

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 433 127**

51 Int. Cl.:

C12N 15/62 (2006.01)

C07K 14/705 (2006.01)

C07K 14/715 (2006.01)

C07K 14/78 (2006.01)

C12N 15/55 (2006.01)

C12N 9/16 (2006.01)

A61K 38/17 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.10.2004 E 04809864 (4)**

97 Fecha y número de publicación de la concesión europea: **14.08.2013 EP 1671097**

54 Título: **Procedimientos y composiciones para producir análogos de receptores triméricos secretados y proteínas de fusión biológicamente activas**

30 Prioridad:

02.10.2003 US 677877

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

09.12.2013

73 Titular/es:

**GENHUNTER CORPORATION (100.0%)
624 GRASSMERE PARK DRIVE, SUITE 17
NASHVILLE, TN 37211, US**

72 Inventor/es:

LIANG, PENG

74 Agente/Representante:

ARPE FERNÁNDEZ, Manuel

ES 2 433 127 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

Procedimientos y composiciones para producir análogos de receptores triméricos secretados y proteínas de fusión biológicamente activas

5

CAMPO DEL INVENTO

El presente invento se refiere a procedimientos para la expresión de proteínas y, más específicamente, para crear y expresar proteínas triméricas secretadas y biológicamente activas, tales como receptores triméricos solubles.

10

ANTECEDENTES DEL INVENTO

En los organismos pluricelulares, como por ejemplo los seres humanos, las células se comunican entre sí por la llamada ruta de transducción de la señal, en la cual un ligando secretado (como por ejemplo citoquinas, factores de crecimiento u hormonas) se une a su/s receptor/es de la superficie celular, lo que lleva a la activación de dicho/s receptor/es. Los receptores son proteínas de membrana que consisten de un dominio extracelular responsable de la unión al ligando, y de una región central transmembrana seguida de un dominio citoplasmático encargado de enviar la señal hacia abajo en la ruta. La transducción de la señal puede tener lugar de tres formas distintas dependiendo de la fuente de la señal secretada y de la ubicación de la célula diana que expresa el/los receptor/es: de forma paracrina (comunicación entre células vecinas), de forma autocrina (comunicación hacia la propia célula) o de forma endocrina (comunicación entre células distantes a través de la circulación). Uno de los mecanismos generales que subyace a la activación del receptor y que pone en marcha la cascada de eventos bajo la membrana celular, incluyendo la activación de la expresión génica, es un ligando polipeptídico, como por ejemplo una citoquina, presente en su forma oligomérica, como un homo-dímero o trímero, que cuando se une a su receptor monomérico en la superficie exterior de las células lleva a la oligomerización del receptor. Las rutas de transducción de la señal desempeñan un papel clave en el desarrollo normal de las células y en su diferenciación, así como en la respuesta a ataques externos tales como infecciones bacterianas o víricas. Determinadas anomalías en estas rutas de transducción de la señal, ya sea en forma de subactivación (por ejemplo con la falta de ligando) o bien de sobreactivación (por ejemplo con un exceso de ligando), son las causas subyacentes de condiciones patológicas y enfermedades tales como la artritis, el cáncer, el SIDA o la diabetes.

15

20

25

30

Una de las actuales estrategias para el tratamiento de estas enfermedades debilitantes se basa en el uso de receptores señuelo, tales como receptores solubles que consisten tan solo del dominio de unión al ligando extracelular, para interceptar el ligando y por lo tanto superar la sobreactivación del receptor. El mejor ejemplo de esta estrategia es la creación de Enbrel, una proteína de fusión dimérica soluble receptor TNF- α - inmunoglobulina (IgG) por Immunex (Mohler et al., 1993; Jacobs et al., 1997), que ahora forma parte de Amgen. La familia TNF de citoquinas es una de las principales señales proinflamatorias producidas por el cuerpo en respuesta a una infección o una lesión tisular. Sin embargo, se ha demostrado que la producción anormal de estas citoquinas, por ejemplo en ausencia de infección o lesión de los tejidos, es una de las causas subyacentes de enfermedades como la artritis y la psoriasis. En la superficie celular está presente de forma natural un receptor TNF- α en su forma monomérica antes de la unión a su ligando, el TNF- α , que existe, por el contrario, como un homotrímero (Locksley et al., 2001). En consecuencia, la fusión de un receptor TNF- α soluble con la región Fc de la inmunoglobulina G1, que es capaz de una dimerización espontánea a través de enlaces disulfuro (Sledziewski et al., 1992 y 1998), permite la secreción de un receptor TNF- α dimérico soluble (Mohler et al., 1993; Jacobs et al., 1997). En comparación con el receptor soluble monomérico, la unión dimérica receptor TNF- α - Fc tiene una afinidad mucho mayor al ligando homotrímérico. Esto proporciona una base molecular para su uso clínico en el tratamiento de la artritis reumatoide (AR), una enfermedad autoinmune en la cual niveles constitutivamente altos de TNF- α , una citoquina proinflamatoria importante, juegan un papel clave en la causa de la enfermedad. Aunque Enbrel ha demostrado tener valores de Ki para TNF- α que se encuentran en el rango de pM (ng/mL) (Mohler et al., 1993), en los pacientes con AR son necesarias inyecciones subcutáneas de 25 mg dos veces por semana para lograr beneficios clínicos, lo que se traduce en niveles del receptor soluble de mg/mL (www.enbrel.com). El alto nivel de consumo recurrente de Enbrel por los pacientes con AR ha generado una gran presión así como un alto coste en lo que respecta al suministro del medicamento, lo que por sí solo limita el acceso al mismo a millones de pacientes potenciales ya solo en este país.

35

40

45

50

Además de la familia TNF- α de las importantes citoquinas proinflamatorias, el virus del VIH que causa el SIDA también utiliza una proteína de cubierta homo-trimérica, la gp120, para poder entrar en las células T colaboradoras CD-4 positivas en nuestro cuerpo (Kwong et al., 1998). Uno de los primeros eventos durante la infección por VIH implica la unión de la gp120 a su receptor CD-4, expresado únicamente en la superficie celular de las células T colaboradoras (Clapham et al., 2001). Hace una década se demostró que el CD-4 monomérico soluble es un agente potente contra la infección por VIH (Clapham et al., 1989). Sin embargo, por desgracia, el entusiasmo cayó cuando se demostró que su capacidad estaba limitada únicamente a cepas de VIH del laboratorio (Daar et al., 1990). Resultó que las cepas de VIH de pacientes con SIDA tenían una afinidad mucho menor al CD-4 monomérico soluble que las cepas de laboratorio, probablemente debido a la variación de la secuencia en la gp120 (Daar et al., 1990). Aunque se han desarrollado proteínas de fusión diméricas solubles Fc - CD-4, estos receptores señuelo CD-4 del VIH mostraron poco efecto antiviral, tanto en los laboratorios como en las clínicas, contra los VIH que se dan de forma natural en los pacientes con SIDA, debido a la baja afinidad por la gp120 (Daar et al., 1990).

60

Es por ello que existe una gran necesidad de crear receptores homotriméricos solubles secretados o proteínas biológicamente activas, las cuales pueden tener sitios de unión perfectamente acoplados, y por lo tanto una mayor afinidad, a sus ligandos homotriméricos presentes en la naturaleza, tales como la familia TNF de las citoquinas y las proteínas de cubierta del VIH. Tales receptores señuelo triméricos deberían tener teóricamente una afinidad mucho mayor por su ligando trimérico que sus homólogos diméricos. Estos análogos de receptores triméricos solubles diseñados de forma racional podrían aumentar significativamente los beneficios clínicos, así como reducir la cantidad o la frecuencia de las inyecciones de medicamentos para cada paciente. Para ser terapéuticamente factible, como la inmunoglobulina Fc, idealmente la fracción de la proteína de trimerización deseada debe ser parte de una proteína secretada de forma natural que sea a su vez abundante en el cuerpo y capaz de autotrimerizarse de forma eficiente.

El colágeno es una familia de proteínas fibrosas que representan los componentes principales de la matriz extracelular. Es la proteína más abundante de los mamíferos, y constituye casi el 25% de las proteínas totales del cuerpo. El colágeno juega un papel estructural importante en la formación de huesos, tendones, piel, córneas, cartílagos, vasos sanguíneos y dientes (Stryer, 1988). Los tipos fibrilares de colágeno I, II, III, IV, V y XI se sintetizan como precursores triméricos más grandes denominados procolágenos, en los que el dominio de triple hélice central no interrumpida que consta de cientos de repeticiones "GXY" (o repeticiones de glicina) está acompañado de los dominios no colágeno (NC), el N- propéptido y el C-propéptido (Stryer, 1988). Tanto la extensión C- como la N-terminal se procesan proteolíticamente después de la secreción del procolágeno, un evento que desencadena el ensamblaje de la proteína madura en las fibrillas de colágeno, las cuales forman una matriz celular insoluble (Prockop et al., 1998). El trímero C-propéptido de colágeno de tipo I se encuentra en la sangre de las personas sanas en una concentración que se encuentra entre 100 y 600 ng/mL, presentando niveles más altos en los niños, los cuales son indicativos de la formación activa de hueso.

Los colágenos de tipo I, IV, V y XI están ensamblados principalmente en formas heterotriméricas que consisten en dos cadenas α -1 y una cadena α -2 (para los tipos I, IV y V), o de tres cadenas α diferentes (para el tipo XI), las cuales son altamente homólogas en su secuencia. Los colágenos de tipo II y III son homotrímeros de cadena α -1. En el caso del colágeno tipo I, la forma más abundante de colágeno, también se forma el homotrímero estable α -1(I) que está presente en niveles variables en diferentes tejidos (Álvarez et al., 1999). La mayoría de estas cadenas C-propéptido de colágeno se pueden autoensamblar en homotrímeros cuando se sobreexpresan solas en una célula. Aunque primero se sintetizan los dominios N-propéptido, el ensamblaje molecular en colágeno trimérico comienza con la asociación de los C-propéptidos. Se cree que el complejo C-propéptido se estabiliza por la formación de enlaces disulfuro intercatenarios, pero la necesidad de que se formen enlaces disulfuro para la adecuada obtención de la cadena no está clara. La triple hélice de glicina se repite y se propaga en forma de cremallera desde los extremos C-terminal asociados hasta los extremos N-terminal. Esto ha permitido la creación de tipos no naturales de matriz de colágeno mediante el intercambio de los C-propéptidos de diferentes cadenas de colágeno por medio del uso de tecnología de ADN recombinante (Bulleid et al., 2001). Las proteínas no colágenas, como las citoquinas y los factores de crecimiento, también se han fusionado con el extremo N-terminal, tanto del procolágeno como del colágeno maduro, para permitir la formación de una nueva matriz de colágeno que está destinada a permitir la lenta liberación de las proteínas no colágenas de la matriz celular (Tomita et al., 2001). Sin embargo, en los dos casos, es necesario que los C-propéptidos se escindan antes de que las fibrillas de colágeno recombinante se ensamblen en una matriz celular insoluble.

Mohler et al., J Immunol, 1993, 151(3): 1548-61, mostraron que los receptores solubles del factor de necrosis tumoral (TNF por sus siglas en inglés 'tumor necrosis factor') son agentes terapéuticos eficaces en la endotoxemia letal y funcionan simultáneamente como portadores y antagonistas de TNF.

El documento EP 0985732 A2 da a conocer una proteína quimera trimérica que ha conservado suficiente actividad fisiológica de un péptido fisiológicamente activo junto con un colágeno con capacidad para la formación de fibrillas. Una proteína quimera trimérica que comprende un colágeno y un péptido biológicamente activo fusionado con el colágeno, en cuyo caso el péptido biológicamente activo se fusiona en el extremo amino terminal del colágeno y el colágeno es un colágeno formador de fibrillas que tiene una estructura de triple hélice en su región de triple hélice.

El documento WO 02/090553 A2 da a conocer proteínas de fusión recombinantes que tienen la propiedad de formar trímeros. Dichas proteínas de fusión recombinantes comprenden al menos un componente A y al menos un componente B. El componente B posee propiedades de trimerización mientras que el componente A presenta propiedades biológicas. El documento WO 02/090553 A2 da a conocer trímeros de dichas proteínas de fusión recombinantes. El documento WO 02/090553 A2 describe también el uso de dichos trímeros en la producción de un medicamento o el uso de los mismos para el diagnóstico *in vitro* o para la producción de un agente de diagnóstico *in vitro*. El documento WO 02/090553 A2 también se refiere a secuencias de ADN que codifican para dicha proteína de fusión y vectores de expresión y células huésped que contienen dichas secuencias de ADN o vectores de expresión.

Hayashi et al., Biochim Biophys Acta, 1528 (2001), 187-95, describen la producción de la proteína quimérica de EGF - colágeno que muestra la actividad mitogénica.

Kishore et al., J. Immunol (2001) 166: 559-565, describen que un homotrímero recombinante, compuesto de una región de cuello de hélice alfa de la proteína D surfactante humana y del dominio globular C1qB, es un inhibidor de la ruta complementaria clásica.

Holler et al., J Immunol Methods (2000) 237(1-2): 159-73, describen el desarrollo de inhibidores solubles de FasL y CD40L mejorados y basados en receptores oligomerizados.

RESUMEN DEL INVENTO

Se presenta aquí un invento que permite que cualquier receptor soluble o polipéptido biológicamente activo se pueda desarrollar en su forma trimérica como una proteína secretada. La esencia del invento es fusionar, por medio de tecnología de ADN recombinante, receptores solubles y proteínas biológicamente activas al dominio C-propéptido del colágeno fibrilar que es capaz de llevar a cabo la autotrimerización. Las proteínas de fusión resultantes, cuando se expresan en células eucariotas, son secretadas como proteínas solubles principalmente en sus formas triméricas fuertemente unidas de forma covalente por enlaces disulfuro intermoleculares formados entre los tres C-propéptidos. En un aspecto del invento, se da a conocer un procedimiento para producir proteínas de fusión triméricas secretadas, que consiste de las siguientes fases: (a) la creación de un constructo de ADN (AND recombinado) que incluye un promotor de transcripción ligado a una matriz que codifique una secuencia de péptido señal seguido de una unión en fase al polipéptido no colágeno que se va a trimerizar, el cual es un polipéptido biológicamente activo o un receptor soluble que consiste de un/unos dominio/s de unión al ligando, y que a su vez está unido en fase a una porción C-terminal de procolágeno que es capaz de llevar a cabo la autotrimerización en una forma soluble con enlaces disulfuro intermoleculares unidos fuertemente de forma covalente; (b) introducir dicho constructo de ADN en una célula eucariota; (c) hacer crecer la célula huésped bajo condiciones fisiológicas en un medio de crecimiento adecuado para permitir la secreción de una proteína de fusión trimerizada soluble codificada por dicha secuencia de ADN, y (d) aislar la proteína de fusión trimerizada del medio de cultivo de la célula huésped.

