de la secuencia SEC ID nº 34,

* la secuencia SEC ID nº 143 delimitada por los nucleótidos situados entre las posiciones 61 y 201 de la secuencia SEC ID nº 37, codificante de la secuencia polipeptídica SEC ID nº 144 correspondiente al fragmento de la región SR de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 21 y 67 de la secuencia SEC ID nº 38.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque los péptidos CT, TMD, SR o fragmento de SR comprendidos en el segundo ácido nucleico, son secuencias homólogas derivadas de la misma glucosiltransferasa nativa, siendo ésta última diferente de los péptidos en la glucosiltransferasa nativa a partir de la que se derivó el fragmento CD con actividad glucosiltransferasa óptima tal como se ha definido anteriormente.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el segundo ácido nucleico se selecciona de entre las secuencias siguientes:

- la secuencia SEC ID nº 145 delimitada por los nucleótidos situados entre las posiciones 1 y 222 de la secuencia SEC ID nº 5, que contiene las secuencias SEC ID nº 49, 91 y 129, y codificantes de la secuencia polipeptídica SEC ID nº 146 correspondiente al fragmento de hST6GalNAc I delimitado por los aminoácidos situados entre las posiciones 1 y 74 de la secuencia SEC ID nº 6, y que contiene las regiones CT, TMD y fragmento SR de hST6GalNAc I correspondiente a las secuencias SEC ID nº 50, 92 y 130, respectivamente,

- la secuencia SEC ID nº 147 delimitada por los nucleótidos situados entre las posiciones 1 y 420 de la secuencia SEC ID nº 5, que contiene las secuencias SEC ID nº 49, 91 y 133, y codificantes de la secuencia polipeptídica SEC ID nº 148 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 1 y 140 de la secuencia SEC ID nº 6, y que contiene las regiones CT, TMD y fragmento SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 134, respectivamente,

- la secuencia SEC ID nº 149 delimitada por los nucleótidos situados entre las posiciones 1 y 774 de la secuencia SEC ID nº 5, que contiene las secuencias SEC ID nº 49, 91 y 135, y codificantes de la secuencia polipeptídica SEC ID nº 150 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados

entre las posiciones 1 y 258 de la secuencia SEC ID nº 6, y que contiene las regiones CT, TMD y fragmento SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 136, respectivamente,

- la secuencia SEC ID nº 151 delimitada por los nucleótidos situados entre las posiciones 1 y 138 de la secuencia SEC ID nº 21, que contiene las secuencias SEC ID nº 65, 107 y 137, y codificantes de la secuencia polipeptídica SEC ID nº 152 correspondiente al fragmento de la región SR de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 1 y 46 de la secuencia SEC ID nº 22, y que contiene las regiones CT, TMD y fragmento SR de hST3Gal III correspondientes a las secuencias SEC ID nº 66, 108 y 138, respectivamente,

- la secuencia SEC ID nº 153 delimitada por los nucleótidos situados entre las posiciones 1 y 138 de la secuencia SEC ID nº 29, que contiene las secuencias SEC ID nº 73, 115 y 139, y codificantes de la secuencia polipeptídica SEC ID nº 154 correspondiente al fragmento de la región SR de rST3Gal III delimitada por los aminoácidos situados entre las posiciones 1 y 46 de la secuencia SEC ID nº 30, y que contiene las regiones CT, TMD y fragmento SR de hST3Gal III correspondientes a las secuencias SEC ID nº 74, 116 y 140, respectivamente,

- la secuencia SEC ID nº 155 delimitada por los nucleótidos situados entre las posiciones 1 y 237 de la secuencia SEC ID nº 33, que contiene las secuencias SEC ID nº 77, 119 y 141, y codificantes de la secuencia polipeptídica SEC ID nº 156 correspondiente al fragmento de la región SR de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 1 y 79 de la secuencia SEC ID nº 34, y que contiene las regiones CT, TMD y fragmento SR de hST8Sia II correspondientes a las secuencias SEC ID nº 78, 120 y 142, respectivamente,

- la secuencia SEC ID nº 157 delimitada por los nucleótidos situados entre las posiciones 1 y 201 de la secuencia SEC ID nº 37, que contiene las secuencias SEC ID nº 81, 123 y 143, y codificantes de la secuencia polipeptídica SEC ID nº 158 correspondiente al fragmento de la región SR de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 1 y 67 de la secuencia SEC ID nº 38, y que contiene las regiones CT, TMD y SR de hST8Sia IV correspondientes a las secuencias SEC ID nº 82, 124 y 144, respectivamente.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque los péptidos CT, TMD, SR o fragmento de SR comprendidos en el segundo ácido nucleico, son secuencias heterólogas derivadas de diferentes genes o transcritos de glucosiltransferasa natural.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el segundo ácido nucleico es la secuencia SEC ID nº 159 correspondiente a la fusión de la secuencia de nucleótidos SEC ID nº 177 que contiene las secuencias SEC ID nº 65 y 107 codificantes de las regiones CT y TMD de hST3Gal III correspondientes a las secuencias SEC ID nº 66 y 108, respectivamente, con la secuencia de nucleótidos SEC ID nº 131 codificante de la secuencia polipeptídica SEC ID nº 132 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 37 y 74 de la secuencia SEC ID nº 6, codificando dicha secuencia SEC ID nº 159 el polipéptido de fusión SEC ID nº 160 entre las regiones CT y TMD de hST3Gal III, por una parte, y el fragmento 37 a 74 de la región SR de hST6GalNAc I, por la otra.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque el segundo ácido nucleico es la secuencia SEC ID nº 179 correspondiente a la fusión de la secuencia de nucleótidos SEC ID nº 65 codificante de la CT de hST3Gal III correspondiente a la secuencia SEC ID nº 66, con la secuencia de nucleótidos SEC ID nº 119 codificante de la región TMD de hST8Sia II correspondiente a la secuencia SEC ID nº 120, y con la secuencia de nucleótidos SEC ID nº 129 codificante de la secuencia polipeptídica SEC ID nº 130 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 74 de la secuencia SEC ID nº 6, codificando dicha secuencia SEC ID nº 179 el polipéptido de fusión SEC ID nº 180 entre la región CT de hST3Gal III, la región TMD de hST8Sia II y el fragmento 36 a 74 de la región SR de hST6GalNAc I.

La invención se refiere más particularmente a un procedimiento tal como se ha definido anteriormente, caracterizado porque comprende la fusión de la secuencia SEC ID nº 43 como el primer ácido nucleico, con un segundo ácido nucleico seleccionado de entre:

- la secuencia SEC ID nº 145, que conduce a la secuencia de nucleótidos SEC ID nº 161 codificante de la proteína de fusión SEC ID nº 162 que contiene las regiones de CT, TMD y fragmento SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 130, respectivamente, unidas mediante un conector GS con el fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 147, que conduce a la secuencia de nucleótidos SEC ID nº 163 codificante de la proteína de fusión SEC ID nº 164 que contiene las regiones CT, TMD y fragmento SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 134, respectivamente, unida mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 149, que conduce a la secuencia de nucleótidos SEC ID nº 165 codificante de la proteína de fusión SEC ID nº 166 que contiene las regiones de los fragmentos CT, TMD y SR de hST6GalNAc I

correspondientes a las secuencias SEC ID nº 50, 92 y 136, respectivamente, unidas mediante un conector SR al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,

- 5 - la secuencia SEC ID nº 151, que conduce a la secuencia de nucleótidos SEC ID nº 167 codificante de la proteína de fusión SEC ID nº 168 que contiene las regiones de los fragmentos CT, TMD y SR de hST3Gal III correspondientes a las secuencias SEC ID nº 66, 108 y 138, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,
 - 10 - la secuencia SEC ID nº 153, que conduce a la secuencia de nucleótidos SEC ID nº 169 codificante de la proteína de fusión SEC ID nº 170 que contiene las regiones de los fragmentos CT, TMD y SR de ST3Gal III de rata correspondientes a las secuencias SEC ID nº 74, 116 y 140, respectivamente, unidas mediante un conector SR al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,
 - 15 - la secuencia SEC ID nº 155, que conduce a la secuencia de nucleótidos SEC ID nº 171 codificante de la proteína de fusión SEC ID nº 172 que contiene las regiones de los fragmentos CT, TMD y SR de hST8Sia II correspondientes a las secuencias SEC ID nº 78, 120 y 142, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,
 - 20 - la secuencia SEC ID nº 157, que conduce a la secuencia de nucleótidos SEC ID nº 173 codificante de la proteína de fusión SEC ID nº 174 que contiene las regiones de los fragmentos CT, TMD y SR de hST8Sia IV correspondientes a las secuencias SEC ID nº 82, 124 y 144, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,
 - 25 - la secuencia SEC ID nº 159, que conduce a la secuencia de nucleótidos SEC ID nº 175 codificante de la proteína de fusión SEC ID nº 176 que contiene las regiones de los fragmentos CT, TMD y SR de hST3Gal III correspondientes a las secuencias SEC ID nº 66 y 108, respectivamente, y el fragmento 37 a 74 de la región SR de hST6GalNAc I correspondiente a la secuencia SEC ID nº 132, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I,
 - 30 - la secuencia SEC ID nº 179, que conduce a la secuencia de nucleótidos SEC ID nº 181 codificante de la proteína de fusión SEC ID nº 182 que contiene la región CT de hST3Gal III, la región TMD de hST8Sia II y el fragmento 36 a 74 de la región SR de hST6GalNAc I, correspondientes a las secuencias SEC ID nº 66, 120 y 130, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I.
 - 35 La invención también se refiere a secuencias génicas codificantes de glucosiltransferasas quiméricas tales como las obtenidas siguiendo el procedimiento definido anteriormente.
- La invención se refiere más particularmente a secuencias génicas tal como se ha definido anteriormente, seleccionadas de entre:
- 40 - la secuencia SEC ID nº 161 codificante de la proteína de fusión SEC ID nº 162 que contiene las regiones de los fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 130, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 161 a la fusión de la secuencia SEC ID nº 43 y la secuencia
 - 45 SEC ID nº 145,
 - la secuencia SEC ID nº 163 codificante de la proteína de fusión SEC ID nº 164 que contiene las regiones de los fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 134, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 163 a la fusión de la secuencia SEC ID nº 43 y la secuencia
 - 50 SEC ID nº 147,
 - la secuencia SEC ID nº 165 codificante de la proteína de fusión SEC ID nº 166 que contiene las regiones de los fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 136, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 165 a la fusión de la secuencia SEC ID nº 43 y la secuencia
 - 55 SEC ID nº 149,
 - la secuencia SEC ID nº 167 codificante de la proteína de fusión SEC ID nº 168 que contiene las regiones de los fragmentos CT, TMD y SR de hST6GalNAc III correspondientes a las secuencias SEC ID nº 66, 108 y 138, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 167 a la fusión de la secuencia SEC ID nº 43 y la secuencia
 - 60 SEC ID nº 151,
 - la secuencia SEC ID nº 169 codificante de la proteína de fusión SEC ID nº 170 que contiene las regiones de los fragmentos CT, TMD y SR de ST3Gal III de rata correspondientes a las secuencias SEC ID nº 74, 116 y 140,
 - 65

respectivamente, unidas mediante un conector SR al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 169 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 153,

- la secuencia SEC ID nº 171 codificante de la proteína de fusión SEC ID nº 172 que contiene las regiones de los fragmentos CT, TMD y SR de hST8Sia II correspondientes a las secuencias SEC ID nº 78, 120 y 142, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 171 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 155,

- la secuencia SEC ID nº 173 codificante de la proteína de fusión SEC ID nº 174 que contiene las regiones de los fragmentos CT, TMD y SR de hST8Sia IV correspondientes a las secuencias SEC ID nº 82, 124 y 144, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 173 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 157,

- la secuencia SEC ID nº 175 codificante de la proteína de fusión SEC ID nº 176 que contiene las regiones de los fragmentos CT, TMD y SR de hST3Gal III correspondientes a las secuencias SEC ID nº 66 y 108, respectivamente, y el fragmento 37 a 74 de la región SR de hST6GalNAc I correspondiente a la secuencia SEC ID nº 132, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 175 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 159,

- la secuencia SEC ID nº 181 codificante de la proteína de fusión SEC ID nº 182 que contiene la región CT de hST3Gal III, la región TMD de hST8Sia II y el fragmento 36 a 74 de la región SR de hST6GalNAc I, correspondientes a las secuencias SEC ID nº 66, 120 y 130, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 181 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 179.

La invención también se refiere a vectores, tales como plásmidos, constructos víricos o bacterianos, que comprenden por lo menos una secuencia génica tal como se ha definido anteriormente.

La invención también se refiere a células eucarióticas huésped de origen en levaduras, hongos, insectos, plantas, mamíferos o seres humanos, transformadas con por lo menos una secuencia génica tal como se ha definido anteriormente, utilizando por lo menos un vector indicado anteriormente.

La invención también se refiere a la utilización de por lo menos una secuencia génica tal como se ha definido anteriormente, o de un vector tal como se ha indicado anteriormente, para la transformación de células tal como se ha definido anteriormente, o a la utilización de animales transgénicos obtenidos a partir de dichas células transformadas, en el marco de la producción de proteínas recombinantes de interés.

La invención también se refiere a un método para la preparación de proteínas recombinantes de interés, que comprende la transformación de células tal como se ha definido anteriormente, con un vector que contiene por lo menos una secuencia de nucleótidos codificante de dichas proteínas recombinantes de interés.

Las proteínas recombinantes preferentes de interés que pueden prepararse según un método indicado anteriormente según la invención se seleccionan de entre hormonas, enzimas, factores de coagulación, antígenos carbohidrato/biomarcadores séricos, citocinas, factores de crecimiento, anticuerpos o receptores.

Las células huésped preferentes para la preparación de proteínas recombinantes de interés tal como se ha indicado anteriormente, se seleccionan de entre células u organismos aprobados farmacológicamente, preferentemente de entre células de roedor, mamífero o ser humano.

Descripción de las figuras

La figura 1 representa la topología membranar de las glucosiltransferasas. Las glucosiltransferasas son proteínas membranales de tipo II que incluyen una cola N-terminal citoplasmática (CT), una señal de anclaje transmembranar (TMD) seguido de una región de tallo (SR) y un gran dominio catalítico C-terminal (CD).

Figura 2: secuencia de aminoácidos de los motivos sialilo L, S y VS en 20 sialiltransferasas humanas. Los residuos aminoácidos de consenso en todas las sialiltransferasas se muestran en negrita.

La figura 3 representa la distribución esquemática de CT, TMD, SR y CD de la ST6Gal I de rata, mostrando los residuos clave, que son tyr123 y cys conservadas. CT: cola citoplasmática (9 aminoácidos), TMD: dominio transmembranar (17 aminoácidos), SR: región de tallo (70 aminoácidos), CD: dominio catalítico (307 aminoácidos); L: motivo sialilo L; S: motivo sialilo S y VS: motivo sialilo VS.

La figura 4 representa constructos de ADN para producir quimeras y para evaluar el papel de la cola citoplasmática y el dominio transmembranal en dos isoformas enzimáticas (rST6Gal I Tyr o Cys 123), respectivamente (Fenteany y Colley, 2005).

5 La figura 5 representa la alineación de secuencias de la ST6Gal I de rata y humana. Puede observarse que la secuencia N-terminal que comprende la CT(1-9) y el TMD(10-26) se encuentra totalmente conservada, mientras que la parte SR yuxtamembranal (27-89) es más variable, aunque el péptido yuxtamembrana y especialmente positivamente los residuos lisina y cisteína, se encuentran bien conservados.

10 La figura 6 representa la ruta del ácido siálico en células humanas en el contexto de la glucosilación celular global. Los enzimas que participan en este proceso secuencial son: (a y b) UDP-N-acetilglucosamina/2-epimerasa/N-acetilmanosamina quinasa (UDP-GlcNAc 2-epimerasa/ManNAc 6-quinasa). Reacciones: (a) epimerasa y (b) quinasa, c) reacción de fosfato de ácido N-acetilneuramínico sintasa (NANS SAS; fosfato de ácido N-acetilneuramínico sintasa) y reacción de la N-acetilneuraminato piruvato liasa de *Homo sapiens* (NPL, N-acetilneuraminato piruvato liasa), (d) NeuAc 9-fosfatasa, (e) citidina-5'-monofosfato de ácido N-acetilneuramínico sintetasa (CMP-NeuAc sintetasa), (f) transportador de citidina monofosfato-ácido siálico (transportador de CMP-NeuAc del Golgi), y (g) sialiltransferasas.

20 Cabe mencionar que el derivado del ácido siálico mostrado en el diagrama, denominado "Neu5Ac", se considera ampliamente la forma "humana" del ácido siálico. En todos los demás animales, con la excepción del pollo, existe una etapa adicional en la ruta (mostrada en este diagrama), en la que CMP-Neu5Ac se hidroxila adicionalmente y se convierte en CMP-Neu5Gc por la acción enzimática de la ácido CMP-N-acetilneuramínico hidroxilasa.

25 La figura 7 representa una vista general esquemática de los diversos constructos quiméricos sintéticos generados mediante la invención.

La figura 7A representa la construcción general de una quimera sintética que incluye el dominio sintético N-terminal etiquetado con el epítipo FLAG fusionado con el dominio catalítico mínimo (CD) mediante la adición de un sitio de restricción.

30 La figura 7B muestra la ligación entre los dominios N-terminal sintético y catalítico utilizando un sitio de restricción *BamHI*. Observar que los constructos se introducen en el vector utilizando sitios de restricción, en particular *AflIII* y *XbaI*, que son diferentes del sitio de ligación.

35 La figura 8 muestra la topología y características del vector pcDNA3.1 (+) (Invitrogen).

La figura 9 representa las construcciones del CD de FLAG-hST6Gal I y de las formas quiméricas de este CD fusionadas con varios fragmentos N-terminales de otras sialiltransferasas de longitud variable.

40 La figura 9A corresponde a FLAG-CD.

La figura 9B corresponde a hST8SialI-79/CD.

La figura 9C corresponde a hST8SialIV-67/CD.

45 La figura 9D corresponde a hST6GalNAc I-258/CD.

La figura 9E corresponde a r/hST3Gal III-46/CD.

50 La figura 10 representa el dominio catalítico mínimo digerido de hST6Gal I y su producto de PCR amplificado destinado a la utilización en cada construcción de la quimera sintética. Se cargaron las muestras en un gel de agarosa al 1,5% con el marcador de ácidos nucleicos SmartLadder (SL).

55 La figura 10A muestra el dominio catalítico de hST6Gal I digerido del vector CMV, producido a partir de clonación en el laboratorio.

La figura 10B muestra el producto de PCR del dominio catalítico mínimo en 966 pb.

60 La figura 10C muestra el producto de PCR concentrado del dominio catalítico mínimo de hST6Gal I con 5 µl cargados en el gel.

La figura 11 muestra un gel de agarosa (al 2%) que muestra la banda de ADN de la región N-terminal sintética reconstituida de hST3Gal III. El carril 1 corresponde a un producto de PCR de 174 pb; el carril 2 corresponde al SmartLadder SF; el carril 3 corresponde a 5 µl del producto de PCR concentrado de 174 pb de longitud.

65 La figura 12 representa el gen de enzima reconstruido total de hST3Gal III/CD.

En la figura 12A, un gel de agarosa (al 1,5%) muestra la banda de ADN de hST3Gal III/CD sintético reconstituido amplificado mediante PCR.

5 La figura 12B corresponde a 5 µl de los productos de PCR concentrados de aproximadamente 1.200 pb de longitud.

La figura 12C representa el producto de la digestión del vector recombinante con los enzimas de restricción *AflII* y *XbaI*, que muestra la inserción del gen quimérico (tamaño esperado: 1.200 pb).

10 La figura 13 muestra un gel de agarosa (al 2%) que muestra la banda de ADN de la región N-terminal sintética reconstituida de hST6GalNAc I-74. El carril 1 corresponde a un producto de PCR de 270 pb; el carril 2 corresponde al SmartLadder SF; el carril 3 corresponde a 5 µl de los productos de PCR concentrados de 270 pb de longitud.

15 La figura 14 representa el gen sintético reconstruido completo de hST6GalNAc I-74/CD.

La figura 14A muestra un gel de agarosa (al 1,5%) con la banda de ADN de hST6GalNAc I-74/CD sintético reconstituido amplificado mediante PCR.

20 La figura 14B muestra 5 µl de los productos de PCR concentrados de la quimera de aproximadamente 1.200 pb de longitud.

La figura 14C muestra el producto de la digestión del vector recombinante con los enzimas de restricción *AflII* y *XbaI*, que muestra la inserción del gen quimera (tamaño esperado: 1.225 pb).

25 La figura 15 muestra un gel de agarosa (al 1,5%) que muestra la banda de ADN de la región N-terminal sintética reconstituida de hST6GalNAc I-40. Carril 1: SmartLadder SF; carril 2: producto de PCR de 468 pb.

30 La figura 16 muestra un gel de agarosa (al 1,5%) que muestra la banda de ADN del gen híbrido sintético reconstituido compuesto de hST3Gal III-29/37-hST6GalNAc I-74. El carril 1 corresponde al SmartLadder SF; el carril 2 corresponde a un producto de PCR de 249 pb.

35 La figura 17 muestra un gel de agarosa (al 1,5%) que muestra la banda de ADN del gen híbrido sintético reconstituido compuesto de hST3Gal III-29/37-hST6GalNAc I-74 ligado con el CD. El carril 1 corresponde al SmartLadder SF; el carril 2 corresponde a un producto de PCR de 1.197 pb.

40 La figura 18 representa la expresión del dominio catalítico mínimo (CD) (panel A) de hST6Gal I en células CHO, basado en un doble marcaje con FITC-SNA ligante de ácido siálico unido en posición 6 y con un mAb anti-FLAG, utilizando microscopía confocal. La expresión transitoria del CD se analizó mediante marcaje anti-FLAG (panel B) y para la unión de SNA-FITC (panel C). El panel D muestra la superposición de la señal de los paneles B y C (Donadio *et al.*, 2003).

45 La figura 19 muestra la expresión de dos enzimas quiméricas: hST8SialIV-67/CD, hST8SialII-79/CD en células CHO. Las células transfectadas se marcaron doblemente con un mAb anti-FLAG (A y D) y con FITC-SNA (B y E). Se representan las superposiciones en C y F.

La figura 20 muestra la expresión de ST3GalIII/CD y ST3GalIII-hST6GalNAc I-74/CD en células CHO. Las células transfectadas se marcaron doblemente con un mAb anti-FLAG (A y D) y con FITC-SNA (B y E). Las superposiciones se representan en C y F.

50 La figura 21 muestra la expresión de hST6GalNAc I-74/CD y 258/CD en células CHO. Las células transfectadas se marcaron doblemente con un mAb anti-FLAG (A y D) y con FITC-SNA (B y E). Las superposiciones se representan en C y F.

55 La figura 22 muestra un gel de agarosa (al 2%) que muestra la banda de ADN de la región N-terminal sintética reconstituida de hST3Gal III/hST8Sia II/hST6GalNAc I. El carril 1 corresponde a un producto de PCR de 240 pb; el carril 2 corresponde al SmartLadder SF; el carril 3 corresponde a 3 µl del producto de PCR concentrado de 240 pb.

60 La figura 23 representa el producto de la digestión del vector recombinante con los enzimas de restricción *XbaI* y *BamHI*, que muestra la inserción del gen del dominio catalítico (tamaño esperado: 960 pb). La muestra se cargó en un gel de agarosa al 1,5% con el marcador de ácidos nucleicos SL.

La figura 24 muestra el gen de enzima reconstruido de hST3Gal III/hST8Sia II/hST6GalNAc I/CD.

65 En la figura 24a, un gel de agarosa (al 1,5%) muestra el producto de la digestión del vector recombinante con los enzimas de restricción *XbaI* y *AflII*, que muestra la inserción del gen quimera (tamaño esperado: 1.200 pb).

La figura 24b representa el producto de PCR obtenido tras la amplificación de una muestra miniprep, que confirma la inserción del gen quimera.

- 5 La figura 25 muestra la expresión del enzima quimérico hST3Gal III/hST8Sia II/hST6GalNAc I/CD en células CHO. Las células transfectadas se marcaron doblemente con un mAb anti-FLAG (A) y con FITC-SNA (B). Se representa la superposición en C.

10 Leyendas de las Tabla 1 a 5

Tabla 1: secuencias y números de acceso de las 20 sialiltransferasas conocidas en el ser humano.

Tabla 2: sustratos aceptores y sitios de expresión de las sialiltransferasas humanas tal como se describe en Harduin-Lepers *et al.*, 2001, y en Jeanneau *et al.*, 2003.

15 Tabla 3: distribución de los 4 dominios peptídicos compartidos por las ST humanas. Una cola citoplasmática N-terminal corta (CT; aproximadamente 10 aminoácidos), un dominio transmembranal (TMD, aproximadamente 20 aminoácidos), una región de tallo (SR), de longitud muy variable, y un dominio catalítico (CD, aproximadamente 300 aminoácidos), incluyendo los motivos sialilo L, S y VS.

20 Tabla 4: péptidos conservados propuestos de aproximadamente 31 a 85 aminoácidos identificados dentro de la región SR hipervariable. Se encuentran 5 motivos: motivo A común a ST6Gal y ST6GalNAc I, motivo B común a ST6GalNAc I y ST6GalNAc II, motivos C y D comunes a todos los ST8Sia y motivo E, presente en ST3Gal excepto en ST3Gal I y II (Donadio *et al.*, 2003).

25 Tabla 5: diseño de las cuatro parejas de cebadores utilizadas para la amplificación de PCR.

Tabla 1

SIALILTRANSFERASAS HUMANAS	Nº DE ACCESO DE GenBank	Nº DE ACCESO DE SwissProt
ST6Gal I	A17362	P15907
ST6Gal II	NM_032528	Q8IUG7
ST6GalNAc I	NM_018414	Q9NSC7
ST6GalNAc II	NM_006456	Q9UJ37
ST6GalNAc III	NM_152996	Q8NDV1
ST6GalNAc IV	NM_014403	Q9H4F1
ST6GalNAc V	NM_030965	Q9BVH7
ST6GalNAc VI	AB035173	Q969X2
ST3Gal I	L29555	Q11201
ST3Gal II	U63090	Q16842
ST3Gal III	L23768	Q11203
ST3Gal IV	L23767	Q11206
ST3Gal V	AF105026	Q9UNP4
ST3Gal VI	AF119391	Q9Y274
ST8Sia I	L32867	Q92185
ST8Sia II	U33551	Q92186
ST8Sia III	AF004668	043173
ST8Sia IV	L41680	Q92187
ST8Sia V	U91641	015466
ST8Sia VI	AJ621583	P61647

Tabla 2

SIALILTRANSFERASAS HUMANAS	Aceptor	Expresión
ST6Gal I	Gal β 1-4GlcNAc	Ubicuo
ST6Gal II	Gal β 1-4GlcNAc	Cerebro, testículos, tiroide, tejido fetal, ganglios linfáticos
ST6GalNAc I	Gal β 1-3GalNAc GalNAc α -O-Ser/Thr Sia α 2-3 Gal β 1-3GalNAc α -O-Ser/Thr	Submaxilar y mamaria Glándulas, bazo, colon
ST6GalNAc II	Gal β 1-3GalNAc-O-Ser/Thr	Glándulas mamarias lactantes, testículo
ST6GalNAc III	Sia α 2-3Gal β 1-3GalNAc α -O-Ser/Thr	
ST6GalNAc IV	Sia α 2-3Gal β 1-3GalNAc	Ubicuo
ST6GalNAc V	Sia α 2-3Gal β 1-3GalNAc, GM1b	
ST6GalNAc VI	GM1b, GT1b	
ST3Gal I	Gal β 1-3GalNAc	Ubicuo
ST3Gal II	Gal β 1-3GalNAc, GM1,	Corazón, hígado, músculo esquelético
	Asialo-GM1	Timo, nódulo linfático, apéndice, glándulas salivares, bazo
ST3Gal III	Gal β 1-3GlcNAc	Músculo esquelético, tejido fetal
ST3Gal IV	Gal β 1-4GlcNAc	
ST3Gal V	Gal β 1-4GlcNAc-O	Placenta, testículo, ovario
ST3Gal VI	Gal β 1-4Glc β -Cer	Ubicuo
	Gal β 1-4GlcNAc	Corazón, placenta, hígado y la mayoría de los demás tejidos
ST8Sia I	Sia α 2-3Gal β 1-4Glc β 1-O-Cer, GM3	Tejidos embrionarios, cerebro
ST8Sia II	SiaGal β 1-4GalNAc	
ST8Sia III	Sia2-3Gal β 1-4GlcNAc	
ST8Sia IV	SiaGal β 1-4GalNAc	Cerebro, tejidos fetales, corazón adulto, bazo, timo
ST8Sia V	GM1b, GD1a, GT1b, GD3	
ST8Sia VI	NeuAc α 2, 3(6)Gal β	

Tabla 3
ESTIMACIÓN DE LA LONGITUD DE LOS DOMINIOS

Sialiltransferasas humanas	Total	Cola citoplasmática (CT)	Dominio transmembranal (TMD)	Región de tallo (SR; +/- 10 aa)	Dominio catalítico (CD; +/- 10 aa)	Motivo sialilo L (47 aa)	Motivo sialilo S (24 aa)	Motivo sialilo VS (12 aa)
ST6Gal I	406	1-9 (9 AA)	10-26 (17 AA)	27-100 (74 AA)	101-406 (306 AA)	181-227	320-343	366-378
ST6Gal II	529	1-10 (10 AA)	11-30 (20 AA)	31-112 (82 AA)	113-529 (417 AA)	293-339	433-456	479-491
ST6GalNAc I	600	1-14 (14 AA)	15-35 (41 AA)	36-281 (246 AA)	282-600 (319 AA)	362-408	518-541	563-575
ST6GalNAc II	374	1-7 (7 AA)	8-28 (21 AA)	29-67 (39 AA)	68-374 (307 AA)	148-194	302-325	347-359
ST6GalNAc III	305	1-8 (8 AA)	9-28 (20 AA)	29-34 (6 AA)	35-305 (271 AA)	77-123	214-237	273-285
ST6GalNAc IV	302	1-6 (6 AA)	7-27 (21 AA)	28-30 (3 AA)	31-302 (272 AA)	73-119	210-233	270-282
ST6GalNAc V	336	1-8 (8 AA)	9-29 (21 AA)	30-50 (21 AA)	51-336 (286 AA)	93-139	230-253	291-303
ST6GalNAc VI	333	1-43 (43 AA)	44-59 (16 AA)	60-61 (2 AA)	62-333 (272 AA)	105-151	241-264	303-315
ST3Gal I	340	1-13 (13 AA)	14-34 (21 AA)	35-58 (24 AA)	59-340 (282 AA)	139-185	266-289	312-324
ST3Gal II	350	1-6 (6 AA)	7-27 (21 AA)	28-68 (41 AA)	69-350 (282 AA)	149-195	276-299	322-334
ST3Gal III	444	1-8 (8 AA)	9-28 (20 AA)	29-76 (48 AA)	17-444 (368 AA)	157-203	299-322	346-358
ST3Gal IV	329	1-8 (8 AA)	9-26 (18 AA)	27-36 (10 AA)	37-329 (293 AA)	117-163	258-281	307-319
ST3Gal V	362	1-5 (5 AA)	6-26 (21 AA)	27-55 (29 AA)	56-362 (307 AA)	136-182	282-305	329-341
ST3Gal VI	331	1-4 (4 AA)	5-25 (21 AA)	26-34 (9 AA)	35-331 (297 AA)	115-161	256-279	303-315
ST8Sia I	356	1-29 (29 AA)	30-48 (19 AA)	49-54 (6 AA)	55-356 (302 AA)	135-182	272-295	318-330
ST8Sia II	375	1-6 (6 AA)	7-23 (17 AA)	24-73 (50 AA)	74-375 (302 AA)	157-200	292-315	342-354
ST8Sia III	380	1-9 (9 AA)	10-33 (24 AA)	34-78 (45 AA)	79-380 (302 AA)	159-205	298-321	350-362
ST8Sia IV	359	1-7 (7 AA)	8-20 (13 AA)	21-58 (38 AA)	59-359 (301 AA)	139-185	277-300	327-339
ST8Sia V	376	1-17 (17 AA)	18-38 (21 AA)	39-80 (42 AA)	81-376 (296 AA)	93-208	298-321	344-350
ST8Sia VI	398	1-3 (3 AA)	4-24 (21 AA)	25-102 (78 AA)	103-398 (296 AA)	183-230	320-343	360-378

Tabla 4

ENZIMAS	Región conservada	Motivo A	Motivo B	Motivo C	Motivo D	Motivo E
hST6Gal I	93-165	159-165	X	X	X	X
hST6GalNAc I	259-331	310-315	324-331	X	X	X
hST8Sia IV	71-133	X	X	71-90	119-133	X
hST3Gal IV	75-106	X	X	X	X	84-91

Tabla 5

Fragmentos amplificados	Conjunto de cebadores	Tamaño del producto de PCR (pb)
hST3Gal III	5' AflII-ST3 3'ST3-BamHI	186
hST6GalNAc I-74	5' AflII-ST6 3'ST6-BamHI	270
hST6GalNAc I-140	5' AflII-ST6 3'ST6-BamHI ^{n°2}	468
hST3Gal III-29/37-hST6GalNAc-74	5' AflII-ST3 3'ST6-BamHI	246

Descripción de la invención

La invención se describe con mayor detalle en la descripción detallada a continuación.

1. Introducción

Bases de la invención:

Se han centrado en el estudio de hST6Gal I debido a que esta actividad enzimática no se encuentra presente en todas las células huésped utilizadas hasta hoy para la producción de proteínas heterólogas, especialmente en los sistemas celulares aprobados farmacológicamente. El enzima humano fue clonado en Ronin, 2001.

En contraste con la literatura, se pudo demostrar inicialmente que el enzima de longitud completa puede sialilar diferencialmente las glucoproteínas aceptoras de estructura de glicano biantenaria, triantenaria y/o tetraantenaria (Ronin, 2001) y, a partir de ello, se planteó la hipótesis de que, dentro de la estructura de transferasa, debe ejercerse un control estéril situado en la región SR hipervariable sobre el CD para regular y restringir naturalmente la especificidad enzimática. Para el propósito de producir proteínas sialiladas de interés biomédico, este control regulador debería eliminarse y, como resultado, ampliarse la especificidad. En el interior de la región conservada (93-165) de hST6Gal I, se ha identificado una secuencia hidrofóbica: 93-QVWxKDS-100. La importancia de esta secuencia ha sido estudiada mediante la delección progresiva en hST6Gal I de su parte N-terminal dentro de la región conservada recién identificada. Los mutantes por delección $\Delta 35$, $\Delta 48$, $\Delta 60$ y $\Delta 82$ muestran una eficiencia de transferencia incrementada. Las delecciones realizadas antes y después del dominio conservado (93-165 de ST6Gal I) demostraron que el mutante $\Delta 100$ pierde su actividad catalítica, mientras que $\Delta 89$ (que contiene la secuencia hidrofóbica QVWxKDS) presenta una actividad catalítica óptima. Los resultados indican claramente que esta secuencia corta resulta crucial para estimular la actividad. Esta secuencia hidrofóbica podría contribuir a cambios conformacionales locales que resultan esenciales para que el sitio activo estimule la transferencia de ácido siálico.

Se ha llevado a cabo un segundo estudio destinado a identificar la base molecular y celular que controla la preferencia de aceptor de las ST (Ronin, 2003). Debido a que resultaba difícil delinear el CD de la región hipervariable de la SR, plantearon la hipótesis de que la SR de las ST podría contener características estructurales relacionadas con la preferencia de aceptor. Se analizaron 53 enzimas animales y humanos de especificidad conocida mediante análisis bioinformático con el fin de determinar si las ST podían compartir una parte peptídica con el fin de restringir el reconocimiento de aceptor a una subfamilia de enzimas. Se ha identificado una región altamente conservada de aproximadamente 31 a 85 aminoácidos y, en el interior de esta región, se han identificado 5 motivos: el motivo A, común a ST6Gal y ST6GalNAc I; el motivo B, común a ST6GalNAc I y ST6GalNAc II, los motivos C y D comunes a todas las ST8Sia, y el motivo E, presente en ST3Gal excepto en ST3Gal I y II (Tabla 4).

Dichos 5 motivos son típicos de un subgrupo de ST que comparte una actividad catalítica similar y que de esta manera participan de la misma manera en la transferencia de ácido siálico. Se encuentran situados al final de la SR en proximidad al motivo sialilo L y pueden considerarse parte del CD como característica clave del reconocimiento de aceptor.

También resulta de interés especial el resultado de que el mutante $\Delta 89$ es muy eficiente en las células CHO y que sigue la ruta intracelular desde el Golgi temprano hasta la red trans-Golgi/trans-Golgi. El mutante $\Delta 89$ se expresa a lo largo de la ruta intracelular de las glucoproteínas y glucolípidos en los dictiosomas completos del aparato de Golgi (Ronin, 2003). Este estudio permitió delinear el CD mínimo para hST6Gal I que contenía la secuencia hidrofóbica crucial (QVWxKDS) y que desplazaba el reconocimiento de aceptor de los aceptores residentes intracelulares a los glucos conjugados de superficie celular. Estos resultados proporcionaron la base para el diseño de nuevas quimeras ancladas a membrana que podrían mostrar un tráfico intracelular similar para encontrar glucoproteínas nuevamente

sinetizadas dentro de la ruta secretoria a medida que son empaquetadas en células huésped artificiales durante su producción como fármaco.

- 5 El CD mínimo de hST6Gal I Δ 89 ha sido utilizado en la invención como "CD optimizado" de especificidad ampliada y eficiencia de transferencia incrementada y en lo sucesivo se denominará "CD".

Distribución de las sialiltransferasas en el aparato de Golgi

- 10 La localización en el Golgi de las GT no se encuentra organizada estrictamente; los enzimas no se distribuyen y aíslan en un único compartimento del aparato de Golgi, más frecuentemente se solapan y cocompartimentalizan (Colley, 1997; Berger *et al.*, 1998; Berger, 2002). Todavía no se han identificado señales de retención claras, pero los enzimas forman gradientes solapantes a través de los dictiosomas (Opat *et al.*, 2001) y se han identificado únicamente las regiones cruciales.

- 15 Se han planteado dos hipótesis respecto a los mecanismos de retención en el Golgi: i) la longitud del TMD controla la retención en el Golgi (modelo de grosor de bicapa; Bretscher y Munro, 1993; Colley, 1997; Munro, 1998), e ii) la oligomerización de las proteínas conduce a la retención en el Golgi (hipótesis de oligomerización/reconocimiento de afines; Machamer, 1991; Colley, 1997), en la que los autores postulan que los enzimas interactúan en las cisternas del Golgi formando estructuras excesivamente grandes para entrar en las vesículas de transporte (Munro, 1998).

- 20 La longitud del TMD aparentemente también resulta crucial como señal de retención. El TMD representa un factor clave en el proceso de retención; en la mayoría de casos resulta suficiente estimular una localización en el Golgi, es decir a través de los compartimentos cis, mediales y trans. El alargamiento del TMD de las ST resulta en una retención reducida y el TMD sintético (creado mediante mutagénesis) de 17 leucinas confiere retención en el Golgi, mientras que 23 leucinas no lo consiguen (Munro, 1991, 1995, 1998). Respecto a las señales de retención de las ST, participan varias regiones de un modo independiente para retener enzimas en el aparato de Golgi. El TMD con sus regiones flanqueantes resultan suficientes para la retención en el Golgi (Colley, 1997). Otros estudios sugieren que no resulta necesaria ninguna secuencia específica en el TMD para la retención en el Golgi, especialmente en presencia de secuencias flanqueantes citoplasmáticas y lumenales espaciadas apropiadas y/o SR (Colley, 1997). Se encontró que la presencia de secuencias de aminoácidos de carga negativa en las ST, particularmente en proximidad al anclaje membranal, localizaba erróneamente las proteínas en el aparato de Golgi. La presencia de epítipo FLAG o MYC en el CT interfiere en la localización en el Golgi de ST6Gal I, mientras que la utilización de una secuencia espaciadora adicional para espaciar las cargas negativas del TMD, muestra una fuerte ventaja en la retención en el Golgi (Yang *et al.*, 1996). Las secuencias de la SR por sí solas aparentemente funcionan como otro tipo de retención en el Golgi (Colley, 1997).

- 35 Probablemente existe más de un mecanismo de retención que explique la colocalización de las ST en el interior del trans-Golgi y en la red del trans-Golgi. Además, la localización de los enzimas también depende del tipo celular en el que se expresan (Colley, 1997). El TMD de las ST por sí solo podría resultar suficiente para la retención en el Golgi en las células MDCK, mientras que este mismo TMD y sus secuencias flanqueantes resultan claramente necesarias para la retención en el Golgi en las células CHO (Colley, 1997). Resulta interesante que los datos disponibles revelan que la modificación del CT, del TMD y de la SR no altera la actividad catalítica de las ST, sino que permite deslocalizarlos en el interior del aparato de Golgi (Grabenhorst y Condradt, 1999).

- 45 El enzima más estudiado para la señal de retención en el Golgi es la ST6Gal I de rata (rST6Gal I), que ha sido localizada en el trans-Golgi y en la red del trans-Golgi de los hepatocitos (Roth *et al.*, 1985; Opat *et al.*, 2001) y en los compartimentos post-Golgi en otras células (Colley, 1997). Existen dos isoformas naturales de rST6Gal I, que difieren únicamente en un aminoácido en la posición 123 en el CD. Esto se debe a un único cambio de nucleótido, de A a G, resultando en un cambio de aminoácido Tyr a Cys. No se han observado diferencias de actividad catalítica entre las dos isoformas del enzima (Chen *et al.*, 2003). La forma STcys se encuentra en el Golgi en células de expresión moderada y en ningún caso en las superficies celulares o escindido o secretado al medio de células COS-1 o HeLa, mientras que la forma STtyr se encuentra en el Golgi, a niveles bajos sobre la superficie celular y se escinde y se secreta de las célula COS-1 y HeLa (Ma *et al.*, 1997). Con el fin de explicar las dos localizaciones diferentes de las proteínas, los autores han propuesto dos hipótesis sobre el mecanismo de retención: el grosor de la bicapa y la oligomerización. En primer lugar, el análisis del alargamiento del TMD de STcys y tyr sugiere que el grosor de la bicapa no es un proceso prioritario para la retención en el Golgi de ST6Gal I. En segundo lugar, el análisis de los oligómeros revela que el 100% y el 13% de STcys y STtyr, respectivamente, son insolubles a pH 6,3 (pH de Golgi tardío), pero que a pH 8,0, ambas isoformas son solubles. Los resultados sugieren que los cambios conformacionales y formación de enlaces disulfuro en el CD representan la base para la capacidad incrementada de la STcys de formar oligómeros y su localización estable en el aparato de Golgi. La naturaleza del aminoácido 123 influye sobre la formación de los oligómeros (Chen *et al.*, 2000). Además, el proceso de oligomerización y la actividad catalítica también depende de los otros siete residuos de cisteína conservados (C24, 139, 181, 332, 359, 361 y 403). La Cys24, en el interior del TMD, resulta necesaria para la dimerización, mientras que los residuos de cisteína presentes en el CD resultan necesarios para el tráfico y la actividad catalítica. Los enzimas con Cys181 y Cys332 (en los motivos sialilo L y S, respectivamente) resultan retenidos en el retículo endoplasmático y presentan una actividad mínima o nula, respectivamente. Los enzimas con Cys359 y 361 (entre los motivos sialilo S y VS) son

enzimas inactivos sin comprometer la localización y tráfico. Los enzimas con mutaciones Cys139 ó Cys403 no producen ningún cambio en la actividad catalítica y en la localización en el Golgi. La sustitución de estos dos residuos cisteína reduce la escisión y secreción, sugiriendo que resultan necesarios para la escisión de una señal (Qian *et al.*, 2001) (figura 3).

Se han caracterizado los mutantes del enzima STtyr que comparten una delección en el SR: STtyr Δ 4 y STtyr Δ 5 (con delecciones de los aminoácidos 32 a 104 y 86 a 104, respectivamente) no son activos y/o no resultan secretados; STtyr Δ 1, Δ 2 y Δ 3 (delecionados entre los aminoácidos 32 y 86) representan formas no escindibles y muestran una expresión creciente en la superficie celular (Kitazume *et al.*, 1999).

Un estudio reciente (Fenteany y Colley, 2005) demuestra que en efecto resultan necesarias múltiples señales para la oligomerización de rST6Gal I y para la localización en el Golgi. Los autores pretendían reevaluar el papel de la CT y del TMD de las dos isoformas del enzima (Tyr o Cys 123). Prepararon varios constructos de ADN para producir quimeras tales como las representadas en la figura 4. El alargamiento del dominio transmembranal (con más de 17 aminoácidos) no intensifica la salida del Golgi, no modifica el tráfico ni la localización en el Golgi. Respecto al papel del CT y del TMD con la isoforma STcys, existe una señal compuesta de 3 residuos lisina en el CT que resulta reconocida como señal de exportación al exterior del RE. Los autores demuestran que la alteración de tanto el CT como el TMD altera el enrutamiento del enzima. La única diferencia entre las isoformas de rST6Gal I es su capacidad de formar oligómeros según el pH. En el caso de la STcys, la formación de oligómeros según el pH resulta comprometida únicamente en el caso de que se alteren el CT y el TMD. En el caso de la STtyr, incrementa ligeramente la tasa de salida del Golgi (Fenteany y Colley, 2005).

Todos estos nuevos resultados referentes a rST6Gal I pueden considerarse aplicables a la ST6Gal I humana debido a que ambos enzimas comparten una homología de 85% y a que ambos CD son prácticamente idénticos. En efecto, la alineación de las dos secuencias (figura 5) demuestra claramente que las siete cisteínas anteriormente indicadas se encuentran conservadas en hST6Gal I y que también se encuentra presente la tirosina 123.

A pesar de los considerables esfuerzos de los científicos, no ha podido identificarse ninguna señal en la secuencia de las ST o FuT. Se cree que un proceso multifactorial podría enlentecer los enzimas en su camino a la superficie para direccionarlos a los compartimentos correctos, que son la red del trans-Golgi en el caso de ST6Gal. Por consiguiente, los inventores plantearon la hipótesis de que el CT, el TMD y la SR podrían intercambiarse alternativamente para anclar un CD definido en las células huésped que no presentan el gen relevante y no expresan la actividad de transferasa relevante. De esta manera, resulta posible, por lo tanto, generar un amplio abanico de combinaciones debido a que estas 3 partes pueden seleccionarse de entre 55 genes animales conocidos. Cada uno de ellos potencialmente puede fusionarse con 20 dominios catalíticos diferentes conocidos para expresar la capacidad de transferir ácido siálico. Lo anterior resulta claramente de interés para la humanización de levaduras (*Pichia pastoris* o *Schizosaccharomyces piombae*), plantas e insectos que no expresan estas actividades y que, sin embargo, se utilizan para producir fármacos recombinantes. Aunque sin respaldo experimental, este argumento también podría aplicarse perfectamente a FuTs, GalT, GalNAcT y GlcNAc, debido a que controlan la biosíntesis de glicotopos de elevada relevancia clínica.

Importancia de algunos residuos

Tal como se ha indicado anteriormente, se ha identificado un cuarto motivo (motivo 3) en la familia de las ST, entre los motivos S y VS. Está compuesto de cuatro aminoácidos altamente conservados con la secuencia de consenso siguiente: (H/y)Y(Y/F/W/h)(E/D/q/g). La mutagénesis sitio-dirigida en hST3Gal I ha demostrado la importancia funcional de dos aminoácidos en este motivo: His299 y Tyr300. Los resultados sugieren la importancia de los residuos aromáticos, su posible participación en el reconocimiento de aceptor y su contribución a una eficiencia catalítica óptima. En particular la Tyr300 invariante desempeña un papel conformacional importante. El análisis de mutaciones ha demostrado que los mutantes H299A e Y300A no muestran actividades catalíticas, mientras que el mutante Y300F restaura parcialmente la actividad (Jeanneau *et al.*, 2004). Este motivo también se encuentra presente en el dominio catalítico de hST6Gal I. Además, varios residuos cisteína resultan de gran importancia en la dimerización y actividad catalítica tal como se ha indicado anteriormente (Qian *et al.*, 2001). Se ha identificado por lo menos un enlace disulfuro entre dos residuos cisteína conservados en el interior de los motivos sialilo L y S. Este enlace resulta esencial para mantener la conformación de la proteína. Las mismas observaciones han revelado la existencia de esta unión en la PST (Angata *et al.*, 2001). Además, existe un segundo enlace disulfuro entre los motivos sialilo y la región C-terminal. Se ha demostrado la dimerización de ST6Gal I, mediante un enlace disulfuro. Aproximadamente 20% a 30% de este enzima en el interior del aparato de Golgi se encuentra en forma de dímero, que presenta una menor afinidad debido a su menor afinidad para CMP-NeuAc (Jeanneau, 2003).

"Autoglucosilaciones"

Las GT mismas con frecuencia se encuentran glucosiladas. Se dispone de pocos datos sobre el estado de glucosilación de las ST y su importancia para la función biológica de los enzimas. Aparentemente la N-glucosilación en ST6Gal I no resulta necesaria para la actividad biológica del enzima *in vivo*. Por otra parte, en ensayos *in vitro* con mutantes, únicamente se observa actividad enzimática para la proteína mutada en el primer sitio de

glucosilación. Estos resultados sugieren la existencia de dos N-glicanos (Asn146 y Asn158) que pueden estabilizar y/o prevenir la degradación de la proteína. La mutación en el segundo sitio de glucosilación puede conducir a la agregación o degradación de la proteína (Chen y Colley, 2000; Chen *et al.*, 2000). Sin embargo, la N-glucosilación en ST8Sia I puede afectar a su actividad y a su localización subcelular (Martina *et al.*, 1998). La eliminación de los N-glicanos redujo su actividad *in vivo* a menos del 10% de la actividad inicial. Recientemente se ha demostrado que ST8Sia II y ST8Sia IV presentan, además de su actividad clásica, una actividad de autopolisialilación (Muhlenhoff *et al.*, 2001). Esta autopolisialilación se produce en los N-glicanos en la posición Asn 74 en la PST y en las posiciones Asn 69 y 219 en la STX. La mutagénesis sitio-dirigida en estos sitios inactiva el enzima tanto *in vivo* como *in vitro*. En consecuencia, se ha procurado mantener los sitios de glucosilación en Asn 74 y en Asn 69 al utilizar la SR de ST8Sia IV y ST8Sia II, respectivamente, para construir genes quiméricos.

II. Utilización de las secuencias de la invención para introducir nueva actividad de transferasa en sistemas de expresión

Ila. Papel de los glicanos y su importancia en el sistema inmunológico

Los glicanos son señales de reconocimiento en la mayoría de organismos vivos, aunque también son determinantes estructurales clave para la potencia biológica de las proteínas.

Desempeñan un papel crucial en el plegamiento y oligomerización de las proteínas, el control de calidad, la clasificación y el transporte (Helenius y Aebi, 2001). Los glicanos representan epítomos oligosídicos que pueden considerarse antígenos carbohidrato. La naturaleza de los glicanos modifica la inmunorreactividad. Existen varias familias de antígenos oligosídicos: los determinantes de grupo sanguíneo ABH, los grupos de Lewis de tejidos, los miembros antígenos de la familia del antígeno T y los antígenos específicos de adhesión celular polisialilados o sulfatados (Legaigreur *et al.*, 1999). Los glicanos presentan múltiples funciones: también participan en mecanismos de interacciones entre células y entre células y matriz. Algunas proteínas celulares denominadas lectinas reconocen específicamente las estructuras glicánicas y actúan como receptores específicos (Gabius *et al.*, 2004). En la mayoría de casos, son componentes de glucoconjugados de superficie celular.

Los grupos carbohidrato actúan como señales de reconocimiento en el sistema inmunológico y influyen sobre el reconocimiento inmunológico de por lo menos dos maneras: i) alteran la conformación de la proteína y modulan su función biológica, ii) los oligosacáridos actúan de determinantes de reconocimiento. Particularmente, los ácidos siálicos contribuyen en gran medida a ambos efectos debido a que su naturaleza altamente electronegativa e hidrofílica puede influir sobre la conformación de las macromoléculas sialiladas, y ello resulta particularmente relevante para las glucoproteínas. Los ácidos siálicos presentan la capacidad de actuar como máscaras biológicas que evitan o reducen la accesibilidad y desempeñan un papel significativo sobre la superficie celular en el proceso de reconocimiento de auto/no autorreconocimiento (Pilatte *et al.*, 1993; adición que debe realizarse: Glycobiology 2006).

Debido a que los N-glicanos se encuentran bien representados en las glucoproteínas séricas y en muchas otras proteínas de tejidos del cuerpo humano, resulta altamente deseable que no sean antigénicos al encontrarse presentes en fármacos glucosilados recombinantes. Sin embargo, debido a que la HuEPO producida en las células CHO se ha desarrollado como primer fármaco aprobado, aparentemente varios azúcares terminales podrían ser antigénicos en el ser humano: poliLacNAc, alfa1,3-Gal y ácido N-glicolilneuramínico. Según los organismos reguladores más recientes, no deberían encontrarse presentes en el futuro en los terapéuticos recombinantes.

IIb. Eliminación

Aparte de su influencia sobre las propiedades físicoquímicas y sobre las funciones biológicas de las proteínas, los glicanos también presentan una papel prevalente sobre la duración de las glucoproteínas en la sangre. Este fenómeno se denomina eliminación metabólica y representa la tasa a la que un compuesto natural o fármaco resulta eliminado del cuerpo por el hígado o el riñón. Se define como el volumen plasmático purificado por unidad de tiempo (ml/minuto). Este mecanismo permite que el organismo elimine los fármacos.

La mayor parte de las proteínas plasmáticas circulantes pertenecen a la familia de las glucoproteínas que comparten oligosacáridos N-ligados y la totalidad de ellos terminan en ácido siálico, particularmente en el ser humano. La eliminación del ácido siálico por parte de una neuraminidasa reduce drásticamente la duración en el plasma de las glucoproteínas séricas y estimula su incorporación por el hígado (van den Hamer *et al.*, 1970; Morell *et al.*, 1971). Por ejemplo, en el caso de que la hormona gestante (más recientemente la HuEPO recombinante) no presente ácido siálico, su vida media se reduce a 2 minutos, mientras que habitualmente su vida es de aproximadamente 48 horas. Es ampliamente conocido en la actualidad que el ácido siálico resulta necesario para mantener las glucoproteínas en la sangre debido a que evita la eliminación de las proteínas circulantes por parte de un receptor sialoglucoproteína (ASGPR o lectina hepática) compuesto de dos subunidades e incluido en las membranas de los hepatocitos (Hudgin *et al.*, 1974; Kawasaki y Ashwell, 1976; Bianucci y Chiellini, 2000). Este receptor se une a galactosa o residuos N-acetilglucosamina de los N-glicanos desialilados (Meier *et al.*, 2000). Las glucoproteínas reconocidas seguidamente

se internalizan en vesículas endocíticas cubiertas de clatrina y se redirigen al interior de los lisosomas, en donde resultan degradados (Ashwell y Hardford, 1982).

Este etapa aparentemente resulta crucial para las propiedades farmacocinéticas de las glucoproteínas recombinantes terapéuticas, debido a que los organismos inferiores no pueden sialilar proteínas y en las células CHO, la adición de ácido siálico es extremadamente sensible al estado energético de las células en cultivo (se sintetiza NeuAc mediante condensación de piruvato y ácido láctico) y con frecuencia no es completa. Todos los fármacos superventas sialilados (EPO, IFN, GM-CSF, FSH, Ab...) deberían purificarse para que el fabricante presente un elevado contenido de ácido siálico y controlarse para la consistencia entre lotes.

IIC. Control de la glucosilación y alergenidad

Los residuos de ácido siálico en la posición terminal de los N-glicanos resultan de gran importancia para las proteínas terapéuticas debido a que la sialilación de las proteínas confiere importantes propiedades a las glucoproteínas. La maquinaria requerida para la síntesis, activación e introducción de los residuos sialilados se encuentra pobremente representada en los diferentes sistemas de expresión de proteínas recombinantes. Además, las proteínas humanas recombinantes producidas con frecuencia se encuentran infrasialiladas o incluso no sialiladas en absoluto, en comparación con sus contrapartidas nativas. Adicionalmente, en el caso de que se expresen en células de mamífero, la sialilación puede producirse a través del ácido N-glucolilneuramínico (NeuGc), que difiere significativamente del derivado N-acetilado (NeuAc) debido a que se encuentra presente en las células de mamífero pero no en células humanas.

Debido a que existe un problema importante con el ácido siálico en el mercado de los fármacos recombinantes, muchos equipos de investigación están trabajando en la modificación del proceso de N-glucosilación en los diferentes sistemas de expresión existentes. Se está realizando un trabajo exhaustivo destinado a la humanización del patrón de glucosilación de las proteínas recombinantes para aproximarse al patrón observado en las glucoproteínas naturales al máximo y para mejorar la farmacocinética así como la seguridad del producto. Simultáneamente, el procedimiento tenderá a reducir la presencia de determinantes inmunogénicos. A este respecto, un estudio muy reciente ha confirmado que el ácido siálico unido en posición 6 y el ácido siálico unido en posición 3 evitan que las glucoproteínas se unan a receptores del sistema inmunológico y generen activación (Glycobiology, 2006). Por lo tanto, la adición de ácido siálico unido en posición 6 a proteínas glucosiladas recombinantes presentaría un gran beneficio en la seguridad del fármaco y sería un gran avance en el campo.

En primer lugar se extrajeron las proteínas de interés terapéutico a partir de fuentes naturales tales como sangre, placenta o tejidos humanos o animales. Sin embargo, este enfoque se encuentra limitado por la cantidad de los tejidos humanos disponibles y podría comportar riesgos importantes de contaminación (virus, priones, oncogenes) y/o generar reacciones alérgicas debidas a la presencia de trazas de proteínas animales o toxinas. Con los avances de la biología molecular, se han desarrollado nuevos enfoques para la producción de proteínas utilizando sistemas de expresión bastante diferentes. Se encuentra disponible un amplio abanico de sistemas de expresión heterólogos (Andersen y Krummen, 2002) y cada uno de ellos presenta ventajas y desventajas, con atención particular al patrón de N-glucosilación de las proteínas recombinantes producidas en estos sistemas.

IId. Biosíntesis de N-glicanos en los sistemas de expresión conocidos: necesidades para la humanización de la ruta de glucosilación

Bacterias

Uno de los sistemas más utilizados es la bacteria *Escherichia coli* (*E. coli*) (Swartz, 2001; Baneyx, 1999), pero su inconveniente principal es que las modificaciones postraduccionales humanas, particularmente las glucosilaciones, no son llevadas a cabo por este procariota debido a que no se expresa este tipo de glucosiltransferasa en *E. coli*. Esto puede conducir a un plegamiento incorrecto y al posterior rechazo de las proteínas terapéuticas de interés por parte del sistema inmunológico, a la reducción de su tiempo de vida y de su actividad biológica.

Hasta el momento no se ha apreciado N-glucosilación del tipo humano en dicho microorganismo.

Levaduras y hongos filamentosos

Las levaduras y hongos filamentosos también son sistemas de expresión eucarióticos bien establecidos y presentan una maquinaria celular similar a la de las células humanas. Las levaduras producen proteínas complejas y llevan a cabo varias modificaciones postraduccionales, incluyendo glucosilaciones. Tanto las levaduras como los hongos típicamente producen glicanos ricos en manosa mediante la adición de hasta 100 residuos de manosa (levaduras) sobre el núcleo pentasacárido (Tanner y Lehele, 1987; Herscovics y Orlean, 1993). Esta hipermanosilación definitivamente induce una respuesta inmunológica en el ser humano. Además, hasta hoy, no se ha encontrado ningún oligosacárido complejo que contenga ácido siálico, galactosa, fucosa y N-acetilgalactosamina en el interior de las glucoproteínas producidas por estos organismos (Blanchard, 2004).

El proceso de N-glucosilación llevado a cabo por las levaduras y hongos es similar al proceso en mamífero con respecto a las etapas iniciales en el RE, pero la presencia de polimananos añadidos en el aparato de Golgi evita su aprobación por parte de los organismos reguladores.

5 Células de insecto

Las proteínas expresadas en las células de insecto se pliegan correctamente, se secretan y pueden recibir modificaciones postraduccionales. Las modificaciones postraduccionales llevadas a cabo por estas células son similares a las llevadas a cabo por las células de mamífero, en particular la N-glucosilación de la proteína. Las estructuras de glicano obtenidas en este caso, sin embargo, son incompletas y se denominan paucimanosa debido a la presencia de una actividad de N-acetilglucosaminidasa no deseable que degrada las nuevas glucoproteínas expresadas durante la expresión baculovírica (Blanchard, 2004).

La falta de ácido neuramínico en los insectos todavía es objeto de activo debate (Marchal *et al.*, 2001; Leuroge *et al.*, 2005). En algunas líneas celulares de insecto, se encuentran residuos de fucosa unidos en $\alpha 1,3$, y ello puede inducir una respuesta inmunológica en el ser humano. En la actualidad, por lo tanto, se restringe la utilización de este sistema a la producción de antígenos vacunales.

20 Plantas transgénicas

Al igual que con otras células eucarióticas, las células vegetales muestran una compleja y sofisticada maquinaria celular que puede utilizarse para producir proteínas terapéuticas. Las proteínas recombinantes presentan una calidad farmacológica muy buena debido a que las plantas expresan los enzimas necesarios para la maduración de las proteínas.

Sin embargo, el proceso de glucosilación requiere varios ajustes para no producir proteínas alergénicas. En efecto, las N-glucosilaciones en plantas (Lerouge *et al.*, 2000) son similares a las realizadas en el ser humano en cuanto a la glucosilación nuclear. Las glucosilación todavía no presentan antenas sialiladas y se produce una adición de $\beta 1,2$ -xilosa y $\alpha 1,3$ -fucosa. Ambos residuos son altamente inmunogénicos para el ser humano y en la actualidad comprometen considerablemente la aprobación de las plantas transgénicas como sistemas de expresión para terapéuticos.

Células de mamífero

El sistema de expresión de las células de mamífero, es decir el de las células CHO, en la actualidad es el único sistema farmacológicamente aprobado para la producción de terapéuticos recombinantes. Estas células muestran una ventaja importante debido a que pueden sintetizar proteínas complejas, conseguir una N-glucosilación de tipo complejo de alta masa molecular. Las células de mamífero utilizan naturalmente no sólo los enzimas que participan en la síntesis y transporte de los azúcares de nucleótido, sino también las glucosiltransferasas requeridas para garantizar una glucosilación compleja de las proteínas recombinantes heterólogas con un elevado nivel de $\alpha 2,3$ -sialilación. Sin embargo, existe una ausencia de otros enzimas, tales como las $\alpha 1,3/4$ -fucosiltransferasas y las $\alpha 2,6$ -sialiltransferasas, que transfieran motivos glucosídicos específicos de los N-glicanos de tejidos humanos. En los roedores, el ácido N-glucolilneuramínico es sustancialmente preferente en los N-glicanos; la O-acetilación del ácido N-acetil-neuramínico en las posiciones 4, 7, 8 y con más frecuencia, 9, también se produce con frecuencia, y cada uno de estos derivados potencialmente puede ser inmunogénico en el ser humano. Esto representa una limitación a la utilización de las células de ratón o de ratones/conejos lactantes para la expresión de proteínas recombinantes (Blanchard, 2004).

50 Ile. Manipulación mediante glucosilación

Tal como se ha expuesto anteriormente, la sialilación terminal de las glucoproteínas aprobadas es la etapa más difícil de llevar a cabo en todos los sistemas de expresión disponibles en la actualidad. A título de ejemplo, no ha podido añadirse ácido siálico unido en $\alpha 2,6$, aunque es el enlace prevalente en la sangre humana. En todos los organismos simplemente no existe la maquinaria enzimática necesaria (figura 6).

El objetivo final de la investigación en este campo consiste en mejorar el proceso de glucosilación, el rendimiento productivo y la calidad de las proteínas recombinantes. La utilización de sistemas de expresión modificados genéticamente para la glucosilación permitiría producir glucoproteínas con una homogeneidad excepcional de sus estructuras glicánicas. En este caso podrían utilizarse estos sistemas para desarrollar un nivel elevado de producción de proteínas de interés biomédico de estructuras bien definidas. La manipulación mediante glucosilación en las levaduras (Hamilton *et al.*, 2003; Roy *et al.*, 2000) inicialmente ha impedido la adición de cadenas polimanosídicas. Los enzimas que resultan necesarios para la galactosilación y N-acetilglucosaminilación han sido añadidos posteriormente a estos sistemas (Maras *et al.*, 1999; Bretthauer, 2003; Vervecken *et al.*, 2004). Sin embargo, la adición del ácido siálico terminal todavía es difícil de llevar a cabo debido al número y localización de los enzimas que participan en el proceso, aunque en la actualidad se está realizando en Japón (figura 6).

En las células de insecto, se han eliminado las GT que catalizan la transferencia de azúcares inmunogénicos, mediante la adición de 3 genes, codificantes de la N-acetilglucosamina-2-epimerasa, N-acetilneuraminil-lasa y la CMP-Neu5Ac-sintasa (Jarvis *et al.*, 1998; Aumiller *et al.*, 2003). Los autores han creado nuevas líneas celulares de insecto (SfSWT-3) diseñadas para sintetizar su propio CMP-ácido siálico. Las células resultantes expresan la totalidad de los 7 genes de mamífero necesarios para formar un N-glicano, producir CMP-ácido siálico y sialilar una proteína recombinante en un cultivo en un medio sin suero (Aumiller *et al.*, 2003). Este trabajo ha sido patentado.

En plantas, la primera estrategia utilizada presentaba el objetivo de evitar la adición de azúcares alergénicos mediante almacenamiento de proteínas en el RE. Aunque en este caso los glicanos no pueden ser del tipo complejo. Otra estrategia se basa en la inhibición de varias GT en el interior del aparato de Golgi. Esta inhibición puede completarse y/o puede entrar en competencia con la maquinaria endógena de maduración. También se está llevando a cabo la adición de la maquinaria de sialilación.

En el caso de las células de mamífero, se ha llevado a cabo la sobreexpresión de una $\alpha 2,3$ -ST y una $\beta 1,4$ -Gal (Weikert *et al.*, 1999); estos dos enzimas se encuentran presentes en el genoma pero sus actividades varían dependiendo de las condiciones de cultivo. Esto conduce a una amplia variabilidad respecto a la presencia de Gal y ácido siálico terminales, además de una extensa microheterogeneidad en las estructuras de glicano de las proteínas secretoras.

La investigación se ha orientado mayoritariamente hacia la optimización de la galactosilación y sialilación (Granbenhorst *et al.*, 1999) mediante la introducción de una $\alpha 2,6$ -ST (Bragonzi *et al.*, 2000). Se ha establecido una línea celular CHO que expresa establemente una $\alpha 2,6$ -ST de longitud completa desde hace mucho tiempo, aunque también ha proporcionado resultados decepcionantes. La proporción observada entre los residuos terminales de ácido siálico $\alpha 2,6$ y $\alpha 2,3$ -ligados transportados era de 40,4% de residuos de ácido $\alpha 2,6$ -siálico y 59,6% de ácido $\alpha 2,3$ -siálico, mejorando la farmacocinética en estudios de eliminación (Bragonzi *et al.*, 2000). A pesar de la mejora de la humanización de las células, no puede controlarse la proporción entre los residuos terminales de ácido $\alpha 2,6$ -siálico y ácido $\alpha 2,3$ -siálico y ni siquiera resulta favorable para la actividad en posición 6.

En resumen, la sialilación es un etapa crítica en el control de las estructuras de glicano en las proteínas producidas mediante ingeniería genética. La dificultad principal reside no tanto en expresar la actividad faltante, sino en conseguir que el sistema heterólogo de producción funcione correctamente. En efecto, deben alcanzarse 3 objetivos principales para humanizar la glucosilación de los sistemas de expresión actuales: 1) conseguir el funcionamiento del sustrato donador y el transportador relevante necesario para este enzima dentro del Golgi, 2) conseguir que el enzima correctamente compartimentalizado finalmente compita con la actividad endógena de sialiltransferasa, 3) conseguir que el enzima pueda catalizar la transferencia de ácido siálico a cualquier sustrato aceptor relevante.

IIf. Antecedentes de la invención:

La estrategia se basa en la consideración de que el enzima ST6 de longitud completa hasta el momento no ha podido cumplir los objetivos anteriormente indicados, especialmente la competición con la actividad endógena de ST3 existente en la mayoría de células de mamífero, probablemente debido a que el producto génico no se inserta en compartimentos intracelulares implicados en la ruta secretoria. De esta manera, los inventores desarrollaron un procedimiento que pudiese proporcionar una oportunidad óptima para que el CD de una transferasa alcance los compartimentos del Golgi, en los que las neoglicoproteínas se clasifican en vesículas secretorias y/o en donde entran en contacto con la ST manipulada antes/en sustitución de la ST3 existente.

III. Diseño de ST membranales sintéticas

La presente invención consiste en un procedimiento que proporciona un panel de ST quiméricas de actividad catalítica conocida, que presenta un anclaje membranar para dirigir las al aparato de Golgi de las células eucarióticas. Ninguna de ellas existe en las células vivas y se consideran "ST sintéticas" debido a que su construcción no requiere ninguna de las secuencias de ADN codificantes de la CT, el TMD o el SR. Los oligonucleótidos relevantes presentan menos de 200 pb (60 aminoácidos) de longitud y pudieron diseñarse informáticamente y obtenerse comercialmente. Mediante la utilización de métodos de PCR, se construyó un anclaje sintético etiquetado que codificase la mitad N-terminal de la proteína ST y se fusionó con el dominio catalítico optimizado de hST6Gall.

La invención describe diversos métodos moleculares para construir 3 tipos de ST quiméricas membranales:

- *constructo homólogo*: anclaje de ST fusionado con el CD Δ 89 de ST6Gal sintético o finalmente copiado mediante PCR. En este caso, las porciones CT+TMD+SR son de la misma ST: ST3Gal/ST6GalNAc/ST8Sia. El ejemplo describe una secuencia de 200 pb.
- *constructo híbrido*: CT+TMD son de una ST y la SR de otra ST. El tamaño mínimo también es de 200 pb.
- *constructo heterólogo*: los fragmentos CT, TMD y SR proceden de diferentes ST. El tamaño mínimo también es de 200 pb.
- *constructo largo*: se preparó un constructo CT+TMD+SR1 de 200 pb y se fusionó con un segundo constructo SR2 de como máximo 200 pb de manera que SR2 estuviese cadena abajo de SR1. El constructo duplicado puede ligarse adicionalmente en el extremo 3' con cualquier otro constructo de SR y el número de repeticiones que resulte necesario.

