



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 275 358**

51 Int. Cl.:
C12N 15/12 (2006.01)
C07K 14/705 (2006.01)
C12Q 1/68 (2006.01)
C07K 16/28 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **99960200 .6**
86 Fecha de presentación : **05.11.1999**
87 Número de publicación de la solicitud: **1127126**
87 Fecha de publicación de la solicitud: **29.08.2001**

54 Título: **Receptores acoplados a proteína G, homólogos al GPCR 2 (EB1-2) inducido por EBV. Procedimientos para buscar ligandos de los mismos.**

30 Prioridad: **06.11.1998 US 187134**
25.08.1999 US 382918

73 Titular/es: **Bayer HealthCare AG.**
51368 Leverkusen, DE

45 Fecha de publicación de la mención BOPI:
01.06.2007

72 Inventor/es: **Glucksmann, Maria, Alexandra;**
Gu, Wei y
Weich, Nadine, S.

45 Fecha de la publicación del folleto de la patente:
01.06.2007

74 Agente: **Carpintero López, Francisco**

ES 2 275 358 T3

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

DESCRIPCIÓN

Receptores acoplados a proteína G, homólogos al GPCR 2 (EB1-2) inducido por EBV. Procedimientos para buscar ligandos de los mismos.

5 **Campo de la invención**

La presente invención se refiere a un receptor recientemente identificado que pertenece a la superfamilia de receptores acoplados a proteína G. La invención también se refiere a los polinucleótidos que codifican el receptor. La invención además se refiere a procedimientos que usan los polipéptidos y polinucleótidos del receptor como diana para el diagnóstico y tratamiento de trastornos mediados por el receptor. La invención además se refiere a procedimientos de selección de fármacos usando los polipéptidos y polinucleótidos del receptor para identificar agonistas y antagonistas para el diagnóstico y tratamiento. La invención además abarca agonistas y antagonistas basados en los polipéptidos y polinucleótidos del receptor. La invención además se refiere a procedimientos para producir los polipéptidos y polinucleótidos del receptor.

Antecedentes de la invención

Receptores acoplados a proteína G

Los receptores acoplados a proteína G (GPCR) constituyen una clase importante de proteínas responsables de transducir una señal dentro de una célula. Los GPCR tienen tres dominios estructurales: un dominio extracelular amino terminal, un dominio transmembranal que contiene siete segmentos transmembranales, tres bucles extracelulares y tres bucles intracelulares y un dominio intracelular carboxilo terminal. Tras la unión de un ligando a una porción extracelular de un GPCR, se transduce una señal dentro de la célula que da lugar a un cambio en una propiedad biológica o fisiológica de la célula. Los GPCR, junto con las proteínas G y sus efectores (enzimas intracelulares y canales modulados por proteínas G), son los componentes de un sistema de señalización modular que relaciona la condición de segundos mensajeros intracelulares con los estímulos extracelulares.

Los genes GPCR y los productos génicos son potenciales agentes causales de enfermedades (Spiegel y col., *J. Clin. Invest.* 92:1119-1125 (1993); McKusick y col., *J. Med. Genet.* 30:1-26 (1993)). Se han demostrado que defectos específicos en el gen de la rodopsina y en el gen del receptor de la vasopresina V2 causan diversas formas de retinitis pigmentosa (Nathans y col., *Annu. Rev. Genet.* 26:403-424 (1992)) y diabetes insípida nefrogénica (Holtzman y col., *Hum. Mol. Genet.* 2:1201-1204 (1993)). Estos receptores son de importancia crítica tanto para el sistema nervioso central como para procesos fisiológicos periféricos. Los análisis evolutivos sugieren que el ancestro de estas proteínas originalmente se desarrolló de acuerdo con planos corporales y sistemas nerviosos complejos.

La superfamilia de proteínas GPCR puede dividirse en cinco familias: Familia I, receptores tipificados por rodopsina y el receptor β 2-adrenérgico y representada actualmente por más de 200 miembros únicos (Dohlman y col., *Annu. Rev. Biochem.* 60:653-688 (1991)); Familia II, la familia de receptores de la hormona paratiroidea, calcitonina y secretina (Juppner y col., *Science* 254: 1024-1026 (1991); Lin y col., *Science* 254:1022-1024 (1991)), Familia III, la familia del receptor de glutamato metabotrópico (Nakanishi, *Science* 258:597-603 (1992)); Familia IV, la familia del receptor de AMPc, importante en la quimiotaxis y en el desarrollo de *D. discoideum* (Klein y col., *Science* 241:1467-1472 (1988)); y Familia V, los receptores de feromonas de apareamiento de hongos, tal como STE2 (Kurjan, *Annu. Rev. Biochem.* 61:1097-1129 (1992)).

Existe también un pequeño número de otras proteínas que presentan siete probables segmentos hidrófobos y que parecen no estar relacionadas con los GPCR; no se ha demostrado que se acoplen a proteínas G. La *Drosophila* expresa una proteína específica de fotorreceptor, boss (del inglés, *bride of sevenless*), una proteína con siete segmentos transmembranales que se ha estudiado a fondo y no muestra evidencia de ser un GPCR (Hart y col., *Proc. Natl. Acad. Sci. USA* 90:5047-5051 (1993)). También se pensó que el gen *frizzled* (*fz*) de *Drosophila* era una proteína con siete segmentos transmembranales. Al igual que boss, no se ha demostrado que *fz* se acople a proteínas G (Vinson *et al.*, *Nature* 338:263-264 (1989)).

Las proteínas G representan una familia de proteínas heterotriméricas compuestas de subunidades α , β y γ , que se unen a nucleótidos de guanina. Estas proteínas están ligadas generalmente a receptores de la superficie celular, por ejemplo, receptores que contienen siete segmentos transmembranales. Tras la unión del ligando al GPCR, se transmite un cambio conformacional a la proteína G, que hace que la subunidad α intercambie una molécula de GDP unida por una molécula de GTP y que se disocie de las subunidades $\beta\gamma$. La forma unida a GTP de la subunidad α funciona generalmente como un resto modulador del efector, lo que conduce a la producción de segundos mensajeros, tal como el AMPc (por ejemplo, mediante la activación de la adenilato ciclasa), el diacilglicerol o los inositol fosfatos. En humanos se conocen más de 20 tipos diferentes de subunidades α . Estas subunidades se asocian con un grupo más pequeño de subunidades β y γ . Los ejemplos de proteínas G de mamíferos incluyen Gi, Go, Gq, Gs y Gt. Las proteínas G se describen ampliamente en Lodish y col., *Molecular Cell Biology*, (Scientific American Books Inc., Nueva York, N. Y., 1995). Los GPCR, las proteínas G y los sistemas de efectores ligados a proteína G y de segundos mensajeros se han revisado en *The G-Protein Linked Receptor Fact Book*, Watson y compañía, Academic Press editores (1994).

ES 2 275 358 T3

Los GPCR son una diana principal de la acción y desarrollo de fármacos. Por consiguiente, es valioso en el campo del desarrollo farmacéutico identificar y caracterizar GPCR previamente desconocidos. La presente invención avanza el estado de la técnica proporcionando un GPCR humano no identificado previamente.

5 Resumen de la invención

Un objeto de la invención es identificar nuevos GPCR.

10 Un objeto adicional de la invención es proporcionar nuevos polipéptidos de GPCR que sean útiles como reactivos o dianas en ensayos de receptor aplicables al tratamiento y diagnóstico de trastornos mediados por GPCR.

15 Un objeto adicional de la invención es proporcionar polinucleótidos que corresponden a los polipéptidos del nuevo receptor GPCR que sean útiles como dianas y reactivos en ensayos de receptores aplicables al tratamiento y diagnóstico de trastornos mediados por GPCR y útiles para producir nuevos polipéptidos del receptor mediante procedimientos de recombinación.

Un objeto específico de la invención es identificar compuestos que actúen como agonistas y antagonistas y modulen la expresión del nuevo receptor.

20 Un objeto específico adicional de la invención es proporcionar compuestos que modulen la expresión del receptor para el tratamiento y diagnóstico de trastornos relacionados con GPCR.

Por tanto, la invención se basa en la identificación de un nuevo GPCR denominado receptor 15625.

25 La invención proporciona un procedimiento para identificar un agente que module el nivel o la actividad de un polipéptido en una célula o que interactúe con un polipéptido en una célula, en el que dicho polipéptido se selecciona entre:

(a) la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

30 (b) la secuencia de aminoácidos de una variante alélica de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

35 (c) la secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en la que la variante de secuencia es codificada por una molécula de ácido nucleico que hibrida con la molécula de ácido nucleico mostrada en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;

40 (d) un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos;

(e) la secuencia de aminoácidos del polipéptido maduro del receptor de aproximadamente el aminoácido 6 a aproximadamente el aminoácido 342, mostrada en la ID SEC N° 1 ó 3;

45 (f) la secuencia de aminoácidos del polipéptido mostrado en la ID SEC N° 1 ó 3, de aproximadamente el aminoácido 1 a aproximadamente el aminoácido 342; y

(g) la secuencia de aminoácidos de una región portadora de un epítipo de uno cualquiera de los polipéptidos de (a)-(f);

50 y dicho procedimiento comprende:

(i) poner en contacto dicho agente con una célula capaz de expresar dicho polipéptido, de modo que dicho nivel o actividad del polipéptido pueda modularse en dicha célula mediante dicho agente, en el que dicha célula deriva de cerebro, células CD34⁺, células B, músculo esquelético, ganglios linfáticos, bazo, timo, hígado, amígdala, colon, corazón, granulocitos, eritroblastos o páncreas; y

(ii) medir el nivel o la actividad de dicho polipéptido.

60 La invención también proporciona un procedimiento para detectar la presencia de un polipéptido en una muestra, comprendiendo dicho procedimiento:

(i) poner en contacto dicha muestra con un anticuerpo que permita específicamente la detección de la presencia del polipéptido en la muestra; y

65 (ii) detectar la presencia del polipéptido, en el que dicho polipéptido se selecciona entre:

(a) la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

ES 2 275 358 T3

(b) la secuencia de aminoácidos de una variante alélica de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

5 (c) la secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en la que la secuencia variante es codificada por una molécula de ácido nucleico que hibrida con la molécula de ácido nucleico mostrada en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;

10 (d) un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos;

(e) la secuencia de aminoácidos del polipéptido maduro del receptor de aproximadamente el aminoácido 6 a aproximadamente el aminoácido 342, mostrada en la ID SEC N° 1 ó 3;

15 (f) la secuencia de aminoácidos del polipéptido mostrado en la ID SEC N° 1 ó 3, de aproximadamente el aminoácido 1 a aproximadamente el aminoácido 342; y

20 (g) la secuencia de aminoácidos de una región portadora de un epítipo de uno cualquiera de los polipéptidos de (a)-(f);

y en el que dicha muestra deriva de una célula seleccionada entre células de cerebro, CD34⁺, B, músculo esquelético, ganglio linfático, bazo, timo, hígado, amígdala, colon, corazón, granulocitos, eritroblastos y pancreáticas.

25 La invención también proporciona un procedimiento para identificar un agente que module el nivel o actividad de una molécula de ácido nucleico en una célula o que interaccione con una molécula de ácido nucleico en una célula, en el que dicha molécula de ácido nucleico tiene una secuencia de ácido nucleico seleccionada entre:

(a) la secuencia nucleotídica mostrada en la ID SEC N° 2 ó 4;

30 (b) una secuencia nucleotídica que codifica la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

(c) una secuencia nucleotídica complementaria a cualquiera de las secuencias nucleotídicas de (a) o (b);

35 (d) una secuencia nucleotídica que codifica una secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3 que hibrida con la secuencia nucleotídica mostrada en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;

(e) una secuencia nucleotídica complementaria a la secuencia nucleotídica de (d);

40 (f) una secuencia nucleotídica que codifica un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos; y

(g) una secuencia nucleotídica complementaria a la secuencia nucleotídica de (f);

45 y dicho procedimiento comprende:

(i) poner en contacto dicho agente con una célula capaz de expresar dicha molécula de ácido nucleico, de modo que dicho nivel o actividad de la molécula de ácido nucleico se puede modular en dicha célula mediante dicho agente, en el que dicha célula deriva de cerebro, células CD34⁺, células B, músculo esquelético, ganglios linfáticos, bazo, timo, hígado, amígdala, colon, corazón, granulocitos, eritroblastos o páncreas; y

(ii) medir el nivel o la actividad de dicha molécula de ácido nucleico.

55 La invención también proporciona un procedimiento para detectar la presencia de una molécula de ácido nucleico en una muestra, comprendiendo dicho procedimiento:

(i) poner en contacto dicha muestra con un agente que permita específicamente la detección de la presencia de la molécula de ácido nucleico en la muestra; y

60 (ii) detectar la presencia de la molécula de ácido nucleico o un fragmento de la misma, en la que la molécula de ácido nucleico tiene una secuencia nucleotídica seleccionada entre:

(a) la secuencia nucleotídica mostrada en la ID SEC N° 2 ó 4;

65 (b) una secuencia nucleotídica que codifica la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

(c) una secuencia nucleotídica complementaria a cualquiera de las secuencias nucleotídicas de (a) o (b);

- (d) una secuencia nucleotídica que codifica una secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3 que hibrida con la secuencia nucleotídica mostrada en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;
- 5 (e) una secuencia nucleotídica complementaria a la secuencia nucleotídica de (d);
- (f) una secuencia nucleotídica que codifica un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos; y
- 10 (g) una secuencia nucleotídica complementaria a la secuencia nucleotídica de (f);

y en el que dicha muestra deriva de un tejido seleccionado entre cerebro, células CD34+, células B, músculo esquelético, ganglio linfático, bazo, timo, hígado, amígdala, colon, corazón, granulocitos, eritroblastos o páncreas, y en el que dicho agente comprende un oligonucleótido que hibrida con una cualquiera de las secuencias de ácidos nucleicos de (a)-(g) en condiciones rigurosas y la etapa (ii) comprende determinar si el oligonucleótido se une a la secuencia de ácidos nucleicos en la muestra.

15

Descripción de los dibujos

20 La Figura 1 muestra un análisis de la secuencia de aminoácidos de 15625: regiones de giro $\alpha\beta$ y de superhélice; hidrofiliidad; regiones anfipáticas; regiones flexibles; índice antigénico y gráfico de probabilidad de superficie.

La Figura 2 muestra un gráfico de hidrofobicidad del receptor 15625. Los aminoácidos corresponden a 1-342 y muestran los siete segmentos transmembranales.

25

Descripción detallada de la invención

Función del receptor y ruta de señalización

30 La proteína receptora 15625 es un GPCR que participa en las rutas de señalización. Según se usa en este documento, una “ruta de señalización” se refiere a la modulación (por ejemplo, estimulación o inhibición) de una función o actividad celular tras la unión de un ligando al GPCR (proteína 15625). Ejemplos de estas funciones incluyen la movilización de moléculas intracelulares que participan en una ruta de transducción de señal, por ejemplo, fosfatidilinositol 4,5-bifosfato (PIP₂), inositol 1,4,5-trifosfato (IP₃) y adenilato ciclasa; la polarización de la membrana plasmática; la producción o secreción de moléculas; la alteración de la estructura de un componente de la célula; la proliferación de la célula, por ejemplo, la síntesis de ADN; la migración de la célula; la diferenciación celular y la supervivencia de la célula. La proteína receptora 15625 se expresa en todo el cerebro, principalmente en las células de la glía. También se expresa de forma elevada en las células CD34⁺ de la médula ósea que incluyen, pero sin limitaciones, megacariocitos. También se expresa moderadamente en linfocitos B en reposo, disminuyendo el nivel cuando estos linfocitos están activados, y en músculo esquelético. El gen también se expresa en ganglios linfáticos, bazo, timo, hígado, amígdalas, colon, corazón, granulocitos y eritroblastos. También se expresa en placenta y en páncreas. Por consiguiente, las células que participan en una ruta de señalización de la proteína receptora 15625 incluyen, pero sin limitaciones, células derivadas de estos tejidos.

35

40

45 La respuesta mediada por la proteína receptora depende del tipo de célula. Por ejemplo, en algunas células, la unión de un ligando a la proteína receptora puede estimular una actividad, tal como la liberación de compuestos, apertura de un canal, adhesión celular, migración, diferenciación, etc., a través del metabolismo y recambio del fosfatidilinositol o del AMPc, mientras que en otras células la unión del ligando producirá un resultado diferente. Independientemente de la actividad o respuesta celular modulada por la proteína receptora, es universal que la proteína es un GPCR e interacciona con proteínas G para producir una o más señales secundarias, en diversas rutas intracelulares de transducción de la señal, por ejemplo, a través del metabolismo y recambio del fosfatidilinositol o del AMPc, en una célula.

50

Según se usa en este documento, el “recambio y metabolismo del fosfatidilinositol” se refiere a las moléculas implicadas en el recambio y metabolismo del fosfatidilinositol 4,5-bifosfato (PIP₂), así como las actividades de estas moléculas. PIP₂ es un fosfolípido que se encuentra en la cara citosólica de la membrana plasmática. La unión del ligando al receptor activa, en algunas células, la enzima fosfolipasa C de la membrana plasmática que a su vez puede hidrolizar el PIP₂ para producir 1,2-diacilglicerol (DAG) y el inositol 1,4,5-trifosfato (IP₃). Una vez formado el IP₃ puede difundir a la superficie del retículo endoplásmico donde se une a un receptor de IP₃, por ejemplo, una proteína del canal de calcio que contiene un sitio de unión a IP₃. La unión de IP₃ puede inducir la apertura del canal, permitiendo que los iones de calcio se liberen al citoplasma. El IP₃ también puede fosforilarse mediante una cinasa específica para formar inositol 1,3,4,5-tetrafosfato (IP₄), una molécula que puede causar la entrada de calcio en el citoplasma desde el medio extracelular. Posteriormente, IP₃ e IP₄ pueden hidrolizarse muy rápidamente en los productos inactivos inositol 1,4-bifosfato (IP₂) e inositol 1,3,4-trifosfato, respectivamente. Estos productos inactivos pueden ser reciclados por la célula para sintetizar PIP₂. El otro segundo mensajero producido por la hidrólisis de PIP₂, concretamente, 1,2-diacilglicerol (DAG), permanece en la membrana celular donde puede servir para activar la enzima proteína cinasa C. La proteína cinasa C se encuentra habitualmente soluble en el citoplasma de la célula, pero tras un aumento de la concentración de calcio intracelular, esta enzima puede trasladarse a la membrana plasmática donde puede ser activada por el DAG. La activación de la proteína cinasa C en diferentes células da lugar a diversas respuestas celulares, tales

55

60

65

ES 2 275 358 T3

como la fosforilación de glucógeno sintetasa o la fosforilación de diversos factores de transcripción, por ejemplo, NF- κ B. La expresión “actividad fosfatidilinositol”, según se usa en este documento, se refiere a una actividad de PIP₂ o de uno de sus metabolitos.

5 Otra ruta de señalización en la que puede participar el receptor es la ruta de recambio del AMPc. Según se usa en este documento, el “recambio y metabolismo del AMP cíclico” se refiere a las moléculas implicadas en el recambio y metabolismo del AMP cíclico (AMPc), así como a las actividades de estas moléculas. El AMP cíclico es un segundo mensajero producido en respuesta a la estimulación inducida por ligando de ciertos receptores acoplados a proteína G. En la ruta de señalización del AMPc, la unión de un ligando a un GPCR puede conducir a la activación de la enzima
10 adenilato ciclasa, que cataliza la síntesis del AMPc. El AMPc recién sintetizado puede a su vez activar una proteína cinasa dependiente de AMPc. Esta cinasa activada puede fosforilar una proteína del canal de potasio dependiente de voltaje, o una proteína asociada, y conduce a la incapacidad del canal de potasio para abrirse durante un potencial de acción. La incapacidad del canal de potasio para abrirse da lugar a un descenso en el flujo de salida del potasio, lo que normalmente repolariza la membrana de una neurona, conduciendo a la despolarización prolongada de la membrana.

15 *Polipéptidos*

La invención se basa en el descubrimiento de un nuevo receptor acoplado a proteína G. Específicamente, se seleccionó una etiqueta de secuencia expresada (EST) en función de la homología con las secuencias del receptor acoplado a proteína G. Esta EST se usó para diseñar cebadores basados en las secuencias que contiene y se usó para identificar
20 un ADNc a partir de una biblioteca de ADNc humano. Los clones positivos se secuenciaron y se ensamblaron los fragmentos solapantes. El análisis de la secuencia ensamblada reveló que la molécula de ADNc clonada codifica un receptor acoplado a proteína G.

25 La invención también abarca una secuencia nucleotídica de ADNc (ID SEC N° 4) y la secuencia de aminoácidos deducida (ID SEC N° 3) de un primate no humano (cerebro de macaco), que se corresponden con las secuencias de aminoácidos y nucleotídica del receptor 15625 humano.

De este modo, la invención se refiere a un nuevo GPCR que tiene la secuencia de aminoácidos deducida mostrada
30 en la (ID SEC N° 1).

Se predice que los aminoácidos 1-25 constituyen el dominio extracelular amino terminal, los aminoácidos 26-302 constituyen la región que abarca el dominio transmembranal y los aminoácidos 303-342 constituyen el dominio intracelular carboxilo terminal. El dominio transmembranal contiene siete segmentos transmembranales, tres bucles
35 extracelulares y tres bucles intracelulares. Los segmentos transmembranales se encuentran de aproximadamente el aminoácido 26 a aproximadamente el aminoácido 47, de aproximadamente el aminoácido 59 a aproximadamente el aminoácido 79, de aproximadamente el aminoácido 99 a aproximadamente el aminoácido 120, de aproximadamente el aminoácido 143 a aproximadamente el aminoácido 162, de aproximadamente el aminoácido 189 a aproximadamente el aminoácido 212, de aproximadamente el aminoácido 238 a aproximadamente el aminoácido 255 y de aproxima-
40 damente el aminoácido 284 a aproximadamente el aminoácido 302. Dentro de la región que abarca el dominio transmembranal completo hay tres bucles intracelulares y tres extracelulares. Los tres bucles intracelulares se encuentran de aproximadamente el aminoácido 48 a aproximadamente el aminoácido 58, de aproximadamente el aminoácido 121 a aproximadamente el aminoácido 142 y de aproximadamente el aminoácido 213 a aproximadamente el aminoácido 237. Los tres bucles extracelulares se encuentran de aproximadamente el aminoácido 80 a aproximadamente el aminoácido 98, de aproximadamente el aminoácido 163 a aproximadamente el aminoácido 188 y de aproximadamente el aminoácido 256 a aproximadamente el aminoácido 283.

El dominio transmembranal incluye un motivo característico de la transducción de la señal de GPCR, DRY, en los restos 121-123. La secuencia incluye una arginina en el resto 122, un aminoácido invariable en los GPCR.
50

Una comparación del receptor 15625 con la base de datos Prosite de patrones de proteínas muestra una alta puntuación frente a la superfamilia de la rodopsina con siete segmentos transmembranales. El área subrayada muestra un motivo característico de GPCR, y específicamente la posición de un resto de arginina, conservado en los GPCR. La secuencia conservada con más frecuencia es un triplete aspartato, arginina, tirosina (DRY). DRY está implicada en
55 la transducción de la señal. La arginina no varía. El aspartato se sitúa de forma conservada en varios GPCR. En el caso presente, la arginina se encuentra en la secuencia DRY, que coincide con la posición de DRY o de la arginina invariable en los GPCR de la superfamilia de receptores de la rodopsina.

Un análisis de la fase de lectura abierta de 15625 muestra los aminoácidos correspondientes a los sitios funcionales
60 específicos. Se encuentra un sitio de N-glucosilación aproximadamente en los aminoácidos 6-9 y 13-16. Se encuentra un sitio de fosforilación de proteína cinasa dependiente de AMPc y GMPc aproximadamente en los aminoácidos 173-176. Se encuentran sitios de fosforilación de proteína cinasa C aproximadamente en los aminoácidos 126-128, 163-165 y 304-306. Se encuentra sitios de N-miristilación aproximadamente en los aminoácidos 39-44 y 333-338. Además, los aminoácidos correspondientes en posición con el motivo característico de GPCR y que contienen la arginina invariable
65 se encuentran en la secuencia DRY en los aminoácidos 121-123.

El “polipéptido receptor 15625” o la “proteína receptora 15625” se refieren al polipéptido de la ID SEC N° 1. Sin embargo, el término “proteína del receptor” o “polipéptido del receptor” incluye además las numerosas variantes

ES 2 275 358 T3

descritas en este documento, así como los fragmentos derivados del polipéptido 15625 de longitud completa y sus variantes, por ejemplo, la ID SEC N° 3.

De este modo, la presente invención proporciona un polipéptido receptor 15625 aislado o purificado y variantes y fragmentos del mismo.

El polipéptido 15625 es una proteína de 342 restos que muestra tres dominios estructurales principales. Se identifica que el dominio extracelular amino terminal está en los restos 1 a aproximadamente 25 de la ID SEC N° 1. Se identifica que el dominio transmembranal está en los restos de aproximadamente 26 a aproximadamente 302 de la ID SEC N° 1. Se identifica que el dominio intracelular carboxilo terminal está en los restos de aproximadamente 303 al 342 de la ID SEC N° 1. El dominio transmembranal contiene siete segmentos que atraviesan la membrana. Los segmentos transmembranales se encuentran de aproximadamente el aminoácido 26 a aproximadamente el aminoácido 47, de aproximadamente el aminoácido 59 a aproximadamente el aminoácido 79, de aproximadamente el aminoácido 99 a aproximadamente el aminoácido 120, de aproximadamente el aminoácido 143 a aproximadamente el aminoácido 162, de aproximadamente el aminoácido 189 a aproximadamente el aminoácido 212, de aproximadamente el aminoácido 238 a aproximadamente el aminoácido 255 y de aproximadamente el aminoácido 284 a aproximadamente el aminoácido 302. Dentro de la región que abarca el dominio transmembranal completo hay tres bucles intracelulares y tres extracelulares. Los tres bucles intracelulares se encuentran de aproximadamente el aminoácido 48 a aproximadamente el aminoácido 58, de aproximadamente el aminoácido 121 a aproximadamente el aminoácido 142 y de aproximadamente el aminoácido 213 a aproximadamente el aminoácido 237. Los tres bucles extracelulares se encuentran de aproximadamente el aminoácido 80 a aproximadamente el aminoácido 98, de aproximadamente el aminoácido 163 a aproximadamente el aminoácido 188 y de aproximadamente el aminoácido 256 a aproximadamente el aminoácido 283.