En un ejemplo de ejecución, la secuencia del péptido señal es la secuencia nativa de la proteína que se va a trimerizar. En otro ejemplo de ejecución, la secuencia del péptido señal es de una proteína secretada distinta de aquella que se va a trimerizar. En un ejemplo de ejecución, el polipéptido no colágeno que se va a trimerizar es un receptor soluble que consiste de un/os dominio/s de unión al ligando. En un ejemplo de ejecución, la porción C-terminal del colágeno es el C-propéptido que no tiene ninguna región de triple hélice de colágeno (IDs de Secuencia: 3-4). En otro ejemplo de ejecución, el extremo C-terminal del colágeno consiste de una fracción de la región de triple hélice de colágeno que funciona como un conector con las proteínas no colágenas que se van a trimerizar (IDs de Secuencia: 1-2). En otro ejemplo de ejecución, la porción C-terminal del colágeno tiene un sitio de reconocimiento para la proteasa BMP-1 que está mutado o suprimido (IDs de secuencia: 3-4).

Según uno de los aspectos del invento, se da a conocer un procedimiento para producir una proteína de fusión trimérica secretada que consiste de las siguientes fases: (a) introducir en una célula huésped eucariota un constructo de ADN que comprende un promotor que dirija la transcripción de un marco de lectura abierto que consiste en una secuencia de péptido señal que está unida en fase a un polipéptido no colágeno que se va a trimerizar, el cual se une en fase a su vez a la porción C-terminal de colágeno que es capaz de llevar a cabo la autotrimerización, y que es seleccionada de pro.alfa.1 (I), pro.alfa.2 (I), pro.alfa.1 (II), pro.alfa.1 (III), pro.alfa.1 (V), pro.alfa.2 (V), pro.alfa.1 (XI), pro.alfa.2 (XI) y pro.alfa.3 (XI); (b) cultivar la célula huésped bajo condiciones fisiológicas en un medio de crecimiento apropiado para permitir la secreción de una proteína de fusión trimerizada codificada por dicha secuencia de ADN, y (c) aislar la proteína de fusión trimérica secretada a partir de una célula huésped.

En un ejemplo preferido de ejecución, el polipéptido no colágeno que se va a trimerizar es el TNF-RII soluble (p75) (ID de Secuencia: 9-12). En otro ejemplo de ejecución preferido, el polipéptido no colágeno que se va a trimerizar es el CD-4 soluble, el correceptor del VIH (ID de Secuencia: 13-16). En un nuevo ejemplo preferido de ejecución, el polipéptido no colágeno que se va a trimerizar es una fosfatasa alcalina secretada de la placenta (ID de Secuencia: 5-8).

En un aspecto del presente invento, se da a conocer un procedimiento para producir una proteína de fusión trimérica secretada, que consiste de las siguientes fases: (a) introducir en una célula huésped eucariota un primer constructo de ADN que comprende un promotor que dirige la transcripción de un marco de lectura abierto que consiste en una secuencia de péptido señal que está unida en fase a un polipéptido no colágeno que se va a trimerizar, el cual está unido a su vez a la porción C-terminal del colágeno que es capaz de llevar a cabo la autotrimerización, y que es seleccionada de pro.alfa.1 (I), pro.alfa.2 (I), pro.alfa.1 (II), pro.alfa.1 (III), pro.alfa.1 (V), pro.alfa.2 (V), pro.alfa.1 (XI), pro.alfa.2 (XI) y pro.alfa.3 (XI), (b) introducir en una célula huésped eucariota un segundo constructo de ADN que comprende un promotor que dirige la transcripción de un marco de lectura abierto que consiste de una segunda secuencia de péptido señal que está unida en fase a un segundo polipéptido no colágeno que se va a trimerizar, el cual está unido en fase, a su vez, a la segunda porción C-terminal del colágeno que es capaz de llevar a cabo la autotrimerización y que es seleccionada de pro.alfa.1 (I), pro.alfa.2 (I), pro.alfa.1 (II), pro.alfa.1 (III), pro.alfa.1 (V), pro.alfa.2 (V), pro.alfa.1 (XI), pro.alfa.2 (XI) y pro.alfa.3 (XI), (c) hacer crecer la célula huésped bajo condiciones fisiológicas en un medio de crecimiento adecuado para permitir la secreción de una proteína de fusión trimerizada codificada por las susodichas primera y segunda secuencias de ADN y (d) aislar la proteína de fusión trimérica secretada por la célula huésped.

En un aspecto del presente invento, se da a conocer un procedimiento para producir una proteína de fusión trimérica secretada, que consiste de las siguientes fases: (a) introducir en una célula huésped eucariota un primer constructo de ADN que comprende un promotor que dirige la transcripción de un marco de lectura abierto que consiste de una secuencia de péptido señal que está unida en fase a un polipéptido no colágeno que se va a trimerizar, el cual está unido en fase, a su vez, a la porción C-terminal del colágeno que es capaz de llevar a cabo la autotrimerización y que es seleccionada de pro.alfa.1 (I), pro.alfa.2 (I), pro.alfa.1 (11), pro.alfa.1 (III), pro.alfa.1 (V), pro.alfa.2 (V), pro.alfa.1 (XI), pro.alfa.2 (XI) y pro.alfa.3 (XI), (b) introducir en una célula huésped eucariota un segundo constructo

de ADN que comprende un promotor que impulsa la transcripción de un marco de lectura abierto que consiste de una segunda secuencia de péptido señal que está unida en fase a un segundo polipéptido no colágeno que se va a trimerizar, el cual está unido en fase, a su vez, a una segunda porción C-terminal del colágeno que es capaz de llevar a cabo la autotrimerización y que es seleccionada de pro.alfa.1 (I), pro.alfa.2 (I), pro.alfa.1 (II), pro.alfa.1 (III), pro.alfa.1 (V), pro.alfa.2 (V), pro.alfa.1 (XI), pro.alfa.2 (XI) y pro.alfa.3 (XI), (c) introducir en una célula huésped eucariota un tercer constructo de ADN que comprende un promotor que dirige la transcripción de un marco de lectura abierto que consiste de una tercera secuencia de péptido señal que está unida en fase a un tercer polipéptido no colágeno que se va a trimerizar, el cual está unido en fase, a su vez, a una tercera porción C-terminal del colágeno que es capaz de llevar a cabo la autotrimerización y que es seleccionada de pro.alfa.1 (I), pro.alfa.2 (I), pro.alfa.1 (II), pro.alfa.1 (III), pro.alfa.1 (V), pro.alfa.2 (V), pro.alfa.1 (XI), pro.alfa.2 (XI) y pro.alfa.3 (XI), (d) hacer crecer la célula hospedadora bajo condiciones fisiológicas en un medio de crecimiento adecuado para permitir la secreción de una proteína de fusión trimerizada codificada por las susodichas primera y segunda secuencias de ADN, y (e) aislar la proteína de fusión trimérica secretada por la célula huésped.

Este invento presenta las siguientes ventajas: (1) el colágeno es la proteína más abundante secretada en el cuerpo de un mamífero y constituye cerca del 25% de las proteínas totales del cuerpo; (2) las principales formas de colágeno se producen naturalmente en forma de hélices triméricas, y sus C-propéptidos globulares son los responsables del inicio de la trimerización, (3) el C-propéptido de colágeno trimérico liberado proteolíticamente del colágeno maduro se encuentra naturalmente en niveles por debajo del microgramo/mL en la sangre de los mamíferos y no se ha demostrado que sea tóxico para el cuerpo; (4) la región de triple hélice lineal de colágeno puede estar incluida como un enlazador con una separación esperada de 2.9 Å por residuo o excluida como parte de la proteína de fusión, por lo que la distancia entre la proteína que se va a trimerizar y el C-propéptido de colágeno puede ser ajustada de forma precisa para lograr una actividad biológica óptima; (5) el sitio de reconocimiento del BMP1, el cual escinde el C-propéptido del procolágeno, se puede mutar o suprimir para evitar la alteración de la proteína de fusión trimérica; (6) el dominio C-propéptido proporciona una etiqueta de afinidad universal, que se puede utilizar para la purificación de cualquiera de las proteínas de fusión secretadas creadas por el presente invento.

En contraste con la tecnología Tag Fc (Sledziewski et al., 1992 y 1998), con la que se pueden crear proteínas de fusión diméricas secretadas, el presente invento aquí descrito permite por primera vez la creación y secreción de proteínas de fusión solubles triméricas. Teniendo en cuenta el hecho de que un homotrímero tiene simetría ternaria mientras que un homodímero tiene tan solo simetría binaria, teóricamente estas dos formas estructurales distintas nunca pueden ser perfectamente superpuestas (Fig. 1). De esta forma, ni el homodímero soluble TNF-R-Fc (como por ejemplo, Enbrel), ni las proteínas de fusión CD4-Fc solubles, podrían haber tenido una interfaz óptima para la unión a sus correspondientes ligandos homotriméricos, TNF- α y la gp120 del VIH, respectivamente. Por el contrario, el CD4 y los receptores TNF solubles homotriméricos creados por el presente invento son estructuralmente trivalentes y tienen el potencial de acoplarse perfectamente a los ligandos homotriméricos correspondientes. Es por ello que estos análogos de receptores triméricos solubles pueden ser mucho más eficaces neutralizando las actividades biológicas de sus ligandos triméricos. Con este oportuno invento se pueden diseñar fácilmente y racionalmente medicamentos más eficaces y aún por encima menos costosos, tales como el CD4 y el TNF-R soluble trimérico descritos en los ejemplos preferidos de ejecución, para combatir las enfermedades debilitantes, tales como la artritis y el SIDA. Con el presente invento se puede crear también la gp120 trimérica soluble, que puede mimetizar mejor el complejo de la proteína de cubierta gp120 trimérica nativa que se encuentra en los virus del VIH, y se utiliza así como una vacuna más eficaz en comparación con la antígenos gp120 no triméricos utilizados anteriormente. Anticuerpos quimera en forma trimérica también pueden ser creados con el presente invento, lo que le proporcionaría a un anticuerpo una mayor intensidad en la neutralización de su antígeno.

45 BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las figuras 1A, 1B, 1C y 1D son una representación esquemática del procedimiento de acuerdo con el presente invento en comparación con la anterior fusión de la inmunoglobulina Fc dimérica.

50 Las figuras 1A y 1B son una vista en alzado lateral y una vista en planta desde arriba, respectivamente, de las características estructurales de una fusión Fc - receptor TNF-RII homodimérico soluble, como por ejemplo Enbrel de Amgen, ya sea en forma de ligando libre o unido tal y como se indica.

Los dominios marcados con un 1 indican el TNF-RII soluble. Nótese que la proteína de fusión Fc (marcada con un 2 con enlaces disulfuro intercatenarios 3) es una estructura dimérica. Dada su simetría binaria, la proteína de fusión Fc dimérica es bivalente y por lo tanto teóricamente no tiene una conformación óptima para unirse a un ligando homotrimérico, como por ejemplo el TNF- α (marcado con un 4) que tiene una simetría ternaria.

55 Las figuras 1C y 1D son una vista en alzado lateral y una vista en planta desde arriba, respectivamente, de las características estructurales de la fusión de un receptor sTNF-RII soluble trimérico y un C-propéptido.

Dada su simetría ternaria, una proteína de fusión sTNF-RII - Trimer es trivalente en la naturaleza, por lo que puede acoplarse perfectamente a su ligando trimérico TNF- α . El C-propéptido de colágeno que es capaz de llevar a cabo la autotrimerización está etiquetado con un 5 con enlaces disulfuro intercatenarios 3.

60 Las figuras 2A y 2B ilustran las estructuras de vectores plasmídicos pTRIMER para la creación de proteínas de fusión triméricas secretadas. Cualquier ADNc que codifica para receptores solubles o polipéptidos biológicamente activos puede ser clonado en los sitios únicos Hind III o Bgl II para permitir la unión en fase en los extremos C-terminales del colágeno α (I) que contiene la secuencia C-propéptido para la trimerización.

La figura 2A muestra que el constructo pTRIMER (T0) contiene parte de las repeticiones de glicina (GXY)_n por encima del C-propéptido.

La figura 2B muestra que el pTRIMER (T2) contiene solo el dominio C-propéptido con un sitio de reconocimiento de la proteasa BMP-1 mutado.

5 Las figuras 3A y 3B ilustran la expresión y secreción de proteínas de fusión de colágeno triméricas unidas por enlaces disulfuro.

10 La figura 3A muestra un Western blot de la trimerización de la fosfatasa alcalina de placenta humana (AP) cuando se fusiona con los C-propéptidos de colágeno α (I). Los vectores de expresión que codifican ya sea para AP sola o para las uniones de AP - C-propéptido en vectores pTRIMER fueron transitoriamente transfectados en células HEK293T. Cuarenta y ocho horas más tarde, el medio condicionado (20 mL) de cada una de las células transfectadas en la forma indicada fue hervido durante 5 minutos en el mismo volumen de 2X tampón SDS, con o sin agente reductor (mercaptoetanol), separado en un 10% de SDS -PAGE y analizado por Western blot utilizando un anticuerpo policlonal para AP (GenHunter Corporation). Nótese que las AP de 67 kDa secretadas no forman por sí solas enlaces disulfuro intermoleculares, mientras que las fusiones AP-T0 y AP-T2 secretadas son ensambladas de manera eficiente en trímeros unidos por enlaces disulfuro.

15 La figura 3B muestra un Western blot de la trimerización del TNF-RII humano soluble cuando se une a los C-propéptidos del colágeno α (I). Los vectores de expresión que codifican ya sea para la unión AP - C-propéptido (T2) (como un control negativo para la especificidad del anticuerpo), o para uniones C-propéptido - TNF-RII solubles humanos en la forma indicada en los vectores de pTRIMER fueron transitoriamente transfectados en células HEK293T. Cuarenta y ocho horas más tarde, el medio acondicionado (20 mL) de cada una de las células transfectadas y no transfectadas en la forma indicada fue hervido durante 5 minutos en un volumen igual de 2X tampón SDS, con o sin agente reductor (mercaptoetanol), separado en un 10% de SDS-PAGE y analizado por Western blot utilizando un anticuerpo monoclonal para el TNF-RII humano (clon 226, R & D Systems, Inc.). Nótese que el anticuerpo monoclonal solo puede reconocer el TNF-RII secretado con enlaces disulfuro. Tanto las uniones TNF-RII - T0 como TNF-RII - T2 solubles son ensambladas de manera eficiente en trímeros unidos por enlaces disulfuro.

20 Las figuras 4 y 5 ilustran los bioensayos que muestran la potente actividad neutralizante de la proteína de fusión soluble trimérica humana TNF-RII - C-propéptido frente a la apoptosis mediada por TNF- α .