El tamaño de 200 pb se basa en la observación de que la parte N-terminal más corta de todas las ST humanas es ST6GalNAcIII con CT+TMD+SR=200 pb. Este enzima prácticamente no presenta tallo.

En la actualidad, se ha validado un método para construir un fragmento sintético de ADN y fusionarlo con un dominio catalítico mínimo con el fin de generar un nuevo enzima transferasa de la actividad deseada. Se encuentran disponibles varios constructos representativos para la expresión adicional. Las quimeras sintéticas son activas y se están estudiando en células CHO utilizando la expresión transitoria y la microscopía confocal.

Ejemplos

Ejemplo 1: producción de dominio catalítico mínimo (CD) de hST6Gal I (SEC ID nº 43 codificante de SEC ID nº 44)

Se utilizó la ST6Gal I humana (hST6Gal I) clonada por Legaigreur *et al.*, 2001, para amplificar el CD mínimo clonado en el vector pFLAG-CMV (Sigma) (Donadio *et al.*, 2003), que se utilizó en todos los constructos quiméricos.

En primer lugar, se digirió el vector recombinante para verificar la presencia del CD de hST6Gal I. La electroforesis en gel de agarosa reveló la presencia de una banda de nucleótidos aproximadamente en 1.000 pb, correspondiente al tamaño esperado para el dominio catalítico mínimo (figura 10 A-C).

El CD correspondiente a los aminoácidos 90 a 406 de hST6Gal I se obtuvo mediante PCR utilizando los cebadores siguientes:

5'*Bam*HI- Δ 4: 5'-GAGCCCGGATCCGAGGCCTCCTTC-3', y
3' Δ 4-XbaI: 5'TAACCTCTAGATTAGCAGTGAATGGTCCGGAAGC-3'.

Dichos cebadores contienen los sitios de restricción *Bam*HI (5'*Bam*HI- Δ 4) y *Xba*I (3' Δ 4-XbaI) y el codón de parada natural de la secuencia. Se utilizó el sitio de restricción *Bam*HI para ligar el CD de hST6Gal 90-406 a la parte N-terminal sintética de otras ST para formar enzimas quiméricos. Se utilizó el sitio de restricción *Xba*I en el extremo 3' para introducir la quimera en el interior del vector pcDNA3.1 (Invitrogen) (figura 11).

Se llevó a cabo una PCR en un aparato I-Cycler (Biorad) utilizando 2,5 unidades de ADN polimerasa ProofStart (Qiagen), 300 μ M de cada dNTP (Sigma Aldrich), 1 μ M de cada cebador, 1X de tampón del fabricante de ProofStart (Qiagen), 1X de solución Q del fabricante (Qiagen) y 1,5 mM de Mg^{2+} en un volumen final de 50 μ l con 200 ng de pFlag-CMV-hST6Gal I-90-406. Se llevó a cabo la reacción de la manera siguiente: 95°C durante 5 minutos, seguido de 40 ciclos de 30 segundos a 94°C, 30 segundos a 55°C y 1 minuto a 72°C. Se analizó el producto de PCR amplificado mediante electroforesis en un gel de agarosa al 1,5% (amplificación de un producto de PCR en el tamaño esperado de 1.000 pb; figura 17B) y se purificó utilizando el kit de extracción en gel de Qiagen siguiendo las instrucciones del fabricante.

Se agruparon varios productos de amplificación de PCR y se concentraron para estimar la cantidad de ADN en un gel de agarosa (figura 10C). Los productos de PCR purificados se mantuvieron a -20°C hasta la utilización.

Ejemplo 2: ensamblaje de dominios N-terminales sintéticos

Para crear un amplio abanico de ST según la clasificación realizada mediante análisis bioinformático, se sintetizaron y ensamblaron la CT, el TMD y la SR de diferentes ST.

Todas las partes N-terminales seleccionadas de la quimera contenían 3 regiones típicas de la familia de las ST: la CT, el TMD y la SR. Cabe indicar que la SR puede presentar una longitud variable y podría originarse de la duplicación génica sintética. Todos estos fragmentos, es decir, CT+TMD+SR se reconstituyeron por completo utilizando las etapas de hibridación de complementarios y fosforilación tal como se describe posteriormente.

a. Construcción del dominio no catalítico de hST3Gal III (SEC ID nº 151 codificante de SEC ID nº 152)

Se diseñaron y sintetizaron (Eurogentec) seis oligonucleótidos de sentido y cinco oligonucleótidos antisentido correspondientes a parte de la región N-terminal (1 a 138 pb) de hST3Gal III. El presente ejemplo muestra el método para sintetizar una secuencia natural de interés. Se añadió un epítipo FLAG (en negrita y subrayado) entre el codón de inicio ATG y el segundo codón (GGA) de la secuencia de hST3Gal III:

ATG **GACTACAAAGACGATGACGACAAG** GGA.

Se fosforiló un microgramo de cada nucleótido interno utilizando 1 unidad de polinucleótido quinasa (Eurogentec) en un tampón quinasa que contenía 2 mM de ATP (Sigma). Se llevó a cabo la reacción durante 60 minutos a 37°C, y la incubación finalmente se inactivó durante 10 minutos a 65°C. Todos los oligonucleótidos se desnaturalizaron separadamente durante 10 minutos a 80°C y después se mezclaron para encontrar los complementarios. Este procedimiento se llevó a cabo durante la noche con un gradiente decreciente de temperaturas de 80°C a 20°C. A continuación, el fragmento resultante se sometió a PCR.

b. Construcción del dominio no catalítico de hST6GalNAc 1-74 (SEC ID nº 145 codificante de SEC ID nº 146)

Uno de los tallos más cortos dentro de la familia de las ST es el tallo de hST6GalNAc III, que está compuesto de aproximadamente 6 aminoácidos. Un posible dominio N-terminal de esta ST presenta 74 aminoácidos (CT, TMD y SR), que corresponden a 222 pb. Por lo tanto, con el fin de delinear un constructo con la SR más corta, se reconstituyó un dominio N-terminal sintético de 222 pb correspondiente a los primeros 74 aminoácidos de hST6GalNAc I.

Se diseñaron y sintetizaron (Eurogentec) nueve oligonucleótidos de sentido y ocho oligonucleótidos antisentido correspondientes a la secuencia de la región N-terminal completa (1 a 222 pb) de hST6GalNAc I. Se añadió una secuencia FLAG (en negrita y subrayada) entre el codón de inicio ATG y el segundo codón (GGA) de la secuencia de hST6GalNAc I:

ATG **GACTACAAAGACGATGACGACAAG** GGA.

Las reacciones de fosforilación y de apareamiento se llevaron a cabo tal como se ha indicado anteriormente.

c. Duplicación de oligonucleótidos sintéticos para ensamblar el dominio no catalítico hST6GalNAc I 1-140 (SEC ID nº 148)

Se seleccionó hST6GalNAc I como la sialiltransferasa que mostraba la SR más larga (246 aa). En el presente ejemplo, la longitud de la parte CT+TMD+SR reconstituida sintéticamente (residuos aminoácidos 1 a 35) se unió a otro tramo de la SR (residuos aminoácidos 36 a 140) para duplicar la longitud y poder añadir un anclaje de membrana de 140 residuos aminoácidos tal como se informa en Donadio *et al.*, 2003.

d. Construcción de un dominio sintético híbrido: hST3Gal III 1-28-hST6GalNAc I 37-74 (SEC ID nº 160)

Se construyó una región N-terminal híbrida que contenía el primer CT y el TMD de hST3Gal III (1 a 28 residuos aminoácidos) y después el inicio de la SR de hST6GalNAc I (aminoácidos 37 a 74).

La longitud total de este híbrido sintético era de 66 aminoácidos, correspondiente a 198 nucleótidos (SEC ID nº 159). Se diseñaron y se sintetizaron (Eurogentec) nueve oligonucleótidos de sentido y ocho oligonucleótidos antisentido correspondientes a la región N-terminal híbrida indicada anteriormente (1 a 225 pb) utilizando las secuencias de hST3Gal III/hST6GalNAc I. Se añadió un epítipo FLAG (en negrita) entre el codón de inicio ATG y el segundo codón (GGA) de la secuencia de hST3Gal III:

5'-ATG **GACTACAAAGACGATGACGACAAG** GGA-3'.

Las reacciones de fosforilación y de apareamiento se llevaron a cabo tal como se ha indicado anteriormente.

e. Amplificación por PCR de los constructos sintéticos

Tras la etapa de reconstitución sintética, el producto se amplificó mediante la técnica de PCR utilizando cebadores específicos que introducen los sitios de restricción deseados en cada extremo del fragmento de ADN reconstituido con el fin de ligarlos en primer lugar con el CD seleccionado y después en el vector pcDNA3.1.

Se utilizaron los cebadores siguientes:

5'AflII-ST3: 5'-GAGCCCCTTAAGATGGACTACAAAGACGACGATGACG-3'

y 3'ST3-*Bam*HI: 5'-TAAGGGGGATCCGCTAGAGTGAAGTATACTTACTGGA-3',

5'AflII-ST6: 5'-GAGCCCCCTTAAGATGGACTACAAAGACGACGATGACG-3'

5 y 3'ST6-*Bam*HI: 5'-TAAGGGGGATCCGGTTGTGGAGGAACGGGA-3',

3'ST6-*Bam*HI nº 2: 5'-TAAGGGGGATCCTCTGGGTGACAGTGTGTTTAC-3'.

10 Estos cebadores contienen los sitios de restricción *Afl*II (5'AflII-ST3 y 5'AflII-ST6) y *Bam*HI (3' ST3-*Bam*HI, 3' ST6-*Bam*HI y 3' ST6-*Bam*HI nº 2) para garantizar las ligaciones posteriores.

15 Se llevaron a cabo PCR con el conjunto de parejas de cebadores indicado en la Tabla 5. Se llevaron a cabo en un aparato de PCR (I-Cycler, Biorad) con 5 µl de cada uno de los fragmentos sintéticos en una solución que contenía 2,5 unidades de ADN polimerasa ProofStart (Qiagen), 300 µM de cada dNTP (Sigma Aldrich), 1 µM de cada cebador, 1X de tampón del fabricante de ProofStart (Qiagen), 1X de solución Q del fabricante (Qiagen) y 1,5 mM de Mg²⁺ en un volumen final de 50 µl.

20 La reacción se llevó a cabo del modo siguiente: 95°C durante 5 minutos, seguido de 40 ciclos de 30 segundos a 94°C, 30 segundos a 55°C, 53°C, 53°C y 55°C, respectivamente para hST3Gal III (SEC ID nº 151), hST6GalNAc I-74 (SEC ID nº 145), hST6GalNAc I-140 (SEC ID nº 147) y hST3Gal III-28/37-hST6GalNAc I-74 (SEC ID nº 159) y 1 minuto a 72°C. El producto de PCR amplificado se analizó mediante electroforesis en un gel de agarosa al 2% según el tamaño del producto de PCR (Tabla 5):

25 - hST3 Gal III: se amplificó únicamente un producto de PCR con un tamaño estimado de 170 pb, que corresponde al tamaño esperado de 174 pb (figura 11).

- hST6 Gal NAc I: la figura 13 muestra un producto de PCR amplificado con un tamaño estimado de 270 pb, que corresponde al tamaño esperado de 270 pb.

30 - hST6GalNAc I-140: se amplificó una secuencia nucleótida de 468 pb (figura 15).

- hST3Gal III-28-hST6GalNAc I37-74: se amplificó una secuencia de nucleótidos de 249 pb (figura 17).

35 Se purificaron los productos de PCR utilizando el kit de extracción en gel de Qiagen siguiendo las instrucciones del fabricante.

- Se sometió el producto de PCR de hST6GalNAc I-140 a secuenciación directa de ambas cadenas de ADN mediante Genome Express (Meylan, Francia).

40 - Se ligaron los productos de PCR de hST3Gal III y hST6GalNAc I-74 al CD (SEC ID nº 43) antes de introducirlos en el vector pcDNA3.1 (Invitrogen) para que también se secuenciasen adicionalmente mediante Genome Express (Meylan, Francia).

45 - Se ligó el fragmento sintético hST3Gal III-28-hST6GalNAc I 37-74 con el CD (SEC ID nº 43) y se secuenció directamente en ambas cadenas previamente a su introducción en el vector pcDNA3.1 (Invitrogen).

f. Cuantificación de los productos de PCR

50 Se agruparon todos los productos de PCR y se concentraron según cada constructo sintético. Se añadieron dos volúmenes de etanol y 0,1 volúmenes de acetato sódico 3 M, pH 5,2, a muestras de productos de PCR. La mezcla se incubó durante 30 minutos a -20°C y se centrifugó durante 20 minutos a 10.000 g a 4°C. Se lavaron los sedimentos con 100 µl de etanol al 70% y se realizó una centrifugación durante 10 minutos a 10.000 g a 4°C. A continuación, se eliminaron los sobrenadantes y se secaron los sedimentos y se resuspendieron en 50 µl de agua purificada. Se conservaron las muestras a -20°C hasta la utilización.

55 Se cargaron cinco microlitros de los cuatro fragmentos sintéticos concentrados en un gel de agarosa al 2% para estimar su cantidad (figuras 11, 12, 13 y 14).

Ejemplo 3: ensamblaje de dominios membranales N-terminales no catalíticos con actividad de siliiltransferasa

a. Digestiones y purificación

65 Se digirieron todos los productos de PCR (los constructos N-terminales sintéticos o el dominio catalítico) con el enzima de restricción *Bam*HI.

Se digirieron 200 nanogramos de los constructos N-terminales sintéticos en una solución que contenía 3 unidades del enzima de restricción *BamHI* y $0,1 \mu\text{g}\cdot\mu\text{l}^{-1}$ de albúmina de suero bovino y 1X del tampón E apropiado del fabricante (Promega) en un volumen final de 20 μl .

- 5 Se digirieron 500 nanogramos del CD con 5 unidades de *BamHI*, $0,1 \mu\text{g} \mu\text{l}^{-1}$ de albúmina de suero bovino y 1X de tampón E (Promega) en un volumen final de 20 μl .

Las digestiones se llevaron a cabo durante 90 minutos a 37°C y la incubación finalmente se inactivó durante 15 minutos a 65°C. Se purificó el ADN digerido utilizando el kit de purificación de PCR (Qiagen).

10

b. Ligación

Cada uno de los fragmentos N-terminales se ligó al fragmento CD en una solución que contenía 1,5 unidades de ADN ligasa de T4, 1X del tampón de ligación, ambos comercializados (Promega), 62,5 ng, 78,2 ng y 75 ng, respectivamente, para hST3Gal III, hST6GalNAc I-74 y el fragmento híbrido hST3Gal III-28-hST6GalNAc I 37-74 y 100 ng del CD en un volumen final de 20 μl . La mezcla se incubó a 15°C durante la noche y después se realizó una etapa de inactivación durante 10 minutos a 70°C (figura 17).

15

c. Amplificación de la inserción sintética etiquetada

20

Con el fin de verificar la ligación correcta, los productos de ligación se sometieron directamente a PCR utilizando las parejas de cebadores siguientes: 5'AflII-ST3/3' Δ 4-XbaI, 5'AflII-ST6/3' Δ 4-XbaI y 5'AflII-ST3/3' Δ 4-XbaI, respectivamente, para ST3/CD, ST6GalNAc-74/CD y el híbrido/CD.

25

Las reacciones se llevaron a cabo utilizando 2,5 unidades de ADN polimerasa ProofStart (Qiagen), 300 μM de cada dNTP (Sigma Aldrich), 1 μM de cada cebador, 1X de tampón del fabricante de ProofStart (Qiagen), 1X de solución Q del fabricante (Qiagen) y 3 mM de Mg^{2+} en un volumen final de 50 μl . La reacción se llevó a cabo del modo siguiente: 95°C durante 5 minutos, seguido de 40 ciclos de 1 minuto a 94°C, 1 minuto a 57°C y 1 minuto 30 segundos a 72°C.

30

Se analizaron los productos de PCR amplificados mediante electroforesis en un gel de agarosa al 1,5%:

- Inserción sintética de hST3Gal III/CD (SEC ID nº 167): se amplificó un producto de PCR de 1.200 pb (figura 12), que corresponde al tamaño esperado del fragmento de ADN ligado: ST3Gal III más CD.

35

- inserción de hST6GalNAc I-74/CD (SEC ID nº 161): se amplificó un producto de PCR de 1.200 pb (figura 14), que corresponde al tamaño esperado del fragmento de ADN ligado: ST6GalNAc I más CD, es decir, de 1.225 pb exactamente.

40

- inserción de hST3Gal III-28/37hST6GalNAc-74/CD (SEC ID nº 175): se amplificó un producto de PCR de 1.200 pb (figura 17), que corresponde al tamaño esperado del fragmento de ADN ligado hST3Gal III-29/hST6GalNAc37-74 más CD.

Se purificaron los productos de PCR amplificados utilizando el kit de extracción en gel de Qiagen siguiendo las instrucciones del fabricante.

45

Se agruparon los productos de PCR amplificados de la misma ST sintética, se concentraron y se cuantificaron en un gel de agarosa al 1,5% (figuras 12, 13 y 16), tal como se ha indicado anteriormente, con el fin de obtener suficiente cantidad de ADN para ligar la quimera en el vector pcDNA3.1.

50

d. Clonación de la inserción de ST sintética en un vector de expresión

1. Digestión y purificación

Se ligaron inserciones sintéticas, tales como hST3Gal III/CD o hST6GalNAc I-74/CD en el vector pcDNA3.1 en los sitios de restricción *AflII* y *XbaI*.

55

La inserción sintética y el vector pcDNA3.1 (figura 12) (Invitrogen) en primer lugar se digirieron con el enzima de restricción *XbaI*:

60

- para cada construcción, se digirieron 500 ng de vector utilizando 2,5 unidades de *XbaI* (Promega), $0,1 \mu\text{g}\cdot\mu\text{l}^{-1}$ de BSA y 1X del tampón D apropiado, suministrado por el fabricante (Promega), en un volumen final de 20 μl , siguiendo las instrucciones del fabricante,

65

- la inserción se digirió con 2 a 2,5 unidades de *XbaI* (Promega), dependiendo de la cantidad de inserción disponible tras la PCR (por ejemplo 200 ng de hST3Gal III/CD, hST3Gal III-28/hST6GalNAc I 37-74/CD y 500 ng de

hST6GalNAc I-74/CD se digirieron respectivamente con 2 y 2,5 unidades de *Xba*I, 0,1 µg·µl⁻¹ de BSA y 1X de tampón D (Promega) en un volumen final de 20 µl.

Las digestiones se llevaron a cabo a 37°C durante 60 minutos y se inactivaron a 65°C durante 15 minutos.

A continuación, los productos digeridos se sometieron a una digestión con el enzima de restricción *Afl*III siguiendo las instrucciones del fabricante (Biolabs): utilizando el número apropiado de unidades de enzima, dependiendo de la cantidad de ADN (2, 2, 2,5 y 2,5 unidades, respectivamente, para hST3Gal III/CD, hST3Gal III-28/hST6GalNAc I 37-74/CD, hST6GalNAc I-74/CD y pcDNA3.1), 0,1 µg·µl⁻¹ de BSA y 1X de tampón 2 (Biolabs) en un volumen final de 50 µl. Las digestiones se llevaron a cabo a 37°C durante 60 minutos y se inactivaron a 65°C durante 20 minutos. Las digestiones se purificaron con el kit de purificación de PCR (Qiagen).

2. Ligación

Se calcularon las condiciones de ligación según la fórmula siguiente:

$$X \text{ ng de inserción} = \frac{(\text{Y pb de inserción}) (\text{x ng de vector pcDNA3.1})}{\text{tamaño en pb del vector pcDNA3.1} \approx 5428} \times \text{proporción molar} \frac{\text{Inserción}}{\text{Vector}}$$

en la que el tamaño de la inserción es de aproximadamente 1.200 pb para el constructo sintético (por ejemplo hST3Gal III/CD, hST3Gal III-28/hST6GalNAc I 37-74/CD o hST6GalNAc I-74/CD), el tamaño del vector pcDNA3.1 es de 5.428 pb, la cantidad de vector utilizada es de 50 ó 100 ng, y finalmente, la proporción molar de inserción/vector es de 3/1.

- La construcción del vector recombinante pcDNA3.1/ST3/CD se llevó a cabo utilizando 33 ng de la quimera digerida hST3Gal III/CD (SEC ID nº 167), 50 ng de vector digerido pcDNA3.1, 3 unidades de enzima ADN ligasa de T4 (Promega) y 1X del tampón de ligación en un volumen final de 15 µl. La mezcla se incubó durante la noche a 4°C y la reacción se inactivó durante 10 minutos a 70°C.

- La segunda construcción de vector recombinante pcDNA3.1/ST6/CD se llevó a cabo utilizando 66,5 ng de la quimera digerida hST6GalNAc I-74/CD (SEC ID nº 161), 100 ng de vector digerido pcDNA3.1, 3 unidades de enzima ADN ligasa de T4 (Promega) y 1X del tampón de ligación en un volumen final de 15 µl. La mezcla se incubó durante la noche a 15°C y la reacción se inactivó durante 10 minutos a 70°C.

- La tercera construcción de vector recombinante pcDNA3.1/HYB/CD se llevó a cabo utilizando 66,5 ng de la quimera digerida hST3Gal III-28/hST6GalNAc I 37-74/CD (SEC ID nº 175), 100 ng de vector pcDNA3.1, 3 unidades de enzima ADN ligasa de T4 (Promega) y 1X de tampón de ligación en un volumen final de 15 µl. La mezcla se incubó durante la noche a 15°C y la reacción se inactivó durante 10 minutos a 70°C.

3. Clonación

Transformación química de células competentes

Se transformaron *E. coli* químicamente competentes One Shot® TOP10 (Invitrogen) con el vector recombinante siguiendo las instrucciones del fabricante. Se añadieron dos microlitros de cada reacción de ligación a 25 µl de células químicamente competentes y se mezclaron mediante golpes suaves. Los viales se incubaron sobre hielo durante 30 minutos. Las transformaciones químicas se llevaron a cabo durante exactamente 40 segundos en un baño de agua a 42°C. Se extrajeron los viales del baño a 42°C y se dejaron sobre hielo durante 2 minutos. Se añadieron doscientos cincuenta microlitros de medio S.O.C. precalentado, proporcionado por Invitrogen, a cada vial bajo condiciones estériles. A continuación, se agitaron las mezclas a 37°C durante exactamente 1 hora a 225 rpm en un incubador-agitador. Se extendió cada transformación en placas de agar-LB (10 g de bacto-triptona, 5 g de bacto-extracto de levadura, 10 g de NaCl y 15 g de agar por litro de solución de GibcoBRL) que contenía 50 µg/ml de ampicilina bajo condiciones estériles. Se invirtieron las placas, se incubaron a 37°C durante la noche y se mantuvieron a 4°C el día siguiente hasta la selección de los transformantes.

Amplificación

Al final del día, se aislaron colonias individuales y se inocularon en 3 ml de LB (10 g de bacto-triptona, 5 g de bacto-extracto de levadura y 10 g de NaCl de Difco) que contenía 50 µg/ml de ampicilina. Se aseguró el crecimiento a 37°C en un incubador bajo agitación durante la noche. Al día siguiente, se prepararon cultivos madre en glicerol mediante la mezcla de 0,85 ml de cultivo y 0,15 ml de glicerol estéril y se transfirieron a un vial criogénico. Los cultivos en glicerol se almacenaron a -80°C hasta la verificación de las secuencias y la utilización posterior.

Minipreparación de plásmidos

Los vectores recombinantes se purificaron mediante el procedimiento de Minipreparación (Maniatis), se aislaron 1,5 ml de los cultivos y se centrifugaron durante 5 minutos a 10.000 g para sedimentar las bacterias. Se añadió solución I fría (glucosa 5 mM, Tris-HCl 25 mM, pH 8,0, EDTA 10 mM, pH 8,0) (100 µl) para resuspender el sedimento y también se añadieron 200 µl de solución II (NaOH 0,2 N y SDS al 1%) y 150 µl de solución III (potasio 3 M y acetato 5 M). Las mezclas se agitaron suavemente con vórtex, se dejaron sobre hielo durante 5 minutos y se centrifugaron a 4°C durante 5 minutos a 15.000 g. Se obtuvieron los sobrenadantes y se añadió 1 volumen de isopropanol. Las minipreparaciones se incubaron durante 5 minutos a temperatura ambiente y se centrifugaron a 4°C durante 5 minutos a 15.000 g. A continuación, se separaron los sobrenadantes. Se secó el sedimento de ADN y se resuspendió en 50 µl de agua. Se midió la cantidad de ADN utilizando un espectrofotómetro DO (Biophotometer, Eppendorf). Se llevaron a cabo por lo menos 30 minipreparaciones con el fin de cribar para la presencia de la inserción de interés para cada construcción.

Un microgramo de cada una de las 30 minipreparaciones se digirió doblemente utilizando los enzimas de restricción *AflII* y *XbaI* con el fin de verificar la presencia de la inserción. En primer lugar, la solución de digestión contenía 10 unidades de *XbaI* (Promega), 0,1 µg·µl⁻¹ de BSA y 1X del tampón D apropiado del fabricante (Promega) en una solución final de 20 µl. Se llevaron a cabo digestiones en un baño de agua a 37°C durante 1 hora y se inactivaron a 65°C durante 15 minutos. En segundo lugar se digirieron las muestras utilizando 10 unidades de enzima de restricción *AflII* (Biolabs), 0,1X de BSA y 1X de tampón 2 (Biolabs) en un volumen final de 50 µl. Las reacciones se llevaron a cabo en un baño de agua a 37°C durante 1 hora y se inactivaron a 65°C durante 20 minutos. Se purificó la totalidad de las muestras digeridas utilizando el kit de purificación de PCR (Qiagen). Los productos de digestión se cargaron en un gel de agarosa al 1,5% para detectar la presencia de la inserción quimérica. Se llevó a cabo la electroforesis a 100 V durante aproximadamente 35 minutos.

En el caso de pcDNA3.1/ST3/CD y pcDNA3.1/ST6/CD, se detectó una inserción de 1.200 pb (figura 19C ó 23C). El tamaño de esta inserción correspondía exactamente al tamaño esperado de la quimera reconstituida ST3Gal III/CD o ST6GalNAc I-74/CD.

Se secuenciaron por completo los clones positivos con el fin de identificar la secuencia esperada de ADN insertado.

Secuenciación

Se verificaron las etapas de clonación mediante secuenciación de ADN de dos cadenas utilizando el Genome Express (Meylan, Francia) con cebadores universales presentes en el interior del vector: promotor de T7 y BGH Reverse.

Se obtuvieron las secuencias esperadas finales de las tres quimeras: ST3Gal III/CD (SEC ID nº 167), hST3Gal III-28/hST6GalNAc I 37-74/CD (SEC ID nº 175) y ST6GalNAc I-74/CD (SEC ID nº 161). La secuencia nucleótida obtenida se alineó con la secuencia esperada y la alineación reveló una identidad de 100% entre estas tres secuencias. También se alineó la secuencia de aminoácidos deducida con la secuencia esperada de la quimera (http://www.infobiogen.fr/services/analyse/cgi-bin/alignp_in.pl), demostrando una identidad de 100% entre dos de estas secuencias de aminoácidos.

La secuencia de nucleótidos obtenida para hST6GalNAc I-140 mostraba una identidad de 84% con la secuencia teórica. Las partes que faltaban de la secuencia eran los 36 primeros y los 39 últimos nucleótidos, correspondiendo ambos a los cebadores diseñados para la secuenciación del fragmento de ADN en ambas cadenas. Además, la alineación de los aminoácidos entre las secuencias esperadas y las resultantes mostró una identidad de 76,9%. No se encontraban los 12 primeros y los 24 últimos aminoácidos debido a sus correspondencias con los cebadores específicos utilizados para la secuenciación.

Minipreparación

Tras la verificación de la secuencia de la inserción de ADN, se amplificaron los plásmidos recombinantes en 100 ml de cultivo con el fin de obtener una cantidad elevada de cada construcción (pcDNA3.1/ST3/CD y pcDNA3.1/ST6/CD y pcDNA3.1/HYB/CD). Se recolectó una única colonia de cada plásmido de una placa selectiva recién sembrada y se inoculó en un cultivo inicial de 3 ml en medio LB que contenía el antibiótico selectivo ampicilina (50 µg·µl⁻¹). Los cultivos se incubaron durante la noche a 37°C bajo agitación vigorosa. Los cultivos iniciales se diluyeron en 100 ml de medio LB selectivo y se incubaron nuevamente durante la noche a 37°C bajo agitación vigorosa. Se recolectaron las células bacterianas y se purificaron los vectores recombinantes utilizando el kit QIAGEN® Plasmid Midi (Qiagen) siguiendo las instrucciones del fabricante. Se determinó el rendimiento con un espectrofotómetro de UV para cuantificar la concentración de ADN (biofotómetro, Eppendorf). Se almacenó cada constructo a -20°C hasta su utilización.

Ejemplo 4: expresión funcional de los constructos sintéticos en células CHO

Las inserciones expresadas se representan en la figura 15.

5 A. Expresión transitoria de ST6 sintéticas en células CHO

Se utilizó la línea celular CHO-K1 para expresar las construcciones, tal como se describe en Donadio *et al.* (2003). Brevemente, se cultivaron células en medio de Ham suplementado con suero de feto bovino al 10% (FCS), fungizona (2,5 µg/ml) y gentamicina (50 µg/ml) a 37°C en un incubador con 5% de CO₂. Las transfecciones con constructos de vector FLAG-CMV se llevaron a cabo utilizando el reactivo lipofectamina siguiendo el procedimiento recomendado por el fabricante, con 3 µg de ADN de plásmido recombinante. Los experimentos de inmunofluorescencia se llevaron a cabo tras 36 a 48 horas de la transfección.

Se llevó a cabo el doble marcaje con FITC-SNA y mAb anti-FLAG tal como se describe en Donadio *et al.* (2003). Las células se fijaron en paraformaldehído al 1% (solución de PFA, saturada con BSA al 1%, lavado con solución salina tamponada con fosfato (PBS) y se incubaron con FITC-SNA. Las células que expresaban una actividad de ST6 se marcaron en verde. Las muestras se lavaron dos veces con PBS y se fijaron con PFA al 1%, se lavaron adicionalmente con PBS, se incubaron en NH₄Cl 0,05 M y se lavaron con PBS. Las células se permeabilizaron con Triton al 0,1% en PBS y se incubaron adicionalmente en PBS que contenía suero de cabra al 10%. Se llevó a cabo el marcaje anti-FLAG utilizando el mAb anti-FLAG M2 (1:600) en PBS que contenía suero de cabra o de caballo inactivo al 5% seguido de la incubación con un anticuerpo secundario anti-IgG de ratón con Alexa Fluor 568 (1:200) en el mismo tampón. Las células que expresaban la sialiltransferasa etiquetada con FLAG se marcaron en rojo. Tras los lavados, las células se montaron en Mowiol y se mantuvieron a -20°C en la oscuridad. Se llevó a cabo microscopía confocal con un instrumento Olympus o Leica. Las imágenes confocales se trataron con un sistema de imágenes Metamorph. Los volúmenes se trazaron originalmente como imágenes de color verdadero de 24-bit y se transfirieron a Adobe Photoshop en forma de archivos TIFF en RGB o JPG. Bajo estas condiciones, las células CHO-K1 no transfectadas no se encuentran marcadas con ninguno de los reactivos anteriormente indicados.

B. Actividad enzimática de ST6 sintéticas en células CHO**B1. Actividad del dominio catalítico optimizado de hST6Gal I**

Se clonó el CD de hST6Gal I en el vector pFLAG-citomegalovirus (pFLAG-CMV1) de Sigma utilizando un péptido de señal preprotripsinógeno. Esta construcción y estos plásmidos fueron caracterizados anteriormente *in vitro* por Legaigneur *et al.* (2001), demostrando que proporcionaban una proteína enzimática del tamaño esperado que se liberaba al medio celular. Se encontró que este CD soluble era igualmente activo con glucoproteínas aceptoras exógenas de diversos grado de ramificación.

Para generar la forma soluble del CD mínimo de hST6Gal I, se generó un fragmento de PCR codificado entre:

5' (5'-TAATAAAGCTTGAAGCCTCCTTCCAG-3') introduciendo un sitio de restricción *HindIII* y

3' (5'-CTATTGGATCCTTAGCAGTGAATGGT-3') codificante de un sitio *BamHI*. A continuación, el fragmento se digirió con *HindIII/BamHI* y se insertó en el vector de expresión de mamífero pFLAG-CMV1 y se subclonó en el vector pBluescript II KS (Stratagene) para construcciones posteriores.

El CD mínimo de ST6Gal I (SEC ID nº 44) etiquetado con el epitopo FLAG (figura 18) se expresó a nivel elevado en células CHO (figura 18 B, D). Se caracterizó su actividad y localización utilizando el doble marcaje con anticuerpo monoclonal anti-FLAG y FITC-SNA. Significativamente, el CD-FLAG mutante era altamente eficiente en la sialilación de los aceptores de superficie celular, tal como se muestra en las figuras 18C y D.

B2. Expresión funcional de ST8/CD

Se transfectaron transitoriamente células CHO-K1 con formas quiméricas de hST8SialIV (1-67)/CD (SEC ID nº 173), hST8SialI (1-79)/CD (SEC ID nº 171). Se caracterizó su actividad y localización utilizando el doble marcaje con anticuerpo monoclonal anti-FLAG y FITC-SNA, tal como se muestra en la figura 19.

Se encontró que las quimeras tanto hST8SialIV (1-67)/CD (SEC ID nº 174) como hST8SialI (1-79)/CD (SEC ID nº 172) eran activas sobre los aceptores celulares endógenos, ya que la lectina SNA reveló un marcaje intenso de la superficie celular (figura 19 B, C, E y F), indicando que el CD era activo con independencia del origen de la parte N-terminal fusionada con el dominio catalítico. Se mantuvo la actividad de α 2,6-sialiltransferasa con ambas quimeras, indicando que en el enzima quimérico la información contenida en el CD mínimo resultaba necesaria y suficiente para el reconocimiento de aceptor y la eficiencia de la transferencia.

B3. Expresión funcional de ST3/CD y del híbrido ST3-ST6GalNAc/CD

Se transfectaron transitoriamente células CHO-K1 con los constructos quiméricos de hST3Gal III/CD (SEC ID nº 167) y ST3 1-28-ST6GalNAc37-74/CD (SEC ID nº 175). Se caracterizó su actividad y localización utilizando el doble marcaje con anticuerpo monoclonal anti-FLAG y FITC-SNA, tal como se muestra en la figura 19.

Nuevamente, se encontró que ambas quimeras eran activas sobre aceptores endógenos, ya que la unión de SNA revelaba un marcaje intenso de la superficie celular (figura 19 B, C, E y F). Lo anterior indica que el CD es activo con independencia del origen de CT+TMD+SR añadido. Debido a que la actividad de la α 2,6-sialiltransferasas se mantenía en ambas quimeras, también se concluyó que la información contenida en el CD mínimo resultaba necesaria y suficiente para una eficiencia máxima de la transferencia.

Por lo tanto, la fusión de las secuencias de ST3/8 cadena arriba del CD no altera la expresión de la actividad catalítica de la ST6. Debido a que la ST3 Gal III de rata y humana, y más generalmente la ST3Gal III de roedor y de mamífero, son idénticas en su parte N-terminal (CT+TMD+SR yuxtamembrana), aparentemente el intercambio de cualquier trozo de esta parte peptídica por trozos heterólogos no humanos tampoco presenta ningún efecto sobre la actividad catalítica de ST6. Además, la sustitución de CT, TMD y SR por partes heterólogas seleccionadas de entre enzimas de O-glucosilación como ST GalNAcI no altera el reconocimiento de los aceptores intracelulares ni la actividad enzimática.

B4. Expresión funcional de hST6GalNAc I-74/CD y de hST6GalNAc 258/CD

Se utilizó un enfoque similar para estimar la expresión y actividad de los enzimas hST6GalNAc I (1-74)/CD (SEC ID nº 162) y de hST6GalNAc I-258/CD (SEC ID nº 166). Ambas transferasas quiméricas eran activas: la unión de SNA reveló una sialilación intensa similar de la superficie celular y el marcaje con FLAG mostró una localización idéntica dentro del Golgi (figura 21G, H e I).

De esta manera, puede concluirse que la longitud de la región SR puede ser ampliamente variable y extenderse hasta 258 residuos sin afectar a la expresión funcional y localización intracelular de la actividad catalítica de una ST6Gal quimérica. Lo anterior también confirma adicionalmente que la longitud del anclaje membranal N-terminal heterólogo no afecta a la conformación biológicamente activa de la sialiltransferasa de diseño nuevo.

Basándose en los ejemplos anteriores, puede afirmarse que el procedimiento descrito en la invención puede eliminar el control regulador de la región SR sobre la actividad del CD. En consecuencia, puede seleccionarse cualquier combinación posible de CT+TMD+SR entre las glucosiltransferasas (sialiltransferasas) de la ruta de N-glucosilación u O-glucosilación para dotar a células huésped de una actividad de transferasa necesaria, mediante un CD relevante. Además, el procedimiento descrito en la invención también permite dirigir una actividad necesaria en compartimentos intracelulares en los que migran glucoproteínas recién sintetizadas hasta la superficie celular. Este resultado es de crucial importancia para la correcta clasificación y secreción de proteínas sialiladas de interés terapéutico en las células (de levaduras a humanas) manipuladas tal como se describe en la invención.

Ejemplo 5: construcción de un dominio sintético híbrido: hST3Gal III/hST8Sia II/hST6GalNAc I y fusión con el CD de hST6Gal I

Se construyó una región N-terminal híbrida (SEC ID nº 180) que contenía en primer lugar el CT de hST3Gal III (residuos aminoácidos 1 a 8, SEC ID nº 66), después el TMD de hST8Sia II (residuos aminoácidos 7 a 23, SEC ID nº 120) y finalmente la SR de hST6GalNAc I (aminoácidos 36 a 74, SEC ID nº 130). Se añadió un epítipo FLAG (en negrita) entre el codón de inicio ATG y el segundo codón (GGA) de la secuencia de hST3Gal III.

5'-ATG **GACTACAAAGACGATGACGACAAG** GGA- 3'

La longitud total de este híbrido era de 72 aminoácidos (SEC ID nº 180), correspondiente a 216 nucleótidos (SEC ID nº 179). Se diseñaron y sintetizaron (Eurogentec) ocho oligonucleótidos de sentido y siete oligonucleótidos antisentido correspondientes a la región N-terminal híbrida indicada anteriormente (nucleótidos 1 a 216) utilizando las secuencias de hST3Gal III/hST8Sia II/hST6GalNAc I.

Se llevaron a cabo reacciones de fosforilación y apareamiento tal como se ha indicado anteriormente.

Amplificación por PCR del constructo sintético

Tras la etapa de reconstitución sintética, se amplificó el producto utilizando una PCR de alta fidelidad utilizando cebadores específicos que introducen los sitios de restricción deseados en cada extremo del fragmento de ADN reconstituido con el fin de ligarlo al vector pcDNA3.1. Se introdujeron los sitios de restricción *Afl*III y *Bam*HI respectivamente en los extremos 5' y 3'.

Se utilizaron los cebadores siguientes:

5' AflII-ST3: 5'-GAGCCCCTTAAGATGGACTACAAAGACGATGACG-3'
y 3' ST6-BamHI n° 3: 5'-TAAGGGGGATCCGCTTGTCTCCTTGCCT-3'

Se llevó a cabo la PCR en un aparato de PCR (I-Cycler, Biorad) con 5 µl del fragmento sintético en una solución que contenía 1X de tampón ProofStart del fabricante y 1X de solución Q del fabricante, 1,5 mM de Mg²⁺, 300 µM de cada dNTP (Sigma Aldrich), 1 µM del cebador directo AflII-ST3, 1 µM de cebador inverso ST6-BamHI n° 3, 2,5 unidades de ADN polimerasa ProofStart (Qiagen) en un volumen final de 50 µl. El perfil de termociclado utilizado fue el siguiente: 95°C durante 5 minutos, seguido de 40 ciclos de 30 segundos a 94°C, 30 segundos a 55°C y 1 minuto a 72°C.

Los productos de PCR amplificados se analizaron mediante electroforesis en un gel de agarosa al 2%. Se amplificó una secuencia nucleótida de 240 pb, que corresponde al tamaño esperado del fragmento (figura 22, carril 1). Los productos de PCR se purificaron utilizando el kit de extracción en gel (QIAGEN) siguiendo las instrucciones del fabricante.

Cuantificación del producto de PCR

Se agruparon todos los productos de PCR y se agruparon. Se añadieron dos volúmenes de etanol y 0,1 volúmenes de acetato sódico 3 M, pH 5,2, a los productos de PCR. La mezcla se incubó durante 30 minutos a -20°C y se centrifugó durante 20 minutos a 10.000 g a 4°C. Se lavó el sedimento con 100 µl de etanol al 70% y se centrifugó nuevamente durante 10 minutos a 10.000 g a 4°C. A continuación, se separó el sobrenadante y se secó el sedimento y se resuspendió en 30 µl de agua doblemente destilada. La muestra de ADN se conservó a -20°C hasta su utilización.

Se cargaron tres microlitros del fragmento concentrado en un gel de agarosa al 2% para estimar su cantidad (figura 22, carril 3).

Ligación del dominio catalítico a un vector de expresión, pcDNA3.1(+)

1. Digestiones y purificaciones

La secuencia de nucleótidos del dominio catalítico CD de hST6Gal I (SEC ID n° 43) se ligó en el vector pcDNA3.1 en los sitios de restricción *Bam*HI y *Xba*I.

En primer lugar se digirieron tanto el vector pcDNA3.1(+) como el CD con el enzima de restricción *Bam*HI:

Se digirieron 500 ng del CD con 5 unidades de *Bam*HI (Promega), 0,1 µg µl⁻¹ de albúmina de suero bovino y 1X de tampón E (Promega) en un volumen final de 20 µl.

Se digirió 1 µg del vector pcDNA3.1 con 10 unidades de *Bam*HI (Promega), 0,1 µg µl⁻¹ de albúmina de suero bovino y 1X de tampón E (Promega) en un volumen final de 20 µl.

Se llevaron a cabo las digestiones durante 60 minutos a 37°C y las incubaciones se inactivaron finalmente durante 15 minutos a 65°C. Los productos digeridos se purificaron utilizando el kit de purificación de PCR (Qiagen).

A continuación, los productos digeridos se sometieron a una digestión con el enzima de restricción *Xba*I (Promega) siguiendo las instrucciones del fabricante: utilizando el número correcto de unidades enzimáticas, dependiendo de la cantidad de ADN (5 y 6, respectivamente para el CD y el pcDNA3.1), 0,1 µg µl⁻¹ de BSA y 1X de tampón D (Promega) en un volumen final de 50 µl. Las digestiones se llevaron a cabo a 37°C durante 60 minutos. Las digestiones se purificaron utilizando el kit de purificación de PCR (Qiagen).

2. Ligación

Se calcularon las condiciones de ligación según la fórmula anteriormente indicada. El tamaño del CD insertado era de aproximadamente 960 pb, el tamaño del vector pcDNA3.1 era de 5.428 pb, la cantidad de vector utilizada era de 100 ng y finalmente la proporción molar inserción/vector era de 3/1.

La construcción del vector recombinante pcDNA3.1/CD se llevó a cabo utilizando 53 ng de CD doblemente digerido, 100 ng de vector pcDNA3.1 doblemente digerido, 6 unidades de enzima ADN ligasa de T4 (Promega) y 1X de tampón de ligación en un volumen final de 20 µl. La mezcla se incubó durante la noche a 15°C y la reacción se inactivó durante 10 minutos a 70°C.

3. Clonación

Transformación química de células competentes

- 5 Se transformaron *E. coli* químicamente competentes One Shot® TOP10 (Invitrogen, Life Technologies) con el vector recombinante siguiendo las instrucciones del fabricante. Se añadieron cinco microlitros del producto de ligación a una alícuota de 50 µl de células TOP10 y se mezclaron mediante golpecitos suaves. La mezcla se incubó sobre hielo durante 30 minutos y después se sometió a choque térmico durante exactamente 30 segundos en un baño de agua a 42°C y se dejaron sobre hielo durante 2 minutos. Se añadieron doscientos cincuenta microlitros de medio SOC precalentado (Invitrogen) a la reacción de transformación y se incubaron a 37°C durante una hora bajo agitación a 250 rpm. La reacción de transformación se extendió sobre placas de agar LB (composición indicada anteriormente) que contenía 50 µg·ml⁻¹ de ampicilina. Las placas se incubaron durante la noche a 37°C.

Amplificación

- 15 Al final del día siguiente, se recolectaron colonias individuales y se incubaron con 3 ml de caldo LB (composición indicada anteriormente) que contenía 50 µg ml⁻¹ de ampicilina, durante la noche a 37°C bajo agitación a 250 rpm.

Minipreparación de plásmidos

- 20 Los vectores recombinantes se purificaron a partir de alícuotas de 1,5 ml de los cultivos anteriores utilizando el kit QIAprep Spin Miniprep proporcionado por Qiagen y siguiendo las instrucciones del fabricante. Se llevaron a cabo veinticuatro minipreparaciones con el fin de cribar para la presencia de la inserción.
- 25 Cinco microlitros de cada una de las 24 minipreparaciones se digirieron doblemente utilizando los enzimas de restricción *Xba*I y *Bam*HI con el fin de verificar la presencia de la inserción.

- En primer lugar, la solución de digestión contenía 12 unidades de *Xba*I (Promega), 0,1 µg·µl⁻¹ de BSA y 1X de tampón D (Promega) en un volumen final de 10 µl. Las digestiones se llevaron a cabo durante 2 horas a 37°C.
- 30 En segundo lugar, las muestras se digirieron utilizando 20 unidades de *Bam*HI (Promega) y 1X de tampón E (Promega) en un volumen final de 20 µl. Las reacciones se llevaron a cabo durante la noche a 37°C. Los productos de digestión se cargaron en un gel de agarosa al 1,5% para detectar la presencia del CD. Se detectó una inserción de aproximadamente 1.000 pb (figura 23) que correspondía al tamaño esperado del CD (960 pb).

- 35 Se seleccionó un clon positivo para las etapas siguientes de construcción.

Ligación de la región N-terminal en el vector recombinante pcDNA3.1/CD

1. Digestiones y purificaciones

- 40 Se ligó la secuencia de nucleótidos de la región N-terminal (SEC ID nº 179) en el vector recombinante pcDNA3.1/CD en los sitios de restricción *Bam*HI y *Afl*II. En primer lugar se digirieron tanto el fragmento N-terminal como el vector recombinante pcDNA3.1/CD con el enzima de restricción *Bam*HI:

- 45 se digirieron 300 ng del fragmento N-terminal con 5 unidades de *Bam*HI (Promega), 0,1 µg·µl⁻¹ de BSA y 1X de tampón E (Promega) en un volumen final de 20 µl.

- Se digirieron 2 µg del vector recombinante pcDNA3.1/CD con 20 unidades de *Bam*HI (Promega), 0,1 µg·µl⁻¹ de BSA y 1X de tampón E (Promega) en un volumen final de 20 µl.

- 50 Las digestiones se llevaron a cabo durante 100 minutos a 37°C y las incubaciones se inactivaron finalmente durante 15 minutos a 65°C. Los productos digeridos se purificaron utilizando el kit de purificación de PCR (Qiagen).

- A continuación, los productos digeridos se sometieron a una digestión con el enzima de restricción *Afl*II (New England BIOLABS), siguiendo las instrucciones del fabricante: utilizando el número correcto de unidades enzimáticas, dependiendo de la cantidad de ADN (10 y 20, respectivamente para el fragmento N-terminal y el vector recombinante pcDNA3.1/CD), 1X de BSA y 1X de tampón 2 (New England BIOLABS) en un volumen final de 50 µl. Las digestiones se llevaron a cabo a 37°C durante 60 minutos y finalmente las incubaciones se inactivaron durante 20 minutos a 65°C. Las digestiones se purificaron utilizando el kit de purificación de PCR (Qiagen).

2. Ligación

- Se calcularon las condiciones de ligación según la fórmula anteriormente indicada. El tamaño del fragmento N-terminal insertado era de 240 pb, el tamaño del vector recombinante pcDNA3.1/CD era de 6.388 pb, la cantidad de vector utilizada era de 100 ng y finalmente la proporción molar de inserción/vector era de 1/1.

Se llevó a cabo la construcción del vector recombinante pcDNA3.1/fragmento N-terminal/CD utilizando 5 ng del fragmento N-terminal digerido, 100 ng de vector recombinante pcDNA3.1/CD digerido, 6 unidades del enzima ADN ligasa de T4 (Promega) y 1X de tampón de ligación en un volumen final de 20 µl. La mezcla se incubó durante la noche a 15°C y la reacción se inactivó durante 10 minutos a 70°C.

3. Clonación

Transformación química de células competentes

Se transformaron *E. coli* químicamente competentes One Shot® TOP10 (Invitrogen) con el producto de ligación siguiendo las instrucciones del fabricante y tal como se ha indicado anteriormente.

Amplificación

Al final del día siguiente, se recolectaron colonias individuales y se incubaron con 3 ml de caldo LB (composición indicada anteriormente) que contenía 50 µg·ml⁻¹ de ampicilina, durante la noche a 37°C bajo agitación a 250 rpm.

Minipreparación de plásmidos

Se purificaron los vectores recombinantes a partir de alícuotas de 1,5 ml de los cultivos anteriores utilizando el kit QIAprep Spin Miniprep proporcionado por Qiagen, siguiendo las instrucciones del fabricante. Se llevaron a cabo veinticuatro minipreparaciones con el fin de cribar para la presencia de la inserción.

Se digirieron doblemente cinco microlitros de cada una de las 24 minipreparaciones utilizando los enzimas de restricción *Xba*I y *Afl*III con el fin de verificar la presencia de la inserción (fragmento N-terminal + CD, SEC ID nº 181).

La solución de digestión contenía 20 unidades de *Xba*I (Biolabs), 20 unidades de *Afl*III (Biolabs), 1X de BSA (Biolabs) y 1X de tampón 2 (Biolabs) en un volumen final de 20 µl. Las digestiones se llevaron a cabo durante 2 horas y 30 minutos a 37°C. Los productos de digestión se cargaron en un gel de agarosa al 1,5% para detectar la presencia de tanto la región N-terminal como el CD. Se detectó una inserción de aproximadamente 1.200 pb (figura 24a) que correspondía al tamaño esperado de la inserción.

Finalmente se llevó a cabo una PCR en cinco muestras minipreparación para verificar la presencia del ADN insertado (fragmento N-terminal seguido del CD: SEC ID nº 181). Se llevó a cabo la PCR en un aparato de PCR (I-Cycler, Biorad) con 1 µl de cada una de las muestras minipreparación en una solución que contenía 1X de tampón del fabricante de ProofStart y 1X de solución Q del fabricante, 3 mM de Mg²⁺, 300 µM de cada dNTP (Sigma Aldrich), 1 µM de cebador directo *Afl*II-ST3, 1 µM de cebador inverso Δ 4-*Xba*I, 2,5 unidades de ADN polimerasa ProofStart (Qiagen) en un volumen final de 50 µl.

Los oligonucleótidos incluidos en la reacción fueron los siguientes:

5' *Afl*II-ST3: 5'-GAGCCCCTTAAGATGGACTACAAAGACGATGACG-3'

y 3' Δ 4-*Xba*I: 5'-TAACCCTCTAGATTAGCAGTGAATGGTCCGGAAGC-3'

El perfil de termociclado utilizado fue el siguiente: 95°C durante 5 minutos seguido de 40 ciclos de 1 minuto a 94°C, 1 minuto a 57°C y 1 minuto 30 segundos a 72°C. Los productos de PCR amplificados se analizaron mediante electroforesis en un gel de agarosa al 1,5%: se amplificó una secuencia de nucleótidos de 1.200 pb para cada muestra minipreparación, que corresponde al tamaño esperado de la inserción total (figura 24b).

Secuenciación

Se verificaron las etapas de clonación mediante secuenciación de ADN de las dos cadenas utilizando GENOME Express (Meylan, Francia) con cebadores universales presentes en el interior del vector: promotor de T7 y BGH inverso.

La secuencia nucleótida obtenida se alineó con la secuencia esperada y la alineación reveló una identidad de 100%. También se alineó la secuencia de aminoácidos deducida con la secuencia esperada SEC ID nº 181 de la quimera y mostraba una identidad de 100%.

Midipreparación

Tras la verificación de la secuencia de la inserción de ADN, se amplificó el plásmido recombinante en 100 ml de cultivo con el fin de obtener una cantidad elevada de la construcción. Se inocularon cuatrocientos microlitros del cultivo en 3 ml de caldo LB correspondiente indicado anteriormente, en 100 ml de caldo LB que contenía 50 µg·ml⁻¹ de ampicilina, y se incubaron a 37°C bajo agitación a 250 rpm. Se recolectaron las células bacterianas y se

purificaron los vectores recombinantes utilizando el kit QIAGEN Plasmid Midi (Qiagen) siguiendo las instrucciones del fabricante. Se determinó el rendimiento con un espectrofotómetro de UV para cuantificar la concentración del ADN. Se almacenó el constructo a -20°C.

- 5 **Cultivo celular:** expresión del enzima quimérico hST3Gal III/hST8Sia II/hST6GalNAc I/CD (SEC ID nº 182) en células CHO

Se cultivaron rutinariamente células CHO-K1 a 37°C con 5% de CO₂ en DMEM suplementado con FBS al 10%, 2,5 µg/ml de fungizona y 50 µg/ml de gentamicina. El día antes de la transfección, se sembraron células a una densidad de 100.000 células/ml en una placa de 6 pocillos que contenía 3 cubreobjetos de vidrio en cada pocillo. Se transfectaron transitoriamente células CHO-K1 con reactivo de transfección FuGENE 6 (Roche) en una proporción 3:1.

Transcurrido un día de cultivo, las células se fijaron con paraformaldehído al 1% durante 15 minutos a temperatura ambiente (RT). Las células se lavaron tres veces con PBS, se bloquearon durante 30 minutos a RT con BSA al 1% y después se incubaron con 10 µg/ml de SNA-FITC (Vector Laboratories) durante una hora a 4°C. Las células se lavaron sucesivamente tres veces con PBS, se fijaron con paraformaldehído al 3% durante 20 minutos a RT, se lavaron adicionalmente con PBS y se incubaron con NH₄Cl 0,05 M durante 10 minutos a RT. Tras los lavados con PBS, las células se permeabilizaron con Triton X-100 al 0,25% durante 5 minutos a RT y se bloquearon con suero de cabra inactivado al 10% y BSA al 1% durante 30 minutos a RT. Las células se incubaron durante una hora a RT con 1 µg/ml de anticuerpo monoclonal anti-FLAG M2 (Sigma) en PBS que contenía suero de cabra inactivado al 5% y BSA al 0,5%. Las células se lavaron tres veces con PBS y se incubaron durante 45 minutos con 5,7 µg/ml de anticuerpo secundario de cabra antirratón acoplado a Alexa Fluor®594 (Molecular Probes) en PBS que contenía suero de cabra inactivado al 5% y BSA al 0,5%. Tras los lavados con PBS, se montaron las células en Mowiol y se analizaron utilizando un microscopio Olympus (figura 25).

Referencias

- Andersen D. C. & Krummen L. (2002) Recombinant protein expression for therapeutic applications. *Curr. Opin. Biotechnol.*, 13(2): 117-23.
- Angata K., Nakayama J., Fredette B., Chong K., Ranscht B. & Fukuda M. (1997) Human STX polysialyltransferase forms the embryonic form of the neural cell adhesion molecule. Tissue-specific expression, neurite outgrowth, and chromosomal localization in comparison with another polysialyltransferase, PST. *J. Biol. Chem.*, 272(11):7182-90.
- Angata K., Suzuki M., McAuliffe J., Ding Y., Hindsgaul O. & Fukuda M. (2000) Differential biosynthesis of polysialic acid on neural cell adhesion molecule (NCAM) and oligosaccharide acceptors by three distinct alpha 2,8-sialyltransferases, ST8Sia IV (PST), ST8Sia II (STX), and ST8Sia III. *J. Biol. Chem.*, 275(24): 18594-601.
- Angata K., Yen T. Y., El-Battari A., Macher B. A. & Fukuda M. (2001) Unique disulfide bond structures found in ST8Sia IV polysialyltransferase are required for its activity. *J. Biol. Chem.*, 276(18): 15369-77.
- Ashwell G. & Harford J. (1982) Carbohydrate-specific receptors of the liver. *Annu. Rev. Biochem.*, 51:531-54.
- Aumiller J. J., Hollister J. R., & Jarvis D. L. (2003) A transgenic insect cell line engineered to produce CMP-sialic acid and sialylated glycoproteins. *Glycobiology*, 13(6):497-507.
- Baneyx F. (1999) Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.*, 10(5):411-21.
- Berger E. G. (2002) Ectopic localizations of Golgi glycosyltransferases. *Glycobiology*, 12(2):29R-36R.
- Berger E.G., Burger P., Borsig L., Malissard M., Felner K.M., Zeng S., Dinter A. (1998) Immunodetection of glycosyltransferases: prospects and pitfalls. *Adv. Exp. Med. Biol.*, 435:119-32.
- Blanchard S. (2004) Ingenierie de glicoside hydrolases pour la glycosylation des proteines recombinantes. These de l'université Joseph Fourier, 272p.
- Bragonzi A., Distefano G., Buckberry L.D., Acerbis G., Foglieni C, Lamotte D., Campi G., Marc A., Soria M.R., Jenkins N., Monaco L. (2000) A new Chinese hamster ovary cell line expressing alpha2,6-sialyltransferase used as universal host for the production of human-like sialylated recombinant glycoproteins. *Biochim. Biophys. Acta*, 1474(3):273-82.
- Breton C. & Imberty A. (1999) Structure/function studies of glycosyltransferases. *Curr. Opin. Struct. Biol*, 9(5):563-71.

- Breton C, Mucha J. & Jeanneau C. (2001) Structural and functional features of glycosyltransferases. *Biochimie*, 83(8):713-8.
- 5 Bretscher M. S. & Munro S. (1993) Cholesterol and the Golgi apparatus. *Science*, 261(5126):1280-1.
- Bretthauer R.K. (2003) Genetic engineering of *Pichia pastoris* to humanize N-glycosylation of proteins. *Trends Biotechnol.*, 21(11):459-62.
- 10 Bianucci A. M. & Chiellini F. (2000) A 3D model for the human hepatic asialoglycoprotein receptor (ASGP-R). *J Biomol. Struct. Dyn.*, 18(3):435-51.
- Chang M. L., Eddy R.L., Shows T. B., Lau J.T. (1995) Three genes that encode human beta-galactoside alpha 2,3-sialyltransferases. Structural analysis and chromosomal mapping studies. *Glycobiology*, 5(3):319-25.
- 15 Chen C. & Colley K. J. (2000) Minimal structural and glycosylation requirements for ST6Gal I activity and trafficking. *Glycobiology*, 10(5):531-83.
- Chen C, Ma J., Lazic A., Backovic M., Colley KJ. (2000) Formation of insoluble oligomers correlates with ST6Gal I stable localization in the golgi. *J. Biol. Chem.*, 275(18):13819-26. Chen T.L., Chen C, Bergeron N.Q., Close B.E., Bohrer T.J., Vertel B.M. & Colley KJ. (2003) The two rat alpha 2,6-sialyltransferase (ST6Gal I) isoforms: evaluation of catalytic activity and intra-Golgi localization. *Glycobiology*, 13(2): 109-17.
- 20 Colley K. J. (1997) Golgi localization of glycosyltransferases: more questions than answers. *Glycobiology*, 7(1): 1-13.
- 25 Crocker P. R. & Feizi T. (1996) Carbohydrate recognition systems: functional triads in cell-cell interactions. *Curr. Opin. Struct. Biol*, 6(5):679-91.
- Datta A. K., Sinha A. & Paulson J. C. (1998) Mutation of the sialyltransferase S-sialylmotif alters the kinetics of the donor and acceptor substrates. *J. Biol. Chem.*, 273(16):9608-14.
- 30 Datta A. K. & Paulson J. C. (1995) Sialylmotifs of sialyltransferases. *Indian J. Biochem. Biophys.*, 34(1-2):157-65.
- Donadio S., Dubois C, Fichant G., Roybon L., Guillemot J.C., Breton C. & Ronin C. (2003) Recognition of cell surface acceptors by two human alpha-2,6-sialyltransferases produced in CHO cells. *Biochimie*, 85(3-4):311-21.
- 35 Drickamer K. (1993) A conserved disulphide bond in sialyltransferases. *Glycobiology*, 3(1):2-3.
- Feizi T. (1993) Carbohydrate—protein interactions in capillary morphogenesis? *Glycobiology*, 3(6):547-8.
- 40 Fenteany F. H. & Colley K. J. (2005) Multiple signals are required for alpha2,6-sialyltransferase (ST6Gal I) oligomerization and Golgi localization. *J. Biol. Chem.*, (7):5423-9.
- Gabius H.J., Siebert H.C., Andre S., Jimenez-Barbero J. & Rudiger H. (2004) Chemical biology of the sugar code. *Chembiochem.*, 5(6):740-64.
- 45 Geremia R. A., Harduin-Lepers A. & Delannoy P. (1997) Identification of two novel conserved amino acid residues in eukaryotic sialyltransferases: implications for their mechanism of action. *Glycobiology*, 7(2):v-vii.
- Giordanengo V., Bannwarth S., Laffont C, Van Miegem V., Harduin-Lepers A., Delannoy P. & Lefebvre J. C. (1997) Cloning and expression of cDNA for a human Gal(beta1-3)GalNAc alpha2,3-sialyltransferase from the CEM T-cell line. *Eur. J. Biochem.*, 247(2):558-66.
- 50 Grabenhorst E. & Conradt H. S. (1999) The cytoplasmic, transmembrane, and stem regions of glycosyltransferases specify their in vivo functional sub localization and stability in the Golgi. *J. Biol. Chem.*, 274(51):36107-16.
- 55 Grabenhorst E., Schlenke P., Pohl S., Nimtz M. & Conradt H.S. (1999) Genetic engineering of recombinant glycoproteins and the glycosylation pathway in mammalian host cells. *Glycoconj. J.*, 16(2):81-97.
- Hamilton S. R., Bobrowicz P., Bobrowicz B., Davidson R. C, Li H., Mitchell T., Nett J. H., Rausch S., Stadheim T. A., Wischniewski H., Wildt S. & Gerngross, T. U. (2003) Production of complex human glycoproteins in yeast. *Science*, 301, 1244-1246.
- 60 Harduin-Lepers A., Stokes D. C, Steelant W.F., Samyn-Petit B., Krzewinski-Recchi M.A., Vallejo-Ruiz V., Zanetta J.P., Auge C. & Delannoy P. (2000) Cloning, expression and gene organization of a human Neu5Ac alpha 2-3Gal beta 1-3GalNAc alpha 2,6-sialyltransferase: hST6GalNAcIV. *Biochem. J.*, 352(Pt 1):37-48.
- 65

- Harduin-Lepers A., Vallejo-Ruiz V., Krzewinski-Recchi M.A., Samyn-Petit B., Julien S. & Delannoy P. (2001) The human sialyltransferase family. *Biochimie*, 83(8):727-37.
- 5 Helenius A. & Aebi M. (2001) Intracellular functions of N-linked glycans. *Science*, 291(5512):2364-9.
- Herscovics A. & Orlean P. (1993) Glycoprotein biosynthesis in yeast. *FASEB J.*, 7(6):540-50.
- Hudgin R.L., Pricer W.E. Jr, Ashwell G., Stockert R.J. & Morell A.G. (1974) The isolation and properties of a rabbit liver binding protein specific for asialoglycoproteins. *J. Biol. Chem.*, 249(17):5536-43.
- 10 Ikehara Y., Shimizu N., Kono M., Nishihara S., Nakanishi H., Kitamura T., Narimatsu H., Tsuji S. & Tatematsu M. (1999) A novel glycosyltransferase with a polyglutamine repeat; a new candidate for GD1alpha synthase (SToGalNAc V)(I). *FEBS Lett.*, 463(1-2):92-6.
- 15 Ishii A., Ohta M., Watanabe Y., Matsuda K., Ishiyama K., Sakoe K., Nakamura M., Inokuchi J., Sanai Y. & Saito M. (1998) Expression cloning and functional characterization of human cDNA for ganglioside GM3 synthase. *J. Biol. Chem.*, 273(48):31652-5.
- 20 Jarvis D.L., Kwar Z. S. & Hollister J.R. (1998) Engineering N-glycosylation pathways in the baculo virus-insect cell system. *Curr. Opin. Biotechnol.*, 9(5):528-33.
- Jeanneau C. (2003) Bioanalyse et ingenierie de glycosyltransferase. These de l'université de Paris 7, 179p.
- 25 Jeanneau C, Chazalet V., Auge C, Soumpasis D. M., Harduin-Lepers A., Delannoy P., Imberty A. & Breton C. (2004) Structure-function analysis of the human sialyltransferase ST3Gal I: role of n-glycosylation and a novel conserved sialylmotif. *J. Biol. Chem.*, 279(14):13461-8.
- Kawasaki G. & Ashwell T. (1976) Carbohydrate structure of glycopeptides isolated from an hepatic membrane-binding protein specific for asialoglycoproteins. *J. Biol. Chem.*, 251(17):5292-9.
- 30 Kim Y.J., Kim K.S., Kim S.H., Kim C.H., Ko J.H., Choe I.S., Tsuji S. & Lee Y.C. (1996) Molecular cloning and expression of human Gal beta 1,3GalNAc alpha 2,3-sialyltransferase (hST3Gal II). *Biochem. Biophys. Res. Commun.*, 228(2):324-7.
- 35 Kitagawa H. & Paulson J. C. (1993) Cloning and expression of human Gal beta 1,3(4)GlcNAc alpha 2,3-sialyltransferase. *Biochem. Biophys. Res. Commun.*, 194(1):375-82.
- Kitagawa H. & Paulson J. C. (1994a) Differential expression of five sialyltransferase genes in human tissues. *J. Biol. Chem.*, 269(27): 17872-8.
- 40 Kitagawa H. & Paulson J. C. (1994b) Cloning of a novel alpha 2,3-sialyltransferase that sialylates glycoprotein and glyco lipid carbohydrate groups. *J. Biol. Chem.*, 269(2): 1394-401.
- Kitagawa H., Mattei M. G. & Paulson J. C. (1996) Genomic organization and chromosomal mapping of the Gal beta 1,3GalNAc/Gal beta 1,4GlcNAc alpha 2,3-sialyltransferase. *J. Biol. Chem.*, 271(2):931-8.
- 45 Kitazume-Kawaguchi S., Dohmae N., Takio K., Tsuji S. & Colley KJ. (1999) The relationship between ST6Gal I Golgi retention and its cleavage-secretion. *Glycobiology*, 9(12): 1397-406.
- 50 Kitazume-Kawaguchi S., Kabata S. & Arita M. (2001) Differential biosynthesis of polysialic or disialic acid Structure by ST8Sia II and ST8Sia IV. *J. Biol. Chem.*, 276(19): 15696-703.
- Kono M., Yoshida Y., Kojima N. & Tsuji S. (1996) Molecular cloning and expression of a fifth type of alpha2,8-sialyltransferase (ST8Sia V). Its substrate specificity is similar to that of SAT-V/III, which synthesize GD1c, GT1a, GQ1b and GT3. *J. Biol. Chem.*, 271(46):29366-71.
- 55 Kono M., Tsuda T., Ogata S., Takashima S., Liu H., Hamamoto T., Itzkowitz S. H., Nishimura S. & Tsuji S. (2000) Redefined substrate specificity of ST6GalNAc II: a second candidate sialyl- Tn synthase. *Biochem. Biophys. Res. Commun.*, 272(1):94-7.
- 60 Kornfeld R. & Kornfeld S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.*, 1985;54:631-64.
- 65 Krzewinski-Recchi M.A., Julien S., Juliant S., Teinturier-Lelievre M., Samyn-Petit B., Montiel M.D., Mir A.M., Cerutti M., Harduin-Lepers A. & Delannoy P. (2003) Identification and functional expression of a second human beta-galactoside alpha2,6-sialyltransferase, SToGal II. *Eur. J. Biochem.*, 270(5):950-61.