El dominio transmembranal incluye un motivo característico de la transducción de la señal de GPCR, DRY, en los restos 121-123. La secuencia incluye una arginina en el resto 122, un aminoácido invariable en los GPCR.

Basándose en una búsqueda en BLAST, mostró la mayor homología con la superfamilia de GPCR de la rodopsina.

Según se usa en este documento, un polipéptido se dice que está “aislado” o “purificado” cuando está sustancialmente libre de material celular cuando se aísla a partir de células recombinantes y no recombinantes, o libre de precursores químicos u otros compuestos químicos cuando se sintetiza químicamente. Sin embargo, un polipéptido puede estar unido a otro polipéptido con el que normalmente no está asociado en una célula y seguir siendo considerado “aislado” o “purificado”.

Los polipéptidos receptores pueden purificarse a homogeneidad. Se entiende, sin embargo, que las preparaciones en las que el polipéptido no está purificado a homogeneidad son útiles y se considera que contienen una forma aislada del polipéptido. El aspecto crítico es que la preparación permita la función deseada del polipéptido, incluso en presencia de cantidades considerables de otros componentes. De este modo, la invención abarca diversos grados de pureza.

En una realización, la expresión “sustancialmente libre de material celular” incluye preparaciones del polipéptido receptor con menos de aproximadamente el 30% (en peso seco) de otras proteínas (es decir, proteínas contaminantes), menos de aproximadamente el 20% de otras proteínas, menos de aproximadamente el 10% de otras proteínas o menos de aproximadamente el 5% de otras proteínas. Cuando el polipéptido receptor se produce de forma recombinante, también puede estar sustancialmente libre de medio de cultivo, es decir, el medio de cultivo representa menos de aproximadamente el 20%, menos de aproximadamente el 10% o menos de aproximadamente el 5% del volumen de la preparación proteica.

La expresión “sustancialmente libre de precursores químicos u otros compuestos químicos” incluye preparaciones del polipéptido receptor en las que se separa de precursores químicos u otros compuestos químicos que están implicados en su síntesis. En una realización, la expresión “sustancialmente libre de precursores químicos u otros compuestos químicos” incluye preparaciones del polipéptido con menos de aproximadamente el 30% (en peso seco) de precursores químicos u otros compuestos químicos, menos de aproximadamente el 20% de precursores químicos u otros compuestos químicos, menos de aproximadamente el 10% de precursores químicos u otros compuestos químicos o menos de aproximadamente el 5% de precursores químicos u otros compuestos químicos.

En una realización, el polipéptido receptor comprende la secuencia de aminoácidos mostrada en la ID SEC N° 1. Sin embargo, la invención también abarca variantes de la secuencia. Las variantes incluyen una proteína sustancialmente homóloga codificada por el mismo locus genético en un organismo, es decir, una variante alélica. El receptor 15625 se ha mapeado en el cromosoma 3, próximo al marcador AFM 64YG9. Las variantes también abarcan proteínas derivadas de otros loci genéticos en un organismo, pero que tienen homología sustancial con la proteína receptora 15625 de la ID SEC N° 1. Las variantes también incluyen proteínas sustancialmente homólogas con la proteína receptora 15625, pero derivadas de otro organismo, es decir, un ortólogo, tal como la ID SEC N° 3. Las variantes también incluyen proteínas que son sustancialmente homólogas a la proteína receptora 15625 que se producen mediante síntesis química. Las variantes también incluyen proteínas que son sustancialmente homólogas a la proteína receptora 15625 que se producen mediante procedimientos de recombinación. Se entiende, sin embargo, que las variantes excluyen cualquier secuencia de aminoácidos descrita antes de la invención.

ES 2 275 358 T3

Según se usa en este documento, dos proteínas (o una región de las proteínas) son sustancialmente homólogas cuando las secuencias de aminoácidos son al menos aproximadamente el 50-55%, el 55-60%, típicamente al menos aproximadamente el 70-75%, más típicamente al menos aproximadamente el 80-85% y más típicamente al menos aproximadamente el 90-95% o más homólogas. Según la presente invención, una secuencia de aminoácidos sustancialmente homóloga estará codificada por una secuencia de ácido nucleico que hibrida con una secuencia de ácido nucleico, o porción de la misma, de la secuencia mostrada en la ID SEC N° 2 en condiciones rigurosas, como se describe más ampliamente a continuación.

Para determinar el porcentaje de identidad de dos secuencias de aminoácidos o de dos secuencias de ácido nucleico, las secuencias se alinean para una comparación óptima (por ejemplo, para el alineamiento óptimo pueden introducirse huecos en una o en ambas secuencias de aminoácidos o de ácido nucleico primera y segunda y pueden excluirse las secuencias no homólogas para la comparación). En una realización preferida, la longitud de una secuencia de referencia alineada para la comparación es al menos el 30%, preferiblemente al menos el 40%, más preferiblemente al menos el 50%, incluso más preferiblemente al menos el 60% y aún más preferiblemente al menos el 70%, 80% ó 90% de la longitud de la secuencia de referencia (por ejemplo, en la alineación de una segunda secuencia con las secuencias de aminoácidos de este documento, que tiene 342 restos de aminoácido, se alinean al menos 80, preferiblemente al menos 100, más preferiblemente al menos 120, incluso más preferiblemente al menos 140 y aún más preferiblemente al menos 200, 250, 300 ó 340 restos de aminoácidos). A continuación se comparan los restos de aminoácidos o nucleótidos en las correspondientes posiciones de aminoácidos o de nucleótidos. Cuando una posición en la primera secuencia está ocupada por el mismo resto de aminoácido o nucleótido que la posición correspondiente de la segunda secuencia, entonces las moléculas son idénticas en esa posición (según se usa en este documento, la "identidad" de aminoácido o ácido nucleico es equivalente a la "homología" de aminoácido o de ácido nucleico). El porcentaje de identidad entre las dos secuencias está en función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos y la longitud de cada hueco, que es necesario introducir para un alineamiento óptimo de las dos secuencias.

La invención también abarca polipéptidos que tengan un grado de identidad menor pero similitud suficiente como para realizar una o más de las mismas funciones realizadas por el polipéptido 15625. La similitud se determina mediante la sustitución de aminoácidos conservados. Estas sustituciones son aquellas que sustituyen un aminoácido determinado en un polipéptido por otro aminoácido de características similares. Las sustituciones conservadoras son probablemente fenotípicamente imperceptibles. Generalmente, las que se ven como sustituciones conservadoras son las sustituciones, de uno por otro, entre los aminoácidos alifáticos Ala, Val, Leu e Ile; el intercambio de los restos hidroxilo Ser y Thr, el intercambio de los restos ácidos Asp y Glu, la sustitución entre los restos amida Asn y Gln, el intercambio de los restos básicos Lys y Arg y sustituciones entre los restos aromáticos Phe y Tyr. Las directrices con respecto a qué cambios de aminoácidos probablemente son fenotípicamente imperceptibles se encuentran en Bowie y col., *Science* 247:1306-1310 (1990).

TABLA 1

Sustituciones conservadoras de aminoácidos

| | |
|------------|---|
| Aromáticos | Fenilalanina Triptófano Tirosina |
| Hidrófobos | Leucina Isoleucina Valina |
| Polares | Glutamina Asparagina |
| Básicos | Arginina Lisina Histidina |
| Ácidos | Ácido aspártico Ácido glutámico |
| Pequeños | Alanina Serina Treonina Metionina Glicina |

La comparación de secuencias y la determinación del porcentaje de identidad y similitud entre dos secuencias pueden realizarse usando un algoritmo matemático (*Computational Molecular Biology*, Lesk, A.M., editor Oxford University Press, Nueva York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., editor, Academic Press, Nueva York, 1993; *Computer Analysis of Sequence Data, Parte I*, Griffin, A.M., y Griffin, H.G., editores, Humana Press, Nueva Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; y *Sequence Analysis Primer*, Gribskov, M. y Devereux, J., editores, M Stockton Press, Nueva York, 1991). En una realización preferida, el porcentaje de identidad entre dos secuencias de aminoácidos se determina usando el algoritmo de Needleman y Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) que se ha incorporado en el programa GAP en el paquete de programas GCG (disponible en <http://www.gcg.com>), usando una matriz Blossom 62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 ó 4 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. Aún en otra realización preferida, el porcentaje de identidad entre dos secuencias de nucleótidos se determina usando el programa GAP en el paquete de programas GCG (Devereux, J., y col., *Nucleic Acids Res.* 12 (1):387 (1984)) (disponible en <http://www.gcg.com>), usando una matriz NWSgapdna.CMP y un peso de hueco de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. En otra realización, el porcentaje de identidad entre dos secuencias de aminoácidos o nucleotídicas se determina usando el algoritmo de E. Meyers y W. Miller (CABIOS, 4:11-17 (1989)) que se ha incorporado al programa ALIGN (versión 2.0), usando una tabla de pesos de restos PAM120, una penalización por longitud de hueco de 12 y una penalización por hueco de 4.

Las secuencias de ácido nucleico y de proteína de la presente invención puede usarse además como una “secuencia problema” para realizar una búsqueda en bases de datos públicas, por ejemplo, para identificar otros miembros de la familia o secuencias relacionadas. Estas búsquedas pueden realizarse usando los programas NBLAST y XBLAST (versión 2.0) de Altschul, y col. (*J. Mol. Biol.* 215:403-10 (1990)). Las búsquedas de nucleótidos de BLAST pueden realizarse con el programa NBLAST, valor = 100, longitud de palabra = 12 para obtener secuencias de nucleótidos homólogas a las moléculas de ácido nucleico de la invención. Las búsquedas de proteínas de BLAST pueden realizarse con el programa XBLAST, valor = 50, longitud de palabra = 3 para obtener secuencias de aminoácidos homólogas a las proteínas de la invención. Para obtener alineamientos con huecos con fines de comparación, se puede utilizar Gapped BLAST como se describe en Altschul y col. (*Nucleic Acids Res.* 25 (17):3389-3402 (1997)). Cuando se utilizan los programas BLAST y Gapped BLAST, se pueden usar los parámetros por defecto de los respectivos programas (por ejemplo, XBLAST y NBLAST). Véase <http://www.ncbi.nlm.nih.gov>.

Una variante polipeptídica puede diferir en la secuencia de aminoácidos en una o más sustituciones, deleciones, inserciones, inversiones, fusiones y truncamientos o una combinación de cualquiera de éstos.

Las variantes polipeptídicas pueden ser completamente funcionales o pueden carecer de función en una o más actividades. De este modo, en el caso presente, las variaciones pueden afectar a la función, por ejemplo, de una o más de las regiones correspondientes a la unión al ligando, asociación a la membrana, unión a proteína G y transducción de señales.

Generalmente, las variantes completamente funcionales contienen sólo una variación conservadora o una variación en restos no críticos o en regiones no críticas. Las variantes funcionales también pueden contener una sustitución de aminoácidos similares que no produce ningún cambio o un cambio insignificante en la función. Alternativamente, estas sustituciones pueden afectar positiva o negativamente en cierto grado a la función.

Generalmente las variantes no funcionales contienen una o más sustituciones, deleciones, inserciones, inversiones o truncamientos no conservadores de aminoácidos o una sustitución, inserción, inversión o deleción en un resto crítico o en una región crítica.

Como se indica, las variantes pueden ser naturales o pueden estar hechas por medios de recombinación o síntesis química para proporcionar características útiles y nuevas al polipéptido receptor. Esto incluye prevenir la inmunogenicidad de formulaciones farmacéuticas mediante la prevención de la agregación proteica.

Las variaciones útiles además incluyen alteraciones de las características de unión al ligando. Por ejemplo, una realización implica una variación en el sitio de unión que da lugar a la unión pero no a la liberación, o a una liberación más lenta, del ligando. Una variación útil adicional en los mismos sitios puede dar lugar a una mayor afinidad por el ligando. Las variaciones útiles también incluyen cambios que proporcionan afinidad por otro ligando. Otra variación útil incluye aquella que permita la unión, pero previene la activación mediada por el ligando. Otra variación útil incluye la variación en el dominio transmembranal de unión a proteína G y de transducción de señales que proporciona una disminución o aumento de la unión mediante la proteína G apropiada o unión mediante una proteína G diferente a la que normalmente se asocia el receptor. Otra variación útil proporciona una proteína de fusión en la que uno o más dominios o subregiones se fusionan de forma operativa a uno o más dominios o subregiones de otro receptor acoplado a proteína G.

Los aminoácidos que son esenciales para la función pueden identificarse mediante procedimientos conocidos en la técnica, tales como mutagénesis dirigida a sitio o detección sistemática de alaninas por mutagénesis (Cunningham y col., *Science* 244:1081-1085 (1989)). Este último procedimiento introduce mutaciones únicas de alanina en cada resto de la molécula. A continuación se analiza la actividad biológica de las moléculas mutantes resultantes, tal como la unión al receptor o *in vitro* o la actividad proliferativa *in vitro*. Los sitios que son críticos para la unión ligando-receptor pueden determinarse también mediante análisis estructural, tal como cristalización, resonancia magnética nuclear o marcaje por fotoafinidad (Smith y col., *J Mol. Biol.* 224:899-904 (1992); de Vos y col. *Science* 255:306-312 (1992)).

ES 2 275 358 T3

Puede existir homología sustancial con la secuencia de ácido nucleico o de aminoácidos completa o con fragmentos de estas secuencias.

De este modo, la invención también incluye fragmentos de polipéptidos de la proteína receptora 15625. Los fragmentos pueden derivar de la secuencia de aminoácidos mostrada en la ID SEC N° 1. Sin embargo, la invención también abarca fragmentos de las variantes de la proteína receptora 15625 como se describe en este documento.

Sin embargo, no debe interpretarse que los fragmentos a los que se refiere la invención abarcan a fragmentos que puedan haberse descrito antes de la presente invención.

Como se usa en este documento, un fragmento comprende al menos 10 aminoácidos contiguos desde el aminoácido 1 al aminoácido 280 y desde el aminoácido 291 al aminoácido 342. Los fragmentos retienen una o más de las actividades biológicas de la proteína, por ejemplo, la capacidad para unirse a una proteína G o a ligando, así como fragmentos que pueden usarse como un inmunógeno para generar anticuerpo frente al receptor.

Los fragmentos biológicamente activos (péptidos que tienen una longitud de, por ejemplo, 10, 12, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 o más aminoácidos) pueden comprender un dominio o motivo, por ejemplo, un dominio o bucle extracelular o intracelular, uno o más segmentos transmembranales o partes de los mismos, sitios de unión a proteína G o un motivo característico de GPCR, sitios de glucosilación, sitios de fosforilación de la proteína cinasa dependiente de AMPc y GMPc y de la proteína cinasa C y sitios de miristilación.

Los posibles fragmentos incluyen, pero sin limitaciones: 1) péptidos solubles que comprenden el dominio extracelular amino terminal completo, desde aproximadamente el aminoácido 1 a aproximadamente el aminoácido 25 de la ID SEC N° 1 o partes del mismo; 2) péptidos que comprenden el dominio intracelular carboxilo terminal completo desde aproximadamente el aminoácido 303 al aminoácido 342 de la ID SEC N° 1 o partes del mismo; 3) péptidos que comprenden la región que abarca el dominio transmembranal completo desde aproximadamente el aminoácido 26 al aminoácido 302; 4) cualquiera de los segmentos transmembranales específicos o partes de los mismos; 5) cualquiera de los tres bucles intracelulares o tres extracelulares o partes de los mismos.

Los fragmentos pueden retener una o más de las actividades biológicas de la proteína, por ejemplo la capacidad para unirse a una proteína G o al ligando. Los fragmentos también pueden ser útiles como inmunógenos para generar anticuerpos frente al receptor.

Los fragmentos biológicamente activos pueden comprender un dominio o motivo, por ejemplo, un dominio o bucle extracelular o intracelular, uno o más segmentos transmembranales o partes de los mismos, un sitio de unión a proteína G o un motivo característico de GPCR, sitios de glucosilación, sitios de fosforilación de proteína cinasa dependiente de AMPc y GMPc, sitios de fosforilación de proteína cinasa C y sitios de N-miristilación. Estos péptidos pueden tener una longitud de, por ejemplo, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 o más aminoácidos.

Los posibles fragmentos incluyen, pero sin limitaciones: 1) péptidos solubles que comprenden el dominio extracelular amino terminal completo de aproximadamente el aminoácido 1 a aproximadamente el aminoácido 25 de la ID SEC N° 1 o partes del mismo; 2) péptidos que comprenden el dominio intracelular carboxilo terminal completo de aproximadamente el aminoácido 303 al aminoácido 342 de la ID SEC N° 1 o partes del mismo; 3) péptidos que comprenden la región que abarca el dominio transmembranal completo de aproximadamente el aminoácido 26 a aproximadamente el aminoácido 302 o partes de misma; 4) cualquiera de los segmentos transmembranales específicos o partes de los mismos, de aproximadamente el aminoácido 26 a aproximadamente el aminoácido 47, de aproximadamente el aminoácido 59 a aproximadamente el aminoácido 79, de aproximadamente el aminoácido 99 a aproximadamente el aminoácido 120, de aproximadamente el aminoácido 143 a aproximadamente el aminoácido 162, de aproximadamente el aminoácido 189 a aproximadamente el aminoácido 212, de aproximadamente el aminoácido 238 a aproximadamente el aminoácido 255 y de aproximadamente el aminoácido 284 a aproximadamente el aminoácido 302; 5) cualquiera de los tres bucles intracelulares o los tres extracelulares o partes de los mismos, de aproximadamente el aminoácido 48 a aproximadamente el aminoácido 58, de aproximadamente el aminoácido 121 a aproximadamente el aminoácido 142, de aproximadamente el aminoácido 213 a aproximadamente el aminoácido 237, de aproximadamente el aminoácido 80 a aproximadamente el aminoácido 98, de aproximadamente el aminoácido 163 a aproximadamente el aminoácido 188 y de aproximadamente el aminoácido 256 a aproximadamente el aminoácido 283. Los fragmentos además incluyen combinaciones de los fragmentos anteriores, tales como un dominio amino terminal combinado con uno o más segmentos transmembranales y los bucles extra o intracelulares relacionados o uno o más segmentos transmembranales y los bucles intra o extracelulares relacionados más el dominio carboxilo terminal. De este modo, pueden combinarse cualquiera de los fragmentos anteriores. Otros fragmentos incluyen la proteína madura de aproximadamente el aminoácido 6 al 342. Otros fragmentos contienen los diversos sitios funcionales descritos en este documento, tales como de glucosilación, sitios de fosforilación de la proteína cinasa dependiente de AMPc y GMPc, sitios de fosforilación de la proteína cinasa C, sitios de N-miristilación y una secuencia que contiene la secuencia del motivo característico de GPCR. Los fragmentos, por ejemplo, pueden extenderse en uno o ambos sentidos desde el sitio funcional hasta abarcar 5, 10, 15, 20, 30, 40, 50 o hasta 100 aminoácidos. Además, los fragmentos pueden incluir subfragmentos de los dominios específicos mencionados anteriormente, cuyos subfragmentos retienen la función del dominio del cual derivan.

ES 2 275 358 T3

Los fragmentos también pueden incluir fragmentos antigénicos y específicamente aquellos de la Figura 1 que se ha demostrado tener un alto índice antigénico.

Por consiguiente, los posibles fragmentos incluyen fragmentos que definen un sitio de unión a ligando, fragmentos que definen la asociación a la membrana, fragmentos que definen la interacción con las proteínas G y la transducción de la señal y fragmentos que definen los sitios de glucosilación, los sitios de fosforilación de la proteína cinasa dependiente de AMPc y GMPc, los sitios de fosforilación de la proteína cinasa C y los sitios de N-miristilación. Mediante esto se pretende identificar un fragmento discreto que proporcione la función relevante o que permita la función relevante. En una realización preferida, el fragmento contiene el sitio de unión al ligando.

La invención también proporciona fragmentos con propiedades inmunogénicas. Estos contienen una porción portadora de un epítipo de la proteína receptora 15625 y sus variantes. Estos péptidos portadores de epítipo son útiles para obtener anticuerpos que se unan específicamente al polipéptido, región o fragmento del receptor. Estos péptidos pueden contener al menos 10, 12, al menos 14 ó entre al menos aproximadamente 15 y aproximadamente 30 aminoácidos.

Ejemplos no limitantes de polipéptidos antigénicos que pueden usarse para generar anticuerpos incluyen péptidos derivados del dominio extracelular amino terminal o cualquiera de los bucles extracelulares. En la Figura 1 se muestran las regiones que tienen un alto índice de antigenicidad. Sin embargo, también se abarcan los anticuerpos intracelulares (“intracuerpos”) que podrían reconocer regiones peptídicas intracelulares del receptor.

Los polipéptidos del receptor (incluyendo variantes y fragmentos que pueden haberse descrito antes de la presente invención) son útiles para ensayos biológicos relacionados con los GPCR. Estos ensayos implican cualquiera de las funciones, actividades o propiedades conocidas de los GPCR útiles para el diagnóstico y tratamiento de afecciones relacionadas con GPCR.

El receptor portador de epítopos y los polipéptidos pueden obtenerse por cualquiera de los medios convencionales (Houghten, R. A., *Proc. Natl. Acad. Sci. USA* 82: 5131-5135 (1985)). En la Patente de EE.UU. N° 4.631.211 se describe la síntesis simultánea de péptidos múltiples.

Los fragmentos pueden ser discretos (sin fusionar con otros aminoácidos o polipéptidos) o pueden estar dentro de un polipéptido mayor. Además, un único polipéptido más grande puede comprender varios fragmentos. En una realización, un fragmento diseñado para la expresión en un huésped puede tener regiones heterólogas del pre y pro-polipéptido fusionadas al extremo amino terminal del fragmento del receptor y una región adicional fusionada al extremo carboxilo terminal del fragmento.

De este modo, la invención proporciona proteínas quiméricas o de fusión. Éstas comprenden una proteína receptora unida de forma operativa a una proteína heteróloga que tiene una secuencia de aminoácidos no sustancialmente homóloga a la proteína receptora. “Unido de forma operativa” indica que la proteína receptora y la proteína heteróloga se fusionan en fase. La proteína heteróloga puede fusionarse al extremo N-terminal o C-terminal de la proteína receptora.

En una realización, la proteína de fusión no afecta a la función del receptor *per se*. Por ejemplo, la proteína de fusión puede ser una proteína de fusión con GST en la que las secuencias del receptor se fusionan con el extremo C-terminal de las secuencias de GST. Otros tipos de proteínas de fusión incluyen, pero sin limitaciones, proteínas de fusión enzimáticas, por ejemplo, fusiones de beta-galactosidasa, fusiones GAL de doble híbrido en levadura, fusiones de poli-His y fusiones de Ig. Estas proteínas de fusión, especialmente las fusiones de poli-His, pueden facilitar la purificación de la proteína receptora recombinante. En ciertas células huésped (por ejemplo, células huésped de mamífero), la expresión y/o secreción de una proteína puede aumentarse usando una secuencia señal heteróloga. Por tanto, en otra realización, la proteína de fusión contiene una secuencia señal heteróloga en su extremo N-terminal.

El documento EP-A-O 464 533 describe las proteínas de fusión que comprenden diversas porciones de regiones constantes de inmunoglobulina. El Fc es útil en terapia y diagnóstico y, de este modo, da lugar por ejemplo a la mejora de las propiedades farmacocinéticas (documento EP-A 0232 262). En el descubrimiento de fármacos, por ejemplo, las proteínas humanas se han fusionado con las porciones Fc para ensayos de selección de alto rendimiento para identificar antagonistas. Bennett y col. (*J. Mol. Recog.* 8:52-58 (1995)) y Johanson y col. (*J. Biol. Chem.* 270. 16:9459-9471 (1995)). De este modo, esta invención también abarca proteínas de fusión solubles que contienen un polipéptido del receptor y diversas porciones de las regiones constantes de las cadenas pesadas o ligeras de las inmunoglobulinas de diversas subclases (IgG, IgM, IgA, IgE). Se prefiere como inmunoglobulina la parte constante de la cadena pesada de la IgG humana, en particular la IgG1, donde la fusión tiene lugar en la región bisagra. Para algunos usos es deseable eliminar el Fc después de que la proteína de fusión se haya usado para el propósito al que está destinada, por ejemplo, cuando la proteína de fusión se usa como antígeno para inmunizaciones. En una realización en particular, la parte Fc puede eliminarse de una forma sencilla mediante una secuencia de escisión que también se incorpora y que puede ser escindida con el factor Xa.

Una proteína quimérica o de fusión puede producirse mediante técnicas de recombinación de ADN convencionales. Por ejemplo, los fragmentos de ADN que codifican las diferentes secuencias proteicas se unen en fase de acuerdo con técnicas convencionales. En otra realización, el gen de fusión puede sintetizarse mediante técnicas convencionales que incluyen sintetizadores de ADN automáticos. Alternativamente, la amplificación por PCR de los fragmentos

génicos puede realizarse usando cebadores anclados que dan lugar a salientes complementarios entre dos fragmentos génicos consecutivos que pueden, posteriormente, hibridarse y reamplificarse para generar secuencias génicas quiméricas (véase Ausubel y col., *Current Protocols in Molecular Biology*, 1992). Además, hay disponibles en el mercado muchos vectores de expresión que ya codifican un resto de fusión (por ejemplo, una proteína GST). Puede clonarse un ácido nucleico que codifique la proteína receptora en dicho vector de expresión, de modo que el resto de fusión se une en fase a la proteína receptora.