25 La figura 4, las células WEHI - 13VAR sensibles a TNF- α (ATCC) se resuspendieron en 1 millón de células/mL en medio RPMI conteniendo un 10% de SFB. Una alícuota de 100 mL de la suspensión celular se sembró en cada pocillo en una placa de microtitulación de 96 pocillos y a cada pocillo se le añadió Actinomicina D en una concentración de 500 ng/mL seguida por 500 pg/mL de TNF- α humano (R & D Systems) en presencia o ausencia de TNF-RII - T2 trimérico humano soluble en la forma indicada. Como control negativo se añadió la AP-T2 trimérica en lugar de TNF-RII - T2. Después de 16 horas de incubación en un incubador de cultivo tisular, la viabilidad de las células se examinó por medio de un microscopio invertido con un aumento de 20X o mediante azul de Alamar (Biosource, Inc.), que es la tinción indicadora de la viabilidad celular y que fue añadida en cada pocillo a una concentración del 10% (v/v). Las células vivas son capaces de convertir el color de la tinción del azul al rosa. Nótese que el TNF-RII - T2 soluble trimérico humano muestra una potente actividad neutralizante frente al TNF- α , protegiendo a las células de la apoptosis mediada por TNF- α .

30 La figura 5 muestra los análisis cuantitativos de la actividad neutralizante del TNF-RII - T2 soluble trimérico humano frente al TNF- α humano. El experimento se llevó a cabo como se muestra en la Fig. 4A. Dos horas después de añadir el colorante azul de Alamar, el medio de cultivo de cada pocillo fue analizado a una absorbancia (OD) de 575 nm. Las lecturas fueron normalizadas con los valores de los pocillos a los que no se les había añadido TNF- α (100% de viabilidad) o de los que tenían TNF- α pero no se les había añadido el agente neutralizante (0% de viabilidad).

45 BREVE DESCRIPCIÓN DE LAS LISTAS DE SECUENCIAS

Número de ID de secuencia: 1 (963 bases)

50 Secuencia de nucleótidos que codifica para el C-propéptido del constructo T0 del colágeno α (I) humano. El constructo de ADNc fue clonado en el vector pAPtag2, reemplazando a la región de codificación para AP. Las secuencias subrayadas indican los sitios de las enzimas de restricción, utilizados en la construcción del vector pTRIMER correspondiente. Los codones en negrita indican el inicio y la terminación de la región codificante T0.

Número de ID de secuencia: 2 (311 aa)

55 La secuencia proteica esperada para el C-propéptido T0 de colágeno α (I) humano. La secuencia subrayada indica la región de las repeticiones de glicina "por encima" del C-propéptido. Los residuos de aminoácidos que están en rojo indican el sitio de reconocimiento de la proteasa BMP-1.

Número de ID de secuencia: 3 (771 bases)

60 La secuencia de nucleótidos que codifica para el C-propéptido del constructo T2 de colágeno α (I) humano. El constructo de ADNc fue clonado en un vector pAPtag2, en sustitución de la región codificante AP. Las secuencias subrayadas indican los sitios de las enzimas de restricción, utilizados en la construcción del vector pTRIMER correspondiente. Los codones en negrita indican el inicio y la terminación de la región codificante T2.

Número de ID de secuencia: 4 (247 aa)

65 La secuencia proteica esperada para el C-propéptido T2 de colágeno α (I) humano. El residuo de aminoácido que se muestra en rojo indica la ubicación del sitio de reconocimiento de la proteasa BMP-1 mutada.

Número de ID de secuencia: 5 (2487 bases)

- Secuencia de nucleótidos que codifica para la fosfatasa alcalina de la placenta humana (AP) unida al C-propéptido T0 del colágeno α (I) humano (AP-T0). Las secuencias subrayadas indican los sitios de restricción utilizados para llevar a cabo la unión. El sitio de restricción, que marca el sitio de unión mostrado en mitad de la secuencia, es Bgl II. Número de ID de secuencia: 6 (819 aa)
- 5 La secuencia proteica esperada para la proteína de fusión AP-T0. Los residuos de aminoácidos en azul indican los sitios de unión entre las fosfatasas alcalinas de placenta humana (AP) y el polipéptido T0 de colágeno α (I). Los codones en negrita indican el inicio y la terminación de la proteína de fusión. La secuencia subrayada indica la región de las repeticiones de glicina "por encima" del C-propéptido de colágeno α (I) humano. Los residuos de aminoácidos en rojo indican la secuencia de reconocimiento de la proteasa BMP-1.
- 10 Número de ID de secuencia: 7 (2294 bases)
 Secuencia de nucleótidos que codifica para las fosfatasas alcalinas de placenta humana (AP) unidas al C-propéptido T2 de colágeno α (I) humano (AP-T2). Los codones en negrita indican el inicio y la terminación de la proteína de fusión. Las secuencias subrayadas indican los sitios de restricción utilizados para llevar a cabo la fusión. El sitio de restricción, que marca el sitio de fusión mostrado en mitad de la secuencia, es Bgl II.
- 15 Número de ID de secuencia: 8 (755 aa)
 La secuencia proteica esperada para la fusión AP - T2. Los residuos de aminoácidos en azul indican los sitios de fusión entre las fosfatasas alcalinas de placenta humana (AP) y el polipéptido T2 de colágeno α (I). El residuo de aminoácido en rojo indica la ubicación del sitio de reconocimiento de la proteasa BMP-1 mutada.
- 20 Número de ID de secuencia: 9 (1734 bases)
 Secuencia de nucleótidos que codifica para el TNF-RII soluble humano fusionado al C-propéptido T0 de colágeno α (I) humano (sTNF-RII - T0). Los codones en negrita indican el inicio y la terminación de la proteína de fusión. Las secuencias subrayadas indican los sitios de restricción utilizados para llevar a cabo la fusión. La secuencia subrayada, que marca el sitio de fusión mostrado en mitad de la secuencia, es la unión BamH I / Bgl II.
- 25 Número de ID de secuencia: 10 (566 aa)
 Secuencia proteica esperada para la fusión humana soluble TNF-RII - T0. Los residuos de aminoácidos en azul indican los sitios de fusión entre el TNF-RII humano soluble y el polipéptido T0 de colágeno α (I). La secuencia subrayada indica la región de las repeticiones de glicina "por encima" del C-propéptido de colágeno α (I) humano. Los residuos de aminoácidos en rojo indican el sitio de reconocimiento de la proteasa BMP-1.
- 30 Número de ID de secuencia: 11 (1.542 bases)
 Secuencia de nucleótidos que codifica para el TNF-RII humano soluble fusionado al C-propéptido T2 de colágeno α (I) humano (sTNF-RII - T2). Los codones en negrita indican el inicio y la terminación de la proteína de fusión. Las secuencias subrayadas indican los sitios de restricción utilizados para la construcción de la fusión. La secuencia subrayada, que marca el sitio de fusión mostrado en mitad de la secuencia, es la unión BamH I / Bgl II.
- 35 Número de ID de secuencia: 12 (502 aa)
 La secuencia proteica esperada de la proteína de fusión humana soluble TNF-RII - T2. Los residuos de aminoácidos en azul indican los sitios de fusión entre el TNF-RII humano soluble y el polipéptido T2 de colágeno α (I). El residuo de aminoácido que se muestra en rojo indica la ubicación del sitio de reconocimiento de la proteasa BMP-1 mutada.
- 40 Número de ID de secuencia: 13 (2.139 bases)
 Secuencia de nucleótidos que codifica el CD4 humano soluble fusionado al C-propéptido T0 de colágeno α (I) humano. Las secuencias subrayadas indican los sitios de restricción utilizados para la construcción de la fusión. La secuencia subrayada, que marca el sitio de fusión mostrado en mitad de la secuencia, es el sitio Bgl II.
- 45 Número de ID de secuencia: 14 (699 aa)
 La secuencia proteica esperada de la fusión humana soluble CD4 - T0. Los residuos de aminoácidos en azul indican los sitios de fusión entre el polipéptido T0 de colágeno α (I) y el CD4 soluble humano. La secuencia subrayada indica la región de las repeticiones de glicina "por encima" del C-propéptido de colágeno α (I) humano. Los residuos de aminoácidos que se muestran en rojo indican el sitio de reconocimiento de la proteasa BMP-1.
- 50 Número de ID de secuencia: 15 (1.947 bases)
 Secuencia de nucleótidos que codifica el CD4 humano soluble fusionado al C-propéptido T2 de colágeno α (I) humano. Las secuencias subrayadas indican los sitios de restricción utilizados para la construcción de la fusión. La secuencia subrayada, que marca el sitio de fusión mostrado en mitad de la secuencia, es el sitio Bgl II.
- 55 Número de ID de secuencia: 16 (635 aa)
 La secuencia proteica esperada de la fusión humana soluble CD4 - T2. Los residuos de aminoácidos en azul indican los sitios de fusión entre el CD4 humano soluble y el polipéptido T2 de colágeno α (I). El residuo de aminoácido en rojo indica la ubicación del sitio de reconocimiento de la proteasa BMP-1 mutada.

DESCRIPCIÓN DEL INVENTO

60 Antes de presentar el invento, puede resultar útil para su comprensión presentar las definiciones de ciertos términos que se utilizarán de aquí en adelante.

Constructo de ADN: una molécula de ADN, generalmente en forma de un vector plasmídico o vírico, ya sea de una hebra doble o sencilla que ha sido modificado por medio de la tecnología de ADN recombinante para contener segmentos de ADN unidos de tal forma que en su conjunto no existirían en la naturaleza. Los constructos de ADN contienen la información necesaria para dirigir la expresión y/o secreción de la proteína codificante de interés.

Secuencia de péptido señal: tramo de una secuencia de aminoácidos que actúa dirigiendo la secreción de un polipéptido maduro o una proteína de una célula. Los péptidos señal se caracterizan por un núcleo de aminoácidos hidrófobos y se encuentran típicamente en los extremos amino de las proteínas recién sintetizadas que van a ser secretadas o ancladas a la superficie celular. El péptido señal se escinde a menudo de la proteína madura durante la secreción. Dichos péptidos señal contienen sitios que permiten la escisión de los péptidos señal de las proteínas maduras a medida que pasan a través de la ruta de secreción proteica. Cuando una secuencia de péptido señal está unida al extremo amino terminal de otra proteína que no tiene péptido señal puede dirigir la secreción de la proteína fusionada. La mayoría de las proteínas secretadas, como los factores de crecimiento, las hormonas peptídicas, las citoquinas y las proteínas de membrana, como por ejemplo los receptores de la superficie celular, contienen una secuencia de péptido señal cuando se sintetizan como una proteína naciente.

Receptor soluble: una parte o la totalidad de un dominio extracelular de un receptor de la superficie celular que es capaz de unirse a su ligando. Por lo general, no contiene ningún tramo interno de la secuencia de aminoácidos hidrófobos responsables del anclaje a la membrana.

C-propéptido de colágeno: el C-terminal globular, y el dominio no triple hélice de colágeno, que es capaz de autoensamblarse en trímeros. A diferencia de lo que ocurre con la región de triple hélice de colágeno, el C-propéptido no contiene ninguna secuencia de repetición de glicina y normalmente es eliminado proteolíticamente a partir del precursor procolágeno después de la secreción de procolágeno antes de la formación de las fibrillas de colágeno.

Repeticiones de glicina: región central lineal de triple hélice formadora de colágeno que contiene cientos de repeticiones de (Gly -XY)_n en la secuencia de aminoácidos. Estas repeticiones también son ricas en prolina en las posiciones X y/o Y. Tras la eliminación de los N- y C-propéptidos, las repeticiones de glicina que contienen triples hélices de colágeno pueden ensamblarse en fibrillas de colágeno insolubles de mayor orden, lo que constituye el componente principal de la matriz celular.

ADNc: hebras para el ADN complementario o secuencia de ADN complementaria al ARN mensajero. Por lo general, las secuencias de ADNc no contienen ninguna secuencias de intrones (codificación no proteica).

Antes de este invento, casi todos los anticuerpos terapéuticos y las proteínas de fusión de receptor Fc soluble, tales como Enbrel, eran de estructura dimérica (Fig. 1). Aunque se ha demostrado que estas moléculas, en comparación con sus homólogos monoméricos, se unen a sus antígenos diana o ligandos con una mayor intensidad, es de esperar que todavía sean imperfectas, debido a las limitaciones estructurales, para unir las dianas que tienen una estructura homotrimérica. Entre los ejemplos de ligandos triméricos terapéuticamente importantes se encuentran la familia TNF de citoquinas y la proteína de cubierta gp120 del VIH. Por lo tanto, desde un punto de vista estructural, también es deseable que se puedan generar receptores o anticuerpos solubles triméricos que se que puedan acoplar perfectamente a sus ligandos triméricos o antígenos diana (Fig. 1), y bloquear de este modo completamente la actividad del ligando. Se espera que dichos receptores solubles o anticuerpos quiméricos triméricos tengan una afinidad más alta por sus dianas y que por lo tanto se puedan utilizar de una forma más eficaz y eficiente para el tratamiento de enfermedades como la artritis y el SIDA.

Este invento describe formas para la generación de dichos receptores triméricos secretados y proteínas biológicamente activas por fusión a los C-propéptidos de colágeno, que son capaces de autoensamblarse en trímeros. El invento presenta las siguientes ventajas: (1) el colágeno es la proteína más abundante secretada en el cuerpo de un mamífero, constituyendo cerca del 25% de las proteínas totales del cuerpo; (2) las principales formas de colágeno se producen de forma natural como hélices triméricas, con sus C-propéptidos globulares como responsables del inicio de la trimerización, los cuales se escinden posteriormente mediante proteólisis tras la formación de la triple hélice; (3) el C-propéptido trimérico soluble escindido de colágeno se encuentra de forma natural en niveles inferiores a un microgramo/mL en la sangre de los mamíferos; (4) la región de triple hélice lineal de colágeno puede ser incluida como un enlazador o excluida como parte de la proteína de fusión de tal modo que la distancia entre una proteína que va a ser trimerizada y el C-propéptido de colágeno se puede ajustar de forma precisa para lograr una actividad biológica óptima; (5) el sitio de reconocimiento de BMP1 que escinde el C-propéptido del pro-colágeno se puede mutar o suprimir para evitar la alteración de una proteína de fusión trimérica; (6) el dominio C-propéptido proporciona un marcador de afinidad universal, que puede ser utilizado para la purificación de cualquiera de las proteínas de fusión secretadas creadas por este invento; (7) a diferencia del marcador Fc IgG1 que es conocido por tener otras funciones biológicas, tales como la unión a sus propios receptores de superficie celular, la única función biológica del C-propéptido de colágeno es la capacidad de iniciar la trimerización de cadenas nacentes de procolágeno y mantener el trímero recién hecho de pro-colágeno soluble antes del ensamblaje en la matriz celular insoluble. Por estas propiedades únicas del C-propéptido de colágeno es de esperar que sea poco probable que esta etiqueta única de trimerización sea tóxica o inmunogénica, por lo que es un candidato ideal para aplicaciones terapéuticas.