- Kurosawa N., Kojima N., Inoue M., Hamamoto T. & Tsuji S. (1994) Cloning and expression of Gal beta 1,3 GaDSf Ac-specific GalNAc alpha 2,6-sialyltransferase. *J Biol. Chem.*, 269(29): 19048-53.
- 5 Kurosawa N., Inoue M., Yoshida Y. & Tsuji S. (1996) Molecular cloning and genomic analysis of mouse Galbetal, 3 GaDSf Ac-specific GalNAc alpha2,6-sialyltransferase. *J Biol. Chem.*, 1996 Jun 21 ;271(25): 15109-16.
- Kurosawa N., Takashima S., Kono M., Ikehara Y., Inoue M., Tachida Y., Narimatsu H. & Tsuji S. (2000) Molecular cloning and genomic analysis of mouse GalNAc alpha2, 6-sialyltransferase (ST6GalNAc I). *J Biochem.*, 127(5):845-54.
- 10 Laroy W., Ameloot P., & Contreras R. (2001) Characterization of sialyltransferase mutants using surface plasmon resonance. *Glycobiology*, 11 (3): 175-82.
- 15 Lee Y.C., Kim Y.J., Lee K.Y., Kim K.S., Kim B.U., Kim H.N., Kim CH. & Do S.I. (1998) Cloning and expression of cDNA for a human Sia alpha 2,3Gal beta 1, 4GlcNA:alpha 2,8-sialyltransferase (hST8Sia III). *Arch. Biochem. Biophys.*, 360(I):41-6.
- 20 Lee Y. C, Kaufmann M., Kitazume-Kawaguchi S , Kono M , Takashima S , Kurosawa N , Liu H., Pircher H. & Tsuji S. (1999) Molecular cloning and functional expression of two members of mouse NeuAcalpha2,3Galbetal, 3GalNAc GalNAcalpha2,6-sialyltransferase family, ST6GalNAc III and IV. *J. Biol. Chem.*, 274(17):11958-67.
- Legaigneur P., Brioude B. & Ronin C. (1999) Antigenes glycaniques recombinants. *Immunoanal. Biol. Spec.*, 14 :297-307.
- 25 Legaigneur P., Breton C, El Battari A., Guillemot J.C., Auge C, Malissard M., Berger E.G. & Ronin C. (2001) Exploring the acceptor substrate recognition of the human beta-galactoside alpha 2,6-sialyltransferase. *J. Biol. Chem.* , 276(24):21608- 17.
- 30 Lerouge P., Bardor M., Pagny S., Gomord V. & Faye L. (2000) N-glycosylation of recombinant pharmaceutical glycoproteins produced in transgenic plants: towards an humanisation of plant N-glycans. *Curr. Pharm. Biotechnol*, I(4):347-54.
- 35 Livingston B. D. & Paulson J. C. (1993) Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family. *J. Biol. Chem.*, 268(16):11504-7.
- Ma J., Qian R., Rausa F.M. 3rd & Colley KJ. (1997) Two naturally occurring alpha2,6-sialyltransferase forms with a single amino acid change in the catalytic domain differ in their catalytic activity and proteolytic processing. *J. Biol. Chem.*, 272(I):672-9.
- 40 Machamer CE. (1991) Golgi retention signals: do membranes hold the key? *Trends Cell Biol*, 1(6): 141-4.
- Maras M., Van Die L, Contreras R. & Van den Hondel C. A. M. J. J. (1999) Filamentous fungi as production organisms for glycoproteins of bio-medical interest. *Glycoconj. J.*, 16:99-107.
- 45 Marchal L, Jarvis D. L., Cacan R. & Verbert A. (2001) Glycoproteins from insect cells: sialylated or not? *Biol. Chem.*, 382, 151-159.
- Martina J.A., Daniotti J.L. & Maccioni HJ. (1998) Influence of N-glycosylation and N-glycan trimming on the activity and intracellular traffic of GD3 synthase. *J. Biol. Chem.*, 273(6):3725-31.
- 50 Meier M., Bider M.D., Malashkevich V.N., Spiess M. & Burkhard P. (2000) Crystal structure of the carbohydrate recognition domain of the H1 subunit of the asialoglycoprotein receptor. *J. Mol. Biol*, 300(4):857-65.
- 55 Morell A.G., Gregoriadis G., Scheinberg I.H., Hickman J. & Ashwell G. (1971) The role of sialic acid in determining the survival of glycoproteins in the circulation. *J. Biol. Chem.*, 246(5):1461-7.
- Muhlenhoff M., Manegold A., Windfuhr M., Gotza B. & Gerardy-Schahn R. (2001) The impact of N-glycosylation on the functions of polysialyltransferases. *J. Biol. Chem.*, 276(36):34066-73.
- 60 Munro S. (1991) Sequences within and adjacent to the transmembrane segment of alpha-2,6-sialyltransferase specify Golgi retention. *EMBO J.*, 10(12):3577-88.
- Bretscher M.S. & Munro S. (1993) Cholesterol and the Golgi apparatus. *Science*, 261(5126):1280-I.
- 65

- Munro S. (1995) An investigation of the role of transmembrane domains in Golgi protein retention. *EMBO J.*, 14(19):4695-704.
- 5 Munro S. (1998) Localization of proteins to the Golgi apparatus. *Trends Cell Biol*, 8(1): 11-5.
- Nakayama J., Fukuda M.N., Fredette B., Ranscht B. & Fukuda M. (1995) Expression cloning of a human polysialyltransferase that forms the polysialylated neural cell adhesion molecule present in embryonic brain. *Proc. Natl. Acad. Sci. U. S. A.*, 92(15):7031-5.
- 10 Nakayama J., Fukuda M.N., Hirabayashi Y., Kanamori A., Sasaki K., Nishi T. & Fukuda M. (1996) Expression cloning of a human GT3 synthase. GD3 AND GT3 are synthesized by a single enzyme. *J. Biol. Chem.*, 271(7):3684-91.
- Nara K., Watanabe Y., Kawashima L, Tai T., Nagai Y. & Sanai Y. (1996) Acceptor substrate specificity of a cloned GD3 synthase that catalyzes the biosynthesis of both GD3 and GDlc/GTla/GQlb. *Eur. J Biochem.*, 238(3):647-52.
- 15 Narimatsu H. (2004) Construction of a human glycogene library and comprehensive functional analysis. *Glycoconj. J.*, 21(1-2):17-24.
- Okajima T., Fukumoto S., Miyazaki H., Ishida H., Kiso M., Furukawa K., Urano T. & Furukawa K. (1999) Molecular cloning of a novel alpha2,3-sialyltransferase (ST3Gal VI) that sialylates type II lactosamine structures on glycoproteins and glyco lipids. *J Biol. Chem.*, 274(17): 11479-86.
- 20 Okajima T., Chen H. H., Ito H., Kiso M., Tai T., Furukawa K., Urano T. & Furukawa K. (2000) Molecular cloning and expression of mouse GDlalpha/GTlaalpha/GQlbalpha synthase (SToGalNAc VI) gene. *J. Biol. Chem.*, 275(10):6717-23.
- 25 Opat A.S., van Vliet C. & Gleeson P.A. (2001) Trafficking and localisation of resident Golgi glycosylation enzymes. *Biochimie*, 83(8):763-73.
- 30 Paulson J. C. (1989) Glycoproteins: what are the sugar chains for? *Trends Biochem. Sci.*, 14(7):272-6.
- Paulson J.C., Weinstein J., Ujita EX., Riggs KJ. & Lai PH. (1987) The membrane-binding domain of a rat liver Golgi sialyltransferase. *Biochem. Soc. Trans.*, 15(4):618-20.
- 35 Paulson J. C. & Colley KJ. (1989) Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J. Biol. Chem.*, 264(30):17615-8.
- Pilatte Y., Bignon J. & Lambre CR. (1993) Sialic acids as important molecules in the regulation of the immune system: pathophysiological implications of sialidases in immunity. *Glycobiology*, 3(3):201-18.
- 40 Qian R., Chen C. & Colley KJ. (2001) Location and mechanism of alpha 2,6-sialyltransferase dimer formation. Role of cysteine residues in enzyme dimerization, localization, activity, and processing. *J. Biol. Chem.*, 276(31):28641-9.
- 45 Reid M. E. & Lomas-Francis C. (2002) Molecular approaches to blood group identification. *Curr. Opin. Hematol*, 9(2): 152-9.
- Roth J., Taatjes DJ., Lucocq J.M., Weinstein J. & Paulson J. C. (1985) Demonstration of an extensive trans-tubular network continuous with the Golgi apparatus stack that may function in glycosylation. *Cell*, 43(1):287-95.
- 50 Roth J. (1987) Subcellular organization of glycosylation in mammalian cells. *Biochim. Biophys. Acta*, 906(3):405-36.
- Roth J. (1991) Localization of glycosylation sites in the Golgi apparatus using immuno labeling and cytochemistry. *J. Electron. Microsc. Tech.*, 17(2): 121-31.
- 55 Roy S. K., Chiba Y. & Jigami, Y. (2000) Production of Therapeutic Glycoproteins through the Engineering of Glycosylation Pathway in Yeast. *Biotechnol. Bioprocess Eng.*, 5, 219-226.
- Sasaki K. (1996) Molecular cloning and characterization of sialyltransferases. *Trends Glycosci. Glycotechnol.*, 8:195-215.
- 60 Sasaki K., Kurata K., Kojima N., Kurosawa N., Ohta S., Hanai N., Tsuji S. & Nishi T. (1994) Expression cloning of a GM3 -specific alpha-2,8-sialyltransferase (GD3 synthase). *J Biol. Chem.*, 269(22): 15950-6.
- Scheidegger E.P., Sternberg L.R., Roth J. & Lowe JB. (1995) A human STX cDNA confers polysialic acid expression in mammalian cells. *J. Biol. Chem.*, 270(39):22685-8.
- 65

- Sinnott M. L. (1990) Catalytic mechanism of enzymic glycosyl transfer. *Chem. Rev.*, 90, 1171-1202.
- Sjoberg E.R., Kitagawa H., Glushka J., van Halbeek H. & Paulson J.C. (1996) Molecular cloning of a developmentally regulated N-acetylgalactosamine alpha2,6-sialyltransferase specific for sialylated glycoconjugates. *J. Biol. Chem.*, 271(13):7450-9.
- Sujino K., Jackson R.J., Chan N.W., Tsuji S. & Palcic M.M. (2000) A novel viral alpha2,3-sialyltransferase (v-ST3Gal I): transfer of sialic acid to fucosylated acceptors. *Glycobiology*, 10(3):313-20.
- Swartz J. R. (2001) Advances in Escherichia coli production of therapeutic proteins. *Curr. Opin. Biotechnol*, 12, 195-201.
- Takashima S., Ishida H.K., Inazu T., Ando T., Ishida H., Kiso M., Tsuji S. & Tsujimoto M. (2002) Molecular cloning and expression of a sixth type of alpha 2,8-sialyltransferase (ST8Sia VI) that sialylates O-glycans. *J. Biol. Chem.*, 277(27):24030-8.
- Taniguchi A., Kaneta R., Morishita K. & Matsumoto K. (2001) Gene structure and transcriptional regulation of human Gal beta1,4(3) GlcNAc alpha2,3-sialyltransferase VI (hST3Gal VI) gene in prostate cancer cell line. *Biochem. Biophys. Res. Commun.*, (5): 1148-56.
- Tanner W. & Lehele L. (1987) Protein glycosylation in yeast. *Biochim. Biophys. Acta*, 906(l):81-99.
- Tetteroo P.A., de Heij H. T., Van den Eijnden D.H., Visser F.J. , Schoenmarker E. & Geurts van Kessel AH. (1987) A GDP-fucose:[Gal beta 1 — 4]GlcNAc alpha 1— -3-fucosyltransferase activity is correlated with the presence of human chromosome 11 and the expression of the Lex, Ley, and sialyl-Lex antigens in human-mouse cell hybrids. *J. Biol. Chem.*, 262(33): 15984-9.
- Vallejo-Ruiz V., Haque R., Mir A.M., Schwientek T., Mandel U., Cacan R., Delannoy P. & Harduin-Lepers A. (2001) Delineation of the minimal catalytic domain of human Galbeta1-3GalNAc alpha2,3-sialyltransferase (hST3Gal I). *Biochim. Biophys. Acta*, 1549(2): 161-73.
- van den Eijnden D.H., Joiasse D. H., Dorland L., van Halbeek H., Vliegthart J.F. & Schmid K. (1980) Specificity in the enzymic transfer of sialic acid to the oligosaccharide branches of bi- and triantennary glycopeptides of alpha 1-acid glycoprotein. *Biochem. Biophys. Res. Commun.*, 92(3):839-45.
- Van Den Hamer C.J., Morell A.G., Scheinberg I.H., Hickman J. & Ashwell G. (1970) Physical and chemical studies on ceruloplasmin. IX. The role of galactosyl residues in the clearance of ceruloplasmin from the circulation. *J. Biol. Chem.*, 245(17):4397-402.
- Vervecken W., Kaigorodov V., Callewaert N., Geysens S., De Vusser K. & Contreras R. (2004) In vivo synthesis of mammalian-like, hybrid-type N-glycans in *Pichia pastoris*. *Appl. Environ. Microbiol*, 70, 2639-2646.
- Watanabe Y., Nara K., Takahashi H., Nagai Y. & Sanai Y. (1996) The molecular cloning and expression of alpha 2,8-sialyltransferase (GD3 synthase) in a rat brain. *J. Biochem.*, 120(5): 1020-7.
- Weikert S., Papac D., Briggs J., Cowfer D., Tom S., Gawlitzek M., Lofgren J., Mehta S., Chisholm V., Modi N., Eppler S., Carroll K., Chamow S., Peers D., Berman P. & Krummen L. (1999) Engineering Chinese hamster ovary cells to maximize sialic acid content of recombinant glycoproteins. *Nat. Biotechnol*, 17(11): 1116-21.
- Weinstein J., de Souza-e-Silva U. & Paulson J. C. (1982) Sialylation of glycoprotein oligosaccharides N-linked to asparagine. Enzymatic characterization of a Gal beta 1 to 3(4)GlcNAc alpha 2 to 3 sialyltransferase and a Gal beta 1 to 4GlcNAc alpha 2 to 6 sialyltransferase from rat liver. *J. Biol Chem.*, 257(22): 13845-53.
- Wickner W. T. & Lodisch H. F. (1985) Multiple mechanisms of protein insertion into and across membranes. *Science*, 230(4724):400-7.
- Yang W., Pepperkok R., Bender P., Kreis TE. & Storrie B. (1996) Modification of the cytoplasmic domain affects the subcellular localization of Golgi glycosyl-transferases. *Eur. J. Cell Biol*, 71(1):53-61.
- Yoshida Y., Kojima N. & Tsuji S. (1995a) Molecular cloning and characterization of a third type of N-glycan alpha 2,8-sialyltransferase from mouse lung. *J. Biochem.*, 118(3):658-64.
- Yoshida Y., Kojima N., Kurosawa N., Hamamoto T. & Tsuji S. (1995b) Molecular cloning of Sia alpha 2,3Gal beta 1,4GlcNAc alpha 2,8-sialyltransferase from mouse brain. *J. Biol. Chem.*, 270(24): 14628-33.

ES 2 368 267 T3

LISTADO DE SECUENCIAS

<110> CNRS
Université de Provence

<120> PREPARACIÓN Y USOS DE SECUENCIAS GÉNICAS CODIFICANTES
DE GLUCOSILTRANSFERASAS QUIMÉRICAS CON ACTIVIDAD DE
GLUCOSILACIÓN OPTIMIZADA

<130> IOB 05 DI PRO CREA

<160> 176

<170> PatentIn versión 3.3

<210> 1
<211> 1221
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1221)
<400> 1

atg att cac acc aac ctg aag aaa aag ttc agc tgc tgc gtc ctg gtc	48
Met Ile His Thr Asn Leu Lys Lys Lys Phe Ser Cys Cys Val Leu Val	
1 5 10 15	
ttt ctt ctg ttt gca gtc atc tgt gtg tgg aag gaa aag aag aaa ggg	96
Phe Leu Leu Phe Ala Val Ile Cys Val Trp Lys Glu Lys Lys Lys Gly	
20 25 30	
agt tac tat gat tcc ttt aaa ttg caa acc aag gaa ttc cag gtg tta	144
Ser Tyr Tyr Asp Ser Phe Lys Leu Gln Thr Lys Glu Phe Gln Val Leu	
35 40 45	
aag agt ctg ggg aaa ttg gcc atg ggg tct gat tcc cag tct gta tcc	192
Lys Ser Leu Gly Lys Leu Ala Met Gly Ser Asp Ser Gln Ser Val Ser	
50 55 60	
tca agc agc acc cag gac ccc cac agg ggc cgc cag acc ctc ggc agt	240
Ser Ser Ser Thr Gln Asp Pro His Arg Gly Arg Gln Thr Leu Gly Ser	
65 70 75 80	
ctc aga ggc cta gcc aag gcc aaa cca gag gcc tcc ttc cag gtg tgg	288
Leu Arg Gly Leu Ala Lys Ala Lys Pro Glu Ala Ser Phe Gln Val Trp	
85 90 95	
aac aag gac agc tct tcc aaa aac ctt atc cct agg ctg caa aag atc	336
Asn Lys Asp Ser Ser Ser Lys Asn Leu Ile Pro Arg Leu Gln Lys Ile	
100 105 110	
tgg aag aat tac cta agc atg aac aag tac aaa gtg tcc tac aag ggg	384
Trp Lys Asn Tyr Leu Ser Met Asn Lys Tyr Lys Val Ser Tyr Lys Gly	
115 120 125	
cca gga cca ggc atc aag ttc agt gca gag gcc ctg cgc tgc cac ctc	432
Pro Gly Pro Gly Ile Lys Phe Ser Ala Glu Ala Leu Arg Cys His Leu	
130 135 140	

ES 2 368 267 T3

cgg gac cat gtg aat gta tcc atg gta gag gtc aca gat ttt ccc ttc Arg Asp His Val Asn Val Ser Met Val Glu Val Thr Asp Phe Pro Phe 145 150 155 160	480
aat acc tct gaa tgg gag ggt tat ctg ccc aag gag agc att agg acc Asn Thr Ser Glu Trp Glu Gly Tyr Leu Pro Lys Glu Ser Ile Arg Thr 165 170 175	528
aag gct ggg cct tgg ggc agg tgt gct gtt gtg tcg tca gcg gga tct Lys Ala Gly Pro Trp Gly Arg Cys Ala Val Val Ser Ser Ala Gly Ser 180 185 190	576
ctg aag tcc tcc caa cta ggc aga gaa atc gat gat cat gac gca gtc Leu Lys Ser Ser Gln Leu Gly Arg Glu Ile Asp Asp His Asp Ala Val 195 200 205	624
ctg agg ttt aat ggg gca ccc aca gcc aac ttc caa caa gat gtg ggc Leu Arg Phe Asn Gly Ala Pro Thr Ala Asn Phe Gln Gln Asp Val Gly 210 215 220	672
aca aaa act acc att cgc ctg atg aac tct cag ttg gtt acc aca gag Thr Lys Thr Thr Ile Arg Leu Met Asn Ser Gln Leu Val Thr Thr Glu 225 230 235 240	720
aag cgc ttc ctc aaa gac agt ttg tac aat gaa gga atc cta att gta Lys Arg Phe Leu Lys Asp Ser Leu Tyr Asn Glu Gly Ile Leu Ile Val 245 250 255	768
tgg gac cca tct gta tac cac tca gat atc cca aag tgg tac cag aat Trp Asp Pro Ser Val Tyr His Ser Asp Ile Pro Lys Trp Tyr Gln Asn 260 265 270	816
ccg gat tat aat ttc ttt aac aac tac aag act tat cgt aag ctg cac Pro Asp Tyr Asn Phe Phe Asn Asn Tyr Lys Thr Tyr Arg Lys Leu His 275 280 285	864
ccc aat cag ccc ttt tac atc ctc aag ccc cag atg cct tgg gag cta Pro Asn Gln Pro Phe Tyr Ile Leu Lys Pro Gln Met Pro Trp Glu Leu 290 295 300	912
tgg gac att ctt caa gaa atc tcc cca gaa gag att cag cca aac ccc Trp Asp Ile Leu Gln Glu Ile Ser Pro Glu Glu Ile Gln Pro Asn Pro 305 310 315 320	960
cca tcc tct ggg atg ctt ggt atc atc atc atg atg acg ctg tgt gac Pro Ser Ser Gly Met Leu Gly Ile Ile Ile Met Met Thr Leu Cys Asp 325 330 335	1008
cag gtg gat att tat gag ttc ctc cca tcc aag cgc aag act gac gtg Gln Val Asp Ile Tyr Glu Phe Leu Pro Ser Lys Arg Lys Thr Asp Val 340 345 350	1056
tgc tac tac tac cag aag ttc ttc gat agt gcc tgc acg atg ggt gcc Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp Ser Ala Cys Thr Met Gly Ala 355 360 365	1104
tac cac ccg ctg ctc tat gag aag aat ttg gtg aag cat ctc aac cag Tyr His Pro Leu Leu Tyr Glu Lys Asn Leu Val Lys His Leu Asn Gln 370 375 380	1152
ggc aca gat gag gac atc tac ctg ctt gga aaa gcc aca ctg cct ggc	1200

ES 2 368 267 T3

Gly Thr Asp Glu Asp Ile Tyr Leu Leu Gly Lys Ala Thr Leu Pro Gly
 385 390 395 400

ttc cgg acc att cac tgc taa 1221
 Phe Arg Thr Ile His Cys
 405

<210> 2
 <211> 406
 <212> PRT
 <213> Homo sapiens

<400> 2

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

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

ES 2 368 267 T3

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>	3
<211>	1590
<212>	ADN

ES 2 368 267 T3

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1590)

<400> 3

atg aaa cca cac ttg aag caa tgg aga caa cga atg ctt ttc gga ata	48
Met Lys Pro His Leu Lys Gln Trp Arg Gln Arg Met Leu Phe Gly Ile	
1 5 10 15	
ttc gct tgg ggg ctc ctc ttt ttg ctg att ttc atc tac ttc acc gac	96
Phe Ala Trp Gly Leu Leu Phe Leu Leu Ile Phe Ile Tyr Phe Thr Asp	
20 25 30	
agc aac ccc gct gag cct gta ccc agc tcc ctc tcc ttc ctg gag acc	144
Ser Asn Pro Ala Glu Pro Val Pro Ser Ser Leu Ser Phe Leu Glu Thr	
35 40 45	
agg agg ctc ctg ccg gtg cag ggg aag cag cgg gcc atc atg ggc gcc	192
Arg Arg Leu Leu Pro Val Gln Gly Lys Gln Arg Ala Ile Met Gly Ala	
50 55 60	
gca cat gag ccc tcc ccg cct ggg ggc ctg gac gca cgc cag gcg ctg	240
Ala His Glu Pro Ser Pro Pro Gly Gly Leu Asp Ala Arg Gln Ala Leu	
65 70 75 80	
ccc cgc gcc cac cca gcc ggt tcc ttt cat gcg ggg cct gga gac ctg	288
Pro Arg Ala His Pro Ala Gly Ser Phe His Ala Gly Pro Gly Asp Leu	
85 90 95	
cag aaa tgg gcc cag tcc caa gat ggg ttt gaa cat aaa gag ttt ttt	336
Gln Lys Trp Ala Gln Ser Gln Asp Gly Phe Glu His Lys Glu Phe Phe	
100 105 110	
tca tcc cag gtg ggg aga aaa tct caa agt gct ttc tac ccg gag gat	384
Ser Ser Gln Val Gly Arg Lys Ser Gln Ser Ala Phe Tyr Pro Glu Asp	
115 120 125	
gac gac tac ttt ttt gct gct ggt cag cca ggg tgg cac agc cac act	432
Asp Asp Tyr Phe Phe Ala Ala Gly Gln Pro Gly Trp His Ser His Thr	
130 135 140	
cag ggg aca ttg gga ttc cct tcc ccc ggg gag cca ggc cca cgg gag	480
Gln Gly Thr Leu Gly Phe Pro Ser Pro Gly Glu Pro Gly Pro Arg Glu	
145 150 155 160	
ggg gct ttt ccg gct gca cag gtc cag agg agg cgg gtg aag aag agg	528
Gly Ala Phe Pro Ala Ala Gln Val Gln Arg Arg Arg Val Lys Lys Arg	
165 170 175	
cac cgg agg cag aga agg agc cac gtg ttg gag gag ggc gac gac ggc	576
His Arg Arg Gln Arg Arg Ser His Val Leu Glu Glu Gly Asp Asp Gly	
180 185 190	
gac agg ctg tac tcc tcc atg tcc agg gcc ttc ctg tac cgg ctc tgg	624
Asp Arg Leu Tyr Ser Ser Met Ser Arg Ala Phe Leu Tyr Arg Leu Trp	
195 200 205	
aag ggg aac gtc tct tcc aaa atg ctg aac ccg cgc ctg cag aag gcg	672

ES 2 368 267 T3

Lys 210	Gly	Asn	Val	Ser	Ser	Lys 215	Met	Leu	Asn	Pro	Arg 220	Leu	Gln	Lys	Ala	
atg	aag	gat	tac	ctg	acc	gcc	aac	aag	cac	ggg	gtg	cgc	ttc	cgc	ggg	720
Met 225	Lys	Asp	Tyr	Leu	Thr	Ala	Asn	Lys	His	Gly	Val	Arg	Phe	Arg	Gly	240
aag	cgg	gag	gcc	ggg	ctg	agc	agg	gca	cag	ctg	ctg	tgc	cag	ctg	cgg	768
Lys	Arg	Glu	Ala	Gly	Leu	Ser	Arg	Ala	Gln	Leu	Leu	Cys	Gln	Leu	Arg	255
agg	cgc	gcg	cgc	gtg	cgg	acg	ctg	gac	ggc	acc	gag	gcg	ccc	ttt	tct	816
Ser	Arg	Ala	Arg	Val	Arg	Thr	Leu	Asp	Gly	Thr	Glu	Ala	Pro	Phe	Ser	270
gcg	ctg	ggc	tgg	cgg	cgc	ctg	gtg	ccc	gcc	gtg	ccc	ctg	agc	cag	ctg	864
Ala	Leu	Gly	Trp	Arg	Arg	Leu	Val	Pro	Ala	Val	Pro	Leu	Ser	Gln	Leu	285
cac	ccc	cgc	ggc	ctg	cgc	agc	tgc	gct	gtc	gtc	atg	tct	gca	ggc	gca	912
His	Pro	Arg	Gly	Leu	Arg	Ser	Cys	Ala	Val	Val	Met	Ser	Ala	Gly	Ala	300
atc	ctc	aac	tct	tcc	ttg	ggc	gag	gaa	ata	gat	tct	cat	gat	gcg	gtt	960
Ile	Leu	Asn	Ser	Ser	Leu	Gly	Glu	Glu	Ile	Asp	Ser	His	Asp	Ala	Val	320
ttg	aga	ttt	aac	tct	gct	cct	aca	cgt	ggt	tat	gag	aaa	gat	gtt	ggg	1008
Leu	Arg	Phe	Asn	Ser	Ala	Pro	Thr	Arg	Gly	Tyr	Glu	Lys	Asp	Val	Gly	335
aat	aaa	acc	acc	ata	cgc	atc	att	aat	tgc	cag	att	ctg	acc	aac	ccc	1056
Asn	Lys	Thr	Thr	Ile	Arg	Ile	Ile	Asn	Ser	Gln	Ile	Leu	Thr	Asn	Pro	350
agc	cat	cac	ttc	att	gac	agt	tca	ctg	tat	aaa	gac	gtc	att	ttg	gtg	1104
Ser	His	His	Phe	Ile	Asp	Ser	Ser	Leu	Tyr	Lys	Asp	Val	Ile	Leu	Val	365
gcc	tgg	gac	cct	gcc	cca	tat	tcc	gca	aat	ctt	aac	ctg	tgg	tac	aaa	1152
Ala	Trp	Asp	Pro	Ala	Pro	Tyr	Ser	Ala	Asn	Leu	Asn	Leu	Trp	Tyr	Lys	380
aaa	cgc	gat	tac	aac	ctg	ttc	act	cca	tat	att	cag	cat	cgt	cag	aga	1200
Lys	Pro	Asp	Tyr	Asn	Leu	Phe	Thr	Pro	Tyr	Ile	Gln	His	Arg	Gln	Arg	400
aac	cca	aat	cag	cca	ttt	tac	att	ctt	cat	cct	aaa	ttt	ata	tgg	cag	1248
Asn	Pro	Asn	Gln	Pro	Phe	Tyr	Ile	Leu	His	Pro	Lys	Phe	Ile	Trp	Gln	415
ctc	tgg	gat	att	atc	cag	gag	aac	act	aaa	gag	aag	att	caa	cca	aac	1296
Leu	Trp	Asp	Ile	Ile	Gln	Glu	Asn	Thr	Lys	Glu	Lys	Ile	Gln	Pro	Asn	430
cca	cca	tct	tct	ggt	ttc	att	gga	atc	ctc	atc	atg	atg	tcc	atg	tgc	1344
Pro	Pro	Ser	Ser	Gly	Phe	Ile	Gly	Ile	Leu	Ile	Met	Met	Ser	Met	Cys	445
aga	gag	gtg	cac	gtg	tat	gaa	tat	atc	cca	tcc	gtg	cgg	cag	acg	gag	1392
Arg	Glu	Val	His	Val	Tyr	Glu	Tyr	Ile	Pro	Ser	Val	Arg	Gln	Thr	Glu	

ES 2 368 267 T3

450	455	460	
ctg tgc cac tac cac gag ctg tac tac gac gca gcc tgc acc ctc ggg			1440
Leu Cys His Tyr His Glu Leu Tyr Tyr Asp Ala Ala Cys Thr Leu Gly			
465	470	475	480
gcg tac cac cca cta ctc tat gag aag ctc ctg gtg cag cgc ctg aac			1488
Ala Tyr His Pro Leu Leu Tyr Glu Lys Leu Leu Val Gln Arg Leu Asn			
	485	490	495
atg ggc acg cag ggg gat ttg cat cgc aag ggc aag gtg gtt ctt cct			1536
Met Gly Thr Gln Gly Asp Leu His Arg Lys Gly Lys Val Val Leu Pro			
	500	505	510
ggc ttc cag gcg gtg cac tgc cct gca cca agt cca gtc att cca cac			1584
Gly Phe Gln Ala Val His Cys Pro Ala Pro Ser Pro Val Ile Pro His			
	515	520	525
tct taa			1590
Ser			
<210>	4		
<211>	529		
<212>	PRT		
<213>	Homo sapiens		
<400>	4		
Met Lys Pro His Leu Lys Gln Trp Arg Gln Arg Met Leu Phe Gly Ile			
1	5	10	15
Phe Ala Trp Gly Leu Leu Phe Leu Leu Ile Phe Ile Tyr Phe Thr Asp			
	20	25	30
Ser Asn Pro Ala Glu Pro Val Pro Ser Ser Leu Ser Phe Leu Glu Thr			
	35	40	45
Arg Arg Leu Leu Pro Val Gln Gly Lys Gln Arg Ala Ile Met Gly Ala			
	50	55	60
Ala His Glu Pro Ser Pro Pro Gly Gly Leu Asp Ala Arg Gln Ala Leu			
65	70	75	80
Pro Arg Ala His Pro Ala Gly Ser Phe His Ala Gly Pro Gly Asp Leu			
	85	90	95
Gln Lys Trp Ala Gln Ser Gln Asp Gly Phe Glu His Lys Glu Phe Phe			
	100	105	110
Ser Ser Gln Val Gly Arg Lys Ser Gln Ser Ala Phe Tyr Pro Glu Asp			
	115	120	125

ES 2 368 267 T3

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

ES 2 368 267 T3

Ala Trp Asp Pro Ala Pro Tyr Ser Ala Asn Leu Asn Leu Trp Tyr Lys
 370 375 380

Lys Pro Asp Tyr Asn Leu Phe Thr Pro Tyr Ile Gln His Arg Gln Arg
 385 390 395 400

Asn Pro Asn Gln Pro Phe Tyr Ile Leu His Pro Lys Phe Ile Trp Gln
 405 410 415

Leu Trp Asp Ile Ile Gln Glu Asn Thr Lys Glu Lys Ile Gln Pro Asn
 420 425 430

Pro Pro Ser Ser Gly Phe Ile Gly Ile Leu Ile Met Met Ser Met Cys
 435 440 445

Arg Glu Val His Val Tyr Glu Tyr Ile Pro Ser Val Arg Gln Thr Glu
 450 455 460

Leu Cys His Tyr His Glu Leu Tyr Tyr Asp Ala Ala Cys Thr Leu Gly
 465 470 475 480

Ala Tyr His Pro Leu Leu Tyr Glu Lys Leu Leu Val Gln Arg Leu Asn
 485 490 495

Met Gly Thr Gln Gly Asp Leu His Arg Lys Gly Lys Val Val Leu Pro
 500 505 510

Gly Phe Gln Ala Val His Cys Pro Ala Pro Ser Pro Val Ile Pro His
 515 520 525

Ser

<210> 5
 <211> 1803
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1803)

<400> 5
 atg agg tcc tgc ctg tgg aga tgc agg cac ctg agc caa ggc gtc cag 48
 Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val Gln
 1 5 10 15

tgg tcc ttg ctt ctg gct gtc ctg gtc ttc ttt ctc ttc gcc ttg ccc 96
 Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala Leu Pro

ES 2 368 267 T3

20					25					30						
tct	ttt	att	aag	gag	cct	caa	aca	aag	cct	tcc	agg	cat	caa	cgc	aca	144
Ser	Phe	Ile	Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr	
		35					40				45					
gag	aac	att	aaa	gaa	agg	tct	cta	cag	tcc	ctg	gca	aag	cct	aag	tcc	192
Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser	Leu	Ala	Lys	Pro	Lys	Ser	
	50					55					60					
cag	gca	ccc	aca	agg	gca	agg	agg	aca	acc	atc	tat	gca	gag	cca	gtg	240
Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr	Ile	Tyr	Ala	Glu	Pro	Val	
65					70					75					80	
cca	gag	aac	aat	gcc	ctc	aac	aca	caa	acc	cag	ccc	aag	gcc	cac	acc	288
Pro	Glu	Asn	Asn	Ala	Leu	Asn	Thr	Gln	Thr	Gln	Pro	Lys	Ala	His	Thr	
				85					90					95		
acc	gga	gac	aga	gga	aag	gag	gcc	aac	cag	gca	ccg	ccg	gag	gag	cag	336
Thr	Gly	Asp	Arg	Gly	Lys	Glu	Ala	Asn	Gln	Ala	Pro	Pro	Glu	Glu	Gln	
			100					105					110			
gac	aag	gtg	ccc	cac	aca	gca	cag	agg	gca	gca	tgg	aag	agc	cca	gaa	384
Asp	Lys	Val	Pro	His	Thr	Ala	Gln	Arg	Ala	Ala	Trp	Lys	Ser	Pro	Glu	
		115					120					125				
aaa	gag	aaa	acc	atg	gtg	aac	aca	ctg	tca	ccc	aga	ggg	caa	gat	gca	432
Lys	Glu	Lys	Thr	Met	Val	Asn	Thr	Leu	Ser	Pro	Arg	Gly	Gln	Asp	Ala	
	130					135					140					
ggg	atg	gcc	tct	ggc	agg	aca	gag	gca	caa	tca	tgg	aag	agc	cag	gac	480
Gly	Met	Ala	Ser	Gly	Arg	Thr	Glu	Ala	Gln	Ser	Trp	Lys	Ser	Gln	Asp	
145					150					155					160	
aca	aag	acg	acc	caa	gga	aat	ggg	ggc	cag	acc	agg	aag	ctg	acg	gcc	528
Thr	Lys	Thr	Thr	Gln	Gly	Asn	Gly	Gly	Gln	Thr	Arg	Lys	Leu	Thr	Ala	
				165					170					175		
tcc	agg	acg	gtg	tca	gag	aag	cac	cag	ggc	aaa	gcg	gca	acc	aca	gcc	576
Ser	Arg	Thr	Val	Ser	Glu	Lys	His	Gln	Gly	Lys	Ala	Ala	Thr	Thr	Ala	
			180					185						190		
aag	acg	ctc	att	ccc	aaa	agt	cag	cac	aga	atg	ctg	gct	ccc	aca	gga	624
Lys	Thr	Leu	Ile	Pro	Lys	Ser	Gln	His	Arg	Met	Leu	Ala	Pro	Thr	Gly	
		195					200					205				
gca	gtg	tca	aca	agg	acg	aga	cag	aaa	gga	gtg	acc	aca	gca	gtc	atc	672
Ala	Val	Ser	Thr	Arg	Thr	Arg	Gln	Lys	Gly	Val	Thr	Thr	Ala	Val	Ile	
	210					215					220					
cca	cct	aag	gag	aag	aaa	cct	cag	gcc	acc	cca	ccc	cct	gcc	cct	ttc	720
Pro	Pro	Lys	Glu	Lys	Lys	Pro	Gln	Ala	Thr	Pro	Pro	Pro	Ala	Pro	Phe	
225					230					235					240	
cag	agc	ccc	acg	acg	cag	aga	aac	caa	aga	ctg	aag	gcc	gcc	aac	ttc	768
Gln	Ser	Pro	Thr	Thr	Gln	Arg	Asn	Gln	Arg	Leu	Lys	Ala	Ala	Asn	Phe	
				245					250					255		
aaa	tct	gag	cct	cgg	tgg	gat	ttt	gag	gaa	aaa	tac	agc	ttc	gaa	ata	816
Lys	Ser	Glu	Pro	Arg	Trp	Asp	Phe	Glu	Glu	Lys	Tyr	Ser	Phe	Glu	Ile	
			260					265					270			

ES 2 368 267 T3

gga ggc ctt cag acg act tgc cct gac tct gtg aag atc aaa gcc tcc	864
Gly Gly Leu Gln Thr Thr Cys Pro Asp Ser Val Lys Ile Lys Ala Ser	
275 280 285	
aag tcg ctg tgg ctc cag aaa ctc ttt ctg ccc aac ctc act ctc ttc	912
Lys Ser Leu Trp Leu Gln Lys Leu Phe Leu Pro Asn Leu Thr Leu Phe	
290 295 300	
ctg gac tcc aga cac ttc aac cag agt gag tgg gac cgc ctg gaa cac	960
Leu Asp Ser Arg His Phe Asn Gln Ser Glu Trp Asp Arg Leu Glu His	
305 310 315 320	
ttt gca cca ccc ttt ggc ttc atg gag ctc aac tac tcc ttg gtg cag	1008
Phe Ala Pro Pro Phe Gly Phe Met Glu Leu Asn Tyr Ser Leu Val Gln	
325 330 335	
aag gtc gtg aca cgc ttc cct cca gtg ccc cag cag cag ctg ctc ctg	1056
Lys Val Val Thr Arg Phe Pro Pro Val Pro Gln Gln Gln Leu Leu Leu	
340 345 350	
gcc agc ctc ccc gct ggg agc ctc cgg tgc atc acc tgt gcc gtg gtg	1104
Ala Ser Leu Pro Ala Gly Ser Leu Arg Cys Ile Thr Cys Ala Val Val	
355 360 365	
ggc aac ggg ggc atc ctg aac aac tcc cac atg ggc cag gag ata gac	1152
Gly Asn Gly Gly Ile Leu Asn Asn Ser His Met Gly Gln Glu Ile Asp	
370 375 380	
agt cac gac tac gtg ttc cga ttg agc gga gct ctc att aaa ggc tac	1200
Ser His Asp Tyr Val Phe Arg Leu Ser Gly Ala Leu Ile Lys Gly Tyr	
385 390 395 400	
gaa cag gat gtg ggg act cgg aca tcc ttc tac ggc ttt acc gcc ttc	1248
Glu Gln Asp Val Gly Thr Arg Thr Ser Phe Tyr Gly Phe Thr Ala Phe	
405 410 415	
tcc ctg acc cag tca ctc ctt ata ttg ggc aat cgg ggt ttc aag aac	1296
Ser Leu Thr Gln Ser Leu Leu Ile Leu Gly Asn Arg Gly Phe Lys Asn	
420 425 430	
gtg cct ctt ggg aag gac gtc cgc tac ttg cac ttc ctg gaa ggc acc	1344
Val Pro Leu Gly Lys Asp Val Arg Tyr Leu His Phe Leu Glu Gly Thr	
435 440 445	
cgg gac tat gag tgg ctg gaa gca ctg ctt atg aat cag acg gtg atg	1392
Arg Asp Tyr Glu Trp Leu Glu Ala Leu Leu Met Asn Gln Thr Val Met	
450 455 460	
tca aaa aac ctt ttc tgg ttc agg cac aga ccc cag gaa gct ttt cgg	1440
Ser Lys Asn Leu Phe Trp Phe Arg His Arg Pro Gln Glu Ala Phe Arg	
465 470 475 480	
gaa gcc ctg cac atg gac agg tac ctg ttg ctg cac cca gac ttt ctc	1488
Glu Ala Leu His Met Asp Arg Tyr Leu Leu Leu His Pro Asp Phe Leu	
485 490 495	
cga tac atg aag aac agg ttt ctg agg tct aag acc ctg gat ggt gcc	1536
Arg Tyr Met Lys Asn Arg Phe Leu Arg Ser Lys Thr Leu Asp Gly Ala	
500 505 510	

ES 2 368 267 T3

cac tgg agg ata tac cgc ccc acc act ggg gcc ctc ctg ctg ctc act His Trp Arg Ile Tyr Arg Pro Thr Thr Gly Ala Leu Leu Leu Leu Thr 515 520 525 gcc ctt cag ctc tgt gac cag gtg agt gct tat ggc ttc atc act gag Ala Leu Gln Leu Cys Asp Gln Val Ser Ala Tyr Gly Phe Ile Thr Glu 530 535 540 ggc cat gag cgc ttt tct gat cac tac tat gat aca tca tgg aag cgg Gly His Glu Arg Phe Ser Asp His Tyr Tyr Asp Thr Ser Trp Lys Arg 545 550 555 560 ctg atc ttt tac ata aac cat gac ttc aag ctg gag aga gaa gtc tgg Leu Ile Phe Tyr Ile Asn His Asp Phe Lys Leu Glu Arg Glu Val Trp 565 570 575 aag cgg cta cac gat gaa ggg ata atc cgg ctg tac cag cgt cct ggt Lys Arg Leu His Asp Glu Gly Ile Ile Arg Leu Tyr Gln Arg Pro Gly 580 585 590 ccc gga act gcc aaa gcc aag aac tga Pro Gly Thr Ala Lys Ala Lys Asn 595 600	1584 1632 1680 1728 1776 1803
---	---

<210> 6

<211> 600

<212> PRT

<213> Homo sapiens

<400> 6

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

ES 2 368 267 T3

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

ES 2 368 267 T3

Gly	Asn	Gly	Gly	Ile	Leu	Asn	Asn	Ser	His	Met	Gly	Gln	Glu	Ile	Asp
370						375					380				
Ser	His	Asp	Tyr	Val	Phe	Arg	Leu	Ser	Gly	Ala	Leu	Ile	Lys	Gly	Tyr
385					390					395					400
Glu	Gln	Asp	Val	Gly	Thr	Arg	Thr	Ser	Phe	Tyr	Gly	Phe	Thr	Ala	Phe
				405					410					415	
Ser	Leu	Thr	Gln	Ser	Leu	Leu	Ile	Leu	Gly	Asn	Arg	Gly	Phe	Lys	Asn
			420					425					430		
Val	Pro	Leu	Gly	Lys	Asp	Val	Arg	Tyr	Leu	His	Phe	Leu	Glu	Gly	Thr
		435					440					445			
Arg	Asp	Tyr	Glu	Trp	Leu	Glu	Ala	Leu	Leu	Met	Asn	Gln	Thr	Val	Met
	450					455					460				
Ser	Lys	Asn	Leu	Phe	Trp	Phe	Arg	His	Arg	Pro	Gln	Glu	Ala	Phe	Arg
465					470					475					480
Glu	Ala	Leu	His	Met	Asp	Arg	Tyr	Leu	Leu	Leu	His	Pro	Asp	Phe	Leu
				485					490					495	
Arg	Tyr	Met	Lys	Asn	Arg	Phe	Leu	Arg	Ser	Lys	Thr	Leu	Asp	Gly	Ala
			500					505					510		
His	Trp	Arg	Ile	Tyr	Arg	Pro	Thr	Thr	Gly	Ala	Leu	Leu	Leu	Leu	Thr
		515					520					525			
Ala	Leu	Gln	Leu	Cys	Asp	Gln	Val	Ser	Ala	Tyr	Gly	Phe	Ile	Thr	Glu
	530					535					540				
Gly	His	Glu	Arg	Phe	Ser	Asp	His	Tyr	Tyr	Asp	Thr	Ser	Trp	Lys	Arg
545					550					555					560
Leu	Ile	Phe	Tyr	Ile	Asn	His	Asp	Phe	Lys	Leu	Glu	Arg	Glu	Val	Trp
				565					570					575	
Lys	Arg	Leu	His	Asp	Glu	Gly	Ile	Ile	Arg	Leu	Tyr	Gln	Arg	Pro	Gly
			580					585					590		
Pro	Gly	Thr	Ala	Lys	Ala	Lys	Asn								
		595					600								

ES 2 368 267 T3

<210> 7
 <211> 1125
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1125)

<400> 7
 atg ggg ctc ccg cgc ggg tcg ttc ttc tgg gtg ctg ctc ctg ctc acg 48
 Met Gly Leu Pro Arg Gly Ser Phe Phe Trp Val Leu Leu Leu Leu Thr
 1 5 10 15
 gct gcc tgc tcg ggg ctc ctc ttt gcc ctg tac ttc tcg gcg gtg cag 96
 Ala Ala Cys Ser Gly Leu Leu Phe Ala Leu Tyr Phe Ser Ala Val Gln
 20 25 30
 cgg tac ccg ggg cca gcg gcc gga gcc agg gac acc aca tca ttt gaa 144
 Arg Tyr Pro Gly Pro Ala Ala Gly Ala Arg Asp Thr Thr Ser Phe Glu
 35 40 45
 gca ttc ttt caa tcc aag gca tcg aat tct tgg aca gga aag ggc cag 192
 Ala Phe Phe Gln Ser Lys Ala Ser Asn Ser Trp Thr Gly Lys Gly Gln
 50 55 60
 gcc tgc cga cac ctg ctt cac ctg gcc att cag cgg cac ccc cac ttc 240
 Ala Cys Arg His Leu Leu His Leu Ala Ile Gln Arg His Pro His Phe
 65 70 75 80
 cgt ggc ctg ttc aat ctc tcc att cca gtg ctg ctg tgg ggg gac ctc 288
 Arg Gly Leu Phe Asn Leu Ser Ile Pro Val Leu Leu Trp Gly Asp Leu
 85 90 95
 ttc acc cca gcg ctc tgg gac cgc ctg agc caa cac aaa gcc ccg tat 336
 Phe Thr Pro Ala Leu Trp Asp Arg Leu Ser Gln His Lys Ala Pro Tyr
 100 105 110
 ggc tgg cgg ggg ctc tct cac caa gtc atc gcc tcc acc ctg agc ctt 384
 Gly Trp Arg Gly Leu Ser His Gln Val Ile Ala Ser Thr Leu Ser Leu
 115 120 125
 ctg aac ggc tca gag agt gcc aag ctg ttt gcc ccg ccc agg gac acc 432
 Leu Asn Gly Ser Glu Ser Ala Lys Leu Phe Ala Pro Pro Arg Asp Thr
 130 135 140
 cct cca aag tgt atc cgg tgt gcc gtg gtg ggc aac gga ggc att ctg 480
 Pro Pro Lys Cys Ile Arg Cys Ala Val Val Gly Asn Gly Gly Ile Leu
 145 150 155 160
 aat ggg tcc cgc cag ggt ccc aac atc gat gcc cat gac tat gta ttc 528
 Asn Gly Ser Arg Gln Gly Pro Asn Ile Asp Ala His Asp Tyr Val Phe
 165 170 175
 aga ctc aat gga gct gtg atc aaa ggc ttc gag cgc gat gtg ggc acc 576
 Arg Leu Asn Gly Ala Val Ile Lys Gly Phe Glu Arg Asp Val Gly Thr
 180 185 190
 aag act tcc ttc tat ggt ttc act gtg aac acg atg aag aac tcc ctc 624

ES 2 368 267 T3

Lys	Thr	Ser	Phe	Tyr	Gly	Phe	Thr	Val	Asn	Thr	Met	Lys	Asn	Ser	Leu		
		195					200					205					
gtc	tcc	tac	tgg	aat	ctg	ggc	ttc	acc	tcc	gtg	cca	caa	gga	cag	gac	672	
Val	Ser	Tyr	Trp	Asn	Leu	Gly	Phe	Thr	Ser	Val	Pro	Gln	Gly	Gln	Asp		
	210					215					220						
ctg	cag	tat	atc	ttc	atc	ccc	tca	gac	atc	cgc	gac	tat	gtg	atg	ctg	720	
Leu	Gln	Tyr	Ile	Phe	Ile	Pro	Ser	Asp	Ile	Arg	Asp	Tyr	Val	Met	Leu		
225					230					235					240		
aga	tcg	gcc	att	ctg	ggc	gtg	cct	gtc	cct	gag	ggc	cta	gat	aaa	ggg	768	
Arg	Ser	Ala	Ile	Leu	Gly	Val	Pro	Val	Pro	Glu	Gly	Leu	Asp	Lys	Gly		
				245					250					255			
gac	agg	ccg	cac	gcc	tat	ttt	gga	cca	gaa	gcc	tct	gcc	agt	aaa	ttc	816	
Asp	Arg	Pro	His	Ala	Tyr	Phe	Gly	Pro	Glu	Ala	Ser	Ala	Ser	Lys	Phe		
			260					265					270				
aag	ctg	cta	cat	ccg	gac	ttc	atc	agc	tac	ctg	aca	gaa	agg	ttc	ttg	864	
Lys	Leu	Leu	His	Pro	Asp	Phe	Ile	Ser	Tyr	Leu	Thr	Glu	Arg	Phe	Leu		
	275						280					285					
aaa	tca	aag	ttg	att	aac	aca	cat	ttt	gga	gac	cta	tat	atg	cct	agt	912	
Lys	Ser	Lys	Leu	Ile	Asn	Thr	His	Phe	Gly	Asp	Leu	Tyr	Met	Pro	Ser		
	290					295					300						
acc	ggg	gct	ctc	atg	ctg	ctg	aca	gct	ttg	cat	acc	tgt	gac	cag	gtc	960	
Thr	Gly	Ala	Leu	Met	Leu	Leu	Thr	Ala	Leu	His	Thr	Cys	Asp	Gln	Val		
305					310					315					320		
agt	gcc	tat	gga	ttc	atc	aca	agc	aac	tac	tgg	aaa	ttt	tcc	gac	cac	1008	
Ser	Ala	Tyr	Gly	Phe	Ile	Thr	Ser	Asn	Tyr	Trp	Lys	Phe	Ser	Asp	His		
				325					330					335			
tat	ttc	gaa	cga	aaa	atg	aag	cca	ttg	ata	ttt	tat	gca	aac	cac	gat	1056	
Tyr	Phe	Glu	Arg	Lys	Met	Lys	Pro	Leu	Ile	Phe	Tyr	Ala	Asn	His	Asp		
			340					345					350				
ctg	tcc	ctg	gaa	gct	gcc	ctg	tgg	agg	gac	ctg	cac	aag	gcc	ggc	atc	1104	
Leu	Ser	Leu	Glu	Ala	Ala	Leu	Trp	Arg	Asp	Leu	His	Lys	Ala	Gly	Ile		
		355					360					365					
ctt	cag	ctg	tac	cag	cgc	tga										1125	
Leu	Gln	Leu	Tyr	Gln	Arg												
	370																

<210> 8
 <211> 374
 <212> PRT
 <213> Homo sapiens

<400> 8

Met	Gly	Leu	Pro	Arg	Gly	Ser	Phe	Phe	Trp	Val	Leu	Leu	Leu	Leu	Thr
1				5					10					15	
Ala	Ala	Cys	Ser	Gly	Leu	Leu	Phe	Ala	Leu	Tyr	Phe	Ser	Ala	Val	Gln
			20					25					30		

ES 2 368 267 T3

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

ES 2 368 267 T3

Lys Leu Leu His Pro Asp Phe Ile Ser Tyr Leu Thr Glu Arg Phe Leu
 275 280 285

Lys Ser Lys Leu Ile Asn Thr His Phe Gly Asp Leu Tyr Met Pro Ser
 290 295 300

Thr Gly Ala Leu Met Leu Leu Thr Ala Leu His Thr Cys Asp Gln Val
 305 310 315 320

Ser Ala Tyr Gly Phe Ile Thr Ser Asn Tyr Trp Lys Phe Ser Asp His
 325 330 335

Tyr Phe Glu Arg Lys Met Lys Pro Leu Ile Phe Tyr Ala Asn His Asp
 340 345 350

Leu Ser Leu Glu Ala Ala Leu Trp Arg Asp Leu His Lys Ala Gly Ile
 355 360 365

Leu Gln Leu Tyr Gln Arg
 370

<210> 9
 <211> 918
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(918)

<400> 9
 atg gcc tgc atc ctg aag aga aag tct gtg att gct gtg agc ttc ata 48
 Met Ala Cys Ile Leu Lys Arg Lys Ser Val Ile Ala Val Ser Phe Ile
 1 5 10 15
 gca gcg ttc ctt ttc ctg ctg gtt gtg cgt ctt gta aat gaa gtg aat 96
 Ala Ala Phe Leu Phe Leu Leu Val Val Arg Leu Val Asn Glu Val Asn
 20 25 30
 ttc cca ttg cta cta aac tgc ttt gga caa cct ggt aca aag tgg ata 144
 Phe Pro Leu Leu Leu Asn Cys Phe Gly Gln Pro Gly Thr Lys Trp Ile
 35 40 45
 cca ttc tcc tac aca tac agg cgg ccc ctt cga act cac tat gga tac 192
 Pro Phe Ser Tyr Thr Tyr Arg Arg Pro Leu Arg Thr His Tyr Gly Tyr
 50 55 60
 ata aat gtg aag aca caa gag cct ttg caa ctg gac tgt gac ctt tgt 240
 Ile Asn Val Lys Thr Gln Glu Pro Leu Gln Leu Asp Cys Asp Leu Cys
 65 70 75 80
 gcc ata gtg tca aac tca ggt cag atg gtt ggc cag aag gtg gga aat 288

ES 2 368 267 T3

Ala	Ile	Val	Ser	Asn	Ser	Gly	Gln	Met	Val	Gly	Gln	Lys	Val	Gly	Asn		
				85					90					95			
gag	ata	gat	cga	tcc	tcc	tgc	att	tgg	aga	atg	aac	aat	gcc	ccc	acc	336	
Glu	Ile	Asp	Arg	Ser	Ser	Cys	Ile	Trp	Arg	Met	Asn	Asn	Ala	Pro	Thr		
			100					105					110				
aaa	ggt	tat	gaa	gaa	gat	gtc	ggc	cgc	atg	acc	atg	att	cga	gtt	gtg	384	
Lys	Gly	Tyr	Glu	Glu	Asp	Val	Gly	Arg	Met	Thr	Met	Ile	Arg	Val	Val		
		115					120					125					
tcc	cat	acc	agc	gtt	cct	ctt	ttg	cta	aaa	aac	cct	gat	tat	ttt	ttc	432	
Ser	His	Thr	Ser	Val	Pro	Leu	Leu	Leu	Lys	Asn	Pro	Asp	Tyr	Phe	Phe		
	130					135					140						
aag	gaa	gcg	aat	act	act	att	tat	gtt	att	tgg	gga	cct	ttc	cgc	aat	480	
Lys	Glu	Ala	Asn	Thr	Thr	Ile	Tyr	Val	Ile	Trp	Gly	Pro	Phe	Arg	Asn		
145				150						155					160		
atg	agg	aaa	gat	ggc	aat	ggc	atc	att	tac	aac	atg	ttg	aaa	aag	aca	528	
Met	Arg	Lys	Asp	Gly	Asn	Gly	Ile	Ile	Tyr	Asn	Met	Leu	Lys	Lys	Thr		
				165					170					175			
gtt	ggt	atc	tat	ccg	aat	gcc	caa	ata	tac	gtg	acc	aca	gag	aag	cgc	576	
Val	Gly	Ile	Tyr	Pro	Asn	Ala	Gln	Ile	Tyr	Val	Thr	Thr	Glu	Lys	Arg		
			180					185					190				
atg	agt	tac	tgt	gat	gga	gtt	ttt	aag	aag	gaa	act	ggg	aag	gac	aga	624	
Met	Ser	Tyr	Cys	Asp	Gly	Val	Phe	Lys	Lys	Glu	Thr	Gly	Lys	Asp	Arg		
		195					200					205					
gtc	cag	tct	ggc	tca	tat	ctc	agc	aca	ggg	tgg	ttt	acc	ttc	ctt	ctg	672	
Val	Gln	Ser	Gly	Ser	Tyr	Leu	Ser	Thr	Gly	Trp	Phe	Thr	Phe	Leu	Leu		
	210					215					220						
gcc	atg	gac	gcc	tgt	tat	ggc	att	cac	gtc	tac	ggg	atg	ata	aat	gac	720	
Ala	Met	Asp	Ala	Cys	Tyr	Gly	Ile	His	Val	Tyr	Gly	Met	Ile	Asn	Asp		
225				230					235					240			
acc	tac	tgc	aag	aca	gaa	ggg	tat	aga	aaa	gtc	ccc	tac	cat	tat	tat	768	
Thr	Tyr	Cys	Lys	Thr	Glu	Gly	Tyr	Arg	Lys	Val	Pro	Tyr	His	Tyr	Tyr		
				245					250					255			
gaa	caa	gga	aga	gat	gag	tgt	gat	gaa	tat	ttt	ctt	cat	gaa	cat	gcc	816	
Glu	Gln	Gly	Arg	Asp	Glu	Cys	Asp	Glu	Tyr	Phe	Leu	His	Glu	His	Ala		
			260					265					270				
cca	tat	ggg	ggt	cat	agg	ttt	atc	act	gaa	aag	aaa	gtg	ttt	gct	aaa	864	
Pro	Tyr	Gly	Gly	His	Arg	Phe	Ile	Thr	Glu	Lys	Lys	Val	Phe	Ala	Lys		
		275					280					285					
tgg	gcc	aag	aag	cac	agg	ata	ata	ttt	aca	cat	cca	aac	tgg	aca	ttg	912	
Trp	Ala	Lys	Lys	His	Arg	Ile	Ile	Phe	Thr	His	Pro	Asn	Trp	Thr	Leu		
	290					295					300						
tct	tga															918	
Ser																	
305																	

<210> 10

ES 2 368 267 T3

<211> 305
 <212> PRT
 <213> Homo sapiens

<400> 10

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

Ala Ala Phe Leu Phe Leu Leu Val Val Arg Leu Val Asn Glu Val Asn
 20 25 30

Phe Pro Leu Leu Leu Asn Cys Phe Gly Gln Pro Gly Thr Lys Trp Ile
 35 40 45

Pro Phe Ser Tyr Thr Tyr Arg Arg Pro Leu Arg Thr His Tyr Gly Tyr
 50 55 60

Ile Asn Val Lys Thr Gln Glu Pro Leu Gln Leu Asp Cys Asp Leu Cys
 65 70 75 80

Ala Ile Val Ser Asn Ser Gly Gln Met Val Gly Gln Lys Val Gly Asn
 85 90 95

Glu Ile Asp Arg Ser Ser Cys Ile Trp Arg Met Asn Asn Ala Pro Thr
 100 105 110

Lys Gly Tyr Glu Glu Asp Val Gly Arg Met Thr Met Ile Arg Val Val
 115 120 125

Ser His Thr Ser Val Pro Leu Leu Leu Lys Asn Pro Asp Tyr Phe Phe
 130 135 140

Lys Glu Ala Asn Thr Thr Ile Tyr Val Ile Trp Gly Pro Phe Arg Asn
 145 150 155 160

Met Arg Lys Asp Gly Asn Gly Ile Ile Tyr Asn Met Leu Lys Lys Thr
 165 170 175

Val Gly Ile Tyr Pro Asn Ala Gln Ile Tyr Val Thr Thr Glu Lys Arg
 180 185 190

Met Ser Tyr Cys Asp Gly Val Phe Lys Lys Glu Thr Gly Lys Asp Arg
 195 200 205

Val Gln Ser Gly Ser Tyr Leu Ser Thr Gly Trp Phe Thr Phe Leu Leu
 210 215 220

ES 2 368 267 T3

Ala Met Asp Ala Cys Tyr Gly Ile His Val Tyr Gly Met Ile Asn Asp
225 230 235 240

Thr Tyr Cys Lys Thr Glu Gly Tyr Arg Lys Val Pro Tyr His Tyr Tyr
245 250 255

Glu Gln Gly Arg Asp Glu Cys Asp Glu Tyr Phe Leu His Glu His Ala
260 265 270

Pro Tyr Gly Gly His Arg Phe Ile Thr Glu Lys Lys Val Phe Ala Lys
275 280 285

Trp Ala Lys Lys His Arg Ile Ile Phe Thr His Pro Asn Trp Thr Leu
290 295 300

Ser
305

<210> 11
<211> 909
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(909)

<400> 11
atg aag gct ccg ggt cgg ctc gtg ctc atc atc ctg tgc tcc gtg gtc 48
Met Lys Ala Pro Gly Arg Leu Val Leu Ile Ile Leu Cys Ser Val Val
1 5 10 15

ttc tct gcc gtc tac atc ctc ctg tgc tgc tgg gcc ggc ctg ccc ctc 96
Phe Ser Ala Val Tyr Ile Leu Leu Cys Cys Trp Ala Gly Leu Pro Leu
20 25 30

tgc ctg gcc acc tgc ctg gac cac cac ttc ccc aca ggc tcc agg ccc 144
Cys Leu Ala Thr Cys Leu Asp His His Phe Pro Thr Gly Ser Arg Pro
35 40 45

act gtg ccg gga ccc ctg cac ttc agt gga tat agc agt gtg cca gat 192
Thr Val Pro Gly Pro Leu His Phe Ser Gly Tyr Ser Ser Val Pro Asp
50 55 60

ggg aag ccg ctg gtc cgc gag ccc tgc cgc agc tgt gcc gtg gtg tcc 240
Gly Lys Pro Leu Val Arg Glu Pro Cys Arg Ser Cys Ala Val Val Ser
65 70 75 80

agc tcc ggc caa atg ctg ggc tca ggc ctg ggt gct gag atc gac agt 288
Ser Ser Gly Gln Met Leu Gly Ser Gly Leu Gly Ala Glu Ile Asp Ser
85 90 95

gcc gag tgc gtg ttc cgc atg aac cag gcg ccc acc gtg ggc ttt gag 336

ES 2 368 267 T3

Ala Glu Cys Val Phe Arg Met Asn Gln Ala Pro Thr Val Gly Phe Glu	
100 105 110	
gcg gat gtg ggc cag cgc agc acc ctg cgt gtc gtc tca cac aca agc	384
Ala Asp Val Gly Gln Arg Ser Thr Leu Arg Val Val Ser His Thr Ser	
115 120 125	
gtg ccg ctg ctg ctg cgc aac tat tca cac tac ttc cag aag gcc cga	432
Val Pro Leu Leu Leu Arg Asn Tyr Ser His Tyr Phe Gln Lys Ala Arg	
130 135 140	
gac acg ctc tac atg gtg tgg ggc cag ggc agg cac atg gac cgg gtg	480
Asp Thr Leu Tyr Met Val Trp Gly Gln Gly Arg His Met Asp Arg Val	
145 150 155 160	
ctc ggc ggc cgc acc tac cgc acg ctg ctg cag ctc acc agg atg tac	528
Leu Gly Gly Arg Thr Tyr Arg Thr Leu Leu Gln Leu Thr Arg Met Tyr	
165 170 175	
ccc ggc ctg cag gtg tac acc ttc acg gag cgc atg atg gcc tac tgc	576
Pro Gly Leu Gln Val Tyr Thr Phe Thr Glu Arg Met Met Ala Tyr Cys	
180 185 190	
gac cag atc ttc cag gac gag acg ggc aag aac cgg agg cag tgc ggc	624
Asp Gln Ile Phe Gln Asp Glu Thr Gly Lys Asn Arg Arg Gln Ser Gly	
195 200 205	
tcc ttc ctc agc acc ggc tgg ttc acc atg atc ctc gcg ctg gag ctg	672
Ser Phe Leu Ser Thr Gly Trp Phe Thr Met Ile Leu Ala Leu Glu Leu	
210 215 220	
tgt gag gag atc gtg gtc tat ggg atg gtc agc gac agc tac tgc agg	720
Cys Glu Glu Ile Val Val Tyr Gly Met Val Ser Asp Ser Tyr Cys Arg	
225 230 235 240	
gag aag agc cac ccc tca gtg cct tac cac tac ttt gag aag ggc cgg	768
Glu Lys Ser His Pro Ser Val Pro Tyr His Tyr Phe Glu Lys Gly Arg	
245 250 255	
cta gat gag tgt cag atg tac ctg gca cac gag cag gcg ccc cga agc	816
Leu Asp Glu Cys Gln Met Tyr Leu Ala His Glu Gln Ala Pro Arg Ser	
260 265 270	
gcc cac cgc ttc atc act gag aag gcg gtc ttc tcc cgc tgg gcc aag	864
Ala His Arg Phe Ile Thr Glu Lys Ala Val Phe Ser Arg Trp Ala Lys	
275 280 285	
aag agg ccc atc gtg ttc gcc cat ccg tcc tgg agg act gag tag	909
Lys Arg Pro Ile Val Phe Ala His Pro Ser Trp Arg Thr Glu	
290 295 300	
<210> 12	
<211> 302	
<212> PRT	
<213> Homo sapiens	
<400> 12	
Met Lys Ala Pro Gly Arg Leu Val Leu Ile Ile Leu Cys Ser Val Val	
1 5 10 15	

ES 2 368 267 T3

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

ES 2 368 267 T3

Leu Asp Glu Cys Gln Met Tyr Leu Ala His Glu Gln Ala Pro Arg Ser
260 265 270

Ala His Arg Phe Ile Thr Glu Lys Ala Val Phe Ser Arg Trp Ala Lys
275 280 285

Lys Arg Pro Ile Val Phe Ala His Pro Ser Trp Arg Thr Glu
290 295 300

```
<210> 13
<211> 1011
<212> ADN
<213> Homo sapiens
```

```
<220>
<221> CDS
<222> (1)..(1011)
```

<400>	13															
atg	aag	acc	ctg	atg	cgc	cat	ggt	ctg	gca	gtg	tgt	tta	gcg	ctc	acc	48
Met	Lys	Thr	Leu	Met	Arg	His	Gly	Leu	Ala	Val	Cys	Leu	Ala	Leu	Thr	
1				5					10					15		
acc	atg	tgc	acc	agc	ttg	ttg	cta	gtg	tac	agc	agc	ctc	ggc	ggc	cag	96
Thr	Met	Cys	Thr	Ser	Leu	Leu	Leu	Val	Tyr	Ser	Ser	Leu	Gly	Gly	Gln	
			20					25					30			
aag	gag	cgg	ccc	ccg	cag	cag	cag	cag	cag	cag	cag	caa	cag	cag	cag	144
Lys	Glu	Arg	Pro	Pro	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	
		35					40					45				
cag	gcg	tcg	gcc	acc	ggc	agc	tcg	cag	ccg	gcg	gcg	gag	agc	agc	acc	192
Gln	Ala	Ser	Ala	Thr	Gly	Ser	Ser	Gln	Pro	Ala	Ala	Glu	Ser	Ser	Thr	
	50					55					60					
cag	cag	cgc	ccc	ggg	gtc	ccc	gcg	gga	ccg	cgg	cca	ctg	gac	gga	tac	240
Gln	Gln	Arg	Pro	Gly	Val	Pro	Ala	Gly	Pro	Arg	Pro	Leu	Asp	Gly	Tyr	
65				70					75					80		
ctc	gga	gtg	gcg	gac	cac	aag	ccc	ctg	aaa	atg	cac	tgc	agg	gac	tgt	288
Leu	Gly	Val	Ala	Asp	His	Lys	Pro	Leu	Lys	Met	His	Cys	Arg	Asp	Cys	
				85					90					95		
gcc	ctg	gtg	acc	agc	tca	ggg	cat	ctg	ctg	cac	agt	cgg	caa	ggc	tcc	336
Ala	Leu	Val	Thr	Ser	Ser	Gly	His	Leu	Leu	His	Ser	Arg	Gln	Gly	Ser	
			100					105					110			
cag	att	gac	cag	aca	gag	tgt	gtc	atc	cgc	atg	aat	gac	gcc	ccc	aca	384
Gln	Ile	Asp	Gln	Thr	Glu	Cys	Val	Ile	Arg	Met	Asn	Asp	Ala	Pro	Thr	
		115					120					125				
cgc	ggc	tat	ggg	cgt	gac	gtg	ggc	aat	cgc	acc	agc	ctg	agg	gtc	atc	432
Arg	Gly	Tyr	Gly	Arg	Asp	Val	Gly	Asn	Arg	Thr	Ser	Leu	Arg	Val	Ile	
	130					135					140					
gcg	cat	tcc	agc	atc	cag	agg	atc	ctc	cgc	aac	cgc	cat	gac	ctg	ctc	480

ES 2 368 267 T3

Ala	His	Ser	Ser	Ile	Gln	Arg	Ile	Leu	Arg	Asn	Arg	His	Asp	Leu	Leu	
145					150					155					160	
aac	gtg	agc	cag	ggc	acc	gtg	ttc	atc	ttc	tgg	ggc	ccc	agc	agc	tac	528
Asn	Val	Ser	Gln	Gly	Thr	Val	Phe	Ile	Phe	Trp	Gly	Pro	Ser	Ser	Tyr	
				165					170					175		
atg	cgg	cgg	gac	ggc	aag	ggc	cag	gtc	tac	aac	aac	ctg	cat	ctc	ctg	576
Met	Arg	Arg	Asp	Gly	Lys	Gly	Gln	Val	Tyr	Asn	Asn	Leu	His	Leu	Leu	
			180					185					190			
agc	cag	gtg	ctg	ccc	cgg	ctg	aag	gcc	ttc	atg	att	act	cgc	cac	aag	624
Ser	Gln	Val	Leu	Pro	Arg	Leu	Lys	Ala	Phe	Met	Ile	Thr	Arg	His	Lys	
		195					200					205				
atg	ctg	cag	ttt	gat	gag	ctc	ttc	aag	cag	gag	act	ggc	aaa	gac	agg	672
Met	Leu	Gln	Phe	Asp	Glu	Leu	Phe	Lys	Gln	Glu	Thr	Gly	Lys	Asp	Arg	
	210					215					220					
aag	ata	tcc	aac	act	tgg	ctc	agc	act	ggc	tgg	ttt	aca	atg	aca	att	720
Lys	Ile	Ser	Asn	Thr	Trp	Leu	Ser	Thr	Gly	Trp	Phe	Thr	Met	Thr	Ile	
225					230					235					240	
gca	ctg	gag	ctc	tgt	gac	agg	atc	aat	gtt	tat	ggc	atg	gtg	ccc	cca	768
Ala	Leu	Glu	Leu	Cys	Asp	Arg	Ile	Asn	Val	Tyr	Gly	Met	Val	Pro	Pro	
				245					250					255		
gac	ttc	tgc	agg	gat	ccc	aat	cac	cct	tca	gta	cct	tat	cat	tat	tat	816
Asp	Phe	Cys	Arg	Asp	Pro	Asn	His	Pro	Ser	Val	Pro	Tyr	His	Tyr	Tyr	
			260					265					270			
gaa	cct	ttt	gga	cct	gat	gaa	tgt	aca	atg	tac	ctc	tcc	cat	gag	cga	864
Glu	Pro	Phe	Gly	Pro	Asp	Glu	Cys	Thr	Met	Tyr	Leu	Ser	His	Glu	Arg	
		275					280					285				
gga	cgc	aag	ggc	agt	cat	cac	cgc	ttt	atc	aca	gag	aaa	cga	gtc	ttt	912
Gly	Arg	Lys	Gly	Ser	His	His	Arg	Phe	Ile	Thr	Glu	Lys	Arg	Val	Phe	
	290					295					300					
aag	aac	tgg	gca	cgg	aca	ttc	aat	att	cac	ttt	ttt	caa	cca	gac	tgg	960
Lys	Asn	Trp	Ala	Arg	Thr	Phe	Asn	Ile	His	Phe	Phe	Gln	Pro	Asp	Trp	
305					310					315					320	
aaa	cca	gaa	tca	ctt	gct	ata	aat	cat	cct	gag	aat	aaa	cct	gtg	ttc	1008
Lys	Pro	Glu	Ser	Leu	Ala	Ile	Asn	His	Pro	Glu	Asn	Lys	Pro	Val	Phe	
				325					330					335		
taa																1011
<210>	14															
<211>	336															
<212>	PRT															
<213>	Homo sapiens															
<400>	14															
Met	Lys	Thr	Leu	Met	Arg	His	Gly	Leu	Ala	Val	Cys	Leu	Ala	Leu	Thr	
1				5					10					15		