Otra forma de proteína de fusión es aquella que afecta directamente a las funciones del receptor. Por consiguiente, la presente invención abarca un polipéptido del receptor en el que uno o más de los dominios del receptor (o partes de los mismos) se han sustituido por dominios homólogos (o partes de los mismos) de otro receptor acoplado a proteína G o de otro tipo de receptor. Por consiguiente, son posibles diversas permutaciones. El dominio extracelular amino terminal, o una subregión del mismo, (por ejemplo, de unión al ligando) puede sustituirse por el dominio o subregión de otra proteína receptora de unión al ligando. Alternativamente, puede sustituirse el dominio transmembranal completo, o cualquiera de los siete segmentos o bucles, o partes de los mismos, por ejemplo, de unión a proteína G/transducción de la señal. Finalmente, puede sustituirse el dominio intracelular carboxilo terminal o una subregión. De este modo, pueden formarse receptores quiméricos en los que se han sustituido uno o más de los dominios o subregiones nativos.

La proteína receptora aislada puede purificarse a partir de células que la expresan de forma natural, como se describe en este documento, tales como de células de cerebro, y especialmente células de la glía, células CD34⁺ y líneas tales como HEK 293 y Jurkat, purificarse a partir de células que se han alterado para que la expresen (recombinante) o sintetizarse usando procedimiento de síntesis de proteína conocidos.

En una realización, la proteína se produce por técnicas de ADN recombinante. Por ejemplo, se clona una molécula de ácido nucleico que codifica el polipéptido del receptor en un vector de expresión, se introduce el vector de expresión en una célula huésped y se expresa la proteína en la célula huésped. La proteína puede aislarse a continuación a partir de las células por un esquema de purificación apropiado usando técnicas convencionales de purificación de proteínas.

Los polipéptidos a menudo contienen aminoácidos distintos a los 20 aminoácidos frecuentemente denominados como los 20 aminoácidos naturales. Además, muchos aminoácidos, incluyendo los aminoácidos terminales, pueden modificarse mediante procedimientos naturales, tales como el procesamiento y otras modificaciones postraduccionales, o mediante técnicas de modificación químicas bien conocidas en la materia. Las modificaciones comunes que se producen de forma natural en los polipéptidos se describen en textos básicos, monografías detalladas y en la bibliografía científica y son bien conocidas para los expertos en la materia.

Por consiguiente, los polipéptidos también abarcan derivados o análogos en los que un resto aminoacídico sustituido no es uno codificado por el código genético, en los que se incluye un grupo sustituyente, en los que el polipéptido maduro se fusiona con otro compuesto, como un compuesto para aumentar la semivida del polipéptido (por ejemplo, polietilenglicol) o en los que los aminoácidos adicionales se fusionan con el polipéptido maduro, tal como una secuencia líder o secretora o una secuencia para la purificación del polipéptido maduro o una secuencia de pro-proteína.

Las modificaciones conocidas incluyen, pero sin limitaciones, acetilación, acilación, ADP-ribosilación, amidación, unión covalente de flavina, unión covalente de un grupo hemo, unión covalente de un nucleótido o derivado de nucleótido, unión covalente de un lípido o derivado lipídico, unión covalente de fosfatidilinositol, entrecruzamiento, ciclación, formación de puentes disulfuro, desmetilación, formación de entrecruzamientos covalentes, formación de cistina, formación de piroglutamato, formilación, gamma-carboxilación, glucosilación, formación de anclajes GPI, hidroxilación, yodinación, metilación, miristilación, oxidación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenoilación, sulfatación, adición mediada por ARN de transferencia de aminoácidos a proteínas, tales como arginilación y ubiquitinación.

Estas modificaciones son bien conocidas por los expertos en la materia y se han descrito con gran detalle en la bibliografía científica. Varias modificaciones especialmente frecuentes, glucosilación, unión de lípidos, sulfatación, gamma-carboxilación de restos de ácido glutámico, hidroxilación y ADP-ribosilación, por ejemplo, se describen en la mayoría de los textos básicos, tales como *Proteins-Structure and Molecular Properties*, 2^a edición, T. E. Creighton, W. H. Freeman y Compañía, Nueva York (1993). Sobre este tema se dispone de muchas revisiones detalladas, tales como por Wold, F., *Posttranslational Covalent Modification of Proteins*, B. C. Johnson, editor, Academic Press, Nueva York 1-12 (1983); Seifter y col. (*Meth. Enzymol.* 182:626-646 (1990)) y Rattan y col. (*Ann. N.Y. Acad. Sci.* 663: 48-62 (1992)).

Como es bien sabido, los polipéptidos no siempre son completamente lineales. Por ejemplo, los polipéptidos pueden estar ramificados como resultado de la ubiquitinación, y pueden ser circulares, con o sin ramificación, generalmente como resultado de los sucesos postraduccionales, que incluyen sucesos de procesamiento natural y sucesos realizados por la manipulación humana que no ocurren de forma natural. Se pueden sintetizar polipéptidos circulares, ramificados y circulares ramificados mediante procesos naturales no traduccionales y mediante procedimientos sintéticos.

Las modificaciones pueden tener lugar en cualquier parte de un polipéptido, incluyendo el esqueleto peptídico, las cadenas laterales de los aminoácidos y los extremos amino o carboxilo terminales. El bloqueo de los grupos amino o carboxilo o de ambos en un polipéptido, mediante una modificación covalente, es frecuente en polipéptidos naturales y

5 sintéticos. Por ejemplo, el resto amino terminal de polipéptidos hechos en *E. coli*, antes del procesamiento proteolítico, será casi invariablemente N-formilmetionina.

Las modificaciones pueden ser una función de cómo se hace la proteína. Para polipéptidos recombinantes, por ejemplo, las modificaciones se determinarán mediante la capacidad de modificación postraduccional de la célula huésped y por las señales de modificación en la secuencia de aminoácidos del polipéptido. Por consiguiente, cuando se desea la glucosilación, debe expresarse el polipéptido en un huésped capaz de glicosilar, generalmente una célula eucariota. Las células de insecto a menudo realizan las mismas glucosilaciones postraduccionales que las células de mamíferos y, por esta razón, se han desarrollado sistemas de expresión de células de insecto para expresar eficazmente proteínas de mamífero con un patrón nativo de glucosilación. Consideraciones similares se aplican a otras modificaciones.

El mismo tipo de modificación puede estar presente en el mismo grado o en grado variable en varios sitios de un polipéptido determinado. Además, un polipéptido determinado puede contener más de un tipo de modificación.

La descripción anterior generalmente también se aplica a la secuencia del polipéptido mostrada en la ID SEC N° 3. Los dominios y sitios funcionales predichos se pueden identificar fácilmente mediante programas informáticos bien conocidos y fácilmente disponibles para los expertos en la materia (por ejemplo, análisis PROSITE).

20 *Usos de los polipéptidos*

Los polipéptidos del receptor son útiles para producir anticuerpos específicos para la proteína, regiones o fragmentos del receptor 15625. En la Figura 1 se muestran regiones que tienen un valor del índice de antigenicidad alto.

Los polipéptidos del receptor (incluyendo variantes y fragmentos que pueden haberse descrito antes de la presente invención) son útiles para ensayos biológicos relacionados con los GPCR. Estos ensayos implican cualquiera de las funciones o actividades de GPCR conocidas o propiedades útiles para el diagnóstico y tratamiento de afecciones relacionadas con GPCR.

Los polipéptidos del receptor también son útiles en los ensayos de selección de fármacos, en sistemas basados en células y sin células. Los sistemas basados en células pueden ser nativos, es decir, células que normalmente expresan la proteína receptora, como de una biopsia o expandidas en cultivo celular. En una realización, sin embargo, los ensayos basados en células implican células huésped recombinantes que expresan la proteína receptora.

Los polipéptidos pueden usarse para identificar compuestos que modulen la actividad del receptor. Tanto la proteína 15625 como las variantes y fragmentos apropiados pueden usarse en selecciones de alto rendimiento para analizar la capacidad de unión al receptor de compuestos candidatos. Estos compuestos pueden seleccionarse, además, frente a un receptor funcional para determinar el efecto del compuesto sobre la actividad del receptor. Pueden identificarse los compuestos que activen (agonista) o inactiven (antagonista) el receptor en un grado deseado.

Los polipéptidos del receptor pueden usarse para seleccionar compuestos con capacidad para estimular o inhibir la interacción entre la proteína receptora y una molécula diana que normalmente interacciona con la proteína receptora. La diana puede ser el ligando o un componente de la ruta de señalización con el que interaccione normalmente la proteína receptora (por ejemplo, una proteína G u otros componentes que interactúan implicados en el recambio de AMPc o de fosfatidilinositol y/o en la activación de la adenilato ciclasa o de la fosfolipasa C). El ensayo incluye las etapas de combinar la proteína receptora con un compuesto candidato en condiciones que permitan a la proteína receptora o un fragmento interactuar con la molécula diana y detectar la formación de un complejo entre la proteína y la diana o detectar la consecuencia bioquímica de la interacción de la proteína receptora y la diana, tal como cualquiera de los efectos asociados a la transducción de la señal, tal como la fosforilación de la proteína G, el recambio de AMPc o de fosfatidilinositol y la activación de la adenilato ciclasa o de la fosfolipasa C.

Los compuestos candidatos incluyen, por ejemplo, 1) péptidos, tales como péptidos solubles, que incluyen péptidos de fusión con cola de Ig y miembros de bibliotecas aleatorias de péptidos (véase, por ejemplo, Lam y col., *Nature* 354:82-84 (1991); Houghten y col., *Nature* 354:84-86 (1991)) y bibliotecas moleculares derivadas mediante química combinatoria hechas de aminoácidos de configuración D y/o L; 2) fosfopéptidos (por ejemplo, miembros de bibliotecas de fosfopéptidos dirigidos aleatorios o parcialmente degenerados, véase, por ejemplo, Songyang y col., *Cell* 72:767-778 (1993)); 3) anticuerpos (por ejemplo, anticuerpos policlonales, monoclonales, humanizados, antiidiotípicos, quiméricos y de cadena sencilla, así como Fab, F(ab')₂, fragmentos Fab de bibliotecas de expresión y fragmentos de anticuerpos de unión a epítope) y 4) moléculas orgánicas e inorgánicas pequeñas (por ejemplo, moléculas obtenidas de bibliotecas combinatorias y de productos naturales).

Un compuesto candidato es un receptor de longitud completa o fragmento soluble que compite por la unión al ligando. Otros compuestos candidatos incluyen receptores mutantes o fragmentos apropiados que contienen mutaciones que afectan a la función del receptor y, de este modo, compite por el ligando. Por consiguiente, la invención abarca un fragmento que compite por el ligando, por ejemplo con una afinidad mayor, o un fragmento que se une al ligando, pero no permite que se libere.

ES 2 275 358 T3

La invención proporciona otros puntos finales para identificar compuestos que modulen (estimulen o inhiban) la actividad del receptor. Generalmente, los ensayos implican un ensayo de sucesos en la ruta de transcripción de la señal que indica la actividad del receptor. De este modo, puede analizarse la expresión de genes que están regulados por incremento o por disminución en respuesta a la cascada de señalización dependiente de la proteína receptora. En una realización, la región reguladora de estos genes se puede unir de forma operativa a un marcador que sea fácilmente detectable, como la luciferasa. Alternativamente, también podría medirse la fosforilación de la proteína receptora, o una diana de la proteína receptora.

Cualquiera de las funciones biológicas o bioquímicas mediadas por el receptor puede usarse como ensayo de punto final. Estas incluyen todos los sucesos bioquímicos o bioquímicos/biológicos descritos en este documento, en las referencias citadas en este documento.

También pueden seleccionarse compuestos de unión y/o activación usando proteínas receptoras quiméricas en las que el dominio extracelular amino terminal, o partes del mismo, el dominio transmembranal completo o subregiones, tales como cada uno de los siete segmentos transmembranales o cualquiera de los bucles intracelulares o extracelulares y el dominio intracelular carboxilo terminal, o partes del mismo, pueden sustituirse por dominios o subregiones heterólogos. Por ejemplo, puede usarse una región de unión a proteína G que interaccione con una proteína G diferente a la que es reconocida por el receptor nativo. Por consiguiente, se dispone de un grupo diferente de componentes de transducción de la señal como ensayo de punto final para la activación. Alternativamente, puede sustituirse la porción transmembranal completa o subregiones (tales como segmentos transmembranales o bucles intracelulares o extracelulares) por la porción transmembranal completa o subregiones específicas para una célula huésped que es diferente de la célula huésped de la que deriva el dominio extracelular amino terminal y/o la región de unión a proteína G. Esto permite realizar los ensayos en otra célula diferente a la célula huésped específica de la cual deriva el receptor. Alternativamente, podría sustituirse el dominio extracelular amino terminal (y/u otra región de unión a ligando) por un dominio (y/u otra región de unión) que una un ligando diferente proporcionando, de este modo, un ensayo para ensayar compuestos que interaccionen con el dominio (o región) extracelular amino terminal heterólogo, pero que siga induciendo la transducción de la señal. Finalmente, la activación puede detectarse mediante un gen indicador que contenga una región codificante fácilmente detectable unido de forma operativa a una secuencia reguladora de la transcripción que es parte de la ruta nativa de transducción de la señal.

Los polipéptidos del receptor también son útiles en ensayos competitivos de unión en procedimientos diseñados para descubrir compuestos que interaccionen con el receptor. De este modo, se expone un compuesto a un polipéptido del receptor en condiciones que permitan unirse al compuesto o, por el contrario, interaccionar con el polipéptido. También se añaden a la mezcla polipéptidos del receptor solubles. Si el compuesto de ensayo interacciona con el polipéptido soluble del receptor, disminuye la cantidad de complejo formado o la actividad de la diana del receptor. Este tipo de ensayo es especialmente útil en casos en los que se buscan compuestos que interaccionen con regiones específicas del receptor. De este modo, el polipéptido soluble que compite con la región diana del receptor se diseña para que contenga secuencias peptídicas correspondientes a la región de interés.

Para realizar ensayos sin células para la selección de fármacos, es deseable inmovilizar la proteína receptora, o un fragmento, o sus moléculas diana para facilitar la separación de los complejos de las formas no acoplejadas de una o ambas proteínas, así como adaptar la automatización del ensayo.

En los ensayos de selección de fármacos pueden usarse técnicas para inmovilizar proteínas sobre matrices. En una realización, puede proporcionarse una proteína de fusión que añada un dominio que permita a la proteína unirse a una matriz. Por ejemplo, pueden adsorberse proteínas de fusión glutatión-S-transferasa/15625 a perlas de glutatión-sefara (Sigma Chemical, San Luis, MO) o placas de microvaloración derivatizadas con glutatión, que a continuación se mezclan con los lisados celulares (por ejemplo, marcados con ³⁵S) y el compuesto candidato, y la mezcla se incuba en condiciones que promueven la formación de complejos (por ejemplo, en condiciones fisiológicas de salinidad y pH). Tras la incubación, las perlas se lavan para eliminar cualquier marcaje no unido a la matriz inmovilizada y el radio-marcaje se determina directamente, o en el sobrenadante después de que se disocian los complejos. Alternativamente, los complejos pueden disociarse de la matriz, separarse por SDS-PAGE y a partir del gel, puede cuantificarse el nivel de proteína unida al receptor que se encuentra en la fracción de las perlas usando técnicas electroforéticas convencionales. Por ejemplo, el polipéptido o su molécula diana pueden inmovilizarse utilizando la conjugación de biotina y estreptavidina usando técnicas bien conocidas en la técnica. Alternativamente, los anticuerpos que reaccionan con la proteína, pero que no interfieren con la unión de la proteína a su molécula diana, pueden derivatizarse en los pocillos de la placa para atrapar la proteína en los pocillos mediante la conjugación con el anticuerpo. Las preparaciones de una proteína de unión al receptor y de un compuesto candidato se incuban en los pocillos de presentación de proteína al receptor y puede cuantificarse la cantidad de complejos atrapados en el pocillo. Los procedimientos para detectar estos complejos, además de los descritos anteriormente para los complejos GST inmovilizados, incluyen la inmunodetección de complejos usando anticuerpos reactivos con la molécula diana de la proteína receptora, o que son reactivas con la proteína receptora y compiten con la molécula diana, así como ensayos ligados a enzimas que depende de la detección de una actividad enzimática asociada con la molécula diana.

Pueden usarse moduladores de la actividad de la proteína receptora identificados según estos ensayos de selección de fármacos para tratar a un sujeto con un trastorno mediado por la ruta del receptor, tratando las células que expresan la proteína 15625, tal como en el cerebro, especialmente células de la glía, células CD34⁺ de la médula ósea, que incluyen, pero sin limitaciones, megacariocitos, linfocitos B en reposo y músculo esquelético. Otros tejidos incluyen

ganglios linfáticos, bazo, corazón, timo, hígado, amígdalas, colon, granulocitos, placenta, páncreas y eritroblastos. Estos procedimientos de tratamiento incluyen las etapas de administrar los moduladores de la actividad de la proteína en una composición farmacéutica como se describe en este documento, a un sujeto que necesita este tratamiento.

5 Los trastornos que implican al bazo incluyen, pero sin limitaciones, esplenomegalia, incluyendo esplenitis aguda inespecífica, esplenomegalia congestiva e infartos esplénicos; neoplasmas, anomalías congénitas y rotura. Los trastornos asociados con la esplenomegalia incluyen infecciones, tales como esplenitis inespecífica, mononucleosis infecciosa, tuberculosis, fiebre tifoidea, brucelosis, citomegalovirus, sífilis, malaria, histoplasmosis, toxoplasmosis, leishmaniasis visceral, tripanosomiasis, esquistosomiasis, leishmaniasis y equinococosis, estados congestivos relacionados con hipertensión parcial, tales como cirrosis hepática, trombosis de la vena porta o esplénica e insuficiencia cardíaca, trastornos linfohematogénicos, tales como la enfermedad de Hodgkin, linfomas/leucemia no Hodgkin, mieloma múltiple, trastornos mieloproliferativos, anemias hemolíticas y púrpura trombocitopénica; dolencias inmunológicas-inflamatorias, tales como artritis reumatoide y lupus eritematoso sistémico; enfermedades de almacenamiento, tales como la enfermedad de Gaucher, enfermedad de Nieman-Pick, y mucopolisacaridosis y otras dolencias, tales como amiloidosis, neoplasmas primarios y quistes y neoplasmas secundarias.

15 Los trastornos que implican al colon incluyen, pero sin limitaciones, anomalías congénitas, tales como atresia y estenosis, divertículo de Meckel, megacolon agangliónico congénito - enfermedad de Hirschsprung; enterocolitis, tales como diarrea y disentería, enterocolitis infecciosa, incluyendo gastroenteritis vírica, enterocolitis bacteriana, enterocolitis necrotizante, colitis asociada a antibióticos (colitis pseudomembranosa) y colitis colagénica y linfocítica, trastornos inflamatorios intestinales diversos, incluyendo parásitos y protozoos, síndrome de inmunodeficiencia adquirida, trasplante, daño intestinal inducido por fármacos, enterocolitis por radiación, colitis neutropénica (tiflitis) y colitis desviada, enfermedad intestinal inflamatoria idiopática, tal como la enfermedad de Crohn y colitis ulcerosa, tumores del colon, tales como pólipos no neoplásicos, adenomas, síndromes familiares, carcinogénesis colorrectal, carcinoma colorrectal y tumores carcinoides.

20 Los trastornos que implican al hígado incluyen, pero sin limitaciones, daño hepático, ictericia y colestasis, tales como formación de bilirrubina y ácidos biliares, insuficiencia hepática y cirrosis, tal como cirrosis, hipertensión portal, incluyendo ascitis, derivación portosistémica y esplenomegalia; trastornos infecciosos, tales como hepatitis vírica, incluyendo infecciones por virus de la hepatitis A-E e infección por otros virus de la hepatitis, síndromes clínico-patológicos, tales como la enfermedad del vehículo, infección asintomática, hepatitis vírica aguda, hepatitis vírica crónica y hepatitis fulminante; hepatitis autoinmune, enfermedad hepática inducida por fármacos y toxinas; tales como enfermedad hepática alcohólica, enfermedad hepática pediátrica y por errores congénitos del metabolismo, tales como hemocromatosis, enfermedad de Wilson, deficiencia de antitripsina-1 y hepatitis neonatal; enfermedad del tracto biliar intrahepático, tal como cirrosis biliar secundaria, cirrosis biliar primaria, colangitis esclerosante primaria y anomalías del árbol biliar; enfermedades circulatorias, tales como alteraciones del flujo sanguíneo dentro del hígado, incluyendo compromiso de la arteria hepática y obstrucción y trombosis de la vena porta, alteración de flujo sanguíneo a través del hígado, incluyendo congestión pasiva y necrosis centrilobular, y peliosis hepática, obstrucción de flujo de salida de la vena hepática, incluyendo trombosis de la vena hepática (síndrome de Budd-Chiari) y enfermedad venooclusiva; enfermedad hepática asociada con el embarazo, tal como preeclampsia y eclampsia, hígado graso agudo del embarazo, y colestasis intrahepática del embarazo; complicaciones hepáticas del trasplante de órgano o de médula ósea, tal como toxicidad por fármacos tras el trasplante de médula ósea, enfermedad injerto contra huésped y rechazo hepático, y daño no inmunológico a aloinjertos de hígado; dolencias tumorales y tumorosas, tales como hiperplasias nodulares, adenomas y tumores malignos, incluyendo carcinoma primario de hígado y tumores metastáticos.

30 Los trastornos que implican al cerebro incluyen, pero sin limitaciones, trastornos que implican a neuronas y trastornos que implican a la glía, tal como astrocitos, oligodendrocitos, células ependimales y microglía; edema cerebral, presión intracraneal elevada y hemiación, e hidrocefalia; malformaciones y enfermedades del desarrollo, tales como defectos del tubo neural, anomalías del cerebro frontal, anomalías de la fosa posterior y siringomielia e hidromielia; daño cerebral perinatal; enfermedades cerebrovasculares, tales como hipoxia, isquemia e infarto, incluyendo hipotensión, hipoperfusión y estados de bajo flujo (isquemia cerebral global e isquemia cerebral focal) infarto por obstrucción del aporte sanguíneo local, hemorragia intracraneal, incluyendo hemorragia intracerebral (intraparenquimal), hemorragia subaracnoidea y rotura de aneurismas congénitos y malformaciones vasculares, enfermedad cerebrovascular hipertensiva, incluyendo infartos de tipo lacunar, hemorragias por colapso, y encefalopatía hipertensiva; infecciones, tales como meningitis aguda, incluyendo meningitis piogénica (bacteriana) aguda y meningitis aséptica (vírica) aguda, infecciones supurativas focales agudas, incluyendo abscesos cerebrales, empiema subdural y abscesos extradurales, meningoencefalitis bacteriana crónica, incluyendo tuberculosis y micobacteriosis, neurosífilis y neuroborreliosis (enfermedad de Lyme), meningoencefalitis vírica, incluyendo encefalitis vírica transmitida por artrópodos (Arbo), virus *Herpes simplex* de tipo 1, virus *Herpes simplex* de tipo 2, virus de la varicela-zoster (*Herpes zoster*), citomegalovirus, poliomielitis, rabia y virus de la inmunodeficiencia humana de tipo 1, incluyendo meningoencefalitis por VIH-1 (encefalitis subaguda), mielopatía vacuolar, miopatía asociada al SIDA, neuropatía periférica y SIDA en niños, leucoencefalopatía multifocal progresiva, panencefalitis esclerosante subaguda, meningoencefalitis fúngica, otras enfermedades infecciosas del sistema nervioso; encefalopatías espongiiformes transmisibles (enfermedades de priones); enfermedades desmielinizantes, incluyendo esclerosis múltiple, variantes de la esclerosis múltiple, encefalomiелitis diseminada aguda y encefalomiелitis hemorrágica necrosante aguda, y otras enfermedades con desmielinización; enfermedades degenerativas, tal como enfermedades degenerativas que afectan a la corteza cerebral, incluyendo la enfermedad de Alzheimer y la enfermedad de Pick, enfermedades degenerativas de los ganglios basales y del tronco cerebral, incluyendo parkinsonismo, enfermedad de Parkinson idiopática (parálisis agitans), parálisis supranuclear progresiva,

degeneración corticobasal, atrofia sistémica múltiple, incluyendo degeneración estriatonigral, síndrome de Shy-Dra-
 ger y atrofia olivopontocerebelar y enfermedad de Huntington; degeneraciones espinocerebelosas, incluyendo ataxias
 espinocerebelosas, incluyendo ataxia de Friedreich y telangiectasia-ataxia, enfermedades degenerativas que afectan a
 las neuronas motoras, incluyendo esclerosis lateral amiotrófica (enfermedad de neuronas motoras), atrofia bulboespinal
 5 (síndrome de Kennedy) y atrofia muscular espinal; errores congénitos del metabolismo, tal como leucodistrofias,
 incluyendo la enfermedad de Krabbe, leucodistrofia metacromática, adremoleucodistrofia, enfermedad de Pelizaeus-
 Merzbacher y enfermedad Canavan, encefalomiopatías mitocondriales, incluyendo enfermedad de Leigh y otras en-
 cefalomiopatías mitocondriales; enfermedades metabólicas adquiridas y tóxicas, incluyendo deficiencias de vitamina,
 tales como deficiencia de tiamina (vitamina B₁) y deficiencia de vitamina B₁₂; secuela neurológicas de alteracio-
 10 nes metabólicas, incluyendo hipoglucemia, hiperglucemia y encefalopatía hepática, enfermedades tóxicas, incluyendo
 monóxido de carbono, metanol, etanol y radiación, incluyendo metotrexato combinado y daño inducido por radia-
 ción; tumores, tales como gliomas, incluyendo astrocitoma, incluyendo astrocitoma fibrilar (difuso) y glioblastoma
 multiforme, astrocitoma policítico, xantastrocitoma pleomórfico y glioma del tronco cerebral, oligodendroglioma y
 ependimoma y lesiones de la masa paraventricular relacionadas, tumores neuronales, neoplasmas poco diferenciados,
 15 incluyendo meduloblastoma, otros tumores parenquimales, incluyendo linfoma cerebral primario, tumores de células
 germinales y tumores del parénquima pineal, meningiomas, tumores metastásicos, síndromes paraneoplásicos, tumo-
 res de vaina del nervio periférico, incluyendo schwannoma, neurofibroma y tumores de vaina del nervio periférico
 maligno (schwannoma maligno) y síndromes neurocutáneos (facomatosos), incluyendo neurofibromatosis, incluyen-
 20 do neurofibromatosis de tipo 1 (NF1) y neurofibromatosis de tipo 2 (NF2), esclerosis tuberosa y enfermedad de Von
 Hippel-Lindau.