Para demostrar la viabilidad de la fabricación de proteínas de fusión trimérica secretadas, las secuencias de ADNc que codifican todos los C-propéptidos de colágeno $\alpha 1(I)$ humano que contienen alguna región de triple hélice con repeticiones de glicina (constructo T0, números de ID de secuencia 1-2), o bien sin repeticiones de glicina con un sitio de reconocimiento de BMP-1 mutado (constructo T2, número de ID de Secuencia 3-4), fueron amplificadas por RT-PCR usando los clones EST obtenidos de la American Type Culture Collection (ATCC). Los ADNc amplificados se clonaron cada uno como un fragmento XbaI - Bgl II en el vector de expresión de mamíferos pAptag2 (GenHunter Corporation; Leder et al., 1996 y 1998), en sustitución de la región de codificación AP (Fig. 2). Los vectores resultantes se denominan pTRIMER, versiones T2 y T0 respectivamente. Los vectores permiten la conveniente

fusión en fase de cualquier matriz de ADNc que codifica un receptor soluble o una proteína biológicamente activa en los sitios Hind III y Bgl II únicos. Dichas proteínas de fusión tienen las etiquetas de trimerización de colágeno situadas en los extremos C-terminales, de forma similar a los procolágenos nativos.

5 Ejemplo 1:

10 Para demostrar la viabilidad de este invento, un ADNc que codifica la fosfatasa alcalina de placenta humana secretada (AP), incluyendo su secuencia nativa de péptido señal, fue cortada como un fragmento Bgl II – Hind III del vector de pAPtag4 (GenHunter Corporation; Leder et al., 1996 y 1998) y se clonó en los sitios correspondientes de los vectores pTRIMER - T0 y pTRIMER - T2. Los constructos de fusión AP - colágeno resultantes (número de ID de secuencia 5-8) se expresaron en células HEK293T (GenHunter Corporation) después de la transfección. El éxito de la secreción de las proteínas de fusión AP - colágeno puede ser determinada fácilmente mediante un ensayo de actividad AP utilizando un medio acondicionado de las células transfectadas. La actividad AP alcanzó aproximadamente 1 unidad/mL (o el equivalente a aproximadamente 1 mg/mL de la proteína de fusión) 2 días después de la transfección. Para obtener células HEK293T que expresan de forma estable las proteínas de fusión, se seleccionaron clones estables después de la co-transfección con un vector resistente a puomicina pBabe - Puro (GenHunter Corporation). Los clones que expresan actividad AP se ampliaron y se guardaron para la producción a largo plazo de las proteínas de fusión.

20 Para determinar si las proteínas de fusión AP - colágeno están ensambladas en trímeros unidos por enlaces disulfuro, los medios acondicionados que contienen las AP solas o bien las fusiones AP - T0 y AP - T2 son hervidos en tampones de SDS que contienen β - mercaptoetanol (reductor) o no (no reductor), separados por medio de SDS-PAGE y analizados por Western blot utilizando un anticuerpo de policlonación anti-AP (GenHunter Corporation). Las AP solas, sin fusión, mostraron una banda de 67 kDa, tanto bajo condiciones reductoras como no reductoras, lo que, tal y como era de esperar, es consistente con la falta de enlaces disulfuro intermoleculares (Fig. 3A). Por el contrario, tanto la proteína de fusión secretada AP - T0 como la AP - T2 mostraron un tamaño tres veces mayor (alrededor de 300 kDa) en condiciones no reductoras que en condiciones reductoras (90-100 kDa), lo que indica que las dos proteínas de fusión fueron completamente ensambladas en homotrímeros (Fig. 3A). Este resultado reduce esencialmente el concepto de este invento en la práctica.

30 Ejemplo 2:

35 Para proporcionar una prueba de que se pueden proporcionar funciones biológicas nuevas y terapéuticamente beneficiosas a una proteína de fusión trimérica, se construyó un receptor de TNF-RII trimérico soluble humano (p75) por medio del correspondiente clon EST obtenido de la ATCC. Como se ha descrito en el ejemplo 1, la región N - terminal del TNF-RII humano, incluyendo toda la región de unión al ligando pero excluyendo el dominio transmembrana, fue clonado en fase, como un fragmento Bam HI, en el sitio Bgl II tanto del vector pTRIMER - T0 como del vector pTRIMER - T2 (número de ID de secuencia 9-12). Los constructos de fusión resultantes se expresaron en células HEK293T después de la transfección. Los clones estables fueron obtenidos por coselección con puomicina tal y como ha sido descrito en el ejemplo 1. Se llevó a cabo un análisis por Western blot tanto bajo condiciones reductoras como no reductoras para determinar si las resultantes proteínas de fusión solubles TNF-RII - colágeno fueron en efecto expresadas, secretadas y ensambladas en formas triméricas. Como era de esperar, el anticuerpo monoclonal del TNF-RII humano (clon 226 de R&D Systems, Inc.) reconoció de forma precisa las proteínas de fusión - TNF triméricas solubles expresadas por los dos vectores de fusión T0 y T2 como bandas de 220-240 kDa, que son alrededor tres veces más grande que las correspondientes proteínas de fusión monoméricas (Fig. 3B). El anticuerpo TNF-RII no pudo detectar las proteínas de fusión monoméricas en condiciones reductoras, en consonancia con la propiedad especificada por el fabricante del anticuerpo. Como un control negativo para la especificidad del anticuerpo, ni las células HEK293T solas, ni las células que expresaban la proteína de fusión AP - T2 expresaron el TNF-RII (Fig. 3B).

40 Para determinar si los receptores triméricos solubles de TNF-RII son inhibidores potentes de su ligando trimérico TNF- α , se llevó a cabo un bioensayo de TNF- α usando la línea celular sensible a la citoquina WEHI - 13VAR (ATCC) tal y como ha sido descrito por Mohler et al. (1993). El resultado mostrado en la Fig. 4 indica claramente que las proteínas de fusión solubles triméricas de TNF-RII - C-propéptido son altamente potentes neutralizando la apoptosis mediada por TNF- α de las células WEHI - 13VAR en presencia de actinomicina D (500 ng/mL) (Sigma). Cuando el TNF- α humano (R&D Systems) fue utilizado a una concentración de 0,5 ng/mL, el TNF-RII - T2 trimérico soluble (ya sea de medio libre de suero o en su forma purificada) presentó una Ki-50 aparente (50% de inhibición) de aproximadamente 2 ng/mL o 8×10^{-12} M (asumiendo un Pm de 240 kDa para un homotrímero). Esta afinidad al TNF- α es de un orden de magnitud 4 veces mayor que la del monómero de TNF-RII y al menos de 10 a 100 veces mayor que la de la fusión dimérica soluble de TNF-RII - Fc, como es el caso de Enbrel (Mohler et al., 1993).

50 En este ejemplo fundamental demuestra que este invento puede crear proteínas de fusión triméricas con nuevas propiedades biológicas que pueden tener importantes aplicaciones terapéuticas. Este tipo de receptores triméricos solubles de TNF humanos pueden llegar a ser mucho más eficaces que el receptor TNF dimérico soluble (por ejemplo, Enbrel) que se encuentra actualmente en el mercado para el tratamiento de enfermedades autoinmunes como la AR. El drástico aumento de la potencia de los receptores de TNF triméricos podría reducir en gran medida la cantidad de inhibidores del TNF que se le tienen que inyectar semanalmente a cada paciente, al mismo tiempo que mejora el tratamiento y rebaja significativamente el coste que supone dicho tratamiento a los pacientes.

La mejora en la potencia de los receptores de TNF triméricos también puede aliviar el cuello de botella en la producción actual del receptor de TNF dimérico, ya que actualmente solo es posible satisfacer las demandas en el tratamiento de alrededor de 100.000 pacientes en los Estados Unidos.

5 Ejemplo 3.

10 El virus del VIH, causante del sida, infecta y destruye en nuestro cuerpo todo un linaje especial de linfocitos T. Estos linfocitos conocidos como células T CD4+ expresan una proteína de la superficie celular CD4 doblada, que es el receptor del VIH. El VIH reconoce a las células CD4+ con su proteína de cubierta viral gp120, la cual se une a la CD4. En particular, la gp120 existe como un complejo homotrimérico gigante en la superficie viral, mientras que las CD4 son monoméricas en la superficie celular. El modelo actual para la infección por VIH es el de un acoplamiento completo del VIH a las células T CD4+, en el cual es necesaria la unión de cada una de las tres subunidades de los trímeros de gp120 a la CD4 para que el ARN viral entre en las células. Obviamente, una de las estrategias más directas para detener la infección por VIH es el uso de CD4 soluble para neutralizar el virus. En efecto, tal enfoque utilizando tanto las fusiones de CD4 - Fc como las CD4 solubles monoméricas ha demostrado ser muy eficaz en la reducción de infecciones por el VIH en cepas aisladas de laboratorio (Clapham et al., 1989; Daar et al., 1990). Desafortunadamente, estas CD4 solubles fueron menos eficaces en la detención de la infección de cepas virales de VIH que se encuentran en pacientes con SIDA (Daar et al., 1990), posiblemente debido a las variaciones de la secuencia de amino ácidos de la gp120, lo que disminuye la afinidad para las CD4 monoméricas y diméricas solubles.

20 Para aumentar de manera significativa la afinidad de una CD4 soluble a cualquiera de las variantes de gp120 en los virus VIH, es preferible que la CD4 soluble esté en forma trimérica para que se pueda acoplar perfectamente a su ligando trimérico, los homotrímeros gp120. Uno de los principales retos en la lucha contra el SIDA ha sido la alta tasa de mutación del genoma viral, que lleva a una mayor resistencia a los medicamentos. Es por ello que cualquier fármaco que afecte directamente a genes virales del tipo de la transcriptasa inversa del VIH (por ejemplo, AZT) o la proteasa, son probablemente ineficaces debido a la alta tasa de mutaciones virales. Por el contrario, e independientemente de cuál sea la tasa de mutación, el virus del VIH tiene que unirse a un receptor celular CD4 para iniciar la infección. Por lo tanto, un trímero de CD4 soluble de alta afinidad debería de ser inmune a las mutaciones virales ya que dichas mutaciones en los genes gp120 harán que el virus sea incapaz de unirse no solo a un CD4 trimérico soluble sino también al receptor CD4 de las células.

25 Para crear estos análogos triméricos solubles del receptor CD4 del VIH, se amplificó un ADNc que codifica la totalidad de los CD4 solubles humanos, incluyendo su secuencia nativa de péptido señal pero excluyendo los dominios citoplásmicos cortos y de transmembrana, utilizando un clon EST obtenido de la ATCC. A continuación, el ADNc resultante se clonó como un fragmento Bgl II - Hind III en los sitios correspondientes de los vectores de expresión pTRIMER - T0 y pTRIMER - T2. Los constructos de fusión resultantes de colágeno - CD4 solubles (número de ID de secuencia 13-16) se expresaron en células HEK293T (GenHunter Corporation) después de la transfección. Para obtener células HEK293T que expresaran de forma estable las proteínas de fusión, se seleccionaron clones estables después de la cotransfección con un vector resistente a la puomicina, pBabe - Puro (GenHunter Corporation). Los clones que expresan las proteínas de fusión se multiplicaron y guardaron para obtener posteriormente las proteínas de fusión.

30 Para determinar si las proteínas de fusión colágeno humano - CD4 solubles se ensamblan en trímeros unidos por enlaces disulfuro, los medios acondicionados que contienen las fusiones de CD4 - T0 y CD4 - T2 solubles se llevaron a ebullición en un tampón de SDS sin (no reductor) o con β -mercaptoetanol (reductor), se separaron por medio de una SDS-PAGE y se analizaron por Western blot utilizando un anticuerpo monoclonal para CD4 humano (R & D Systems). Los resultados mostraron que las dos proteínas de fusión solubles secretadas CD4 - T0 y CD4 - T2 eran tres veces más grandes (alrededor de 300 kDa) en condiciones no reductoras que aquellas en condiciones reductoras (90-100 kDa), lo que indica que fueron ensambladas completamente en homotrímeros (datos no mostrados). Ahora bien, estos CD4 soluble triméricos pueden ser fácilmente probados para la unión con la gp120 y la infección anti-VIH.

50 LISTADO DE SECUENCIAS

<110> GENHUNTER CORPORATION
 <120> PROCEDIMIENTOS Y COMPOSICIONES PARA PRODUCIR ANÁLOGOS DE RECEPTORES TRIMÉRICOS
 55 SECRETADOS Y PROTEÍNAS DE FUSIÓN BIOLÓGICAMENTE ACTIVAS
 <130> 04-066-PL
 <150> 10/677,877
 <160> 16
 <210> 1
 60 <211> 963
 <212> ADNc
 <213> HOMO SAPIENS
 <220>
 <221>CDS
 65 <222>12 947

<400> 1

Hind III BglII

5 AAGCTTACGTAAGATCTAACGGTCTCCCTGGCCCCATTGGGCCCCCTGGTCCT
 CGCGGTGCGACTGGTGATGCTGGTCCTGTTGGTCCCCCGGCCCTCCTGGACC
 TCCTGGTCCCCTGGTCCTCCCAGCGCTGGTTTCGACTTCAGCTTCCTGCCCC
 AGCCACCTCAAGAGAAGGCTCACGATGGTGGCCGCTACTACCGGGCTGATGAT
 GCCAATGTGGTTCGTGACCGTGACCTCGAGGTGGACACCACCCTCAAGAGCCT
 10 GAGCCAGCAGATCGAGAACATCCGGAGCCCAGAGGGAAGCCGCAAGAACCCCG
 CCCGCACCTGCCGTGACCTCAAGATGTGCCACTCTGACTGGAAGAGTGGAGAG
 TACTGGATTGACCCCAACCAAGGCTGCAACCTGGATGCCATCAAAGTCTTCTG
 CAACATGGAGACTGGTGAGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCC
 AGAAGAAGTGGTACATCAGCAAGAACCCCAAGGACAAGAGGCATGTCTGGTTC
 GCGGAGACTGACCGATGGATTCCAGTTCGAGTATGGCGGCCAGGGCTCCGA
 15 CCCTGCCGATGTGGCCATCCAGCTGACCTTCCTGCGCCTGATGTCCACCGAGG
 CCTCCCAGAACATCACCTACCACTGCAAGAACAGCGTGGCCTACATGGACCAG
 CAGACTGGCAACCTCAAGAAGGCCCTGCTCCTCAAGGGCTCCAACGAGATCGA
 GATCCGCGCCGAGGGCAACAGCCGCTTACCTACAGCGTCACTGTGATGGCT
 GCACGAGTCACACCGGAGCCTGGGGCAAGACAGTGATTGAATACAAAACACC
 20 AAGTCCTCCCGCCTGCCCATCATCGATGTGGCCCCCTTGGACGTTGGTGGCCC
 AGACCAGGAATTCGGCTTCGACGTTGGCCCTGTCTGCTTCTGTAAACTCCCT
 CCATCTAGA