ES 2 368 267 T3

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

ES 2 368 267 T3

260	265	270	
Glu Pro Phe Gly Pro Asp Glu Cys Thr Met Tyr Leu Ser His Glu Arg			
275	280	285	
Gly Arg Lys Gly Ser His His Arg Phe Ile Thr Glu Lys Arg Val Phe			
290	295	300	
Lys Asn Trp Ala Arg Thr Phe Asn Ile His Phe Phe Gln Pro Asp Trp			
305	310	315	320
Lys Pro Glu Ser Leu Ala Ile Asn His Pro Glu Asn Lys Pro Val Phe			
	325	330	335
<210> 15			
<211> 1002			
<212> ADN			
<213> Homo sapiens			
<220>			
<221> CDS			
<222> (1)..(1002)			
<400> 15			
atg gct tgc tcg agg ccc ccc agc cag tgt gaa ccc aca tcc ctg ccc			48
Met Ala Cys Ser Arg Pro Pro Ser Gln Cys Glu Pro Thr Ser Leu Pro			
1	5	10	15
cca ggg cca cct gca gga cgc cga cac cta ccc ctc agc aga cgc cgg			96
Pro Gly Pro Pro Ala Gly Arg Arg His Leu Pro Leu Ser Arg Arg Arg			
	20	25	30
aga gaa atg agt agc aac aaa gag cag cgg tca gca gtg ttc gtg atc			144
Arg Glu Met Ser Ser Asn Lys Glu Gln Arg Ser Ala Val Phe Val Ile			
	35	40	45
ctc ttt gcc ctc atc acc atc ctc atc ctc tac agc tcc aac agt gcc			192
Leu Phe Ala Leu Ile Thr Ile Leu Ile Leu Tyr Ser Ser Asn Ser Ala			
	50	55	60
aat gag gtc ttc cat tac ggc tcc ctg cgg ggc cgt agc cgc cga cct			240
Asn Glu Val Phe His Tyr Gly Ser Leu Arg Gly Arg Ser Arg Arg Pro			
65	70	75	80
gtc aac ctc aag aag tgg agc atc act gac ggc tat gtc ccc att ctc			288
Val Asn Leu Lys Lys Trp Ser Ile Thr Asp Gly Tyr Val Pro Ile Leu			
	85	90	95
ggc aac aag aca ctg ccc tct cgg tgc cac cag tgt gtg att gtc agc			336
Gly Asn Lys Thr Leu Pro Ser Arg Cys His Gln Cys Val Ile Val Ser			
	100	105	110
agc tcc agc cac ctg ctg ggc acc aag ctg ggc cct gag atc gag cgg			384
Ser Ser Ser His Leu Leu Gly Thr Lys Leu Gly Pro Glu Ile Glu Arg			
	115	120	125

ES 2 368 267 T3

gct gag tgt aca atc cgc atg aat gat gca ccc acc act ggc tac tca	432
Ala Glu Cys Thr Ile Arg Met Asn Asp Ala Pro Thr Thr Gly Tyr Ser	
130 135 140	
gct gat gtg ggc aac aag acc acc tac cgc gtc gtg gcc cat tcc agt	480
Ala Asp Val Gly Asn Lys Thr Thr Tyr Arg Val Val Ala His Ser Ser	
145 150 155 160	
gtg ttc cgc gtg ctg agg agg ccc cag gag ttt gtc aac cgg acc cct	528
Val Phe Arg Val Leu Arg Arg Pro Gln Glu Phe Val Asn Arg Thr Pro	
165 170 175	
gaa acc gtg ttc atc ttc tgg ggg ccc ccg agc aag atg cag aag ccc	576
Glu Thr Val Phe Ile Phe Trp Gly Pro Pro Ser Lys Met Gln Lys Pro	
180 185 190	
cag ggc agc ctc gtg cgt gtg atc cag cga gcg ggc ctg gtg ttc ccc	624
Gln Gly Ser Leu Val Arg Val Ile Gln Arg Ala Gly Leu Val Phe Pro	
195 200 205	
aac atg gaa gca tat gcc gtc tct ccc ggc cgc atg cgg caa ttt gac	672
Asn Met Glu Ala Tyr Ala Val Ser Pro Gly Arg Met Arg Gln Phe Asp	
210 215 220	
gac ctc ttc cgg ggt gag acg ggc aag gac agg gag aag tct cat tcg	720
Asp Leu Phe Arg Gly Glu Thr Gly Lys Asp Arg Glu Lys Ser His Ser	
225 230 235 240	
tgg ttg agc aca ggc tgg ttt acc atg gtg atc gcg gtg gag ttg tgt	768
Trp Leu Ser Thr Gly Trp Phe Thr Met Val Ile Ala Val Glu Leu Cys	
245 250 255	
gac cac gtg cat gtc tat ggc atg gtc ccc ccc aac tac tgc agc cag	816
Asp His Val His Val Tyr Gly Met Val Pro Pro Asn Tyr Cys Ser Gln	
260 265 270	
cgg ccc cgc ctc cag cgc atg ccc tac cac tac tac gag ccc aag ggg	864
Arg Pro Arg Leu Gln Arg Met Pro Tyr His Tyr Tyr Glu Pro Lys Gly	
275 280 285	
ccg gac gaa tgt gtc acc tac atc cag aat gag cac agt cgc aag ggc	912
Pro Asp Glu Cys Val Thr Tyr Ile Gln Asn Glu His Ser Arg Lys Gly	
290 295 300	
aac cac cac cgc ttc atc acc gag aaa agg gtc ttc tca tcg tgg gcc	960
Asn His His Arg Phe Ile Thr Glu Lys Arg Val Phe Ser Ser Trp Ala	
305 310 315 320	
cag ctg tat ggc atc acc ttc tcc cac ccc tcc tgg acc tag	1002
Gln Leu Tyr Gly Ile Thr Phe Ser His Pro Ser Trp Thr	
325 330	

<210> 16
 <211> 333
 <212> PRT
 <213> Homo sapiens

<400> 16

ES 2 368 267 T3

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

ES 2 368 267 T3

	245		250		255	
Asp His Val His Val Tyr Gly Met Val Pro Pro Asn Tyr Cys Ser Gln	260	265	270			
Arg Pro Arg Leu Gln Arg Met Pro Tyr His Tyr Tyr Glu Pro Lys Gly	275	280	285			
Pro Asp Glu Cys Val Thr Tyr Ile Gln Asn Glu His Ser Arg Lys Gly	290	295	300			
Asn His His Arg Phe Ile Thr Glu Lys Arg Val Phe Ser Ser Trp Ala	305	310	315	320		
Gln Leu Tyr Gly Ile Thr Phe Ser His Pro Ser Trp Thr	325	330				
<210> 17						
<211> 1023						
<212> ADN						
<213> Homo sapiens						
<220>						
<221> CDS						
<222> (1)..(1023)						
<400> 17						
atg gtg acc ctg cgg aag agg acc ctg aaa gtg ctc acc ttc ctc gtg						48
Met Val Thr Leu Arg Lys Arg Thr Leu Lys Val Leu Thr Phe Leu Val	1	5	10	15		
ctc ttc atc ttc ctc acc tcc ttc ttc ctg aac tac tcc cac acc atg						96
Leu Phe Ile Phe Leu Thr Ser Phe Phe Leu Asn Tyr Ser His Thr Met	20	25	30			
gtg gcc acc acc tgg ttc ccc aag cag atg gtc ctg gag ctc tcc gag						144
Val Ala Thr Thr Trp Phe Pro Lys Gln Met Val Leu Glu Leu Ser Glu	35	40	45			
aac ctg aag aga ctg atc aag cac agg cct tgc acc tgc acc cac tgc						192
Asn Leu Lys Arg Leu Ile Lys His Arg Pro Cys Thr Cys Thr His Cys	50	55	60			
atc ggg cag cgc aag ctc tcg gcc tgg ttc gat gag agg ttc aac cag						240
Ile Gly Gln Arg Lys Leu Ser Ala Trp Phe Asp Glu Arg Phe Asn Gln	65	70	75	80		
acc atg cag ccg ctg ctg acc gcc cag aac gcg ctc ttg gag gac gac						288
Thr Met Gln Pro Leu Leu Thr Ala Gln Asn Ala Leu Leu Glu Asp Asp	85	90	95			
acc tac cga tgg tgg ctg agg ctc cag cgg gag aag aag ccc aat aac						336
Thr Tyr Arg Trp Trp Leu Arg Leu Gln Arg Glu Lys Lys Pro Asn Asn	100	105	110			

ES 2 368 267 T3

ttg aat gac acc atc aag gag ctg ttc aga gtg gtg cct ggg aat gtg Leu Asn Asp Thr Ile Lys Glu Leu Phe Arg Val Val Pro Gly Asn Val 115 120 125	384
gac cct atg ctg gag aag agg tcg gtg ggc tgc cgg cgc tgc gcc gtt Asp Pro Met Leu Glu Lys Arg Ser Val Gly Cys Arg Arg Cys Ala Val 130 135 140	432
gtg ggc aac tcg ggc aac ctg agg gag tct tct tat ggg cct gag ata Val Gly Asn Ser Gly Asn Leu Arg Glu Ser Ser Tyr Gly Pro Glu Ile 145 150 155 160	480
gac agt cac gac ttt gtc ctc agg atg aac aag gcg ccc acg gca ggg Asp Ser His Asp Phe Val Leu Arg Met Asn Lys Ala Pro Thr Ala Gly 165 170 175	528
ttt gaa gct gat gtt ggg acc aag acc acc cac cat ctg gtg tac cct Phe Glu Ala Asp Val Gly Thr Lys Thr Thr His His Leu Val Tyr Pro 180 185 190	576
gag agc ttc cgg gag ctg gga gat aat gtc agc atg atc ctg gtg ccc Glu Ser Phe Arg Glu Leu Gly Asp Asn Val Ser Met Ile Leu Val Pro 195 200 205	624
ttc aag acc atc gac ttg gag tgg gtg gtg agc gcc atc acc acg ggc Phe Lys Thr Ile Asp Leu Glu Trp Val Val Ser Ala Ile Thr Thr Gly 210 215 220	672
acc att tcc cac acc tac atc ccg gtt cct gca aag atc aga gtg aaa Thr Ile Ser His Thr Tyr Ile Pro Val Pro Ala Lys Ile Arg Val Lys 225 230 235 240	720
cag gat aag atc ctg atc tac cac cca gcc ttc atc aag tat gtc ttt Gln Asp Lys Ile Leu Ile Tyr His Pro Ala Phe Ile Lys Tyr Val Phe 245 250 255	768
gac aac tgg ctg caa ggg cac ggg cga tac cca tct acc ggc atc ctc Asp Asn Trp Leu Gln Gly His Gly Arg Tyr Pro Ser Thr Gly Ile Leu 260 265 270	816
tcg gtc atc ttc tca atg cat gtc tgc gat gag gtg gac ttg tac ggc Ser Val Ile Phe Ser Met His Val Cys Asp Glu Val Asp Leu Tyr Gly 275 280 285	864
ttc ggg gca gac agc aaa ggg aac tgg cac cac tac tgg gag aac aac Phe Gly Ala Asp Ser Lys Gly Asn Trp His His Tyr Trp Glu Asn Asn 290 295 300	912
cca tcc gcg ggg gct ttt cgc aag acg ggg gtg cac gat gca gac ttt Pro Ser Ala Gly Ala Phe Arg Lys Thr Gly Val His Asp Ala Asp Phe 305 310 315 320	960
gag tct aac gtg acg gcc acc ttg gcc tcc atc aat aaa atc cgg atc Glu Ser Asn Val Thr Ala Thr Leu Ala Ser Ile Asn Lys Ile Arg Ile 325 330 335	1008
ttc aag ggg aga tga Phe Lys Gly Arg 340	1023

ES 2 368 267 T3

<210> 18
 <211> 340
 <212> PRT
 <213> Homo sapiens

<400> 18

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 30

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

Phe Lys Thr Ile Asp Leu Glu Trp Val Val Ser Ala Ile Thr Thr Gly

ES 2 368 267 T3

<210> 18
 <211> 340
 <212> PRT
 <213> Homo sapiens

<400> 18

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 30

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

Phe Lys Thr Ile Asp Leu Glu Trp Val Val Ser Ala Ile Thr Thr Gly

ES 2 368 267 T3

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> 19
 <211> 1053
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1053)

<400> 19	
atg aag tgc tcc ctg cgg gtg tgg ttc ctc tcc gtg gcc ttc ctg ctg	48
Met Lys Cys Ser Leu Arg Val Trp Phe Leu Ser Val Ala Phe Leu Leu	
1 5 10 15	
gtg ttc atc atg tcc ctg ctc ttc acc tac tcg cac cac agc atg gcc	96
Val Phe Ile Met Ser Leu Leu Phe Thr Tyr Ser His His Ser Met Ala	
20 25 30	
acg ctc ccc tac ctg gac tca ggg gcc ctg gat ggg acg cac cgg gtg	144
Thr Leu Pro Tyr Leu Asp Ser Gly Ala Leu Asp Gly Thr His Arg Val	
35 40 45	
aag ctg gtg ccc ggc tat gcc ggc ctg cag cgc ctc agc aag gag agg	192
Lys Leu Val Pro Gly Tyr Ala Gly Leu Gln Arg Leu Ser Lys Glu Arg	
50 55 60	

ES 2 368 267 T3

ctc	tcg	ggc	aag	agc	tgt	gcc	tgt	cgc	cgc	tgc	atg	ggc	gat	gcc	ggt	240
Leu	Ser	Gly	Lys	Ser	Cys	Ala	Cys	Arg	Arg	Cys	Met	Gly	Asp	Ala	Gly	
65					70					75					80	
gcc	tcc	gac	tgg	ttt	gac	agc	cac	ttt	gac	ggt	aac	att	tcc	ccc	gtc	288
Ala	Ser	Asp	Trp	Phe	Asp	Ser	His	Phe	Asp	Gly	Asn	Ile	Ser	Pro	Val	
				85					90					95		
tgg	acc	cga	gag	aac	atg	gat	ctt	cca	ccg	gac	gtc	cag	agg	tgg	tgg	336
Trp	Thr	Arg	Glu	Asn	Met	Asp	Leu	Pro	Pro	Asp	Val	Gln	Arg	Trp	Trp	
			100					105					110			
atg	atg	ctg	cag	ccc	cag	ttc	aag	tca	cac	aac	acc	aat	gag	gtg	ctg	384
Met	Met	Leu	Gln	Pro	Gln	Phe	Lys	Ser	His	Asn	Thr	Asn	Glu	Val	Leu	
		115					120					125				
gag	aag	ctg	ttc	cag	ata	gtg	cct	ggc	gag	aac	ccc	tac	cgc	ttc	cgg	432
Glu	Lys	Leu	Phe	Gln	Ile	Val	Pro	Gly	Glu	Asn	Pro	Tyr	Arg	Phe	Arg	
	130					135					140					
gac	ccc	cac	cag	tgc	cgg	cgc	tgt	gcc	gtg	gtg	ggg	aac	tcg	ggc	aac	480
Asp	Pro	His	Gln	Cys	Arg	Arg	Cys	Ala	Val	Val	Gly	Asn	Ser	Gly	Asn	
145					150					155					160	
ctg	cgg	ggc	tct	ggc	tat	ggg	cag	gac	gtg	gac	ggg	cac	aac	ttc	atc	528
Leu	Arg	Gly	Ser	Gly	Tyr	Gly	Gln	Asp	Val	Asp	Gly	His	Asn	Phe	Ile	
				165					170					175		
atg	agg	atg	aat	cag	gcg	cca	acc	gtg	ggc	ttt	gag	cag	gat	gtt	ggc	576
Met	Arg	Met	Asn	Gln	Ala	Pro	Thr	Val	Gly	Phe	Glu	Gln	Asp	Val	Gly	
			180					185					190			
agc	cga	acc	acc	cac	cat	ttc	atg	tac	cct	gag	agt	gcc	aag	aac	ctg	624
Ser	Arg	Thr	Thr	His	His	Phe	Met	Tyr	Pro	Glu	Ser	Ala	Lys	Asn	Leu	
			195				200					205				
ccc	gcc	aac	gtc	agc	ttc	gtg	ctg	gtg	ccc	ttc	aag	gtc	ctg	gac	ctt	672
Pro	Ala	Asn	Val	Ser	Phe	Val	Leu	Val	Pro	Phe	Lys	Val	Leu	Asp	Leu	
	210					215					220					
ctg	tgg	atc	gcc	agc	gcc	ttg	tcc	acg	ggg	cag	atc	cga	ttc	acc	tac	720
Leu	Trp	Ile	Ala	Ser	Ala	Leu	Ser	Thr	Gly	Gln	Ile	Arg	Phe	Thr	Tyr	
225					230					235				240		
gcc	cca	gtg	aag	tcc	ttc	ctt	cga	gtg	gat	aaa	gaa	aag	gtc	cag	atc	768
Ala	Pro	Val	Lys	Ser	Phe	Leu	Arg	Val	Asp	Lys	Glu	Lys	Val	Gln	Ile	
				245					250					255		
tac	aac	cca	gcc	ttc	ttc	aag	tat	atc	cac	gac	agg	tgg	aca	gag	cat	816
Tyr	Asn	Pro	Ala	Phe	Phe	Lys	Tyr	Ile	His	Asp	Arg	Trp	Thr	Glu	His	
			260					265					270			
cac	ggg	cgg	tac	cct	tcc	acg	ggg	atg	ctg	gtg	ctt	ttc	ttt	gcc	ctg	864
His	Gly	Arg	Tyr	Pro	Ser	Thr	Gly	Met	Leu	Val	Leu	Phe	Phe	Ala	Leu	
			275				280					285				
cat	gtg	tgt	gat	gag	gtg	aac	gtg	tac	ggg	ttc	ggg	gcc	gac	agc	cgg	912
His	Val	Cys	Asp	Glu	Val	Asn	Val	Tyr	Gly	Phe	Gly	Ala	Asp	Ser	Arg	
	290					295					300					

ES 2 368 267 T3

ggc aac tgg cac cac tac tgg gag aac aac cgg tac gcg ggc gag ttc 960
 Gly Asn Trp His His Tyr Trp Glu Asn Asn Arg Tyr Ala Gly Glu Phe
 305 310 315 320

cgg aag act ggc gtg cac gac gcg gac ttc gag gcc cac atc atc gac 1008
 Arg Lys Thr Gly Val His Asp Ala Asp Phe Glu Ala His Ile Ile Asp
 325 330 335

atg ctg gcc aag gcc agc aag atc gaa gtc tac cgg ggc aac tga 1053
 Met Leu Ala Lys Ala Ser Lys Ile Glu Val Tyr Arg Gly Asn
 340 345 350

<210> 20
 <211> 350
 <212> PRT
 <213> Homo sapiens

<400> 20

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

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
 115 120 125

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

ES 2 368 267 T3

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> 21
<211> 1335
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1335)

<400> 21
atg gga ctc ttg gta ttt gtg cgc aat ctg ctg cta gcc ctc tgc ctc
Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
1 5 10 15

48

ES 2 368 267 T3

ttt ctg gta ctg gga ttt ttg tat tat tct gcg tgg aag cta cac tta	96
Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu	
20 25 30	
ctc cag tgg gag gag gac tcc agt aag tat agt cac tct agc tca ccc	144
Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His Ser Ser Ser Pro	
35 40 45	
cag gag aag cct gtt gca gat tca gtg gtt ctt tcc ttt gac tcc gct	192
Gln Glu Lys Pro Val Ala Asp Ser Val Val Leu Ser Phe Asp Ser Ala	
50 55 60	
gga caa aca cta ggc tca gag tat gat cgg ttg ggc ttc ctc ctg aat	240
Gly Gln Thr Leu Gly Ser Glu Tyr Asp Arg Leu Gly Phe Leu Leu Asn	
65 70 75 80	
ctg gac tct aaa ctc tca ccg agg act ctc tgc acg gtg gtt ttt ggc	288
Leu Asp Ser Lys Leu Ser Pro Arg Thr Leu Cys Thr Val Val Phe Gly	
85 90 95	
ctt gac tgc ata ttg gaa tca cct gga gag cct aaa aaa tta ctg atg	336
Leu Asp Cys Ile Leu Glu Ser Pro Gly Glu Pro Lys Lys Leu Leu Met	
100 105 110	
cct gca tcc cac cct cta gag att ttg aag tca ctg agc gag gac gca	384
Pro Ala Ser His Pro Leu Glu Ile Leu Lys Ser Leu Ser Glu Asp Ala	
115 120 125	
gcc ttt gca tta gga ttt tta aag ctg ccc agg cct gct gaa tta gcc	432
Ala Phe Ala Leu Gly Phe Leu Lys Leu Pro Arg Pro Ala Glu Leu Ala	
130 135 140	
acc aag tac gca aac ttt tca gag gga gct tgc aag cct ggc tat gct	480
Thr Lys Tyr Ala Asn Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala	
145 150 155 160	
tca gcc ttg atg acg gcc atc ttc ccc cgg ttc tcc aag cca gca ccc	528
Ser Ala Leu Met Thr Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro	
165 170 175	
atg ttc ctg gat gac tcc ttt cgc aag tgg gct aga atc cgg gag ttc	576
Met Phe Leu Asp Asp Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe	
180 185 190	
gtg ccg cct ttt ggg atc aaa ggt caa gac aat ctg atc aaa gcc atc	624
Val Pro Pro Phe Gly Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala Ile	
195 200 205	
ttg tca gtc acc aaa gag tac cgc ctg acc cct gcc ttg gac agc ctc	672
Leu Ser Val Thr Lys Glu Tyr Arg Leu Thr Pro Ala Leu Asp Ser Leu	
210 215 220	
cgc tgc cgc cgc tgc atc atc gtg ggc aat gga ggc gtt ctt gcc aac	720
Arg Cys Arg Arg Cys Ile Ile Val Gly Asn Gly Gly Val Leu Ala Asn	
225 230 235 240	
aag tct ctg ggg tca cga att gac gac tat gac att gtg gtg aga ctg	768
Lys Ser Leu Gly Ser Arg Ile Asp Asp Tyr Asp Ile Val Val Arg Leu	
245 250 255	

ES 2 368 267 T3

aat tca gca cca gtg aaa ggc ttt gag aag gac gtg ggc agc aaa acg	816
Asn Ser Ala Pro Val Lys Gly Phe Glu Lys Asp Val Gly Ser Lys Thr	
260 265 270	
aca ctg cgc atc acc tac ccc gag ggc gcc atg cag cgg cct gag cag	864
Thr Leu Arg Ile Thr Tyr Pro Glu Gly Ala Met Gln Arg Pro Glu Gln	
275 280 285	
tac gag cgc gat tct ctc ttt gtc ctc gcc ggc ttc aag tgg cag gac	912
Tyr Glu Arg Asp Ser Leu Phe Val Leu Ala Gly Phe Lys Trp Gln Asp	
290 295 300	
ttt aag tgg ttg aaa tac atc gtc tac aag gag aga gtg agt gca tcg	960
Phe Lys Trp Leu Lys Tyr Ile Val Tyr Lys Glu Arg Val Ser Ala Ser	
305 310 315 320	
gat ggc ttc tgg aaa tct gtg gcc act cga gtg ccc aag gag ccc cct	1008
Asp Gly Phe Trp Lys Ser Val Ala Thr Arg Val Pro Lys Glu Pro Pro	
325 330 335	
gag att cga atc ctc aac cca tat ttc atc cag gag gcc gcc ttc acc	1056
Glu Ile Arg Ile Leu Asn Pro Tyr Phe Ile Gln Glu Ala Ala Phe Thr	
340 345 350	
ctc att ggc ctg ccc ttc aac aat ggc ctc atg ggc cgg ggg aac atc	1104
Leu Ile Gly Leu Pro Phe Asn Asn Gly Leu Met Gly Arg Gly Asn Ile	
355 360 365	
cct acc ctt ggc agt gtg gca gtg acc atg gca cta cac ggc tgt gac	1152
Pro Thr Leu Gly Ser Val Ala Val Thr Met Ala Leu His Gly Cys Asp	
370 375 380	
gag gtg gca gtc gca gga ttt ggc tat gac atg agc aca ccc aac gca	1200
Glu Val Ala Val Ala Gly Phe Gly Tyr Asp Met Ser Thr Pro Asn Ala	
385 390 395 400	
ccc ctg cac tac tat gag acc gtt cgc atg gca gcc atc aaa gag tcc	1248
Pro Leu His Tyr Tyr Glu Thr Val Arg Met Ala Ala Ile Lys Glu Ser	
405 410 415	
tgg acg cac aat atc cag cga gag aaa gag ttt ctg cgg aag ctg gtg	1296
Trp Thr His Asn Ile Gln Arg Glu Lys Glu Phe Leu Arg Lys Leu Val	
420 425 430	
aaa gct cgc gtc atc act gat cta agc agt ggc atc tga	1335
Lys Ala Arg Val Ile Thr Asp Leu Ser Ser Gly Ile	
435 440	
<210> 22	
<211> 444	
<212> PRT	
<213> Homo sapiens	
<400> 22	
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	

ES 2 368 267 T3

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	Ala
		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	
Asn	Ser	Ala	Pro	Val	Lys	Gly	Phe	Glu	Lys	Asp	Val	Gly	Ser	Lys	Thr
			260					265					270		

ES 2 368 267 T3

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
 305 310 315 320

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> 23
 <211> 990
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(990)

<400> 23
 atg gtc agc aag tcc cgc tgg aag ctc ctg gcc atg ttg gct ctg gtc
 Met Val Ser Lys Ser Arg Trp Lys Leu Leu Ala Met Leu Ala Leu Val
 1 5 10 15

48

ES 2 368 267 T3

ctg gtc gtc atg gtg tgg tat tcc atc tcc cgg gaa gac agt ttt tat	96
Leu Val Val Met Val Trp Tyr Ser Ile Ser Arg Glu Asp Ser Phe Tyr	
20 25 30	
ttt ccc atc cca gag aag aag gag cgg tgc ctc cag ggt gag gca gag	144
Phe Pro Ile Pro Glu Lys Lys Glu Pro Cys Leu Gln Gly Glu Ala Glu	
35 40 45	
agc aag gcc tct aag ctc ttt ggc aac tac tcc cgg gat cag ccc atc	192
Ser Lys Ala Ser Lys Leu Phe Gly Asn Tyr Ser Arg Asp Gln Pro Ile	
50 55 60	
ttc ctg cgg ctt gag gat tat ttc tgg gtc aag acg cca tct gct tac	240
Phe Leu Arg Leu Glu Asp Tyr Phe Trp Val Lys Thr Pro Ser Ala Tyr	
65 70 75 80	
gag ctg ccc tat ggg acc aag ggg agt gag gat ctg ctc ctc cgg gtg	288
Glu Leu Pro Tyr Gly Thr Lys Gly Ser Glu Asp Leu Leu Leu Arg Val	
85 90 95	
cta gcc atc acc agc tcc tcc atc ccc aag aac atc cag agc ctc agg	336
Leu Ala Ile Thr Ser Ser Ser Ile Pro Lys Asn Ile Gln Ser Leu Arg	
100 105 110	
tgc cgc cgc tgt gtg gtc gtg ggg aac ggg cac cgg ctg cgg aac agc	384
Cys Arg Arg Cys Val Val Val Gly Asn Gly His Arg Leu Arg Asn Ser	
115 120 125	
tca ctg gga gat gcc atc aac aag tac gat gtg gtc atc aga ttg aac	432
Ser Leu Gly Asp Ala Ile Asn Lys Tyr Asp Val Val Ile Arg Leu Asn	
130 135 140	
aat gcc cca gtg gct ggc tat gag ggt gac gtg ggc tcc aag acc acc	480
Asn Ala Pro Val Ala Gly Tyr Glu Gly Asp Val Gly Ser Lys Thr Thr	
145 150 155 160	
atg cgt ctc ttc tac cct gaa tct gcc cac ttc gac ccc aaa gta gaa	528
Met Arg Leu Phe Tyr Pro Glu Ser Ala His Phe Asp Pro Lys Val Glu	
165 170 175	
aac aac cca gac aca ctc ctc gtc ctg gta gct ttc aag gca atg gac	576
Asn Asn Pro Asp Thr Leu Leu Val Leu Val Ala Phe Lys Ala Met Asp	
180 185 190	
ttc cac tgg att gag acc atc ctg agt gat aag aag cgg gtg cga aag	624
Phe His Trp Ile Glu Thr Ile Leu Ser Asp Lys Lys Arg Val Arg Lys	
195 200 205	
ggt ttc tgg aaa cag cct ccc ctc atc tgg gat gtc aat cct aaa cag	672
Gly Phe Trp Lys Gln Pro Pro Leu Ile Trp Asp Val Asn Pro Lys Gln	
210 215 220	
att cgg att ctc aac ccc ttc ttc atg gag att gca gct gac aaa ctg	720
Ile Arg Ile Leu Asn Pro Phe Phe Met Glu Ile Ala Ala Asp Lys Leu	
225 230 235 240	
ctg agc ctg cca atg caa cag cca cgg aag att aag cag aag ccc acc	768
Leu Ser Leu Pro Met Gln Gln Pro Arg Lys Ile Lys Gln Lys Pro Thr	
245 250 255	
acg ggc ctg ttg gcc atc acg ctg gcc ctc cac ctc tgt gac ttg gtg	816

ES 2 368 267 T3

Thr	Gly	Leu	Leu	Ala	Ile	Thr	Leu	Ala	Leu	His	Leu	Cys	Asp	Leu	Val	
			260					265					270			
cac	att	gcc	ggc	ttt	ggc	tac	cca	gac	gcc	tac	aac	aag	aag	cag	acc	864
His	Ile	Ala	Gly	Phe	Gly	Tyr	Pro	Asp	Ala	Tyr	Asn	Lys	Lys	Gln	Thr	
		275					280				285					
att	cac	tac	tat	gag	cag	atc	acg	ctc	aag	tcc	atg	gcg	ggg	tca	ggc	912
Ile	His	Tyr	Tyr	Glu	Gln	Ile	Thr	Leu	Lys	Ser	Met	Ala	Gly	Ser	Gly	
	290					295				300						
cat	aat	gtc	tcc	caa	gag	gcc	ctg	gcc	att	aag	cgg	atg	ctg	gag	atg	960
His	Asn	Val	Ser	Gln	Glu	Ala	Leu	Ala	Ile	Lys	Arg	Met	Leu	Glu	Met	
305					310					315					320	
gga	gct	atc	aag	aac	ctc	acg	tcc	ttc	tga							990
Gly	Ala	Ile	Lys	Asn	Leu	Thr	Ser	Phe								
				325												
<210>	24															
<211>	329															
<212>	PRT															
<213>	Homo sapiens															
<400>	24															
Met	Val	Ser	Lys	Ser	Arg	Trp	Lys	Leu	Leu	Ala	Met	Leu	Ala	Leu	Val	
1				5					10					15		
Leu	Val	Val	Met	Val	Trp	Tyr	Ser	Ile	Ser	Arg	Glu	Asp	Ser	Phe	Tyr	
			20					25					30			
Phe	Pro	Ile	Pro	Glu	Lys	Lys	Glu	Pro	Cys	Leu	Gln	Gly	Glu	Ala	Glu	
		35					40					45				
Ser	Lys	Ala	Ser	Lys	Leu	Phe	Gly	Asn	Tyr	Ser	Arg	Asp	Gln	Pro	Ile	
	50					55					60					
Phe	Leu	Arg	Leu	Glu	Asp	Tyr	Phe	Trp	Val	Lys	Thr	Pro	Ser	Ala	Tyr	
65					70					75					80	
Glu	Leu	Pro	Tyr	Gly	Thr	Lys	Gly	Ser	Glu	Asp	Leu	Leu	Leu	Arg	Val	
				85					90					95		
Leu	Ala	Ile	Thr	Ser	Ser	Ser	Ile	Pro	Lys	Asn	Ile	Gln	Ser	Leu	Arg	
			100					105					110			
Cys	Arg	Arg	Cys	Val	Val	Val	Gly	Asn	Gly	His	Arg	Leu	Arg	Asn	Ser	
			115				120					125				
Ser	Leu	Gly	Asp	Ala	Ile	Asn	Lys	Tyr	Asp	Val	Val	Ile	Arg	Leu	Asn	
	130					135					140					

ES 2 368 267 T3

Asn Ala Pro Val Ala Gly Tyr Glu Gly Asp Val Gly Ser Lys Thr Thr
145 150 155 160

Met Arg Leu Phe Tyr Pro Glu Ser Ala His Phe Asp Pro Lys Val Glu
165 170 175

Asn Asn Pro Asp Thr Leu Leu Val Leu Val Ala Phe Lys Ala Met Asp
180 185 190

Phe His Trp Ile Glu Thr Ile Leu Ser Asp Lys Lys Arg Val Arg Lys
195 200 205

Gly Phe Trp Lys Gln Pro Pro Leu Ile Trp Asp Val Asn Pro Lys Gln
210 215 220

Ile Arg Ile Leu Asn Pro Phe Phe Met Glu Ile Ala Ala Asp Lys Leu
225 230 235 240

Leu Ser Leu Pro Met Gln Gln Pro Arg Lys Ile Lys Gln Lys Pro Thr
245 250 255

Thr Gly Leu Leu Ala Ile Thr Leu Ala Leu His Leu Cys Asp Leu Val
260 265 270

His Ile Ala Gly Phe Gly Tyr Pro Asp Ala Tyr Asn Lys Lys Gln Thr
275 280 285

Ile His Tyr Tyr Glu Gln Ile Thr Leu Lys Ser Met Ala Gly Ser Gly
290 295 300

His Asn Val Ser Gln Glu Ala Leu Ala Ile Lys Arg Met Leu Glu Met
305 310 315 320

Gly Ala Ile Lys Asn Leu Thr Ser Phe
325

<210> 25
<211> 1089
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1089)

<400> 25

ES 2 368 267 T3

atg aga agg ccc agc ttg tta tta aaa gac atc ctc aaa tgt aca ttg Met Arg Arg Pro Ser Leu Leu Leu Lys Asp Ile Leu Lys Cys Thr Leu 1 5 10 15	48
ctt gtg ttt gga gtg tgg atc ctt tat atc ctc aag tta aat tat act Leu Val Phe Gly Val Trp Ile Leu Tyr Ile Leu Lys Leu Asn Tyr Thr 20 25 30	96
act gaa gaa tgt gac atg aaa aaa atg cat tat gtg gac cct gac cat Thr Glu Glu Cys Asp Met Lys Lys Met His Tyr Val Asp Pro Asp His 35 40 45	144
gta aag aga gct cag aaa tat gct cag caa gtc ttg cag aag gaa tgt Val Lys Arg Ala Gln Lys Tyr Ala Gln Gln Val Leu Gln Lys Glu Cys 50 55 60	192
cgt ccc aag ttt gcc aag aca tca atg gcg ctg tta ttt gag cac agg Arg Pro Lys Phe Ala Lys Thr Ser Met Ala Leu Leu Phe Glu His Arg 65 70 75 80	240
tat agc gtg gac tta ctc cct ttt gtg cag aag gcc ccc aaa gac agt Tyr Ser Val Asp Leu Leu Pro Phe Val Gln Lys Ala Pro Lys Asp Ser 85 90 95	288
gaa gct gag tcc aag tac gat cct cct ttt ggg ttc cgg aag ttc tcc Glu Ala Glu Ser Lys Tyr Asp Pro Pro Phe Gly Phe Arg Lys Phe Ser 100 105 110	336
agt aaa gtc cag acc ctc ttg gaa ctc ttg cca gag cac gac ctc cct Ser Lys Val Gln Thr Leu Leu Glu Leu Leu Pro Glu His Asp Leu Pro 115 120 125	384
gaa cac ttg aaa gcc aag acc tgt cgg cgc tgt gtg gtt att gga agc Glu His Leu Lys Ala Lys Thr Cys Arg Arg Cys Val Val Ile Gly Ser 130 135 140	432
gga gga ata ctg cac gga tta gaa ctg ggc cac acc ctg aac cag ttc Gly Gly Ile Leu His Gly Leu Glu Leu Gly His Thr Leu Asn Gln Phe 145 150 155 160	480
gat gtt gtg ata agg tta aac agt gca cca gtt gag gga tat tca gaa Asp Val Val Ile Arg Leu Asn Ser Ala Pro Val Glu Gly Tyr Ser Glu 165 170 175	528
cat gtt gga aat aaa act act ata agg atg act tat cca gag ggc gca His Val Gly Asn Lys Thr Thr Ile Arg Met Thr Tyr Pro Glu Gly Ala 180 185 190	576
cca ctg tct gac ctt gaa tat tat tcc aat gac tta ttt gtt gct gtt Pro Leu Ser Asp Leu Glu Tyr Tyr Ser Asn Asp Leu Phe Val Ala Val 195 200 205	624
tta ttt aag agt gtt gat ttc aac tgg ctt caa gca atg gta aaa aag Leu Phe Lys Ser Val Asp Phe Asn Trp Leu Gln Ala Met Val Lys Lys 210 215 220	672
gaa acc ctg cca ttc tgg gta cga ctc ttc ttt tgg aag cag gtg gca Glu Thr Leu Pro Phe Trp Val Arg Leu Phe Trp Lys Gln Val Ala 225 230 235 240	720
gaa aaa atc cca ctg cag cca aaa cat ttc agg att ttg aat cca gtt	768

ES 2 368 267 T3

Glu	Lys	Ile	Pro	Leu	Gln	Pro	Lys	His	Phe	Arg	Ile	Leu	Asn	Pro	Val	
				245					250					255		
atc	atc	aaa	gag	act	gcc	ttt	gac	atc	ctt	cag	tac	tca	gag	cct	cag	816
Ile	Ile	Lys	Glu	Thr	Ala	Phe	Asp	Ile	Leu	Gln	Tyr	Ser	Glu	Pro	Gln	
			260					265					270			
tca	agg	ttc	tgg	ggc	cga	gat	aag	aac	gtc	ccc	aca	atc	ggg	gtc	att	864
Ser	Arg	Phe	Trp	Gly	Arg	Asp	Lys	Asn	Val	Pro	Thr	Ile	Gly	Val	Ile	
			275				280					285				
gcc	gtt	gtc	tta	gcc	aca	cat	ctg	tgc	gat	gaa	gtc	agt	ttg	gcg	ggg	912
Ala	Val	Val	Leu	Ala	Thr	His	Leu	Cys	Asp	Glu	Val	Ser	Leu	Ala	Gly	
	290					295					300					
ttt	gga	tat	gac	ctc	aat	caa	ccc	aga	aca	cct	ttg	cac	tac	ttc	gac	960
Phe	Gly	Tyr	Asp	Leu	Asn	Gln	Pro	Arg	Thr	Pro	Leu	His	Tyr	Phe	Asp	
305					310				315						320	
agt	caa	tgc	atg	gct	gct	atg	aac	ttt	cag	acc	atg	cat	aat	gtg	aca	1008
Ser	Gln	Cys	Met	Ala	Ala	Met	Asn	Phe	Gln	Thr	Met	His	Asn	Val	Thr	
				325					330					335		
acg	gaa	acc	aag	ttc	ctc	tta	aag	ctg	gtc	aaa	gag	gga	gtg	gtg	aaa	1056
Thr	Glu	Thr	Lys	Phe	Leu	Leu	Lys	Leu	Val	Lys	Glu	Gly	Val	Val	Lys	
			340					345					350			
gat	ctc	agt	gga	ggc	att	gat	cgt	gaa	ttt	tga						1089
Asp	Leu	Ser	Gly	Gly	Ile	Asp	Arg	Glu	Phe							
		355				360										
<210>	26															
<211>	362															
<212>	PRT															
<213>	Homo sapiens															
<400>	26															
Met	Arg	Arg	Pro	Ser	Leu	Leu	Leu	Lys	Asp	Ile	Leu	Lys	Cys	Thr	Leu	
1				5					10					15		
Leu	Val	Phe	Gly	Val	Trp	Ile	Leu	Tyr	Ile	Leu	Lys	Leu	Asn	Tyr	Thr	
			20					25					30			
Thr	Glu	Glu	Cys	Asp	Met	Lys	Lys	Met	His	Tyr	Val	Asp	Pro	Asp	His	
		35					40					45				
Val	Lys	Arg	Ala	Gln	Lys	Tyr	Ala	Gln	Gln	Val	Leu	Gln	Lys	Glu	Cys	
	50					55					60					
Arg	Pro	Lys	Phe	Ala	Lys	Thr	Ser	Met	Ala	Leu	Leu	Phe	Glu	His	Arg	
65					70					75					80	
Tyr	Ser	Val	Asp	Leu	Leu	Pro	Phe	Val	Gln	Lys	Ala	Pro	Lys	Asp	Ser	
				85					90					95		

ES 2 368 267 T3

Glu	Ala	Glu	Ser	Lys	Tyr	Asp	Pro	Pro	Phe	Gly	Phe	Arg	Lys	Phe	Ser
			100					105					110		
Ser	Lys	Val	Gln	Thr	Leu	Leu	Glu	Leu	Leu	Pro	Glu	His	Asp	Leu	Pro
		115					120					125			
Glu	His	Leu	Lys	Ala	Lys	Thr	Cys	Arg	Arg	Cys	Val	Val	Ile	Gly	Ser
	130					135					140				
Gly	Gly	Ile	Leu	His	Gly	Leu	Glu	Leu	Gly	His	Thr	Leu	Asn	Gln	Phe
145					150					155					160
Asp	Val	Val	Ile	Arg	Leu	Asn	Ser	Ala	Pro	Val	Glu	Gly	Tyr	Ser	Glu
				165					170					175	
His	Val	Gly	Asn	Lys	Thr	Thr	Ile	Arg	Met	Thr	Tyr	Pro	Glu	Gly	Ala
			180					185					190		
Pro	Leu	Ser	Asp	Leu	Glu	Tyr	Tyr	Ser	Asn	Asp	Leu	Phe	Val	Ala	Val
		195					200					205			
Leu	Phe	Lys	Ser	Val	Asp	Phe	Asn	Trp	Leu	Gln	Ala	Met	Val	Lys	Lys
	210					215					220				
Glu	Thr	Leu	Pro	Phe	Trp	Val	Arg	Leu	Phe	Phe	Trp	Lys	Gln	Val	Ala
225					230					235					240
Glu	Lys	Ile	Pro	Leu	Gln	Pro	Lys	His	Phe	Arg	Ile	Leu	Asn	Pro	Val
				245					250					255	
Ile	Ile	Lys	Glu	Thr	Ala	Phe	Asp	Ile	Leu	Gln	Tyr	Ser	Glu	Pro	Gln
			260					265					270		
Ser	Arg	Phe	Trp	Gly	Arg	Asp	Lys	Asn	Val	Pro	Thr	Ile	Gly	Val	Ile
		275					280					285			
Ala	Val	Val	Leu	Ala	Thr	His	Leu	Cys	Asp	Glu	Val	Ser	Leu	Ala	Gly
	290					295					300				
Phe	Gly	Tyr	Asp	Leu	Asn	Gln	Pro	Arg	Thr	Pro	Leu	His	Tyr	Phe	Asp
305					310					315					320
Ser	Gln	Cys	Met	Ala	Ala	Met	Asn	Phe	Gln	Thr	Met	His	Asn	Val	Thr
				325					330					335	

ES 2 368 267 T3

Thr Glu Thr Lys Phe Leu Leu Lys Leu Val Lys Glu Gly Val Val Lys
340 345 350

Asp Leu Ser Gly Gly Ile Asp Arg Glu Phe
355 360

<210> 27
<211> 996
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(996)

<400> 27
atg aga ggg tat ctt gtg gcc ata ttc ctg agt gct gtc ttc ctc tat 48
Met Arg Gly Tyr Leu Val Ala Ile Phe Leu Ser Ala Val Phe Leu Tyr
1 5 10 15

tat gta ctg cat tgc ata tta tgg gga acg aat gtc tat tgg gtg gca 96
Tyr Val Leu His Cys Ile Leu Trp Gly Thr Asn Val Tyr Trp Val Ala
20 25 30

cct gtg gaa atg aaa cgg aga aat aag atc cag cct tgt tta tca aag 144
Pro Val Glu Met Lys Arg Arg Asn Lys Ile Gln Pro Cys Leu Ser Lys
35 40 45

cca gct ttt gcc tct ctg ctg agg ttt cat cag ttt cac cct ttt ctg 192
Pro Ala Phe Ala Ser Leu Leu Arg Phe His Gln Phe His Pro Phe Leu
50 55 60

tgt gcg gct gat ttt aga aag att gct tcc ttg tat ggt agc gat aag 240
Cys Ala Ala Asp Phe Arg Lys Ile Ala Ser Leu Tyr Gly Ser Asp Lys
65 70 75 80

ttt gat ttg ccc tat ggg atg aga aca tca gcg gaa tat ttt cga ctt 288
Phe Asp Leu Pro Tyr Gly Met Arg Thr Ser Ala Glu Tyr Phe Arg Leu
85 90 95

gct ctt tca aaa ctg cag agt tgt gat ctc ttt gat gag ttt gac aac 336
Ala Leu Ser Lys Leu Gln Ser Cys Asp Leu Phe Asp Glu Phe Asp Asn
100 105 110

ata ccc tgt aaa aag tgt gtg gtg gtt ggt aat gga gga gtt ttg aag 384
Ile Pro Cys Lys Lys Cys Val Val Val Gly Asn Gly Gly Val Leu Lys
115 120 125

aat aag aca tta gga gaa aaa atc gac tcc tat gat gta ata ata aga 432
Asn Lys Thr Leu Gly Glu Lys Ile Asp Ser Tyr Asp Val Ile Ile Arg
130 135 140

atg aat aat ggt cct gtt tta gga cat gaa gaa gaa gtt ggg aga agg 480
Met Asn Asn Gly Pro Val Leu Gly His Glu Glu Glu Val Gly Arg Arg
145 150 155 160

aca acc ttc cga ctt ttt tat cca gaa tct gtt ttt tca gat cct att 528

ES 2 368 267 T3

Thr	Thr	Phe	Arg	Leu	Phe	Tyr	Pro	Glu	Ser	Val	Phe	Ser	Asp	Pro	Ile		
				165					170					175			
cac	aat	gac	cct	aat	acg	aca	gtg	att	ctc	act	gct	ttt	aag	cca	cat	576	
His	Asn	Asp	Pro	Asn	Thr	Thr	Val	Ile	Leu	Thr	Ala	Phe	Lys	Pro	His		
			180					185					190				
gat	tta	agg	tgg	ctg	ttg	gaa	ttg	ttg	atg	ggg	gac	aaa	ata	aac	act	624	
Asp	Leu	Arg	Trp	Leu	Leu	Glu	Leu	Leu	Met	Gly	Asp	Lys	Ile	Asn	Thr		
		195					200					205					
aat	ggg	ttt	tgg	aag	aaa	cca	gcc	tta	aac	ctg	att	tat	aaa	cct	tat	672	
Asn	Gly	Phe	Trp	Lys	Lys	Pro	Ala	Leu	Asn	Leu	Ile	Tyr	Lys	Pro	Tyr		
	210					215					220						
caa	atc	cga	ata	tta	gat	cct	ttc	att	atc	aga	aca	gca	gct	tat	gaa	720	
Gln	Ile	Arg	Ile	Leu	Asp	Pro	Phe	Ile	Ile	Arg	Thr	Ala	Ala	Tyr	Glu		
225					230					235					240		
ctg	ctt	cat	ttt	cca	aaa	gtg	ttt	ccc	aaa	aat	cag	aaa	cct	aaa	cac	768	
Leu	Leu	His	Phe	Pro	Lys	Val	Phe	Pro	Lys	Asn	Gln	Lys	Pro	Lys	His		
				245					250					255			
cca	aca	aca	gga	att	att	gcc	atc	aca	ttg	gcg	ttt	tac	ata	tgt	cac	816	
Pro	Thr	Thr	Gly	Ile	Ile	Ala	Ile	Thr	Leu	Ala	Phe	Tyr	Ile	Cys	His		
			260					265					270				
gaa	gtt	cac	cta	gct	ggg	ttt	aaa	tac	aac	ttt	tct	gac	ctc	aag	agt	864	
Glu	Val	His	Leu	Ala	Gly	Phe	Lys	Tyr	Asn	Phe	Ser	Asp	Leu	Lys	Ser		
		275					280					285					
cct	ttg	cac	tac	tat	ggg	aat	gcc	acc	atg	tct	ttg	atg	aat	aag	aac	912	
Pro	Leu	His	Tyr	Tyr	Gly	Asn	Ala	Thr	Met	Ser	Leu	Met	Asn	Lys	Asn		
	290					295					300						
gcg	tat	cac	aat	gtg	act	gca	gag	cag	ctc	ttt	ttg	aag	gac	att	ata	960	
Ala	Tyr	His	Asn	Val	Thr	Ala	Glu	Gln	Leu	Phe	Leu	Lys	Asp	Ile	Ile		
305					310					315					320		
gaa	aaa	aac	ctc	gta	atc	aac	ttg	act	caa	gat	tga					996	
Glu	Lys	Asn	Leu	Val	Ile	Asn	Leu	Thr	Gln	Asp							
				325					330								

<210> 28
 <211> 331
 <212> PRT
 <213> Homo sapiens

<400> 28

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		
Pro	Val	Glu	Met	Lys	Arg	Arg	Asn	Lys	Ile	Gln	Pro	Cys	Leu	Ser	Lys
		35					40					45			

ES 2 368 267 T3

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			

ES 2 368 267 T3

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

<210> 29
<211> 1125
<212> ADN
<213> Rattus rattus

<220>
<221> CDS
<222> (1)..(1125)

<400> 29
atg gga ctc ttg gta ttt gta cgc aac ctg ctg cta gcc ctc tgc ctc 48
Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
1 5 10 15
ttt ctg gtc ctg gga ttt ttg tat tat tct gcc tgg aag cta cac tta 96
Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
20 25 30
ctc caa tgg gaa gac tcc aat tca ctg att ctt tcc ctt gac tcc gct 144
Leu Gln Trp Glu Asp Ser Asn Ser Leu Ile Leu Ser Leu Asp Ser Ala
35 40 45
gga caa acc cta ggc aca gag tat gat agg ctg ggt ttc ctc ctg aag 192
Gly Gln Thr Leu Gly Thr Glu Tyr Asp Arg Leu Gly Phe Leu Leu Lys
50 55 60
ctg gac tct aaa ctg cct gca gag ctg gcc acc aag tac gct aac ttt 240
Leu Asp Ser Lys Leu Pro Ala Glu Leu Ala Thr Lys Tyr Ala Asn Phe
65 70 75 80
tcc gag gga gcc tgc aaa ccc ggc tac gct tca gcc atg atg act gcc 288
Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala Ser Ala Met Met Thr Ala
85 90 95
atc ttc ccc agg ttc tcc aag cca gca ccc atg ttc ctg gat gac tcc 336
Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe Leu Asp Asp Ser
100 105 110
ttt cgc aaa tgg gct agg att cgg gag ttt gtg cca ccc ttt ggg atc 384
Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe Gly Ile
115 120 125
aaa ggt caa gac aat ctg atc aaa gcc atc ttg tca gtc acc aaa gaa 432
Lys Gly Gln Asp Asn Leu Ile Lys Ala Ile Leu Ser Val Thr Lys Glu
130 135 140
tac cgc ctg acc cct gcc ttg gac agc ctc cac tgc cgc cgc tgc atc 480

ES 2 368 267 T3

Tyr	Arg	Leu	Thr	Pro	Ala	Leu	Asp	Ser	Leu	His	Cys	Arg	Arg	Cys	Ile	
145					150					155					160	
atc	gta	ggc	aat	gga	ggg	gtc	ctc	gcc	aac	aag	tct	ctg	ggg	tca	cga	528
Ile	Val	Gly	Asn	Gly	Gly	Val	Leu	Ala	Asn	Lys	Ser	Leu	Gly	Ser	Arg	
				165					170					175		
att	gac	gac	tat	gac	att	gtg	atc	aga	ttg	aac	tca	gca	cct	gtg	aag	576
Ile	Asp	Asp	Tyr	Asp	Ile	Val	Ile	Arg	Leu	Asn	Ser	Ala	Pro	Val	Lys	
			180					185					190			
ggc	ttt	gag	aag	gac	gtg	ggc	agc	aag	acc	acc	ctg	cgc	atc	acc	tac	624
Gly	Phe	Glu	Lys	Asp	Val	Gly	Ser	Lys	Thr	Thr	Leu	Arg	Ile	Thr	Tyr	
		195					200					205				
cct	gaa	ggc	gcc	atg	cag	cgg	cct	gag	caa	tat	gaa	cga	gac	tct	ctc	672
Pro	Glu	Gly	Ala	Met	Gln	Arg	Pro	Glu	Gln	Tyr	Glu	Arg	Asp	Ser	Leu	
	210					215					220					
ttt	gta	cta	gct	ggc	ttc	aag	tgg	cag	gac	ttc	aag	tgg	ctg	aag	tac	720
Phe	Val	Leu	Ala	Gly	Phe	Lys	Trp	Gln	Asp	Phe	Lys	Trp	Leu	Lys	Tyr	
225					230					235					240	
atc	gtc	tac	aag	gag	aga	gtg	agc	gca	tcc	gat	ggc	ttc	tgg	aag	tcc	768
Ile	Val	Tyr	Lys	Glu	Arg	Val	Ser	Ala	Ser	Asp	Gly	Phe	Trp	Lys	Ser	
				245					250					255		
gtg	gcc	acc	cga	gtg	ccc	aag	gag	ccc	cct	gag	atc	cgc	atc	ctc	aac	816
Val	Ala	Thr	Arg	Val	Pro	Lys	Glu	Pro	Pro	Glu	Ile	Arg	Ile	Leu	Asn	
			260					265					270			
ccg	tac	ttc	atc	cag	gag	gct	gcc	ttc	acg	ctc	atc	gga	ctg	ccc	ttc	864
Pro	Tyr	Phe	Ile	Gln	Glu	Ala	Ala	Phe	Thr	Leu	Ile	Gly	Leu	Pro	Phe	
		275					280					285				
aac	aat	ggc	ctc	atg	ggc	aga	ggg	aac	atc	cca	acc	ctt	ggc	agt	gtg	912
Asn	Asn	Gly	Leu	Met	Gly	Arg	Gly	Asn	Ile	Pro	Thr	Leu	Gly	Ser	Val	
	290					295					300					
gca	gtg	acc	atg	gca	ctc	gat	ggc	tgt	gat	gaa	gtg	gca	gtc	gcg	ggc	960
Ala	Val	Thr	Met	Ala	Leu	Asp	Gly	Cys	Asp	Glu	Val	Ala	Val	Ala	Gly	
305					310					315					320	
ttt	ggc	tat	gac	atg	aac	aca	ccc	aac	gcc	ccc	ctg	cac	tac	tat	gaa	1008
Phe	Gly	Tyr	Asp	Met	Asn	Thr	Pro	Asn	Ala	Pro	Leu	His	Tyr	Tyr	Glu	
				325					330					335		
act	gtg	cgc	atg	gca	gcc	atc	aaa	gag	tcc	tgg	aca	cac	aac	atc	cag	1056
Thr	Val	Arg	Met	Ala	Ala	Ile	Lys	Glu	Ser	Trp	Thr	His	Asn	Ile	Gln	
			340					345					350			
cga	gag	aaa	gag	ttt	ctg	cgg	aag	cta	gtg	aaa	gca	cgc	gtc	atc	act	1104
Arg	Glu	Lys	Glu	Phe	Leu	Arg	Lys	Leu	Val	Lys	Ala	Arg	Val	Ile	Thr	
		355					360					365				
gac	tta	agc	agt	ggt	atc	tga										1125
Asp	Leu	Ser	Ser	Gly	Ile											
	370															