Los trastornos que implican a células T incluyen, pero sin limitaciones, hipersensibilidad mediada por células,
 tales como hipersensibilidad de tipo retardado y citotoxicidad mediada por células T y rechazo de trasplantes; enfer-
 medades autoinmunes, tales como lupus eritematoso sistémico, síndrome de Sjögren, esclerosis sistémica, miopatías
 25 inflamatorias, enfermedad mixta del tejido conectivo, poliarteritis nodosa y otras vasculitis; síndromes de deficiencia
 inmunológica, incluyendo pero sin limitaciones, inmunodeficiencias primarias, tales como hipoplasia tímica, enfer-
 medades de inmunodeficiencia combinada grave, y SIDA; leucopenia; proliferaciones reactivas (inflamatorias) de
 leucocitos, incluyendo pero sin limitaciones, leucocitosis, linfadenitis inespecífica aguda y linfadenitis inespecífica
 crónica; proliferaciones neoplásicas de leucocitos, incluyendo pero sin limitaciones, neoplasmas linfoides, tales como
 30 neoplasmas de precursores de células T, tales como leucemia/linfoma linfoblástico, neoplasmas de periférico de célu-
 las T y linfocitos citolíticos naturales que incluyen linfoma periférico de célula T, leucemia/linfoma de células T del
 adulto inespecífico, micosis fungoides y síndrome de Sézary, y enfermedad de Hodgkin.

En médula ósea normal, las series mielocíticas (células polimorfonucleares) representan aproximadamente el 60%
 35 de los elementos celulares y las series eritrocíticas el 20-30%. Los linfocitos, monocitos, células reticulares, células
 plasmáticas y megacariocitos en conjunto constituyen el 10-20%. Los linfocitos constituyen del 5-15% de la mé-
 dula ósea del adulto normal. En la médula ósea, los tipos celulares están mezclados de modo que los precursores
 de eritrocitos (eritroblastos), macrófagos (monoblastos), plaquetas (megacariocitos), leucocitos polimorfonucleares
 (mieloblastos) y linfocitos (linfoblastos) de la sangre, pueden verse en el campo de un microscopio. Además, existen
 40 células troncales para las diferentes estirpes celulares, así como una célula troncal precursora para las células proge-
 nitoras comprometidas de los diferentes linajes. Los diversos tipos de células y etapas de cada una serán conocidos
 por el experto en la técnica y se encuentra, por ejemplo, en la página 42 (Figuras 2-8) de *Immunology, Immuno-
 pathology and Immunity*, Quinta Edición, Sell y col. Simon y Schuster (1996), describen sus explicaciones de los tipos
 celulares que se encuentran en la médula ósea. Por consiguiente, la invención se dirige a trastornos producidos por
 45 estas células. Estos trastornos incluyen, pero sin limitaciones, los siguientes: enfermedades que implican a las célu-
 las troncales hematopoyéticas, células linfoides progenitoras comprometidas, células linfoides que incluyen a células
 B y T, progenitores mieloides comprometidos, incluyendo monocitos, granulocitos y megacariocitos y progenitores
 eritroides comprometidos. Estas incluyen, pero sin limitaciones, las leucemias, incluyendo leucemias linfoides B,
 leucemias linfoides T, leucemias indiferenciadas, eritroleucemia, leucemia megacarioblástica, se incluyen leucemias
 50 monocíticas con o sin diferenciación; leucemia linfoblástica crónica y aguda, leucemia linfocítica crónica y aguda, leu-
 cemia mielogénica crónica y aguda, linfoma, síndrome mielodisplásico, leucemia mielóide crónica y aguda, leucemia
 mielomonocítica, leucemia mieloblástica crónica y aguda, leucemia mielogénica crónica y aguda, leucemia promie-
 locítica crónica y aguda, leucemia mielocítica crónica y aguda; enfermedades hematológicas de la estirpe monocito-
 macrófago, tales como leucemia mielogénica crónica juvenil, LMA secundaria, antecedentes de trastorno hematoló-
 55 gico, anemia refractaria, anemia aplásica, angioendoteliomatosis cutánea reactiva, trastornos fibrosos que implican la
 expresión alterada en células dendríticas, trastornos que incluyen esclerosis sistémica, síndrome E-M, síndrome de
 aceite tóxico epidémico, fascitis eosinofílica con formas localizadas de esclerodermia, colonopatía queloide y fibró-
 sa, histiocitoma fibroso maligno angiomaticoide; carcinoma, incluyendo carcinoma de células escamosas de cabella y
 cuello; sarcoma, incluyendo sarcoma de kaposi; fibroadenoma y tumores filoides, incluyendo fibroadenoma mama-
 60 ria; tumores del estroma; tumores filoides, incluyendo histiocitoma; eritroblastosis, neurofibromatosis; enfermedades
 del endotelio vascular; desmielinización, especialmente en lesiones viejas; gliosis, edema vasogénico, enfermedad
 vascular, enfermedad de Alzheimer y de Parkinson; linfomas de células T; linfomas de células B.

Los trastornos que implican al corazón, incluyen pero sin limitaciones, insuficiencia cardíaca, incluyendo pero sin
 65 limitación, hipertrofia cardíaca, insuficiencia cardíaca del lado izquierdo e insuficiencia cardíaca del lado derecho;
 enfermedad cardíaca isquémica, incluyendo pero sin limitaciones, angina de pecho, infarto de miocardio, enfermedad
 cardíaca isquémica crónica y muerte cardíaca súbita, enfermedad cardíaca hipertensiva, incluyendo pero sin limitacio-
 nes, la enfermedad cardíaca hipertensiva sistémico (del bucle izquierdo) y enfermedad cardíaca hipertensiva pulmonar

(del bucle derecho); enfermedad cardíaca valvular, incluyendo pero sin limitaciones, la degeneración valvular causada por calcificación, tal como estenosis aórtica por calcificación, calcificación de una válvula aórtica congénitamente bicúspide y calcificación anular mitral y degeneración mixomatosa de la válvula mitral (prolapso de la válvula mitral), fiebre reumática y enfermedad cardíaca reumática, endocarditis infectiva y vegetaciones no infectadas, tal como endocarditis trombotica no bacteriana y endocarditis de lupus eritematoso sistémico (enfermedad de Libman-Sacks), enfermedad cardíaca carcinóide y complicaciones de válvulas artificiales; enfermedades miocárdicas, incluyendo, pero sin limitaciones, cardiomiopatía dilatada, cardiomiopatía hipertrófica, cardiomiopatía restrictiva y miocarditis; enfermedad pericárdica, incluyendo pero sin limitaciones, efusión pericárdica y hemopericardio y pericarditis, incluyendo pericarditis aguda y pericarditis en curación, y enfermedad cardíaca reumatoide; enfermedad cardíaca neoplásica, incluyendo pero sin limitaciones, tumores cardíacos primarios, tales como mixoma, fibroelastoma papilar, rabiomioma y sarcoma, y efectos cardíacos de neoplasmas no cardíacos; enfermedad cardíaca congénita, incluyendo, pero sin limitación, cianosis tardía - derivación de izquierda a derecha, tal como defecto del septo atrial, defecto del septo ventricular, ducto arterioso patente y defecto del septo atrioventricular, cianosis temprana- derivación de derecha a izquierda, tal como tetralogía de Fallot, transposición de las arterias grandes, tronco arterioso, atresia tricúspide y conexión venosa pulmonar anómala, anomalías congénitas obstructivas, tales como coartación de la aorta, estenosis pulmonar y atresia y estenosis aórtica y atresia y trastornos que implican trasplante cardíaco.

Los trastornos que implican a los eritrocitos incluyen, pero sin limitaciones, anemias, tales como anemias hemolíticas, incluyendo esferocitosis hereditaria, enfermedad hemolítica debida a defectos en las enzimas de eritrocitos: deficiencia de la glucosa-6-fosfato deshidrogenasa, enfermedad de células falciformes, síndromes de talasemia, hemoglobinuria paroxismal nocturna, anemia inmunoheolítica y anemia hemolítica como resultado de un trauma de los eritrocitos y anemias de eritropoyesis disminuida, incluidas anemias megaloblásticas, tales como anemias de vitamina B12: anemia perniciosa y anemia por deficiencia de folato, anemia por deficiencia de hierro, anemia por enfermedad crónica, anemia aplásica, aplasia pura de eritrocitos y otras formas de insuficiencia de la médula ósea.

Los trastornos que implican al timo incluyen trastornos durante el desarrollo, tales como síndrome de DiGeorge con hipoplasia o aplasia tímica; quistes tímicos; hipoplasia tímica, que implica la aparición de folículos linfoides en el timo, creando hiperplasia folicular tímica; y timomas, que incluyen tumores de células germinales, linfomas, enfermedad de Hodgkin y carcinoides. Los timomas pueden incluir timomas benignos o encapsulados, y timoma maligno de tipo I (timoma invasivo) o de tipo II, denominado carcinoma tímico.

Los trastornos que implican a células B incluyen, pero sin limitaciones, neoplasmas de precursores de células B, tales como leucemia/linfoma linfoblástico. Los neoplasmas de células B periféricas incluyen, pero sin limitaciones, leucemia linfocítica crónica/linfoma linfocítico pequeño, linfoma folicular, linfoma difuso de células B grandes, linfoma de Burkitt, neoplasmas de células plasmáticas, mieloma múltiple y entidades relacionadas, linfoma linfoplasma-cítico (macroglubulinemia de Waldenstrom), linfoma de células del manto, linfoma de la zona marginal (MALToma) y leucemia de células peludas.

Los trastornos que implican al músculo esquelético incluyen tumores tales como rabiomioma.

Los trastornos que implican al páncreas incluyen los del páncreas exocrino, tal como anomalías congénitas, incluyendo pero sin limitaciones, páncreas ectópico; pancreatitis, incluyendo pero sin limitaciones, pancreatitis aguda; quistes, incluyendo pero sin limitaciones, pseudoquistes, tumores, incluyendo pero sin limitaciones, tumores quísticos y carcinoma del páncreas, y trastornos del páncreas endocrino tales como, diabetes mellitus, tumores de las células del islote, incluyendo pero sin limitaciones, insulomas, gastrinomas y otros tumores raros de las células del islote.

Los trastornos relacionados con un número de plaqueta reducido, trombocitopenia, incluyen púrpura trombocitopénica idiopática, incluyendo púrpura trombocitopénica idiopática aguda, trombocitopenia inducida por fármacos, trombocitopenia asociada al VIH y microangiopatías trombóticas: púrpura trombocitopénica trombótica y síndrome hemolítico urémico.

Los trastornos que implican a neoplasmas de precursores de células T incluyen leucemia/linfoma linfoblástica de precursores de T. Los trastornos que implican a neoplasmas periféricos de células T y linfocitos citolíticos naturales incluyen leucemia linfocítica crónica de células T, leucemia linfocítica granular grande, micosis fungoides y síndrome de Sézary, linfoma periférico de células T, linfoma de células T angioinmunoblástico inespecífico, linfoma angiocéntrico (linfoma de NK/células T^{4a}), linfoma intestinal de células T, leucemia/linfoma de células T maduras y linfoma de célula grande anaplásico.

El gen se expresa a niveles significativos en todos los progenitores de células sanguíneas analizados por los inventores. Se expresa de forma muy elevada en médula ósea (CD34⁺), en progenitores de la sangre periférica (que contiene progenitores circulantes derivados de la médula ósea) movilizados por G-CSF y de la sangre del cordón. Por consiguiente, la expresión del gen es importante para tratar trastornos asociados con la formación de células sanguíneas diferenciadas y/o maduras. A este respecto, los trastornos que son especialmente importantes incluyen anemia, neutropenia y trombocitopenia.

Los polipéptidos del receptor también son útiles para proporcionar una diana para diagnosticar una enfermedad o la predisposición a la enfermedad mediada por la proteína receptora, como se discute a continuación con respecto al tratamiento, especialmente en el cerebro (especialmente en las células de la glía) y en células de la médula ósea

CD34⁺, incluyendo pero sin limitaciones, los megacariocitos y también en el corazón, placenta, páncreas, linfocitos B en reposo, músculo esquelético, ganglio linfático, bazo, timo, hígado, amígdalas, colon, granulocitos y eritroblastos. Por consiguiente, se proporcionan procedimientos para detectar la presencia o niveles de la proteína receptora en una célula, tejido u organismo. El procedimiento implica poner en contacto una muestra biológica con un compuesto capaz de interactuar con la proteína receptores de modo que pueda detectarse la interacción.

Un agente para la detección de la proteína receptora es un anticuerpo capaz de unirse de forma selectiva a la proteína receptora. Una muestra biológica incluye tejidos, células y fluidos biológicos aislados de un sujeto, así como tejidos, células y fluidos presentes en un sujeto.

La proteína receptora también proporciona una diana para diagnosticar una enfermedad activa o la predisposición a la enfermedad, en un paciente que tiene una variante de la proteína receptora. De este modo, la proteína receptora puede aislarse a partir de una muestra biológica, y evaluarse la presencia de una mutación genética que da lugar a una proteína receptora anómala. Esto incluye la sustitución, delección, inserción, reordenamiento (como resultado de sucesos de ajuste anómalos) de aminoácidos y modificaciones postraduccionales inadecuadas. Los procedimientos analíticos incluyen movilidad electroforética alterada, digestión trípica de péptidos alterada, actividad del receptor alterada en un ensayo con células o sin células, alteración en los patrones de unión al ligando o a un anticuerpo, punto isoelectrónico alterado, secuenciación directa de aminoácidos y cualquier otra de las técnicas de ensayo conocidas útiles para detectar mutaciones en una proteína.

Las técnicas *in vitro* para detectar la proteína receptora incluyen ensayos de inmunoabsorción ligada a enzima (ELISA), inmunotransferencias, inmunoprecipitaciones e inmunofluorescencia. Alternativamente, la proteína puede detectarse *in vivo* en un sujeto introduciendo en el sujeto un anticuerpo frente al receptor marcado. Por ejemplo, el anticuerpo puede estar marcado con un marcador radiactivo cuya presencia y localización en un sujeto puede detectarse mediante técnicas convencionales de adquisición de imágenes. Son especialmente útiles los procedimientos para detectar las variantes alélicas de una proteína receptora expresada en un sujeto y los procedimientos para detectar fragmentos de una proteína receptora en una muestra.

Los polipéptidos del receptor también son útiles en el análisis farmacogenómico. Los parámetros farmacogenómicos tratan de las variaciones hereditarias clínicamente significativas en la respuesta a fármacos debido a una disposición alterada al fármaco y a su acción anómala en las personas afectadas. Véase, por ejemplo, Eichelbaum, M. (*Clin. Exp. Pharmacol. Physiol.* 23 (10-11):983-985 (1996)), y Linder, M. W. (*Clin. Chem.* 43 (2):254-266 (1997)). Los resultados clínicos de estas variaciones dan lugar a una toxicidad grave de los fármacos terapéuticos en ciertos individuos o a fracaso terapéutico de los fármacos en ciertos individuos como resultado de una variación individual del metabolismo. De este modo, el genotipo del individuo puede determinar la manera en que un compuesto terapéutico actúa en el organismo o la manera en que el organismo metaboliza el compuesto. Adicionalmente, la actividad de las enzimas que metabolizan el fármaco lleva a cabo tanto la intensidad como la duración de la acción de la droga. De este modo, los parámetros farmacogenómicos del individuo permiten la selección de compuestos eficaces y de dosificaciones eficaces de estos compuestos para el tratamiento profiláctico o terapéutico en base al genotipo del individuo. El descubrimiento de polimorfismos genéticos en algunas enzimas que metabolizan fármacos ha explicado porque algunos pacientes no obtienen los efectos del fármaco esperados, mostrando un efecto del fármaco exagerado o experimentando toxicidad seria a partir de dosis convencionales del fármaco. Los polimorfismos pueden expresarse en el fenotipo del metabolizador extenso y el fenotipo del metabolizador pobre. Por consiguiente, el polimorfismo genético puede conducir a variantes alélicas proteicas de la proteína receptora en las una o más de las funciones del receptor en una población son diferentes a los de otra población. Por tanto, los polipéptidos constituyen una diana para establecer una predisposición genética que puede afectar a la modalidad del tratamiento. De este modo, en un tratamiento basado en el ligando, el polimorfismo puede dar lugar a dominios extracelulares amino terminales y/u otras regiones de unión al ligando que son más o menos activas para la unión al ligando y la activación del receptor. Por consiguiente, la dosis del ligando podría modificarse necesariamente para maximizar el efecto terapéutico en una población determinada que contiene un polimorfismo. Como alternativa a la determinación del genotipo, pueden identificarse polipéptidos polimórficos específicos.

Los polipéptidos del receptor también son útiles para controlar los efectos terapéuticos durante los ensayos clínicos y para otro tratamiento. De este modo, la eficacia terapéutica de un agente que se diseña para aumentar o disminuir la expresión génica, pueden controlarse los niveles de proteína o la actividad del receptor a lo largo del tratamiento usando los polipéptidos del receptor como un criterio de valoración diana.

Los polipéptidos del receptor también pueden ser útiles para tratar un trastorno asociado con el receptor. Por consiguiente, los procedimientos de tratamiento incluyen el uso de un receptor soluble o de fragmentos de la proteína receptora que compite por la unión al ligando. Estos receptores o fragmentos pueden tener una afinidad mayor por el ligando de modo que proporcionan una competición eficaz.

La descripción anterior generalmente también se aplica a la secuencia mostrada en la ID SEC N° 3,

Anticuerpos

La invención también proporciona anticuerpos que se unen selectivamente a la proteína receptora 15625 y a sus variantes y fragmentos. Se considera que un anticuerpo se une selectivamente, incluso si también se une a otras proteínas que no son sustancialmente homólogas con la proteína del receptor. Estas otras proteínas comparten homología con un

ES 2 275 358 T3

fragmento o dominio de la proteína receptora. Esta conservación de regiones específicas da origen a anticuerpos que se unen a ambas proteínas en virtud de la secuencia homóloga. En este caso, se entenderá que la unión del anticuerpo a la proteína receptora sigue siendo selectiva.

5 Para generar anticuerpos, se usa un polipéptido del receptor aislado como inmunógeno para generar anticuerpos usando técnicas convencionales para la preparación de un anticuerpo policlonal y monoclonal. Pueden usarse la proteína completa o un fragmento peptídico antigénico. En la Figura 1 se muestran las regiones que tienen un alto índice de antigenicidad.

10 Los anticuerpos se preparan preferiblemente a partir de estas regiones o a partir de fragmentos diferenciados de estas regiones. Sin embargo, los anticuerpos pueden prepararse a partir de cualquier región del péptido como se describe en este documento. Un fragmento preferido produce un anticuerpo que disminuye o previene por completo la unión al ligando. Puede desarrollarse anticuerpos frente al receptor completo o a porciones del receptor, por ejemplo, el dominio intracelular carboxilo terminal, el dominio extracelular amino terminal, el dominio completo transmembranal o segmentos específicos, cualquiera de los bucles intra o extracelulares o cualquier porción de las anteriores. También pueden desarrollarse anticuerpo frente a sitios funcionales específicos, tales como el sitio de unión al ligando, el sitio de acoplamiento con la proteína G o los sitios que están glucosilados, fosforilados o miristilados.

15 Un fragmento antigénico típicamente comprenderá al menos 10 restos de aminoácidos contiguos. El péptido antigénico puede comprender, sin embargo, al menos 12, al menos 14 restos de aminoácidos, al menos 15 restos de aminoácidos, al menos 20 restos de aminoácidos, o al menos 30 restos de aminoácidos. En una realización, los fragmentos corresponden a regiones que se localizan en la superficie de la proteína, por ejemplo, regiones hidrófilas. Estos fragmentos no pretenden, sin embargo, abarcar a ningún fragmento que pueda haber sido descrito antes de la invención.

20 Los anticuerpos pueden ser policlonales o monoclonales. Puede usarse un anticuerpo intacto o un fragmento del mismo (por ejemplo, Fab o (F(ab')₂).

25 La detección puede facilitarse acoplado (es decir, uniendo físicamente) el anticuerpo con una sustancia detectable. Los ejemplos de sustancias detectables incluyen diversos enzimas, grupos protéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes y materiales radiactivos. Ejemplos de enzimas adecuadas incluyen peroxidasa de rábano picante, fosfatasa alcalina, β-galactosidasa o acetilcolinesterasa, ejemplos de complejos de grupos prostéticos adecuado incluyen estreptavidina/biotina y avidina/biotina; ejemplos de materiales fluorescentes adecuados incluyen umbeliferona, fluoresceína, isotiocianato de fluoresceína, rodamina, diclorotriazinilamina fluoresceína, cloruro de dansilo o ficoeritrina; un ejemplo de un material luminiscente incluye luminol; ejemplos de materiales bioluminiscentes incluyen luciferasa, luciferina y aequorina, y ejemplos de material radiactivo adecuado incluye ¹²⁵I, ¹³¹I, ³⁵S o ³H.

30 Una preparación inmunogénica apropiada puede derivar de una proteína nativa, recombinantemente expresada, o de péptidos químicamente sintetizados.

Usos de los anticuerpos

35 Los anticuerpos pueden usarse para aislar una proteína receptora mediante técnicas convencionales, tales como cromatografía de afinidad o inmunoprecipitación. Los anticuerpos pueden facilitar la purificación de la proteína receptora natural a partir de las células y de la proteína receptora producida recombinantemente expresada en las células huésped.

40 Los anticuerpos son útiles para detectar la presencia de la proteína receptora en células o tejidos para determinar el patrón de expresión del receptor entre los diversos tejidos de un organismo y a lo largo del desarrollo normal.

Los anticuerpos pueden usarse para detectar la proteína receptora *in situ*, *in vitro* o en un lisado de células o en el sobrenadante, para evaluar la abundancia y el patrón de expresión.

55 Los anticuerpos pueden usarse para evaluar la distribución tisular anómala o la expresión anómala durante el desarrollo.

La detección por anticuerpos de fragmentos circulantes de la proteína receptora de longitud completa puede usarse para identificar el recambio del receptor.

60 Adicionalmente, los anticuerpos pueden usarse para evaluar la expresión del receptor en enfermedades tales como las etapas activas de la enfermedad o en un individuo con una predisposición a la enfermedad, relacionada con la función del receptor. Cuando un trastorno está causado por una distribución tisular inadecuada, expresión durante el desarrollo o nivel de expresión de la proteína receptora, puede prepararse el anticuerpo frente a la proteína receptora normal. Si un trastorno se caracteriza por una mutación específica en la proteína receptora, pueden usarse anticuerpos específicos para esta proteína mutante para ensayar la presencia de la proteína receptora mutante específica. Sin embargo, también se abarcan anticuerpos hechos intracelularmente (“intracuerpos”), que podrían reconocer regiones peptídicas intracelulares del receptor.

ES 2 275 358 T3

Los anticuerpos también pueden usarse para evaluar la localización subcelular normal y anómala de células en los diversos tejidos del organismo. Puede desarrollarse anticuerpos frente al receptor completo o a porciones del receptor, por ejemplo, porciones del dominio aminoterminal extracelular o de los bucles extracelulares.

5 Los usos diagnósticos pueden aplicarse, no sólo en ensayos genéticos, sino también en el control de una modalidad de tratamiento. Por consiguiente, cuando el tratamiento pretende en última estancia corregir el nivel de expresión del receptor o la presencia de receptores anómalos y la distribución tisular o la expresión durante el desarrollo anómalas, los anticuerpos dirigidos frente al receptor o a fragmentos importantes pueden usarse para controlar la eficacia terapéutica.

10 Adicionalmente, los anticuerpos son útiles para análisis farmacogenómicos. De este modo, los anticuerpos preparados frente a las proteínas receptores polimórficas pueden usarse para identificar individuos que requieran modalidades de tratamiento modificadas.

15 Los anticuerpos son también útiles como herramientas de diagnóstico, como un marcador inmunológico para proteínas receptoras anómalas analizadas mediante movilidad electroforética, punto isoeléctrico, digestión trípica peptídica, y otros ensayos físicos conocidas en la técnica.

20 Los anticuerpos también son útiles para identificar tejidos. De este modo, cuando una proteína receptora específica se ha correlacionado con la expresión en un tejido específico, pueden usarse anticuerpos que son específicos de esta proteína receptora para identificar un tipo de tejido.

25 Los anticuerpos también son útiles para la identificación forense. Por consiguiente, cuando se ha correlacionado a un individuo con un polimorfismo genético específico que da lugar a una proteína polimórfica específica, puede usarse un anticuerpo específico de la proteína polimórfica como ayuda a la identificación.

Los anticuerpos también son útiles para inhibir la función del receptor, por ejemplo, bloqueando la unión al ligando.