Xba I

25

<210>2
 <211>311
 <212> PROTEÍNA
 <213> HOMO SAPIENS
 <400> 2

30

1 RSNGLPGPIG PPGPRGRTGD AGPVGPPGPP GPPGPPGPPS AGFDFSFLPQ PPQEKADHGG 60
 61 RYYRADDANV VRDRDLEVDT TLKSLSQQIE NIRSPEGSRK NPARTCRDLK MCHSDWKSQE 120
 121 YWIDPNQGCN LDAIKVFCNM ETGETCVYPT QPSVAQKNWY ISKNPKDKRH VWFGESMTDG 180
 35 181 FQFEYGGQGS DPADVAIQLT FLRLMSTEAS QNITYHCKNS VAYMDQQTGN LKKALLLKGS 240
 241 NEIEIRAEGN SRFTYSVTVD GCTSHTGAWG KTVIEYKTTK SSRLPIIDVA PLDVGAPDQE 300
 301 FGFVGPVCF L

40

<210>3 3
 <211>711
 <212> ADNc
 <213> HOMO SAPIENS
 <220>

45

<221> CDS
 <222> 12 755
 <400> 3

Hind III BglII

50 AAGCTTACGTAAGATCTGATGCCAATGTGGTTCGTGACCGTGACCTCGAGGTGGACACCACC
 CTCAAGAGCCTGAGCCAGCAGATCGAGAACATCCGGAGCCCAGAGGGAAGCCGCAAGAACCC
 CGCCCGCACCTGCCGTGACCTCAAGATGTGCCACTCTGACTGGAAGAGTGGAGAGTACTGGA
 TTGACCCCAACCAAGGCTGCAACCTGGATGCCATCAAAGTCTTCTGCAACATGGAGACTGGT
 GAGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCCAGAAGAACTGGTACATCAGCAAGAA
 CCCC AAGGACAAGAGGCATGTCTGGTTCGGCGAGAGCATGACCGATGGATTCCAGTTCGAGT
 55 ATGGCGGCCAGGGCTCCGACCCTGCCGATGTGGCCATCCAGCTGACCTTCCTGCGCCTGATG
 TCCACCGAGGCTCCCAAGAATCACCTACCCTGCAAGAACAGCGTGGCCTACATGGACCA
 GCAGACTGGCAACCTCAAGAAGGCCCTGCTCCTCAAGGGCTCCAACGAGATCGAGATCCGCG
 CCGAGGGCAACAGCCGCTTACCTACAGCGTCACTGTGATGGCTGCACGAGTCACACCGGA
 GCCTGGGGCAAGACAGTGATTGAATACAAAACCAAGTCTCCCGCCTGCCCATCATCGA
 60 TGTGGCCCCCTTGGACGTTGGTGGCCCCAGACCAGGAATTCGGCTTCGACGTTGGCCCTGTCT
 GCTTCTGTAAACTCCCTCCATCTAGA

Xba I

65

<210>4
 <211> 247

<212> PROTEÍNA
 <213> HOMO SAPIENS
 <400> 4

5 1 RSDANVVRDR DLEVDTTLKS LSQQIENIRS PEGSRKNPAR TCRDLKMCHS DWKSGEYWID 60
 61 PNQGCNLDAI KVFCNMETGE TCVYPTQPSV AQKNWYISKV PKDKRHVWFG ESMTDGFQFE 120
 121 YGGQGSDDPAD VAIQLTFLRL MSTEASQNT YHCKNSVAYM DQQTGNLKKK LLLKGSNEIE 180
 181 IRAEGNSRFT YSVTVDGCTS HTGAWGKTVI EYKTTKSSRL PIIDVAPLDV GAPDQEFQFD 240
 241 VGPVCF

10 <210> 5
 <211> 2487
 <212> ADNc
 <213> HOMO SAPIENS
 15 <220>
 <221> CDS
 <222> 12 2471
 <400> 5

20 Hind III
AAGCTTCCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGGGCCTGAGGCTACAGCTCTCCC
 TGGGCATCATCCCAGTTGAGGAGGAGAACC CGACTTCTGGAACCGCGAGGCAGCCG
 AGGCCCTGGGTGCCGCAAGAAGCTGCAGCCTGCACAGACAGCCGCCAAGAACCTCA
 25 TCATCTTCTGCGGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTAA
 AAGGGCAGAAGAAGGACAACTGGGGCCTGAGATACCCCTGGCCATGGACCGCTTCC
 CATATGTGGCTCTGTCCAAGACATAAATGTAGACAAACATGTGCCAGACAGTGGAG
 CCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGACCATTGGCTTGA
 GTGCAGCCGCCCCTTTAACCAGTGCAACACGACACGCGGCAACGAGGTCATCTCCG
 TGATGAATCGGGCCAAGAAAGCAGGGAAGTCAAGTGGAGTGGTAACCACACACGAG
 30 TGCAGCACGCCTCGCCAGCCGGCACCTACGCCACACGGTGAACCGCAACTGGTACT
 CGGACGCCGACGTGCCTGCCTCGGCCCGCCAGGAGGGGTGCCAGGACATCGCTACGC
 AGCTCATCTCCAACATGGACATTGACGTGATCCTAGGTGGAGGCCGAAAGTACATGT
 TTCCCATGGGAACCCAGACCCTGAGTACCCAGATGACTACAGCCAAGGTGGGACCA
 GGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCGCCAGGGTGCCTGGT
 35 ATGTGTGGAACCGCACTGAGCTGATGCAGGCTTCCCTGGACCCGTCTGTGACCCATC
 TCATGGGTCTTTGAGCCTGGAGACATGAAATACGAGATCCACCGAGACTCCACAC
 TGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGAACC
 CCCGCGGCTTCTTCTTCTCGTGGAGGGTGGTGCATCGACCATGGTTCATCATGAAA
 GCAGGGCTTACCGGGCACTGACTGAGACGATCATGTTTCGACGACGCCATTGAGAGGG
 40 CGGGCCAGCTACCCAGCGAGGAGGACACGCTGAGCCTCGTCACTGCCGACCACTCCC
 ACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCC
 CTGGCAAGGCCCGGACAGGAAGCCCTACACGCTCCTCCTATACGAAACGGTCCAG
 GCTATGTGCTCAAGGACGGCGCCCGGCGGATGTTACCGAGAGCGAGAGCGGGAGCC
 45 CCGAGTATCGGCAGCAGTCAGCAGTGCCCTGGACGAAGAGACCCACGCAGGCGAGG
 ACGTGGCGGTGTTGCGCGCGCGGCCCGCAGGCGCACCTGGTTCACGGCGTGCAGGAGC
 AGACCTTCATAGCGCACGTGATGCTTTCGCCGCTGCTGAGGCCCTACACCGCCT
 GCGACCTGGCGCCCCCGCCGGCACCACCGACGCGCGCACCCGGGTTCCGGAAAGAT
CTAACGGTCTCCCTGGCCCCATTGGCCCCCTGGTCTCGCGGTGCGACTGGTGATG
 CTGGTCTGTTGGTCCCCCGGCCCTCCTGGACCTCCTGGTCCCCCTGGICCTCCCA
 50 GCGCTGGTTTCGACTTCAGCTTCTGCCCCAGCCACCTCAAGAGAAGGCTCACGATG
 GTGGCCGCTACTACCGGGCTGATGATGCCAATGTGGTTCGTGACCGTGACCTCGAGG
 TGGACACCACCCTCAAGAGCCTGAGCCAGCAGATCGAGAACATCCGGAGCCCAGAGG
 GAAGCCGCAAGAACCCCGCCCGCACCTGCCGTGACCTCAAGATGTGCCACTCTGACT
 GGAAGAGTGGAGAGTACTGGATTGACCCCAACCAAGGCTGCAACCTGGATGCCATCA
 55 AAGTCTTCTGCAACATGGAGACTGGTGAGACCTGCGTGTACCCCACTCAGCCCAGTG
 TGGCCCAGAAGAACTGGTACATCAGCAAGAACCCCAAGGACAAGAGGCATGTCTGGT
 TCGCCGAGAGACTGACCGATGGATTCCAGTTTCGAGTATGGCGGCCAGGGCTCCGACC
 CTGCCGATGTGGCCATCCAGCTGACCTTCTGCGCCTGATGTCCACCGAGGCCCTCCC
 AGAACATCACCTACCACTGCAAGAACAGCGTGGCCTACATGGACCAGCAGACTGGCA
 60 ACCTCAAGAAGGCCCTGCTCCTGAAGGGCTCCAACGAGATCGAGATCCGCGCCGAGG
 GCAACAGCCGCTTACCTACAGCGTCACTGTGATGGCTGCACGAGTCACACCGGAG
 CCTGGGGCAAGACAGTGAATGAATACAAAACCAAGTCTCCCGCCTGCCATCA
 TCGATGTGGCCCCCTTGACGTTGGTGGCCCCAGACCAGGAATTCGGCTTCGACGTTG
 65 GCCCTGTCTGCTTCTGTAACCTCCCTCCATCTAGA

Xba I

<210>6
 <211>819
 <212> PROTEÍNA
 5 <213> HOMO SAPIENS
 <400> 6

1 MLLLLLLLGL RLQLSLGIIP VEEENPDFWN REAAEALGAA KKLQPAQTAA KNLHFLGDG 60
 61 MGVSTVTAAR ILKGQKKDKL GPEIPLAMDR FPYVALSKTY NVDKHVPDSG ATATAYLCGV 120
 10 121 KGNFQTIGLS AAARFNQCNT TRGNEVISVM NRAKKAGKSV GVVTTTRVQH ASPAGTYAHT 180
 181 VNRNWYSAD VPASARQEGC QDIATQLISN MDIDVILGGG RKYMFPMGTP DPEYPDDYSQ 240
 241 GGTRLDGKNL VQEWLAKRQG ARYVWNRTTEL MQASLDPSVT HLMGLFEPGD MKYEIHRDST 300
 301 LDPSLMEMTE AALRLLSRNP RGFFLFVEGG RIDHGHESR AYRALTETIM FDDAIERAGQ 360
 361 LTSEEDTSLV VTADHSHVFS FGGYPLRGSS IFGLAPGKAR DRKAYTVLLY GNGPGYVLKD 420
 15 421 GARPDVTESE SGSPEYRQQS AVPLDEETHA GEDVAVFARG PQAHLVHVQV EQTFIAHVMA 480
 481 FAACLEPYTA CDLAPPAGTT DAAHPGSGRS NGLPGPIGPP GPRGRTGDAG PVGPPGPPGP 540
 541 PGPPGPPSAG FDFSFLPQPP QEKAHDGGRY YRADDANVVR DRDLEVDTTL KSLSQIENI 600
 601 RSPEGSRKNP ARTCRDLKMC HSDWKSGEYW IDPNQGCNLD AIKVFCNMET GETCVYPTQP 660
 661 SVAQKNWYIS KNPKDKRHVW FGESMTDGFQ FEYGGQGSDP ADVAIQLTFL RLMSTEASQN 720
 20 721 ITYHCKNSVA YMDQQTGNLK KALLKGSNE IEIRAEGNSR FTYSVTVDGC TSHTGAWGKT 780
 781 VIEYKTTKSS RLPIDVAPL DVGAPDQEFQ FDVGPVCFL

<210> 7
 25 <211> 2294
 <212> ADNc
 <213> HOMO SAPIENS
 <220>
 <221> CDS
 30 <222> 12 2278
 <400> 7

Hind III
 35 AAGCTTCCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGGGCCTGAGGCTACAGCTCTCCC
 TGGGACATCATCCCAGTTGAGGAGGAGAACC CGACTTCTGGAACCGCGAGGCAGCCC
 AGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTGCACAGACAGCCGCCAAGAACCTCA
 TCATCTTCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTGCCAGGATCCTAA
 AAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCCCTGGCCATGGACCGCTTCC
 40 CATATGTGGCTCTGTCCAAGACATAAATGTAGACAAACATGTGCCAGACAGTGGAG
 CCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGACCATTGGCTTGA
 GTGCAGCCGCCCGCTTTAACAGTGCAACACGACACGCGGCAACGAGGTCATCTCCG
 TGATGAATCGGGCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACACACAGC
 45 TGCAGACCGCTCGCCAGCCGACCTACGCCACACGCTGAACCGCAACTGGTACT
 CGGACGCCGACGTGCCTGCCTCGGCCGCCAGGAGGGGTGCCAGGACATCGCTACGC
 AGCTCATCTCCAACATGGACATTGACGTGATCCTAGGTGGAGGCCGAAAGTACATGT
 TTCCCATGGGAACCCAGACCCTGAGTACCCAGATGACTACAGCCAAGGTGGGACCA
 GGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCGCCAGGGTGGCCGGT
 ATGTGTGGAACCGCACTGAGCTCATGCAGGCTTCCCTGGACCCGTCTGTGACCCATC
 50 TCATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCCACCGAGACTCCACAC
 TGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGAACC
 CCCGCGGCTTCTTCTTCTCGTGAGGGTGGTTCGCATCGACCATGGTCATCATGAAA
 GCAGGGCTTACCGGGCACTGACTGAGACGATCATGTTTCGACGACGCCATTGAGAGGG
 CGGGCCAGCTCACCAGCGAGGAGGACACGCTGAGCCTCGTCACTGCCGACCACTCCC
 ACGTCTTCTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCC
 55 CTGGCAAGGCCCGGGACAGGAAGGCCACACGGTCTCTCTATACGGAAACGGTCCAG
 GCTATGTGCTCAAGGACGGCGCCCGCCGGATGTTACCGAGAGCGAGAGCGGGGAGCC
 CCGATATCGGCAGCAGTACGAGTGCCTGACAGCAAGAGACCCACGCAGGGCAGG
 ACGTGGCGGTGTTTCGCGCGCGGCCCGCAGGCGCAGCTGGTTACGGCGTGCAGGAGC
 AGACCTTCATAGCGCACGTGATGGCCTTCGCCGCCTGCCTGGAGCCCTACACCGCCT
 60 GCGACCTGGCGCCCCCGCCGGCACCACCGACGCCGCGCACCCGGGTTCCGGAGATC
IGATGCCAATGTGGTTCTGTGACCGTGACCTCGAGGTGGACACCACCTCAAGAGCCT
GAGCCAGCAGATCGAGAACATCCGGAGCCCAGAGGGAAGCCGCAAGAACCCCGCCCCG
 CACCTGCCGTGACCTCAAGATGTGCCACTCTGACTGGAAGAGTGGAGAGTACTGGAT
 TGACCCCAACCAAGGCTGCAACCTGGATGCCATCAAAGTCTTCTGCAACATGGAGAC
 65 TGGTGAGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCAGAAAGAACTGGTACAT

CAGCAAGAACCCCAAGGACAAGAGGCATGTCTGGTTCGGCGAGAGCATGACCGATGG
 ATTCCAGTTCGAGTATGGCGGCCAGGGCTCCGACCCTGCCGATGTGGCCATCCAGCT
 GACCTTCCTGCGCCTGATGTCCACCGAGGCCTCCGAGAACATCACCTACCACTGCAA
 5 GAACAGCGTGGCCTACATGGACCAGCAGACTGGCAACCTCAAGAAGGCCCTGCTCCT
 CAAGGGCTCCAACGAGATCGAGATCCGCGCCGAGGGCAACAGCCGCTTCACCTACAG
 CGTCACTGTCGATGGCTGCACGAGTACACCCGGAGCCTGGGGCAAGACAGTGATTGA
 ATACAAAACCACCAAGTCTCCCGCCTGCCCATCATCGATGTGGCCCCCTTGACGT
 TGGTGCCCCAGACCAGGAATTCGGCTTCGACGTTGGCCCTGTCTGCTTCTGTA**AAAC**
 TCCCTCCATCTAGA
 10 XbaI

 <210> 8
 <211> 755
 <212> PROTEÍNA
 15 <213> HOMO SAPIENS
 <400> 8

 1 MLLLLLLLGL RLQLSLGIIP VEEENPDFWN REAAEALGAA KKLQPAQTAA KNLIFLGDG 60
 61 MGVSTVTAAR ILKGQKKDKL GPEIPLAMDR FPYVALSKTY NVDKHVPDSG ATATAYLCGV 120
 121 KGNFQTIGLS AAARFNQCNT TRGNEVISVM NRAKKAGKSV GVVTTTRVQH ASPAGTYAHT 180
 181 VNRNWYSDAD VPASARQEGC QDIATQLISN MDIDVILGGG RKYMIFPMGTP DPEYPDDYSQ 240
 241 GGTRLDGKNL VQBWLAKRQG ARYVWNRTTEL MQASLDPSVT HLMGLFEPGD MKYEIHRDST 300
 301 LDPSLMEMTE AALRLLSRNP RGFFLFVEGG RIDHGHESR AYRALTETIM FDDAIERAGQ 360
 361 LTSEEDTSLV VTADHSHVFS FGGYPLRGSS IFGLAPGKAR DRKAYTVLLY GNGPGYVLKD 420
 25 421 GARPDVTESE SGSPEYRQQS AVPLDEETHA GEDVAVFARG PQAHLVHVQ EQTFIAHVMA 480
 481 FAACLEPYTA CDLAPPAGTT DAAHPGSGRS DANVVRDRDL EVDTTLKSL SQQIENIRSP E 540
 541 GSRKNPARTC RDLKMCHSDW KSGEYWIDPN QGCNLDAIKV FCNMETGETC VYPTQPSVAQ 600
 601 KNWYISKNP KDKRHVWFGE MTDGFQFEYG GQSDPADVA IQLTFLRLMS TEASQNITYH 660
 661 CKNSVAYMDQ QTGNLKKALL LKGSNEIEIR AEGNSRFTYS VTDGCTSHT GAWGKTIVIEY 720
 30 721 KTKSSRLPI IDVAPLDVGA PDQEFGFDVG PVCFL

 <210> 9
 <211> 1734
 <212> ADNc
 35 <213> HOMO SAPIENS
 <220>
 <221> CDS
 <222> 18 1718
 <400> 9

 40 Bam HI
 GGATCCCGCCCCGACCC**ATGGCGCCCGT**CGCCGCTCTGGGCCGCGCTGGCCGTCGGACTGGAG
 CTCTGGGCTGCGGCGCACGCCTTGCCCGCCAGGTGGCATTACACCCTACGCCCGGAGCC
 CGGGAGCACATGCCGGCTCAGAGAATACTATGACCAGACAGCTCAGATGTGCTGCAGCAAAT
 45 GCTCGCCGGGCCAACATGCAAAAGTCTTCTGTACCAAGACCTCGGACACCGTGTGTGACTCC
 TGTGAGGACAGCACATACCCAGCTCTGGAAGTGGGTTCCCGAGTGCTTGAGCTGTGGCTC
 CCGCTGTAGCTCTGACCAGGTGAAACTCAAGCCTGCACTCGGGAACAGAACC GCATCTGCA
 CCTGCAGGCCCCGGCTGGTACTGCGCGCTGAGCAAGCAGGAGGGGTGCCGGCTGTGCGCGCCG
 CTGCGCAAGTGCCGCCCGGGCTTCCGGCGTGCCAGACCAGGAAGTGAACATCAGACGTGGT
 50 GTGCAAGCCCTGTGCCCGGGGACGTTCTCCAACACGACTTCATCCACGGATATTTGCAGGC
 CCCACCAGATCTGTAACGTGGTGGCCATCCCTGGGAATGCAAGCATGGATGCAGTCTGCACG
 TCCACGTCCCCACCCGGAGTATGGCCCCAGGGGCAGTACACTTACCCAGCCAGTGTCCAC
 ACGATCCCAACACACGCAGCCAACTCCAGAACCAGCACTGCTCCAAGCACCTCCTTCTGCTGC
 TCCCAATGGGCCCCAGCCCCCAGCTGAAGGGAGCACTGGATCTAACGGTCTCCCTGGCCC
 55 CATTGGGCCCCCTGGTCTCGCGGTGCGACTGGTGTGCTGGTCTGTTGGTCCCCC
 CGGCCCTCTGGACCTCCTGGTCCCCCTGGTCTCCAGCGCTGGTTTCGACTTCAG
 CTTCTCTCCCCACCTCAAGAGAAGGCTACAGATGGTGGCCGCTACTACCGGGC
 TGATGATGCCAATGTGGTTCGTGACCGTGACCTCGAGGTGGACACCACCCTCAAGAG
 CCTGAGCCAGCAGATCGAGAACATCCGGAGCCAGAGGGAAGCCGCAAGAACCCCGC
 60 CCGCACCTGCCGTGACCTCAAGATGTGCCACTCTGACTGGAAGAGTGGAGAGTACTG
 GATTGACCCCAACCAAGGCTGCAACCTGGATGCCATCAAAGTCTTCTGCAACATGGA
 GACTGGTGAGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCAGAAAGTGGTA
 CATCAGCAAGAACCCCAAGGACAAGAGGCATGTCTGGTTCGGCGAGAGCATGACCGA
 TGATTCCAGTTCGAGTATGGCGGCCAGGGCTCCGACCCCTGCCGATGTGGCCATCCA
 65 GCTGACCTTCTGCGCCTGATGTCCACCGAGGCCTCCAGAACATCACCTACCACTG

CAAGAACAGCGTGGCCTACATGGACCAGCAGACTGGCAACCTCAAGAAGGCCCTGCT
 CCTCAAGGGCTCCAACGAGATCGAGATCCGCGCCGAGGGCAACAGCCGTTACCTA
 CAGCGTCACTGTGATGGCTGCACGAGTCACACCGGAGCCTGGGGCAAGACAGTGAT
 TGAATACAAAACCACCAAGTCCCTCCCGCCTGCCATCATCGATGTGGCCCCCTTGGGA
 5 CGTTGGTGCCCCAGACCAGGAATTCGGCTTCGACGTTGGCCCTGTCTGCTTCTGTA
AACTCCCTCCATCTAGA
 Xba I

10 <210> 10
 <211> 566
 <212> PROTEÍNA
 <213> HOMO SAPIENS
 <400> 10

15 1 MAPVAWVAAL AVGLELWAAA HALPAQVAFT PYAPEPGSTC RLREYYDQTA QMCCSKCSPG 60
 61 QHAKVFCTKT SDTVCDSCED STYTQLWNWV PECLSCGSRC SSDQVETQAC TREQNRITC 120
 121 RPYWYCALSK QEGCRLCAPL RKCRPGFGVA RPYTETSDV CKPCAPGTF NTSSTDICR 180
 181 PHQICNVVAI PGNASMDAVC TSTSPTRMSA PGAVHLPQPV STRSQHTQPT PEPSTAPSTS 240
 20 241 FLLPMGPPSP AEGST**GS**NGL PGP~~IG~~PPGPR GRTGDAGPVG PPGPPGPPG PGPPSAGFDF 300
 301 SFLPQPPQEK AHDGGRYRA DDANVVRDRD LEVDTLKSL SQQIENIRSP EGSRKNPART 360
 361 CRDLKMCHSD WKSGEYWIDP NQGCNLDIAK VFCNMETGET CVYPTQPSVA QKNWYISKNP 420
 421 KDKRHVWFGE SMTDGFQFEY GGQSDPADV AIQLTFLRLM STEASQNTY HCKNSVAYMD 480
 481 QQTGNLKKAL LLKGSNEIEI RAEGNSRFTY SVTVDGCTSH TGAWGKTVE YKTTKSSRLP 540
 541 IIDVAPLDVG APDQEFDFD GPVCF

25 <210>11
 <211> 1542
 <212> ADNc
 30 <213> HOMO SAPIENS
 <220>
 <221> CDS
 <222> 18 1526
 <400> 11

35 Bam HI
GGATCCCGCCCGCACCCATGGCGCCCGTCGCCGCTGGGCCGCGCTGGCCGTCGGACTGGAG
 CTCTGGGCTGCGGCGCACGCCTTGCCCGCCAGGTGGCATTACACCCTACGCCCCGAGCC
 40 CGGGAGCACATGCCGGCTCAGAGAATACTATGACCAGACAGCTCAGATGTGCTGCAGCAAAT
 GCTCGCCGGGCCAACATGCAAAAGTCTTCTGTACCAAGACCTCGGACACCGTGTGTGACTCC
 TGTGAGGACAGCACATACCCAGCTCTGGAAGTGGGTTCCCGAGTGCTTGTGACTGTGGCTC
 CCGCTGAGTCTGACCAGGTGAAACTCAAGCCTGCACTCGGGAACAGAACCAGCATCTGCA
 CCTGCAGGCCCGCTGGTACTGCGCGCTGAGCAAGCAGGAGGGGTGCCGGCTGTGCGCGCCG
 45 CTGCGCAAGTGCCGCCCGGGCTTCGGCGTGGCCAGACCAGGAAGTGAACATCAGACGTGGT
 GTGCAAGCCCTGTGCCCGGGGACGTTCTCCAACACGACTTCATCCACGGATATTTGCAGGC
 CCCACCAGATCTGTAACGTGGTGGCCATCCCTGGGAATGCAAGCATGGATGCAGTCTGCACG
 TCCACGTCCCCACCCGGAGTATGGCCCCAGGGGCAGTACACTTACCCAGCCAGTGTCCAC
 ACGATCCCAACACACGCAGCCAACTCCAGAACCAGCACTGCTCCAAGCACCTCCTTCTGC
 TCCCAATGGGCCCCAGCCCCCAGCTGAAGGGAGCACTGGATCTGATGCCAATGTGGTTCCG
 50 TGACCGTGACCTCGAGGTGACACCACCCTCAAGAGCCTGAGCCAGCAGATCGAGAA
 CATCCGGAGCCAGAGGGAAGCCGCAAGAACCCTGCCCCGACCTGCCGTGACCTCAA
 GATGTGCCACTCTGACTGGAAGAGTGGAGAGTACTGGATTGACCCCAACCAAGGCTG
 CAACCTGGATGCCATCAAAGTCTTCTGCAACATGGAGACTGGTGGAGACCTGCGTGTA
 CCCCAGTCCAGCCAGTGTGGCCAGAAAGTGGTACATCAGCAAGAACCCTCAAGGA
 55 CAAGAGGCATGTCTGGTTCCGGCAGAGCATGACCGATGGATTCCAGTTCCAGTATGG
 CGGCCAGGGCTCCGACCTGCCGATGTGGCCATCCAGCTGACCTTCTGCGCCTGAT
 TGCCACCGAGGCTCCAGAACATCACCTACCCTGCAAGAACAGCGTGGCCTACAT
 GGACCAGCAGACTGGCAACCTCAAGAAGGCCCTGCTCCTCAAGGGCTCCAACGAGAT
 CGAGATCCGCGCCGAGGGCAACAGCCGCTTACCTACAGCGTCACTGTGATGGCTG
 60 CACGAGTACACCCGAGCCTGGGGCAAGACAGTGATTGAATACAAAACCACCAAGTC
 CTCCCGCCTGCCATCATCGATGTGGCCCCCTTGACGTTGGTGGCCAGACCAGGA
 ATTCGGCTTCGACGTTGGCCCTGTCTGCTTCTGTA**AACTCCCTCCATCTAGA**
 Xba I

65 <210> 12

<211> 502
 <212> PROTEÍNA
 <213> HOMO SAPIENS
 <400> 12

5
 1 MAPVAVWAAL AVGLELWAAA HALPAQVAFT PYAPEPGSTC RLREYYDQTA QMCCSKCSPG 60
 61 QHAKVFCTKT SDTVCDSCED STYTQLWNWV PECLSCGSRG SSDQVETQAC TREQNRICR 120
 121 RPYWYCALSK QEGCRLCAPL RKCRPGFGVA RPYTETS DVV CKPCAPGTFS NTSSTDICR 180
 181 PHQICNVVAI PGNASMDAVC TSTSPTRMSA PGAVHLPQPV STRSQHTQPT PEPSTAPSTS 240
 241 FLLPMGSPSP AEGSTGSDAN VVRDRDLEVD TTLKLSLQI ENIRSPESGR KNPARTCRDL 300
 301 KMCHSDWKSG EYWIDPNQGC NLDAIKVFCN METGETCVYP TQPSVAQKNW YISKPNPKDKR 360
 361 HVWFGBSMTD GFQFEYOGQG SDPADVAIQL TFLRLMSTEA SQNITYHCKN SVAYMDQQTG 420
 421 NLKKALLLKG SNEIEIRAEG NSRFTYSVTV DGCTSHTGAW GKTVIEYKTT KSSRLPIIDV 480
 481 APLDVGAPDQ EFGFDVGPVC FL