<210> 30

ES 2 368 267 T3

<211> 374
 <212> PRT
 <213> Rattus rattus

<400> 30

```

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

ES 2 368 267 T3

290	295	300
Ala Val Thr Met Ala Leu Asp Gly Cys Asp Glu Val Ala Val Ala Gly		
305	310	315 320
Phe Gly Tyr Asp Met Asn Thr Pro Asn Ala Pro Leu His Tyr Tyr Glu		
	325	330 335
Thr Val Arg Met Ala Ala Ile Lys Glu Ser Trp Thr His Asn Ile Gln		
	340	345 350
Arg Glu Lys Glu Phe Leu Arg Lys Leu Val Lys Ala Arg Val Ile Thr		
	355	360 365
Asp Leu Ser Ser Gly Ile		
370		

<210> 31
 <211> 1071
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1071)

<400> 31	
atg agc ccc tgc ggg cgg gcc cgg cga caa acg tcc aga ggg gcc atg	48
Met Ser Pro Cys Gly Arg Ala Arg Arg Gln Thr Ser Arg Gly Ala Met	
1 5 10 15	
gct gta ctg gcg tgg aag ttc ccg cgg acc cgg ctg ccc atg gga gcc	96
Ala Val Leu Ala Trp Lys Phe Pro Arg Thr Arg Leu Pro Met Gly Ala	
20 25 30	
agt gcc ctc tgt gtc gtg gtc ctc tgt tgg ctc tac atc ttc ccc gtc	144
Ser Ala Leu Cys Val Val Val Leu Cys Trp Leu Tyr Ile Phe Pro Val	
35 40 45	
tac cgg ctg ccc aac gag aaa gag atc gtg cag ggg gtg ctg caa cag	192
Tyr Arg Leu Pro Asn Glu Lys Glu Ile Val Gln Gly Val Leu Gln Gln	
50 55 60	
ggc acg gcg tgg agg agg aac cag acc gcg gcc aga gcg ttc agg aaa	240
Gly Thr Ala Trp Arg Arg Asn Gln Thr Ala Ala Arg Ala Phe Arg Lys	
65 70 75 80	
caa atg gaa gac tgc tgc gac cct gcc cat ctc ttt gct atg act aaa	288
Gln Met Glu Asp Cys Cys Asp Pro Ala His Leu Phe Ala Met Thr Lys	
85 90 95	
atg aat tcc cct atg ggg aag agc atg tgg tat gac ggg gag ttt tta	336
Met Asn Ser Pro Met Gly Lys Ser Met Trp Tyr Asp Gly Glu Phe Leu	
100 105 110	
tac tca ttc acc att gac aat tca act tac tct ctc ttc cca cag gca	384
Tyr Ser Phe Thr Ile Asp Asn Ser Thr Tyr Ser Leu Phe Pro Gln Ala	
115 120 125	

ES 2 368 267 T3

acc cca ttc cag ctg cca ttg aag aaa tgc gcg gtg gtg gga aat ggt	432
Thr Pro Phe Gln Leu Pro Leu Lys Lys Cys Ala Val Val Gly Asn Gly	
130 135 140	
ggg att ctg aag aag agt ggc tgt ggc cgt caa ata gat gaa gca aat	480
Gly Ile Leu Lys Lys Ser Gly Cys Gly Arg Gln Ile Asp Glu Ala Asn	
145 150 155 160	
ttt gtc atg cga tgc aat ctc cct cct ttg tca agt gaa tac act aag	528
Phe Val Met Arg Cys Asn Leu Pro Pro Leu Ser Ser Glu Tyr Thr Lys	
165 170 175	
gat gtt gga tcc aaa agt cag tta gtg aca gct aat ccc agc ata att	576
Asp Val Gly Ser Lys Ser Gln Leu Val Thr Ala Asn Pro Ser Ile Ile	
180 185 190	
cgg caa agg ttt cag aac ctt ctg tgg tcc aga aag aca ttt gtg gac	624
Arg Gln Arg Phe Gln Asn Leu Leu Trp Ser Arg Lys Thr Phe Val Asp	
195 200 205	
aac atg aaa att tat aac cac agt tac atc tac atg cct gcc ttt tct	672
Asn Met Lys Ile Tyr Asn His Ser Tyr Ile Tyr Met Pro Ala Phe Ser	
210 215 220	
atg aag aca gga aca gag cca tct ttg agg gtt tat tat aca ctg tca	720
Met Lys Thr Gly Thr Glu Pro Ser Leu Arg Val Tyr Tyr Thr Leu Ser	
225 230 235 240	
gat gtt ggt gcc aat caa aca gtg ctg ttt gcc aac ccc aac ttt ctg	768
Asp Val Gly Ala Asn Gln Thr Val Leu Phe Ala Asn Pro Asn Phe Leu	
245 250 255	
cgt agc att gga aag ttc tgg aaa agt aga gga atc cat gcc aag cgc	816
Arg Ser Ile Gly Lys Phe Trp Lys Ser Arg Gly Ile His Ala Lys Arg	
260 265 270	
ctg tcc aca gga ctt ttt ctg gtg agc gca gct ctg ggt ctc tgt gaa	864
Leu Ser Thr Gly Leu Phe Leu Val Ser Ala Ala Leu Gly Leu Cys Glu	
275 280 285	
gag gtg gcc atc tat ggc ttc tgg ccc ttc tct gtg aat atg cat gag	912
Glu Val Ala Ile Tyr Gly Phe Trp Pro Phe Ser Val Asn Met His Glu	
290 295 300	
cag ccc atc agc cac cac tac tat gac aac gtc tta ccc ttt tct ggc	960
Gln Pro Ile Ser His His Tyr Tyr Asp Asn Val Leu Pro Phe Ser Gly	
305 310 315 320	
ttc cat gcc atg ccc gag gaa ttt ctc caa ctc tgg tat ctt cat aaa	1008
Phe His Ala Met Pro Glu Glu Phe Leu Gln Leu Trp Tyr Leu His Lys	
325 330 335	
atc ggt gca ctg aga atg cag ctg gac cca tgt gaa gat acc tca ctc	1056
Ile Gly Ala Leu Arg Met Gln Leu Asp Pro Cys Glu Asp Thr Ser Leu	
340 345 350	
cag ccc act tcc tag	1071
Gln Pro Thr Ser	
355	

ES 2 368 267 T3

<210> 32
 <211> 356
 <212> PRT
 <213> Homo sapiens

<400> 32

Met Ser Pro Cys Gly Arg Ala Arg Arg Gln Thr Ser Arg Gly Ala Met
 1 5 10 15

Ala Val Leu Ala Trp Lys Phe Pro Arg Thr Arg Leu Pro Met Gly Ala
 20 25 30

Ser Ala Leu Cys Val Val Val Leu Cys Trp Leu Tyr Ile Phe Pro Val
 35 40 45

Tyr Arg Leu Pro Asn Glu Lys Glu Ile Val Gln Gly Val Leu Gln Gln
 50 55 60

Gly Thr Ala Trp Arg Arg Asn Gln Thr Ala Ala Arg Ala Phe Arg Lys
 65 70 75 80

Gln Met Glu Asp Cys Cys Asp Pro Ala His Leu Phe Ala Met Thr Lys
 85 90 95

Met Asn Ser Pro Met Gly Lys Ser Met Trp Tyr Asp Gly Glu Phe Leu
 100 105 110

Tyr Ser Phe Thr Ile Asp Asn Ser Thr Tyr Ser Leu Phe Pro Gln Ala
 115 120 125

Thr Pro Phe Gln Leu Pro Leu Lys Lys Cys Ala Val Val Gly Asn Gly
 130 135 140

Gly Ile Leu Lys Lys Ser Gly Cys Gly Arg Gln Ile Asp Glu Ala Asn
 145 150 155 160

Phe Val Met Arg Cys Asn Leu Pro Pro Leu Ser Ser Glu Tyr Thr Lys
 165 170 175

Asp Val Gly Ser Lys Ser Gln Leu Val Thr Ala Asn Pro Ser Ile Ile
 180 185 190

Arg Gln Arg Phe Gln Asn Leu Leu Trp Ser Arg Lys Thr Phe Val Asp
 195 200 205

Asn Met Lys Ile Tyr Asn His Ser Tyr Ile Tyr Met Pro Ala Phe Ser
 210 215 220

ES 2 368 267 T3

Met Lys Thr Gly Thr Glu Pro Ser Leu Arg Val Tyr Tyr Thr Leu Ser
225 230 235 240

Asp Val Gly Ala Asn Gln Thr Val Leu Phe Ala Asn Pro Asn Phe Leu
245 250 255

Arg Ser Ile Gly Lys Phe Trp Lys Ser Arg Gly Ile His Ala Lys Arg
260 265 270

Leu Ser Thr Gly Leu Phe Leu Val Ser Ala Ala Leu Gly Leu Cys Glu
275 280 285

Glu Val Ala Ile Tyr Gly Phe Trp Pro Phe Ser Val Asn Met His Glu
290 295 300

Gln Pro Ile Ser His His Tyr Tyr Asp Asn Val Leu Pro Phe Ser Gly
305 310 315 320

Phe His Ala Met Pro Glu Glu Phe Leu Gln Leu Trp Tyr Leu His Lys
325 330 335

Ile Gly Ala Leu Arg Met Gln Leu Asp Pro Cys Glu Asp Thr Ser Leu
340 345 350

Gln Pro Thr Ser
355

<210> 33
<211> 1128
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1128)

<400> 33
atg cag ctg cag ttc cgg agc tgg atg ctg gcc gcg ctc acg ctg ctc 48
Met Gln Leu Gln Phe Arg Ser Trp Met Leu Ala Ala Leu Thr Leu Leu
1 5 10 15
gtg gtc ttc ctc atc ttc gca gac atc tca gag atc gaa gaa gaa atc 96
Val Val Phe Leu Ile Phe Ala Asp Ile Ser Glu Ile Glu Glu Glu Ile
20 25 30
ggg aat tcg gga ggc aga ggt aca atc aga tca gct gtg aac agc tta 144
Gly Asn Ser Gly Gly Arg Gly Thr Ile Arg Ser Ala Val Asn Ser Leu
35 40 45

ES 2 368 267 T3

cat agc aaa tct aat aga gct gaa gtt gta ata aac ggc tcc tca tca	192
His Ser Lys Ser Asn Arg Ala Glu Val Val Ile Asn Gly Ser Ser Ser	
50 55 60	
cca gct gtt gtt gac aga agt aat gaa agc atc aag cac aac atc cag	240
Pro Ala Val Val Asp Arg Ser Asn Glu Ser Ile Lys His Asn Ile Gln	
65 70 75 80	
cca gcc tcg tcc aaa tgg aga cat aac cag acg ctc tct ctg agg atc	288
Pro Ala Ser Ser Lys Trp Arg His Asn Gln Thr Leu Ser Leu Arg Ile	
85 90 95	
agg aag cag att tta aag ttc ttg gat gct gaa aag gac att tct gtc	336
Arg Lys Gln Ile Leu Lys Phe Leu Asp Ala Glu Lys Asp Ile Ser Val	
100 105 110	
cta aag gga acc ctg aag cct gga gat att att cat tac atc ttc gat	384
Leu Lys Gly Thr Leu Lys Pro Gly Asp Ile Ile His Tyr Ile Phe Asp	
115 120 125	
cga gac agc acc atg aat gtg tcc cag aac ctc tac gag ctc ctc ccc	432
Arg Asp Ser Thr Met Asn Val Ser Gln Asn Leu Tyr Glu Leu Leu Pro	
130 135 140	
agg act tcg cca ctg aag aat aag cac ttt ggg act tgt gcc atc gtg	480
Arg Thr Ser Pro Leu Lys Asn Lys His Phe Gly Thr Cys Ala Ile Val	
145 150 155 160	
ggc aac tcg ggg gtc ttg ctg aac agc ggc tgt ggg cag gag att gac	528
Gly Asn Ser Gly Val Leu Leu Asn Ser Gly Cys Gly Gln Glu Ile Asp	
165 170 175	
gcc cac agc ttc gtc atc agg tgc aac ctg gcc cca gta cag gag tat	576
Ala His Ser Phe Val Ile Arg Cys Asn Leu Ala Pro Val Gln Glu Tyr	
180 185 190	
gcc cgg gat gtg ggg ctc aag aca gac ctg gta acc atg aac ccc tcg	624
Ala Arg Asp Val Gly Leu Lys Thr Asp Leu Val Thr Met Asn Pro Ser	
195 200 205	
gtc atc cag cgg gcc ttt gag gac ttg gtc aat gcc acg tgg cgg gag	672
Val Ile Gln Arg Ala Phe Glu Asp Leu Val Asn Ala Thr Trp Arg Glu	
210 215 220	
aag ctg ctg caa cgg ctg cac agc ctc aat ggc agc atc ctg tgg atc	720
Lys Leu Leu Gln Arg Leu His Ser Leu Asn Gly Ser Ile Leu Trp Ile	
225 230 235 240	
cct gcc ttc atg gcc cgg ggc ggc aag gag cgt gtt gag tgg gtc aac	768
Pro Ala Phe Met Ala Arg Gly Gly Lys Glu Arg Val Glu Trp Val Asn	
245 250 255	
gag ctt atc ctg aag cac cac gtc aac gtg cgc act gca tac ccc tcg	816
Glu Leu Ile Leu Lys His His Val Asn Val Arg Thr Ala Tyr Pro Ser	
260 265 270	
ctg cgc ctg ctg cac gcc gtt cgc gga tac tgg ctg acc aac aaa gtc	864
Leu Arg Leu Leu His Ala Val Arg Gly Tyr Trp Leu Thr Asn Lys Val	
275 280 285	
cac atc aaa aga ccc acc acc ggc ctc ttg atg tat acc ctg gcc aca	912

ES 2 368 267 T3

His	Ile	Lys	Arg	Pro	Thr	Thr	Gly	Leu	Leu	Met	Tyr	Thr	Leu	Ala	Thr	
	290						295				300					
cgt	ttc	tgc	aaa	caa	atc	tac	ctc	tac	ggc	ttc	tgg	ccc	ttt	ccg	ctg	960
Arg	Phe	Cys	Lys	Gln	Ile	Tyr	Leu	Tyr	Gly	Phe	Trp	Pro	Phe	Pro	Leu	
305					310				315					320		
gat	cag	aac	cag	aac	cca	gtc	aag	tac	cac	tat	tat	gac	agc	ctc	aag	1008
Asp	Gln	Asn	Gln	Asn	Pro	Val	Lys	Tyr	His	Tyr	Tyr	Asp	Ser	Leu	Lys	
				325					330					335		
tat	ggc	tac	acc	tcc	cag	gcc	agc	ccg	cat	acc	atg	ccc	ttg	gag	ttt	1056
Tyr	Gly	Tyr	Thr	Ser	Gln	Ala	Ser	Pro	His	Thr	Met	Pro	Leu	Glu	Phe	
			340					345					350			
aag	gcc	ctc	aag	agc	cta	cat	gag	cag	ggg	gct	ttg	aaa	ctg	act	gtc	1104
Lys	Ala	Leu	Lys	Ser	Leu	His	Glu	Gln	Gly	Ala	Leu	Lys	Leu	Thr	Val	
	355						360					365				
ggc	cag	tgc	gat	ggg	gcc	acg	tag									1128
Gly	Gln	Cys	Asp	Gly	Ala	Thr										
	370					375										
<210>	34															
<211>	375															
<212>	PRT															
<213>	Homo sapiens															
<400>	34															
Met	Gln	Leu	Gln	Phe	Arg	Ser	Trp	Met	Leu	Ala	Ala	Leu	Thr	Leu	Leu	
1				5					10					15		
Val	Val	Phe	Leu	Ile	Phe	Ala	Asp	Ile	Ser	Glu	Ile	Glu	Glu	Glu	Ile	
			20					25					30			
Gly	Asn	Ser	Gly	Gly	Arg	Gly	Thr	Ile	Arg	Ser	Ala	Val	Asn	Ser	Leu	
	35						40					45				
His	Ser	Lys	Ser	Asn	Arg	Ala	Glu	Val	Val	Ile	Asn	Gly	Ser	Ser	Ser	
	50					55					60					
Pro	Ala	Val	Val	Asp	Arg	Ser	Asn	Glu	Ser	Ile	Lys	His	Asn	Ile	Gln	
65				70						75				80		
Pro	Ala	Ser	Ser	Lys	Trp	Arg	His	Asn	Gln	Thr	Leu	Ser	Leu	Arg	Ile	
				85					90					95		
Arg	Lys	Gln	Ile	Leu	Lys	Phe	Leu	Asp	Ala	Glu	Lys	Asp	Ile	Ser	Val	
		100						105					110			
Leu	Lys	Gly	Thr	Leu	Lys	Pro	Gly	Asp	Ile	Ile	His	Tyr	Ile	Phe	Asp	
		115					120					125				

ES 2 368 267 T3

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

ES 2 368 267 T3

Gly Gln Cys Asp Gly Ala Thr
370 375

<210> 35
<211> 1143
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1143)

<400> 35
atg aga aac tgc aaa atg gcc cgg gtc gcc agt gtg ctg ggg ctg gtc 48
Met Arg Asn Cys Lys Met Ala Arg Val Ala Ser Val Leu Gly Leu Val
1 5 10 15

atg ctc agc gtc gcc ctg ctg att tta tcg ctc atc agc tac gtg tcc 96
Met Leu Ser Val Ala Leu Leu Ile Leu Ser Leu Ile Ser Tyr Val Ser
20 25 30

ctg aaa aag gag aac atc ttc acc act ccc aag tac gcc agc ccg ggg 144
Leu Lys Lys Glu Asn Ile Phe Thr Thr Pro Lys Tyr Ala Ser Pro Gly
35 40 45

gcg ccc cga atg tac atg ttc cac gcg gga ttc cgg tca caa ttt gcg 192
Ala Pro Arg Met Tyr Met Phe His Ala Gly Phe Arg Ser Gln Phe Ala
50 55 60

ctg aag ttt cta gac ccg tca ttc gtg ccc att acg aat tct ctc acc 240
Leu Lys Phe Leu Asp Pro Ser Phe Val Pro Ile Thr Asn Ser Leu Thr
65 70 75 80

cag gaa ctc caa gag aaa cct tct aag tgg aaa ttt aat cgg aca gcg 288
Gln Glu Leu Gln Glu Lys Pro Ser Lys Trp Lys Phe Asn Arg Thr Ala
85 90 95

ttt tta cat caa agg caa gaa att ctt cag cat gtc gat gta ata aaa 336
Phe Leu His Gln Arg Gln Glu Ile Leu Gln His Val Asp Val Ile Lys
100 105 110

aat ttt tct ttg acc aag aat agt gtt cgg att gga caa ctg atg cac 384
Asn Phe Ser Leu Thr Lys Asn Ser Val Arg Ile Gly Gln Leu Met His
115 120 125

tat gat tat tcc agc cat aaa tat gtt ttc tct att agc aat aac ttc 432
Tyr Asp Tyr Ser Ser His Lys Tyr Val Phe Ser Ile Ser Asn Asn Phe
130 135 140

cgg tca ctt ctt cca gat gtg tca ccc att atg aac aag cat tat aat 480
Arg Ser Leu Leu Pro Asp Val Ser Pro Ile Met Asn Lys His Tyr Asn
145 150 155 160

att tgt gct gtg gtt gga aat agt ggg atc ctg aca ttt atc cag tgt 528
Ile Cys Ala Val Val Gly Asn Ser Gly Ile Leu Thr Phe Ile Gln Cys
165 170 175

gga cga gaa ata gat aaa tca gat ttt gtt ttc cgt tgc aat ttc gcc 576

ES 2 368 267 T3

Gly	Arg	Glu	Ile	Asp	Lys	Ser	Asp	Phe	Val	Phe	Arg	Cys	Asn	Phe	Ala		
			180					185					190				
cca	tcg	gag	gct	ttc	caa	aga	gat	gtt	gga	agg	aaa	acc	aat	ctt	acc		624
Pro	Ser	Glu	Ala	Phe	Gln	Arg	Asp	Val	Gly	Arg	Lys	Thr	Asn	Leu	Thr		
		195					200					205					
acc	ttc	aac	ccc	agc	atc	ctg	gaa	aaa	tat	tac	aac	aat	ctc	ttg	act		672
Thr	Phe	Asn	Pro	Ser	Ile	Leu	Glu	Lys	Tyr	Tyr	Asn	Asn	Leu	Leu	Thr		
	210					215					220						
att	cag	gac	cgt	aac	aac	ttt	ttc	ctc	agt	tta	aaa	aag	ctt	gac	ggg		720
Ile	Gln	Asp	Arg	Asn	Asn	Phe	Phe	Leu	Ser	Leu	Lys	Lys	Leu	Asp	Gly		
225					230					235					240		
gcc	att	ctt	tgg	atc	cct	gca	ttt	ttc	ttc	cac	act	tca	gca	act	gtg		768
Ala	Ile	Leu	Trp	Ile	Pro	Ala	Phe	Phe	Phe	His	Thr	Ser	Ala	Thr	Val		
				245					250					255			
acc	agg	aca	tta	gtt	gac	ttt	ttt	gtt	gaa	cac	aga	ggt	cag	tta	aaa		816
Thr	Arg	Thr	Leu	Val	Asp	Phe	Phe	Val	Glu	His	Arg	Gly	Gln	Leu	Lys		
			260					265					270				
gtc	caa	ctg	gct	tgg	ccg	gga	aat	ata	atg	caa	cat	gtc	aac	agg	tac		864
Val	Gln	Leu	Ala	Trp	Pro	Gly	Asn	Ile	Met	Gln	His	Val	Asn	Arg	Tyr		
		275					280					285					
tgg	aaa	aac	aaa	cat	ttg	tca	cct	aaa	cgg	ctg	agc	aca	ggt	att	ctt		912
Trp	Lys	Asn	Lys	His	Leu	Ser	Pro	Lys	Arg	Leu	Ser	Thr	Gly	Ile	Leu		
	290					295					300						
atg	tac	acc	ctt	gca	tca	gca	ata	tgt	gaa	gag	atc	cac	ttg	tat	gga		960
Met	Tyr	Thr	Leu	Ala	Ser	Ala	Ile	Cys	Glu	Glu	Ile	His	Leu	Tyr	Gly		
305					310				315					320			
ttt	tgg	ccg	ttt	gga	ttt	gac	ccc	aac	aca	agg	gaa	gat	ctt	cca	tac		1008
Phe	Trp	Pro	Phe	Gly	Phe	Asp	Pro	Asn	Thr	Arg	Glu	Asp	Leu	Pro	Tyr		
				325					330					335			
cat	tac	tat	gac	aaa	aaa	gga	acc	aaa	ttt	acc	acc	aag	tgg	cag	gag		1056
His	Tyr	Tyr	Asp	Lys	Lys	Gly	Thr	Lys	Phe	Thr	Thr	Lys	Trp	Gln	Glu		
			340					345					350				
tcc	cac	cag	ctg	cct	gct	gag	ttt	cag	ctg	ctg	tac	cga	atg	cat	ggg		1104
Ser	His	Gln	Leu	Pro	Ala	Glu	Phe	Gln	Leu	Leu	Tyr	Arg	Met	His	Gly		
		355				360						365					
gaa	ggg	ctc	acc	aag	ctg	act	ctg	tca	cac	tgt	gcc	taa					1143
Glu	Gly	Leu	Thr	Lys	Leu	Thr	Leu	Ser	His	Cys	Ala						
	370					375					380						

<210> 36
 <211> 380
 <212> PRT
 <213> Homo sapiens

<400> 36

Met	Arg	Asn	Cys	Lys	Met	Ala	Arg	Val	Ala	Ser	Val	Leu	Gly	Leu	Val
1				5					10					15	

ES 2 368 267 T3

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

ES 2 368 267 T3

Thr Arg Thr Leu Val Asp Phe Phe Val Glu His Arg Gly Gln Leu Lys
 260 265 270

Val Gln Leu Ala Trp Pro Gly Asn Ile Met Gln His Val Asn Arg Tyr
 275 280 285

Trp Lys Asn Lys His Leu Ser Pro Lys Arg Leu Ser Thr Gly Ile Leu
 290 295 300

Met Tyr Thr Leu Ala Ser Ala Ile Cys Glu Glu Ile His Leu Tyr Gly
 305 310 315 320

Phe Trp Pro Phe Gly Phe Asp Pro Asn Thr Arg Glu Asp Leu Pro Tyr
 325 330 335

His Tyr Tyr Asp Lys Lys Gly Thr Lys Phe Thr Thr Lys Trp Gln Glu
 340 345 350

Ser His Gln Leu Pro Ala Glu Phe Gln Leu Leu Tyr Arg Met His Gly
 355 360 365

Glu Gly Leu Thr Lys Leu Thr Leu Ser His Cys Ala
 370 375 380

<210> 37
 <211> 1080
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1080)

<400> 37
 atg cgc tcc att agg aag agg tgg acg atc tgc aca ata agt ctg ctc 48
 Met Arg Ser Ile Arg Lys Arg Trp Thr Ile Cys Thr Ile Ser Leu Leu
 1 5 10 15
 ctg atc ttt tat aag aca aaa gaa ata gca aga act gag gag cac cag 96
 Leu Ile Phe Tyr Lys Thr Lys Glu Ile Ala Arg Thr Glu Glu His Gln
 20 25 30
 gag acg caa ctc atc gga gat ggt gaa ttg tct ttg agt cgg tca ctt 144
 Glu Thr Gln Leu Ile Gly Asp Gly Glu Leu Ser Leu Ser Arg Ser Leu
 35 40 45
 gtc aat agc tct gat aaa atc att cga aag gct ggc tct tca atc ttc 192
 Val Asn Ser Ser Asp Lys Ile Ile Arg Lys Ala Gly Ser Ser Ile Phe
 50 55 60
 cag cac aat gta gaa ggt tgg aaa atc aat tcc tct ttg gtc cta gag 240

ES 2 368 267 T3

Gln	His	Asn	Val	Glu	Gly	Trp	Lys	Ile	Asn	Ser	Ser	Leu	Val	Leu	Glu	
65					70					75					80	
ata	agg	aag	aac	ata	ctt	cgt	ttc	tta	gat	gca	gaa	cga	gat	gtg	tca	288
Ile	Arg	Lys	Asn	Ile	Leu	Arg	Phe	Leu	Asp	Ala	Glu	Arg	Asp	Val	Ser	
			85						90					95		
gtg	gtc	aag	agc	agt	ttt	aag	cct	ggc	gat	gtc	ata	cac	tat	gtg	ctt	336
Val	Val	Lys	Ser	Ser	Phe	Lys	Pro	Gly	Asp	Val	Ile	His	Tyr	Val	Leu	
			100					105					110			
gac	agg	cgc	cgg	aca	cta	aac	att	tct	cat	gat	cta	cat	agc	ctc	cta	384
Asp	Arg	Arg	Arg	Thr	Leu	Asn	Ile	Ser	His	Asp	Leu	His	Ser	Leu	Leu	
			115				120					125				
cct	gaa	gtt	tca	cca	atg	aag	aat	cgc	agg	ttt	aag	acc	tgt	gca	gtt	432
Pro	Glu	Val	Ser	Pro	Met	Lys	Asn	Arg	Arg	Phe	Lys	Thr	Cys	Ala	Val	
	130					135					140					
gtt	gga	aat	tct	ggc	att	ctg	tta	gac	agt	gaa	tgt	gga	aag	gag	att	480
Val	Gly	Asn	Ser	Gly	Ile	Leu	Leu	Asp	Ser	Glu	Cys	Gly	Lys	Glu	Ile	
145					150					155					160	
gac	agt	cac	aat	ttt	gta	ata	agg	tgt	aat	cta	gct	cct	gtg	gtg	gag	528
Asp	Ser	His	Asn	Phe	Val	Ile	Arg	Cys	Asn	Leu	Ala	Pro	Val	Val	Glu	
				165					170					175		
ttt	gct	gca	gat	gtg	gga	act	aaa	tca	gat	ttt	att	acc	atg	aat	cca	576
Phe	Ala	Ala	Asp	Val	Gly	Thr	Lys	Ser	Asp	Phe	Ile	Thr	Met	Asn	Pro	
			180					185					190			
tca	gtt	gta	caa	aga	gca	ttt	gga	ggc	ttt	cga	aat	gag	agt	gac	aga	624
Ser	Val	Val	Gln	Arg	Ala	Phe	Gly	Gly	Phe	Arg	Asn	Glu	Ser	Asp	Arg	
		195					200					205				
gaa	aaa	ttt	gtg	cat	aga	ctt	tcc	atg	ctg	aat	gac	agt	gtc	ctt	tgg	672
Glu	Lys	Phe	Val	His	Arg	Leu	Ser	Met	Leu	Asn	Asp	Ser	Val	Leu	Trp	
	210					215					220					
att	cct	gct	ttc	atg	gtc	aaa	gga	gga	gag	aag	cac	gtg	gag	tgg	gtt	720
Ile	Pro	Ala	Phe	Met	Val	Lys	Gly	Gly	Glu	Lys	His	Val	Glu	Trp	Val	
225					230					235					240	
aat	gca	tta	atc	ctt	aag	aat	aaa	ctg	aaa	gtg	cga	act	gcc	tat	ccg	768
Asn	Ala	Leu	Ile	Leu	Lys	Asn	Lys	Leu	Lys	Val	Arg	Thr	Ala	Tyr	Pro	
				245					250					255		
tca	ttg	aga	ctt	att	cat	gct	gtc	aga	ggc	tac	tgg	ctg	acc	aac	aaa	816
Ser	Leu	Arg	Leu	Ile	His	Ala	Val	Arg	Gly	Tyr	Trp	Leu	Thr	Asn	Lys	
			260					265					270			
gtt	cct	atc	aaa	aga	ccc	agc	aca	ggc	ctt	ctc	atg	tat	aca	ctt	gcc	864
Val	Pro	Ile	Lys	Arg	Pro	Ser	Thr	Gly	Leu	Leu	Met	Tyr	Thr	Leu	Ala	
		275					280					285				
aca	aga	ttc	tgt	gat	gaa	att	cac	ctg	tat	gga	ttc	tgg	ccc	ttc	cct	912
Thr	Arg	Phe	Cys	Asp	Glu	Ile	His	Leu	Tyr	Gly	Phe	Trp	Pro	Phe	Pro	
	290					295					300					
aag	gat	tta	aat	gga	aaa	gcg	gtc	aaa	tat	cat	tat	tat	gat	gac	tta	960
Lys	Asp	Leu	Asn	Gly	Lys	Ala	Val	Lys	Tyr	His	Tyr	Tyr	Asp	Asp	Leu	

ES 2 368 267 T3

305		310		315		320	
aaa tat agg tac ttt tcc aat gca agc cct cac aga atg cca tta gaa							1008
Lys Tyr Arg Tyr Phe Ser Asn Ala Ser Pro His Arg Met Pro Leu Glu							
	325			330		335	
ttc aaa aca tta aat gtg cta cat aat aga gga gct cta aaa ctg aca							1056
Phe Lys Thr Leu Asn Val Leu His Asn Arg Gly Ala Leu Lys Leu Thr							
	340			345		350	
aca gga aag tgt gta aag caa taa							1080
Thr Gly Lys Cys Val Lys Gln							
	355						
<210>	38						
<211>	359						
<212>	PRT						
<213>	Homo sapiens						
<400>	38						
Met Arg Ser Ile Arg Lys Arg Trp Thr Ile Cys Thr Ile Ser Leu Leu							
1		5			10		15
Leu Ile Phe Tyr Lys Thr Lys Glu Ile Ala Arg Thr Glu Glu His Gln							
	20			25		30	
Glu Thr Gln Leu Ile Gly Asp Gly Glu Leu Ser Leu Ser Arg Ser Leu							
	35			40		45	
Val Asn Ser Ser Asp Lys Ile Ile Arg Lys Ala Gly Ser Ser Ile Phe							
	50			55		60	
Gln His Asn Val Glu Gly Trp Lys Ile Asn Ser Ser Leu Val Leu Glu							
65		70		75		80	
Ile Arg Lys Asn Ile Leu Arg Phe Leu Asp Ala Glu Arg Asp Val Ser							
	85			90		95	
Val Val Lys Ser Ser Phe Lys Pro Gly Asp Val Ile His Tyr Val Leu							
	100			105		110	
Asp Arg Arg Arg Thr Leu Asn Ile Ser His Asp Leu His Ser Leu Leu							
	115			120		125	
Pro Glu Val Ser Pro Met Lys Asn Arg Arg Phe Lys Thr Cys Ala Val							
	130			135		140	
Val Gly Asn Ser Gly Ile Leu Leu Asp Ser Glu Cys Gly Lys Glu Ile							
145		150		155		160	

ES 2 368 267 T3

Asp Ser His Asn Phe Val Ile Arg Cys Asn Leu Ala Pro Val Val Glu
 165 170 175

Phe Ala Ala Asp Val Gly Thr Lys Ser Asp Phe Ile Thr Met Asn Pro
 180 185 190

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

Glu Lys Phe Val His Arg Leu Ser Met Leu Asn Asp Ser Val Leu Trp
 210 215 220

Ile Pro Ala Phe Met Val Lys Gly Gly Glu Lys His Val Glu Trp Val
 225 230 235 240

Asn Ala Leu Ile Leu Lys Asn Lys Leu Lys Val Arg Thr Ala Tyr Pro
 245 250 255

Ser Leu Arg Leu Ile His Ala Val Arg Gly Tyr Trp Leu Thr Asn Lys
 260 265 270

Val Pro Ile Lys Arg Pro Ser Thr Gly Leu Leu Met Tyr Thr Leu Ala
 275 280 285

Thr Arg Phe Cys Asp Glu Ile His Leu Tyr Gly Phe Trp Pro Phe Pro
 290 295 300

Lys Asp Leu Asn Gly Lys Ala Val Lys Tyr His Tyr Tyr Asp Asp Leu
 305 310 315 320

Lys Tyr Arg Tyr Phe Ser Asn Ala Ser Pro His Arg Met Pro Leu Glu
 325 330 335

Phe Lys Thr Leu Asn Val Leu His Asn Arg Gly Ala Leu Lys Leu Thr
 340 345 350

Thr Gly Lys Cys Val Lys Gln
 355

<210> 39
 <211> 1131
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS

ES 2 368 267 T3

<222> (1)..(1131)

<400> 39

atg cgc tac gcg gac ccc tcg gcc aac cgg gat ttg ttg ggg agc cga	48
Met Arg Tyr Ala Asp Pro Ser Ala Asn Arg Asp Leu Leu Gly Ser Arg	
1 5 10 15	
act ttg ctc ttc atc ttc atc tgc gcc ttt gcc ttg gtg acc ttg ctg	96
Thr Leu Leu Phe Ile Phe Ile Cys Ala Phe Ala Leu Val Thr Leu Leu	
20 25 30	
caa cag atc ctg tat ggc agg aac tac att aag agg tac ttt gaa ttt	144
Gln Gln Ile Leu Tyr Gly Arg Asn Tyr Ile Lys Arg Tyr Phe Glu Phe	
35 40 45	
tat gag ggg cct ttt gaa tat aac tcc aca aga tgc ctg gag ctg agg	192
Tyr Glu Gly Pro Phe Glu Tyr Asn Ser Thr Arg Cys Leu Glu Leu Arg	
50 55 60	
cac gaa ata ttg gaa gtg aag gtg ctg tcc atg gtg aag cag tca gag	240
His Glu Ile Leu Glu Val Lys Val Leu Ser Met Val Lys Gln Ser Glu	
65 70 75 80	
ctg ttc gac agg tgg aag agc ctc cag atg tgc aaa tgg gcg atg aac	288
Leu Phe Asp Arg Trp Lys Ser Leu Gln Met Cys Lys Trp Ala Met Asn	
85 90 95	
atc tct gag gcc aac cag ttc aag tct act ctg tcc agg tgc tgc aac	336
Ile Ser Glu Ala Asn Gln Phe Lys Ser Thr Leu Ser Arg Cys Cys Asn	
100 105 110	
gcc cct gcc ttt ctc ttc acc acc cag aag aac act ccc ctg ggg aca	384
Ala Pro Ala Phe Leu Phe Thr Thr Gln Lys Asn Thr Pro Leu Gly Thr	
115 120 125	
aag ctc aag tat gag gtg gac acc agt ggc atc tac cac atc aac cag	432
Lys Leu Lys Tyr Glu Val Asp Thr Ser Gly Ile Tyr His Ile Asn Gln	
130 135 140	
gag atc ttc cgc atg ttt ccc aag gac atg ccc tac tac cgg tcc cag	480
Glu Ile Phe Arg Met Phe Pro Lys Asp Met Pro Tyr Tyr Arg Ser Gln	
145 150 155 160	
ttt aag aag tgt gct gta gtg ggc aac gga ggc atc ttg aag aac agc	528
Phe Lys Lys Cys Ala Val Val Gly Asn Gly Gly Ile Leu Lys Asn Ser	
165 170 175	
cgc tgc ggg agg gag atc aac agc gcc gac ttc gtc ttc cgg tgc aac	576
Arg Cys Gly Arg Glu Ile Asn Ser Ala Asp Phe Val Phe Arg Cys Asn	
180 185 190	
ctg ccc ccc atc tca gag aag tac acc atg gat gtg ggg gtg aag acg	624
Leu Pro Pro Ile Ser Glu Lys Tyr Thr Met Asp Val Gly Val Lys Thr	
195 200 205	
gat gtg gtc act gtg aac ccc agc atc atc aca gag agg ttc cac aag	672
Asp Val Val Thr Val Asn Pro Ser Ile Ile Thr Glu Arg Phe His Lys	
210 215 220	
ctg gag aag tgg cgg cgg ccg ttc tat cgc gtg ctg cag gtg tac gag	720
Leu Glu Lys Trp Arg Arg Pro Phe Tyr Arg Val Leu Gln Val Tyr Glu	

ES 2 368 267 T3

225	230	235	240	
aac gcg tcg gtg ctg ctg cct gcc ttc tac aac acg cgc aac acc gac				768
Asn Ala Ser Val Leu Leu Pro Ala Phe Tyr Asn Thr Arg Asn Thr Asp				
245	250	255		
gtg tcc atc cgc gtc aag tac gtg ctg gac gac ttc gaa tcg ccg caa				816
Val Ser Ile Arg Val Lys Tyr Val Leu Asp Asp Phe Glu Ser Pro Gln				
260	265	270		
gct gtc tac tac ttc cat ccg cag tac ctg gtc aac gtg tcg cgc tac				864
Ala Val Tyr Tyr Phe His Pro Gln Tyr Leu Val Asn Val Ser Arg Tyr				
275	280	285		
tgg ctc agc ctg ggg gtg cgc gcc aag cgc atc agc acc ggc ctc att				912
Trp Leu Ser Leu Gly Val Arg Ala Lys Arg Ile Ser Thr Gly Leu Ile				
290	295	300		
ctg gtc act gcg gcg ctg gag ctc tgt gag gag gtg cac ctc ttt ggc				960
Leu Val Thr Ala Ala Leu Glu Leu Cys Glu Glu Val His Leu Phe Gly				
305	310	315	320	
ttc tgg gcc ttc ccc atg aac ccc tcg ggc ctc tac atc act cac cac				1008
Phe Trp Ala Phe Pro Met Asn Pro Ser Gly Leu Tyr Ile Thr His His				
325	330	335		
tac tat gac aac gtc aag ccg cgt ccc ggc ttc cac gcc atg ccc tct				1056
Tyr Tyr Asp Asn Val Lys Pro Arg Pro Gly Phe His Ala Met Pro Ser				
340	345	350		
gag atc ttc aac ttc ctg cac ttg cac agc cga ggc atc ctc cgc gtg				1104
Glu Ile Phe Asn Phe Leu His Leu His Ser Arg Gly Ile Leu Arg Val				
355	360	365		
cac acg ggc acc tgc agc tgc tgc tga				1131
His Thr Gly Thr Cys Ser Cys Cys				
370	375			
<210> 40				
<211> 376				
<212> PRT				
<213> Homo sapiens				
<400> 40				
Met Arg Tyr Ala Asp Pro Ser Ala Asn Arg Asp Leu Leu Gly Ser Arg				
1	5	10	15	
Thr Leu Leu Phe Ile Phe Ile Cys Ala Phe Ala Leu Val Thr Leu Leu				
20	25	30		
Gln Gln Ile Leu Tyr Gly Arg Asn Tyr Ile Lys Arg Tyr Phe Glu Phe				
35	40	45		
Tyr Glu Gly Pro Phe Glu Tyr Asn Ser Thr Arg Cys Leu Glu Leu Arg				
50	55	60		

ES 2 368 267 T3

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

ES 2 368 267 T3

Leu Val Thr Ala Ala Leu Glu Leu Cys Glu Glu Val His Leu Phe Gly
305 310 315 320

Phe Trp Ala Phe Pro Met Asn Pro Ser Gly Leu Tyr Ile Thr His His
325 330 335

Tyr Tyr Asp Asn Val Lys Pro Arg Pro Gly Phe His Ala Met Pro Ser
340 345 350

Glu Ile Phe Asn Phe Leu His Leu His Ser Arg Gly Ile Leu Arg Val
355 360 365

His Thr Gly Thr Cys Ser Cys Cys
370 375

<210> 41
<211> 1197
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1197)

<400> 41
atg cgg ccg ggg ggc gca ctg ctc gcc ctg ctc gcc agc ctg ctg ctg 48
Met Arg Pro Gly Gly Ala Leu Leu Ala Leu Leu Ala Ser Leu Leu Leu
1 5 10 15

ctg ctg ctg ctg cgc ctg ctc tgg tgc ccg gca gac gcg ccc ggc cgc 96
Leu Leu Leu Leu Arg Leu Leu Trp Cys Pro Ala Asp Ala Pro Gly Arg
20 25 30

gcc agg att ctg gtg gag gaa agc agg gag gcc acc cac ggc acc ccc 144
Ala Arg Ile Leu Val Glu Glu Ser Arg Glu Ala Thr His Gly Thr Pro
35 40 45

gca gcg ctg agg acg ctc cgg agc ccg gcg acc gcg gta ccg cgc gcc 192
Ala Ala Leu Arg Thr Leu Arg Ser Pro Ala Thr Ala Val Pro Arg Ala
50 55 60

act aac agc aca tat ctg aat gag aag tcg ctc caa ctg acg gag aaa 240
Thr Asn Ser Thr Tyr Leu Asn Glu Lys Ser Leu Gln Leu Thr Glu Lys
65 70 75 80

tgc aaa aat ctg caa tat ggc att gag tct ttc tct aac aaa acg aaa 288
Cys Lys Asn Leu Gln Tyr Gly Ile Glu Ser Phe Ser Asn Lys Thr Lys
85 90 95

ggg tat tca gag aac gac tac ctt cag att atc aca gat ata cag agt 336
Gly Tyr Ser Glu Asn Asp Tyr Leu Gln Ile Ile Thr Asp Ile Gln Ser
100 105 110

tgt cca tgg aaa cgg caa gca gaa gaa tat gca aat ttt aga gcc aaa 384
Cys Pro Trp Lys Arg Gln Ala Glu Glu Tyr Ala Asn Phe Arg Ala Lys

ES 2 368 267 T3

115	120	125	
ctt gct tcc tgc tgt gat gct gtt caa aac ttt gtt gtt tct cag aat Leu Ala Ser Cys Cys Asp Ala Val Gln Asn Phe Val Val Ser Gln Asn 130 135 140			432
aac act cca gtt ggg act aat atg agt tac gag gtg gaa agc aaa aaa Asn Thr Pro Val Gly Thr Asn Met Ser Tyr Glu Val Glu Ser Lys Lys 145 150 155 160			480
gaa atc cca att aag aag aac att ttt cat atg ttt cca gtg tcc cag Glu Ile Pro Ile Lys Lys Asn Ile Phe His Met Phe Pro Val Ser Gln 165 170 175			528
cct ttt gtg gac tac cct tat aat cag tgt gca gtg gtc gga aat ggg Pro Phe Val Asp Tyr Pro Tyr Asn Gln Cys Ala Val Val Gly Asn Gly 180 185 190			576
gga att ctg aat aag tct ctc tgt gga act gaa ata gat aaa tcc gac Gly Ile Leu Asn Lys Ser Leu Cys Gly Thr Glu Ile Asp Lys Ser Asp 195 200 205			624
ttc gtt ttt agg tgt aac cta ccc cca acc aca gga gat gtt agt aaa Phe Val Phe Arg Cys Asn Leu Pro Pro Thr Thr Gly Asp Val Ser Lys 210 215 220			672
gat gtt ggc agt aaa aca aat ctt gtg act ata aat cca agc atc ata Asp Val Gly Ser Lys Thr Asn Leu Val Thr Ile Asn Pro Ser Ile Ile 225 230 235 240			720
act ctg aaa tat ggg aac tta aag gaa aaa aaa gcc cta ttt ctg gag Thr Leu Lys Tyr Gly Asn Leu Lys Glu Lys Lys Ala Leu Phe Leu Glu 245 250 255			768
gac att gca acc tat gga gat gca ttt ttt ctt ctg cca gca ttt tcc Asp Ile Ala Thr Tyr Gly Asp Ala Phe Phe Leu Leu Pro Ala Phe Ser 260 265 270			816
ttc agg gcc aac acg ggt acc tct ttc aaa gta tac tac acg ctc gaa Phe Arg Ala Asn Thr Gly Thr Ser Phe Lys Val Tyr Tyr Thr Leu Glu 275 280 285			864
gag tct aaa gca aga caa aag gtt cta ttt ttc cat ccc aag tac ctg Glu Ser Lys Ala Arg Gln Lys Val Leu Phe Phe His Pro Lys Tyr Leu 290 295 300			912
aaa gat ctg gcc ctt ttc tgg aga act aaa ggt gtg act gca tac cgc Lys Asp Leu Ala Leu Phe Trp Arg Thr Lys Gly Val Thr Ala Tyr Arg 305 310 315 320			960
ttg tcc acc ggc ttg atg atc aca agt gtt gca gtg gaa ctg tgt aaa Leu Ser Thr Gly Leu Met Ile Thr Ser Val Ala Val Glu Leu Cys Lys 325 330 335			1008
aat gtg aag ctg tat gga ttc tgg ccc ttc tct aaa act gta gaa gac Asn Val Lys Leu Tyr Gly Phe Trp Pro Phe Ser Lys Thr Val Glu Asp 340 345 350			1056
ata cct gtc agc cat cac tat tat gac aac aag cta cct aaa cat ggt Ile Pro Val Ser His His Tyr Tyr Asp Asn Lys Leu Pro Lys His Gly 355 360 365			1104

ES 2 368 267 T3

ttc	cat	cag	atg	ccc	aaa	gaa	tac	agc	cag	atc	ctc	caa	ctt	cac	atg	1152
Phe	His	Gln	Met	Pro	Lys	Glu	Tyr	Ser	Gln	Ile	Leu	Gln	Leu	His	Met	
	370					375					380					

aaa	gga	atc	ctc	aaa	ctg	caa	ttt	agc	aaa	tgt	gaa	gtc	gcc	taa	1197	
Lys	Gly	Ile	Leu	Lys	Leu	Gln	Phe	Ser	Lys	Cys	Glu	Val	Ala			
385					390					395						

<210> 42
 <211> 398
 <212> PRT
 <213> Homo sapiens

<400> 42

Met	Arg	Pro	Gly	Gly	Ala	Leu	Leu	Ala	Leu	Leu	Ala	Ser	Leu	Leu	Leu
1				5					10					15	

Leu	Leu	Leu	Leu	Arg	Leu	Leu	Trp	Cys	Pro	Ala	Asp	Ala	Pro	Gly	Arg
			20					25					30		

Ala	Arg	Ile	Leu	Val	Glu	Glu	Ser	Arg	Glu	Ala	Thr	His	Gly	Thr	Pro
		35					40					45			

Ala	Ala	Leu	Arg	Thr	Leu	Arg	Ser	Pro	Ala	Thr	Ala	Val	Pro	Arg	Ala
50						55					60				

Thr	Asn	Ser	Thr	Tyr	Leu	Asn	Glu	Lys	Ser	Leu	Gln	Leu	Thr	Glu	Lys
65					70					75					80

Cys	Lys	Asn	Leu	Gln	Tyr	Gly	Ile	Glu	Ser	Phe	Ser	Asn	Lys	Thr	Lys
			85						90					95	

Gly	Tyr	Ser	Glu	Asn	Asp	Tyr	Leu	Gln	Ile	Ile	Thr	Asp	Ile	Gln	Ser
			100					105					110		

Cys	Pro	Trp	Lys	Arg	Gln	Ala	Glu	Glu	Tyr	Ala	Asn	Phe	Arg	Ala	Lys
		115					120					125			

Leu	Ala	Ser	Cys	Cys	Asp	Ala	Val	Gln	Asn	Phe	Val	Val	Ser	Gln	Asn
130						135					140				

Asn	Thr	Pro	Val	Gly	Thr	Asn	Met	Ser	Tyr	Glu	Val	Glu	Ser	Lys	Lys
145					150					155					160

Glu	Ile	Pro	Ile	Lys	Lys	Asn	Ile	Phe	His	Met	Phe	Pro	Val	Ser	Gln
				165				170						175	

ES 2 368 267 T3

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

<210> 43
 <211> 951
 <212> ADN
 <213> Secuencia Artificial

ES 2 368 267 T3

<220>

<223> Nucleótidos 268-1218 de SEC ID NO : 1

<220>

<221> CDS

<222> (1)..(951)

<400> 43

gag gcc tcc ttc cag gtg tgg aac aag gac agc tct tcc aaa aac ctt	48
Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu	
1 5 10 15	

atc cct agg ctg caa aag atc tgg aag aat tac cta agc atg aac aag	96
Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys	
20 25 30	

tac aaa gtg tcc tac aag ggg cca gga cca ggc atc aag ttc agt gca	144
Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala	
35 40 45	

gag gcc ctg cgc tgc cac ctc cgg gac cat gtg aat gta tcc atg gta	192
Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met Val	
50 55 60	

gag gtc aca gat ttt ccc ttc aat acc tct gaa tgg gag ggt tat ctg	240
Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr Leu	
65 70 75 80	

ccc aag gag agc att agg acc aag gct ggg cct tgg ggc agg tgt gct	288
Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys Ala	
85 90 95	

gtt gtg tcg tca gcg gga tct ctg aag tcc tcc caa cta ggc aga gaa	336
Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg Glu	
100 105 110	

atc gat gat cat gac gca gtc ctg agg ttt aat ggg gca ccc aca gcc	384
Ile Asp Asp His Asp Ala Val Leu Arg Phe Asn Gly Ala Pro Thr Ala	
115 120 125	

aac ttc caa caa gat gtg ggc aca aaa act acc att cgc ctg atg aac	432
Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met Asn	
130 135 140	

tct cag ttg gtt acc aca gag aag cgc ttc ctc aaa gac agt ttg tac	480
Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu Tyr	
145 150 155 160	

aat gaa gga atc cta att gta tgg gac cca tct gta tac cac tca gat	528
Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser Asp	
165 170 175	

atc cca aag tgg tac cag aat ccg gat tat aat ttc ttt aac aac tac	576
Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn Tyr	
180 185 190	

aag act tat cgt aag ctg cac ccc aat cag ccc ttt tac atc ctc aag	624
Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu Lys	
195 200 205	

ES 2 368 267 T3

ccc cag atg cct tgg gag cta tgg gac att ctt caa gaa atc tcc cca	672
Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser Pro	
210 215 220	
gaa gag att cag cca aac ccc cca tcc tct ggg atg ctt ggt atc atc	720
Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile Ile	
225 230 235 240	
atc atg atg acg ctg tgt gac cag gtg gat att tat gag ttc ctc cca	768
Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu Pro	
245 250 255	
tcc aag cgc aag act gac gtg tgc tac tac tac cag aag ttc ttc gat	816
Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp	
260 265 270	
agt gcc tgc acg atg ggt gcc tac cac ccg ctg ctc tat gag aag aat	864
Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys Asn	
275 280 285	
ttg gtg aag cat ctc aac cag ggc aca gat gag gac atc tac ctg ctt	912
Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu Leu	
290 295 300	
gga aaa gcc aca ctg cct ggc ttc cgg acc att cac tgc	951
Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys	
305 310 315	

<210> 44
 <211> 317
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Constructo sintético

<400> 44

Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu	
1 5 10 15	

Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys	
20 25 30	

Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala	
35 40 45	

Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met Val	
50 55 60	

Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr Leu	
65 70 75 80	

Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys Ala	
85 90 95	

ES 2 368 267 T3

Val	Val	Ser	Ser	Ala	Gly	Ser	Leu	Lys	Ser	Ser	Gln	Leu	Gly	Arg	Glu
			100					105					110		
Ile	Asp	Asp	His	Asp	Ala	Val	Leu	Arg	Phe	Asn	Gly	Ala	Pro	Thr	Ala
		115					120					125			
Asn	Phe	Gln	Gln	Asp	Val	Gly	Thr	Lys	Thr	Thr	Ile	Arg	Leu	Met	Asn
	130					135					140				
Ser	Gln	Leu	Val	Thr	Thr	Glu	Lys	Arg	Phe	Leu	Lys	Asp	Ser	Leu	Tyr
145					150					155					160
Asn	Glu	Gly	Ile	Leu	Ile	Val	Trp	Asp	Pro	Ser	Val	Tyr	His	Ser	Asp
				165					170					175	
Ile	Pro	Lys	Trp	Tyr	Gln	Asn	Pro	Asp	Tyr	Asn	Phe	Phe	Asn	Asn	Tyr
			180					185					190		
Lys	Thr	Tyr	Arg	Lys	Leu	His	Pro	Asn	Gln	Pro	Phe	Tyr	Ile	Leu	Lys
		195					200					205			
Pro	Gln	Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Leu	Gln	Glu	Ile	Ser	Pro
	210					215					220				
Glu	Glu	Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	Met	Leu	Gly	Ile	Ile
225					230					235					240
Ile	Met	Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	Tyr	Glu	Phe	Leu	Pro
				245					250					255	
Ser	Lys	Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	Tyr	Gln	Lys	Phe	Phe	Asp
			260					265					270		
Ser	Ala	Cys	Thr	Met	Gly	Ala	Tyr	His	Pro	Leu	Leu	Tyr	Glu	Lys	Asn
		275					280					285			
Leu	Val	Lys	His	Leu	Asn	Gln	Gly	Thr	Asp	Glu	Asp	Ile	Tyr	Leu	Leu
	290					295					300				
Gly	Lys	Ala	Thr	Leu	Pro	Gly	Phe	Arg	Thr	Ile	His	Cys			
305					310					315					

<210> 45
 <211> 27
 <212> ADN

ES 2 368 267 T3

<213> Secuencia artificial

<220>

<223> Nucleótidos 1-27 de SEC ID n°: 1

<220>

<221> CDS

<222> (1)..(27)

<400> 45

atg att cac acc aac ctg aag aaa aag

Met Ile His Thr Asn Leu Lys Lys Lys

1 5

27

<210> 46

<211> 9

<212> PRT

<213> Secuencia artificial

<220>

<223> Constructo sintético

<400> 46

Met Ile His Thr Asn Leu Lys Lys Lys

1 5

<210> 47

<211> 30

<212> ADN

<213> Secuencia artificial

<220>

<223> Nucleótidos 1-30 de SEC ID n°: 3

<220>

<221> CDS

<222> (1)..(30)

<400> 47

atg aaa cca cac ttg aag caa tgg aga caa

Met Lys Pro His Leu Lys Gln Trp Arg Gln

1 5 10

30

<210> 48

<211> 10

<212> PRT

<213> Secuencia artificial

<220>

<223> Constructo sintético

<400> 48

Met Lys Pro His Leu Lys Gln Trp Arg Gln

1 5 10

ES 2 368 267 T3

<210> 49
 <211> 42
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Nucleótidos 1-42 de SEC ID n°: 5

<220>
 <221> CDS
 <222> (1)..(42)

<400> 49
 atg agg tcc tgc ctg tgg aga tgc agg cac ctg agc caa ggc 42
 Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly
 1 5 10

<210> 50
 <211> 14
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Constructo sintético

<400> 50
 Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly
 1 5 10

<210> 51
 <211> 21
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Nucleótidos 1-21 de SEC ID n°: 7

<220>
 <221> CDS
 <222> (1)..(21)

<400> 51
 atg ggg ctc ccg cgc ggg tcg 21
 Met Gly Leu Pro Arg Gly Ser
 1 5

<210> 52
 <211> 7
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Constructo sintético

ES 2 368 267 T3

<400> 52

Met Gly Leu Pro Arg Gly Ser
1 5

<210> 53

<211> 24

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 1-24 de SEC ID n°: 9

<220>

<221> CDS

<222> (1)..(24)

<400> 53

atg gcc tgc atc ctg aag aga aag
Met Ala Cys Ile Leu Lys Arg Lys
1 5

24

<210> 54

<211> 8

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 54

Met Ala Cys Ile Leu Lys Arg Lys
1 5

<210> 55

<211> 18

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 1-18 de SEC ID n°: 11

<220>

<221> CDS

<222> (1)..(18)

<400> 55

atg aag gct ccg ggt cgg
Met Lys Ala Pro Gly Arg
1 5

18

<210> 56

<211> 6

ES 2 368 267 T3

<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 56

Met Lys Ala Pro Gly Arg
1 5

<210> 57
<211> 24
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-24 de SEC ID nº: 13

<220>
<221> CDS
<222> (1)..(21)

<400> 57
atg aag acc ctg atg cgc cat ggt
Met Lys Thr Leu Met Arg His
1 5

24

<210> 58
<211> 7
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 58

Met Lys Thr Leu Met Arg His
1 5

<210> 59
<211> 129
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-129 de SEC ID nº: 15

<220>
<221> CDS
<222> (1)..(129)

<400> 59
atg gct tgc tcg agg ccc ccc agc cag tgt gaa ccc aca tcc ctg ccc
Met Ala Cys Ser Arg Pro Pro Ser Gln Cys Glu Pro Thr Ser Leu Pro

48

ES 2 368 267 T3

1	5	10	15	
cca ggg cca cct gca gga cgc cga cac cta ccc ctc agc aga cgc cgg				96
Pro Gly Pro Pro Ala Gly Arg Arg His Leu Pro Leu Ser Arg Arg Arg				
	20	25	30	

aga gaa atg agt agc aac aaa gag cag cgg tca				129
Arg Glu Met Ser Ser Asn Lys Glu Gln Arg Ser				
	35	40		

<210> 60
 <211> 43
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 60

Met Ala Cys Ser Arg Pro Pro Ser Gln Cys Glu Pro Thr Ser Leu Pro			
1	5	10	15

Pro Gly Pro Pro Ala Gly Arg Arg His Leu Pro Leu Ser Arg Arg Arg			
	20	25	30

Arg Glu Met Ser Ser Asn Lys Glu Gln Arg Ser	
	35 40

<210> 61
 <211> 39
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-39 de SEC ID n°: 17

<220>
 <221> CDS
 <222> (1)..(30)

<400> 61				
atg gtg acc ctg cgg aag agg acc ctg aaa gtgctcacc				39
Met Val Thr Leu Arg Lys Arg Thr Leu Lys				
1	5	10		

<210> 62
 <211> 10
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 62

ES 2 368 267 T3

Met Val Thr Leu Arg Lys Arg Thr Leu Lys
1 5 10

<210> 63
<211> 18
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-18 de SEC ID nº: 19

<220>
<221> CDS
<222> (1)..(18)

<400> 63
atg aag tgc tcc ctg cgg 18
Met Lys Cys Ser Leu Arg
1 5

<210> 64
<211> 6
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 64

Met Lys Cys Ser Leu Arg
1 5

<210> 65
<211> 24
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-24 de SEC ID nº: 21

<220>
<221> CDS
<222> (1)..(24)

<400> 65
atg gga ctc ttg gta ttt gtg cgc 24
Met Gly Leu Leu Val Phe Val Arg
1 5

<210> 66
<211> 8
<212> PRT
<213> Artificial

ES 2 368 267 T3

<220>
 <223> Constructo sintético
 <400> 66
 Met Gly Leu Leu Val Phe Val Arg
 1 5

<210> 67
 <211> 24
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-24 de SEC ID nº: 23

<220>
 <221> CDS
 <222> (1)..(24)

<400> 67
 atg gtc agc aag tcc cgc tgg aag
 Met Val Ser Lys Ser Arg Trp Lys
 1 5

24

<210> 68
 <211> 8
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 68

Met Val Ser Lys Ser Arg Trp Lys
 1 5

<210> 69
 <211> 15
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-15 de SEC ID nº: 25

<220>
 <221> CDS
 <222> (1)..(15)

<400> 69
 atg aga agg ccc agc
 Met Arg Arg Pro Ser
 1 5

15

ES 2 368 267 T3

<210> 70
 <211> 5
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 70

Met Arg Arg Pro Ser
 1 5

<210> 71
 <211> 12
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-12 de SEC ID n°: 27

<220>
 <221> CDS
 <222> (1)..(12)

<400> 71
 atg aga ggg tat
 Met Arg Gly Tyr
 1

12

<210> 72
 <211> 4
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 72

Met Arg Gly Tyr
 1

<210> 73
 <211> 24
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-24 de SEC ID n°: 29

<220>
 <221> CDS
 <222> (1)..(24)

ES 2 368 267 T3

<400> 73
 atg gga ctc ttg gta ttt gta cgc
 Met Gly Leu Leu Val Phe Val Arg
 1 5

<210> 74
 <211> 8
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 74
 Met Gly Leu Leu Val Phe Val Arg
 1 5

<210> 75
 <211> 87
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-87 de SEC ID n°: 31

<220>
 <221> CDS
 <222> (1)..(87)

<400> 75
 atg agc ccc tgc ggg cgg gcc cgg cga caa acg tcc aga ggg gcc atg
 Met Ser Pro Cys Gly Arg Ala Arg Arg Gln Thr Ser Arg Gly Ala Met
 1 5 10 15

gct gta ctg gcg tgg aag ttc ccg cgg acc cgg ctg ccc
 Ala Val Leu Ala Trp Lys Phe Pro Arg Thr Arg Leu Pro
 20 25

<210> 76
 <211> 29
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 76
 Met Ser Pro Cys Gly Arg Ala Arg Arg Gln Thr Ser Arg Gly Ala Met
 1 5 10 15

Ala Val Leu Ala Trp Lys Phe Pro Arg Thr Arg Leu Pro
 20 25

ES 2 368 267 T3

<210> 77
 <211> 18
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-18 de SEC ID nº: 33

<220>
 <221> CDS
 <222> (1)..(18)

<400> 77
 atg cag ctg cag ttc cgg
 Met Gln Leu Gln Phe Arg
 1 5

18

<210> 78
 <211> 6
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 78
 Met Gln Leu Gln Phe Arg
 1 5

<210> 79
 <211> 27
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-27 de SEC ID nº: 35

<220>
 <221> CDS
 <222> (1)..(27)

<400> 79
 atg aga aac tgc aaa atg gcc cgg gtc
 Met Arg Asn Cys Lys Met Ala Arg Val
 1 5

27

<210> 80
 <211> 9
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 80

ES 2 368 267 T3

Met Arg Asn Cys Lys Met Ala Arg Val
1 5

<210> 81
<211> 21
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-21 de SEC ID nº: 37

<220>
<221> CDS
<222> (1)..(21)

<400> 81
atg cgc tcc att agg aag agg
Met Arg Ser Ile Arg Lys Arg
1 5

21

<210> 82
<211> 7
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 82

Met Arg Ser Ile Arg Lys Arg
1 5

<210> 83
<211> 51
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-51 of SEQ ID NO : 39

<220>
<221> CDS
<222> (1)..(51)

<400> 83
atg cgc tac gcg gac ccc tcg gcc aac cgg gat ttg ttg ggg agc cga
Met Arg Tyr Ala Asp Pro Ser Ala Asn Arg Asp Leu Leu Gly Ser Arg
1 5 10 15

48

act
Thr

51

ES 2 368 267 T3

<210> 84
 <211> 17
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 84

Met Arg Tyr Ala Asp Pro Ser Ala Asn Arg Asp Leu Leu Gly Ser Arg
 1 5 10 15

Thr

<210> 85
 <211> 9
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-9 de SEC ID n°: 41

<220>
 <221> CDS
 <222> (1)..(9)

<400> 85
 atg cgg ccg
 Met Arg Pro
 1

9

<210> 86
 <211> 3
 <212> PRT
 <213> Artificial

<220>
 <223> Aminoácidos 1-3 de SEC ID n°: 42

<400> 86

Met Arg Pro
 1

<210> 87
 <211> 51
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 28-78 de SEC ID n°: 1

<220>

ES 2 368 267 T3

<221> CDS
<222> (1)..(51)

<400> 87
ttc agc tgc tgc gtc ctg gtc ttt ctt ctg ttt gca gtc atc tgt gtg 48
Phe Ser Cys Cys Val Leu Val Phe Leu Leu Phe Ala Val Ile Cys Val
1 5 10 15

tgg 51
Trp

<210> 88
<211> 17
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 88
Phe Ser Cys Cys Val Leu Val Phe Leu Leu Phe Ala Val Ile Cys Val
1 5 10 15

Trp

<210> 89
<211> 60
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 31-90 de SEC ID n°: 3

<220>
<221> CDS
<222> (1)..(60)

<400> 89
cga atg ctt ttc gga ata ttc gct tgg ggg ctc ctc ttt ttg ctg att 48
Arg Met Leu Phe Gly Ile Phe Ala Trp Gly Leu Leu Phe Leu Leu Ile
1 5 10 15

ttc atc tac ttc 60
Phe Ile Tyr Phe
20

<210> 90
<211> 20
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

ES 2 368 267 T3

<221> CDS
<222> (1)..(51)

<400> 87
ttc agc tgc tgc gtc ctg gtc ttt ctt ctg ttt gca gtc atc tgt gtg 48
Phe Ser Cys Cys Val Leu Val Phe Leu Leu Phe Ala Val Ile Cys Val
1 5 10 15

tgg 51
Trp

<210> 88
<211> 17
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 88
Phe Ser Cys Cys Val Leu Val Phe Leu Leu Phe Ala Val Ile Cys Val
1 5 10 15

Trp

<210> 89
<211> 60
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 31-90 de SEC ID n°: 3

<220>
<221> CDS
<222> (1)..(60)

<400> 89
cga atg ctt ttc gga ata ttc gct tgg ggg ctc ctc ttt ttg ctg att 48
Arg Met Leu Phe Gly Ile Phe Ala Trp Gly Leu Leu Phe Leu Leu Ile
1 5 10 15

ttc atc tac ttc 60
Phe Ile Tyr Phe
20

<210> 90
<211> 20
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

ES 2 368 267 T3

```

<220>
<221> CDS
<222> (1)..(63)

<400> 93
ttc ttc tgg gtg ctg ctc ctg ctc acg gct gcc tgc tcg ggg ctc ctc      48
Phe Phe Trp Val Leu Leu Leu Leu Thr Ala Ala Cys Ser Gly Leu Leu
1          5          10          15

ttt gcc ctg tac ttc      63
Phe Ala Leu Tyr Phe
          20

<210> 94
<211> 21
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 94

Phe Phe Trp Val Leu Leu Leu Leu Thr Ala Ala Cys Ser Gly Leu Leu
1          5          10          15

Phe Ala Leu Tyr Phe
          20

<210> 95
<211> 60
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 25-84 de SEC ID n°: 9

<220>
<221> CDS
<222> (1)..(60)

<400> 95
tct gtg att gct gtg agc ttc ata gca gcg ttc ctt ttc ctg ctg gtt      48
Ser Val Ile Ala Val Ser Phe Ile Ala Ala Phe Leu Phe Leu Leu Val
1          5          10          15

gtg cgt ctt gta      60
Val Arg Leu Val
          20

<210> 96
<211> 20
<212> PRT
<213> Artificial

<220>

```

ES 2 368 267 T3

<223> Constructo sintético

<400> 96

Ser Val Ile Ala Val Ser Phe Ile Ala Ala Phe Leu Phe Leu Leu Val
1 5 10 15

Val Arg Leu Val
20

<210> 97

<211> 63

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 19-81 de SEC ID nº: 11

<220>

<221> CDS

<222> (1)..(63)

<400> 97

ctc gtg ctc atc atc ctg tgc tcc gtg gtc ttc tct gcc gtc tac atc 48
Leu Val Leu Ile Ile Leu Cys Ser Val Val Phe Ser Ala Val Tyr Ile
1 5 10 15

ctc ctg tgc tgc tgg 63
Leu Leu Cys Cys Trp
20

<210> 98

<211> 21

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 98

Leu Val Leu Ile Ile Leu Cys Ser Val Val Phe Ser Ala Val Tyr Ile
1 5 10 15

Leu Leu Cys Cys Trp
20

<210> 99

<211> 63

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 25-87 de SEC ID nº: 13

ES 2 368 267 T3

```

<220>
<221> CDS
<222> (1)..(63)

<400> 99
ctg gca gtg tgt tta gcg ctc acc acc atg tgc acc agc ttg ttg cta      48
Leu Ala Val Cys Leu Ala Leu Thr Thr Met Cys Thr Ser Leu Leu Leu
1          5          10          15

gtg tac agc agc ctc      63
Val Tyr Ser Ser Leu
          20

<210> 100
<211> 21
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 100

Leu Ala Val Cys Leu Ala Leu Thr Thr Met Cys Thr Ser Leu Leu Leu
1          5          10          15

Val Tyr Ser Ser Leu
          20

<210> 101
<211> 48
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 130-177 de SEC ID nº: 15

<220>
<221> CDS
<222> (1)..(48)

<400> 101
gca gtg ttc gtg atc ctc ttt gcc ctc atc acc atc ctc atc ctc tac      48
Ala Val Phe Val Ile Leu Phe Ala Leu Ile Thr Ile Leu Ile Leu Tyr
1          5          10          15

<210> 102
<211> 16
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 102

```

ES 2 368 267 T3

Ala Val Phe Val Ile Leu Phe Ala Leu Ile Thr Ile Leu Ile Leu Tyr
 1 5 10 15

<210> 103
 <211> 63
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 40-102 de SEC ID nº: 17

<220>
 <221> CDS
 <222> (1)..(63)

<400> 103
 ttc ctc gtg ctc ttc atc ttc ctc acc tcc ttc ttc ctg aac tac tcc 48
 Phe Leu Val Leu Phe Ile Phe Leu Thr Ser Phe Phe Leu Asn Tyr Ser
 1 5 10 15

cac acc atg gtg gcc 63
 His Thr Met Val Ala
 20

<210> 104
 <211> 21
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 104

Phe Leu Val Leu Phe Ile Phe Leu Thr Ser Phe Phe Leu Asn Tyr Ser
 1 5 10 15

His Thr Met Val Ala
 20

<210> 105
 <211> 63
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 19-81 de SEC ID nº: 19

<220>
 <221> CDS
 <222> (1)..(63)

<400> 105
 gtg tgg ttc ctc tcc gtg gcc ttc ctg ctg gtg ttc atc atg tcc ctg 48

ES 2 368 267 T3

Val Trp Phe Leu Ser Val Ala Phe Leu Leu Val Phe Ile Met Ser Leu
1 5 10 15

ctc ttc acc tac tcg 63
Leu Phe Thr Tyr Ser
20

<210> 106
<211> 21
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 106

Val Trp Phe Leu Ser Val Ala Phe Leu Leu Val Phe Ile Met Ser Leu
1 5 10 15

Leu Phe Thr Tyr Ser
20

<210> 107
<211> 60
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 25-84 de SEC ID nº: 21

<220>
<221> CDS
<222> (1)..(60)

<400> 107
aat ctg ctg cta gcc ctc tgc ctc ttt ctg gta ctg gga ttt ttg tat 48
Asn Leu Leu Leu Ala Leu Cys Leu Phe Leu Val Leu Gly Phe Leu Tyr
1 5 10 15

tat tct gcg tgg 60
Tyr Ser Ala Trp
20

<210> 108
<211> 20
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 108

Asn Leu Leu Leu Ala Leu Cys Leu Phe Leu Val Leu Gly Phe Leu Tyr
1 5 10 15

ES 2 368 267 T3

Tyr Ser Ala Trp
20

<210> 109
<211> 54
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 25-78 de SEC ID nº: 23

<220>
<221> CDS
<222> (1)..(54)

<400> 109
ctc ctg gcc atg ttg gct ctg gtc ctg gtc gtc atg gtg tgg tat tcc 48
Leu Leu Ala Met Leu Ala Leu Val Leu Val Val Met Val Trp Tyr Ser
1 5 10 15

atc tcc 54
Ile Ser

<210> 110
<211> 18
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 110

Leu Leu Ala Met Leu Ala Leu Val Leu Val Val Met Val Trp Tyr Ser
1 5 10 15

Ile Ser

<210> 111
<211> 63
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 16-78 de SEC ID nº: 25

<220>
<221> CDS
<222> (1)..(63)

<400> 111

ES 2 368 267 T3

ttg	tta	tta	aaa	gac	atc	ctc	aaa	tgt	aca	ttg	ctt	gtg	ttt	gga	gtg	48
Leu	Leu	Leu	Lys	Asp	Ile	Leu	Lys	Cys	Thr	Leu	Leu	Val	Phe	Gly	Val	
1				5				10					15			

tgg atc ctt tat atc 63
Trp Ile Leu Tyr Ile
20

```
<210> 112
<211> 21
<212> PRT
<213> Artificial
```

<220>
<223> Constructo sintético

<400> 112

Leu Leu Leu Lys Asp Ile Leu Lys Cys Thr Leu Leu Val Phe Gly Val
1 5 10 15

Trp Ile Leu Tyr Ile
20

```
<210> 113
<211> 63
<212> ADN
<213> Artificial
```

<220>
<223> Nucleótidos 13-75 de SEC ID n°: 27

```
<220>
<221> CDS
<222> (1)..(63)
```

```
<400> 113
ctt gtg gcc ata ttc ctg agt gct gtc ttc ctc tat tat gta ctg cat      48
Leu Val Ala Ile Phe Leu Ser Ala Val Phe Leu Tyr Tyr Val Leu His
1          5          10          15
```

tgc ata tta tgg gga 63
Cys Ile Leu Trp Gly
20

```
<210> 114
<211> 21
<212> PRT
<213> Artificial
```

<220>
<223> Constructo sintético

<400> 114

Leu Val Ala Ile Phe Leu Ser Ala Val Phe Leu Tyr Tyr Val Leu His

ES 2 368 267 T3

1 5 10 15

Cys Ile Leu Trp Gly
20

<210> 115
<211> 60
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 25-84 de SEC ID n°: 29

<220>
<221> CDS
<222> (1)..(60)

<400> 115
aac ctg ctg cta gcc ctc tgc ctc ttt ctg gtc ctg gga ttt ttg tat 48
Asn Leu Leu Leu Ala Leu Cys Leu Phe Leu Val Leu Gly Phe Leu Tyr
1 5 10 15

tat tct gcc tgg 60
Tyr Ser Ala Trp
20

<210> 116
<211> 20
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 116

Asn Leu Leu Leu Ala Leu Cys Leu Phe Leu Val Leu Gly Phe Leu Tyr
1 5 10 15

Tyr Ser Ala Trp
20

<210> 117
<211> 57
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 88-144 de SEC ID n°: 31

<220>
<221> CDS
<222> (1)..(57)

ES 2 368 267 T3

<400> 117
 atg gga gcc agt gcc ctc tgt gtc gtg gtc ctc tgt tgg ctc tac atc 48
 Met Gly Ala Ser Ala Leu Cys Val Val Val Leu Cys Trp Leu Tyr Ile
 1 5 10 15

ttc ccc gtc 57
 Phe Pro Val

<210> 118
 <211> 19
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 118

Met Gly Ala Ser Ala Leu Cys Val Val Val Leu Cys Trp Leu Tyr Ile
 1 5 10 15

Phe Pro Val

<210> 119
 <211> 51
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 19-69 de SEC ID nº: 33

<220>
 <221> CDS
 <222> (1)..(51)

<400> 119
 agc tgg atg ctg gcc gcg ctc acg ctg ctc gtg gtc ttc ctc atc ttc 48
 Ser Trp Met Leu Ala Ala Leu Thr Leu Leu Val Val Phe Leu Ile Phe
 1 5 10 15

gca 51
 Ala

<210> 120
 <211> 17
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 120

ES 2 368 267 T3

Ser Trp Met Leu Ala Ala Leu Thr Leu Leu Val Val Phe Leu Ile Phe
1 5 10 15

Ala

<210> 121
<211> 72
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 28-99 de SEC ID nº: 35

<220>
<221> CDS
<222> (1)..(72)

<400> 121
gcc agt gtg ctg ggg ctg gtc atg ctc agc gtc gcc ctg ctg att tta 48
Ala Ser Val Leu Gly Leu Val Met Leu Ser Val Ala Leu Leu Ile Leu
1 5 10 15

tcg ctc atc agc tac gtg tcc ctg 72
Ser Leu Ile Ser Tyr Val Ser Leu
20

<210> 122
<211> 24
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 122

Ala Ser Val Leu Gly Leu Val Met Leu Ser Val Ala Leu Leu Ile Leu
1 5 10 15

Ser Leu Ile Ser Tyr Val Ser Leu
20

<210> 123
<211> 39
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 22-60 de SEC ID nº: 37

<220>
<221> CDS
<222> (1)..(39)

ES 2 368 267 T3

<400> 123
 tgg acg atc tgc aca ata agt ctg ctc ctg atc ttt tat 39
 Trp Thr Ile Cys Thr Ile Ser Leu Leu Leu Ile Phe Tyr
 1 5 10

<210> 124
 <211> 13
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 124
 Trp Thr Ile Cys Thr Ile Ser Leu Leu Leu Ile Phe Tyr
 1 5 10

<210> 125
 <211> 63
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 52-114 de SEC ID nº: 39

<220>
 <221> CDS
 <222> (1)..(63)

<400> 125
 ttg ctc ttc atc ttc atc tgc gcc ttt gcc ttg gtg acc ttg ctg caa 48
 Leu Leu Phe Ile Phe Ile Cys Ala Phe Ala Leu Val Thr Leu Leu Gln
 1 5 10 15

cag atc ctg tat ggc 63
 Gln Ile Leu Tyr Gly
 20

<210> 126
 <211> 21
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 126
 Leu Leu Phe Ile Phe Ile Cys Ala Phe Ala Leu Val Thr Leu Leu Gln
 1 5 10 15

Gln Ile Leu Tyr Gly
 20

ES 2 368 267 T3

<210> 127
 <211> 63
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 10-72 de SEC ID nº: 41

<220>
 <221> CDS
 <222> (1)..(63)

<400> 127
 ggg ggc gca ctg ctc gcc ctg ctc gcc agc ctg ctg ctg ctg ctg ctg 48
 Gly Gly Ala Leu Leu Ala Leu Leu Ala Ser Leu Leu Leu Leu Leu Leu
 1 5 10 15
 ctg cgc ctg ctc tgg 63
 Leu Arg Leu Leu Trp
 20

<210> 128
 <211> 21
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 128
 Gly Gly Ala Leu Leu Ala Leu Leu Ala Ser Leu Leu Leu Leu Leu Leu
 1 5 10 15
 Leu Arg Leu Leu Trp
 20

<210> 129
 <211> 117
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 106-222 de SEC ID nº: 5

<220>
 <221> CDS
 <222> (1)..(117)

<400> 129
 aag gag cct caa aca aag cct tcc agg cat caa cgc aca gag aac att 48
 Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile
 1 5 10 15
 aaa gaa agg tct cta cag tcc ctg gca aag cct aag tcc cag gca ccc 96

ES 2 368 267 T3

Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro
20 25 30

aca agg gca agg agg aca acc 117
Thr Arg Ala Arg Arg Thr Thr
35

<210> 130
<211> 39
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 130

Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile
1 5 10 15

Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro
20 25 30

Thr Arg Ala Arg Arg Thr Thr
35

<210> 131
<211> 114
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 109-222 de SEC ID n°: 5

<220>
<221> CDS
<222> (1)..(114)

<400> 131
gag cct caa aca aag cct tcc agg cat caa cgc aca gag aac att aaa 48
Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile Lys
1 5 10 15

gaa agg tct cta cag tcc ctg gca aag cct aag tcc cag gca ccc aca 96
Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro Thr
20 25 30

agg gca agg agg aca acc 114
Arg Ala Arg Arg Thr Thr
35

<210> 132
<211> 38
<212> PRT
<213> Artificial

ES 2 368 267 T3

<220>

<223> Constructo sintético

<400> 132

Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr	Glu	Asn	Ile	Lys
1				5					10					15	

Glu	Arg	Ser	Leu	Gln	Ser	Leu	Ala	Lys	Pro	Lys	Ser	Gln	Ala	Pro	Thr
			20					25					30		

Arg	Ala	Arg	Arg	Thr	Thr
				35	

<210> 133

<211> 315

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 106-420 de SEC ID n°: 5

<220>

<221> CDS

<222> (1)..(315)

<400> 133

aag	gag	cct	caa	aca	aag	cct	tcc	agg	cat	caa	cgc	aca	gag	aac	att	
Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr	Glu	Asn	Ile	
1				5					10					15		48

aaa	gaa	agg	tct	cta	cag	tcc	ctg	gca	aag	cct	aag	tcc	cag	gca	ccc	
Lys	Glu	Arg	Ser	Leu	Gln	Ser	Leu	Ala	Lys	Pro	Lys	Ser	Gln	Ala	Pro	
			20					25					30			96

aca	agg	gca	agg	agg	aca	acc	atc	tat	gca	gag	cca	gtg	cca	gag	aac	
Thr	Arg	Ala	Arg	Arg	Thr	Thr	Ile	Tyr	Ala	Glu	Pro	Val	Pro	Glu	Asn	
		35					40					45				144

aat	gcc	ctc	aac	aca	caa	acc	cag	ccc	aag	gcc	cac	acc	acc	gga	gac	
Asn	Ala	Leu	Asn	Thr	Gln	Thr	Gln	Pro	Lys	Ala	His	Thr	Thr	Gly	Asp	
	50						55				60					192

aga	gga	aag	gag	gcc	aac	cag	gca	ccg	ccg	gag	gag	cag	gac	aag	gtg	
Arg	Gly	Lys	Glu	Ala	Asn	Gln	Ala	Pro	Pro	Glu	Glu	Gln	Asp	Lys	Val	
65					70					75				80		240

ccc	cac	aca	gca	cag	agg	gca	gca	tgg	aag	agc	cca	gaa	aaa	gag	aaa	
Pro	His	Thr	Ala	Gln	Arg	Ala	Ala	Trp	Lys	Ser	Pro	Glu	Lys	Glu	Lys	
				85					90					95		288

acc	atg	gtg	aac	aca	ctg	tca	ccc	aga								
Thr	Met	Val	Asn	Thr	Leu	Ser	Pro	Arg								
					100			105								315

ES 2 368 267 T3

<210> 134
 <211> 105
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 134

Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile
 1 5 10 15

Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro
 20 25 30

Thr Arg Ala Arg Arg Thr Thr Ile Tyr Ala Glu Pro Val Pro Glu Asn
 35 40 45

Asn Ala Leu Asn Thr Gln Thr Gln Pro Lys Ala His Thr Thr Gly Asp
 50 55 60

Arg Gly Lys Glu Ala Asn Gln Ala Pro Pro Glu Glu Gln Asp Lys Val
 65 70 75 80

Pro His Thr Ala Gln Arg Ala Ala Trp Lys Ser Pro Glu Lys Glu Lys
 85 90 95

Thr Met Val Asn Thr Leu Ser Pro Arg
 100 105

<210> 135
 <211> 669
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 106-774 de SEC ID n°: 5

<220>
 <221> CDS
 <222> (1)..(669)

<400> 135

aag gag cct caa aca aag cct tcc agg cat caa cgc aca gag aac att 48
 Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile
 1 5 10 15

aaa gaa agg tct cta cag tcc ctg gca aag cct aag tcc cag gca ccc 96
 Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro
 20 25 30

aca agg gca agg agg aca acc atc tat gca gag cca gtg cca gag aac 144

ES 2 368 267 T3

Thr	Arg	Ala	Arg	Arg	Thr	Thr	Ile	Tyr	Ala	Glu	Pro	Val	Pro	Glu	Asn	
		35					40					45				
aat	gcc	ctc	aac	aca	caa	acc	cag	ccc	aag	gcc	cac	acc	acc	gga	gac	192
Asn	Ala	Leu	Asn	Thr	Gln	Thr	Gln	Pro	Lys	Ala	His	Thr	Thr	Gly	Asp	
	50					55				60						
aga	gga	aag	gag	gcc	aac	cag	gca	ccg	ccg	gag	gag	cag	gac	aag	gtg	240
Arg	Gly	Lys	Glu	Ala	Asn	Gln	Ala	Pro	Pro	Glu	Glu	Gln	Asp	Lys	Val	
65					70					75				80		
ccc	cac	aca	gca	cag	agg	gca	gca	tgg	aag	agc	cca	gaa	aaa	gag	aaa	288
Pro	His	Thr	Ala	Gln	Arg	Ala	Ala	Trp	Lys	Ser	Pro	Glu	Lys	Glu	Lys	
				85					90					95		
acc	atg	gtg	aac	aca	ctg	tca	ccc	aga	ggg	caa	gat	gca	ggg	atg	gcc	336
Thr	Met	Val	Asn	Thr	Leu	Ser	Pro	Arg	Gly	Gln	Asp	Ala	Gly	Met	Ala	
			100					105					110			
tct	ggc	agg	aca	gag	gca	caa	tca	tgg	aag	agc	cag	gac	aca	aag	acg	384
Ser	Gly	Arg	Thr	Glu	Ala	Gln	Ser	Trp	Lys	Ser	Gln	Asp	Thr	Lys	Thr	
		115					120					125				
acc	caa	gga	aat	ggg	ggc	cag	acc	agg	aag	ctg	acg	gcc	tcc	agg	acg	432
Thr	Gln	Gly	Asn	Gly	Gly	Gln	Thr	Arg	Lys	Leu	Thr	Ala	Ser	Arg	Thr	
	130					135					140					
gtg	tca	gag	aag	cac	cag	ggc	aaa	gcg	gca	acc	aca	gcc	aag	acg	ctc	480
Val	Ser	Glu	Lys	His	Gln	Gly	Lys	Ala	Ala	Thr	Thr	Ala	Lys	Thr	Leu	
145					150					155					160	
att	ccc	aaa	agt	cag	cac	aga	atg	ctg	gct	ccc	aca	gga	gca	gtg	tca	528
Ile	Pro	Lys	Ser	Gln	His	Arg	Met	Leu	Ala	Pro	Thr	Gly	Ala	Val	Ser	
				165				170						175		
aca	agg	acg	aga	cag	aaa	gga	gtg	acc	aca	gca	gtc	atc	cca	cct	aag	576
Thr	Arg	Thr	Arg	Gln	Lys	Gly	Val	Thr	Thr	Ala	Val	Ile	Pro	Pro	Lys	
			180					185					190			
gag	aag	aaa	cct	cag	gcc	acc	cca	ccc	cct	gcc	cct	ttc	cag	agc	ccc	624
Glu	Lys	Lys	Pro	Gln	Ala	Thr	Pro	Pro	Pro	Ala	Pro	Phe	Gln	Ser	Pro	
		195					200					205				
acg	acg	cag	aga	aac	caa	aga	ctg	aag	gcc	gcc	aac	ttc	aaa	tct		669
Thr	Thr	Gln	Arg	Asn	Gln	Arg	Leu	Lys	Ala	Ala	Asn	Phe	Lys	Ser		
	210					215					220					
<210> 136																
<211> 223																
<212> PRT																
<213> Artificial																
<220>																
<223> Constructo sintético																
<400> 136																
Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr	Glu	Asn	Ile	
1				5					10					15		

ES 2 368 267 T3

Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro
 20 25 30

Thr Arg Ala Arg Arg Thr Thr Ile Tyr Ala Glu Pro Val Pro Glu Asn
 35 40 45

Asn Ala Leu Asn Thr Gln Thr Gln Pro Lys Ala His Thr Thr Gly Asp
 50 55 60

Arg Gly Lys Glu Ala Asn Gln Ala Pro Pro Glu Glu Gln Asp Lys Val
 65 70 75 80

Pro His Thr Ala Gln Arg Ala Ala Trp Lys Ser Pro Glu Lys Glu Lys
 85 90 95

Thr Met Val Asn Thr Leu Ser Pro Arg Gly Gln Asp Ala Gly Met Ala
 100 105 110

Ser Gly Arg Thr Glu Ala Gln Ser Trp Lys Ser Gln Asp Thr Lys Thr
 115 120 125

Thr Gln Gly Asn Gly Gly Gln Thr Arg Lys Leu Thr Ala Ser Arg Thr
 130 135 140

Val Ser Glu Lys His Gln Gly Lys Ala Ala Thr Thr Ala Lys Thr Leu
 145 150 155 160

Ile Pro Lys Ser Gln His Arg Met Leu Ala Pro Thr Gly Ala Val Ser
 165 170 175

Thr Arg Thr Arg Gln Lys Gly Val Thr Thr Ala Val Ile Pro Pro Lys
 180 185 190

Glu Lys Lys Pro Gln Ala Thr Pro Pro Pro Ala Pro Phe Gln Ser Pro
 195 200 205

Thr Thr Gln Arg Asn Gln Arg Leu Lys Ala Ala Asn Phe Lys Ser
 210 215 220

<210> 137

<211> 54

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 85-138 de SEC ID n°: 21

ES 2 368 267 T3