30 Estos usos también pueden aplicarse en un contexto terapéutico en el que el tratamiento implica inhibir la función del receptor. Puede usarse un anticuerpo para, por ejemplo, bloquear la unión al ligando. Pueden prepararse anticuerpos frente a fragmentos específicos que contiene sitios requeridos para la función o frente al receptor intacto asociado con una célula.

35 La invención también abarca kits para usar los anticuerpos en la detección de la presencia de una proteína receptora en una muestra biológica. Por ejemplo, el kit puede comprender anticuerpos tales como un anticuerpo marcado o susceptible de marcaje y un compuesto o agente para detectar la proteína receptora en una muestra biológica, medios para determinar la cantidad de proteína receptora en la muestra, y medios para comparar la cantidad de proteína receptora en la muestra con un patrón. El compuesto o agente puede empaquetarse en un recipiente adecuado. El kit además puede comprender instrucciones de uso del kit para detectar la proteína receptora.

40 La descripción anterior generalmente también se aplica a la secuencia mostrada en la ID SEC N° 3.

Polinucleótidos

45 El ADNc descrito específicamente comprende la región codificante y las secuencias no traducidas 5' y 3' (ID SEC N° 2).

50 El ADNc del receptor 15625 humano tiene una longitud de aproximadamente 2.286 nucleótidos y codifica una proteína completa que tiene una longitud de aproximadamente 342 restos aminoácidos. El ácido nucleico se expresa como se describe en este documento, tal como en el cerebro, especialmente en las células de la glía, en células CD34⁺ y en las líneas celulares 293 y Jurkat. En la Figura 2 se proporciona el análisis estructural de la secuencia de aminoácidos de la ID SEC N° 1, un gráfico de hidropatía. La figura muestra la estructura posible de los siete segmentos transmembranal, el dominio extracelular amino terminal y el dominio intracelular carboxilo terminal.

55 Según se usa en este documento, la expresión “segmento transmembranal” se refiere a un motivo estructural aminoácido que incluye una hélice hidrófoba que abarca la membrana plasmática. El dominio transmembranal completo abarca de aproximadamente el aminoácido 26 a aproximadamente el aminoácido 302. Se extienden siete segmentos por la membrana y en este dominio hay tres bucles intracelulares y tres extracelulares.

60 La invención proporciona polinucleótidos aislados que codifican una proteína receptor 15625. La expresión “polinucleótido 15625” o “ácido nucleico 15625” se refiere a la secuencia mostrada en la ID SEC N° 2. La expresión “polinucleótido receptor” o “ácido nucleico receptor” además incluye variantes y fragmentos del polinucleótido 15625, tales como los mostrados en la ID SEC N° 4.

65 Un ácido nucleico receptor “aislado” es aquel que se separa de otro ácido nucleico presente en la fuente natural del ácido nucleico receptor. Preferiblemente, un ácido nucleico “aislado” está libre de secuencias que flanquean de forma natural al ácido nucleico (es decir, secuencias localizadas en los extremos 5' y 3' del ácido nucleico) en el ADN genómico del organismo del cual deriva el ácido nucleico. Sin embargo, puede haber algunas secuencias de

ES 2 275 358 T3

nucleótidos flanqueantes, por ejemplo, de hasta aproximadamente 5 kB. El punto importante es que el ácido nucleico está aislado de las secuencias flanqueantes de modo que puede someterse a las manipulaciones específicas descritas en este documento, tales como expresión recombinante, preparación de sondas y cebadores y otros usos específicos de las secuencias del ácido nucleico del receptor.

Además, una molécula de ácido nucleico "aislada", tal como una molécula de ADNc, puede estar sustancialmente libre de otro material celular, o de medio de cultivo cuando se produce por técnicas recombinantes o de precursores químicos u otras sustancias químicas cuando se sintetizan químicamente. Sin embargo, la molécula de ácido nucleico puede fusionarse con otras secuencias codificantes o reguladoras y seguir considerándose aislada.

Por ejemplo, las moléculas de ADN recombinante contenidas en un vector se consideran aisladas. Otros ejemplos de moléculas de ADN aisladas incluyen moléculas de ADN recombinantes mantenidas en células huésped heterólogas o moléculas de ADN purificadas (parcial o sustancialmente) en solución. Las moléculas de ARN aisladas incluyen transcripciones de ARN *in vivo* o *in vitro* de las moléculas de ADN aisladas de la presente invención. Las moléculas de ácido nucleico aisladas según la presente invención además incluyen estas moléculas producidas sintéticamente.

Los polinucleótidos del receptor pueden codificar la proteína madura más aminoácidos amino o carboxilo terminales adicionales o aminoácidos del interior del polipéptido maduro (cuando la forma madura tiene más de una cadena polipeptídica, por ejemplo). Estas secuencias pueden jugar un papel en el procesamiento de una proteína a partir de un precursor a una forma madura, facilitar el tráfico de la proteína, prolongar o acortar la semivida de la proteína o facilitar la manipulación de una proteína para ensayo o producción, entre otras cosas. Como generalmente es el caso *in situ*, los aminoácidos adicionales pueden eliminarse de la proteína madura mediante enzimas celulares.

Los polinucleótidos del receptor incluyen, pero sin limitaciones, la secuencia que codifica el polipéptido maduro sólo, la secuencia que codifica el polipéptido maduro y secuencias codificantes adicionales, tales como una secuencia líder o secretora (por ejemplo, una secuencia pre-pro o pro-proteína), la secuencia que codifica el polipéptido maduro, con o sin las secuencias codificantes adicionales, además de secuencias no codificantes adicionales, por ejemplos intrones y secuencias 5' y 3' no codificantes, tales como secuencias transcritas pero no traducidas que juegan un papel en la transcripción, en el procesamiento del ARNm (incluyendo las señales de ajuste y poliadenilación), en la unión al ribosoma y la estabilización del ARNm. Además, los polinucleótidos pueden fusionarse con una secuencia marcadora que codifica, por ejemplo, un péptido que facilita la purificación.

Los polinucleótidos del receptor pueden estar en forma de ARN, tal como ARNm, o en forma de ADN, incluyendo ADNc y ADN genómico obtenido por clonación o producidos por técnicas de síntesis química o mediante una combinación de los mismos. El ácido nucleico, especialmente ADN, puede ser de hebra doble o de hebra sencilla. El ácido nucleico de hebra sencilla puede ser la hebra codificante (cadena sentido) o la hebra no codificante (hebra antisentido).

Un ácido nucleico del receptor comprende la secuencia de nucleótidos mostrada en la ID SEC N° 2, que se corresponde con el ADNc de cerebro humano.

En una realización, el ácido nucleico del receptor comprende sólo la región codificante.

La invención además proporciona variantes de los polinucleótidos del receptor y fragmentos de los mismos que difieren de la secuencia de nucleótidos mostrada en la ID SEC N° 2 debido a la degeneración del código genético y, además codifica la misma proteína que la codificada por la secuencia de nucleótidos mostrada en la ID SEC N° 2.

La invención también proporciona moléculas de ácido nucleico del receptor que codifican las variantes polipeptídicas descritas en este documento. Estos polinucleótidos pueden ser naturales, tales como variantes alélicas (mismo locus) (que se sitúan en el cromosoma 3 próximos a AFM164YG9), homólogos (locus diferente) y ortólogos (organismo diferente), tal como la secuencia mostrada en la ID SEC N° 4 o pueden construirse mediante procedimientos de ADN recombinante o mediante síntesis química. Estas variantes no naturales pueden hacerse mediante técnicas de mutagénesis, incluyendo las aplicadas a polinucleótidos, células u organismos. Por consiguiente, como se discute anteriormente, las variantes pueden contener sustituciones, deleciones, inversiones e inserciones de nucleótidos.

La variación puede darse en una o ambas regiones codificante y no codificante. Las variaciones pueden producir sustituciones de aminoácidos tanto conservadoras como no conservadoras.

Las variantes ortólogas, homólogas y alélicas pueden identificarse usando procedimientos bien conocidos en la técnica. Estas variantes comprenden una secuencia de nucleótidos que codifica un receptor que es el 50%, al menos aproximadamente el 55%, típicamente al menos aproximadamente el 70-75%, más típicamente al menos aproximadamente el 80-85% y lo más típicamente al menos aproximadamente el 90-95% o más homólogo con la secuencia de nucleótidos mostrada en la ID SEC N° 2 o un fragmento de esta secuencia. Estas moléculas de ácido nucleico pueden identificarse fácilmente como capaces de hibridar en condiciones rigurosas, con la secuencia de nucleótidos mostrada en la ID SEC N° 2 o un fragmento de la secuencia. Se entiende que la hibridación rigurosa no indica homología sustancial cuando es debida a homología general, tal como secuencias poli A o secuencias comunes a todas o a la mayoría de las proteínas, a todas los GPCR o a toda la familia de GPCR I. Además, se entiende que estas variantes no incluyen a ninguna de las secuencias de ácido nucleico que pueden haber sido descritas antes de la invención.

ES 2 275 358 T3

Según se usa en este documento, se pretende que la expresión “hibrida en condiciones rigurosas” describa las condiciones para la hibridación y el lavado en las que las secuencias de nucleótidos que codifican los receptores al menos el 50%, 55% homólogas entre sí típicamente permanece hibridadas. Las condiciones pueden ser tales que las secuencias al menos aproximadamente el 65%, al menos aproximadamente el 70%, o al menos aproximadamente el 75% o más homólogas entre sí típicamente permanecen hibridadas. Los expertos en la materia conocen estas condiciones rigurosas que puede encontrarse en *Current Protocols in Molecular Biology*, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6. Un ejemplo de condiciones de hibridación rigurosas es la hibridación en cloruro sódico/citrato sódico x 6 (SSC) a aproximadamente 45°C, seguida de uno o más lavados en SSC x 0,2, SDS al 0,1% a 50-65°C. En una realización, una molécula de ácido nucleico del receptor aislado que hibrida en condiciones rigurosas con la ID SEC N° 2 se corresponde con una molécula de ácido nucleico natural. Según se usa en este documento, una molécula de ácido nucleico “natural” se refiere a una molécula de ARN o ADN que tiene una secuencia de nucleótidos que aparece en la naturaleza (por ejemplo, codifica una proteína natural).

Además, la invención proporciona polinucleótidos que comprenden un fragmento de los polinucleótidos del receptor de longitud completa. El fragmento puede ser de hebra sencilla o doble y puede comprender ADN o ARN. El fragmento puede derivar tanto de la secuencia codificante como de la no codificante.

Un fragmento puede comprender una secuencia de nucleótidos contiguos mayor de 12 nucleótidos, del nucleótido 1 a aproximadamente el nucleótido 500, mayor de 24 nucleótidos de aproximadamente el nucleótido 476 a aproximadamente el nucleótido 1.096 y mayor de 12 nucleótidos de aproximadamente el nucleótido 1.147 al nucleótido 1.715.

Los fragmentos de ácido nucleico del receptor aislados hibridan en condiciones rigurosas con la molécula de ácido nucleico que comprende la secuencia de nucleótidos de la ID SEC N° 2. En otras realizaciones, el ácido nucleico tiene una longitud de al menos 30, 40, 50, 100, 250 ó 500 nucleótidos.

En otra realización, un ácido nucleico del receptor aislado codifica la región codificante completa del aminoácido 1 al aminoácido 342. En otra realización, el ácido nucleico del receptor aislado codifica una secuencia que se corresponde con la proteína madura de aproximadamente el aminoácido 6 al aminoácido 342. Otros fragmentos incluyen secuencias de nucleótidos que codifican los fragmentos de aminoácidos descritos en este documento. Fragmentos adicionales pueden incluir subfragmentos de los dominios o sitios específicos descritos en este documento. No se interpreta que los fragmentos de ácido nucleico, según la presente invención, incluyan a aquellos fragmentos que puedan haber sido descritos antes de la invención.

Los fragmentos de ácido nucleico del receptor incluyen además secuencias que se corresponden con los dominios descritos en este documento; subregiones también descritas y sitios funcionales específicos. Los fragmentos de ácido nucleico del receptor también incluyen combinaciones de los dominios, segmentos, bucles y otros sitios funcionales descritos anteriormente. De este modo, por ejemplo, un ácido nucleico del receptor podría incluir secuencias correspondientes con el dominio extracelular amino terminal y un fragmento transmembranal. Una persona experta en la materia conocerá las numerosas permutaciones posibles.

Sin embargo, se entiende que un fragmento del receptor incluye cualquier secuencia de ácido nucleico que no incluya el gen completo.

Los fragmentos de ácido nucleico del receptor incluyen moléculas de ácido nucleico que codifican un polipéptido que comprende el dominio extracelular amino terminal que incluye los restos de aminoácidos del 1 a aproximadamente el 25, un polipéptido que comprende la región que abarca el dominio transmembranal (restos de aminoácidos de aproximadamente el 26 a aproximadamente el 302), un polipéptido que comprende el dominio intracelular carboxilo terminal (restos de aminoácidos de aproximadamente el 303 a aproximadamente el 342) y un polipéptido que codifica el motivo distintivo del receptor de la proteína G (restos de aminoácidos 121-123 o adyacentes de aproximadamente el 110 a aproximadamente el 133); moléculas de ácido nucleico que codifica cualquiera de los siete segmentos transmembrana, bucles extracelulares o intracelulares y sitios de glucosilación, fosforilación por proteínas cinasas dependientes de AMPc y GMPc, fosforilación por proteína cinasa C y N-miristilación. Cuando se ha predicho la localización de los dominios mediante análisis por ordenador, un experto apreciará que los restos de aminoácidos que constituyen estos dominios pueden variar dependiendo de los criterios usados para definir los dominios.

La invención también proporciona fragmentos de ácido nucleico del receptor que codifica regiones portadoras de epítopes de las proteínas receptoras descritas en este documento.

Las secuencias de polinucleótido del receptor aisladas y, especialmente los fragmentos, son útiles como sondas y cebadores de ADN.

Por ejemplo, la región codificante de un gen del receptor puede aislarse usando la secuencia de nucleótidos conocida para sintetizar una sonda oligonucleotídica. A continuación puede usarse una sonda marcada para analizar una biblioteca de ADNc, una biblioteca de ADN genómico o el ARNm para aislar el ácido nucleico que se corresponde con la región codificante. Adicionalmente, pueden usarse cebadores en reacciones de PCR para clonar regiones específicas de los genes del receptor.

ES 2 275 358 T3

Una sonda/cebador típicamente comprende un oligonucleótido sustancialmente purificado. El oligonucleótido típicamente comprende una región de secuencia nucleotídica que hibrida en condiciones rigurosas con al menos aproximadamente 12, típicamente aproximadamente 25, más típicamente aproximadamente 40, 50 ó 75 nucleótidos consecutivos de la hebra sentido o complementaria de la ID SEC N° 2 u otros polinucleótidos del receptor. Una sonda además comprende un marcador, por ejemplo, un radioisótopo, compuesto fluorescente, encima o cofactor de una enzima.

La descripción anterior generalmente también se aplica a la secuencia mostrada en la ID SEC N° 4,

10 Usos de los polinucleótidos

Los polinucleótidos del receptor son útiles como sondas, cebadores y en ensayos biológicos. Cuando los polinucleótidos se usan para evaluar las propiedades del GPCR, propiedades o funciones, tales como en los ensayos descritos en este documento, puede ser útil todo o casi todo el ADNc completo. En este caso, se incluyen incluso los fragmentos que puedan ser conocidos antes de la invención. De este modo, por ejemplo, los ensayos específicamente dirigidos a las funciones de los GPCR, tales como la evaluación de la actividad agonista o antagonista, abarca el uso de fragmentos conocidos. Adicionalmente, también pueden practicarse procedimientos diagnósticos para evaluar la función del receptor con cualquier fragmento, incluyendo los fragmentos que puedan ser conocidos antes de la invención. De forma similar, en procedimientos que implican el tratamiento de la disfunción del receptor, están abarcados todos los fragmentos incluyendo aquellos que puedan ser conocidos en la técnica.

Los polinucleótidos receptores son útiles como sonda de hibridación del ADNc y del ADN genómico para aislar un ADNc de longitud completa y clones genómicos que codifican el polipéptido descrito en la ID SEC N° 1 y para aislar clones de ADNc y genómicos que se correspondan con variantes que producen el mismo polipéptido mostrado en la ID SEC N° 1 o las otras variantes descritas en este documento. Las variantes pueden aislarse a partir del mismo tejido y organismo del que se ha aislado el polipéptido mostrado en la ID SEC N° 1, de tejidos diferentes del mismo organismo o de organismos diferentes. Este procedimiento es útil para aislar genes y ADNc que están controlados por el desarrollo y, por tanto, puede expresarse en el mismo tejido o en tejidos diferentes en puntos diferentes del desarrollo de un organismo.

La sonda puede corresponderse con cualquier secuencia a lo largo de la longitud completa del gen que codifica el receptor. Por consiguiente, podría derivar de las regiones 5' no codificantes, de la región codificante y de las regiones 3' no codificantes. Sin embargo, como se describe, los fragmentos no deben interpretarse como aquellos que puedan abarcar fragmentos descritos antes de la presente invención.

La sonda de ácido nucleico puede ser, por ejemplo, el ADNc de longitud completa de la ID SEC N° 1, o un fragmento del mismo, tal como un oligonucleótido con una longitud de al menos 12, 15, 30, 50, 100, 250 ó 500 nucleótidos y suficiente para hibridar específicamente en condiciones rigurosas con ARNm o ADN.

Los fragmentos de los polinucleótidos descritos en este documento también pueden ser útiles para sintetizar fragmentos mayores o polinucleótidos de longitud completa descritos en este documento. Por ejemplo, un fragmento puede hibridar con cualquier porción de un ARNm y puede producir un ADNc mayor o de longitud completa.

Los fragmentos también son útiles para sintetizar moléculas complementarias de longitud y secuencia deseadas.

Los polinucleótidos receptores también son útiles como cebadores de PCR para amplificar cualquier región determinada de un polinucleótido del receptor.

Los polinucleótidos del receptor también son útiles para construir vectores recombinantes. Estos vectores incluyen vectores de expresión que expresan una porción de, o todos, los polipéptidos del receptor. Los vectores también incluyen vectores de inserción, usados para integrarse en otra secuencia polinucleotídica, tal como dentro del genoma celular, y alterar la expresión *in situ* de los genes del receptor y de los productos génicos. Por ejemplo, una secuencia de codificación del receptor endógeno puede sustituirse mediante recombinación homóloga con toda o parte de la región codificante que contiene una o más mutaciones introducidas específicamente.

Los polinucleótidos del receptor también son útiles para expresar porciones antigénicas de las proteínas receptoras.

Los polinucleótidos receptores también son útiles como sondas para determinar las posiciones cromosómicas de los polinucleótidos receptores mediante procedimientos de hibridación *in situ*.

Las sondas polinucleótidos del receptor también son útiles para determinar los patrones de presencia del gen que codifica los receptores y sus variantes con respecto a la distribución tisular, por ejemplo, si ha tenido lugar una duplicación génica y si la duplicación tiene lugar en todos o sólo en un subgrupo de tejidos. Los genes pueden ser naturales o pueden introducirse en una célula, tejido, organismo de forma exógena.

Los polinucleótidos del receptor también son útiles para diseñar ribozimas correspondientes con todo, o parte, del ARNm producido a partir de los genes que codifican los polinucleótidos descritos en este documento.

ES 2 275 358 T3

Los polinucleótidos del receptor también son útiles para construir células huésped que expresen una parte o todos los polinucleótidos y polipéptidos del receptor.

5 Los polinucleótidos del receptor también son útiles para construir animales transgénicos que expresen una parte o todos los polinucleótidos y polipéptidos del receptor.

Los polinucleótidos del receptor también son útiles para construir vectores que expresen una parte o todos los polipéptidos del receptor.

10 Los polinucleótidos del receptor también son útiles como sondas de hibridación para determinar el nivel de expresión del ácido nucleico del receptor. Por consiguiente, las sondas pueden usarse para detectar la presencia o para determinar los niveles de ácido nucleico del receptor en células, tejidos y en organismos. El ácido nucleico cuyo nivel se determina puede ser ADN o ARN. Por lo tanto, las sondas, que se corresponden con los polipéptidos descritos en este documento, pueden usarse para evaluar el número de copias de un gen en una célula, tejido u organismo determinado. Esto es especialmente importante en casos en los que ha habido una amplificación de los genes del receptor.

20 Alternativamente, la sonda puede usarse en un contexto de hibridación *in situ* para evaluar la posición de copias extra de los genes del receptor, como en elementos extracromosómicos, o integrados en cromosomas en los que normalmente no se encuentra el gen del receptor, por ejemplo, como una región teñida homogéneamente.

Estos usos son importantes para diagnosticar trastornos que implican un aumento o descenso en la expresión del receptor en relación con resultados normales.

25 Las técnicas *in vitro* para la detección de ARNm incluyen hibridaciones de transferencia de ARN total e hibridaciones *in situ*. Las técnicas *in vitro* para la detección de ADN incluyen hibridaciones de transferencia de ADN e hibridaciones *in situ*.

30 Las sondas pueden usarse como parte de un kit de ensayo diagnóstico para identificar células o tejidos que expresan una proteína receptora, tal como midiendo el nivel de un ácido nucleico que codifica el receptor en una muestra de células de un sujeto, por ejemplo, ARNm o ADN genómico, o determinando si un gen del receptor ha sido mutado.

Los ensayos de expresión de ácido nucleico son útiles en el análisis de fármacos para identificar compuestos que modulen la expresión del ácido nucleico del receptor.

35 Por tanto, la invención proporciona un procedimiento para identificar un compuesto que pueda usarse para tratar un trastorno asociado con la expresión del ácido nucleico del gen del receptor. El procedimiento típicamente incluye evaluar la capacidad del compuesto para modular la expresión del ácido nucleico del receptor y, por tanto, identificar un compuesto que pueda usarse para tratar un trastorno caracterizado por la expresión no deseada de un ácido nucleico del receptor.

40 Los ensayos pueden realizarse en sistemas basados en células o sin células. Los ensayos basados en células incluyen células que expresan de forma natural el ácido nucleico del receptor o células recombinantes modificadas genéticamente para expresar secuencias específicas de ácido nucleico.

45 Alternativamente, los compuestos candidatos pueden evaluarse *in vivo* en pacientes o en animales transgénicos.

50 El ensayo para la expresión del ácido nucleico del receptor puede implicar un ensayo directo de los niveles de ácido nucleico, tales como los niveles de ARNm, o de compuestos colaterales implicados en la ruta de señalización (tales como recambio de AMPc o de fosfatidilinositol). Adicionalmente, también pueden evaluarse la expresión de genes que se regulan por incremento o por descenso en respuesta a la ruta de señalización de la proteína receptora. En esta realización, las regiones reguladoras de estos genes pueden unirse de forma operativa a un gen indicador, tal como la luciferasa.

55 De este modo, pueden identificarse moduladores de la expresión del gen del receptor en un procedimiento en el que una célula se pone en contacto con un compuesto candidato y se determina la expresión del ARNm. El nivel de expresión del ARNm del receptor en presencia del compuesto candidato se compara con el nivel de expresión del ARNm del receptor en ausencia del compuesto candidato. A continuación, el compuesto candidato puede identificarse como un modulador de la expresión de ácido nucleico en base a esta comparación y usarse, por ejemplo, para tratar un trastorno caracterizado por la expresión anómala del ácido nucleico. Cuando estadísticamente la expresión del ARNm es significativamente mayor en presencia del compuesto candidato que en su ausencia, el compuesto candidato se identifica como un estimulador de la expresión del ácido nucleico. Cuando estadísticamente la expresión del ácido nucleico es significativamente menor en presencia del compuesto candidato que en su ausencia, el compuesto candidato se identifica como un inhibidor de la expresión del ácido nucleico.

65 Por consiguiente, la invención proporciona procedimientos de tratamiento, con el ácido nucleico como objetivo, usando un compuesto identificado a través del análisis de fármacos como un modulador del gen para modular la expresión del ácido nucleico del receptor. La modulación incluye tanto la regulación por incremento (es decir, activación o agonización) como la regulación por descenso (supresión o antagonización) o la expresión del ácido nucleico.

ES 2 275 358 T3

Alternativamente, un modulador para la expresión del ácido nucleico del receptor puede ser una molécula pequeña o un fármaco identificado usando los ensayos de análisis descritos en este documento siempre que el fármaco o la molécula pequeña inhiba la expresión del ácido nucleico del receptor.

5 Los polinucleótidos del receptor también se usan para controlar la eficacia de los compuestos moduladores en la expresión o actividad del gen del receptor en ensayos clínicos o en un régimen de tratamiento. De este modo, el patrón de expresión del gen puede servir como barómetro de la eficacia continuada del tratamiento con el compuesto, especialmente con compuestos a los que un paciente puede desarrollar resistencia. El patrón de expresión del gen también puede servir como marcador indicativo de una respuesta fisiológica de las células afectadas al compuesto.
10 Por consiguiente, este control permitirá aumentar la administración del compuesto o la administración de compuestos alternativos a los que el paciente no se ha hecho resistente. De forma similar, si el nivel de expresión del ácido nucleico cae por debajo de un nivel deseado, podría disminuirse de forma proporcional la administración del compuesto.

15 Los polinucleótidos del receptor también son útiles en los ensayos diagnósticos de cambios cualitativos en el ácido nucleico del receptor, y especialmente en cambios cualitativos que lleven a una patología. Los polinucleótidos pueden usarse para detectar mutaciones en los genes del receptor y en los productos de expresión del gen, tales como ARNm. Los polinucleótidos pueden usarse como sondas de hibridación para detectar mutaciones genéticas naturales en el gen del receptor y para determinar de este modo si un sujeto con la mutación sufre el riesgo de desarrollar un trastorno causado por la mutación. Las mutaciones incluyen delección, adición o sustitución de uno o más nucleótidos en el
20 gen, reordenamiento cromosómico, tales como inversión o transposición, modificación del ADN genómico, tal como patrones de metilación anómalos o cambios en el número de copias del gen, tal como amplificación. La detección de una forma mutada del gen del receptor asociada con una disfunción proporciona una herramienta diagnóstica de una enfermedad activa o de la susceptibilidad a una enfermedad cuando la enfermedad es el resultado de sobreexpresión, infraexpresión o expresión alterada de una proteína receptora.