15
 <210> 13
 <211>2139
 <212> ADNc
 <213> HOMO SAPIENS
 20
 <220>
 <221> CDS
 <222> 24 2123
 <400> 13

25 Hind III
AAGCTTCCCTCGGCAAGGCCACAATGAACCGGGGAGTCCCTTTTAGGCACTTGCTTC
 TGGTGTGCAACTGGCGCTCCTCCCAGCAGCCACTCAGGGAAAGAAAGTGGTGCTGG
 GCAAAAAAGGGGATACAGTGGAACTGACCTGTACAGCTTCCCAGAAGAAGAGCATA
 AATTCCACTGGAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCT
 30 TAACTAAAGGTCCATCCAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCCTTTGGG
 ACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTT
 ACATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTGCTAGTGTTCCGGATTGA
 CTGCCAACTCTGACACCCACCTGCTTCCAGGGGAGAGCCTGACCCTGACCTTGGAGA
 GCCCCCCTGGTAGTAGCCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAACAAACATA
 35 AGGGGGGGAAGACCCCTCCTCGTGTCTCAGCTGGAGCTCCAGGATAGTGGCACCTGGA
 CATGCCTGTCTTGCAGAACCAGAAAGGTGGAGTTCAAATAGACATCGTGGTGC
 TAGCTTTCCAGAAGGCCTCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGT
 TCTCCTTCCCCTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGT
 40 GGCAGGCGGAGAGGGCTTCTCCTCCAAGTCTTGGATCACCTTTGACCTGAAGAACA
 AGGAAGTGTCTGTAACCGGGTTACCCAGGACCCTAAGCTCCAGATGGGCAAGAAGC
 TCCCCTCCACCTCACCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACC
 TCACCCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCATCAGGAAGTGAACCTGGTGG
 TGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCT
 45 CCCCTAAGCTGATGCTGAGCTTGAACCTGGAGAACAAGGAGGCAAAGGTCTCGAAGC
 GGGAGAAGGCGGTGTGGGTGCTGAACCTGAGGCGGGGATGTGGCAGTGTCTGCTGA
 GTGACTCGGGACAGGTCTGCTGGAATCCAACATCAAGTTCTGCCAGATCTAACG
 GTCTCCCTGGCCCCATTGGGCCCTGGTCTCGCGGTGCGACTGGTGATGCTGGTC
 CTGTTGGTCCCCCGGCCCTCCTGGACCTCCTGGTCCCCCTGGTCTCCAGCGCTG
 50 GTTTCGACTTCAGCTTCTGCCCCAGCCACCTCAAGAGAAGGCTCACGATGGTGGCC
 GCTACTACCGGGCTGATGATGCCAATGTGGTTCGTGACCGTGACCTCGAGGTGGACA
 CCACCTCAAGAGCCTGAGCCAGCAGATCGAGAACATCCGGAGCCCAGAGGGAAGCC
 GCAAGAACCCCGCCCGCACCTGCCGTGACCTCAAGATGTGCCACTCTGACTGGAAGA
 GTGGAGAGTACTGGATTGACCCCAACCAAGGCTGCAACCTGGATGCCATCAAAGTCT
 55 TCTGCAACATGGAGACTGGTGAGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCC
 AGAAGAAGTGGTACATCAGCAAGAACCCCAAGGACAAGAGGCATGTCTGGTTCCGGC
 AGAGCATGACCGATGGATTCCAGTTCGAGTATGGCGGCCAGGGCTCCGACCCCTGCC
 ATGTGGCCACTCCAGCTTCTGCGCCTGATGTCCACCGAGGCCTCCAGAACA
 TCACCTACCACTGCAAGAACAGCGTGGCCTACATGGACCAGCAGACTGGCAACCTCA
 60 AGAAGGCCCTGCTCCTCAAGGGCTCCAACGAGATCGAGATCCGCGCCGAGGGCAACA
 GCCGCTTACCTACAGCGTCACTGTGATGGCTGCACGAGTACACCGGAGCCTGGG
 GCAAGACAGTGATTGAATACAAAACCAAGTCTCCCGCCTGCCATCATCGATG
 TGGCCCCCTTGGACGTTGGTGGCCCCAGACCAGGAATTCGGCTTCGACGTTGGCCCTG
 TCTGCTTCTGTAAACTCCCTCCATCTAGA
 Xba I

<210> 14
 <211> 699
 <212> PROTEÍNA
 <213> HOMO SAPIENS
 5 <400> 14

1 MNRGVPFRHL LLVLQLALLP AATQGGKVVV GKKGDTVELT CTASQKKSQ FHWKNSNIQ 60
 61 ILGNQGSFLT KGPSKLNDR DSRRLWDQG NFPLIKNLK IEDSDTYICE VEDQKEEVQL 120
 121 LVFGLTANS THLLQGQSLT LTLBSPGGSS PSVQCRSPRG KNIQGGKTL VSQLELQDSG 180
 10 181 TWTCTVLQNN KKVEFKIDIV VLAFAQKASSI VYKKEGEQVE FSFPLAFTVE KLTGSGELWW 240
 241 QAERASSSKS WITFDLKNKE VSVKRVQDP KLQMGKKLPL HLTLPQALPQ YAGSGNLTLA 300
 301 LEAKTGKLNHQ EVNLVVMRAT QLQKNLTCEV WGPTSPKMLL SLKLENKEAK VSKREKAVVW 360
 361 LNPEAGMWQC LLSDSGQVLL ESNIKVLPR**S** NGLPGPIGPP GPRGRTGDAG PVGPPGPPGP 420
 421 PGPPGPPSAG FDSFLPQPP QEKAHDGGRY YRADDANVVR DRDLEVDTTL KSLSQQIENI 480
 15 481 RSPEGSRKNP ARTCRDLKMC HSDWKSGEYW IDPNQGCNLD AIKVFCNMET GETCVYPTQP 540
 541 SVAQKNWYIS KNPKDKRHVW FGESMTDGFQ FEYGGQSDP ADVAIQLTFL RLMSTEASQN 600
 601 ITYHCKNSVA YMDQQTGNLK KALLLKGSNE IEIRAEGNSR FTYSVTVDGC TSHTGAWGKT 660
 661 VIEYKTTKSS RLPIIDVAPL DVGAPDQEFG FDVGPVCFL

<210> 15
 <211> 1947
 <212> ADNc
 <213> HOMO SAPIENS
 <220>
 <221> CDS
 25 <222> 24 1931
 <400> 15

Hind III
 30 AAGCTTCCCTCGGCAAGGCCACAATGAACCGGGAGTCCCTTTTAGGCACTTGCTTC
TGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTCAGGGAAAGAAAGTGGTGCTGG
GCAAAAAAGGGGATACAGTGGAACTGACCTGTACAGCTTCCCAGAAGAAGAGCATAC
AATCCACTGGAAAACTCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCT
TAACTAAAGGTCCATCCAAGCTGAATGATCGCGCTGACTCAAGAAGAAGCCTTTGGG
 35 ACCAAGGAAACTTTCCCTGATCATCAAGAATCTTAAGATAGAAGACTCAGATACTT
ACATCTGTGAAGTGGAGGACCAGAAGGAGGTGCAATTGCTAGTGTTCCGATTGA
CTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACCCTGACCTTGGAGA
GCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAAAAACATAC
AGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGCACCTGGA
 40 CATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTGGTGC
TAGCTTTCCAGAAGGCCTCCAGCATAGTCTATAAGAAAGAGGGGGAACAGGTGGAGT
TCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGT
GGCAGGCGGAGAGGGCTTCCCTCCTCCAAGTCTTGATCACCTTTGACCTGAAGAACA
AGGAATGTCTGTAAAAACGGGTTACCCAGGACCCTAAGCTCCAGATGGGCAAGAAGC
 45 TCCCGCTCCACCTCACCTGCCCCAGGCCTTGCCTAGTATGCTGGCTCTGGAAACC
TCACCCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCATCAGGAAGTGAACCTGGTGG
TGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCT
CCCTAAGCTGATGCTGAGCTTGAAACTGGAGAACAAGGAGGCAAAGGTCTCGAAGC
GGGGAAAGGCGGTGTGGGTGCTGAACCTGAGGCGGGGATGTGGCAGTGTCTGCTGA
GTACTCGGGACAGGTCTGCTGGAATCCAACATCAAGGTTCTGCCAGATCTGATG
 50 CCAATGTGGTTCGTGACCGTGACCTCGAGGTGGACACCCACCTCAAGAGCCTGAGCC
AGCAGATCGAGAACATCCGGAGCCCAGAGGGAAGCCGCAAGAACCCCGCCCACCT
GCCGTGACCTCAAGATGTGCCACTCTGACTGGAAGAGTGGAGAGTACTGGATTGACC
CCAACCAAGGCTGCAACCTGGATGCCATCAAAGTCTTCTGCAACATGGAGACTGGTG
AGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCCAGAAGAACTGGTACATCAGCA
 55 AGAACCCCAAGGACAAGAGGCATGTCTGGTTCCGGCGAGAGCATGACCGATGGATTCC
AGTTCGAGTATGGCGGCCAGGGCTCCGACCCTGCCGATGTGGCCATCCAGCTGACCT
TCTCGCCTGATGTGTCACCGAGCCTCCCAAGACATCACCTACCACTGCAAGAACA
GCGTGGCCTACATGGACCAGCAGACTGGCAACCTCAAGAAGGCCCTGCTCCTCAAGG
GCTCCAACGAGATCGAGATCCGCGCCGAGGGCAACAGCCGCTTCACTACAGCGTCA
 60 CTGTCGATGGCTGCACGAGTACACCCGGAGCCTGGGGCAAGACAGTATTGAATACA
AAACCACCAAGTCCCTCCCGCCTGCCATCATCGATGTGGCCCCCTTGGACGTTGGTG
CCCCAGACCAGGAATTCGGCTTCGACGTTGGCCCTGTCTGCTTCCCTGTAACTCCCT
CCATCTAGA

Xba I
 65

ES 2 433 127 T3

<210> 16
<211> 635
<212> PROTEÍNA
<213> HOMO SAPIENS
<400> 16

5

10

15

1 MNRGVPFRHL LLVLQLALLP AATQGKQVVL GKKGDTVELT CTASQKKSQ FHWKNSNIQ 60
61 ILGNQGSFLT KGPSKLNDRS DSRRLWDQDQ NFPLIKNLK IEDSDTYICE VEDQKEEVQL 120
121 LVFGLTANS DTHLLQGQSLT LTLESPPGSS PSVQCRSPRG KNIQGGKTL S VSQLELQDSG 180
181 TWTCTVLQNN KKVEFKIDIV VLAFAQKASSI VYKKEGEQVE FSFPLAFTVE KLTGSGELWW 240
241 QAERASSSKS WITFDLKNKE VSVKRVTDQP KLQMGKKLPL HLTLPQALPQ YAGSGNLTLA 300
301 LEAKTGKLNQ EVNLVVMRAT QLQKNLTCEV WGPTSPKMLL SLKLENKEAK VSKREKAVWV 360
361 LNPEAGMWQC LLSDSGQVLL ESNIKVLP**RS** DANVVRDRDL EVDTTLSLS QQIENIRSPE 420
421 GSRKNPARTC RDLKMHSDW KSGEYWDPN QGCNLDAIKV FCNMETGETC VYPTQPSVAQ 480
481 KNWYISKNP DKRHVWFGES MTDGFQFEYG GQSDPADVA IQLTFLRLMS TEASQNITYH 540
541 CKNSVAYMDQ QTGNLKKALL LKGSNEIEIR AEGNSRFTYS VTDGCTSH T GAWGKTVIEY 600
601 KTKSSRLPI IDVAPLDVGA PDQEFQFDVG PVCFL

REIVINDICACIONES

1. Procedimiento para generar una proteína de fusión trimérica secretada, que implica:
 - 5 (a) crear un constructo de ADN (AND recombinado) que incluya un promotor de transcripción ligado a una matriz que codifique una secuencia de péptido señal seguido de una unión en fase en un polipéptido no colágeno que va a ser trimerizado, el cual es un polipéptido biológicamente activo o un receptor soluble que consiste de un/unos dominio/s de unión al ligando, y que a su vez está unido en fase a una porción C-terminal de procolágeno que es capaz de autotrimerizarse en una forma soluble con enlaces intermoleculares disulfuro unidos fuertemente de forma covalente;
 - 10 (b) introducir dicho constructo de ADN en una célula eucariota;
 - (c) cultivar dicha célula huésped bajo condiciones fisiológicas en un medio de crecimiento adecuado para permitir la secreción de una proteína de fusión trimerizada soluble codificada por dicha secuencia de ADN; y
 - 15 (d) aislar dicha proteína de fusión trimerizada del medio de cultivo de dicha célula huésped.
2. Procedimiento conforme a la reivindicación 1 en el que la proteína de fusión polipeptídica trimerizada es un homotrímero soluble.
3. Procedimiento conforme a la reivindicación 1 en el que el residuo de polipéptido trimerizado comprende la porción C-terminal de colágeno que es capaz de auto-ensamblarse en un trímero soluble seleccionado de un grupo que consiste de pro.alfa.1 (I), pro.alfa.2 (I), pro.alfa.1 (II), pro.alfa.1 (III), pro.alfa.1 (V), pro.alfa.2 (V), pro.alfa.1 (XI), pro.alfa.2 (XI) y pro.alfa.3 (XI).
4. Procedimiento conforme a las reivindicaciones 1 a 3, en el cual la secuencia de péptido señal y el polipéptido no colágeno que va a ser trimerizado proceden de la misma proteína secretada nativa.
5. Procedimiento conforme a las reivindicaciones 1 a 4, en el cual tanto la secuencia de péptido señal como el polipéptido no colágeno que va a ser trimerizado se seleccionan a partir de dos proteínas secretadas distintas.
6. Procedimiento conforme a la reivindicación 1, en el cual la célula eucariota huésped es una célula procedente de hongos o de insectos.
7. Procedimiento conforme a la reivindicación 1, en el cual la célula eucariota huésped es una línea celular de mamíferos cultivada.
8. Procedimiento conforme a las reivindicaciones 1 a 3, en el cual la porción C-terminal del procolágeno incluye una región de colágeno de triple hélice de "glicina - repetida" unida al C-propéptido.
9. Procedimiento conforme a la reivindicación 8, en el cual la porción C-terminal de procolágeno está identificada con los números de ID de secuencia 1-2.
10. Procedimiento conforme a las reivindicaciones 1 a 3, en el cual la porción C-terminal de trimerización de procolágeno comprende únicamente un C-propéptido sin una región de colágeno de triple hélice de "glicina - repetida".
11. Procedimiento conforme a las reivindicaciones 8 a 10, en el cual la porción C-terminal de trimerización del procolágeno comprende una secuencia de reconocimiento de la proteasa BMP-1 mutada o suprimida, confiriendo de este modo a las proteínas de fusión triméricas resistencia a la degradación por las proteasas BMP -1.
12. Procedimiento conforme a las reivindicaciones 10 a 11, en el cual la porción C-terminal de trimerización de procolágeno está identificada por los números de ID de secuencia 3-4.
13. Composiciones de proteínas de fusión secretadas generadas por los procedimientos de las reivindicaciones 1, 2, 3, 8, 9, 10, 11 y 12, siendo el receptor soluble trimérico humano TNF- α II (p75) identificado por los números de ID de secuencia 9-12.
14. Composiciones de proteínas de fusión secretadas generadas por los procedimientos de las reivindicaciones 1, 2, 3, 8, 9, 10, 11 y 12, siendo los CD4 triméricos humanos solubles identificados por los números de ID de secuencia 13-16.
- 60 15. Composiciones de proteínas de fusión secretadas generadas por los procedimientos de las reivindicaciones 1, 2, 3, 8, 9, 10, 11 y 12, siendo la fosfatasa alcalina de placenta humana trimérica soluble identificada por los números de ID de secuencia 5-8.