```

<220>
<221> CDS
<222> (1)..(54)

<400> 137
aag cta cac tta ctc cag tgg gag gag gac tcc agt aag tat agt cac      48
Lys Leu His Leu Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His
1              5              10              15

tct agc
Ser Ser

```

```

<210> 138
<211> 18
<212> PRT
<213> Artificial

```

```

<220>
<223> Constructo sintético

```

```

<400> 138

Lys Leu His Leu Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His
1              5              10              15

Ser Ser

```

```

<210> 139
<211> 54
<212> ADN
<213> Artificial

```

```

<220>
<223> Nucleótidos 85-138 de SEC ID nº: 29

```

```

<220>
<221> CDS
<222> (1)..(54)

```

```

<400> 139
aag cta cac tta ctc caa tgg gaa gac tcc aat tca ctg att ctt tcc      48
Lys Leu His Leu Leu Gln Trp Glu Asp Ser Asn Ser Leu Ile Leu Ser
1              5              10              15

ctt gac
Leu Asp

```

```

<210> 140
<211> 18
<212> PRT
<213> Artificial

```

ES 2 368 267 T3

<220>

<223> Constructo sintético

<400> 140

Lys Leu His Leu Leu Gln Trp Glu Asp Ser Asn Ser Leu Ile Leu Ser
1 5 10 15

Leu Asp

<210> 141

<211> 168

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 70-237 de SEC ID nº: 33

<220>

<221> CDS

<222> (1)..(168)

<400> 141

gac atc tca gag atc gaa gaa gaa atc ggg aat tcg gga ggc aga ggt 48
Asp Ile Ser Glu Ile Glu Glu Glu Ile Gly Asn Ser Gly Gly Arg Gly
1 5 10 15

aca atc aga tca gct gtg aac agc tta cat agc aaa tct aat aga gct 96
Thr Ile Arg Ser Ala Val Asn Ser Leu His Ser Lys Ser Asn Arg Ala
20 25 30

gaa gtt gta ata aac ggc tcc tca tca cca gct gtt gtt gac aga agt 144
Glu Val Val Ile Asn Gly Ser Ser Ser Pro Ala Val Val Asp Arg Ser
35 40 45

aat gaa agc atc aag cac aac atc 168
Asn Glu Ser Ile Lys His Asn Ile
50 55

<210> 142

<211> 56

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 142

Asp Ile Ser Glu Ile Glu Glu Glu Ile Gly Asn Ser Gly Gly Arg Gly
1 5 10 15

Thr Ile Arg Ser Ala Val Asn Ser Leu His Ser Lys Ser Asn Arg Ala
20 25 30

ES 2 368 267 T3

Glu Val Val Ile Asn Gly Ser Ser Ser Pro Ala Val Val Asp Arg Ser
 35 40 45

Asn Glu Ser Ile Lys His Asn Ile
 50 55

<210> 143
 <211> 141
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 61-201 de SEC ID nº: 37

<220>
 <221> CDS
 <222> (1)..(141)

<400> 143
 aag aca aaa gaa ata gca aga act gag gag cac cag gag acg caa ctc 48
 Lys Thr Lys Glu Ile Ala Arg Thr Glu Glu His Gln Glu Thr Gln Leu
 1 5 10 15
 atc gga gat ggt gaa ttg tct ttg agt cgg tca ctt gtc aat agc tct 96
 Ile Gly Asp Gly Glu Leu Ser Leu Ser Arg Ser Leu Val Asn Ser Ser
 20 25 30
 gat aaa atc att cga aag gct ggc tct tca atc ttc cag cac aat 141
 Asp Lys Ile Ile Arg Lys Ala Gly Ser Ser Ile Phe Gln His Asn
 35 40 45

<210> 144
 <211> 47
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 144
 Lys Thr Lys Glu Ile Ala Arg Thr Glu Glu His Gln Glu Thr Gln Leu
 1 5 10 15
 Ile Gly Asp Gly Glu Leu Ser Leu Ser Arg Ser Leu Val Asn Ser Ser
 20 25 30
 Asp Lys Ile Ile Arg Lys Ala Gly Ser Ser Ile Phe Gln His Asn
 35 40 45

<210> 145
 <211> 222
 <212> ADN

ES 2 368 267 T3

<213> Artificial

<220>

<223> Nucleótidos 1-222 de SEC ID nº: 5

<220>

<221> CDS

<222> (1)..(222)

<400> 145

atg	agg	tcc	tgc	ctg	tgg	aga	tgc	agg	cac	ctg	agc	caa	ggc	gtc	cag	48
Met	Arg	Ser	Cys	Leu	Trp	Arg	Cys	Arg	His	Leu	Ser	Gln	Gly	Val	Gln	
1				5					10					15		

tgg	tcc	ttg	ctt	ctg	gct	gtc	ctg	gtc	ttc	ttt	ctc	ttc	gcc	ttg	ccc	96
Trp	Ser	Leu	Leu	Leu	Ala	Val	Leu	Val	Phe	Phe	Leu	Phe	Ala	Leu	Pro	
			20				25						30			

tct	ttt	att	aag	gag	cct	caa	aca	aag	cct	tcc	agg	cat	caa	cgc	aca	144
Ser	Phe	Ile	Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr	
		35				40						45				

gag	aac	att	aaa	gaa	agg	tct	cta	cag	tcc	ctg	gca	aag	cct	aag	tcc	192
Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser	Leu	Ala	Lys	Pro	Lys	Ser	
	50					55				60						

cag	gca	ccc	aca	agg	gca	agg	agg	aca	acc	222
Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr	
65					70					

<210> 146

<211> 74

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 146

Met	Arg	Ser	Cys	Leu	Trp	Arg	Cys	Arg	His	Leu	Ser	Gln	Gly	Val	Gln
1				5					10					15	

Trp	Ser	Leu	Leu	Leu	Ala	Val	Leu	Val	Phe	Phe	Leu	Phe	Ala	Leu	Pro
			20				25						30		

Ser	Phe	Ile	Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr
		35				40						45			

Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser	Leu	Ala	Lys	Pro	Lys	Ser
	50					55				60					

Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr
65					70				

ES 2 368 267 T3

```

<210> 147
<211> 420
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-420 de SEC ID nº: 5

<220>
<221> CDS
<222> (1)..(420)

<400> 147
atg agg tcc tgc ctg tgg aga tgc agg cac ctg agc caa ggc gtc cag      48
Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val Gln
1          5          10          15

tgg tcc ttg ctt ctg gct gtc ctg gtc ttc ttt ctc ttc gcc ttg ccc      96
Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala Leu Pro
          20          25          30

tct ttt att aag gag cct caa aca aag cct tcc agg cat caa cgc aca      144
Ser Phe Ile Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr
          35          40          45

gag aac att aaa gaa agg tct cta cag tcc ctg gca aag cct aag tcc      192
Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser
          50          55          60

cag gca ccc aca agg gca agg agg aca acc atc tat gca gag cca gtg      240
Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr Ile Tyr Ala Glu Pro Val
65          70          75          80

cca gag aac aat gcc ctc aac aca caa acc cag ccc aag gcc cac acc      288
Pro Glu Asn Asn Ala Leu Asn Thr Gln Thr Gln Pro Lys Ala His Thr
          85          90          95

acc gga gac aga gga aag gag gcc aac cag gca ccg ccg gag gag cag      336
Thr Gly Asp Arg Gly Lys Glu Ala Asn Gln Ala Pro Pro Glu Glu Gln
          100          105          110

gac aag gtg ccc cac aca gca cag agg gca gca tgg aag agc cca gaa      384
Asp Lys Val Pro His Thr Ala Gln Arg Ala Ala Trp Lys Ser Pro Glu
          115          120          125

aaa gag aaa acc atg gtg aac aca ctg tca ccc aga      420
Lys Glu Lys Thr Met Val Asn Thr Leu Ser Pro Arg
          130          135          140

<210> 148
<211> 140
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 148

```

ES 2 368 267 T3

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

<210> 149

<211> 774

<212> ADN

<213> Artificial

<220>

<223> Nucleótidos 1-774 de SEC ID nº: 5

<220>

<221> CDS

<222> (1)..(774)

<400> 149

atg	agg	tcc	tgc	ctg	tgg	aga	tgc	agg	cac	ctg	agc	caa	ggc	gtc	cag	48
Met	Arg	Ser	Cys	Leu	Trp	Arg	Cys	Arg	His	Leu	Ser	Gln	Gly	Val	Gln	
1				5					10					15		

tgg	tcc	ttg	ctt	ctg	gct	gtc	ctg	gtc	ttc	ttt	ctc	ttc	gcc	ttg	ccc	96
Trp	Ser	Leu	Leu	Leu	Ala	Val	Leu	Val	Phe	Phe	Leu	Phe	Ala	Leu	Pro	
			20					25					30			

tct	ttt	att	aag	gag	cct	caa	aca	aag	cct	tcc	agg	cat	caa	cgc	aca	144
Ser	Phe	Ile	Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr	

ES 2 368 267 T3

35	40	45	
gag aac att aaa gaa agg tct cta cag tcc ctg gca aag cct aag tcc			192
Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser			
50	55	60	
cag gca ccc aca agg gca agg agg aca acc atc tat gca gag cca gtg			240
Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr Ile Tyr Ala Glu Pro Val			
65	70	75	80
cca gag aac aat gcc ctc aac aca caa acc cag ccc aag gcc cac acc			288
Pro Glu Asn Asn Ala Leu Asn Thr Gln Thr Gln Pro Lys Ala His Thr			
	85	90	95
acc gga gac aga gga aag gag gcc aac cag gca ccg ccg gag gag cag			336
Thr Gly Asp Arg Gly Lys Glu Ala Asn Gln Ala Pro Pro Glu Glu Gln			
	100	105	110
gac aag gtg ccc cac aca gca cag agg gca gca tgg aag agc cca gaa			384
Asp Lys Val Pro His Thr Ala Gln Arg Ala Ala Trp Lys Ser Pro Glu			
	115	120	125
aaa gag aaa acc atg gtg aac aca ctg tca ccc aga ggg caa gat gca			432
Lys Glu Lys Thr Met Val Asn Thr Leu Ser Pro Arg Gly Gln Asp Ala			
	130	135	140
ggg atg gcc tct ggc agg aca gag gca caa tca tgg aag agc cag gac			480
Gly Met Ala Ser Gly Arg Thr Glu Ala Gln Ser Trp Lys Ser Gln Asp			
145	150	155	160
aca aag acg acc caa gga aat ggg ggc cag acc agg aag ctg acg gcc			528
Thr Lys Thr Thr Gln Gly Asn Gly Gly Gln Thr Arg Lys Leu Thr Ala			
	165	170	175
tcc agg acg gtg tca gag aag cac cag ggc aaa gcg gca acc aca gcc			576
Ser Arg Thr Val Ser Glu Lys His Gln Gly Lys Ala Ala Thr Thr Ala			
	180	185	190
aag acg ctc att ccc aaa agt cag cac aga atg ctg gct ccc aca gga			624
Lys Thr Leu Ile Pro Lys Ser Gln His Arg Met Leu Ala Pro Thr Gly			
	195	200	205
gca gtg tca aca agg acg aga cag aaa gga gtg acc aca gca gtc atc			672
Ala Val Ser Thr Arg Thr Arg Gln Lys Gly Val Thr Thr Ala Val Ile			
	210	215	220
cca cct aag gag aag aaa cct cag gcc acc cca ccc cct gcc cct ttc			720
Pro Pro Lys Glu Lys Lys Pro Gln Ala Thr Pro Pro Pro Ala Pro Phe			
225	230	235	240
cag agc ccc acg acg cag aga aac caa aga ctg aag gcc gcc aac ttc			768
Gln Ser Pro Thr Thr Gln Arg Asn Gln Arg Leu Lys Ala Ala Asn Phe			
	245	250	255
aaa tct			774
Lys Ser			

<210> 150

<211> 258

ES 2 368 267 T3

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 150

Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val Gln
1 5 10 15

Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala Leu Pro
20 25 30

Ser Phe Ile Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr
35 40 45

Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser
50 55 60

Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr Ile Tyr Ala Glu Pro Val
65 70 75 80

Pro Glu Asn Asn Ala Leu Asn Thr Gln Thr Gln Pro Lys Ala His Thr
85 90 95

Thr Gly Asp Arg Gly Lys Glu Ala Asn Gln Ala Pro Pro Glu Glu Gln
100 105 110

Asp Lys Val Pro His Thr Ala Gln Arg Ala Ala Trp Lys Ser Pro Glu
115 120 125

Lys Glu Lys Thr Met Val Asn Thr Leu Ser Pro Arg Gly Gln Asp Ala
130 135 140

Gly Met Ala Ser Gly Arg Thr Glu Ala Gln Ser Trp Lys Ser Gln Asp
145 150 155 160

Thr Lys Thr Thr Gln Gly Asn Gly Gly Gln Thr Arg Lys Leu Thr Ala
165 170 175

Ser Arg Thr Val Ser Glu Lys His Gln Gly Lys Ala Ala Thr Thr Ala
180 185 190

Lys Thr Leu Ile Pro Lys Ser Gln His Arg Met Leu Ala Pro Thr Gly
195 200 205

Ala Val Ser Thr Arg Thr Arg Gln Lys Gly Val Thr Thr Ala Val Ile

ES 2 368 267 T3

210

215

220

Pro Pro Lys Glu Lys Lys Pro Gln Ala Thr Pro Pro Pro Ala Pro Phe
225 230 235 240

Gln Ser Pro Thr Thr Gln Arg Asn Gln Arg Leu Lys Ala Ala Asn Phe
245 250 255

Lys Ser

<210> 151
<211> 138
<212> ADN
<213> Artificial

<220>
<223> Nucleótidos 1-138 de SEC ID n°: 21,

<220>
<221> CDS
<222> (1)..(138)

<400> 151
atg gga ctc ttg gta ttt gtg cgc aat ctg ctg cta gcc ctc tgc ctc 48
Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
1 5 10 15
ttt ctg gta ctg gga ttt ttg tat tat tct gcg tgg aag cta cac tta 96
Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
20 25 30
ctc cag tgg gag gag gac tcc agt aag tat agt cac tct agc 138
Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His Ser Ser
35 40 45

<210> 152
<211> 46
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 152

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

ES 2 368 267 T3

35

40

45

<210> 153
 <211> 138
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-138 de SEC ID n°: 29

<220>
 <221> CDS
 <222> (1)..(138)

<400> 153
 atg gga ctc ttg gta ttt gta cgc aac ctg ctg cta gcc ctc tgc ctc 48
 Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15
 ttt ctg gtc ctg gga ttt ttg tat tat tct gcc tgg aag cta cac tta 96
 Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu
 20 25 30
 ctc caa tgg gaa gac tcc aat tca ctg att ctt tcc ctt gac 138
 Leu Gln Trp Glu Asp Ser Asn Ser Leu Ile Leu Ser Leu Asp
 35 40 45

<210> 154
 <211> 46
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 154
 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 Asp Ser Asn Ser Leu Ile Leu Ser Leu Asp
 35 40 45

<210> 155
 <211> 237
 <212> ADN
 <213> Artificial

<220>
 <223> Nucleótidos 1-237 de SEC ID n°: 33

ES 2 368 267 T3

```

<220>
<221> CDS
<222> (1)..(237)

<400> 155
atg cag ctg cag ttc cgg agc tgg atg ctg gcc gcg ctc acg ctg ctc      48
Met Gln Leu Gln Phe Arg Ser Trp Met Leu Ala Ala Leu Thr Leu Leu
1          5          10          15

gtg gtc ttc ctc atc ttc gca gac atc tca gag atc gaa gaa gaa atc      96
Val Val Phe Leu Ile Phe Ala Asp Ile Ser Glu Ile Glu Glu Glu Ile
          20          25          30

ggg aat tcg gga ggc aga ggt aca atc aga tca gct gtg aac agc tta      144
Gly Asn Ser Gly Gly Arg Gly Thr Ile Arg Ser Ala Val Asn Ser Leu
          35          40          45

cat agc aaa tct aat aga gct gaa gtt gta ata aac ggc tcc tca tca      192
His Ser Lys Ser Asn Arg Ala Glu Val Val Ile Asn Gly Ser Ser Ser
          50          55          60

cca gct gtt gtt gac aga agt aat gaa agc atc aag cac aac atc      237
Pro Ala Val Val Asp Arg Ser Asn Glu Ser Ile Lys His Asn Ile
65          70          75

<210> 156
<211> 79
<212> PRT
<213> Artificial

<220>
<223> Constructo sintético

<400> 156

Met Gln Leu Gln Phe Arg Ser Trp Met Leu Ala Ala Leu Thr Leu Leu
1          5          10          15

Val Val Phe Leu Ile Phe Ala Asp Ile Ser Glu Ile Glu Glu Glu Ile
          20          25          30

Gly Asn Ser Gly Gly Arg Gly Thr Ile Arg Ser Ala Val Asn Ser Leu
          35          40          45

His Ser Lys Ser Asn Arg Ala Glu Val Val Ile Asn Gly Ser Ser Ser
          50          55          60

Pro Ala Val Val Asp Arg Ser Asn Glu Ser Ile Lys His Asn Ile
65          70          75

<210> 157
<211> 201
<212> ADN
<213> Artificial

```

ES 2 368 267 T3

<220>

<223> Nucleótidos 1-201 de SEC ID nº: 37

<220>

<221> CDS

<222> (1)..(201)

<400> 157

atg	cgc	tcc	att	agg	aag	agg	tgg	acg	atc	tgc	aca	ata	agt	ctg	ctc	48
Met	Arg	Ser	Ile	Arg	Lys	Arg	Trp	Thr	Ile	Cys	Thr	Ile	Ser	Leu	Leu	
1				5				10						15		

ctg	atc	ttt	tat	aag	aca	aaa	gaa	ata	gca	aga	act	gag	gag	cac	cag	96
Leu	Ile	Phe	Tyr	Lys	Thr	Lys	Glu	Ile	Ala	Arg	Thr	Glu	Glu	His	Gln	
			20				25						30			

gag	acg	caa	ctc	atc	gga	gat	ggt	gaa	ttg	tct	ttg	agt	cgg	tca	ctt	144
Glu	Thr	Gln	Leu	Ile	Gly	Asp	Gly	Glu	Leu	Ser	Leu	Ser	Arg	Ser	Leu	
		35				40						45				

gtc	aat	agc	tct	gat	aaa	atc	att	cga	aag	gct	ggc	tct	tca	atc	ttc	192
Val	Asn	Ser	Ser	Asp	Lys	Ile	Ile	Arg	Lys	Ala	Gly	Ser	Ser	Ile	Phe	
	50					55					60					

cag	cac	aat														201
Gln	His	Asn														
65																

<210> 158

<211> 67

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 158

Met	Arg	Ser	Ile	Arg	Lys	Arg	Trp	Thr	Ile	Cys	Thr	Ile	Ser	Leu	Leu
1				5				10						15	

Leu	Ile	Phe	Tyr	Lys	Thr	Lys	Glu	Ile	Ala	Arg	Thr	Glu	Glu	His	Gln
			20				25						30		

Glu	Thr	Gln	Leu	Ile	Gly	Asp	Gly	Glu	Leu	Ser	Leu	Ser	Arg	Ser	Leu
		35				40						45			

Val	Asn	Ser	Ser	Asp	Lys	Ile	Ile	Arg	Lys	Ala	Gly	Ser	Ser	Ile	Phe
	50					55					60				

Gln	His	Asn													
65															

<210> 159

ES 2 368 267 T3

<211> 198
 <212> ADN
 <213> Artificial

<220>
 <223> Fusión de la secuencia nucleotídica SEC ID n°:161 con la
 secuencia nucleotídica SEC ID n°:131

<220>
 <221> CDS
 <222> (1)..(198)

<400> 159
 atg gga ctc ttg gta ttt gtg cgc aat ctg ctg cta gcc ctc tgc ctc 48
 Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu
 1 5 10 15
 ttt ctg gta ctg gga ttt ttg tat tat tct gcg tgg gag cct caa aca 96
 Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Glu Pro Gln Thr
 20 25 30
 aag cct tcc agg cat caa cgc aca gag aac att aaa gaa agg tct cta 144
 Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile Lys Glu Arg Ser Leu
 35 40 45
 cag tcc ctg gca aag cct aag tcc cag gca ccc aca agg gca agg agg 192
 Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro Thr Arg Ala Arg Arg
 50 55 60
 aca acc 198
 Thr Thr
 65

<210> 160
 <211> 66
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 160
 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 Glu Pro Gln Thr
 20 25 30
 Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile Lys Glu Arg Ser Leu
 35 40 45
 Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro Thr Arg Ala Arg Arg
 50 55 60

ES 2 368 267 T3

Thr Thr
65

<210> 161
<211> 1179
<212> ADN
<213> Artificial

<220>
<223> Fusión de la secuencia de nucleótidos SEC ID n°: 145 con
la secuencia de nucleótidos SEC ID n°: 43

<220>
<221> CDS
<222> (1)..(1179)

```

<400> 161
atg agg tcc tgc ctg tgg aga tgc agg cac ctg agc caa ggc gtc cag      48
Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val Gln
1          5          10          15

tgg tcc ttg ctt ctg gct gtc ctg gtc ttc ttt ctc ttc gcc ttg ccc      96
Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala Leu Pro
          20          25          30

tct ttt att aag gag cct caa aca aag cct tcc agg cat caa cgc aca      144
Ser Phe Ile Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr
          35          40          45

gag aac att aaa gaa agg tct cta cag tcc ctg gca aag cct aag tcc      192
Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser
          50          55          60

cag gca ccc aca agg gca agg agg aca acc gga tcc gag gcc tcc ttc      240
Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr Gly Ser Glu Ala Ser Phe
65          70          75          80

cag gtg tgg aac aag gac agc tct tcc aaa aac ctt atc cct agg ctg      288
Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu Ile Pro Arg Leu
          85          90          95

caa aag atc tgg aag aat tac cta agc atg aac aag tac aaa gtg tcc      336
Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys Tyr Lys Val Ser
          100          105          110

tac aag ggg cca gga cca ggc atc aag ttc agt gca gag gcc ctg cgc      384
Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala Glu Ala Leu Arg
          115          120          125

tgc cac ctc cyg gac cat gtg aat gta tcc atg gta gag gtc aca gat      432
Cys His Leu Arg Asp His Val Asn Val Ser Met Val Glu Val Thr Asp
          130          135          140

ttt ccc ttc aat acc tct gaa tgg gag ggt tat ctg ccc aag gag agc      480
Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr Leu Pro Lys Glu Ser
145          150          155          160

att agg acc aag gct ggg cct tgg ggc agg tgt gct gtt gtg tcg tca      528
Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys Ala Val Val Ser Ser

```

ES 2 368 267 T3

	165	170	175	
gcg gga tct ctg aag tcc tcc caa cta ggc aga gaa atc gat gat cat				576
Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg Glu Ile Asp Asp His				
	180	185	190	
gac gca gtc ctg agg ttt aat ggg gca ccc aca gcc aac ttc caa caa				624
Asp Ala Val Leu Arg Phe Asn Gly Ala Pro Thr Ala Asn Phe Gln Gln				
	195	200	205	
gat gtg ggc aca aaa act acc att cgc ctg atg aac tct cag ttg gtt				672
Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met Asn Ser Gln Leu Val				
	210	215	220	
acc aca gag aag cgc ttc ctc aaa gac agt ttg tac aat gaa gga atc				720
Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu Tyr Asn Glu Gly Ile				
	225	230	235	240
cta att gta tgg gac cca tct gta tac cac tca gat atc cca aag tgg				768
Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser Asp Ile Pro Lys Trp				
	245	250	255	
tac cag aat ccg gat tat aat ttc ttt aac aac tac aag act tat cgt				816
Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn Tyr Lys Thr Tyr Arg				
	260	265	270	
aag ctg cac ccc aat cag ccc ttt tac atc ctc aag ccc cag atg cct				864
Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu Lys Pro Gln Met Pro				
	275	280	285	
tgg gag cta tgg gac att ctt caa gaa atc tcc cca gaa gag att cag				912
Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser Pro Glu Glu Ile Gln				
	290	295	300	
cca aac ccc cca tcc tct ggg atg ctt ggt atc atc atc atg atg acg				960
Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile Ile Ile Met Met Thr				
	305	310	315	320
ctg tgt gac cag gtg gat att tat gag ttc ctc cca tcc aag cgc aag				1008
Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu Pro Ser Lys Arg Lys				
	325	330	335	
act gac gtg tgc tac tac tac cag aag ttc ttc gat agt gcc tgc acg				1056
Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp Ser Ala Cys Thr				
	340	345	350	
atg ggt gcc tac cac ccg ctg ctc tat gag aag aat ttg gtg aag cat				1104
Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys Asn Leu Val Lys His				
	355	360	365	
ctc aac cag ggc aca gat gag gac atc tac ctg ctt gga aaa gcc aca				1152
Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu Leu Gly Lys Ala Thr				
	370	375	380	
ctg cct ggc ttc cgg acc att cac tgc				1179
Leu Pro Gly Phe Arg Thr Ile His Cys				
	385	390		
<210> 162				
<211> 393				

ES 2 368 267 T3

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 162

Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val Gln
1 5 10 15

Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala Leu Pro
20 25 30

Ser Phe Ile Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr
35 40 45

Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser
50 55 60

Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr Gly Ser Glu Ala Ser Phe
65 70 75 80

Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu Ile Pro Arg Leu
85 90 95

Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys Tyr Lys Val Ser
100 105 110

Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala Glu Ala Leu Arg
115 120 125

Cys His Leu Arg Asp His Val Asn Val Ser Met Val Glu Val Thr Asp
130 135 140

Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr Leu Pro Lys Glu Ser
145 150 155 160

Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys Ala Val Val Ser Ser
165 170 175

Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg Glu Ile Asp Asp His
180 185 190

Asp Ala Val Leu Arg Phe Asn Gly Ala Pro Thr Ala Asn Phe Gln Gln
195 200 205

Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met Asn Ser Gln Leu Val

ES 2 368 267 T3

210		215		220
Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu Tyr Asn Glu Gly Ile				
225		230		240
Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser Asp Ile Pro Lys Trp				
	245		250	255
Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn Tyr Lys Thr Tyr Arg				
	260		265	270
Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu Lys Pro Gln Met Pro				
	275		280	285
Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser Pro Glu Glu Ile Gln				
	290		295	300
Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile Ile Ile Met Met Thr				
305		310		320
Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu Pro Ser Lys Arg Lys				
	325		330	335
Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp Ser Ala Cys Thr				
	340		345	350
Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys Asn Leu Val Lys His				
	355		360	365
Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu Leu Gly Lys Ala Thr				
	370		375	380
Leu Pro Gly Phe Arg Thr Ile His Cys				
385		390		

<210> 163
 <211> 1377
 <212> DNA
 <213> Artificial

<220>
 <223> Fusión de la secuencia de nucleótidos SEC ID n°:147 con la
 secuencia de nucleótidos SEC ID n°:43

<220>
 <221> CDS
 <222> (1)..(1377)

ES 2 368 267 T3

```

<400> 163
atg agg tcc tgc ctg tgg aga tgc agg cac ctg agc caa ggc gtc cag      48
Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val Gln
1          5          10          15

tgg tcc ttg ctt ctg gct gtc ctg gtc ttc ttt ctc ttc gcc ttg ccc      96
Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala Leu Pro
          20          25          30

tct ttt att aag gag cct caa aca aag cct tcc agg cat caa cgc aca      144
Ser Phe Ile Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr
          35          40          45

gag aac att aaa gaa agg tct cta cag tcc ctg gca aag cct aag tcc      192
Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser
          50          55          60

cag gca ccc aca agg gca agg agg aca acc atc tat gca gag cca gtg      240
Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr Ile Tyr Ala Glu Pro Val
          65          70          75          80

cca gag aac aat gcc ctc aac aca caa acc cag ccc aag gcc cac acc      288
Pro Glu Asn Asn Ala Leu Asn Thr Gln Thr Gln Pro Lys Ala His Thr
          85          90          95

acc gga gac aga gga aag gag gcc aac cag gca ccg ccg gag gag cag      336
Thr Gly Asp Arg Gly Lys Glu Ala Asn Gln Ala Pro Pro Glu Glu Gln
          100          105          110

gac aag gtg ccc cac aca gca cag agg gca gca tgg aag agc cca gaa      384
Asp Lys Val Pro His Thr Ala Gln Arg Ala Ala Trp Lys Ser Pro Glu
          115          120          125

aaa gag aaa acc atg gtg aac aca ctg tca ccc aga gga tcc gag gcc      432
Lys Glu Lys Thr Met Val Asn Thr Leu Ser Pro Arg Gly Ser Glu Ala
          130          135          140

tcc ttc cag gtg tgg aac aag gac agc tct tcc aaa aac ctt atc cct      480
Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu Ile Pro
          145          150          155          160

agg ctg caa aag atc tgg aag aat tac cta agc atg aac aag tac aaa      528
Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys Tyr Lys
          165          170          175

gtg tcc tac aag ggg cca gga cca ggc atc aag ttc agt gca gag gcc      576
Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala Glu Ala
          180          185          190

ctg cgc tgc cac ctc cgg gac cat gtg aat gta tcc atg gta gag gtc      624
Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met Val Glu Val
          195          200          205

aca gat ttt ccc ttc aat acc tct gaa tgg gag ggt tat ctg ccc aag      672
Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr Leu Pro Lys
          210          215          220

gag agc att agg acc aag gct ggg cct tgg ggc agg tgt gct gtt gtg      720
Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys Ala Val Val
          225          230          235          240

```

ES 2 368 267 T3

tcg tca gcg gga tct ctg aag tcc tcc caa cta ggc aga gaa atc gat Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg Glu Ile Asp 245 250 255	768
gat cat gac gca gtc ctg agg ttt aat ggg gca ccc aca gcc aac ttc Asp His Asp Ala Val Leu Arg Phe Asn Gly Ala Pro Thr Ala Asn Phe 260 265 270	816
caa caa gat gtg ggc aca aaa act acc att cgc ctg atg aac tct cag Gln Gln Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met Asn Ser Gln 275 280 285	864
ttg gtt acc aca gag aag cgc ttc ctc aaa gac agt ttg tac aat gaa Leu Val Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu Tyr Asn Glu 290 295 300	912
gga atc cta att gta tgg gac cca tct gta tac cac tca gat atc cca Gly Ile Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser Asp Ile Pro 305 310 315 320	960
aag tgg tac cag aat ccg gat tat aat ttc ttt aac aac tac aag act Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn Tyr Lys Thr 325 330 335	1008
tat cgt aag ctg cac ccc aat cag ccc ttt tac atc ctc aag ccc cag Tyr Arg Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu Lys Pro Gln 340 345 350	1056
atg cct tgg gag cta tgg gac att ctt caa gaa atc tcc cca gaa gag Met Pro Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser Pro Glu Glu 355 360 365	1104
att cag cca aac ccc cca tcc tct ggg atg ctt ggt atc atc atc atg Ile Gln Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile Ile Ile Met 370 375 380	1152
atg acg ctg tgt gac cag gtg gat att tat gag ttc ctc cca tcc aag Met Thr Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu Pro Ser Lys 385 390 395 400	1200
cgc aag act gac gtg tgc tac tac tac cag aag ttc ttc gat agt gcc Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp Ser Ala 405 410 415	1248
tgc acg atg ggt gcc tac cac ccg ctg ctc tat gag aag aat ttg gtg Cys Thr Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys Asn Leu Val 420 425 430	1296
aag cat ctc aac cag ggc aca gat gag gac atc tac ctg ctt gga aaa Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu Leu Gly Lys 435 440 445	1344
gcc aca ctg cct ggc ttc cgg acc att cac tgc Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys 450 455	1377
<210> 164	
<211> 459	
<212> PRT	
<213> Artificial	

ES 2 368 267 T3

<220>

<223> Constructo sintético

<400> 164

Met	Arg	Ser	Cys	Leu	Trp	Arg	Cys	Arg	His	Leu	Ser	Gln	Gly	Val	Gln
1				5					10					15	

Trp	Ser	Leu	Leu	Leu	Ala	Val	Leu	Val	Phe	Phe	Leu	Phe	Ala	Leu	Pro
		20						25					30		

Ser	Phe	Ile	Lys	Glu	Pro	Gln	Thr	Lys	Pro	Ser	Arg	His	Gln	Arg	Thr
	35						40					45			

Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser	Leu	Ala	Lys	Pro	Lys	Ser
50						55					60				

Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr	Ile	Tyr	Ala	Glu	Pro	Val
65					70					75					80

Pro	Glu	Asn	Asn	Ala	Leu	Asn	Thr	Gln	Thr	Gln	Pro	Lys	Ala	His	Thr
				85					90					95	

Thr	Gly	Asp	Arg	Gly	Lys	Glu	Ala	Asn	Gln	Ala	Pro	Pro	Glu	Glu	Gln
			100						105					110	

Asp	Lys	Val	Pro	His	Thr	Ala	Gln	Arg	Ala	Ala	Trp	Lys	Ser	Pro	Glu
		115						120					125		

Lys	Glu	Lys	Thr	Met	Val	Asn	Thr	Leu	Ser	Pro	Arg	Gly	Ser	Glu	Ala
	130						135					140			

Ser	Phe	Gln	Val	Trp	Asn	Lys	Asp	Ser	Ser	Ser	Lys	Asn	Leu	Ile	Pro
145						150					155				160

Arg	Leu	Gln	Lys	Ile	Trp	Lys	Asn	Tyr	Leu	Ser	Met	Asn	Lys	Tyr	Lys
				165						170					175

Val	Ser	Tyr	Lys	Gly	Pro	Gly	Pro	Gly	Ile	Lys	Phe	Ser	Ala	Glu	Ala
			180						185					190	

Leu	Arg	Cys	His	Leu	Arg	Asp	His	Val	Asn	Val	Ser	Met	Val	Glu	Val
		195						200				205			

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

ES 2 368 267 T3

Glu 225	Ser	Ile	Arg	Thr	Lys 230	Ala	Gly	Pro	Trp	Gly 235	Arg	Cys	Ala	Val	Val 240
Ser	Ser	Ala	Gly	Ser	Leu	Lys	Ser	Ser	Gln	Leu	Gly	Arg	Glu	Ile	Asp 255
Asp	His	Asp	Ala	Val	Leu	Arg	Phe	Asn	Gly	Ala	Pro	Thr	Ala	Asn	Phe 270
Gln	Gln	Asp	Val	Gly	Thr	Lys	Thr	Thr	Ile	Arg	Leu	Met	Asn	Ser	Gln 285
Leu	Val	Thr	Thr	Glu	Lys	Arg	Phe	Leu	Lys	Asp	Ser	Leu	Tyr	Asn	Glu 290
Gly	Ile	Leu	Ile	Val	Trp	Asp	Pro	Ser	Val	Tyr	His	Ser	Asp	Ile	Pro 305
Lys	Trp	Tyr	Gln	Asn	Pro	Asp	Tyr	Asn	Phe	Phe	Asn	Asn	Tyr	Lys	Thr 335
Tyr	Arg	Lys	Leu	His	Pro	Asn	Gln	Pro	Phe	Tyr	Ile	Leu	Lys	Pro	Gln 340
Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Leu	Gln	Glu	Ile	Ser	Pro	Glu	Glu 355
Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	Met	Leu	Gly	Ile	Ile	Ile	Met 370
Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	Tyr	Glu	Phe	Leu	Pro	Ser	Lys 385
Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	Tyr	Gln	Lys	Phe	Phe	Asp	Ser	Ala 405
Cys	Thr	Met	Gly	Ala	Tyr	His	Pro	Leu	Leu	Tyr	Glu	Lys	Asn	Leu	Val 420
Lys	His	Leu	Asn	Gln	Gly	Thr	Asp	Glu	Asp	Ile	Tyr	Leu	Leu	Gly	Lys 435
Ala	Thr	Leu	Pro	Gly	Phe	Arg	Thr	Ile	His	Cys					

ES 2 368 267 T3

```

<210> 165
<211> 1731
<212> ADN
<213> Artificial

<220>
<223> Fusión de la secuencia de nucleótidos SEC ID n°:149 con la
      secuencia de nucleótidos SEC ID n°:43

<220>
<221> CDS
<222> (1)..(1731)

<400> 165
atg agg tcc tgc ctg tgg aga tgc agg cac ctg agc caa ggc gtc cag      48
Met Arg Ser Cys Leu Trp Arg Cys Arg His Leu Ser Gln Gly Val Gln
1          5          10          15

tgg tcc ttg ctt ctg gct gtc ctg gtc ttc ttt ctc ttc gcc ttg ccc      96
Trp Ser Leu Leu Leu Ala Val Leu Val Phe Phe Leu Phe Ala Leu Pro
          20          25          30

tct ttt att aag gag cct caa aca aag cct tcc agg cat caa cgc aca      144
Ser Phe Ile Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr
          35          40          45

gag aac att aaa gaa agg tct cta cag tcc ctg gca aag cct aag tcc      192
Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser Leu Ala Lys Pro Lys Ser
          50          55          60

cag gca ccc aca agg gca agg agg aca acc atc tat gca gag cca gtg      240
Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr Ile Tyr Ala Glu Pro Val
65          70          75          80

cca gag aac aat gcc ctc aac aca caa acc cag ccc aag gcc cac acc      288
Pro Glu Asn Asn Ala Leu Asn Thr Gln Thr Gln Pro Lys Ala His Thr
          85          90          95

acc gga gac aga gga aag gag gcc aac cag gca ccg ccg gag gag cag      336
Thr Gly Asp Arg Gly Lys Glu Ala Asn Gln Ala Pro Pro Glu Glu Gln
          100          105          110

gac aag gtg ccc cac aca gca cag agg gca gca tgg aag agc cca gaa      384
Asp Lys Val Pro His Thr Ala Gln Arg Ala Ala Trp Lys Ser Pro Glu
          115          120          125

aaa gag aaa acc atg gtg aac aca ctg tca ccc aga ggg caa gat gca      432
Lys Glu Lys Thr Met Val Asn Thr Leu Ser Pro Arg Gly Gln Asp Ala
          130          135          140

ggg atg gcc tct ggc agg aca gag gca caa tca tgg aag agc cag gac      480
Gly Met Ala Ser Gly Arg Thr Glu Ala Gln Ser Trp Lys Ser Gln Asp
145          150          155          160

aca aag acg acc caa gga aat ggg ggc cag acc agg aag ctg acg gcc      528
Thr Lys Thr Thr Gln Gly Asn Gly Gly Gln Thr Arg Lys Leu Thr Ala
          165          170          175

tcc agg acg gtg tca gag aag cac cag ggc aaa gcg gca acc aca gcc      576
Ser Arg Thr Val Ser Glu Lys His Gln Gly Lys Ala Ala Thr Thr Ala

```

ES 2 368 267 T3

180	185	190	
aag acg ctc att ccc aaa agt cag cac aga atg ctg gct ccc aca gga Lys Thr Leu Ile Pro Lys Ser Gln His Arg Met Leu Ala Pro Thr Gly 195 200 205			624
gca gtg tca aca agg acg aga cag aaa gga gtg acc aca gca gtc atc Ala Val Ser Thr Arg Thr Arg Gln Lys Gly Val Thr Thr Ala Val Ile 210 215 220			672
cca cct aag gag aag aaa cct cag gcc acc cca ccc cct gcc cct ttc Pro Pro Lys Glu Lys Lys Pro Gln Ala Thr Pro Pro Pro Ala Pro Phe 225 230 235 240			720
cag agc ccc acg acg cag aga aac caa aga ctg aag gcc gcc aac ttc Gln Ser Pro Thr Thr Gln Arg Asn Gln Arg Leu Lys Ala Ala Asn Phe 245 250 255			768
aaa tct tct aga gag gcc tcc ttc cag gtg tgg aac aag gac agc tct Lys Ser Ser Arg Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser 260 265 270			816
tcc aaa aac ctt atc cct agg ctg caa aag atc tgg aag aat tac cta Ser Lys Asn Leu Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu 275 280 285			864
agc atg aac aag tac aaa gtg tcc tac aag ggg cca gga cca ggc atc Ser Met Asn Lys Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile 290 295 300			912
aag ttc agt gca gag gcc ctg cgc tgc cac ctc cgg gac cat gtg aat Lys Phe Ser Ala Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn 305 310 315 320			960
gta tcc atg gta gag gtc aca gat ttt ccc ttc aat acc tct gaa tgg Val Ser Met Val Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp 325 330 335			1008
gag ggt tat ctg ccc aag gag agc att agg acc aag gct ggg cct tgg Glu Gly Tyr Leu Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp 340 345 350			1056
ggc agg tgt gct gtt gtg tcg tca gcg gga tct ctg aag tcc tcc caa Gly Arg Cys Ala Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln 355 360 365			1104
cta ggc aga gaa atc gat gat cat gac gca gtc ctg agg ttt aat ggg Leu Gly Arg Glu Ile Asp Asp His Asp Ala Val Leu Arg Phe Asn Gly 370 375 380			1152
gca ccc aca gcc aac ttc caa caa gat gtg ggc aca aaa act acc att Ala Pro Thr Ala Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile 385 390 395 400			1200
cgc ctg atg aac tct cag ttg gtt acc aca gag aag cgc ttc ctc aaa Arg Leu Met Asn Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys 405 410 415			1248
gac agt ttg tac aat gaa gga atc cta att gta tgg gac cca tct gta Asp Ser Leu Tyr Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val 420 425 430			1296

ES 2 368 267 T3

tac cac tca gat atc cca aag tgg tac cag aat ccg gat tat aat ttc Tyr His Ser Asp Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe 435 440 445	1344
ttt aac aac tac aag act tat cgt aag ctg cac ccc aat cag ccc ttt Phe Asn Asn Tyr Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe 450 455 460	1392
tac atc ctc aag ccc cag atg cct tgg gag cta tgg gac att ctt caa Tyr Ile Leu Lys Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln 465 470 475 480	1440
gaa atc tcc cca gaa gag att cag cca aac ccc cca tcc tct ggg atg Glu Ile Ser Pro Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met 485 490 495	1488
ctt ggt atc atc atc atg atg acg ctg tgt gac cag gtg gat att tat Leu Gly Ile Ile Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr 500 505 510	1536
gag ttc ctc cca tcc aag cgc aag act gac gtg tgc tac tac tac cag Glu Phe Leu Pro Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln 515 520 525	1584
aag ttc ttc gat agt gcc tgc acg atg ggt gcc tac cac ccg ctg ctc Lys Phe Phe Asp Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu 530 535 540	1632
tat gag aag aat ttg gtg aag cat ctc aac cag ggc aca gat gag gac Tyr Glu Lys Asn Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp 545 550 555 560	1680
atc tac ctg ctt gga aaa gcc aca ctg cct ggc ttc cgg acc att cac Ile Tyr Leu Leu Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His 565 570 575	1728
tgc Cys	1731

<210> 166

<211> 577

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 166

Met	Arg	Ser	Cys	Leu	Trp	Arg	Cys	Arg	His	Leu	Ser	Gln	Gly	Val	Gln
1				5					10					15	

Trp	Ser	Leu	Leu	Leu	Ala	Val	Leu	Val	Phe	Phe	Leu	Phe	Ala	Leu	Pro
			20					25					30		

Ser Phe Ile Lys Glu Pro Gln Thr Lys Pro Ser Arg His Gln Arg Thr

ES 2 368 267 T3

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

ES 2 368 267 T3

Ser	Met	Asn	Lys	Tyr	Lys	Val	Ser	Tyr	Lys	Gly	Pro	Gly	Pro	Gly	Ile
	290					295					300				
Lys	Phe	Ser	Ala	Glu	Ala	Leu	Arg	Cys	His	Leu	Arg	Asp	His	Val	Asn
305					310					315					320
Val	Ser	Met	Val	Glu	Val	Thr	Asp	Phe	Pro	Phe	Asn	Thr	Ser	Glu	Trp
				325					330					335	
Glu	Gly	Tyr	Leu	Pro	Lys	Glu	Ser	Ile	Arg	Thr	Lys	Ala	Gly	Pro	Trp
			340					345					350		
Gly	Arg	Cys	Ala	Val	Val	Ser	Ser	Ala	Gly	Ser	Leu	Lys	Ser	Ser	Gln
		355					360					365			
Leu	Gly	Arg	Glu	Ile	Asp	Asp	His	Asp	Ala	Val	Leu	Arg	Phe	Asn	Gly
	370					375					380				
Ala	Pro	Thr	Ala	Asn	Phe	Gln	Gln	Asp	Val	Gly	Thr	Lys	Thr	Thr	Ile
385					390					395					400
Arg	Leu	Met	Asn	Ser	Gln	Leu	Val	Thr	Thr	Glu	Lys	Arg	Phe	Leu	Lys
			405						410					415	
Asp	Ser	Leu	Tyr	Asn	Glu	Gly	Ile	Leu	Ile	Val	Trp	Asp	Pro	Ser	Val
			420					425					430		
Tyr	His	Ser	Asp	Ile	Pro	Lys	Trp	Tyr	Gln	Asn	Pro	Asp	Tyr	Asn	Phe
		435					440					445			
Phe	Asn	Asn	Tyr	Lys	Thr	Tyr	Arg	Lys	Leu	His	Pro	Asn	Gln	Pro	Phe
	450					455					460				
Tyr	Ile	Leu	Lys	Pro	Gln	Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Leu	Gln
465					470					475					480
Glu	Ile	Ser	Pro	Glu	Glu	Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	Met
				485					490					495	
Leu	Gly	Ile	Ile	Ile	Met	Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	Tyr
			500					505					510		
Glu	Phe	Leu	Pro	Ser	Lys	Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	Tyr	Gln
		515					520					525			

ES 2 368 267 T3

Lys	Phe	Phe	Asp	Ser	Ala	Cys	Thr	Met	Gly	Ala	Tyr	His	Pro	Leu	Leu
530						535					540				

Tyr	Glu	Lys	Asn	Leu	Val	Lys	His	Leu	Asn	Gln	Gly	Thr	Asp	Glu	Asp
545					550					555					560

Ile	Tyr	Leu	Leu	Gly	Lys	Ala	Thr	Leu	Pro	Gly	Phe	Arg	Thr	Ile	His
				565					570					575	

Cys

<210> 167
 <211> 1095
 <212> ADN
 <213> Artificial

<220>
 <223> Fusión de la secuencia de nucleótidos SEC ID n°: 151 con la
 secuencia de nucleótidos SEC ID n°: 43

<220>
 <221> CDS
 <222> (1)..(1095)

<400>	167	
	atg gga ctc ttg gta ttt gtg cgc aat ctg ctg cta gcc ctc tgc ctc	48
	Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu	
	1 5 10 15	
	ttt ctg gta ctg gga ttt ttg tat tat tct gcg tgg aag cta cac tta	96
	Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu	
	20 25 30	
	ctc cag tgg gag gag gac tcc agt aag tat agt cac tct agc gga tcc	144
	Leu Gln Trp Glu Glu Asp Ser Ser Lys Tyr Ser His Ser Ser Gly Ser	
	35 40 45	
	gag gcc tcc ttc cag gtg tgg aac aag gac agc tct tcc aaa aac ctt	192
	Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu	
	50 55 60	
	atc cct agg ctg caa aag atc tgg aag aat tac cta agc atg aac aag	240
	Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys	
	65 70 75 80	
	tac aaa gtg tcc tac aag ggg cca gga cca ggc atc aag ttc agt gca	288
	Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala	
	85 90 95	
	gag gcc ctg cgc tgc cac ctc cgg gac cat gtg aat gta tcc atg gta	336
	Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met Val	
	100 105 110	
	gag gtc aca gat ttt ccc ttc aat acc tct gaa tgg gag ggt tat ctg	384

ES 2 368 267 T3

Glu	Val	Thr	Asp	Phe	Pro	Phe	Asn	Thr	Ser	Glu	Trp	Glu	Gly	Tyr	Leu	
		115					120					125				
ccc	aag	gag	agc	att	agg	acc	aag	gct	ggg	cct	tgg	ggc	agg	tgt	gct	432
Pro	Lys	Glu	Ser	Ile	Arg	Thr	Lys	Ala	Gly	Pro	Trp	Gly	Arg	Cys	Ala	
	130						135				140					
gtt	gtg	tcg	tca	gcg	gga	tct	ctg	aag	tcc	tcc	caa	cta	ggc	aga	gaa	480
Val	Val	Ser	Ser	Ala	Gly	Ser	Leu	Lys	Ser	Ser	Gln	Leu	Gly	Arg	Glu	
145					150					155					160	
atc	gat	gat	cat	gac	gca	gtc	ctg	agg	ttt	aat	ggg	gca	ccc	aca	gcc	528
Ile	Asp	Asp	His	Asp	Ala	Val	Leu	Arg	Phe	Asn	Gly	Ala	Pro	Thr	Ala	
				165					170					175		
aac	ttc	caa	caa	gat	gtg	ggc	aca	aaa	act	acc	att	cgc	ctg	atg	aac	576
Asn	Phe	Gln	Gln	Asp	Val	Gly	Thr	Lys	Thr	Thr	Ile	Arg	Leu	Met	Asn	
			180					185					190			
tct	cag	ttg	gtt	acc	aca	gag	aag	cgc	ttc	ctc	aaa	gac	agt	ttg	tac	624
Ser	Gln	Leu	Val	Thr	Thr	Glu	Lys	Arg	Phe	Leu	Lys	Asp	Ser	Leu	Tyr	
	195						200					205				
aat	gaa	gga	atc	cta	att	gta	tgg	gac	cca	tct	gta	tac	cac	tca	gat	672
Asn	Glu	Gly	Ile	Leu	Ile	Val	Trp	Asp	Pro	Ser	Val	Tyr	His	Ser	Asp	
	210					215					220					
atc	cca	aag	tgg	tac	cag	aat	ccg	gat	tat	aat	ttc	ttt	aac	aac	tac	720
Ile	Pro	Lys	Trp	Tyr	Gln	Asn	Pro	Asp	Tyr	Asn	Phe	Phe	Asn	Asn	Tyr	
225					230					235					240	
aag	act	tat	cgt	aag	ctg	cac	ccc	aat	cag	ccc	ttt	tac	atc	ctc	aag	768
Lys	Thr	Tyr	Arg	Lys	Leu	His	Pro	Asn	Gln	Pro	Phe	Tyr	Ile	Leu	Lys	
				245					250					255		
ccc	cag	atg	cct	tgg	gag	cta	tgg	gac	att	ctt	caa	gaa	atc	tcc	cca	816
Pro	Gln	Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Leu	Gln	Glu	Ile	Ser	Pro	
			260					265					270			
gaa	gag	att	cag	cca	aac	ccc	cca	tcc	tct	ggg	atg	ctt	ggt	atc	atc	864
Glu	Glu	Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	Met	Leu	Gly	Ile	Ile	
	275					280						285				
atc	atg	atg	acg	ctg	tgt	gac	cag	gtg	gat	att	tat	gag	ttc	ctc	cca	912
Ile	Met	Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	Tyr	Glu	Phe	Leu	Pro	
	290					295					300					
tcc	aag	cgc	aag	act	gac	gtg	tgc	tac	tac	tac	cag	aag	ttc	ttc	gat	960
Ser	Lys	Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	Tyr	Gln	Lys	Phe	Phe	Asp	
305					310					315					320	
agt	gcc	tgc	acg	atg	ggt	gcc	tac	cac	ccg	ctg	ctc	tat	gag	aag	aat	1008
Ser	Ala	Cys	Thr	Met	Gly	Ala	Tyr	His	Pro	Leu	Leu	Tyr	Glu	Lys	Asn	
				325					330					335		
ttg	gtg	aag	cat	ctc	aac	cag	ggc	aca	gat	gag	gac	atc	tac	ctg	ctt	1056
Leu	Val	Lys	His	Leu	Asn	Gln	Gly	Thr	Asp	Glu	Asp	Ile	Tyr	Leu	Leu	
			340					345					350			
gga	aaa	gcc	aca	ctg	cct	ggc	ttc	cgg	acc	att	cac	tgc				1095
Gly	Lys	Ala	Thr	Leu	Pro	Gly	Phe	Arg	Thr	Ile	His	Cys				

ES 2 368 267 T3

355

360

365

<210> 168

<211> 365

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 168

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 Gly Ser
35 40 45

Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu
50 55 60

Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys
65 70 75 80

Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala
85 90 95

Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met Val
100 105 110

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

Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys Ala
130 135 140

Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg Glu
145 150 155 160

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

Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met Asn
180 185 190

ES 2 368 267 T3

Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu Tyr
 195 200 205

Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser Asp
 210 215 220

Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn Tyr
 225 230 235 240

Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu Lys
 245 250 255

Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser Pro
 260 265 270

Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile Ile
 275 280 285

Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu Pro
 290 295 300

Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp
 305 310 315 320

Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys Asn
 325 330 335

Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu Leu
 340 345 350

Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys
 355 360 365

<210> 169
 <211> 1095
 <212> ADN
 <213> Artificial

<220>
 <223> Fusión de la secuencia de nucleótidos SEC ID n°:153 con la
 secuencia de nucleótidos SEC ID n°: 43

<220>
 <221> CDS
 <222> (1)..(1095)

<400> 169
 atg gga ctc ttg gta ttt gta cgc aac ctg ctg cta gcc ctc tgc ctc
 Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu

48

ES 2 368 267 T3

1	5	10	15	
ttt ctg gtc ctg gga ttt ttg tat tat tct gcc tgg aag cta cac tta				96
Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu	20	25	30	
ctc caa tgg gaa gac tcc aat tca ctg att ctt tcc ctt gac tct aga				144
Leu Gln Trp Glu Asp Ser Asn Ser Leu Ile Leu Ser Leu Asp Ser Arg	35	40	45	
gag gcc tcc ttc cag gtg tgg aac aag gac agc tct tcc aaa aac ctt				192
Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu	50	55	60	
atc cct agg ctg caa aag atc tgg aag aat tac cta agc atg aac aag				240
Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys	65	70	75	80
tac aaa gtg tcc tac aag ggg cca gga cca ggc atc aag ttc agt gca				288
Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala	85	90	95	
gag gcc ctg cgc tgc cac ctc cgg gac cat gtg aat gta tcc atg gta				336
Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met Val	100	105	110	
gag gtc aca gat ttt ccc ttc aat acc tct gaa tgg gag ggt tat ctg				384
Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr Leu	115	120	125	
ccc aag gag agc att agg acc aag gct ggg cct tgg ggc agg tgt gct				432
Pro Lys Glu Ser Ile Arg Lys Lys Ala Gly Pro Trp Gly Arg Cys Ala	130	135	140	
gtt gtg tgc tca gcg gga tct ctg aag tcc tcc caa cta ggc aga gaa				480
Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg Glu	145	150	155	160
atc gat gat cat gac gca gtc ctg agg ttt aat ggg gca ccc aca gcc				528
Ile Asp Asp His Asp Ala Val Leu Arg Phe Asn Gly Ala Pro Thr Ala	165	170	175	
aac ttc caa caa gat gtg ggc aca aaa act acc att cgc ctg atg aac				576
Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met Asn	180	185	190	
tct cag ttg gtt acc aca gag aag cgc ttc ctc aaa gac agt ttg tac				624
Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu Tyr	195	200	205	
aat gaa gga atc cta att gta tgg gac cca tct gta tac cac tca gat				672
Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser Asp	210	215	220	
atc cca aag tgg tac cag aat ccg gat tat aat ttc ttt aac aac tac				720
Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn Tyr	225	230	235	240
aag act tat cgt aag ctg cac ccc aat cag ccc ttt tac atc ctc aag				768
Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu Lys	245	250	255	

ES 2 368 267 T3

ccc cag atg cct tgg gag cta tgg gac att ctt caa gaa atc tcc cca	816
Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser Pro	
260 265 270	

gaa gag att cag cca aac ccc cca tcc tct ggg atg ctt ggt atc atc	864
Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile Ile	
275 280 285	

atc atg atg acg ctg tgt gac cag gtg gat att tat gag ttc ctc cca	912
Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu Pro	
290 295 300	

tcc aag cgc aag act gac gtg tgc tac tac tac cag aag ttc ttc gat	960
Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe Asp	
305 310 315 320	

agt gcc tgc acg atg ggt gcc tac cac ccg ctg ctc tat gag aag aat	1008
Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys Asn	
325 330 335	

ttg gtg aag cat ctc aac cag ggc aca gat gag gac atc tac ctg ctt	1056
Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu Leu	
340 345 350	

gga aaa gcc aca ctg cct ggc ttc cgg acc att cac tgc	1095
Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys	
355 360 365	

<210> 170
 <211> 365
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 170

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 Asp Ser Asn Ser Leu Ile Leu Ser Leu Asp Ser Arg	
35 40 45	

Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu	
50 55 60	

Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys	
65 70 75 80	

Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala	
---	--

ES 2 368 267 T3

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

ES 2 368 267 T3

Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu Leu
 340 345 350

Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys
 355 360 365

<210> 171
 <211> 1194
 <212> ADN
 <213> Artificial

<220>
 <223> Fusión de la secuencia de nucleótidos SEC ID n°: 155 con la
 secuencia de nucleótidos SEC ID n°: 43

<220>
 <221> CDS
 <222> (1)..(1194)

<400> 171
 atg cag ctg cag ttc cgg agc tgg atg ctg gcc gcg ctc acg ctg ctc 48
 Met Gln Leu Gln Phe Arg Ser Trp Met Leu Ala Ala Leu Thr Leu Leu
 1 5 10 15
 gtg gtc ttc ctc atc ttc gca gac atc tca gag atc gaa gaa gaa atc 96
 Val Val Phe Leu Ile Phe Ala Asp Ile Ser Glu Ile Glu Glu Glu Ile
 20 25 30
 ggg aat tcg gga ggc aga ggt aca atc aga tca gct gtg aac agc tta 144
 Gly Asn Ser Gly Gly Arg Gly Thr Ile Arg Ser Ala Val Asn Ser Leu
 35 40 45
 cat agc aaa tct aat aga gct gaa gtt gta ata aac ggc tcc tca tca 192
 His Ser Lys Ser Asn Arg Ala Glu Val Val Ile Asn Gly Ser Ser Ser
 50 55 60
 cca gct gtt gtt gac aga agt aat gaa agc atc aag cac aac atc aag 240
 Pro Ala Val Val Asp Arg Ser Asn Glu Ser Ile Lys His Asn Ile Lys
 65 70 75 80
 ctt gag gcc tcc ttc cag gtg tgg aac aag gac agc tct tcc aaa aac 288
 Leu Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn
 85 90 95
 ctt atc cct agg ctg caa aag atc tgg aag aat tac cta agc atg aac 336
 Leu Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn
 100 105 110
 aag tac aaa gtg tcc tac aag ggg cca gga cca ggc atc aag ttc agt 384
 Lys Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser
 115 120 125
 gca gag gcc ctg cgc tgc cac ctc cgg gac cat gtg aat gta tcc atg 432
 Ala Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met
 130 135 140

ES 2 368 267 T3

gta gag gtc aca gat ttt ccc ttc aat acc tct gaa tgg gag ggt tat	480
Val Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr	
145 150 155 160	
ctg ccc aag gag agc att agg acc aag gct ggg cct tgg ggc agg tgt	528
Leu Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys	
165 170 175	
gct gtt gtg tcg tca gcg gga tct ctg aag tcc tcc caa cta ggc aga	576
Ala Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg	
180 185 190	
gaa atc gat gat cat gac gca gtc ctg agg ttt aat ggg gca ccc aca	624
Glu Ile Asp Asp His Asp Ala Val Leu Arg Phe Asn Gly Ala Pro Thr	
195 200 205	
gcc aac ttc caa caa gat gtg ggc aca aaa act acc att cgc ctg atg	672
Ala Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met	
210 215 220	
aac tct cag ttg gtt acc aca gag aag cgc ttc ctc aaa gac agt ttg	720
Asn Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu	
225 230 235 240	
tac aat gaa gga atc cta att gta tgg gac cca tct gta tac cac tca	768
Tyr Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser	
245 250 255	
gat atc cca aag tgg tac cag aat ccg gat tat aat ttc ttt aac aac	816
Asp Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn	
260 265 270	
tac aag act tat cgt aag ctg cac ccc aat cag ccc ttt tac atc ctc	864
Tyr Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu	
275 280 285	
aag ccc cag atg cct tgg gag cta tgg gac att ctt caa gaa atc tcc	912
Lys Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser	
290 295 300	
cca gaa gag att cag cca aac ccc cca tcc tct ggg atg ctt ggt atc	960
Pro Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile	
305 310 315 320	
atc atc atg atg acg ctg tgt gac cag gtg gat att tat gag ttc ctc	1008
Ile Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu	
325 330 335	
cca tcc aag cgc aag act gac gtg tgc tac tac tac cag aag ttc ttc	1056
Pro Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe	
340 345 350	
gat agt gcc tgc acg atg ggt gcc tac cac ccg ctg ctc tat gag aag	1104
Asp Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys	
355 360 365	
aat ttg gtg aag cat ctc aac cag ggc aca gat gag gac atc tac ctg	1152
Asn Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu	
370 375 380	
ctt gga aaa gcc aca ctg cct ggc ttc cgg acc att cac tgc	1194

ES 2 368 267 T3

Leu Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys
 385 390 395

<210> 172
 <211> 398
 <212> PRT
 <213> Artificial

<220>
 <223> Constructo sintético

<400> 172

Met Gln Leu Gln Phe Arg Ser Trp Met Leu Ala Ala Leu Thr Leu Leu
 1 5 10 15

Val Val Phe Leu Ile Phe Ala Asp Ile Ser Glu Ile Glu Glu Glu Ile
 20 25 30

Gly Asn Ser Gly Gly Arg Gly Thr Ile Arg Ser Ala Val Asn Ser Leu
 35 40 45

His Ser Lys Ser Asn Arg Ala Glu Val Val Ile Asn Gly Ser Ser Ser
 50 55 60

Pro Ala Val Val Asp Arg Ser Asn Glu Ser Ile Lys His Asn Ile Lys
 65 70 75 80

Leu Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn
 85 90 95

Leu Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn
 100 105 110

Lys Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser
 115 120 125

Ala Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met
 130 135 140

Val Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr
 145 150 155 160

Leu Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys
 165 170 175

Ala Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg
 180 185 190

ES 2 368 267 T3

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

Ala Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met
 210 215 220

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

Tyr Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val Tyr His Ser
 245 250 255

Asp Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe Phe Asn Asn
 260 265 270

Tyr Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe Tyr Ile Leu
 275 280 285

Lys Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln Glu Ile Ser
 290 295 300

Pro Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met Leu Gly Ile
 305 310 315 320

Ile Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr Glu Phe Leu
 325 330 335

Pro Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln Lys Phe Phe
 340 345 350

Asp Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu Tyr Glu Lys
 355 360 365

Asn Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp Ile Tyr Leu
 370 375 380

Leu Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His Cys
 385 390 395

<210> 173

<211> 1158

<212> ADN

<213> Artificial

<220>

<223> Fusión de la secuencia de nucleótidos SEC ID n°: 157 con la
 secuencia de nucleótidos SEC ID n°: 43

ES 2 368 267 T3

<220>

<221> CDS

<222> (1)..(1158)

<400> 173

atg cgc tcc att agg aag agg tgg acg atc tgc aca ata agt ctg ctc	48
Met Arg Ser Ile Arg Lys Arg Trp Thr Ile Cys Thr Ile Ser Leu Leu	
1 5 10 15	
ctg atc ttt tat aag aca aaa gaa ata gca aga act gag gag cac cag	96
Leu Ile Phe Tyr Lys Thr Lys Glu Ile Ala Arg Thr Glu Glu His Gln	
20 25 30	
gag acg caa ctc atc gga gat ggt gaa ttg tct ttg agt cgg tca ctt	144
Glu Thr Gln Leu Ile Gly Asp Gly Glu Leu Ser Leu Ser Arg Ser Leu	
35 40 45	
gtc aat agc tct gat aaa atc att cga aag gct ggc tct tca atc ttc	192
Val Asn Ser Ser Asp Lys Ile Ile Arg Lys Ala Gly Ser Ser Ile Phe	
50 55 60	
cag cac aat aag ctt gag gcc tcc ttc cag gtg tgg aac aag gac agc	240
Gln His Asn Lys Leu Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser	
65 70 75 80	
tct tcc aaa aac ctt atc cct agg ctg caa aag atc tgg aag aat tac	288
Ser Ser Lys Asn Leu Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr	
85 90 95	
cta agc atg aac aag tac aaa gtg tcc tac aag ggg cca gga cca ggc	336
Leu Ser Met Asn Lys Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly	
100 105 110	
atc aag ttc agt gca gag gcc ctg cgc tgc cac ctc cgg gac cat gtg	384
Ile Lys Phe Ser Ala Glu Ala Leu Arg Cys His Leu Arg Asp His Val	
115 120 125	
aat gta tcc atg gta gag gtc aca gat ttt ccc ttc aat acc tct gaa	432
Asn Val Ser Met Val Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu	
130 135 140	
tgg gag ggt tat ctg ccc aag gag agc att agg acc aag gct ggg cct	480
Trp Glu Gly Tyr Leu Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro	
145 150 155 160	
tgg ggc agg tgt gct gtt gtg tgc tca gcg gga tct ctg aag tcc tcc	528
Trp Gly Arg Cys Ala Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser	
165 170 175	
caa cta ggc aga gaa atc gat gat cat gac gca gtc ctg agg ttt aat	576
Gln Leu Gly Arg Glu Ile Asp Asp His Asp Ala Val Leu Arg Phe Asn	
180 185 190	
ggg gca ccc aca gcc aac ttc caa caa gat gtg ggc aca aaa act acc	624
Gly Ala Pro Thr Ala Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr	
195 200 205	
att cgc ctg atg aac tct cag ttg gtt acc aca gag aag cgc ttc ctc	672
Ile Arg Leu Met Asn Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu	

ES 2 368 267 T3

210	215	220	
aaa gac agt ttg tac aat gaa gga atc cta att gta tgg gac cca tct			720
Lys Asp Ser Leu Tyr Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser			
225	230	235	240
gta tac cac tca gat atc cca aag tgg tac cag aat ccg gat tat aat			768
Val Tyr His Ser Asp Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn			
	245	250	255
ttc ttt aac aac tac aag act tat cgt aag ctg cac ccc aat cag ccc			816
Phe Phe Asn Asn Tyr Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro			
	260	265	270
ttt tac atc ctc aag ccc cag atg cct tgg gag cta tgg gac att ctt			864
Phe Tyr Ile Leu Lys Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu			
	275	280	285
caa gaa atc tcc cca gaa gag att cag cca aac ccc cca tcc tct ggg			912
Gln Glu Ile Ser Pro Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly			
	290	295	300
atg ctt ggt atc atc atc atg atg acg ctg tgt gac cag gtg gat att			960
Met Leu Gly Ile Ile Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile			
305	310	315	320
tat gag ttc ctc cca tcc aag cgc aag act gac gtg tgc tac tac tac			1008
Tyr Glu Phe Leu Pro Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr			
	325	330	335
cag aag ttc ttc gat agt gcc tgc acg atg ggt gcc tac cac ccg ctg			1056
Gln Lys Phe Phe Asp Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu			
	340	345	350
ctc tat gag aag aat ttg gtg aag cat ctc aac cag ggc aca gat gag			1104
Leu Tyr Glu Lys Asn Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu			
	355	360	365
gac atc tac ctg ctt gga aaa gcc aca ctg cct ggc ttc cgg acc att			1152
Asp Ile Tyr Leu Leu Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile			
	370	375	380
cac tgc			1158
His Cys			
385			

<210> 174

<211> 386

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 174

Met Arg Ser Ile Arg Lys Arg Trp Thr Ile Cys Thr Ile Ser Leu Leu
1 5 10 15

ES 2 368 267 T3

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

ES 2 368 267 T3

260							265					270				
Phe	Tyr	Ile	Leu	Lys	Pro	Gln	Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Leu	
		275				280						285				
Gln	Glu	Ile	Ser	Pro	Glu	Glu	Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	
290						295				300						
Met	Leu	Gly	Ile	Ile	Ile	Met	Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	
305					310				315						320	
Tyr	Glu	Phe	Leu	Pro	Ser	Lys	Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	Tyr	
				325				330						335		
Gln	Lys	Phe	Phe	Asp	Ser	Ala	Cys	Thr	Met	Gly	Ala	Tyr	His	Pro	Leu	
		340						345				350				
Leu	Tyr	Glu	Lys	Asn	Leu	Val	Lys	His	Leu	Asn	Gln	Gly	Thr	Asp	Glu	
		355				360						365				
Asp	Ile	Tyr	Leu	Leu	Gly	Lys	Ala	Thr	Leu	Pro	Gly	Phe	Arg	Thr	Ile	
370						375				380						
His	Cys															
385																

<210>	175
<211>	1155
<212>	ADN
<213>	Artificial

<220>
<223> Fusión de la secuencia de nucleótidos SEC ID n°: 159 con la
secuencia de nucleótidos SEC ID n°: 43