25 Los individuos portadores de las mutaciones en el gen del receptor pueden detectarse a nivel del ácido nucleico mediante diversas técnicas. El ADN genómico puede analizarse directamente o puede amplificarse usando PCR antes del análisis. Pueden usarse ARN o ADNc de la misma manera.

30 En ciertas realizaciones, la detección de la mutación implica el uso de una sonda/cebador en una reacción en cadena de la polimerasa (PCR) (véanse, por ejemplo, las Patentes de EE.UU. N° 4.683.195 y 4.683.202), tal como PCR anclada o PCR RACE o, alternativamente, en la reacción en cadena de la ligasa (LCR) (véase, por ejemplo, Landegran y col., *Science* 241:1077-1080 (1988); y Nakazawa y col., *PNAS* 91:360-364 (1994)), la última de las cuales puede ser especialmente útil para detectar mutaciones puntuales en el gen (véase Abravaya y col., *Nucleic Acids Res.* 23:675-682 (1995)). Este procedimiento puede incluir las etapas de recoger una muestra de células de un paciente, aislar el
35 ácido nucleico (por ejemplo, genómico, ARNm o ambos) de las células de la muestra, poner en contacto la muestra de ácido nucleico con uno o más cebadores que hibriden específicamente con un gen en condiciones en las que tenga lugar la hibridación y amplificación del gen (si está presente) y detectar la presencia o ausencia de un producto de amplificación, o detectar el tamaño del producto de amplificación y comparar la longitud con una muestra control. Las delecciones e inserciones pueden detectarse mediante un cambio en el tamaño del producto amplificado en comparación con el genotipo normal. Las mutaciones puntuales pueden identificarse hibridando el ADN amplificado con secuencias de ARN normales o ADN complementarias.

45 Alternativamente, las mutaciones en un gen del receptor pueden identificarse directamente, por ejemplo, mediante alteraciones en los patrones de digestión con enzimas de restricción determinado por electroforesis en gel.

Adicionalmente, las ribozimas específicas de secuencia (Patente de EE.UU. N° 5.498.531) pueden usarse para valorar la presencia de mutaciones específicas desarrollando o quitando un sitio de escisión de ribozimas.

50 Las secuencias que coinciden perfectamente pueden distinguirse de las secuencias con errores de apareamiento mediante ensayos de digestión con nucleasas de escisión o mediante diferencias en la temperatura de fusión.

Los cambios en la secuencia en localizaciones específicas también pueden valorarse mediante ensayos de protección de nucleasa, tales como de protección de ARNasa y de S1 o el procedimiento de escisión química.

55 Adicionalmente, las diferencias de secuencia entre un gen del receptor mutante y un gen de tipo natural pueden determinarse mediante la secuenciación directa del ADN. Pueden utilizarse diversos procedimientos de secuenciación automática cuando se realizan los ensayos diagnósticos ((1995) *Biotechniques* 19:448), incluyendo la secuenciación mediante espectrometría de masas (véase, por ejemplo, la publicación internacional PCT N° WO 94/16101; Cohen y col., *Adv. Chromatogr.* 36:127-162 (1996) y Griffin y col., *Appl. Biochem. Biotechnol.* 38:147-159 (1993)).
60

Otros procedimientos para detectar mutaciones en el gen incluyen procedimientos en los que se usa la protección frente a agentes de escisión para detectar bases desapareadas en los dupletes ARN/ARN o ARN/ADN (Myers y col., *Science* 230:1242 (1985)); Cotton y col., *PNAS* 85:4397 (1988); Saleeba y col., *Meth. Enzymol.* 217:286-295 (1992)), en los que se compara la movilidad electroforética de los ácidos nucleicos mutante y de tipo natural (Orita y col., *PNAS* 86:2766 (1989); Cotton y col., *Mutat. Res.* 285:125-144 (1993); y Hayashi y col., *Genet. Anal. Tech. Appl.* 9:73-79 (1992)) y el movimiento de fragmentos mutantes o de tipo natural en geles de poliacrilamida que contiene un gradiente de desnaturalización, se ensaya usando geles de electroforesis con un gradiente desnaturalizante (Myers y col., *Nature*

313:495 (1985)). Los ejemplos de otras técnicas para detectar mutaciones puntuales incluyen la hibridación selectiva de oligonucleótidos, la amplificación selectiva y la extensión selectiva con un cebador.

5 Los polinucleótidos del receptor también son útiles para ensayar el genotipo de un individuo que, aunque no causa necesariamente la enfermedad, afecta sin embargo a la modalidad de tratamiento. De este modo, los polinucleótidos pueden usarse para estudiar la relación entre un genotipo de un individuo y la respuesta del individuo a un compuesto usado para el tratamiento (relación farmacogenómica). En el caso presente, por ejemplo, una mutación en el gen del receptor que da lugar a una afinidad alterada por el ligando podría dar lugar a un efecto excesivo o disminuido del fármaco con concentraciones convencionales del ligando que activa al receptor. Por consiguiente, los polinucleótidos del receptor descritos en este documento pueden usarse para evaluar el contenido de mutación del gen del receptor en un individuo para seleccionar un compuesto o régimen de dosificación apropiada para el tratamiento.

15 De este modo, los polinucleótidos que muestran variaciones genéticas que afectan al tratamiento proporcionan un objetivo diagnóstico que puede usarse para adaptar el tratamiento a un individuo. Por consiguiente, la producción de células y animales recombinantes que contienen estos polimorfismos permite un diseño clínico eficaz de compuestos de tratamiento y regímenes de dosificación.

20 Los polinucleótidos del receptor también son útiles para la identificación cromosómica cuando la secuencia se identifica con un cromosoma individual y con una localización en particular en el cromosoma. En primer lugar, la secuencia de ADN se hace coincidir con el cromosoma por hibridación *in situ* u otra hibridación específica para cromosomas. Las secuencias también pueden correlacionarse con cromosomas específicos preparando cebadores de PCR que pueden usarse para analizar por PCR híbridos de células somáticas que contienen cromosomas individuales de las especies deseadas. Sólo los híbridos que contengan el cromosoma con el gen homólogo para el cebador producirán un fragmento amplificado. Puede conseguirse una sublocalización usando fragmentos cromosómicos. Otras estrategias incluyen la preselección con cromosomas separados por flujo marcados y la preselección mediante hibridación con bibliotecas específicas de cromosoma. Estrategias de mapeo adicionales incluyen hibridación *in situ* fluorescente que permite la hibridación con sondas más cortas que las usadas tradicionalmente. Los reactivos para el mapeo cromosómico pueden usarse individualmente para marcar un único cromosoma o un único sitio en este cromosoma, o pueden usarse paneles de reactivos para marcar sitios múltiples y/o cromosomas múltiples. Realmente, para mapear se prefieren reactivos que se corresponden con regiones no codificantes de los genes. Es más probable que las secuencias codificantes estén conservadas dentro de las familias de genes, aumentando de este modo la posibilidad de hibridaciones cruzadas durante el mapeo cromosómico.

35 Los polinucleótidos del receptor también pueden usarse para identificar individuos a partir de pequeñas muestras biológicas. Esto puede hacerse, por ejemplo, usando polimorfismo de longitud de fragmentos de restricción (RFLP) para identificar a un individuo. De este modo, los polinucleótidos descritos en este documento son útiles como marcadores de ADN para RFLP (Véase la Patente de EE.UU. N° 5.272.057).

40 Además, la secuencia del receptor puede usarse para proporcionar una técnica alternativa que determine la secuencia de ADN real de fragmentos seleccionados en el genoma de un individuo. De este modo, las secuencias del receptor descritas en este documento pueden usarse para preparar dos cebadores de PCR a partir de los extremos 5' y 3' de las secuencias. Estos cebadores pueden usarse a continuación para amplificar el ADN de un individuo para su secuenciación posterior.

45 Los paneles de secuencias de ADN correspondiente a individuos preparados de esta manera pueden proporcionar identificaciones individuales únicas, de modo que cada individuo tendrá un único conjunto de estas secuencias de ADN. Se estima que la variación alélica en humanos tiene lugar con una frecuencia de aproximadamente una de cada 500 bases. La variación alélica tiene lugar en cierto grado en las regiones codificantes de estas secuencias y en mayor grado en las regiones no codificantes. Las secuencias receptoras pueden usarse para obtener estas secuencias de identificación de individuos y de tejidos. Las secuencias representan fragmentos únicos del genoma humano. Cada una de las secuencias descritas en este documento puede, en cierto grado, usarse como patrón frente al que puede compararse el ADN de un individuo con fines de identificación.

55 Si se usa un panel de reactivos de las secuencias para generar una base de datos de identificación única de un individuo, estos mismos reactivos pueden usarse posteriormente para identificar el tejido de este individuo. Usando la base de datos de identificación única, puede hacerse una identificación del individuo, vivo o muerto, a partir de muestras de tejido extremadamente pequeñas.

60 Los polinucleótidos del receptor también pueden usarse en procedimientos de identificación forense. La tecnología de PCR puede usarse para amplificar secuencias de ADN tomando muestras biológicas muy pequeñas, tales como un único folículo piloso o fluidos corporales (por ejemplo, sangre, saliva o semen). La secuencia amplificada puede compararse a continuación con un patrón lo que permite la identificación del origen de la muestra.

65 Los polinucleótidos del receptor pueden usarse por tanto para proporcionar reactivos polinucleotídicos, por ejemplo, cebadores de PCR, dirigidos hacia loci específicos del genoma humano, que pueden potenciar la fiabilidad de las identificaciones forenses basadas en el ADN proporcionando, por ejemplo, otro "marcador de identificación" (es decir, otra secuencia de ADN que es única de un individuo en particular). Como se describe anteriormente, la información actual de la secuencia de bases puede usarse para la identificación como una alternativa precisa a los patrones forma-

dos por fragmentos generados por enzimas de restricción. Las secuencias dirigidas hacia las regiones no codificantes son especialmente útiles ya que los mayores polimorfismos tienen lugar en las regiones no codificantes, haciendo más fácil diferenciar a los individuos usando esta técnica.

5 Los polinucleótidos del receptor pueden además usarse para proporcionar reactivos polinucleotídicos, por ejemplo, sondas marcadas o susceptibles de marcaje, que pueden usarse, por ejemplo, en una técnica de hibridación *in situ*, para identificar un tejido específico. Esto es útil en casos en que a un patólogo forense se le presente un tejido de origen desconocido. Los paneles de las sondas del receptor pueden usarse para identificar el tejido por especie y/o por tipo de

10

De modo similar, estos cebadores y sondas pueden usarse para analizar la contaminación en cultivos tisulares (es decir, análisis de la presencia de una mezcla de diferentes tipos de células en un cultivo).

15 Alternativamente, los polinucleótidos del receptor pueden usarse directamente para bloquear la transcripción o traducción de las secuencias del gen del receptor mediante construcciones complementarias o ribozimas. De este modo, en un trastorno caracterizado por la expresión del gen del receptor anómalamente elevada o no deseada, los ácidos nucleicos pueden usarse directamente para el tratamiento.

20 Los polinucleótidos del receptor son útiles por tanto como construcciones complementarias para controlar la expresión del gen del receptor en células, tejidos y organismos. Se diseña un polinucleótido de ADN complementario para que sea complementario con una región del gen implicada en la transcripción, que prevenga la transcripción y, por tanto, la producción de la proteína receptora. Un polinucleótido de ARN o ADN complementario hibridaría con el ARNm y, por tanto, bloquearía la traducción del ARNm a la proteína receptora.

25 Ejemplos de moléculas complementarias útiles para inhibir la expresión del ácido nucleico incluyen moléculas complementarias a un fragmento de la región 5' no traducida de la ID SEC N° 2 que también incluye el codón de inicio y moléculas complementarias que son complementarias a un fragmento de la región 3' no traducida de la ID SEC N° 2.

30 Alternativamente, puede usarse una clase de moléculas complementarias para inactivar el ARNm con el fin disminuir la expresión del ácido nucleico del receptor. Por consiguiente, estas moléculas pueden tratar un trastorno caracterizado por una expresión anómala o no deseada del ácido nucleico del receptor. Esta técnica implica la escisión mediante ribozimas que contienen secuencias de nucleótidos complementarias con una o más regiones del ARNm, que atenúan la capacidad del ARNm para ser traducido. Las regiones posibles incluyen regiones codificantes y especialmente regiones codificantes correspondientes con la actividad catalítica y con otras actividades funcionales de la proteína receptora, tales como la unión al ligando.

35 Los polinucleótidos del receptor también proporcionan vectores para terapia génica en pacientes que contienen células anómalas para la expresión del gen del receptor. De este modo, las células recombinantes, que incluyen las células del paciente que han sido manipuladas genéticamente *ex vivo* y devueltas al paciente, se introducen en un individuo en el que las células producen la proteína receptora deseada para tratar al individuo.

45 La invención también abarca kits para la detección de la presencia de un ácido nucleico del receptor en una muestra biológica. Por ejemplo, el kit puede comprender reactivos tales como un ácido nucleico o agente marcado o susceptible de marcaje capaz de detectar el ácido nucleico del receptor en una muestra biológica, medios para determinar la cantidad de ácido nucleico del receptor en la muestra, y medios para comparar la cantidad del ácido nucleico del receptor en la muestra con un patrón. El compuesto o agente puede envasarse en un recipiente adecuado. El kit además puede comprender instrucciones de uso del kit para detectar el ARNm o el ADN del receptor.

50 La descripción anterior también se aplica a la secuencia mostrada en la ID SEC N° 4.

Vectores/células huésped

55 La invención también proporciona vectores que contienen los polinucleótidos del receptor. El término "vector" se refiere a un vehículo, preferiblemente una molécula de ácido nucleico, que puede transportar los polinucleótidos del receptor. Cuando el vector es una molécula de ácido nucleico, los polinucleótidos del receptor se unen covalentemente al ácido nucleico del vector. En este aspecto de la invención, el vector incluye un plásmido, un fago de hebra sencilla o doble, un vector vírico de ARN o ADN de hebra sencilla o doble o un cromosoma artificial, tal como un BAC, PAC, YAC o MAC.

60

Un vector puede mantenerse en la célula huésped como un elemento extracromosómico, donde se replica y producen copias adicionales de los polinucleótidos del receptor. Alternativamente, el vector puede integrarse en el genoma de la célula huésped y producir copias adicionales de los polinucleótidos del receptor cuando la célula huésped se replica.

65

La invención proporciona vectores para el mantenimiento (vectores de clonación) o vectores para la expresión (vectores de expresión) de los polinucleótidos del receptor. Los vectores puede trabajar en células procariotas o eucariotas o en ambas (vectores lanzadera).

ES 2 275 358 T3

Los vectores de expresión contienen regiones reguladoras que actúan en *cis* que se unen en el vector de forma operativa con los polinucleótidos del receptor de modo que se permita la transcripción de los polinucleótidos en la célula huésped. Los polinucleótidos pueden introducirse en la célula huésped con un polinucleótido separado capaz de afectar a la transcripción. De este modo, el segundo polinucleótido puede proporcionar un factor que actúa en *trans* que interacciona con la región control reguladora en *cis* para permitir la transcripción de los polinucleótidos del receptor a partir del vector. Alternativamente, la célula huésped puede suministrar un factor que actúa en *trans*. Finalmente, el vector mismo puede producir un factor que actúa en *trans*.

Se entiende, sin embargo, que en algunas realizaciones, puede darse la transcripción y/o traducción de los polinucleótidos del receptor en un sistema sin células.

La secuencia reguladora a la que los polinucleótidos descritos en este documento pueden unirse de forma operativa incluye promotores para dirigir la transcripción del ARNm. Estos incluyen, pero sin limitaciones, el promotor izquierdo del bacteriófago λ , los promotores lac, TRP y TAC de *E. coli*, los promotores temprano y tardío de SV40, el promotor inmediato temprano del CMV, los promotores temprano y tardío del adenovirus y las repeticiones largas terminales del retrovirus.

Además de las regiones de control que promueven la transcripción, los vectores de expresión también pueden incluir regiones que modulen la transcripción, tales como sitios de unión del represor y potenciadores. Los ejemplos incluyen el potenciador del SV40, el potenciador inmediato temprano de citomegalovirus, el potenciador del polioma, los potenciadores de adenovirus y los potenciadores de la RLT del retrovirus.

Además de contener sitios para el inicio y control de la transcripción, los vectores de expresión también pueden contener secuencias necesarias para finalizar la transcripción, y en la región transcrita, un sitio de unión al ribosoma para la traducción. Otros elementos reguladores de control de la expresión incluyen codones de inicio y parada así como señales de poliadenilación. El experto en la materia conocerá las numerosas secuencias reguladoras que son útiles en los vectores de expresión. Estas secuencias reguladoras se describen, por ejemplo, en Sambrook y col., *A Laboratory Manual*. 2^a ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).

Pueden usarse diversos vectores de expresión para expresar un polinucleótido del receptor. Estos vectores incluyen vectores cromosómicos, episomales y derivados de virus, por ejemplo vectores derivados de plásmidos bacterianos, de bacteriófagos, de episomas de levaduras, de elementos cromosómicos de levaduras, incluyendo cromosomas artificiales de levaduras, de virus tales como baculovirus, papovavirus, tales como SV40, virus vaccinia, adenovirus, poxvirus, virus de pseudorrabias y retrovirus. Los vectores también pueden derivar de combinaciones de estas fuentes, tales como los que derivan de elementos genéticos plasmídicos y bacteriófagos, por ejemplo, cósmidos y fagémidos. En Sambrook y col., *Molecular Cloning, A Laboratory Manual*. 2^a ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)), se describen vectores de clonación y de expresión apropiados para huésped procariontas y eucariotas.

La secuencia reguladora puede proporcionar una expresión constitutiva en una o más células huésped (es decir, específica de tejido) o puede proporcionar una expresión inducible en uno o más tipos de células, tal como mediante temperatura, adición de un nutriente, o un factor exógeno, tal como una hormona u otro ligando. Los expertos en la materia conocen bien diversos vectores que proporcionan expresión constitutiva e inducible en huésped procariontas y eucariotas.

Los polinucleótidos del receptor pueden insertarse en el ácido nucleico del vector mediante metodología bien conocida. Generalmente, la secuencia de ADN que finalmente se expresará se une a un vector de expresión escindiendo la secuencia de ADN y el vector de expresión con una o más enzimas de restricción y, a continuación, ligando los fragmentos entre sí. Los expertos en la materia conocen bien los procedimientos de digestión con enzimas de restricción y de ligamiento.

El vector que contiene el polinucleótido apropiado puede introducirse en una célula huésped apropiada para su propagación o expresión usando técnicas bien conocidas. Las células bacterianas incluyen, pero sin limitaciones, *E. coli*, *Streptomyces* y *Salmonella typhimurium*. Las células eucariotas incluyen, pero sin limitaciones, levaduras, células de insectos tales como *Drosophila*, células animales tales como células COS y CHO y células vegetales.

Como se describe en este documento, puede ser deseable expresar el polipéptido como una proteína de fusión. Por consiguiente, la invención proporciona vectores de fusión que permiten la producción de los polipéptidos del receptor. Los vectores de fusión pueden aumentar la expresión de una proteína recombinante, aumentar la solubilidad de la proteína recombinante y ayudar a la purificación de la proteína actuando, por ejemplo, como ligando para purificación por afinidad. Puede introducirse un sitio de escisión proteolítica en la unión del resto de fusión de modo que el polipéptido deseado pueda separarse finalmente del resto de fusión. Las enzimas proteolíticas incluyen, pero sin limitaciones, el factor Xa, la trombina y la enterocinasa. Los vectores de expresión de fusiones típicos incluyen pGEX (Smith y col., *Gene* 67:31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Pharmacia, Piscataway, NJ) que fusionan la glutatión S-transferasa (GST), la proteína de unión a maltosa E o la proteína A, respectivamente, con la proteína recombinante objetivo. Los ejemplos de vectores de expresión de *E. coli* sin fusión inducibles adecuados incluyen pTrc (Amann y col., *Gene* 69:301-315 (1988)) y pET 11d (Studier y col., *Gene Expression Technology: Methods in Enzymology* 185:60-89 (1990)).

ES 2 275 358 T3

La expresión de la proteína recombinante puede maximizarse en una bacteria huésped proporcionando un fondo genético en el que la célula huésped tenga la capacidad para escindir proteolíticamente la proteína recombinante alterada (Gottesman, S., *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, California (1990) 119-128). Alternativamente, la secuencia del polinucleótido de interés puede alterarse para proporcionar un uso de codones preferencial de una célula huésped específica, por ejemplo *E. coli* (Wada y col., *Nucleic Acids Res.* 20:2111-2118 (1992)).

Los polinucleótidos del receptor también pueden expresarse mediante vectores de expresión que son funcionales en levaduras. Los ejemplos de vectores para expresión en levaduras por ejemplo, *S. cerevisiae* incluyen pYepSecl (Baldari y col., *EMBO J.* 6:229-234 (1987)), pMFa (Kurjan y col., *Cell* 30:933-943 (1982)), pJRY88 (Schultz y col., *Gene* 54:113-123 (1987)), y pYES2 (Invitrogen Corporation, San Diego, CA).

Los polinucleótidos del receptor también pueden expresarse en células de insecto usando, por ejemplo, vectores de expresión de baculovirus. Los vectores de baculovirus disponibles para la expresión de proteínas en células de insecto cultivadas (por ejemplo, células Sf 9) incluyen la serie pAc (Smith y col., *Mol. Cell Biol.* 3:2156-2165 (1983)) y la serie pVL (Lucklow y col., *Virology* 170:31-39 (1989)).

En ciertas realizaciones de la invención, los polinucleótidos descritos en este documento se expresan en células de mamífero usando vectores de expresión de mamíferos. Los ejemplos de vectores de expresión de mamíferos incluyen pCDM8 (Seed, B., *Nature* 329:840 (1987)) y pMT2PC (Kaufman y col., *EMBO J.* 6:187-195 (1987)).

Los vectores de expresión enumerados en este documento se proporcionan a modo de ejemplo sólo de los vectores bien conocidos disponibles para los expertos en la materia que podrían ser útiles para expresar los polinucleótidos del receptor. El experto en la materia conocerá otros vectores adecuados para mantener la propagación o la expresión de los polinucleótidos descritos en este documento. Estos se encuentran por ejemplo en Sambrook, J., Fritsh, E. F. y Maniatis, T., *Molecular Cloning: A Laboratory Manual. 2ª ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

La invención también abarca a vectores en los que las secuencias de ácido nucleico descritas en este documento se han clonado en el vector en orientación inversa, pero unidas de forma operativa a una secuencia reguladora que permite la transcripción del ARN complementario. De este modo, puede producirse una transcripción complementaria de todas, o una parte, de las secuencias polinucleotídicas descritas en este documento, incluyen tanto regiones codificantes como no codificantes. La expresión de este ARN complementario está sometida a cada uno de los parámetros descritos anteriormente en relación con la expresión del ARN sentido (secuencias reguladoras, expresión constitutiva o inducible, expresión específica de tejido).

La invención también se refiere a células huésped recombinantes que contienen los vectores descritos en este documento. Por tanto, las células huésped incluyen células procariotas, células eucariotas inferiores, tales como levaduras, otras células eucariotas, tales como células de insecto y células eucariotas superiores, tales como células de mamíferos.

Las células huésped recombinantes se preparan introduciendo las construcciones del vector descritas en este documento en las células mediante técnicas fácilmente disponibles para el experto en la materia. Estas incluyen, pero sin limitación, transfección con fosfato cálcico, transfección mediada por DEAE-dextrano, transfección mediada por lípidos catiónicos, electroporación, transducción, infección, lipofección y otras técnicas tales como las que se encuentran en Sambrook y col. (*Molecular Cloning: A Laboratory Manual. 2ª ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Las células huésped pueden contener más de un vector. De este modo, pueden introducirse secuencias nucleotídicas diferentes en vectores diferentes de la misma célula. De forma similar, los polinucleótidos del receptor pueden introducirse solos o con otros polinucleótidos que no están relacionados con los polinucleótidos del receptor, tal como los que proporcionar factores de activación en *trans* para vectores de expresión. Cuando se introducen en una célula más de un vector, los vectores pueden introducirse individualmente, conjuntamente o unidos al vector del polinucleótido del receptor.

En el caso de vectores bacteriófagos y víricos, estos pueden introducirse en las células como virus empaquetados o encapsulados mediante procedimientos convencionales para la infección y la transducción. Los vectores víricos pueden ser competentes para la replicación o con replicación defectuosa. En el caso en que la replicación vírica sea defectuosa, la replicación tendrá lugar en las células huésped que proporcionan las funciones que complementan a las deficiencias.

Los vectores generalmente incluyen marcadores seleccionables que permiten la selección de la subpoblación de células que contiene las construcciones de vectores recombinantes. El marcador puede estar contenido en el mismo vector que contiene los polinucleótidos descritos en este documento o puede estar en un vector independiente. Los marcadores incluyen genes de resistencia a tetraciclina o a ampicilina para células huésped procariotas y dihidrofolato reductasa o resistencia a neomicina para células huésped eucariotas. Sin embargo, será eficaz cualquier marcador que proporcione selección de un rasgo fenotípico.

Mientras que las proteínas maduras pueden producirse en bacterias, levaduras, células de mamíferos y otras células

bajo el control de las secuencias reguladoras apropiadas, también pueden usarse sistemas de transcripción y traducción sin células para producir estas proteínas usando ARN derivado de las construcciones de ADN descritas en este documento.