5

16. Receptor soluble TNF- α II trimerizado que comprende tres cadenas polipeptídicas, incluyendo cada una de dichas cadenas polipeptídicas un dominio de unión al ligando del receptor, unido a un C-propéptido de colágeno, en cuyo caso la trimerización de la fusión polipeptídica resulta en la mejora de la actividad biológica hacia TNF- α .

17. Producto o composición para ser usada en un tratamiento médico mediante el bloqueo de la actividad biológica del TNF- α usando un receptor soluble TNF- α II trimerizado generado conforme a la reivindicación 13.

Vista lateral

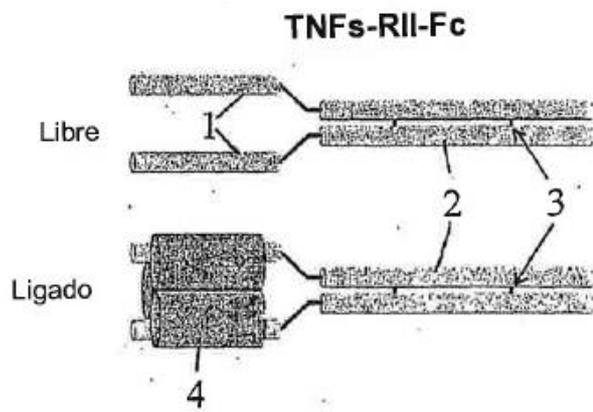


FIG. 1A

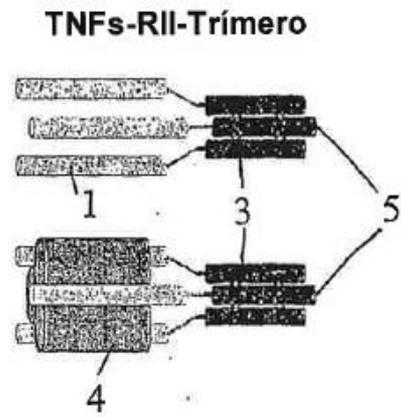


FIG. 1C

Vista en planta

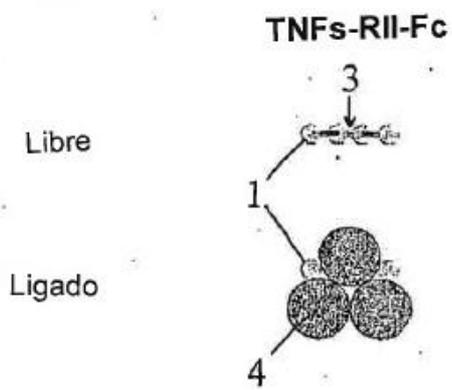


FIG. 1B

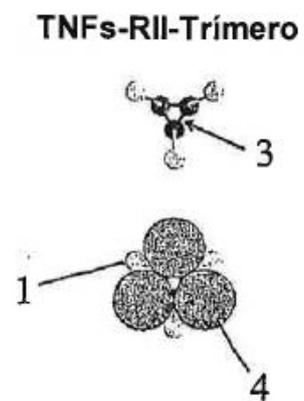


FIG. 1D

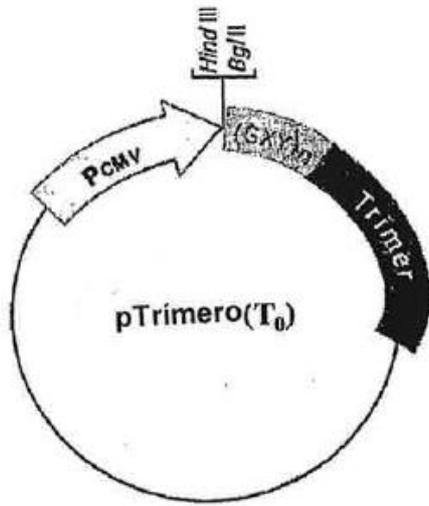


FIG. 2A

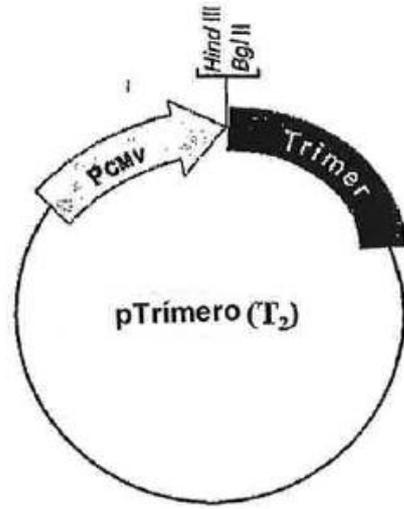


FIG. 2B

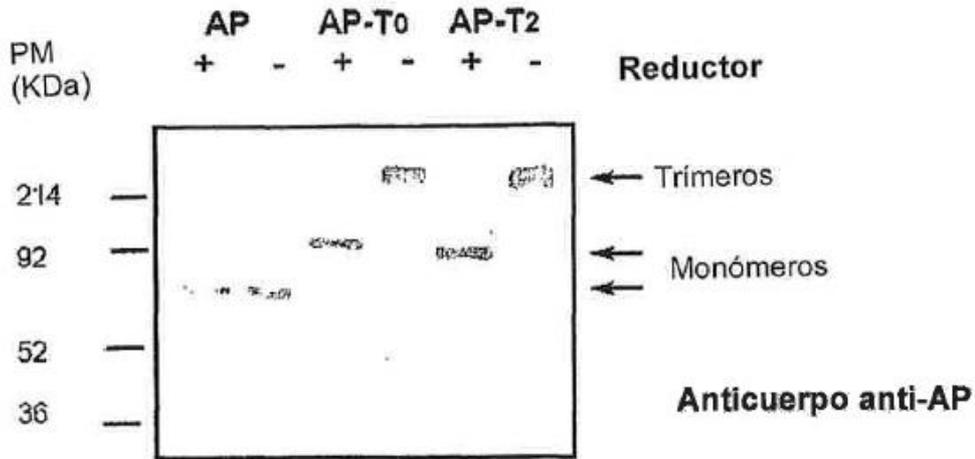


FIG. 3A

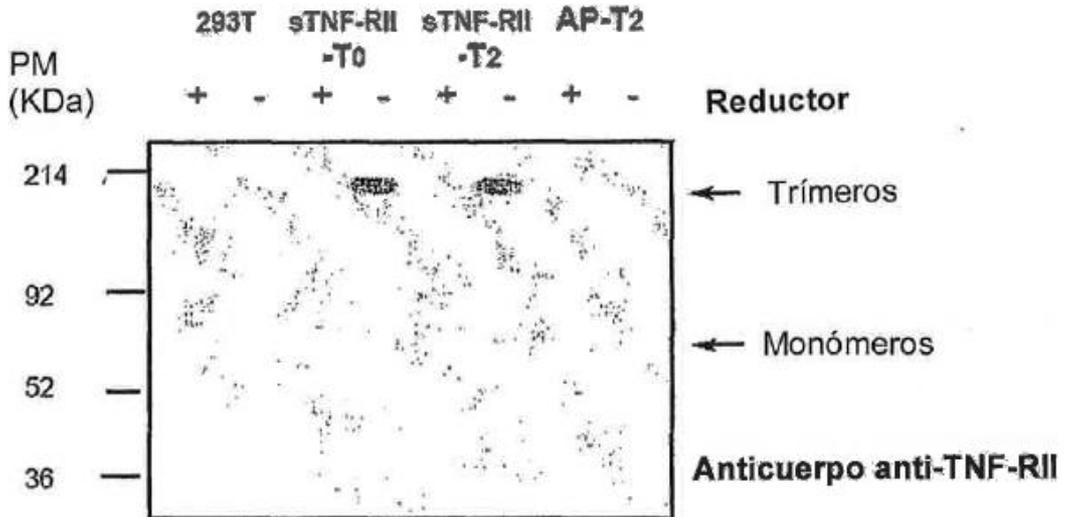


FIG. 3B

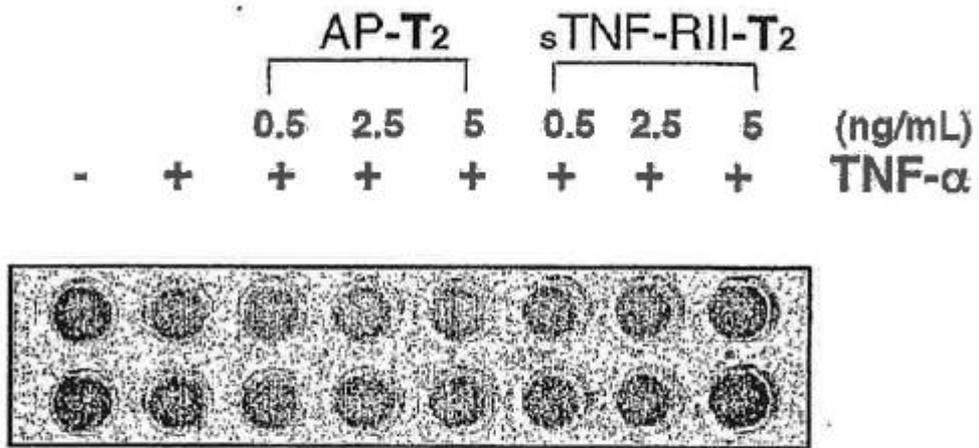


FIG. 4

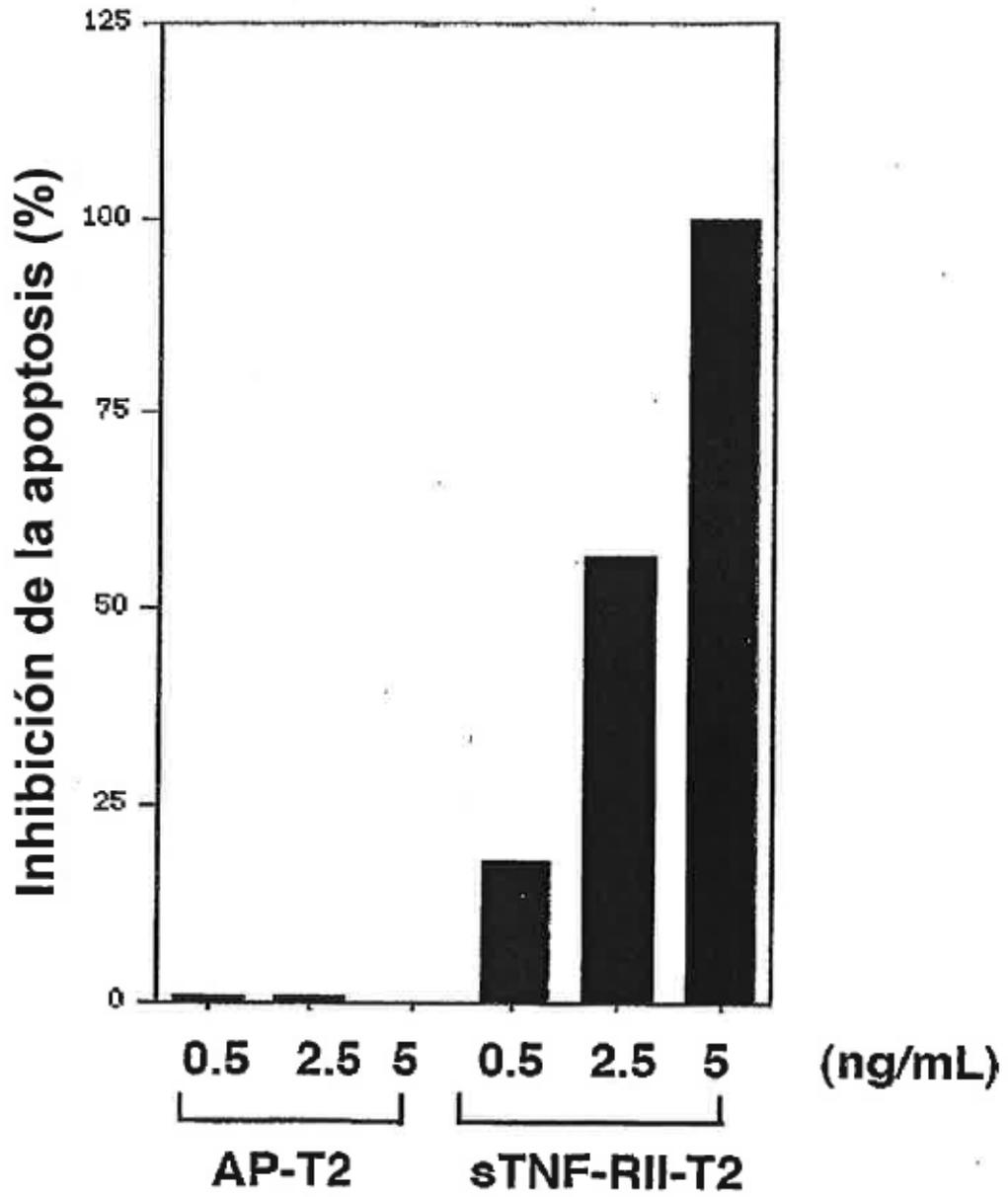


FIG. 5

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

5 La lista de referencias citada por el solicitante lo es solamente para utilidad del lector, no formando parte de los documentos de patente europeos. Aún cuando las referencias han sido cuidadosamente recopiladas, no pueden excluirse errores u omisiones y la OEP rechaza toda responsabilidad a este respecto.

10 Documentos de patente citado en la descripción

- EP 0985732 A2 [0009]
- WO 02090553 A2 [0010]
- US 10677877 B [0038]

Bibliografía de patentes citada en la descripción

- MOHLER et al. *J Immunol.*, 1993, vol. 151 (3), 1548-61 [0008]
 - HAYASHI et al. *Biochim Biophys Acta*, 2001, vol. 1528, 187-95 [0011]
 - KISHORE et al. *J. Immunol*, 2001, vol. 166, 559-565 [0012]
 - HOLLER et al. *J Immunol Methods*, 2000, vol. 237 (1-2), 159-73 [0013]
- 15