```
<220>
<221> CDS
<222> (1) .. (1155)
```

<400>	175																	
atg gga ctc ttg gta ttt gtg cgc aat ctg ctg cta gcc ctc tgc ctc																		48
Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu																		
1				5					10					15				
ttt ctg gta ctg gga ttt ttg tat tat tct gcg tgg gag cct caa aca																		96
Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Glu Pro Gln Thr																		
			20					25						30				
aag cct tcc agg cat caa cgc aca gag aac att aaa gaa agg tct cta																		144
Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile Lys Glu Arg Ser Leu																		
			35				40						45					

ES 2 368 267 T3

cag tcc ctg gca aag cct aag tcc cag gca ccc aca agg gca agg agg Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro Thr Arg Ala Arg Arg 50 55 60	192
aca acc gga tcc gag gcc tcc ttc cag gtg tgg aac aag gac agc tct Thr Thr Gly Ser Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser 65 70 75 80	240
tcc aaa aac ctt atc cct agg ctg caa aag atc tgg aag aat tac cta Ser Lys Asn Leu Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu 85 90 95	288
agc atg aac aag tac aaa gtg tcc tac aag ggg cca gga cca ggc atc Ser Met Asn Lys Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile 100 105 110	336
aag ttc agt gca gag gcc ctg cgc tgc cac ctc cgg gac cat gtg aat Lys Phe Ser Ala Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn 115 120 125	384
gta tcc atg gta gag gtc aca gat ttt ccc ttc aat acc tct gaa tgg Val Ser Met Val Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp 130 135 140	432
gag ggt tat ctg ccc aag gag agc att agg acc aag gct ggg cct tgg Glu Gly Tyr Leu Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp 145 150 155 160	480
ggc agg tgt gct gtt gtg tgc tca gcg gga tct ctg aag tcc tcc caa Gly Arg Cys Ala Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln 165 170 175	528
cta ggc aga gaa atc gat gat cat gac gca gtc ctg agg ttt aat ggg Leu Gly Arg Glu Ile Asp Asp His Asp Ala Val Leu Arg Phe Asn Gly 180 185 190	576
gca ccc aca gcc aac ttc caa caa gat gtg ggc aca aaa act acc att Ala Pro Thr Ala Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile 195 200 205	624
cgc ctg atg aac tct cag ttg gtt acc aca gag aag cgc ttc ctc aaa Arg Leu Met Asn Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys 210 215 220	672
gac agt ttg tac aat gaa gga atc cta att gta tgg gac cca tct gta Asp Ser Leu Tyr Asn Glu Gly Ile Leu Ile Val Trp Asp Pro Ser Val 225 230 235 240	720
tac cac tca gat atc cca aag tgg tac cag aat ccg gat tat aat ttc Tyr His Ser Asp Ile Pro Lys Trp Tyr Gln Asn Pro Asp Tyr Asn Phe 245 250 255	768
ttt aac aac tac aag act tat cgt aag ctg cac ccc aat cag ccc ttt Phe Asn Asn Tyr Lys Thr Tyr Arg Lys Leu His Pro Asn Gln Pro Phe 260 265 270	816
tac atc ctc aag ccc cag atg cct tgg gag cta tgg gac att ctt caa Tyr Ile Leu Lys Pro Gln Met Pro Trp Glu Leu Trp Asp Ile Leu Gln 275 280 285	864

ES 2 368 267 T3

gaa atc tcc cca gaa gag att cag cca aac ccc cca tcc tct ggg atg Glu Ile Ser Pro Glu Glu Ile Gln Pro Asn Pro Pro Ser Ser Gly Met 290 295 300	912
ctt ggt atc atc atc atg atg acg ctg tgt gac cag gtg gat att tat Leu Gly Ile Ile Ile Met Met Thr Leu Cys Asp Gln Val Asp Ile Tyr 305 310 315 320	960
gag ttc ctc cca tcc aag cgc aag act gac gtg tgc tac tac tac cag Glu Phe Leu Pro Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr Tyr Gln 325 330 335	1008
aag ttc ttc gat agt gcc tgc acg atg ggt gcc tac cac ccg ctg ctc Lys Phe Phe Asp Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu 340 345 350	1056
tat gag aag aat ttg gtg aag cat ctc aac cag ggc aca gat gag gac Tyr Glu Lys Asn Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp 355 360 365	1104
atc tac ctg ctt gga aaa gcc aca ctg cct ggc ttc cgg acc att cac Ile Tyr Leu Leu Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His 370 375 380	1152
tgc Cys 385	1155
<210> 176 <211> 385 <212> PRT <213> Artificial	
<220> <223> Constructo sintético	
<400> 176	
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 Glu Pro Gln Thr 20 25 30	
Lys Pro Ser Arg His Gln Arg Thr Glu Asn Ile Lys Glu Arg Ser Leu 35 40 45	
Gln Ser Leu Ala Lys Pro Lys Ser Gln Ala Pro Thr Arg Ala Arg Arg 50 55 60	
Thr Thr Gly Ser Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser 65 70 75 80	
Ser Lys Asn Leu Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu 85 90 95	

ES 2 368 267 T3

Ser	Met	Asn	Lys	Tyr	Lys	Val	Ser	Tyr	Lys	Gly	Pro	Gly	Pro	Gly	Ile	100	105	110
Lys	Phe	Ser	Ala	Glu	Ala	Leu	Arg	Cys	His	Leu	Arg	Asp	His	Val	Asn	115	120	125
Val	Ser	Met	Val	Glu	Val	Thr	Asp	Phe	Pro	Phe	Asn	Thr	Ser	Glu	Trp	130	135	140
Glu	Gly	Tyr	Leu	Pro	Lys	Glu	Ser	Ile	Arg	Thr	Lys	Ala	Gly	Pro	Trp	145	150	155
Gly	Arg	Cys	Ala	Val	Val	Ser	Ser	Ala	Gly	Ser	Leu	Lys	Ser	Ser	Gln	165	170	175
Leu	Gly	Arg	Glu	Ile	Asp	Asp	His	Asp	Ala	Val	Leu	Arg	Phe	Asn	Gly	180	185	190
Ala	Pro	Thr	Ala	Asn	Phe	Gln	Gln	Asp	Val	Gly	Thr	Lys	Thr	Thr	Ile	195	200	205
Arg	Leu	Met	Asn	Ser	Gln	Leu	Val	Thr	Thr	Glu	Lys	Arg	Phe	Leu	Lys	210	215	220
Asp	Ser	Leu	Tyr	Asn	Glu	Gly	Ile	Leu	Ile	Val	Trp	Asp	Pro	Ser	Val	225	230	235
Tyr	His	Ser	Asp	Ile	Pro	Lys	Trp	Tyr	Gln	Asn	Pro	Asp	Tyr	Asn	Phe	245	250	255
Phe	Asn	Asn	Tyr	Lys	Thr	Tyr	Arg	Lys	Leu	His	Pro	Asn	Gln	Pro	Phe	260	265	270
Tyr	Ile	Leu	Lys	Pro	Gln	Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Leu	Gln	275	280	285
Glu	Ile	Ser	Pro	Glu	Glu	Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	Met	290	295	300
Leu	Gly	Ile	Ile	Ile	Met	Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	Tyr	305	310	315
Glu	Phe	Leu	Pro	Ser	Lys	Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	Tyr	Gln	325	330	335

ES 2 368 267 T3

Lys Phe Phe Asp Ser Ala Cys Thr Met Gly Ala Tyr His Pro Leu Leu
340 345 350

Tyr Glu Lys Asn Leu Val Lys His Leu Asn Gln Gly Thr Asp Glu Asp
355 360 365

Ile Tyr Leu Leu Gly Lys Ala Thr Leu Pro Gly Phe Arg Thr Ile His
370 375 380

Cys
385

<210> 177

<211> 84

<212> ADN

<213> Artificial

<220>

<223> Fusión de la secuencia de nucleótidos SEC ID n°: 65 con la
secuencia de nucleótidos SEC ID n°: 107

<220>

<221> CDS

<222> (1)..(84)

<400> 177

atg gga ctc ttg gta ttt gtg cgc aat ctg ctg cta gcc ctc tgc ctc	48
Met Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu	
1 5 10 15	

ttt ctg gta ctg gga ttt ttg tat tat tct gcg tgg	84
Phe Leu Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp	
1 20 25	

<210> 178

<211> 28

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 178

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	
20 25	

<210> 179

<211> 192

<212> ADN

ES 2 368 267 T3

<213> Artificial

<220>

<223> Fusión de la secuencia de nucleótidos SEC ID n°: 65 con la secuencia de nucleótidos SEC ID n°:119 y con la secuencia de nucleótidos SEC ID n°:129

<220>

<221> CDS

<222> (1)..(192)

<400> 179

atg	gga	ctc	ttg	gta	ttt	gtg	cgc	agc	tgg	atg	ctg	gcc	gcg	ctc	acg	48
Met	Gly	Leu	Leu	Val	Phe	Val	Arg	Ser	Trp	Met	Leu	Ala	Ala	Leu	Thr	
1				5					10					15		

ctg	ctc	gtg	gtc	ttc	ctc	atc	ttc	gca	aag	gag	cct	caa	aca	aag	cct	96
Leu	Leu	Val	Val	Phe	Leu	Ile	Phe	Ala	Lys	Glu	Pro	Gln	Thr	Lys	Pro	
		20						25					30			

tcc	agg	cat	caa	cgc	aca	gag	aac	att	aaa	gaa	agg	tct	cta	cag	tcc	144
Ser	Arg	His	Gln	Arg	Thr	Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser	
		35					40					45				

ctg	gca	aag	cct	aag	tcc	cag	gca	ccc	aca	agg	gca	agg	agg	aca	acc	192
Leu	Ala	Lys	Pro	Lys	Ser	Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr	
	50					55				60						

<210> 180

<211> 64

<212> PRT

<213> Artificial

<220>

<223> Constructo sintético

<400> 180

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

Leu	Leu	Val	Val	Phe	Leu	Ile	Phe	Ala	Lys	Glu	Pro	Gln	Thr	Lys	Pro
		20						25					30		

Ser	Arg	His	Gln	Arg	Thr	Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser
		35					40					45			

Leu	Ala	Lys	Pro	Lys	Ser	Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr
	50					55					60				

<210> 181

<211> 1143

<212> ADN

<213> Artificial

<220>

<223> Fusión de la secuencia de nucleótidos SEC ID n°: 179 con la secuencia de nucleótidos SEC ID n°: 43

ES 2 368 267 T3

<220>

<221> CDS

<222> (1)..(1143)

<400> 181

atg gga ctc ttg gta ttt gtg cgc agc tgg atg ctg gcc gcg ctc acg	48
Met Gly Leu Leu Val Phe Val Arg Ser Trp Met Leu Ala Ala Leu Thr	
1 5 10 15	

ctg ctc gtg gtc ttc ctc atc ttc gca aag gag cct caa aca aag cct	96
Leu Leu Val Val Phe Leu Ile Phe Ala Lys Glu Pro Gln Thr Lys Pro	
20 25 30	

tcc agg cat caa cgc aca gag aac att aaa gaa agg tct cta cag tcc	144
Ser Arg His Gln Arg Thr Glu Asn Ile Lys Glu Arg Ser Leu Gln Ser	
35 40 45	

ctg gca aag cct aag tcc cag gca ccc aca agg gca agg agg aca acc	192
Leu Ala Lys Pro Lys Ser Gln Ala Pro Thr Arg Ala Arg Arg Thr Thr	
50 55 60	

gag gcc tcc ttc cag gtg tgg aac aag gac agc tct tcc aaa aac ctt	240
Glu Ala Ser Phe Gln Val Trp Asn Lys Asp Ser Ser Ser Lys Asn Leu	
65 70 75 80	

atc cct agg ctg caa aag atc tgg aag aat tac cta agc atg aac aag	288
Ile Pro Arg Leu Gln Lys Ile Trp Lys Asn Tyr Leu Ser Met Asn Lys	
85 90 95	

tac aaa gtg tcc tac aag ggg cca gga cca ggc atc aag ttc agt gca	336
Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro Gly Ile Lys Phe Ser Ala	
100 105 110	

gag gcc ctg cgc tgc cac ctc cgg gac cat gtg aat gta tcc atg gta	384
Glu Ala Leu Arg Cys His Leu Arg Asp His Val Asn Val Ser Met Val	
115 120 125	

gag gtc aca gat ttt ccc ttc aat acc tct gaa tgg gag ggt tat ctg	432
Glu Val Thr Asp Phe Pro Phe Asn Thr Ser Glu Trp Glu Gly Tyr Leu	
130 135 140	

ccc aag gag agc att agg acc aag gct ggg cct tgg ggc agg tgt gct	480
Pro Lys Glu Ser Ile Arg Thr Lys Ala Gly Pro Trp Gly Arg Cys Ala	
145 150 155 160	

gtt gtg tcg tca gcg gga tct ctg aag tcc tcc caa cta ggc aga gaa	528
Val Val Ser Ser Ala Gly Ser Leu Lys Ser Ser Gln Leu Gly Arg Glu	
165 170 175	

atc gat gat cat gac gca gtc ctg agg ttt aat ggg gca ccc aca gcc	576
Ile Asp Asp His Asp Ala Val Leu Arg Phe Asn Gly Ala Pro Thr Ala	
180 185 190	

aac ttc caa caa gat gtg ggc aca aaa act acc att cgc ctg atg aac	624
Asn Phe Gln Gln Asp Val Gly Thr Lys Thr Thr Ile Arg Leu Met Asn	
195 200 205	

tct cag ttg gtt acc aca gag aag cgc ttc ctc aaa gac agt ttg tac	672
Ser Gln Leu Val Thr Thr Glu Lys Arg Phe Leu Lys Asp Ser Leu Tyr	
210 215 220	

ES 2 368 267 T3

aat Asn 225	gaa Glu	gga Gly	atc Ile	cta Leu	att Ile	gta Val	tgg Trp	gac Asp	cca Pro	tct Ser	gta Val	tac Tyr	cac His	tca Ser	gat Asp 240	720
atc Ile	cca Pro	aag Lys	tgg Trp	tac Tyr 245	cag Gln	aat Asn	ccg Pro	gat Asp	tat Tyr 250	aat Asn	ttc Phe	ttt Phe	aac Asn	aac Asn	tac Tyr 255	768
aag Lys	act Thr	tat Tyr	cgt Arg 260	aag Lys	ctg Leu	cac His	ccc Pro	aat Asn 265	cag Gln	ccc Pro	ttt Phe	tac Tyr	atc Ile 270	ctc Leu	aag Lys	816
ccc Pro	cag Gln	atg Met 275	cct Pro	tgg Trp	gag Glu	cta Leu	tgg Trp 280	gac Asp	att Ile	ctt Leu	caa Gln	gaa Glu 285	atc Ile	tcc Ser	cca Pro	864
gaa Glu 290	gag Glu	att Ile	cag Gln	cca Pro	aac Asn	ccc Pro 295	cca Pro	tcc Ser	tct Ser	ggg Gly 300	atg Met	ctt Leu	ggt Gly	atc Ile	atc Ile	912
atc Ile 305	atg Met	atg Met	acg Thr	ctg Leu	tgt Cys 310	gac Asp	cag Gln	gtg Val	gat Asp	att Ile 315	tat Tyr	gag Glu	ttc Phe	ctc Leu	cca Pro 320	960
tcc Ser	aag Lys	cgc Arg	aag Lys	act Thr 325	gac Asp	gtg Val	tgc Cys	tac Tyr	tac Tyr 330	tac Tyr	cag Gln	aag Lys	ttc Phe	ttc Phe 335	gat Asp	1008
agt Ser	gcc Ala	tgc Cys	acg Thr	atg Met	ggg Gly	gcc Ala	tac Tyr	cac His	ccg Pro	ctg Leu	ctc Leu	tat Tyr	gag Glu	aag Lys	aat Asn	1056
ttg Leu	gtg Val	aag Lys 355	cat His	ctc Leu	aac Asn	cag Gln	ggc Gly 360	aca Thr	gat Asp	gag Glu	gac Asp	atc Ile 365	tac Tyr	ctg Leu	ctt Leu	1104
gga Gly	aaa Lys 370	gcc Ala	aca Thr	ctg Leu	cct Pro	ggc Gly 375	ttc Phe	cgg Arg	acc Thr	att Ile	cac His	tgc Cys				1143