5 Cuando se desea la secreción del polipéptido, se incorporan al vector señales de secreción apropiadas. La secuencia señal puede ser endógena con respecto a los polipéptidos del receptor o heteróloga con respecto a estos polipéptidos.

10 Cuando el polipéptido no se secreta al medio, la proteína puede aislarse a partir de la célula huésped mediante procedimientos de rotura convencionales, incluyendo congelación y descongelación, sonicación, rotura mecánica, uso de agentes de lisis y similares. A continuación, el polipéptido puede recuperarse y purificarse mediante procedimientos de purificación bien conocidos que incluyen precipitación con sulfato amónico, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía en fosfo celulosa, cromatografía de interacción hidrofóbica, cromatografía de afinidad, cromatografía en hidroxilapatita, cromatografía en lectina o cromatografía líquida de alta resolución.

15 También se entiende que dependiendo de la célula huésped en la producción recombinante de los polipéptidos descritos en este documento, los polipéptidos pueden tener diversos patrones de glucosilación, dependiendo de la célula o puede no estar glucosilados cuando se producen en una bacteria. Además, los polipéptidos pueden incluir una metionina modificada inicial en algunos casos como resultado de un procedimiento mediado por el huésped.

20 La descripción anterior también se aplica a la secuencia mostrada en las ID SEC N° 3 y 4.

Usos de los vectores y de las células huésped

25 Las células huésped que expresan los polipéptidos descritos en este documento, y especialmente las células huésped recombinantes, tienen diversos usos. En primer lugar, las células son útiles para producir proteínas o polipéptidos del receptor que pueden purificarse adicionalmente para producir cantidades deseadas de la proteína receptora o de fragmentos. De este modo, las células huésped que contienen vectores de expresión son útiles para la producción del polipéptido.

30 Las células huésped además son útiles para realizar ensayos basados en células que impliquen al receptor o a fragmentos del receptor. De este modo, una célula huésped recombinante que expresa un receptor nativo es útil para ensayar compuestos que estimulan o inhiben la función del receptor. Esto incluye la unión al ligando, la expresión génica a nivel de la transcripción o traducción, la interacción con la proteína G y los componentes de la ruta de transducción de la señal.

35 Las células huésped también son útiles para identificar mutantes del receptor en los que estas funciones están afectadas. Si los mutantes son naturales y dan origen a una patología, las células huésped que contienen las mutaciones son útiles para ensayar compuestos que tienen un efecto deseado sobre el receptor mutante (por ejemplo, función de estimulación o inhibición) que no puede estar indicado mediante su efecto en el receptor nativo.

40 Las células huésped recombinantes también son útiles para la expresión de los polipéptidos quiméricos descritos en este documento para evaluar compuestos que activan o suprimen la activación mediante un dominio extracelular amino terminal (u otra región de unión). Alternativamente, puede usarse una región heteróloga que abarca el dominio transmembranal completo (o partes del mismo) para evaluar el efecto de un dominio extracelular amino terminal deseado (u otra región de unión) sobre cualquier célula huésped determinada. En esta realización, se usa una región que abarca el dominio transmembranal completo (o partes del mismo) compatible con la célula huésped específica para hacer el vector quimérico. Alternativamente, puede introducirse en la célula huésped un dominio intracelular carboxilo terminal heterólogo, por ejemplo, una señal de transducción.

45 Adicionalmente, pueden diseñarse receptores mutantes que se han manipulado genéticamente para aumentar o disminuir una o más de sus diversas funciones (por ejemplo, unión al ligando o unión a la proteína G) y usarse para aumentar o sustituir las proteínas receptoras en un individuo. De este modo, las células huésped pueden proporcionar un beneficio terapéutico sustituyendo un receptor anómalo o proporcionando un receptor anómalo que proporcione un resultado terapéutico. En una realización, las células proporcionan receptores que son anormalmente activos.

50 En otra realización, las células proporcionan receptores que son anormalmente inactivos. Estos receptores pueden competir en el individuo con receptores endógenos.

55 En otra realización, se introducen en un individuo células que expresan receptores que no pueden activarse, para competir por el ligando con receptores endógenos. Por ejemplo, en el caso en el que el ligando excesivo es parte de una modalidad de tratamiento, puede ser necesario inactivar este ligando en un punto específico del tratamiento. Podría ser beneficioso proporcionar células que compitan por el ligando, pero que no puedan verse afectadas por la activación del receptor.

60 También pueden producirse células huésped recombinantes de forma homóloga que permitan la alteración *in situ* de las secuencias polinucleotídicas endógenas del receptor en el genoma en una célula huésped. Esta tecnología se describe más en profundidad en los documentos WO 93/09222, WO 91/12650 y U.S. 5.641.670. Brevemente, se per-

mite que las secuencias polinucleotídicas específicas que se corresponden con los polinucleótidos de receptor o con secuencias próximas o distales a un gen del receptor se integren en el genoma de una célula huésped mediante recombinación homóloga donde la expresión del gen puede verse afectada. En una realización, se introducen secuencias reguladoras que aumentan o disminuyen la expresión de una secuencia endógena. Por consiguiente, puede producirse una proteína receptora en una célula que normalmente no la produce, o la expresión aumentada de la proteína receptora puede dar lugar a una célula que normalmente produce la proteína a un nivel específico. Alternativamente, puede delecionarse el gen completo. Aún más, pueden introducirse mutaciones específicas en cualquier región deseada del gen para producir proteínas receptoras mutantes. Estas mutaciones pueden introducirse, por ejemplo, en las regiones funcionales específicas, tales como el sitio de unión al ligando o el sitio de unión a la proteína G.

En una realización, la célula huésped puede ser un ovocito fertilizado o una célula troncal embrionaria que puede usarse para producir un animal transgénico que contiene el gen del receptor alterado. Alternativamente, la célula huésped puede ser una célula troncal u otro precursor temprano de tejido que origine una subpoblación específica de células y pueden usarse para producir tejidos transgénicos en un animal. Para una descripción de vectores de recombinación homóloga, véase también Thomas y col., *Cell* 51:503 (1987). El vector se introduce en una línea celular troncal embrionaria (por ejemplo, mediante electroporación) y se seleccionan células en las que el gen introducido se recombinaba de forma homóloga con el gen del receptor endógeno (véase, por ejemplo, Li, E. y col. *Cell* 69:915 (1992)). A continuación, las células seleccionadas se inyectan en un blastocisto de un animal (por ejemplo, un ratón) para formar quimeras de agregación (véase por ejemplo, Bradley, A. en *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, editores (IRL, Oxford, 1987) pág. 113-152). A continuación, puede implantarse un embrión quimérico en una hembra adoptiva pseudopreñada adecuada y el embrión se lleva a término. La progenie portadora del ADN recombinado de forma homóloga en sus células germinales puede usarse para criar animales en los que todas las células del animal contienen el ADN recombinado de forma homóloga por la transmisión a la línea germinal del transgen. Los procedimientos para construir vectores de recombinación homóloga y animales recombinantes homólogos se describen adicionalmente en Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 y en las publicaciones internacionales PCT N° WO 90/11354; WO 91/01140 y WO 93/04169.

Las células huésped manipuladas genéticamente pueden usarse para producir animales transgénicos no humanos. Preferiblemente, el animal transgénico es un mamífero, por ejemplo, un roedor, tal como una rata o un ratón, en el que una o más de las células del animal incluye un transgen. Un transgén es un ADN exógeno que se integra en el genoma de una célula a partir de la cual se desarrolla un animal transgénico y que permanece en el genoma del animal maduro en uno o más tipos de células o tejidos del animal transgénico. Estos animales son útiles para estudiar la función de una proteína receptora y para identificar y evaluar moduladores de la actividad de la proteína receptora.

Otros ejemplos de animales transgénicos incluyen primates no humanos, ovejas, perros, vacas, cabras, pollos y anfibios.

En una realización, una célula huésped es un ovocito fertilizado o una célula troncal embrionaria en la cual se han introducido las secuencias polinucleotídicas del receptor.

Un animal transgénico puede producirse introduciendo un ácido nucleico en el pronúcleo masculino de un ovocito fertilizado, por ejemplo, mediante microinyección, infección retroviral y permitiendo que el ovocito se desarrolle en una hembra adoptiva pseudopreñada. Cualquiera de las secuencias de nucleótidos del receptor puede introducirse como un transgen en el genoma de un animal no humano, tal como un ratón.

Cualquiera de las secuencias reguladoras u otras secuencias útiles en los vectores de expresión pueden formar parte de la secuencia transgénica. Esto incluye secuencias intrónicas y señales de poliadenilación, si todavía no se han incluido. Pueden unirse de forma operativa al transgen secuencia(s) reguladora(s) específica(s) de tejido para la expresión directa de la proteína receptora en células en particular.

Los procedimientos para la generación de animales transgénicos por medio de manipulación y microinyección del embrión, especialmente en animales como ratones, se han convertido en convencionales en la técnica y se describen, por ejemplo, en las Patentes de EE.UU. N° 4.736.866 y 4.870.009, ambas de Leder y col., Patente de EE.UU. N° 4.873.191 de Wagner y col. y en Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Procedimientos similares se usan para la producción de otros animales transgénicos. Un animal transgénico fundador puede identificarse en base a la presencia del transgen en su genoma y/o a la expresión del ARNm transgénico en los tejidos o células de los animales. A continuación, puede usarse un animal fundador transgénico para criar animales adicionales portadores del transgen. Además, los animales transgénicos portadores de un transgén puede adicionalmente cruzarse con otros animales transgénicos portadores de otros transgenes. Un animal transgénico también incluye animales en los que el animal completo o tejidos del animal se han producido usando las células huésped recombinantes de forma homóloga descritas en este documento.

En otra realización, pueden producirse animales transgénicos no humanos que contienen sistemas seleccionados que permiten la expresión regulada del transgen. Un ejemplo de este sistema es el sistema de la recombinasa *cre/loxP* del bacteriófago P1. Para una descripción del sistema de la recombinasa *cre/loxP*, véase, por ejemplo, Lakso y col. *PNAS* 89:6232-6236 (1992). Otro ejemplo de un sistema recombinasa es el sistema de la recombinasa FLP de *S. cerevisiae* (O'Gorman y col. *Science* 251:1351-1355 (1991). Si se usa el sistema de la recombinasa *cre/loxP* para regular la expresión del transgen, se requieren animales que contenga transgenes que codifiquen tanto la recombinasa

ES 2 275 358 T3

Cre como una proteína seleccionada. Estos animales pueden proporcionarse mediante la construcción de animales transgénicos “dobles”, por ejemplo, apareando dos animales transgénicos, uno que contiene un transgen que codifica una proteína seleccionada y el otro que contiene un transgen que codifica una recombinasa.

5 También pueden producirse clones de los animales transgénicos no humanos descritos en este documento, según los procedimientos descritos en Wilmut, I. y col. *Nature* 385:810-813 (1997) y las publicaciones internacionales PCT International N° WO 97/07668 y WO 97/07669. Brevemente, puede aislarse una célula, por ejemplo, una célula somática, a partir de un animal transgénico e inducir su salida del ciclo de crecimiento y su entrada en la fase G₀. A continuación, las células en reposo se fusionan, por ejemplo, a través del uso de pulsos eléctricos, con un ovocito enucleado de un animal de la misma especie del que se ha aislado la célula en reposo. El ovocito reconstruido se cultiva a continuación de modo que se desarrolle a mórula o a blastocisto y se transfiere luego a una hembra adoptiva pseudopreñada. La descendencia nacida de esta hembra adoptiva será un clon del animal del cual se aísla la célula, por ejemplo, la célula somática.

15 Los animales transgénicos que contienen células recombinantes que expresan los polipéptidos descritos en este documento son útiles para realizar los ensayos descritos en este documento en un contexto *in vivo*. Por consiguiente, los diversos factores fisiológicos que se presentan *in vivo* y que pueden llevar a cabo la unión al ligando, la activación del receptor y la transducción de la señal, pueden no ser evidentes a partir de los ensayos *in vitro* sin células o basados en células. Por consiguiente, es útil proporcionar animales transgénicos no humanos para el ensayo *in vivo* de la función del receptor, incluyendo la interacción con el ligando, el efecto de receptores mutantes específicos en la función del receptor y la interacción con el ligando, y el efecto de los receptores quiméricos. También es posible evaluar el efecto de mutaciones completas, que son mutaciones que eliminan sustancial o completamente una o más funciones del receptor.

20 La descripción anterior también se aplica a la secuencia mostrada en las ID SEC N° 3 y 4.

25 *Composiciones farmacéuticas*

Las moléculas de ácido nucleico del receptor, la proteína (especialmente fragmentos tales como el dominio extracelular amino terminal), los moduladores de la proteína y anticuerpos (también denominados en este documento como “compuestos activos”) pueden incorporarse a composiciones farmacéuticas adecuadas para su administración a un sujeto, por ejemplo, un humano. Estas composiciones típicamente comprenden la molécula de ácido nucleico, proteína, modulador o anticuerpo y un vehículo farmacéuticamente aceptable.

35 Según se usa en este documento, la expresión “vehículo farmacéuticamente aceptable” pretende incluir cualquiera y todos los disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardadores de la absorción y similares, compatibles con la administración farmacéutica. El uso de estos medios y agentes para sustancias farmacéuticamente activas es bien conocido en la técnica. Excepto en la medida que cualquier medio o agente sea incompatible con el compuesto activo, estos pueden usarse en las composiciones de la invención. También pueden incorporarse a las composiciones compuestos activos suplementarios. Una composición farmacéutica de la invención se formula para que sea compatible con la vía de administración que se pretende. Los ejemplos de vías de administración incluyen la vía parenteral, por ejemplo, administración intravenosa, intradérmica, subcutánea, oral (por ejemplo, inhalación), transdérmica (tópica), transmucosa y rectal. Las soluciones o suspensiones usadas para la aplicación parental, intradérmica o subcutánea pueden incluir los siguientes componentes: un disolvente estéril, tal como agua para inyección, solución salina, aceites fijados, polietilenglicoles y otros disolventes sintéticos; agentes antibacterianos, tales como alcohol bencílico o metilparabenos; antioxidantes, tales como ácido ascórbico o bisulfito sódico; agentes quelantes, tales como ácido etilendiaminotetraacético; tampones, tales como acetatos, citratos o fosfatos, y agentes para ajustar la tonicidad, tal como cloruro sódico o dextrosa. El pH puede ajustarse con ácidos o bases, tales como ácido clorhídrico o hidróxido sódico. La preparación parenteral puede incluirse en ampollas, jeringas desechables o viales de dosis múltiple hechos de vidrio o de plástico.

50 Las composiciones farmacéuticas adecuadas para uso inyectable incluyen soluciones acuosas estériles (cuando son hidrosolubles) o dispersiones y polvos estériles para la preparación extemporánea de soluciones o dispersiones estériles inyectables. Para la administración intravenosa, los vehículos adecuados incluyen solución salina fisiológica, agua bacteriostática, Cremophor ELJ (BASF, Parsippany, NJ) o solución salina tamponada con fosfato (PBS). En todos los casos, la composición debería estar estéril o podría ser fluida en grado tal que pueda cargarse en una jeringa fácilmente. Deberá ser estable en las condiciones de fabricación y conservación y deberá conservarse de la acción contaminante de microorganismos, tales como bacterias y hongos. El vehículo puede ser un medio disolvente o de dispersión que contengan, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares) y mezclas adecuadas de los mismos. La fluidez adecuada puede mantenerse, por ejemplo, usando un recubrimiento tal como la lecitina, manteniendo el tamaño de la partícula requerido en el caso de la dispersión y usando tensioactivos. La prevención de la acción de microorganismos puede conseguirse mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorbutanol, fenol, ácido ascórbico, timerosal y similares. En muchos casos, se preferirá incluir agentes isotónicos en la composición, por ejemplo, azúcares, polialcoholes tal como manitol, sorbitol o cloruro sódico. La absorción prolongada de las composiciones inyectables puede realizarse incluyendo en la composición un agente que retrase la absorción, por ejemplo, monoestearato de aluminio y gelatina.

65 Las soluciones inyectables estériles pueden prepararse incorporando el compuesto activo (por ejemplo, una proteína receptora o un anticuerpo anti-receptor) en la cantidad necesaria en un disolvente apropiado con uno o una

ES 2 275 358 T3

combinación de los ingredientes enumerados anteriormente, según se necesite, seguido de la esterilización por filtración. Generalmente, las dispersiones se preparan incorporando el compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los demás ingredientes necesarios de los enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los procedimientos preferidos de preparación son la desecación al vacío y la liofilización que produce un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de una solución previamente esterilizada por filtración de la misma.

Las composiciones por vía oral, generalmente incluyen un disolvente inerte o un vehículo comestible. Pueden estar incluidas en cápsulas de gelatina o comprimirse en comprimidos. Para la administración por vía oral, el agente puede estar contenido en formas entéricas para sobrevivir al estómago o estar recubierto o mezclado adicionalmente para liberarse en una región del tracto GI en particular, mediante procedimientos conocidos. Para el fin de la administración terapéutica por vía oral, el compuesto activo puede incorporarse con excipientes y usarse en forma de comprimidos, pastillas o cápsulas. Las composiciones por vía oral también pueden prepararse usando un vehículo fluido para su uso como un colutorio, en el que el compuesto en el vehículo fluido se aplica por vía oral, se agita y escupe o traga. Como parte de la composición, también pueden incluirse agentes de unión farmacéuticamente compatibles, y/o materiales adyuvantes. Los comprimidos, píldoras, cápsulas, pastillas y similares, pueden contener cualquiera de los siguientes ingredientes o compuestos de naturaleza similar: un aglutinante, tal como celulosa microcristalina, goma de tragacanto o gelatina; un excipiente, tal como almidón o lactosa; un agente de desintegración, tal como ácido algínico, Primogel o almidón de maíz; un lubricante, tal como estearato de magnesio o Sterotes; un deslizante, tal como dióxido de silicio coloidal; un agente edulcorante, tal como sacarosa o sacarina o un agente aromatizante, tal como menta piperita, salicilato de metilo o aroma de naranja.

Para la administración mediante inhalación, los compuestos se administran en forma de un pulverizado aerosol a partir de un recipiente o distribuidor presurizado que contiene un propulsor adecuado, por ejemplo, un gas, tal como dióxido de carbono, o un nebulizador.

La administración sistémica también puede ser por medios transmucosos o transdérmicos. Para la administración transmucosa o transdérmica, se usan en la formulación agentes penetrantes apropiados para la barrera que se quiere penetrar. Generalmente, estos agentes penetrantes son conocidos en la técnica e incluyen, por ejemplo, para la administración transmucosa, detergentes, sales biliares y derivados del ácido fusídico. La administración transmucosa puede lograrse mediante el uso de pulverizadores nasales o de supositorios. Para la administración transdérmica, los compuestos activos se formulan en pomadas, pomadas balsámicas, geles o cremas, generalmente conocidos en la técnica.

Los compuestos también pueden prepararse en forma de supositorios (por ejemplo, con bases para supositorio convencionales, tales como manteca de cacao y otros glicéridos) o enemas de retención para administración por vía rectal.

En una realización, los compuestos activos se preparan con vehículos que protegerán al compuesto de una rápida eliminación del organismo, tal como una formulación de liberación controlada, incluyendo sistemas de implantes y de administración microencapsulada. Pueden usarse polímeros biocompatibles y biodegradables, tales como acetato de etilvinilo, polianhídridos, ácido poliglicólico, colágeno, poliolefinas y ácido poliláctico. Los procedimientos para la preparación de estas formulaciones serán aparentes para los expertos en la materia. Los materiales también pueden obtenerse comercialmente de Alza Corporation y Nova Pharmaceuticals, Inc.. También pueden usarse como vehículos farmacéuticamente aceptables suspensiones de liposomas (incluyendo liposomas dirigidos hacia células infectadas con anticuerpos monoclonales frente a antígenos víricos). Estos pueden prepararse según procedimientos conocidos por los expertos en la materia, por ejemplo, como se describe en la Patente de EE. UU. N° 4.522.811.

Es especialmente ventajoso formular composiciones por vía oral o parenteral en forma de unidades de dosis para una administración fácil y una dosificación uniforme. La forma de unidad de dosis según se usa en este documento se refiere a unidades físicamente discretas adaptadas como dosificaciones unitarias para el sujeto que se va a tratar, conteniendo cada unidad una cantidad predeterminada de compuesto activo calculado para que produzca el efecto terapéutico deseado en asociación con el vehículo farmacéutico necesario. La especificación para las formas de unidad de dosis de la invención está dictada y depende directamente de las características exclusivas del compuesto activo y del efecto terapéutico en particular que se desea lograr, y las limitaciones inherentes en la técnica de composición de este compuesto activo para el tratamiento de individuos.

Las moléculas de ácido nucleico de la invención pueden insertarse en vectores y usarse como vectores para terapia génica. Los vectores de terapia génica pueden administrarse a un sujeto mediante, por ejemplo, inyección intravenosa, administración local (documento U.S. 5.328.470), o mediante inyección estereotáctica (véase, por ejemplo, Chen y col., *PNAS* 91:3054-3057 (1994)). La preparación farmacéutica del vector de terapia génica puede incluir el vector de terapia génica en un disolvente aceptable, o puede comprender una matriz de liberación lenta en la cual está embebido el vehículo de administración del gen. Alternativamente, cuando el vector de administración del gen completo pueda producirse intacto a partir de las células recombinantes, por ejemplo, vectores retrovirales, la preparación farmacéutica puede incluir una o más células que producen el sistema de liberación del gen.

Las composiciones farmacéuticas pueden incluirse en un recipiente, paquete o dispensador junto con las instrucciones para su administración.

La descripción anterior también se aplica a las secuencias mostradas en las ID SEC N° 3 y 4.

ES 2 275 358 T3

REIVINDICACIONES

5 1. Un procedimiento para identificar un agente que modula el nivel o la actividad de un polipéptido en una célula o que interacciona con un polipéptido en una célula, en el que dicho polipéptido se selecciona entre:

(a) la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3.

10 (b) la secuencia de aminoácidos de una variante alélica de la secuencia de aminoácidos mostrada en las ID SEC N° 1 ó 3.

(c) la secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en la que la variante de secuencia es codificada por una molécula de ácido nucleico que hibrida con la molécula de ácido nucleico mostrada en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;

15 (d) un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos.

20 (e) la secuencia de aminoácidos del polipéptido del receptor maduro de aproximadamente el aminoácido 6 a aproximadamente el aminoácido 342, mostrada en la ID SEC N° 1 ó 3,

(f) la secuencia de aminoácidos del polipéptido mostrada en la ID SEC N° 1 ó 3, de aproximadamente el aminoácido 1 a aproximadamente el aminoácido 342; y

25 (g) la secuencia de aminoácidos de una región portadora de un epítipo de uno cualquiera de los polipéptidos de (a)-(f), y dicho procedimiento comprende:

30 (i) poner en contacto dicho agente con una célula capaz de expresar dicho polipéptido de modo que el nivel o la actividad de dicho polipéptido puede modularse en dicha célula mediante dicho agente, en el que dicha célula deriva del cerebro, células CD34⁺, células B, músculo esquelético, ganglio linfático, bazo, timo, hígado, amígdalas, colon, corazón, granulocitos, eritroblastos o páncreas; y

(ii) medir el nivel o actividad de dicho polipéptido.

35 2. El procedimiento de la reivindicación 1 en el que dicho agente modula una interacción entre dicho polipéptido y una molécula diana para dicho polipéptido, la etapa (i) se lleva a cabo en condiciones que permiten a dicho polipéptido interactuar con dicha molécula diana y la etapa (ii) comprende detectar la formación de un complejo entre dicho polipéptido y dicha molécula diana o detectar la actividad de dicho polipéptido como resultado de la interacción de dicho polipéptido con dicha molécula diana.

40 3. El procedimiento de la reivindicación 1 ó 2, en el que dicho agente aumenta el nivel o la actividad de dicho polipéptido.

45 4. El procedimiento de la reivindicación 1 ó 2, en el que dicho agente disminuye el nivel o la actividad de dicho polipéptido.

5. El procedimiento de una cualquiera de las reivindicaciones precedentes, en el que la etapa (ii) comprende:

(1) añadir un polipéptido competitivo que puede interactuar con dicho agente; y

50 (2) comparar la cantidad de interacción entre dicho agente y dicho polipéptido con la cantidad de interacción en ausencia de dicho polipéptido competitivo.

55 6. El procedimiento de una cualquiera de las reivindicaciones precedentes, en el que dicha célula de cerebro es una célula de la glía.

7. El procedimiento de una cualquiera de las reivindicaciones precedentes en el que dicha célula está *in vivo* y en el que dicha célula no está en un ser humano.

60 8. El procedimiento de la reivindicación 7 en el que dicha célula está en un animal no humano transgénico o no transgénico.

9. El procedimiento de una cualquiera de las reivindicaciones 1 a 6 en el que dicha célula está *in vitro*.

65 10. El procedimiento de la reivindicación 9 en el que dicha célula ha sido alterada.

11. El procedimiento de la reivindicación 9 en el que dicha célula está en una biopsia o es una célula natural o recombinante en cultivo celular.

ES 2 275 358 T3

12. El procedimiento de una cualquiera de las reivindicaciones precedentes en el que dicho agente se selecciona entre un péptido, un fosfopéptido, un anticuerpo, una molécula orgánica y una molécula inorgánica.

5 13. Un procedimiento para detectar la presencia de un polipéptido en una muestra, comprendiendo dicho procedimiento:

(i) poner en contacto dicha muestra con un anticuerpo que específicamente permite la detección de la presencia del polipéptido en la muestra; y

10 (ii) detectar la presencia del polipéptido, en el que dicho polipéptido se selecciona entre:

(a) la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

15 (b) la secuencia de aminoácidos de una variante alélica de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

20 (c) la secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que la variante de secuencia está codificada por una molécula de ácido nucleico que hibrida con la molécula de ácido nucleico mostrado en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;

(d) un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos.