```
<210> 182
<211> 381
<212> PRT
<213> Artificial
```

<220>
<223> Constructo sintético

<400> 182

Met	Gly	Leu	Leu	Val	Phe	Val	Arg	Ser	Trp	Met	Leu	Ala	Ala	Leu	Thr
1				5					10					15	
Leu	Leu	Val	Val	Phe	Leu	Ile	Phe	Ala	Lys	Glu	Pro	Gln	Thr	Lys	Pro
			20					25					30		
Ser	Arg	His	Gln	Arg	Thr	Glu	Asn	Ile	Lys	Glu	Arg	Ser	Leu	Gln	Ser
		35					40					45			
Leu	Ala	Lys	Pro	Lys	Ser	Gln	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Thr	Thr

ES 2 368 267 T3

50		55		60											
Glu	Ala	Ser	Phe	Gln	Val	Trp	Asn	Lys	Asp	Ser	Ser	Ser	Lys	Asn	Leu
65					70					75				80	
Ile	Pro	Arg	Leu	Gln	Lys	Ile	Trp	Lys	Asn	Tyr	Leu	Ser	Met	Asn	Lys
			85						90					95	
Tyr	Lys	Val	Ser	Tyr	Lys	Gly	Pro	Gly	Pro	Gly	Ile	Lys	Phe	Ser	Ala
			100					105					110		
Glu	Ala	Leu	Arg	Cys	His	Leu	Arg	Asp	His	Val	Asn	Val	Ser	Met	Val
		115					120					125			
Glu	Val	Thr	Asp	Phe	Pro	Phe	Asn	Thr	Ser	Glu	Trp	Glu	Gly	Tyr	Leu
	130						135				140				
Pro	Lys	Glu	Ser	Ile	Arg	Thr	Lys	Ala	Gly	Pro	Trp	Gly	Arg	Cys	Ala
145					150					155					160
Val	Val	Ser	Ser	Ala	Gly	Ser	Leu	Lys	Ser	Ser	Gln	Leu	Gly	Arg	Glu
				165					170					175	
Ile	Asp	Asp	His	Asp	Ala	Val	Leu	Arg	Phe	Asn	Gly	Ala	Pro	Thr	Ala
			180					185					190		
Asn	Phe	Gln	Gln	Asp	Val	Gly	Thr	Lys	Thr	Thr	Ile	Arg	Leu	Met	Asn
		195					200					205			
Ser	Gln	Leu	Val	Thr	Thr	Glu	Lys	Arg	Phe	Leu	Lys	Asp	Ser	Leu	Tyr
	210					215					220				
Asn	Glu	Gly	Ile	Leu	Ile	Val	Trp	Asp	Pro	Ser	Val	Tyr	His	Ser	Asp
225					230				235						240
Ile	Pro	Lys	Trp	Tyr	Gln	Asn	Pro	Asp	Tyr	Asn	Phe	Phe	Asn	Asn	Tyr
			245						250					255	
Lys	Thr	Tyr	Arg	Lys	Leu	His	Pro	Asn	Gln	Pro	Phe	Tyr	Ile	Leu	Lys
			260					265					270		
Pro	Gln	Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Leu	Gln	Glu	Ile	Ser	Pro
		275					280					285			
Glu	Glu	Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	Met	Leu	Gly	Ile	Ile
	290					295					300				
Ile	Met	Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	Tyr	Glu	Phe	Leu	Pro
305					310					315					320
Ser	Lys	Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	Tyr	Gln	Lys	Phe	Phe	Asp
				325					330					335	
Ser	Ala	Cys	Thr	Met	Gly	Ala	Tyr	His	Pro	Leu	Leu	Tyr	Glu	Lys	Asn
			340					345					350		
Leu	Val	Lys	His	Leu	Asn	Gln	Gly	Thr	Asp	Glu	Asp	Ile	Tyr	Leu	Leu
		355					360					365			
Gly	Lys	Ala	Thr	Leu	Pro	Gly	Phe	Arg	Thr	Ile	His	Cys			
	370					375					380				

REIVINDICACIONES

1. Procedimiento para la producción de secuencias génicas codificantes de glucosiltransferasas membranales quiméricas que presentan una actividad de glucosilación optimizada en las células transformadas con dichas secuencias, en comparación con la actividad de glucosilación de las glucosiltransferasas nativas correspondientes hacia el sustrato aceptor, siendo dicha actividad de glucosilación optimizada por lo menos 30 veces superior a la actividad inicial de las glucosiltransferasas de longitud completa nativas,

comprendiendo dicho procedimiento la fusión:

- de un primer ácido nucleico codificante de un fragmento mínimo C-terminal del dominio catalítico (CD) de la glucosiltransferasa de longitud completa nativa,

siendo dicha primera secuencia de ácidos nucleicos obtenida mediante la eliminación de nucleótidos codificantes de uno o varios aminoácidos contiguos que se extienden desde el primer aminoácido del extremo N-terminal de dicho CD, y la selección del ácido nucleico codificante de dicho fragmento de CD mínimo C-terminal de manera que si n representa el número de aminoácidos contiguos tal como se ha definido anteriormente, que han sido deletados, el fragmento obtenido al deletar por lo menos n+1 aminoácidos contiguos, tal como se ha definido anteriormente, no presenta ninguna actividad sustancial de transferasa,

- con un segundo ácido nucleico de secuencia variable codificante de una cadena peptídica transmembranal que especifica el anclaje de las glucosiltransferasas en compartimentos intracelulares, y que comprende en su región N-terminal una región de cola citoplasmática (CT) situada cadena arriba de un dominio transmembranal (TMD) situado cadena arriba de una región de tallo (SR) o de un fragmento de por lo menos 3 aminoácidos contiguos de la SR, estando dicha SR o parte de la misma unida a dicho dominio catalítico mediante un conector o péptido de conexión de por lo menos 2 aminoácidos codificados por un sitio de restricción que no existe en la secuencia de nucleótidos codificante del CD nativo mencionado anteriormente,

con la condición de que por lo menos uno de estos péptidos CT, TMD y SR sea diferente de la estructura primaria de las contrapartidas peptídicas naturales presentes en la glucosiltransferasa nativa a partir de la que se obtiene el fragmento de CD con actividad de glucosiltransferasa óptima tal como se ha definido anteriormente,

siendo dicha fusión realizada de manera que el primer ácido nucleico se encuentre situado cadena abajo del segundo ácido nucleico y proporcione un producto proteína en el que el CD se encuentre en la mitad C-terminal.

2. Procedimiento según la reivindicación 1, caracterizado porque los primer y segundo ácidos nucleicos se obtienen de las secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en glucosiltransferasas de origen eucariótico, preferentemente mamíferos o seres humanos, estando dichas glucosiltransferasas implicadas en:

- la O-glucosilación de las proteínas en las células eucarióticas, tales como las N-acetilgalactosaminiltransferasas, N-acetilglucosaminiltransferasas, fucosiltransferasas, galactosiltransferasas o sialiltransferasas,
- la N-glucosilación de las proteínas en las células eucarióticas, tales como glucosaminiltransferasas, galactosiltransferasas, fucosiltransferasas o sialiltransferasas,
- la glucosilación de los lípidos, tales como glucosiltransferasas, N-acetilgalactosaminiltransferasas, glucosaminiltransferasas, fucosiltransferasas, galactosiltransferasas o sialiltransferasas.

3. Procedimiento según la reivindicación 1 ó 2, caracterizado porque los primer y segundo ácidos nucleicos se obtienen de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en glucosiltransferasas, N-acetilgalactosaminiltransferasas, N-acetilglucosaminiltransferasas, galactosiltransferasas, fucosiltransferasas o sialiltransferasas.

4. Procedimiento según cualquiera de las reivindicaciones 1 a 3, caracterizado porque los primer y segundo ácidos nucleicos se obtienen de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en sialiltransferasas.

5. Procedimiento según cualquiera de las reivindicaciones 1 a 4, caracterizado porque los primer y segundo ácidos nucleicos se obtienen de secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en α 2,6-sialiltransferasas, α 2,3-sialiltransferasas o α 2,8-sialiltransferasas.

6. Procedimiento según cualquiera de las reivindicaciones 1 a 5, caracterizado porque los primer y segundo ácidos nucleicos se obtienen de las secuencias de nucleótidos codificantes de dominios CD, o regiones CT, TMD y SR, respectivamente, en:

- α 2,6-sialiltransferasas seleccionadas de entre:

* las β 1,4-galactósido α 2,6-sialiltransferasas I y II humanas (hST6Gal I y hST6Gal II) representadas por las secuencias SEC ID nº 2 y SEC ID nº 4, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 1 y SEC ID nº 3, respectivamente, o

* las N-acetilgalactosaminida- α 2,6-sialiltransferasas I a VI humanas (hST6GalNAc I a VI) representadas por las secuencias SEC ID nº 6, SEC ID nº 8, SEC ID nº 10, SEC ID nº 12, SEC ID nº 14 y SEC ID nº 16, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 5, SEC ID nº 7, SEC ID nº 9, SEC ID nº 11, SEC ID nº 13 y SEC ID nº 15, respectivamente,

- α 2,3-sialiltransferasas seleccionadas de entre galactósido- α 2,3-sialiltransferasas I a VI humanas (hST3Gal I a VI) representadas por las secuencias SEC ID nº 18, SEC ID nº 20, SEC ID nº 22, SEC ID nº 24, SEC ID nº 26 y SEC ID nº 28, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 17, SEC ID nº 19, SEC ID nº 21, SEC ID nº 23, SEC ID nº 25 y SEC ID nº 27, respectivamente, o las galactósido- α 2,3-sialiltransferasas I a VI de rata (ST3Gal I a VI de rata), tales como rST3Gal III representada por la secuencia SEC ID nº 30 codificada por la secuencia de nucleótidos SEC ID nº 29, o cualquier ST de otro origen animal con la condición de que comparta una homología de por lo menos 85% con el enzima humano,

- α 2,8-sialiltransferasas seleccionadas de entre las ácido siálico- α 2,8-sialiltransferasas I a VI (hST8Sia I a VI) humanas representadas por las secuencias SEC ID nº 32, SEC ID nº 34, SEC ID nº 36, SEC ID nº 38, SEC ID nº 40 y SEC ID nº 42, respectivamente, codificadas por las secuencias de nucleótidos SEC ID nº 31, SEC ID nº 33, SEC ID nº 35, SEC ID nº 37, SEC ID nº 39 y SEC ID nº 41, respectivamente.

7. Procedimiento según cualquiera de las reivindicaciones 1 a 6, caracterizado porque la secuencia de nucleótidos codificante del dominio CD comprendido en el primer ácido nucleico se selecciona de entre las secuencias constituidas por o que comprenden:

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 268 y 330, y en su extremo 3' por el nucleótido situado en la posición 1.218 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6Gal I delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 90 y 110 y en su extremo C-terminal por el aminoácido situado en la posición 406 de la secuencia SEC ID nº 2,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 307 y 369, y en su extremo 3' por el nucleótido situado en la posición 1.587 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6Gal II delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 103 y 123 y en su extremo C-terminal por el aminoácido situado en la posición 529 de la secuencia SEC ID nº 4,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 814 y 876, y en su extremo 3' por el nucleótido situado en la posición 1.800 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNAc I delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 272 y 292 y en su extremo C-terminal por el aminoácido situado en la posición 600 de la secuencia SEC ID nº 6,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 172 y 234, y en su extremo 3' por el nucleótido situado en la posición 1.122 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNAc II delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 58 y 78 y en su extremo C-terminal por el aminoácido situado en la posición 374 de la secuencia SEC ID nº 8,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 73 y 135, y en su extremo 3' por el nucleótido situado en la posición 915 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNAc III delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 25 y 45, y en su extremo C-terminal por el aminoácido situado en la posición 305 de la secuencia SEC ID nº 10,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 61 y 123, y en su extremo 3' por el nucleótido situado en la posición 906 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNAc IV delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 21 y 41, y en su extremo C-terminal por el aminoácido situado en la posición 302 de la secuencia SEC ID nº 12,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 121 y 183, y en su extremo 3' por el nucleótido situado en la posición 1.008 de la secuencia SEC ID nº 13, codificando dicha

secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac V delimitado en su extremo N-terminal por el aminoácido situado entre las posiciones 41 y 61, y en su extremo C-terminal por el aminoácido situado en la posición 336 de la secuencia SEC ID nº 14,

5 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 154 y 216, y en su extremo 3' por el nucleótido situado en la posición 999 de la secuencia SEC ID nº 15, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST6GalNac VI delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 52 y 72 y en su extremo C-terminal por el aminoácido situado en la posición 333 de la secuencia SEC ID nº 16,

10 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 145 y 207, y en su extremo 3' por el nucleótido situado en la posición 1.020 de la secuencia SEC ID nº 17, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal I delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 49 y 69, y en su extremo C-terminal por el aminoácido situado en la posición 340 de la secuencia SEC ID nº 18,

15 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 175 y 237, y en su extremo 3' por el nucleótido situado en la posición 1.050 de la secuencia SEC ID nº 19, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal II delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 59 y 79, y en su extremo C-terminal por el aminoácido situado en la posición 350 de la secuencia SEC ID nº 20,

20 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 199 y 261, y en su extremo 3' por el nucleótido situado en la posición 1.332 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal III delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 67 y 87, y en su extremo C-terminal por el aminoácido situado en la posición 444 de la secuencia SEC ID nº 22,

25 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 79 y 141, y en su extremo 3' por el nucleótido situado en la posición 987 de la secuencia SEC ID nº 23, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal IV delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 27 y 47, y en su extremo C-terminal por el aminoácido situado en la posición 329 de la secuencia SEC ID nº 24,

30 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 136 y 198, y en su extremo 3' por el nucleótido situado en la posición 1.086 de la secuencia SEC ID nº 25, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal V delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 46 y 66, y en su extremo C-terminal por el aminoácido situado en la posición 362 de la secuencia SEC ID nº 26,

35 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 73 y 135, y en su extremo 3' por el nucleótido situado en la posición 993 de la secuencia SEC ID nº 27, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST3Gal VI delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 25 y 45, y en su extremo C-terminal por el aminoácido situado en la posición 331 de la secuencia SEC ID nº 28,

40 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 103 y 165, y en su extremo 3' por el nucleótido situado en la posición 1.122 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de ST3Gal III de rata delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 35 y 55, y en su extremo C-terminal por el aminoácido situado en la posición 374 de la secuencia SEC ID nº 30,

45 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 133 y 195, y en su extremo 3' por el nucleótido situado en la posición 1.068 de la secuencia SEC ID nº 31, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia I delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 45 y 65, y en su extremo C-terminal por el aminoácido situado en la posición 356 de la secuencia SEC ID nº 32,

50 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 190 y 252, y en su extremo 3' por el nucleótido situado en la posición 1.125 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia II delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 64 y 84, y en su extremo C-terminal por el aminoácido situado en la posición 375 de la secuencia SEC ID nº 34,

55 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 205 y 267, y en su extremo 3' por el nucleótido situado en la posición 1.140 de la secuencia SEC ID nº 35, codificando dicha

secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia III delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 69 y 89, y en su extremo C-terminal por el aminoácido situado en la posición 380 de la secuencia SEC ID nº 36,

5 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 145 y 207, y en su extremo 3' por el nucleótido situado en la posición 1.077 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia IV delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 49 y 69, y en su extremo C-terminal por el aminoácido situado en la posición 359 de la secuencia SEC ID nº 38,

10 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 211 y 273, y en su extremo 3' por el nucleótido situado en la posición 1.128 de la secuencia SEC ID nº 39, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia V delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 71 y 91, y en su extremo C-terminal por el aminoácido situado en la posición 376 de la secuencia SEC ID nº 40,

15 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado entre las posiciones 277 y 339, y en su extremo 3' por el nucleótido situado en la posición 1.194 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente al dominio CD de hST8Sia VI delimitada en su extremo N-terminal por el aminoácido situado entre las posiciones 93 y 113, y en su extremo C-terminal por el aminoácido situado en la posición 398 de la secuencia SEC ID nº 42.

20 8. Procedimiento según cualquiera de las reivindicaciones 1 a 7, caracterizado porque el primer ácido nucleico es la secuencia de nucleótidos SEC ID nº 43 correspondiente a la secuencia delimitada por los nucleótidos situados en la posición 268 y 1.218 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos SEC ID nº: 43 la secuencia polipeptídica SEC ID nº 44 correspondiente al fragmento mínimo C-terminal del dominio CD de hST6Gal I delimitado por los aminoácidos situados en las posiciones 90 a 406 de la secuencia SEC ID nº 2.

30 9. Procedimiento según cualquiera de las reivindicaciones 1 a 6, caracterizado porque:

- la secuencia de nucleótidos codificante de la región CT comprendida en el segundo ácido nucleico se selecciona de entre:

35 * la secuencia de nucleótidos SEC ID nº 45 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 27 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos SEC ID nº 45 la secuencia polipeptídica SEC ID nº 46 correspondiente a la región CT de hST6Gal I delimitada por los aminoácidos situados entre las posiciones 1 y 9 de la secuencia SEC ID nº 2,

40 * la secuencia de nucleótidos SEC ID nº 47 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 30 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos SEC ID nº 47 la secuencia polipeptídica SEC ID nº 48 correspondiente a la región CT de hST6Gal II delimitada por los aminoácidos situados entre las posiciones 1 y 10 de la secuencia SEC ID nº 4,

45 * la secuencia de nucleótidos SEC ID nº 49 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 42 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos SEC ID nº 50 la secuencia polipeptídica SEC ID nº 50 correspondiente a la región CT de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 1 y 14 de la secuencia SEC ID nº 6,

50 * la secuencia de nucleótidos SEC ID nº 51 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 21 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos SEC ID nº 51 la secuencia polipeptídica SEC ID nº 52 correspondiente a la región CT de hST6GalNAc II delimitada por los aminoácidos situados entre las posiciones 1 y 7 de la secuencia SEC ID nº 8,

55 * la secuencia de nucleótidos SEC ID nº 53 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos SEC ID nº 53 la secuencia polipeptídica SEC ID nº 54 correspondiente a la región CT de hST6GalNAc III delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 10,

60 * la secuencia de nucleótidos SEC ID nº 55 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 18 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos SEC ID nº 55 la secuencia polipeptídica SEC ID nº 56 correspondiente a la región CT de hST6GalNAc IV delimitada por los aminoácidos situados entre las posiciones 1 y 6 de la secuencia SEC ID nº 12,

65 * la secuencia de nucleótidos SEC ID nº 57 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 13, codificando dicha secuencia de

nucleótidos SEC ID nº 83 la secuencia polipeptídica SEC ID nº 84 correspondiente a la región CT de hST8Sia V delimitada por los aminoácidos situados entre las posiciones 1 y 17 de la secuencia SEC ID nº 40,

* la secuencia de nucleótidos SEC ID nº 85 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 9 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos SEC ID nº 85 la secuencia polipeptídica SEC ID nº 86 correspondiente a la región CT de hST8Sia VI delimitada por los aminoácidos situados entre las posiciones 1 y 3 de la secuencia SEC ID nº 42,

- la secuencia de nucleótidos codificante de la región TMD comprendida en el segundo ácido nucleico se selecciona de entre:

* la secuencia de nucleótidos SEC ID nº 87 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 28 y 78 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos SEC ID nº 87 la secuencia polipeptídica SEC ID nº 88 correspondiente a la región TMD de hST6Gal I delimitada por los aminoácidos situados entre las posiciones 10 y 26 de la secuencia SEC ID nº 2,

* la secuencia de nucleótidos SEC ID nº 89 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 31 y 90 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos SEC ID nº 89 la secuencia polipeptídica SEC ID nº 90 correspondiente a la región TMD de hST6Gal II delimitada por los aminoácidos situados entre las posiciones 11 y 30 de la secuencia SEC ID nº 4,

* la secuencia de nucleótidos SEC ID nº 91 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 43 y 105 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos SEC ID nº 91 la secuencia polipeptídica SEC ID nº 92 correspondiente a la región TMD de hST6GalNAC I delimitada por los aminoácidos situados entre las posiciones 15 y 35 de la secuencia SEC ID nº 6,

* la secuencia de nucleótidos SEC ID nº 93 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 22 y 84 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos SEC ID nº 93 la secuencia polipeptídica SEC ID nº 94 correspondiente a la región TMD de hST6GalNAC II delimitada por los aminoácidos situados entre las posiciones 8 y 28 de la secuencia SEC ID nº 8,

* la secuencia de nucleótidos SEC ID nº 95 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos SEC ID nº 95 la secuencia polipeptídica SEC ID nº 96 correspondiente a la región TMD de hST6GalNAC III delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 10,

* la secuencia de nucleótidos SEC ID nº 97 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 19 y 81 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos SEC ID nº 97 la secuencia polipeptídica SEC ID nº 98 correspondiente a la región TMD de hST6GalNAC IV delimitada por los aminoácidos situados entre las posiciones 7 y 27 de la secuencia SEC ID nº 12,

* la secuencia de nucleótidos SEC ID nº 99 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 87 de la secuencia SEC ID nº 13, codificando dicha secuencia de nucleótidos SEC ID nº 99 la secuencia polipeptídica SEC ID nº 100 correspondiente a la región TMD de hST6GalNAC V delimitada por los aminoácidos situados entre las posiciones 9 y 29 de la secuencia SEC ID nº 14,

* la secuencia de nucleótidos SEC ID nº 101 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 130 y 177 de la secuencia SEC ID nº 15, codificando dicha secuencia de nucleótidos SEC ID nº 101 la secuencia polipeptídica SEC ID nº 102 correspondiente a la región TMD de hST6GalNAC VI delimitada por los aminoácidos situados entre las posiciones 44 y 59 de la secuencia SEC ID nº 16,

* la secuencia de nucleótidos SEC ID nº 103 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 40 y 102 de la secuencia SEC ID nº 17, codificando dicha secuencia de nucleótidos SEC ID nº 103 la secuencia polipeptídica SEC ID nº 104 correspondiente a la región TMD de hST3Gal I delimitada por los aminoácidos situados entre las posiciones 14 y 34 de la secuencia SEC ID nº 18,

* la secuencia de nucleótidos SEC ID nº 105 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 19 y 81 de la secuencia SEC ID nº 19, codificando dicha secuencia de nucleótidos SEC ID nº 105 la secuencia polipeptídica SEC ID nº 106 correspondiente a la región TMD de hST3Gal II delimitada por los aminoácidos situados entre las posiciones 7 y 27 de la secuencia SEC ID nº 20,

* la secuencia de nucleótidos SEC ID nº 107 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos SEC ID nº 107 la secuencia polipeptídica SEC ID nº 108 correspondiente a la región TMD de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 22,

* la secuencia de nucleótidos SEC ID nº 109 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 78 de la secuencia SEC ID nº 23, codificando dicha secuencia de nucleótidos SEC ID nº 109 la secuencia polipeptídica SEC ID nº 110 correspondiente a la región TMD de hST3Gal IV delimitada por los aminoácidos situados entre las posiciones 9 y 26 de la secuencia SEC ID nº 24,

* la secuencia de nucleótidos SEC ID nº 111 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 16 y 78 de la secuencia SEC ID nº 25, codificando dicha secuencia de nucleótidos SEC ID nº 111 la secuencia polipeptídica SEC ID nº 112 correspondiente a la región TMD de hST3Gal V delimitada por los aminoácidos situados entre las posiciones 6 y 26 de la secuencia SEC ID nº 26,

* la secuencia de nucleótidos SEC ID nº 113 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 13 y 75 de la secuencia SEC ID nº 27, codificando dicha secuencia de nucleótidos SEC ID nº 113 la secuencia polipeptídica SEC ID nº 114 correspondiente a la región TMD de hST3Gal VI delimitada por los aminoácidos situados entre las posiciones 5 y 25 de la secuencia SEC ID nº 28,

* la secuencia de nucleótidos SEC ID nº 115 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos SEC ID nº 115 la secuencia polipeptídica SEC ID nº 116 correspondiente a la región TMD de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 30,

* la secuencia de nucleótidos SEC ID nº 117 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 88 y 144 de la secuencia SEC ID nº 31, codificando dicha secuencia de nucleótidos SEC ID nº 117 la secuencia polipeptídica SEC ID nº 118 correspondiente a la región TMD de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 30 y 48 de la secuencia SEC ID nº 32,

* la secuencia de nucleótidos SEC ID nº 119 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 19 y 69 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos SEC ID nº 119 la secuencia polipeptídica SEC ID nº 120 correspondiente a la región TMD de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 7 y 23 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos SEC ID nº 121 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 28 y 99 de la secuencia SEC ID nº 35, codificando dicha secuencia de nucleótidos SEC ID nº 121 la secuencia polipeptídica SEC ID nº 122 correspondiente a la región TMD de hST8Sia III delimitada por los aminoácidos situados entre las posiciones 10 y 33 de la secuencia SEC ID nº 36,

* la secuencia de nucleótidos SEC ID nº 123 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 22 y 60 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos SEC ID nº 123 la secuencia polipeptídica SEC ID nº 124 correspondiente a la región TMD de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 8 y 20 de la secuencia SEC ID nº 38,

* la secuencia de nucleótidos SEC ID nº 125 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 52 y 114 de la secuencia SEC ID nº 39, codificando dicha secuencia de nucleótidos SEC ID nº 125 la secuencia polipeptídica SEC ID nº 126 correspondiente a la región TMD de hST8Sia V delimitada por los aminoácidos situados entre las posiciones 18 y 38 de la secuencia SEC ID nº 40,

* la secuencia de nucleótidos SEC ID nº 127 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 10 y 72 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos SEC ID nº 127 la secuencia polipeptídica SEC ID nº 128 correspondiente a la región TMD de hST8Sia VI delimitada por los aminoácidos situados entre las posiciones 4 y 24 de la secuencia SEC ID nº 42,

- la secuencia de nucleótidos codificante de la región SR comprendida en el segundo ácido nucleico, o codificante de un fragmento de por lo menos 3 aminoácidos del mismo, se selecciona de entre las secuencias constituidas por o que comprenden:

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 79, y en su extremo 3' por el nucleótido situado entre las posiciones 267 y 327 de la secuencia SEC ID nº 1, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal I delimitada en su extremo N-terminal por el aminoácido situado en la posición 27 y en su extremo C-terminal por el aminoácido situado entre las posiciones 89 y 109 de la secuencia SEC ID nº 2,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 91, y en su extremo 3' por el nucleótido situado entre las posiciones 306 y 336 de la secuencia SEC ID nº 3, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal II delimitada en su extremo N-terminal por el aminoácido situado en la posición 31 y en su extremo C-terminal por el aminoácido situado entre las posiciones 102 y 112 de la secuencia SEC ID nº 4,

- 5 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 106, y en su extremo 3' por el nucleótido situado entre las posiciones 813 y 873 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc I delimitada en su extremo N-terminal por el aminoácido situado en la posición 36 y en su extremo C-terminal por el aminoácido situado entre las posiciones 271 y 291 de la secuencia SEC ID nº 6,
- 10 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 171 y 231 de la secuencia SEC ID nº 7, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc II delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 57 y 77 de la secuencia SEC ID nº 8,
- 15 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 102 y 132 de la secuencia SEC ID nº 9, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc III delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 34 y 44 de la secuencia SEC ID nº 10,
- 20 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 82, y en su extremo 3' por el nucleótido situado entre las posiciones 90 y 120 de la secuencia SEC ID nº 11, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc IV delimitada en su extremo N-terminal por el aminoácido situado en la posición 28 y en su extremo C-terminal por el aminoácido situado entre las posiciones 30 y 40 de la secuencia SEC ID nº 12,
- 25 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 88, y en su extremo 3' por el nucleótido situado entre las posiciones 120 y 180 de la secuencia SEC ID nº 13, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc V delimitada en su extremo N-terminal por el aminoácido situado en la posición 30 y en su extremo C-terminal por el aminoácido situado entre las posiciones 40 y 60 de la secuencia SEC ID nº 14,
- 30 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 178, y en su extremo 3' por el nucleótido situado en la posición 183 de la secuencia SEC ID nº 15, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6GalNAc VI delimitada en su extremo N-terminal por el aminoácido situado en la posición 60 y en su extremo C-terminal por el aminoácido situado en la posición 61 de la secuencia SEC ID nº 16,
- 35 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 103, y en su extremo 3' por el nucleótido situado entre las posiciones 144 y 204 de la secuencia SEC ID nº 17, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal I delimitada en su extremo N-terminal por el aminoácido situado en la posición 35 y en su extremo C-terminal por el aminoácido situado entre las posiciones 48 y 68 de la secuencia SEC ID nº 18,
- 40 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 82, y en su extremo 3' por el nucleótido situado entre las posiciones 174 y 234 de la secuencia SEC ID nº 19, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST6Gal II delimitada en su extremo N-terminal por el aminoácido situado en la posición 28 y en su extremo C-terminal por el aminoácido situado entre las posiciones 58 y 78 de la secuencia SEC ID nº 20,
- 45 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 198 y 258 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal III delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 66 y 86 de la secuencia SEC ID nº 22,
- 50 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 79, y en su extremo 3' por el nucleótido situado entre las posiciones 108 y 138 de la secuencia SEC ID nº 23, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal IV delimitada en su extremo N-terminal por el aminoácido situado en la posición 27 y en su extremo C-terminal por el aminoácido situado entre las posiciones 36 y 46 de la secuencia SEC ID nº 24,
- 55 * la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 79, y en su extremo 3' por el nucleótido situado entre las posiciones 135 y 195 de la secuencia SEC ID nº 25, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal V delimitada en su extremo N-terminal por el aminoácido situado en la posición 27 y en su extremo C-terminal por el aminoácido situado entre las posiciones 45 y 65 de la secuencia SEC ID nº 26,
- 60
- 65

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 76, y en su extremo 3' por el nucleótido situado entre las posiciones 102 y 132 de la secuencia SEC ID nº 27, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST3Gal VI delimitada en su extremo N-terminal por el aminoácido situado en la posición 26 y en su extremo C-terminal por el aminoácido situado entre las posiciones 34 y 44 de la secuencia SEC ID nº 28,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 85, y en su extremo 3' por el nucleótido situado entre las posiciones 105 y 165 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de ST3Gal III de rata delimitada en su extremo N-terminal por el aminoácido situado en la posición 29 y en su extremo C-terminal por el aminoácido situado entre las posiciones 35 y 55 de la secuencia SEC ID nº 30,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 145, y en su extremo 3' por el nucleótido situado entre las posiciones 162 y 192 de la secuencia SEC ID nº 31, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia I delimitada en su extremo N-terminal por el aminoácido situado en la posición 49 y en su extremo C-terminal por el aminoácido situado entre las posiciones 54 y 64 de la secuencia SEC ID nº 32,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 70, y en su extremo 3' por el nucleótido situado entre las posiciones 189 y 249 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia II delimitada en su extremo N-terminal por el aminoácido situado en la posición 24 y en su extremo C-terminal por el aminoácido situado entre las posiciones 63 y 83 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 100, y en su extremo 3' por el nucleótido situado entre las posiciones 204 y 264 de la secuencia SEC ID nº 35, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia III delimitada en su extremo N-terminal por el aminoácido situado en la posición 34 y en su extremo C-terminal por el aminoácido situado entre las posiciones 68 y 88 de la secuencia SEC ID nº 36,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 61, y en su extremo 3' por el nucleótido situado entre las posiciones 144 y 204 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia IV delimitada en su extremo N-terminal por el aminoácido situado en la posición 21 y en su extremo C-terminal por el aminoácido situado entre las posiciones 48 y 68 de la secuencia SEC ID nº 38,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 115, y en su extremo 3' por el nucleótido situado entre las posiciones 210 y 270 de la secuencia SEC ID nº 39, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia V delimitada en su extremo N-terminal por el aminoácido situado en la posición 39 y en su extremo C-terminal por el aminoácido situado entre las posiciones 70 y 90 de la secuencia SEC ID nº 40,

* la secuencia de nucleótidos delimitada en su extremo 5' por el nucleótido situado en la posición 73, y en su extremo 3' por el nucleótido situado entre las posiciones 276 y 336 de la secuencia SEC ID nº 41, codificando dicha secuencia de nucleótidos la secuencia polipeptídica correspondiente a la región SR de hST8Sia VI delimitada en su extremo N-terminal por el aminoácido situado en la posición 25 y en su extremo C-terminal por el aminoácido situado entre las posiciones 92 y 112 de la secuencia SEC ID nº 42,

* o cualquier fragmento de por lo menos 6 nucleótidos de las secuencias de nucleótidos codificantes de las secuencias polipeptídicas correspondientes a las regiones SR mencionadas anteriormente, y codificantes de por lo menos 2 aminoácidos contiguos de dichas regiones SR, tales como:

** el fragmento SEC ID nº 129 delimitado por los nucleótidos situados entre las posiciones 106 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 130 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 74 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 131 delimitado por los nucleótidos situados entre las posiciones 109 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 132 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 37 y 74 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 133 delimitado por los nucleótidos situados entre las posiciones 106 y 420 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 134 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 140 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 135 delimitado por los nucleótidos situados entre las posiciones 106 y 774 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 136 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 258 de la secuencia SEC ID nº 6,

** el fragmento SEC ID nº 137 delimitado por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 21, codificante de la secuencia polipeptídica SEC ID nº 138 correspondiente al fragmento de la región SR de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 22,

** el fragmento SEC ID nº 139 delimitado por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 29, codificante de la secuencia polipeptídica SEC ID nº 140 correspondiente al fragmento de la región SR de ST3GalII de rata delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 30,

* el fragmento SEC ID nº 141 delimitado por los nucleótidos situados entre las posiciones 70 y 237 de la secuencia SEC ID nº 33, codificante de la secuencia polipeptídica SEC ID nº 142 correspondiente al fragmento de la región SR de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 24 y 79 de la secuencia SEC ID nº 34,

* el fragmento SEC ID nº 143 delimitado por los nucleótidos situados entre las posiciones 61 y 201 de la secuencia SEC ID nº 37, codificante de la secuencia polipeptídica SEC ID nº 144 correspondiente al fragmento de la región SR de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 21 y 67 de la secuencia SEC ID nº 38.

10. Procedimiento según la reivindicación 9, caracterizado porque:

la secuencia de nucleótidos codificante de la región CT comprendida en el segundo ácido nucleico es seleccionada de entre:

* la secuencia de nucleótidos SEC ID nº 49 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 42 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos SEC ID nº 49 la secuencia polipeptídica SEC ID nº 50 correspondiente a la región CT de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 1 y 14 de la secuencia SEC ID nº 6,

* la secuencia de nucleótidos SEC ID nº 65 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos SEC ID nº 65 la secuencia polipeptídica SEC ID nº 66 correspondiente a la región CT de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 22,

* la secuencia de nucleótidos SEC ID nº 73 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 24 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos SEC ID nº 73 la secuencia polipeptídica SEC ID nº 74 correspondiente a la región CT de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 1 y 8 de la secuencia SEC ID nº 30,

* la secuencia de nucleótidos SEC ID nº 77 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 18 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos SEC ID nº 77 la secuencia polipeptídica SEC ID nº 78 correspondiente a la región CT de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 1 y 6 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos SEC ID nº 81 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 1 y 21 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos SEC ID nº 81 la secuencia polipeptídica SEC ID nº 82 correspondiente a la región CT de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 1 y 7 de la secuencia SEC ID nº 38,

- la secuencia de nucleótidos codificante de la región TMD comprendida en el segundo ácido nucleico es seleccionada de entre:

* la secuencia de nucleótidos SEC ID nº 91 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 43 y 105 de la secuencia SEC ID nº 5, codificando dicha secuencia de nucleótidos SEC ID nº 91 la secuencia polipeptídica SEC ID nº 92 correspondiente a la región TMD de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 15 y 35 de la secuencia SEC ID nº 6,

* la secuencia de nucleótidos SEC ID nº 107 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 21, codificando dicha secuencia de nucleótidos SEC ID nº 107 la secuencia polipeptídica SEC ID nº 108 correspondiente a la región TMD de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 22,

* la secuencia de nucleótidos SEC ID nº 115 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 25 y 84 de la secuencia SEC ID nº 29, codificando dicha secuencia de nucleótidos SEC ID nº 115 la secuencia polipeptídica SEC ID nº 116 correspondiente a la región TMD de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 9 y 28 de la secuencia SEC ID nº 30,

* la secuencia de nucleótidos SEC ID nº 119 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 19 y 69 de la secuencia SEC ID nº 33, codificando dicha secuencia de nucleótidos SEC ID nº 119 la secuencia polipeptídica SEC ID nº 120 correspondiente a la región TMD de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 7 y 23 de la secuencia SEC ID nº 34,

* la secuencia de nucleótidos SEC ID nº 123 correspondiente a la secuencia de nucleótidos delimitada por los nucleótidos situados entre las posiciones 22 y 60 de la secuencia SEC ID nº 37, codificando dicha secuencia de nucleótidos SEC ID nº 123 la secuencia polipeptídica SEC ID nº 124 correspondiente a la región TMD de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 8 y 20 de la secuencia SEC ID nº 38,

- la secuencia de nucleótidos codificante de la región SR o su fragmento comprendida en el segundo ácido nucleico es seleccionada de entre:

* la secuencia SEC ID nº 129 delimitada por los nucleótidos situados entre las posiciones 106 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 130 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 74 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 131 delimitada por los nucleótidos situados entre las posiciones 109 y 222 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 132 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 37 y 74 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 133 delimitada por los nucleótidos situados entre las posiciones 106 y 420 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 134 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 140 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 135 delimitada por los nucleótidos situados entre las posiciones 106 y 774 de la secuencia SEC ID nº 5, codificante de la secuencia polipeptídica SEC ID nº 136 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 258 de la secuencia SEC ID nº 6,

* la secuencia SEC ID nº 137 delimitada por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 21, codificante de la secuencia polipeptídica SEC ID nº 138 correspondiente al fragmento de la región SR de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 22,

* la secuencia SEC ID nº 139 delimitada por los nucleótidos situados entre las posiciones 85 y 138 de la secuencia SEC ID nº 29, codificante de la secuencia polipeptídica SEC ID nº 140 correspondiente al fragmento de la región SR de rST3GalIII delimitada por los aminoácidos situados entre las posiciones 29 y 46 de la secuencia SEC ID nº 30,

* la secuencia SEC ID nº 141 delimitada por los nucleótidos situados entre las posiciones 70 y 237 de la secuencia SEC ID nº 33, codificante de la secuencia polipeptídica SEC ID nº 142 correspondiente al fragmento de la región SR de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 24 y 79 de la secuencia SEC ID nº 34,

* la secuencia SEC ID nº 143 delimitada por los nucleótidos situados entre las posiciones 61 y 201 de la secuencia SEC ID nº 37, codificante de la secuencia polipeptídica SEC ID nº 144 correspondiente al fragmento de la región SR de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 21 y 67 de la secuencia SEC ID nº 38.

11. Procedimiento según la reivindicación 9 ó 10, caracterizado porque CT, TMD, SR o péptidos de fragmento de SR comprendidos en el segundo ácido nucleico son secuencias homólogas obtenidas de la misma glucosiltransferasa nativa, siendo ésta última diferente de los péptidos en la glucosiltransferasa nativa a partir de la que se obtiene el fragmento de CD con actividad glucosiltransferasas óptima según la reivindicación 1.

12. Procedimiento según cualquiera de las reivindicaciones 9 a 11, caracterizado porque el segundo ácido nucleico se selecciona de entre las secuencias siguientes:

- la secuencia SEC ID nº 145 delimitada por los nucleótidos situados entre las posiciones 1 y 222 de la secuencia SEC ID nº 5, que contiene las secuencias SEC ID nº 49, 91 y 129, y codificantes de la secuencia polipeptídica SEC ID nº 146 correspondiente al fragmento de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 1 y 74 de la secuencia SEC ID nº 6 y que contiene las regiones de fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 130, respectivamente,

- la secuencia SEC ID nº 147 delimitada por los nucleótidos situados entre las posiciones 1 y 420 de la secuencia SEC ID nº 5, que contiene las secuencias SEC ID nº 49, 91 y 133, y codificantes de la secuencia polipeptídica SEC ID nº 148 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 1 y 140 de la secuencia SEC ID nº 6, y que contiene las regiones de fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 134, respectivamente,

- la secuencia SEC ID nº 149 delimitada por los nucleótidos situados entre las posiciones 1 y 774 de la secuencia SEC ID nº 5, que contiene las secuencias SEC ID nº 49, 91 y 135, y codificantes de la secuencia polipeptídica SEC ID nº 150 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 1 y 258 de la secuencia SEC ID nº 6, y que contiene las regiones de fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 136, respectivamente,

- la secuencia SEC ID nº 151 delimitada por los nucleótidos situados entre las posiciones 1 y 138 de la secuencia SEC ID nº 21, que contiene las secuencias SEC ID nº 65, 107 y 137, y codificantes de la secuencia polipeptídica SEC ID nº 152 correspondiente al fragmento de la región SR de hST3Gal III delimitada por los aminoácidos situados entre las posiciones 1 y 46 de la secuencia SEC ID nº 22, y que contiene las regiones de fragmentos CT, TMD y SR de hST3Gal III correspondientes a las secuencias SEC ID nº 66, 108 y 138, respectivamente,

- la secuencia SEC ID nº 153 delimitada por los nucleótidos situados entre las posiciones 1 y 138 de la secuencia SEC ID nº 29, que contiene las secuencias SEC ID nº 73, 115 y 139, y codificantes de la secuencia polipeptídica SEC ID nº 154 correspondiente al fragmento de la región SR de ST3Gal III de rata delimitada por los aminoácidos situados entre las posiciones 1 y 46 de la secuencia SEC ID nº 30, y que contiene las regiones de fragmentos CT, TMD y SR de hST3Gal III correspondientes a las secuencias SEC ID nº 74, 116 y 140, respectivamente,

- la secuencia SEC ID nº 155 delimitada por los nucleótidos situados entre las posiciones 1 y 237 de la secuencia SEC ID nº 33, que contiene las secuencias SEC ID nº 77, 119 y 141, y codificantes de la secuencia polipeptídica SEC ID nº 156 correspondiente al fragmento de la región SR de hST8Sia II delimitada por los aminoácidos situados entre las posiciones 1 y 79 de la secuencia SEC ID nº 34, y que contiene las regiones CT, TMD y SR de hST8Sia II correspondientes a las secuencias SEC ID nº 78, 120 y 142, respectivamente,

- la secuencia SEC ID nº 157 delimitada por los nucleótidos situados entre las posiciones 1 y 201 de la secuencia SEC ID nº 37, que contiene las secuencias SEC ID nº 81, 123 y 143, y codificantes de la secuencia polipeptídica SEC ID nº 158 correspondiente al fragmento de la región SR de hST8Sia IV delimitada por los aminoácidos situados entre las posiciones 1 y 67 de la secuencia SEC ID nº 38, y que contiene las regiones CT, TMD y SR de hST8Sia IV correspondientes a las secuencias SEC ID nº 82, 124 y 144, respectivamente.

13. Procedimiento según la reivindicación 9 ó 10, caracterizado porque los CT, TMD, SR o péptidos de fragmento de SR comprendidos en el segundo ácido nucleico son secuencias heterólogas obtenidas de glucosiltransferasas nativas diferentes.

14. Procedimiento según la reivindicación 13, caracterizado porque el segundo ácido nucleico es la secuencia SEC ID nº 159 correspondiente a la fusión de la secuencia de nucleótidos SEC ID nº 177 que contiene las secuencias SEC ID nº 65 y SEC ID nº 107 codificantes de las regiones CT y TMD de hST3Gal III correspondientes a las secuencias SEC ID nº 66 y SEC ID nº 108, respectivamente, con la secuencia de nucleótidos SEC ID nº 131 codificante de la secuencia polipeptídica SEC ID nº 132 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 37 y 74 de la secuencia SEC ID nº 6, codificando dicha secuencia SEC ID nº 159 el polipéptido de fusión SEC ID nº 160 entre las regiones CT y TMD de hST3Gal III, por una parte, y el fragmento 37-74 de la región SR de hST6GalNAc I, por la otra.

15. Procedimiento según la reivindicación 13, caracterizado porque el segundo ácido nucleico es la secuencia SEC ID nº 179 correspondiente a la fusión de la secuencia de nucleótidos SEC ID nº 65 codificante de la CT de hST3Gal III correspondiente a la secuencia SEC ID nº 66, con la secuencia de nucleótidos SEC ID nº 119 codificante de la región TMD de hST8Sia II correspondiente a la secuencia SEC ID nº 120 y con la secuencia de nucleótidos SEC ID nº 129 codificante de la secuencia polipeptídica SEC ID nº 130 correspondiente al fragmento de la región SR de hST6GalNAc I delimitada por los aminoácidos situados entre las posiciones 36 y 74 de la secuencia SEC ID nº 6, codificando dicha secuencia SEC ID nº 179 el polipéptido de fusión SEC ID nº 180 entre la región CT de hST3Gal III, la región TMD de hST8Sia II, y el fragmento 36-74 de la región SR de hST6GalNAc I.

16. Procedimiento según cualquiera de las reivindicaciones 1 a 15, caracterizado porque comprende la fusión de la secuencia SEC ID nº 43 como primer ácido nucleico, con un segundo ácido nucleico seleccionado de entre:

- la secuencia SEC ID nº 145, conduciendo a la secuencia de nucleótidos SEC ID nº 161 codificante de la proteína de fusión SEC ID nº 162 que contiene las regiones de fragmento CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 130, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 147, conduciendo a la secuencia de nucleótidos SEC ID nº 163 codificante de la proteína de fusión SEC ID nº 164 que contiene las regiones de fragmento CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 134, respectivamente, unidas mediante un coector GS al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 149, conduciendo a la secuencia de nucleótidos SEC ID nº 165 codificante de la proteína de fusión SEC ID nº 166 que contiene las regiones de fragmento CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 136, respectivamente, unidas mediante un conector SR al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 151, conduciendo a la secuencia de nucleótidos SEC ID nº 167 codificante de la proteína de fusión SEC ID nº 168 que contiene las regiones de fragmento CT, TMD y SR de hST3Gal III correspondientes a las secuencias SEC ID nº 66, 108 y 138, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 153, conduciendo a la secuencia de nucleótidos SEC ID nº 169 codificante de la proteína de fusión SEC ID nº 170 que contiene las regiones CT, TMD y SR de ST3Gal III de rata correspondientes a las secuencias SEC ID nº 74, 116 y 140, respectivamente, unidas mediante un conector SR al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 155, conduciendo a la secuencia de nucleótidos SEC ID nº 171 codificante de la proteína de fusión SEC ID nº 172 que contiene las regiones CT, TMD y SR de hST8Sia II correspondientes a las secuencias SEC ID nº 78, 120 y 142, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 157, conduciendo a la secuencia de nucleótidos SEC ID nº 173 codificante de la proteína de fusión SEC ID nº 174 que contiene las regiones CT, TMD y SR de hST8Sia IV correspondientes a las secuencias SEC ID nº 82, 124 y 144, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 159, conduciendo a la secuencia de nucleótidos SEC ID nº 175 codificante de la proteína de fusión SEC ID nº 176 que contiene las regiones CT y TMD de hST3Gal III correspondientes a las secuencias SEC ID nº 66 y 108, y el fragmento 37-74 de la región SR de hST6GalNAc I correspondiente a la secuencia SEC ID nº 132, unidas mediante un coenctor GS al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

- la secuencia SEC ID nº 179, conduciendo a la secuencia de nucleótidos SEC ID nº 181 codificante de la proteína de fusión SEC ID nº 182 que contiene las regiones CT de hST3Gal III, la región TMD de hST8Sia II y el fragmento 36-74 de la región SR de hST6GalNAc I, correspondientes a las secuencias SEC ID nº 66, 120 y 130, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90-406 del dominio CD de hST6Gal I,

17. Secuencias génicas codificantes de glucosiltransferasas quiméricas obtenidas según el procedimiento definido en cualquiera de las reivindicaciones 1 a 16.

18. Secuencias génicas según la reivindicación 17, seleccionadas de entre:

- la secuencia SEC ID nº 161 codificante de la proteína de fusión SEC ID nº 162 que contiene las regiones de fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 130, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 161 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 145,

- la secuencia SEC ID nº 163 codificante de la proteína de fusión SEC ID nº 164 que contiene las regiones de fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 134, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 163 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 147,

- la secuencia SEC ID nº 165 codificante de la proteína de fusión SEC ID nº 166 que contiene las regiones de fragmentos CT, TMD y SR de hST6GalNAc I correspondientes a las secuencias SEC ID nº 50, 92 y 136, respectivamente, unidas mediante un conector SR al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 165 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 149,

- la secuencia SEC ID nº 167 codificante de la proteína de fusión SEC ID nº 168 que contiene las regiones de fragmentos CT, TMD y SR de hST3Gal III correspondientes a las secuencias SEC ID nº 66, 108 y 138,

respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 167 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 151,

5 - la secuencia SEC ID nº 169 codificante de la proteína de fusión SEC ID nº 170 que contiene las regiones de fragmentos CT, TMD y SR de rST3Gal III correspondientes a las secuencias SEC ID nº 74, 116 y 140, respectivamente, unidas mediante un conector SR al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 169 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 153,

10 - la secuencia SEC ID nº 171 codificante de la proteína de fusión SEC ID nº 172 que contiene las regiones CT, TMD y SR de hST8Sia II correspondientes a las secuencias SEC ID nº 78, 120 y 142, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 171 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 155,

15 - la secuencia SEC ID nº 173 codificante de la proteína de fusión SEC ID nº 174 que contiene las regiones CT, TMD y SR de hST8Sia IV correspondientes a las secuencias SEC ID nº 82, 124 y 144, respectivamente, unidas mediante un conector KL al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 173 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 157,

20 - la secuencia SEC ID nº 175 codificante de la proteína de fusión SEC ID nº 176 que contiene las regiones CT y TMD de hST3Gal III correspondientes a las secuencias SEC ID nº 66 y 108, respectivamente, y el fragmento 37 a 74 de la región SR de hST6GalNAc I correspondiente a la secuencia SEC ID nº 132, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 175 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 159,

25 - la secuencia SEC ID nº 181 codificante de la proteína de fusión SEC ID nº 182 que contiene la región CT de hST3Gal III, la región TMD de hST8Sia II y el fragmento 36 a 74 de la región SR de hST6GalNAc I, correspondientes a las secuencias SEC ID nº 66, 120 y 130, respectivamente, unidas mediante un conector GS al fragmento mínimo C-terminal 90 a 406 del dominio CD de hST6Gal I, correspondiendo dicha secuencia SEC ID nº 181 a la fusión de la secuencia SEC ID nº 43 y la secuencia SEC ID nº 179.

35 19. Vectores, tales como plásmidos, vectores víricos o bacterianos, que comprenden por lo menos una secuencia génica según la reivindicación 17 ó 18.

20. Células huésped, tales como células de levaduras, hongos, vegetales, de insecto, de ave, de mamífero o humanas, transformadas con por lo menos una secuencia génica según la reivindicación 17 ó 18, utilizando un vector según la reivindicación 19.

40 21. Utilización de por lo menos una secuencia génica según la reivindicación 17 ó 18, o de un vector según la reivindicación 19, para la transformación de células según la reivindicación 20, en el marco de la preparación de proteínas recombinantes de interés, siendo dicha proteína de interés seleccionada de entre hormonas, enzimas, factores de coagulación, antígenos carbohidratos/biomarcadores séricos, citocinas, factores de crecimiento, anticuerpos o receptores.

45 22. Procedimiento para la preparación de proteínas recombinantes de interés que comprende la transformación de células según la reivindicación 20, con un vector que contiene una secuencia de nucleótidos codificante de dichas proteínas recombinantes de interés.

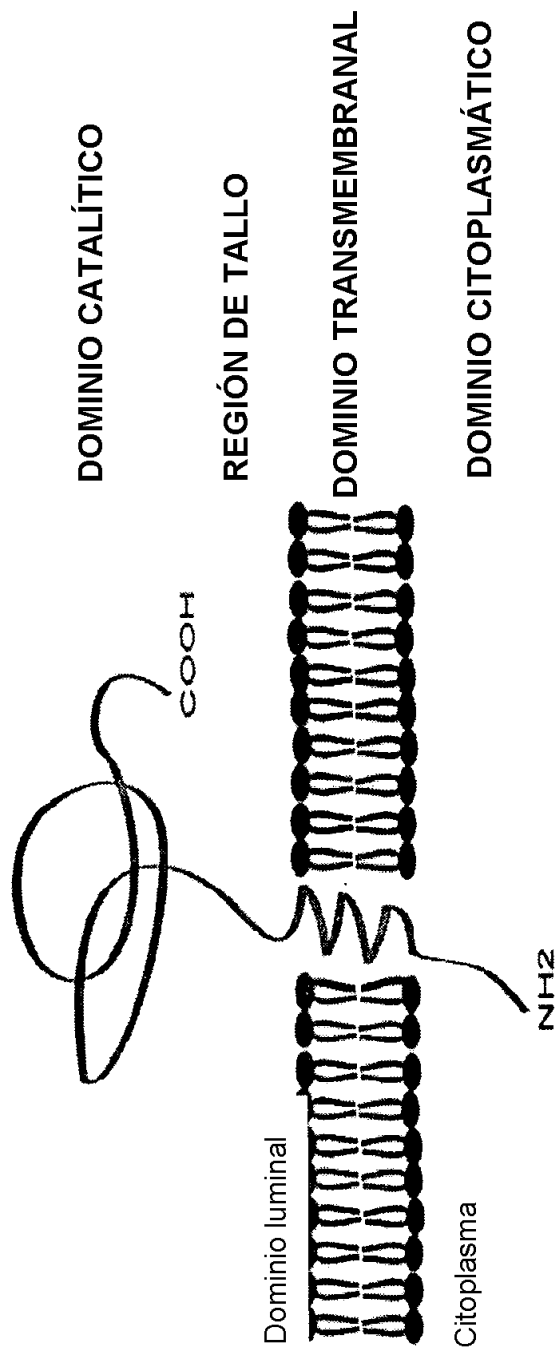


Figura 1

ST8Sia II	TCAIVGNSGVLLNSGCGQEIDAHSFVIRC�LAPV-QEYARDVGLKT
ST8Sia IV	TCAVVGNŠGILLDSECGKEIDSHNFVIRC�LAPV-VEFAADVGTKS
ST3Gal I	RCVVGNŠGNLRESSYGPEDSHDFVLRMNKAPT-AGFEADVGTKT
ST3Gal II	RCVVGNŠGNLRGSGYGQDVGHNFMRMNQAPT-VGFEQDVGSRT
ST6GalNAc III	LCAIVSNSGQMVGQKVGNEIDRSSCIWRMNNAPT-KGYEEDVGRMT
ST6GalNAc IV	SCAVVSSSGQMLGSGLGAEIDSAECVFRMNQAPT-VGFEADVQORS
ST6GalNAc V	DCALVTSSGHLLHSRQGSQIDQTECVIRMNDAPT-RGYGRDVGNRT
ST6GalNAc VI	QCVIVSSSSSHLLGTKLGPETIERAECTIRMNDAPT-TGYSADVGNKT
ST8Sia III	ICAVVGNŠGILTFIQCGREIDKSDVFVRCNFAPS-EAFQRDVGRKT
ST8Sia I	KCAVVGNŠGGILKKSGCGRQIDEANFVMRC�LPPLSSEYTKDVGSKS
ST8Sia VI	QCAVVGNŠGGILNKSLCGTEIDKSDVFVRC�LPPTTGDVSKDVGSKT
ST8Sia V	KCAVVGNŠGGILKNŠRCGREINSADFVFRC�LPPISEKYTMDVGVKT
ST3Gal III	RCIIVGNŠGGVLANKSLGSRIDDYDIVRLNSAPV-KGFEEKDVGSKT
ST3Gal V	RCVVIGSGGILHGLELGHHTLNQFDVVIRLNSAPV-EGYSEHVGNT
ST3Gal IV	RCVVVGNŠGHLRNRSSLDGDAINKYDVVIRLNNAPV-AGYEGDVGSKT
ST3Gal VI	KCVVVGNŠGGVLKNKTLEGEKIDSYDVIIIRMNNGPV-LGHEEEVGRRT
ST6GalNAc I	TCAVVGNŠGGILNŠSHMQEIDSHDYVFRLSGALI-KGYEQDVGTRT
ST6GalNAc II	RCVVGNŠGGILNŠSRQGPNI DAHDYVFRLNGAVI-KGFERDVGTKT
ST6Gal I	RCVVSSAGSLKSSQLGREIDDHDAVLRFNŠAPT-ANFQQDVGTKT
ST6Gal II	SCAVVMSAGAILNŠSLGEEIDSHDAVLRFNŠAPT-RGYEKDVGNKT
	* :: . . : * :: . * . . ** :
ST8Sia II	PTTGLLMYTLATRFCKQIYLYGF
ST8Sia IV	PSTGLLMYTLATRFCDIEHLYGF
ST3Gal I	PSTGILSVIFSMHVCDEVDLYGF
ST3Gal II	PSTGMLVLFFALHVCDEVNVYGF
ST6GalNAc III	LSTGWFTFLLAMDACYGIHVYGF
ST6GalNAc IV	LSTGWFTMILALELC EEIVVYGM
ST6GalNAc V	LSTGWFTMTIALELC DRINVYGM
ST6GalNAc VI	LSTGWFTMVI AVELCDHVHVYGM
ST8Sia III	LSTGILMYTLASAICEEIHLYGF
ST8Sia I	LSTGLFLVSAALGLCEEVALYGF
ST8Sia VI	LSTGLMITSV AVELCKNVKLYGF
ST8Sia V	ISTGLILVTA ALELC EEVHLFGF
ST3Gal III	PTLGSVAVTMALHGCDEVAVAGF
ST3Gal V	PTIGVIAVVLATHLCDEVSLAGF
ST3Gal IV	PTTGLLAITLALHLC DLVHIAGF
ST3Gal VI	PTTGIIAITLAFYICHEVHLAGF
ST6GalNAc I	PTTGALLLLTALQLCDQVSAYGF
ST6GalNAc II	PSTGALMLLTALHTCDQVSAYGF
ST6Gal I	PSSGMLGIIIMMTLCDQVDIYEF
ST6Gal II	PSSGFIGILIMMSMCREVHVYEF
	: * . * :
ST8Sia II	ASPHTMPLEFKA
ST8Sia IV	ASPHRMPL EFKT
ST3Gal I	TGVHDADFEŠNV
ST3Gal II	TGVHDADFEAHI
ST6GalNAc III	YGGHRFITEKKV
ST6GalNAc IV	RSAHRFITEKAV
ST6GalNAc V	GSHHRFITEKRV
ST6GalNAc VI	GNHHRFITEKRV
ST8Sia III	QESHQLPAEFQL
ST8Sia I	SGFHAMPEEFLO
ST8Sia VI	HGFHQMPKEYSQ
ST8Sia V	PGGHAMPSEIFN
ST3Gal III	SWTHNIQREKEF
ST3Gal V	QTMHNVTETKF
ST3Gal IV	GSGHNVSQEALA
ST3Gal VI	NAYHNVTAEQLF
ST6GalNAc I	YINHDFKLEREV
ST6GalNAc II	YANHDLŠLEAAL
ST6Gal I	GAYHPLLYEKNL
ST6Gal II	GAYHPLLYEKNL

Figura 2

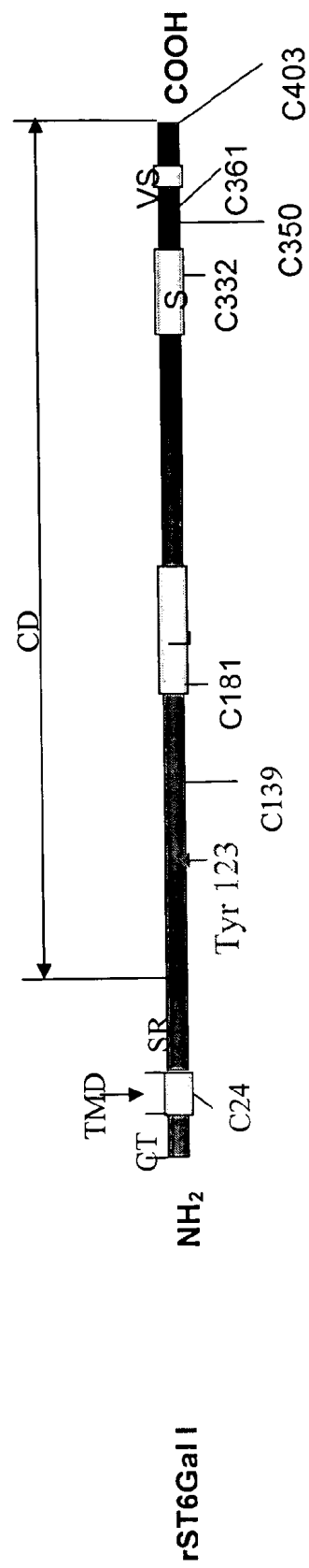


Figura 3

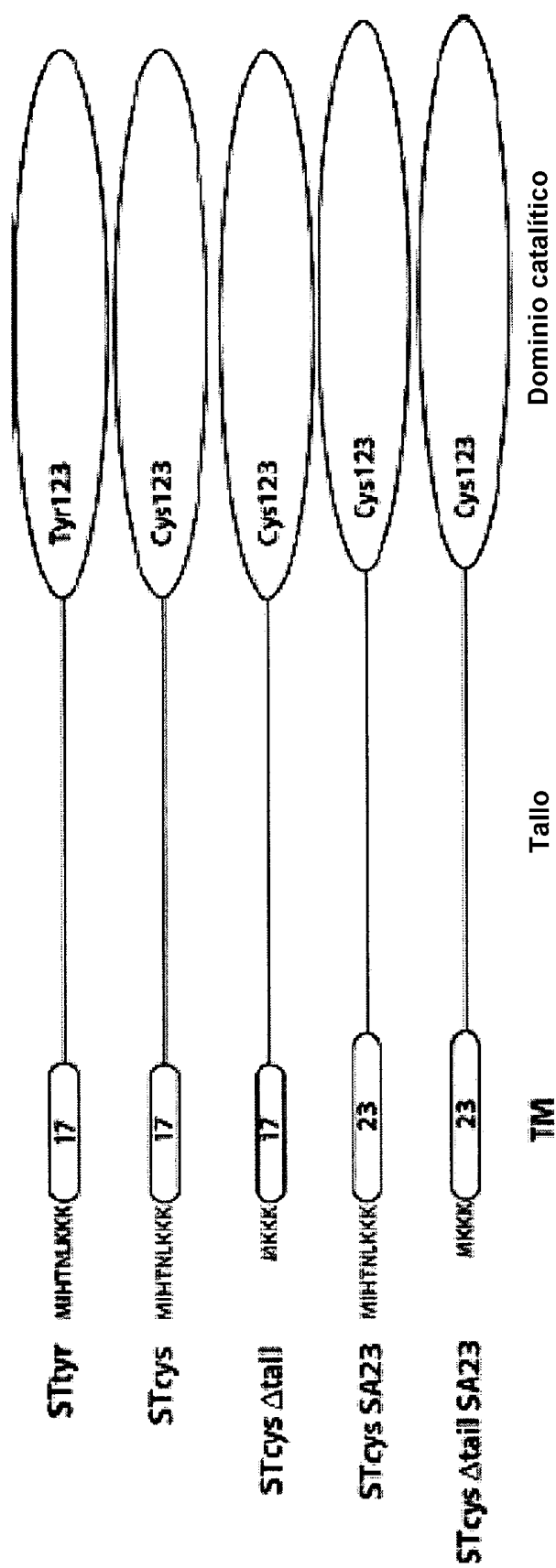


Figura 4

Figura 5

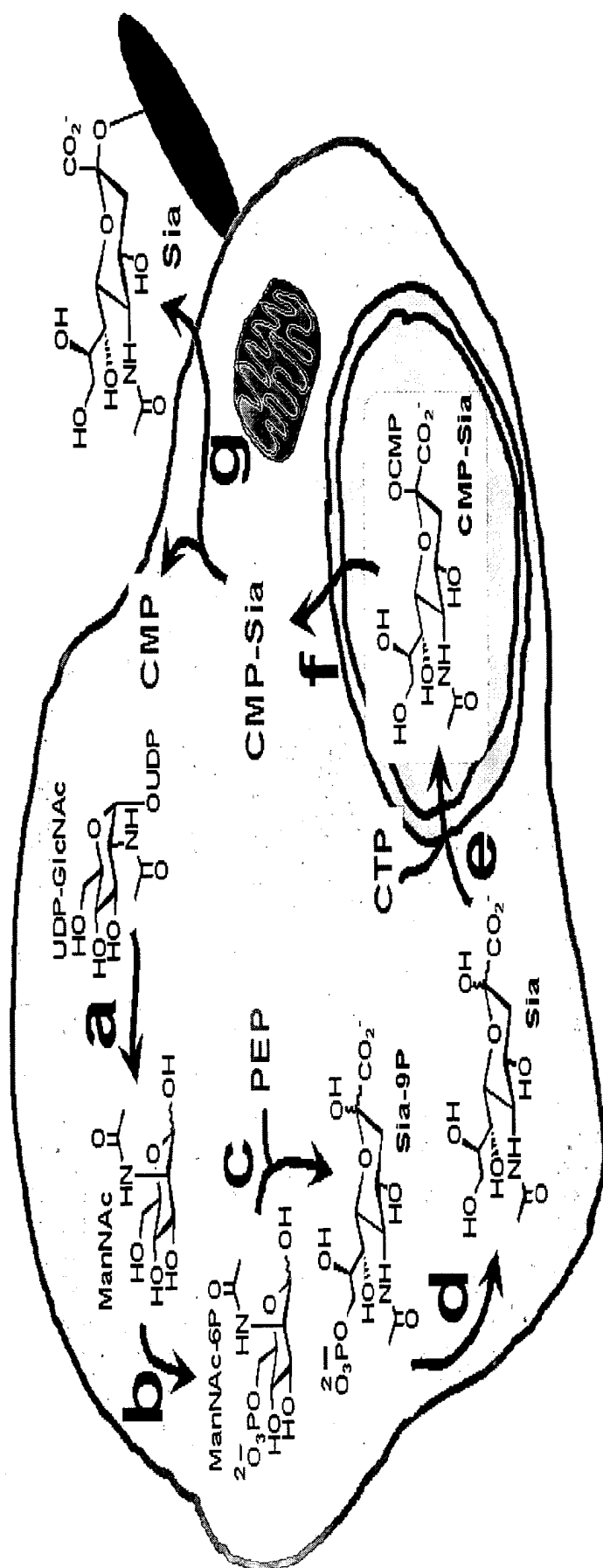


Figura 6

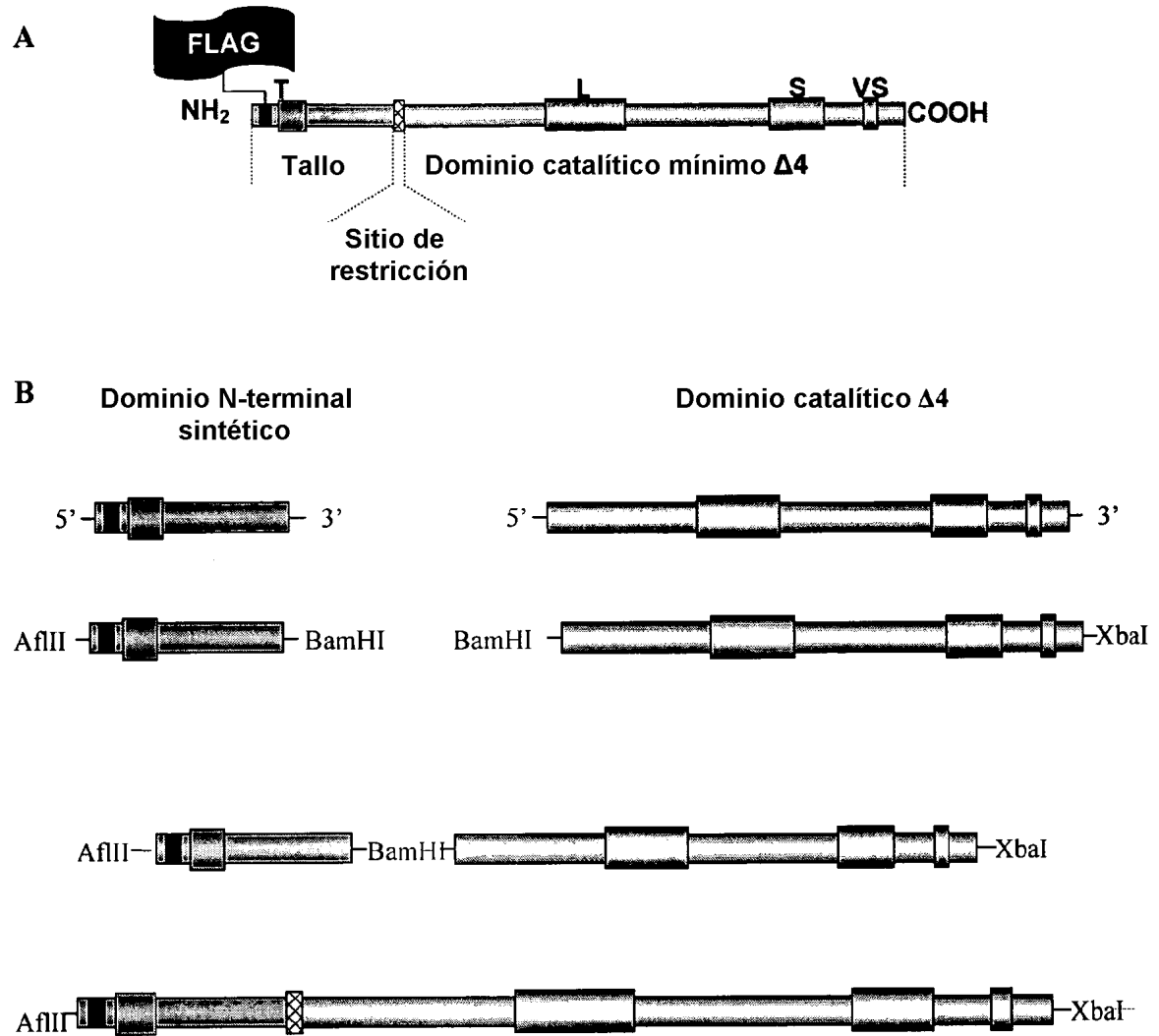


Figura 7

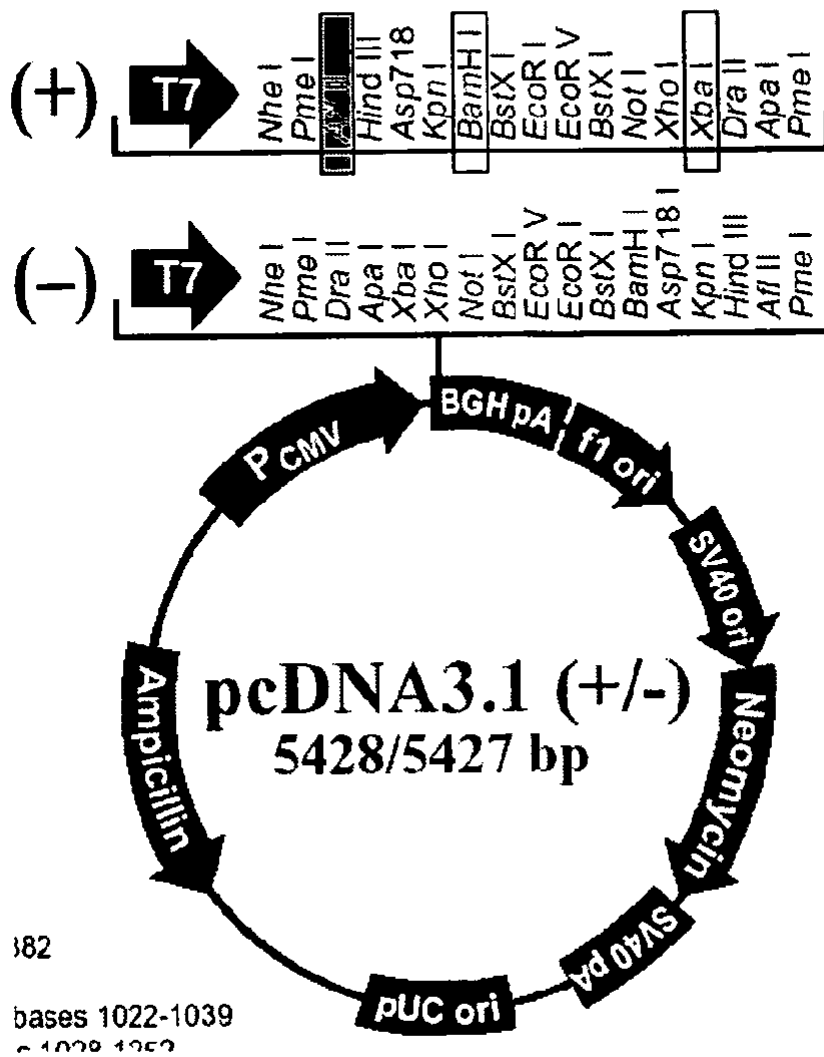


Figura 8

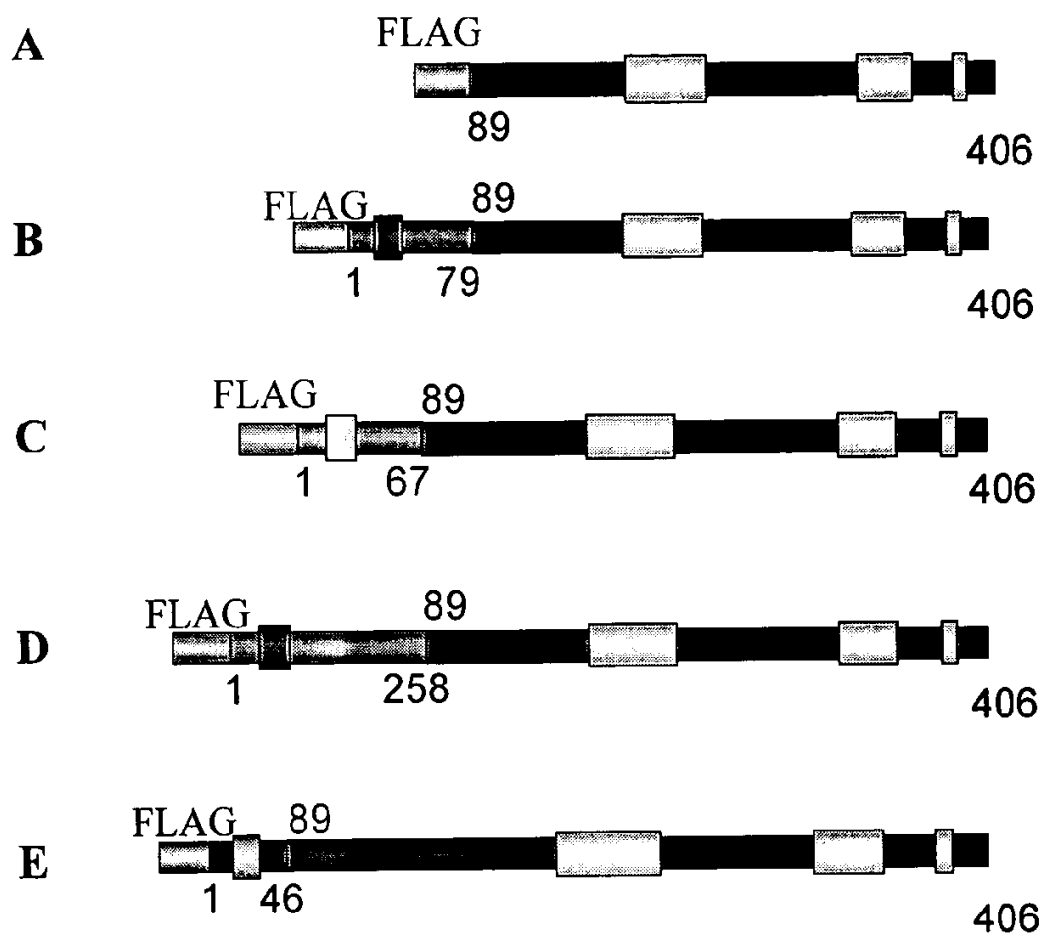


Figura 9

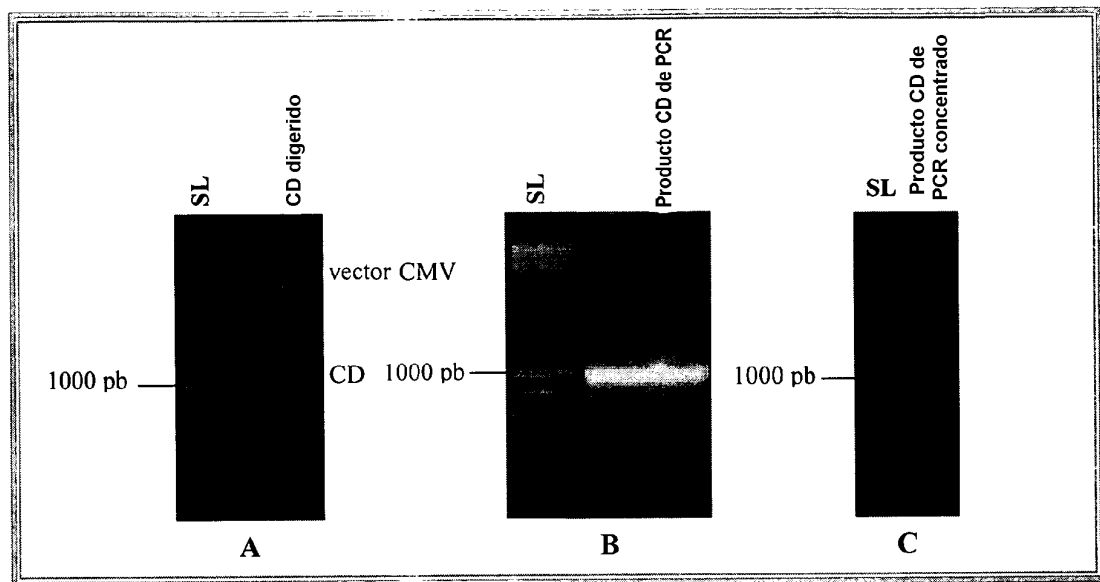


Figura 10

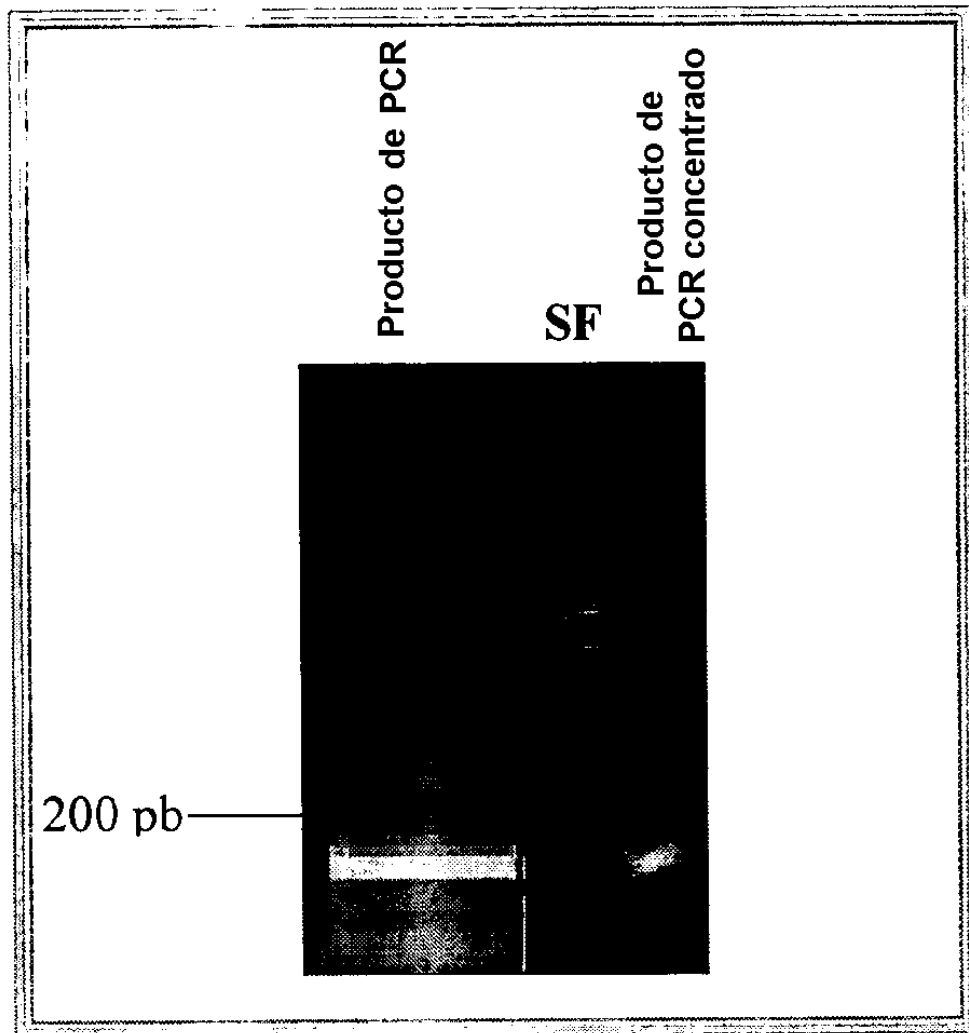


Figura 11

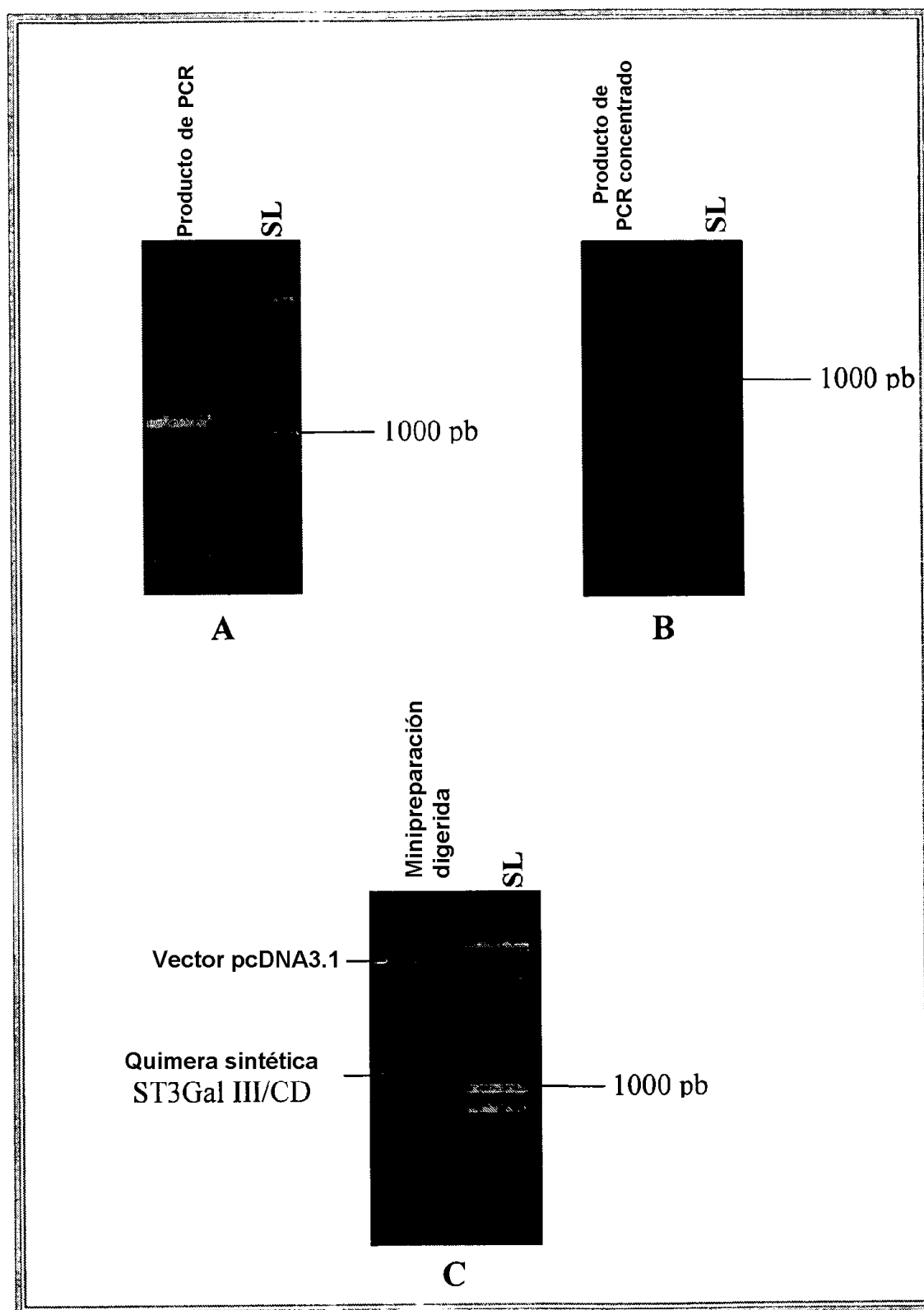


Figura 12

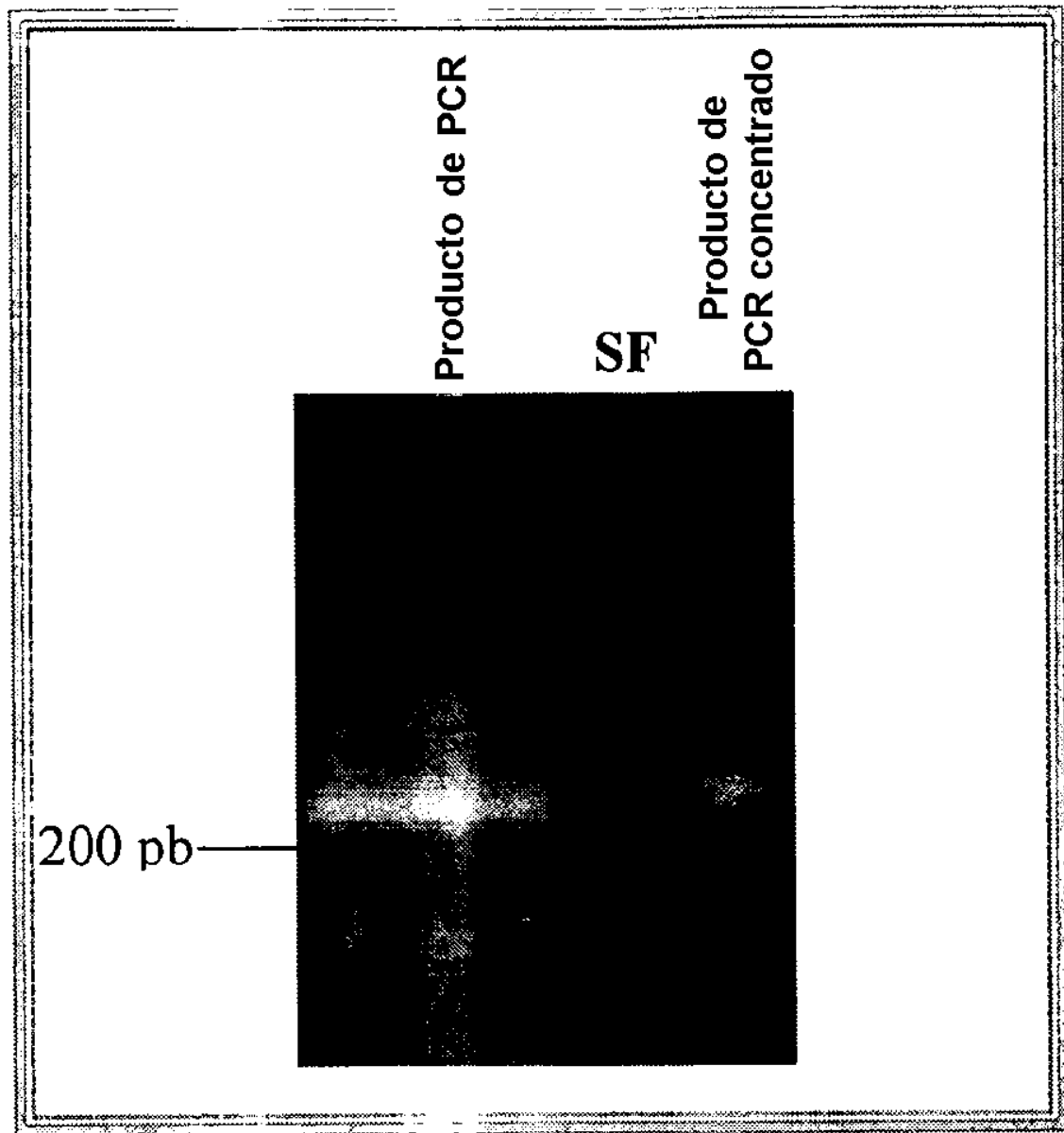


Figura 13

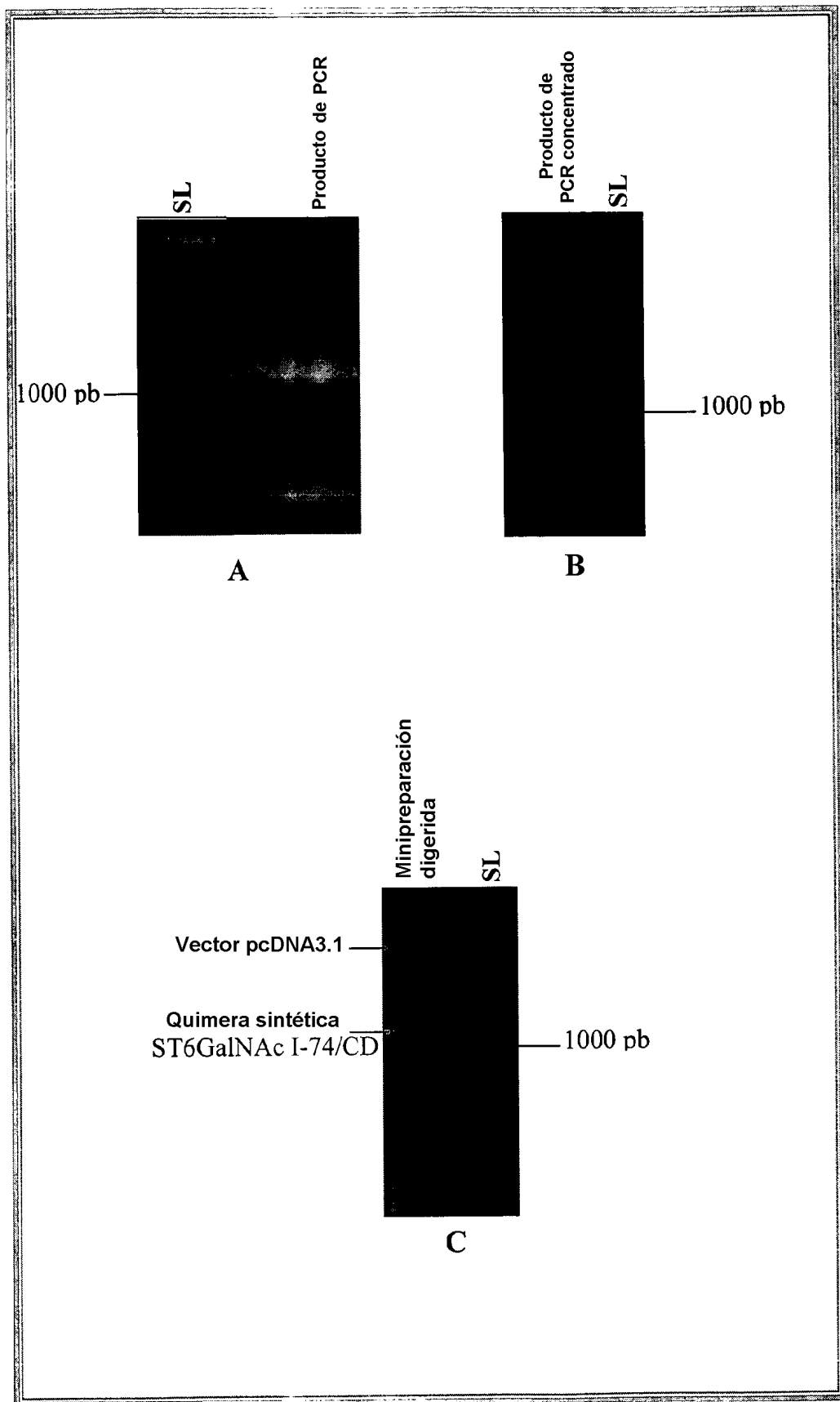


Figura 14

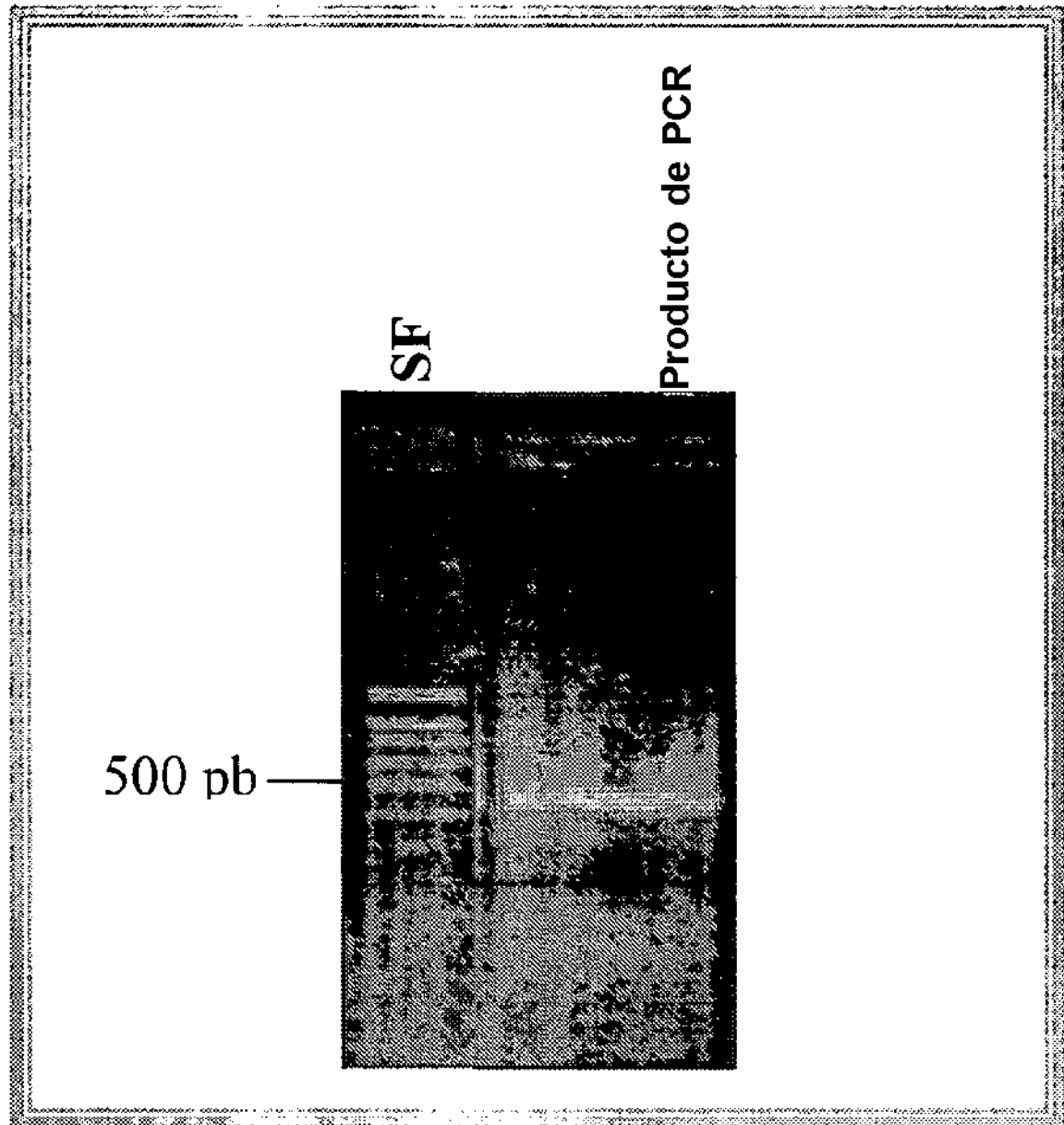


Figura 15

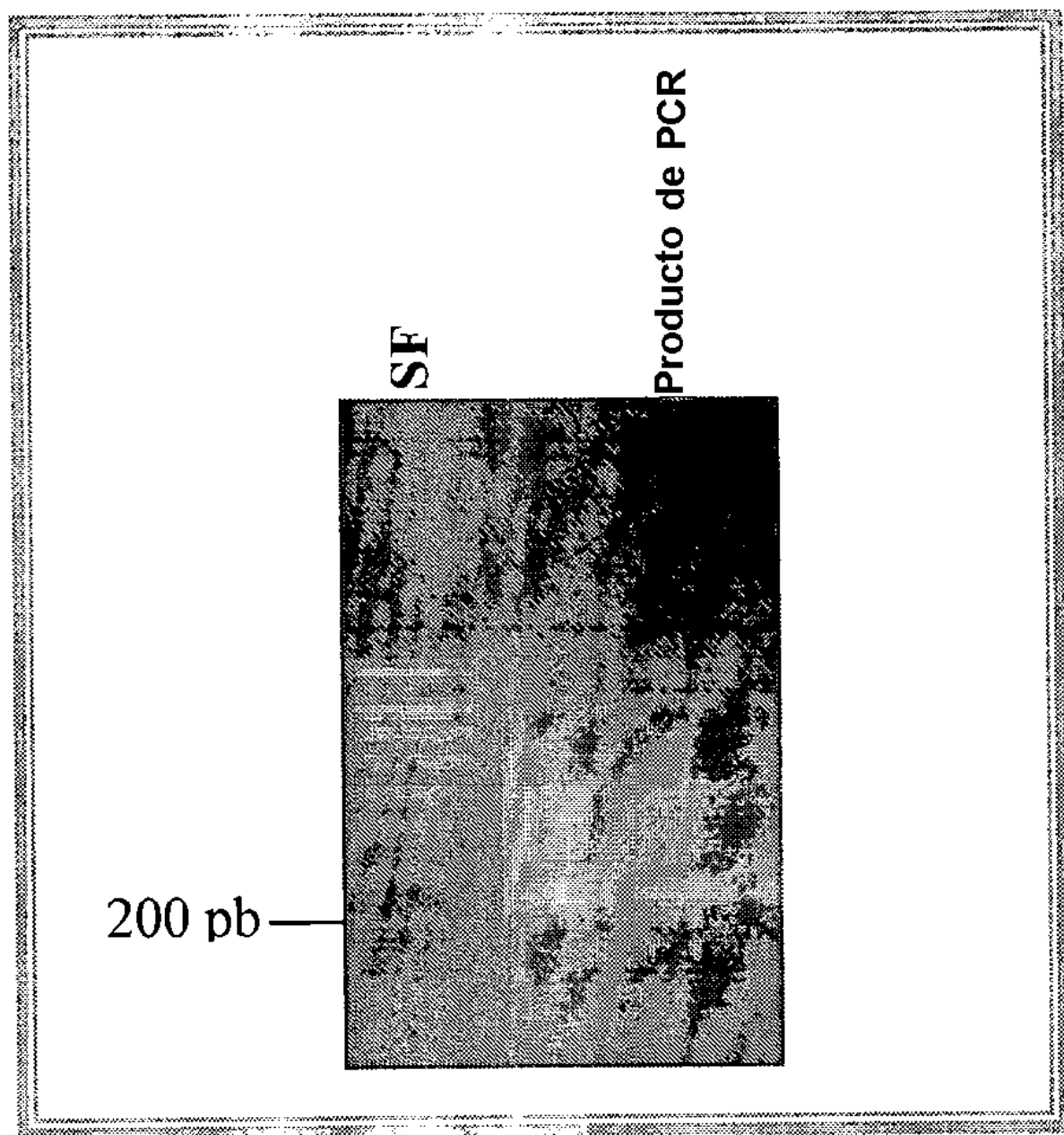


Figura 16

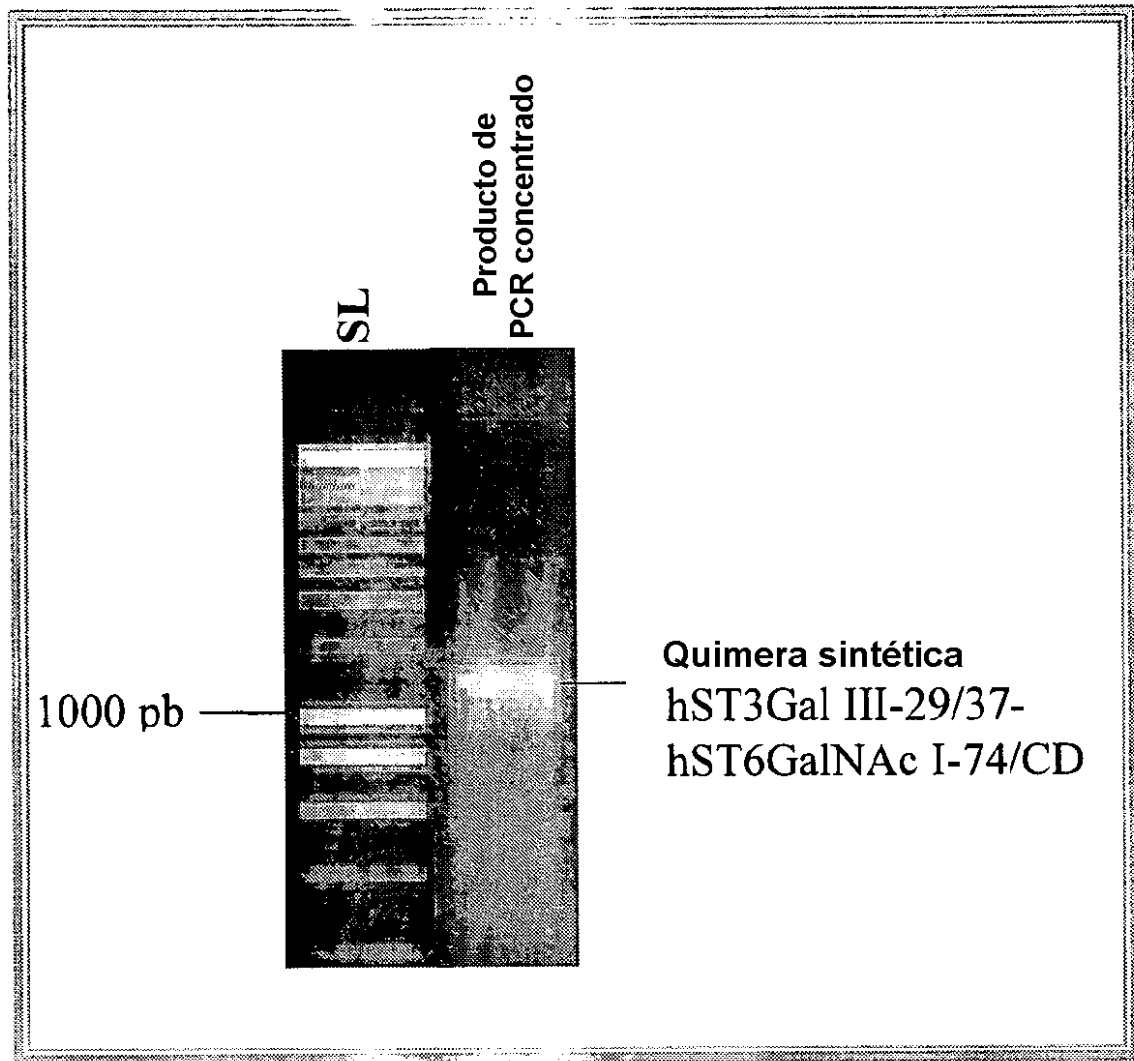


Figura 17

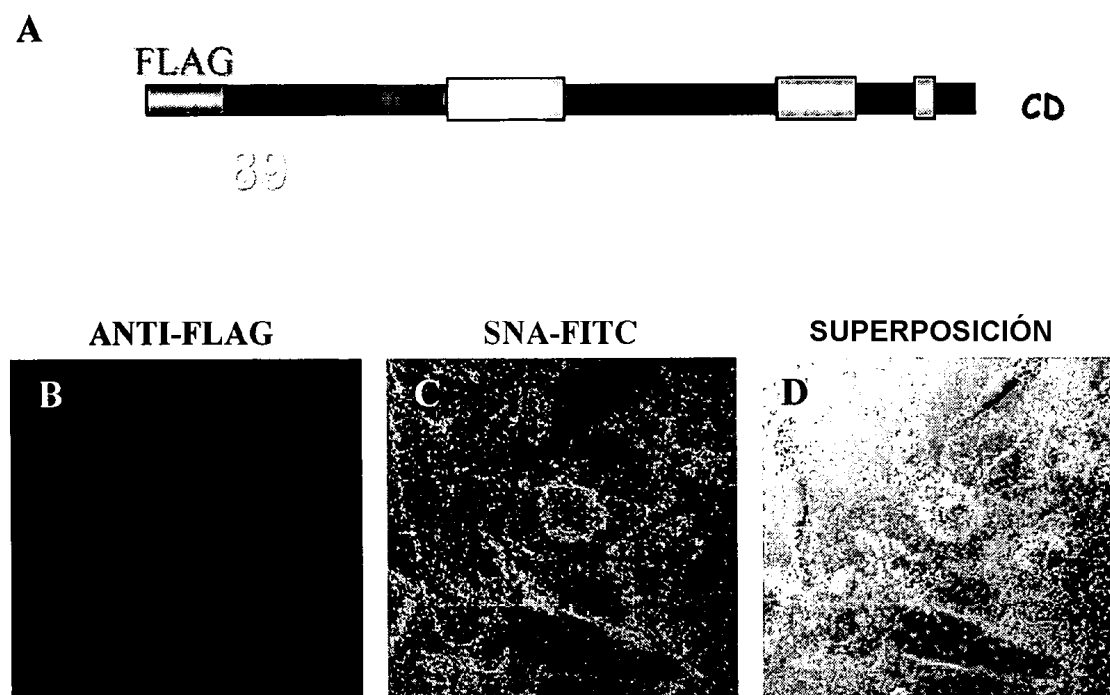


Figura 18

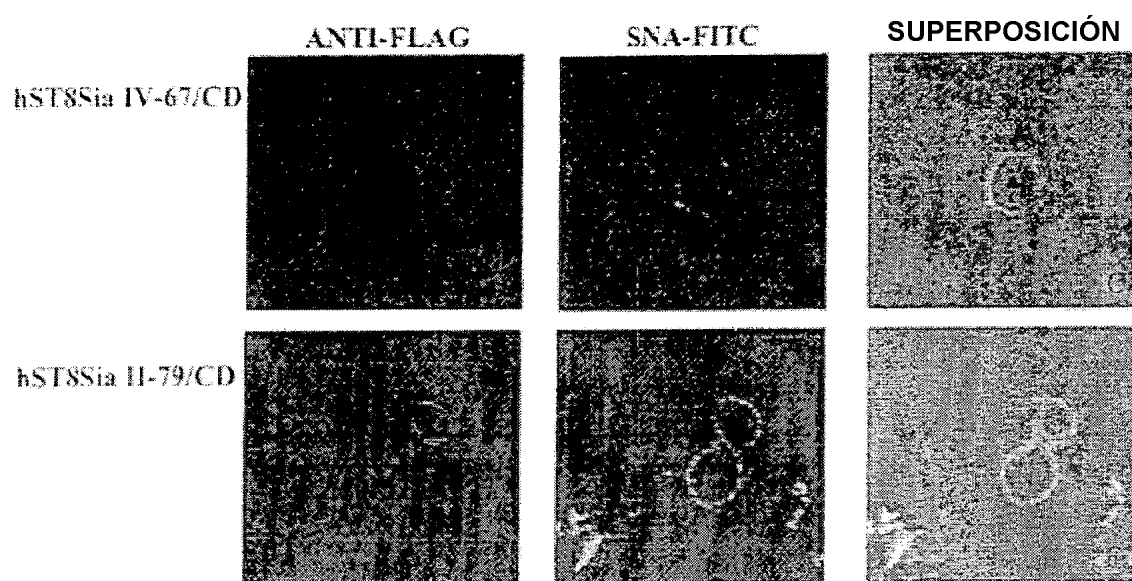


Figura 19

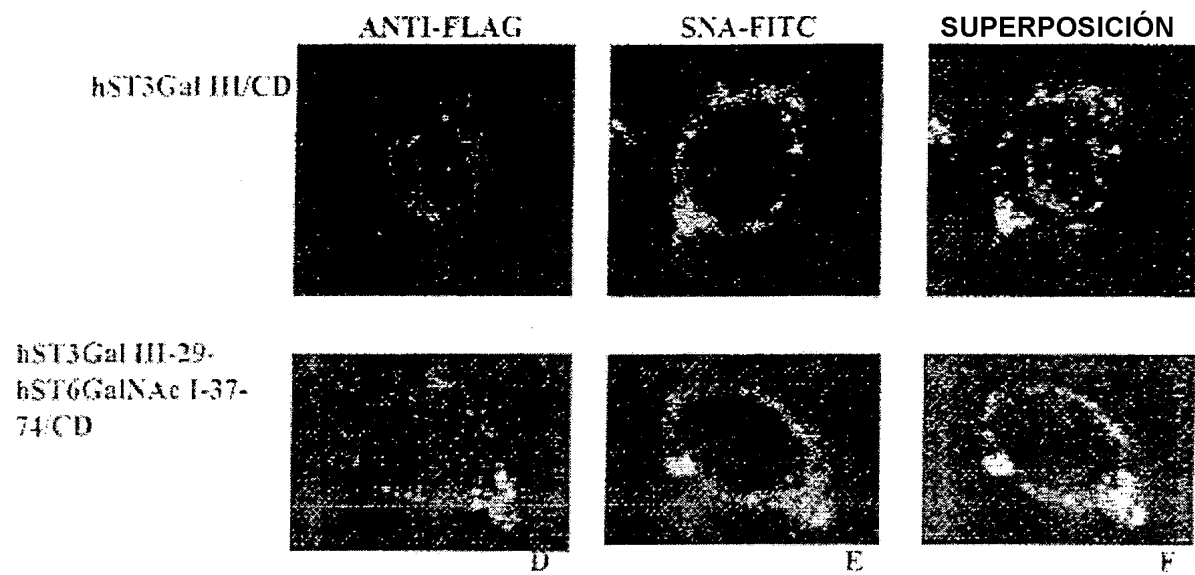


Figura 20

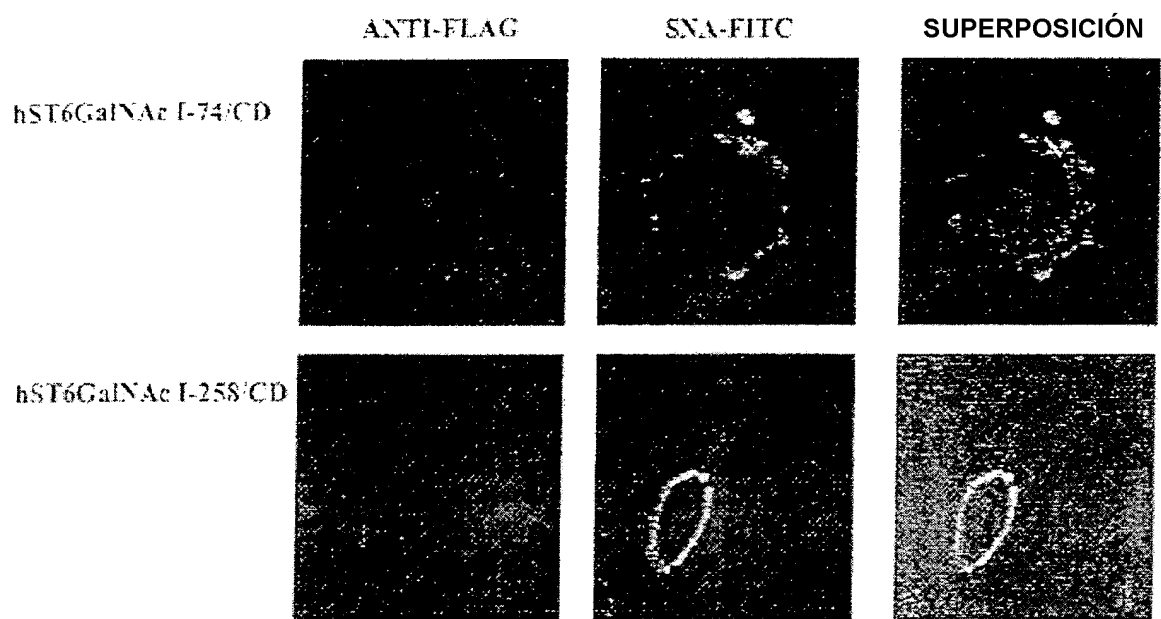


Figura 21

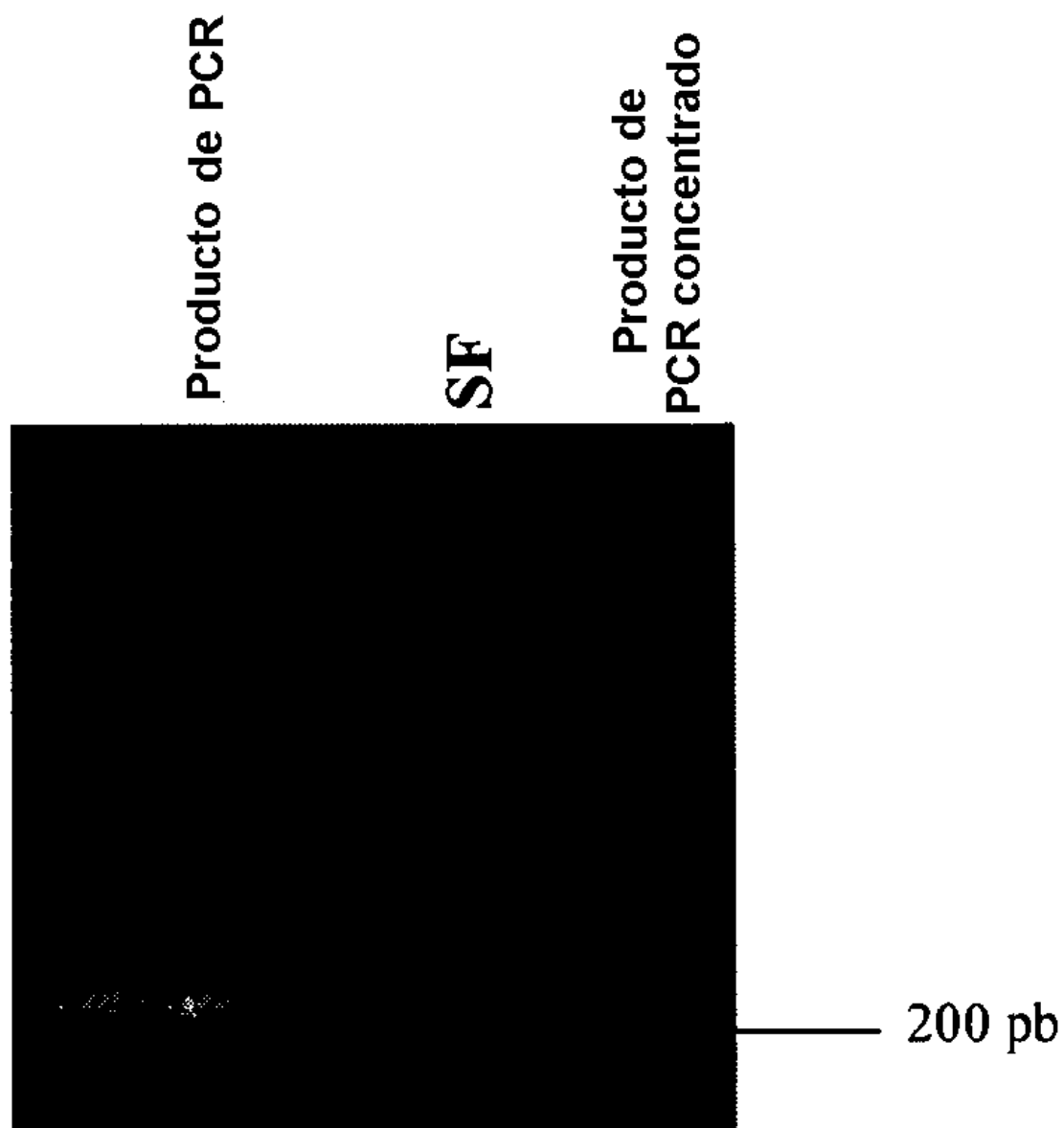


Figura 22

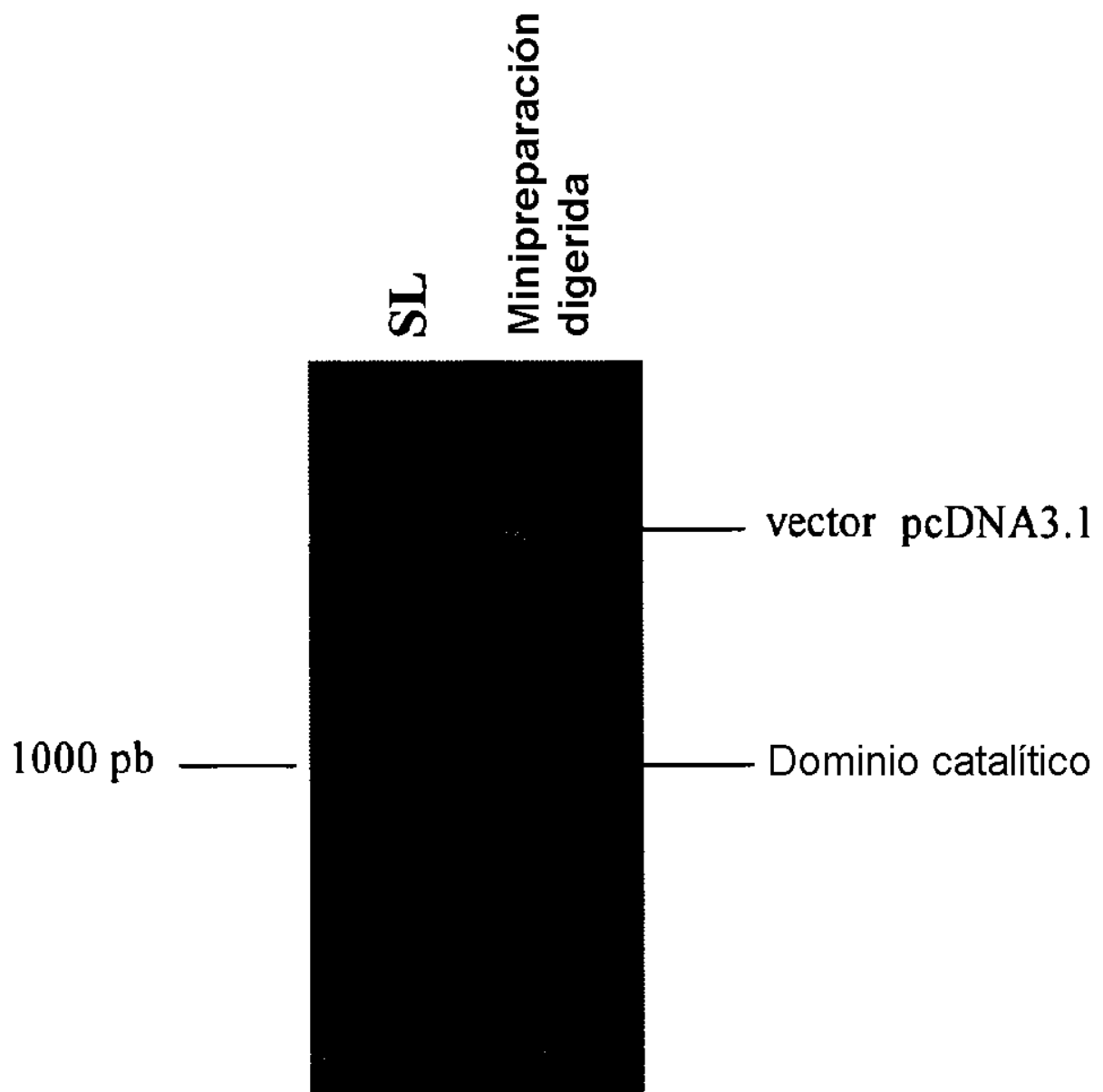


Figura 23

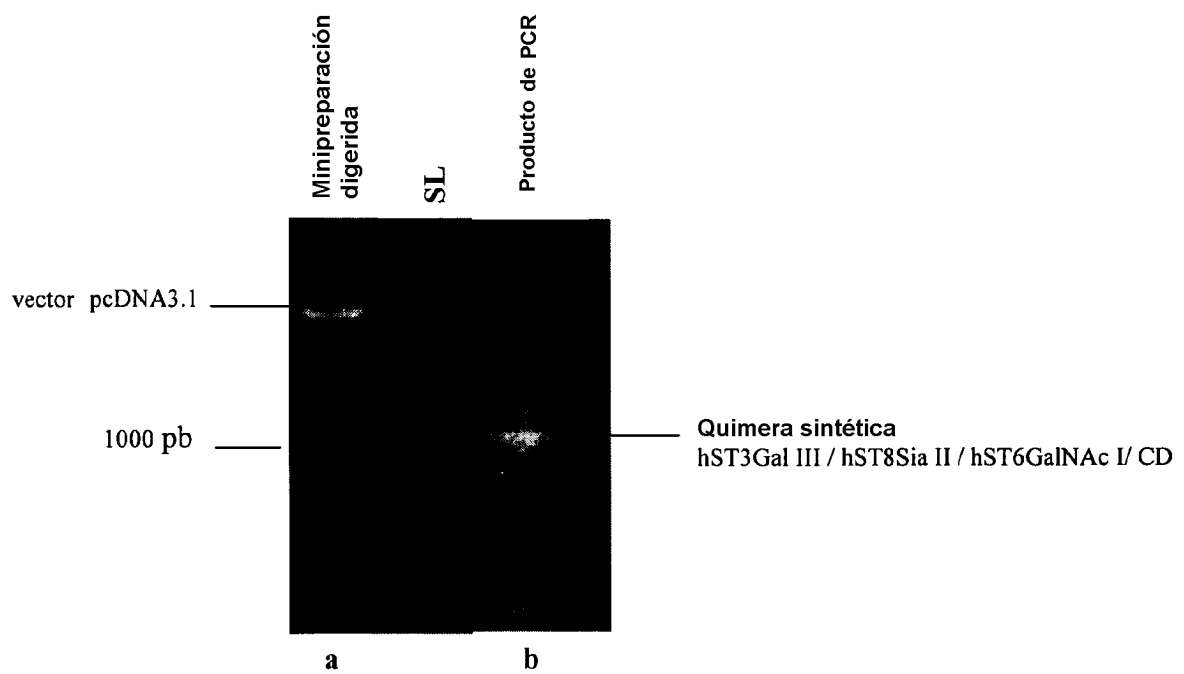


Figura 24

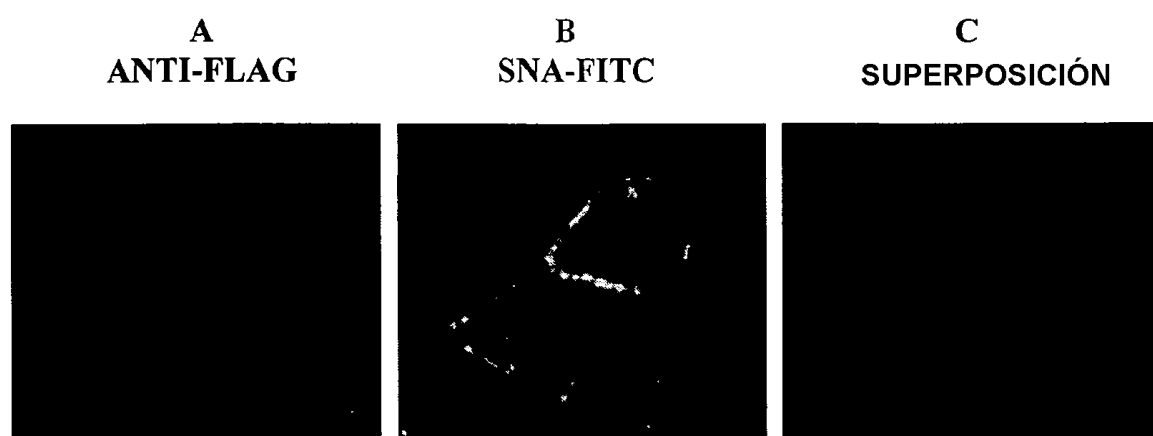


Figura 25