25 (e) la secuencia de aminoácidos del polipéptido del receptor maduro de aproximadamente el aminoácido 6 a aproximadamente el aminoácido 342, mostrada en la ID SEC N° 1 ó 3;

(f) la secuencia de aminoácidos del polipéptido mostrada en la ID SEC N° 1 ó 3, de aproximadamente el aminoácido 1 a aproximadamente el aminoácido 342;

30 (g) la secuencia de aminoácidos de una región portadora de un epítipo de uno cualquiera de los polipéptidos de (a)-(f);

35 y en el que dicha muestra deriva de una célula seleccionada entre cerebro, CD34⁺, B, músculo esquelético, ganglio linfático, bazo, timo, hígado, amígdalas, colon, corazón, granulocitos, eritroblastos y células pancreáticas.

14. El procedimiento de la reivindicación 13 en el que dicha detección se hace en una célula derivada de un sujeto que tiene un trastorno que implica a dicha célula.

40 15. Un procedimiento para identificar un agente que modula el nivel o la actividad de una molécula de ácido nucleico en una célula o que interacciona con una molécula de ácido nucleico en una célula, en el que dicha molécula de ácido nucleico tiene una secuencia de ácido nucleico seleccionada entre:

(a) la secuencia de nucleótidos mostrada en la ID SEC N° 2 ó 4;

45 (b) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

(c) una secuencia de nucleótidos complementaria a cualquiera de las secuencias de nucleótidos de (a) o (b);

50 (d) una secuencia de nucleótidos que codifica una secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3 que hibrida con la secuencia de nucleótidos mostrada en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;

(e) una secuencia de nucleótidos complementaria a la secuencia de nucleótidos de (d);

55 (f) una secuencia de nucleótidos que codifica un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos; y

(g) una secuencia de nucleótidos complementaria a la secuencia de nucleótidos de (f);

60 y dicho procedimiento comprende:

(i) poner en contacto dicho agente con una célula capaz de expresar dicha molécula de ácido nucleico de modo que puede modularse el nivel o la actividad de dicha molécula de ácido nucleico en dicha célula mediante dicho agente, en el que dicha célula deriva de cerebro, células CD34⁺, células B, músculo esquelético, ganglio linfático, bazo, timo, hígado, amígdalas, colon, corazón, granulocitos, eritroblastos o páncreas y;

(ii) medir el nivel o la actividad de dicha molécula de ácido nucleico.

ES 2 275 358 T3

16. Un procedimiento para detectar la presencia de una molécula de ácido nucleico en una muestra, comprendiendo dicho procedimiento:

5 (i) poner en contacto dicha muestra con un agente que permita específicamente la detección de la presencia de la molécula de ácido nucleico en la muestra y;

(ii) detectar la presencia de la molécula de ácido nucleico o fragmento de la misma, en el que la molécula de ácido nucleico tiene una secuencia de nucleótidos seleccionada entre:

10 (a) la secuencia de nucleótidos mostrada en la ID SEC N° 2 ó 4;

(b) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3;

15 (c) una secuencia de nucleótidos complementaria a cualquiera de las secuencias de nucleótidos de (a) o (b);

(d) una secuencia de nucleótidos que codifica una secuencia de aminoácidos de una variante de secuencia de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3 que hibrida con la secuencia de nucleótidos mostrada en la ID SEC N° 2 ó 4, respectivamente, en condiciones rigurosas;

20 (e) una secuencia de nucleótidos complementaria a la secuencia de nucleótidos en (d);

(f) una secuencia de nucleótidos que codifica un fragmento de la secuencia de aminoácidos mostrada en la ID SEC N° 1 ó 3, en el que el fragmento comprende al menos 10 aminoácidos contiguos; y

25 (g) una secuencia de nucleótidos complementaria a la secuencia de nucleótidos de (f);

en el que dicha muestra deriva de un tejido seleccionado entre cerebro, células CD34⁺, células B, músculo esquelético, ganglio linfático, bazo, timo, hígado, amígdalas, colon, corazón, granulocitos, eritroblastos y páncreas, y en el que dicho agente comprende un oligonucleótido que hibrida con una cualquiera de las secuencias de ácidos nucleicos de (a)-(g) en condiciones rigurosas y la etapa (ii) comprende determinar si el oligonucleótido se une a la secuencia de ácido nucleico en la muestra.

30

17. El procedimiento de la reivindicación 16 en el que el ácido nucleico es ARNm.

35

40

45

50

55

60

65

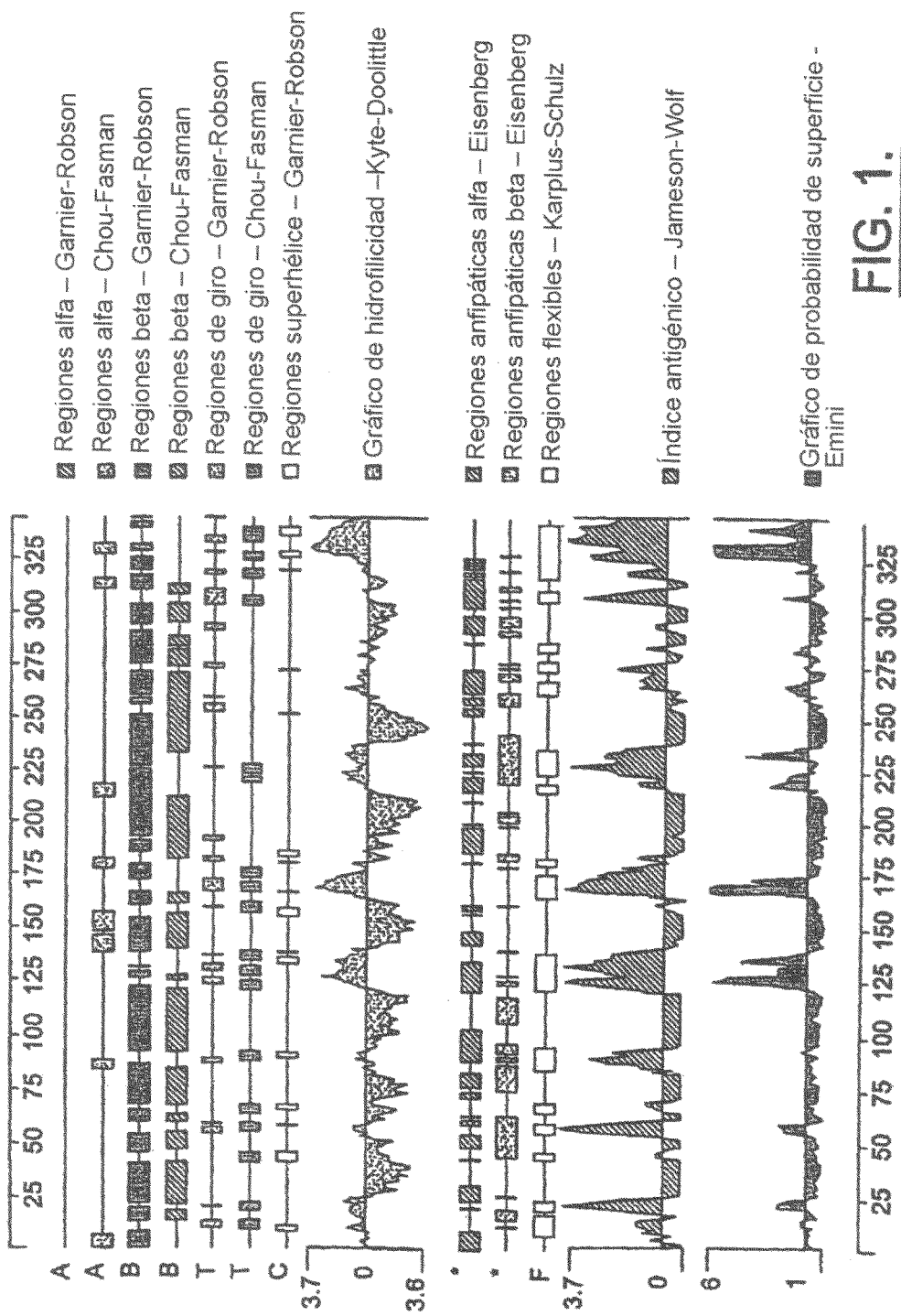


FIG. 1.

aa del orf de 15625 ==> KD hidrofobicidad <11/1>

Secuencia de la proteína 343 aa MQABDNLTSAPG...DGGDPNEETPMZ

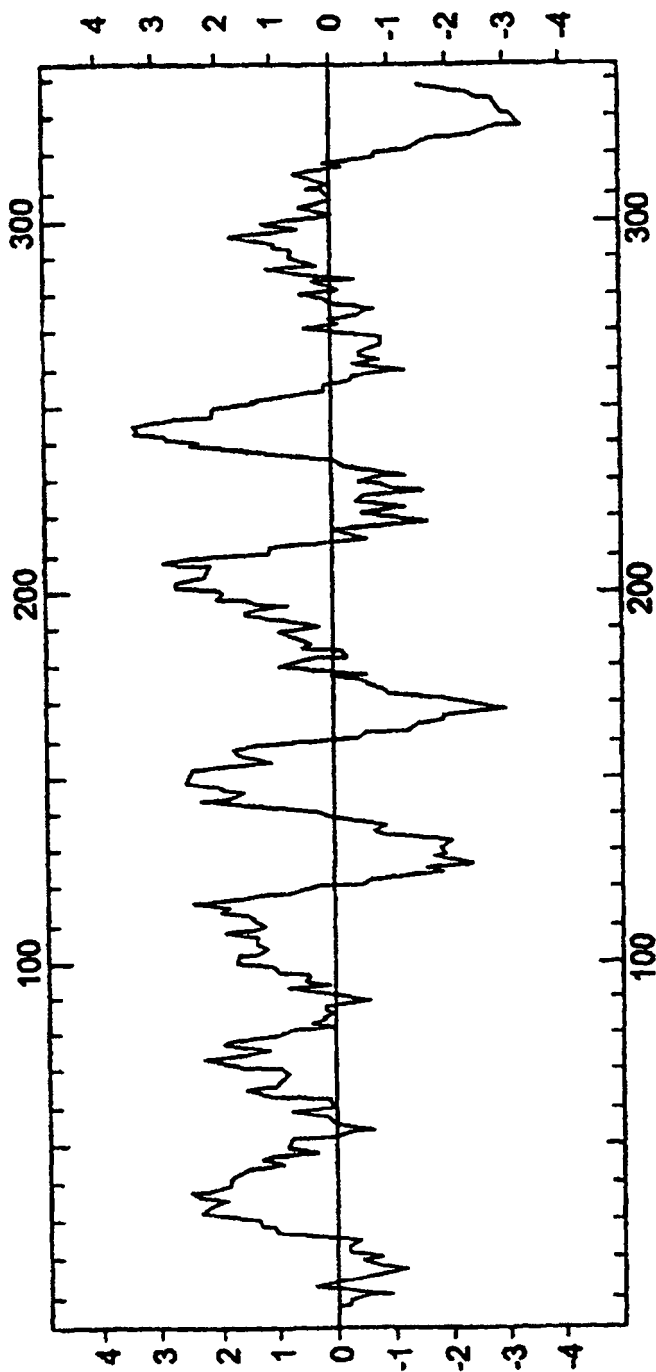


FIG. 2.

ES 2 275 358 T3

LISTA DE SECUENCIAS

<110> Glucksmann, Maria A.
Gu, Wei

5 <120> Receptor 15625, un nuevo receptor ligado a la proteína G

<130> 5800-13A-1

<140>

10 <141>

<160> 5

<170> PatentIn Ver. 2.0

<210> 1

15 <211> 342

<212> PROT

<213> *Homo sapiens*

20 <400> 1

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

25 **Cys Thr Arg Asp Tyr Lys Ile Thr Gln Val Leu Phe Pro Leu Leu Tyr**
20 25 30

30 **Thr Val Leu Phe Phe Val Gly Leu Ile Thr Asn Gly Leu Ala Met Arg**
35 40 45

Ile Phe Phe Gln Ile Arg Ser Lys Ser Asn Phe Ile Ile Phe Leu Lys
50 55 60

35 **Asn Thr Val Ile Ser Asp Leu Leu Met Ile Leu Thr Phe Pro Phe Lys**
65 70 75 80

40 **Ile Leu Ser Asp Ala Lys Leu Gly Thr Gly Pro Leu Arg Thr Phe Val**
85 90 95

Cys Gln Val Thr Ser Val Ile Phe Tyr Phe Thr Met Tyr Ile Ser Ile
100 105 110

45 **Ser Phe Leu Gly Leu Ile Thr Ile Asp Arg Tyr Gln Lys Thr Thr Arg**
115 120 125

50 **Pro Phe Lys Thr Ser Asn Pro Lys Asn Leu Leu Gly Ala Lys Ile Leu**
130 135 140

55

60

65

ES 2 275 358 T3

Ser Val Val Ile Trp Ala Phe Met Phe Leu Leu Ser Leu Pro Asn Met
 145 150 155 160
 5
 Ile Leu Thr Asn Arg Gln Pro Arg Asp Lys Asn Val Lys Lys Cys Ser
 165 170 175
 10
 Phe Leu Lys Ser Glu Phe Gly Leu Val Trp His Glu Ile Val Asn Tyr
 180 185 190
 15
 Ile Cys Gln Val Ile Phe Trp Ile Asn Phe Leu Ile Val Ile Val Cys
 195 200 205
 20
 Tyr Thr Leu Ile Thr Lys Glu Leu Tyr Arg Ser Tyr Val Arg Thr Arg
 210 215 220
 25
 Gly Val Gly Lys Val Pro Arg Lys Lys Val Asn Val Lys Val Phe Ile
 225 230 235 240
 30
 Ile Ile Ala Val Phe Phe Ile Cys Phe Val Pro Phe His Phe Ala Arg
 245 250 255
 35
 Ile Pro Tyr Thr Leu Ser Gln Thr Arg Asp Val Phe Asp Cys Thr Ala
 260 265 270
 40
 Glu Asn Thr Leu Phe Tyr Val Lys Glu Ser Thr Leu Trp Leu Thr Ser
 275 280 285
 45
 Leu Asn Ala Cys Leu Asp Pro Phe Ile Tyr Phe Phe Leu Cys Lys Ser
 290 295 300
 50
 Phe Arg Asn Ser Leu Ile Ser Met Leu Lys Cys Pro Asn Ser Ala Thr
 305 310 315 320
 55
 Ser Leu Ser Gln Asp Asn Arg Lys Lys Glu Gln Asp Gly Gly Asp Pro
 325 330 335
 60
 Asn Glu Glu Thr Pro Met
 340
 65

<210> 2

50 <211> 2.286

<212> ADN

<213> *Homo sapiens*

ES 2 275 358 T3

<400> 2

cgcccgtagc tttgagtcca gtggttgaag acaatctctg attgtgaagc cctcttttcc 60
5 tctccttcta tttctctcta gagcactcaa gactttactg acgaaaacte aggaaatcct 120
ctatcacaaa gaggtttggc aactaaacta agacattaaa aggaaaatac cagatgccac 180
tctgcaggct gcaataacta ctacttactg gatacattca aaccctccag aatcaacagt 240
tatacaggtaa ccaacaagaa atgcaagccg tgcacaacct cacctctgcy cctgggacca 300
ccagtctgtg caccagagac tacaaaatca cccaggtcct cttcccactg ctctacactg 360
10 tctgtttttt tgttggaact atcacaaatg gcctggcgat gaggattttc tttcaaatcc 420
ggagtaaatc aaactttatt atttttctta agaacacagt catttctgat cttctcatga 480
ttctgacttt tccattcaaa attcttagtg atgccaaact gggaacagga cactgagaa 540
cttttggtg tcaagttacc tccgtcatat tttatttcac aatgtatata agtatttcat 600
15 tctgggact gataactatc gatcgctacc agaagaccac caggccattt aaaacatcca 660
accccaaaaa tctcttgggg gctaagattc tctctgtgt catctgggca ttcattgctt 720
tactctctt gcctaacatg attctgacca acaggcagcc gagagacaag aatgtgaaga 780
aatgtctctt ccttaaatca gagttcggtc tagtctggca tgaatatgta aattacatct 840
20 gtcaagtcat tttctggatt aatttcttaa ttgttattgt atgttatata ctcattacaa 900
aagaagtga cgggtcatac gtaagaacga ggggtgtagg taaagtcccc aggaaaagg 960
tgaactcaa agttttcatt atcattgctg tattctttat ttgtttggt ctttccatt 1020
ttgccgaat tccttacacc ctgagccaaa cccgggatgt ctttgactgc actgctgaaa 1080
25 atactctgtt ctatgtgaaa gagagcactc tgtggttaac ttctttaat gcatgcctgg 1140
atcctctcat ctattttttc ctttgcaagt cttcagaaa ttcttgata agtattgctga 1200
agtgcccaa ttctgcaaca tctctgtccc aggacaatag gaaaaaagaa caggatggtg 1260
gtgacccaaa tgaagagact ccaatgtaaa caaattaact aaggaaatat ttcaatctct 1320
30 ttgtgttcag aactcgtaa agcaagcgc taagtaaaaa tattaactga cgaagaagca 1380
actaatgaa taataatgac tctaaagaaa cagaagatta caaaagcaat tttcatttac 1440
ctttccagta tgaagagcta tcttaaaata tagaaaacta atctaactg tagctgtatt 1500
agcagcaaaa caaacgacat ccaattgtca tgctgcatgc aaaactacac agaattcatg 1560
35 ttttgcaag ttttgccaaa atgagtaatc atataatatt tactgtaatt tttaaaatac 1620
attatcgttc acaattttat ttttccataa tcaactaagg aagaacgatc aattggatat 1680
aatttcttac caaaaatgat agttaaaatg tatatatatc ctagtcccct aaccaaatcc 1740
tgacctattg ggatacttat aaaaatttaa gtaagtggga tacacaaaga ataataacta 1800
40 ttaacttttc attattagca aaaacctaag ggatttaaac taattgaaac tgtatttgat 1860
tggacttaat tttttatggt tatttagaag ataaagattt aaagaagacc tttacaataa 1920
agagaagaaa tategaagtc attaaaataa ggagacttac ttttatgaca ttctaatact 1980
aaaaaatata gaaatatttc cttaattcta gagaaactag ttttactaat tttttacaac 2040
ttcaataata ccatcactga cacttacctt tattaattag cttctagaaa atagctgcta 2100
45 attaggttaa tgaacatttt accttagtga aaaaaattaa ttaaatatga ttacaaagtt 2160
gcacagcata actactgaga ggaaagtgat tgatctgttt gtaattactt gtttgattg 2220
gtgtgtataa aatacaaaat ttacattaaa ctctaaaaaa aaaaaaaaaa aaaaaaaaaa 2280
gggcgg 2286

50

<210> 3

<211> 342

55 <212> PROT

<213> *Macaca sp.*

60

65

ES 2 275 358 T3

<400> 3

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

ES 2 275 358 T3

Glu Asn Thr Leu Phe Tyr Val Lys Glu Ser Thr Leu Trp Leu Thr Ser
 275 280 285

5 Leu Asn Ala Cys Leu Asp Pro Phe Thr Tyr Phe Phe Leu Cys Lys Ser
 290 295 300

10 Phe Arg Asn Ser Leu Ile Ser Met Leu Lys Cys Pro Asn Ser Ala Thr
 305 310 315 320

15 Ser Gln Ser Gln Asp Asn Arg Lys Lys Glu Gln Asp Gly Gly Asp Pro
 325 330 335

Asn Glu Glu Thr Pro Met
 340

20 <210> 4
 <211> 2.272
 <212> ADN
 <213> *Macaca sp*
 25 <400> 4

30 acgcgtccgc aatctctgat tgtaaagccc tctcttcttc tccttctatt tctctataga 60
 aactcaaga ctttactgat gaaaactcag gaaattctct atcaciaaga ggtttgcaa 120
 ctaaactaag acattaaaag gaaaatacca gatgccactc tgcacgttgc aataactact 180
 atttactgga tacattcaa tcctccagaa tcaacgggta tcaggtaacc aacaagaaat 240
 gcaagccatc gacaacctca cgtctgcgcc tgggaacacc agtctgtgca ccagagacta 300
 35 caaaatcacc caggctctct tcccactgct ctacactgtc ctgttttttg ttggactcat 360
 cacaaatagc ctggcgatga ggattttctt tcaaattcgg agtaaatcaa actttattat 420
 tttcttaag aacacagtca tttccgatct tctcatgatt ctgactttc cattcaaaat 480
 tcttagtgat gccaaactgg gaacaggacc actgagaact tttgtgtgtc aagttacctc 540
 40 cgtcatattt tatttcacaa tgtatatcag tatttcattc ctgggactga taactatcga 600
 tcgctaccag aagaccacca ggccatttaa aacatccaac cccaaaaatc tcttgggggc 660
 taagattctc tctgttctca tctgggcatt catgttctta ctctctttgc ctaacatgat 720
 tctgactaac aggcggccaa gagacaagaa tgtgaagaaa tgctctttcc ttaaatacaga 780
 gttcggccta gtctggcatg aaatagtaaa ttacatctgt caagtcattt tctggattaa 840
 45 tttcttaatt gtcattgtat gttacacact cattacaaa gaactgtacc ggtcatatgt 900
 aagaacaagg ggtgtaggta aagtccccag gaaaaagggt aacgtcaaag ttttcattat 960
 cattgctgta ttctttattt gttttgttcc tttccatttt gcccgaaatc cttataacct 1020
 gagccaaacc cgggatgtct ttgactgcgc cgctgaaaat actctgttct atgtgaaaga 1080
 50 gagtactctg tggttaactt ccttaaagtc atgcctggat ccgttcacct atttttctct 1140
 ttgcaagtcc ttcagaaatt ccttgataag tatgctgaag tgcaccaatt ctgcaacatc 1200
 tcagtcccag gacaatagga aaaaagaaca ggatggtggt gacccaaag aagagactcc 1260
 aatgtaaca tattaactga ggaaatatgt caatctcttt gcgttcagaa ctcattaaag 1320
 55 caaagcgcta cgtaaaaata ttaactgacg aagaagcaac tgagttaata acaatgactc 1380
 ttaaaacatg taatagaaga tttacaaaag caattttcat ttaccttcc agtatgaaaa 1440
 gctatgttaa aatatagaaa actaatctaa cctgtagctg tatagtatca aaacaatga 1500
 catccaattg gcatgctgca tgcaaaacta cacagaatc acgttttgca gagttttgcc 1560

60

65

ES 2 275 358 T3

5 aaaatgagta atcatataat atctaccgta atgtttaaaa tacattattg ctcacgattt 1620
 tatttcttca taatcaacta aggaagaatt atcaattgga tacaatcttc ttacaaaaaa 1680
 tgacacctaa aatgtatata tatectagcc cctaaccaaa tcctgaccta ttgggatact 1740
 10 tataaaaatt tgagtaagtg ggatacacia agaataataa ctattaactt ttaattatga 1800
 gcaaaaacct aagggttaa tttaaactaa ttgaaactgt atttgattgg acttaatttt 1860
 tttgtttatt aagaagacac ttgaagaaga cttttacaat aaagagaaga aatatcaaag 1920
 tcattaaaat aaggagagt actttttatga tattctaaca ctaaacaata tagaatatt 1980
 tccttaatat tagtttctag agaaactagt ttactaatt ttttacaacc tcaataatac 2040
 catcattgac acttaccttt attaactagc ttctagaaaa tacctgctaa ttaggttaat 2100
 gaacatttta tgtagtgaa aaaaattaat taaatatgat tacaagttg cacagcataa 2160
 ctactgaaag tgattgatcc atttgtaatt atttgtttgt actggtgtgt ataaaataca 2220
 15 aaatttcat taaactctaa atcaccaaaa aaaaaaaaaa aaaaaagggc gg 2272

<210> 5
 20 <211> 269
 <212> PROT
 <213> Desconocido
 25 <220>
 <223> Descripción de organismo desconocido: receptor transmembranal de la familia de la rodopsina
 <400> 5

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

65

ES 2 275 358 T3

| | 130 | | 135 | | | 140 | | | | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 5 | Tyr | Asp | Trp | Pro | Glu | Trp | Met | Trp | Arg | Trp | Tyr | Val | Ile | Leu | Met | Thr | |
| | 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| 10 | Ile | Ile | Met | Gly | Phe | Tyr | Ile | Pro | Met | Ile | Ile | Met | Leu | Phe | Cys | Tyr | |
| | | | | 165 | | | | | | 170 | | | | | 175 | | |
| 15 | Trp | Arg | Ile | Tyr | Arg | Ile | Ala | Arg | Leu | Trp | Met | Arg | Met | Ile | Pro | Ser | |
| | | | | 180 | | | | | 185 | | | | | 190 | | | |
| 20 | Trp | Gln | Arg | Arg | Arg | Arg | Met | Ser | Met | Arg | Arg | Glu | Arg | Arg | Ile | Val | |
| | | | | 195 | | | | 200 | | | | | | 205 | | | |
| 25 | Lys | Met | Leu | Ile | Ile | Ile | Met | Val | Val | Phe | Ile | Ile | Cys | Trp | Leu | Pro | |
| | 210 | | | | | | 215 | | | | | 220 | | | | | |
| 30 | Tyr | Phe | Ile | Val | Met | Phe | Met | Asp | Thr | Leu | Met | Met | Trp | Trp | Phe | Cys | |
| | 225 | | | | | 230 | | | | | 235 | | | | 240 | | |
| 35 | Glu | Phe | Cys | Ile | Trp | Arg | Arg | Leu | Trp | Met | Tyr | Ile | Phe | Glu | Trp | Leu | |
| | | | | 245 | | | | | | 250 | | | | | 255 | | |
| 40 | Ala | Tyr | Val | Asn | Cys | Pro | Cys | Ile | Asn | Pro | Ile | Ile | Tyr | | | | |
| | | | | 260 | | | | | 265 | | | | | | | | |