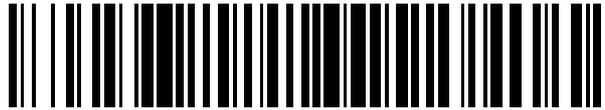


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 415 890**

51 Int. Cl.:

C12N 15/63 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **11.09.2006 E 06803300 (0)**

97 Fecha y número de publicación de la concesión europea: **24.04.2013 EP 1960525**

54 Título: **Procedimiento de preparación de proteínas triméricas**

30 Prioridad:

09.09.2005 US 715883 P

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

29.07.2013

73 Titular/es:

**ZYMOGENETICS, INC. (100.0%)
1201 Eastlake Avenue East
Seattle, WA 98102, US**

72 Inventor/es:

**MOORE, MARGARET, D. y
FOX, BRIAN, A.**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 415 890 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

Procedimiento de preparación de proteínas triméricas.

La presente memoria descriptiva describe polipéptidos capaces de formar multímeros, particularmente trímeros y la fabricación y uso de tales polipéptidos.

5 A. Bobinas en espiral

Un componente básico de la estructura cuaternaria de los presentes polipéptidos multimerizantes es la bobina en espiral (revisado en Müller y col., (2000) Meth. Enzymol. 328: 261-283). Las bobinas en espiral son dominios de proteína que toman la forma de haces similares a sogas suavemente retorcidas. Los haces contienen dos a cinco hélices en orientación paralela o antiparalela. La característica esencial de las secuencias de bobina en espiral es un repetición de siete residuos, o héptada (**abcdefg**)_n con la primera (a) y cuarta (d) posiciones comúnmente ocupadas por aminoácidos hidrófobos. Los restantes aminoácidos de la estructura de bobina en espiral son generalmente polares, excluyéndose normalmente la prolina debido a su efecto perturbador sobre la arquitectura helicoidal.

Esta repetición de héptada característica (también conocida como una repetición 3,4 hidrófoba) es lo que forma la estructura del dominio de la bobina en espiral, barriendo cada residuo aproximadamente 100°. Esto hace que los siete residuos de la repetición de héptada no alcancen dos giros completos de 20°. El retraso forma una ligera tira hidrófoba hacia la izquierda de residuos que transcurren hacia abajo en una hélice y la bobina en espiral se forma cuando se asocian estas tiras hidrófobas. Las desviaciones de la separación 3,4 regular de residuos apolares cambia el ángulo de la tira hidrófoba con respecto al eje de la hélice, alterando el ángulo de cruce de las hélices y desestabilizando la estructura cuaternaria. Un ejemplo de un medio común de representar en diagrama las repeticiones de héptada de bobina en espiral se ilustra en la Figura 1.

Los dímeros y trímeros paralelos son las estructuras de bobina en espiral observadas más comunes. Las características que distinguen dímeros de trímeros de formaciones de oligomerización de mayor orden son relativamente bien entendidas. Los residuos centrales de la repetición de héptada (residuos **a** y **d**) determinan en gran medida el estado de oligomerización, mientras que los residuos en el borde de la hélice (**e** y **g**) desempeñan funciones secundarias. Los trímeros son la organización por defecto para una distribución al azar de residuos centrales, ya que otros estados de oligomerización no pueden tolerar aminoácidos β-ramificados (Ile, Val y Thr) en la posición **d** para dímeros o la posición **a** para tetrámeros. A diferencia, los trímeros generalmente permiten la presencia de aminoácidos β-ramificados y otros hidrófobos en las posiciones centrales.

Se han usado fusiones de bobinas en espiral para lograr diversos fines experimentales. Un uso común es la sustitución de dominios de oligomerización naturales con una secuencia heteróloga para alterar el estado de oligomerización, estabilidad y/o avidéz. Los monómeros de baja afinidad que no se asocian naturalmente pueden oligomerizarse con el fin de unirse efectivamente a otras dianas multiméricas. Adicionalmente, la fusión de dominios de oligomerización puede usarse para imitar el estado activado de la proteína nativa, que es difícil de lograr con la producción de proteínas recombinantes (véase, por ejemplo, Pullen y col. (1999) Biochem. 94:6032). Este enfoque ha sido particularmente eficaz cuando solo se producen dominios específicos tales como la porción extracelular (citoplásmica) o intracelular de una proteína de interés. Comúnmente, las bobinas en espiral se fusionan genéticamente a la proteína de interés mediante un adaptador flexible que proporcionará acceso para la fusión a un gran espacio tridimensional. Se usan fusiones directas para fines experimentales que requieren más moléculas rígidas, tales como aquellas usadas para la cristalización.

Se han desarrollado varios sistemas modelo de bobina en espiral basándose en la información estructural de grandes proteínas estructural tales como miosina y tropomiosina (TM43, Lau y col. J Biol Chem; 259: 13253-13261), un grupo de proteínas conocido como colectinas (Hoppe y col. (1994) Protein Sci; 3:1143-1158), o de la región de dimerización de proteínas reguladoras de ADN tales como la proteína del activador transcripcional de levadura GCN4-p1 (Landschulz y col. (1988) Science; 240:1759-1764). Esta última estructura se denomina frecuentemente una "cremallera de leucina" o LZ. Se han preparado sistemas modelo derivados de TM43, específicamente en los que una leucina por héptada se ha cambiado a fenilalanina. Esta estructura se conoce como una "cremallera de fenilalanina" o FZ (Thomas y col. Prog Colloid Polymer Sci; 99: 24-30). Un tercer tipo de derivado muy conocido de LZ es la cremallera de isoleucina (IZ) (Harbury y col. (1994) Nature 371:80-83).

Un concepto importante de bobinas en espiral modelo es la capacidad para producirse en el huésped de expresión. La falta de enlaces disulfuro en las estructuras de bobina en espiral ayuda en su producción en sistemas de expresión heterólogos. Sin embargo, las secuencias diseñadas *de novo* tienden a ser sensibles a la proteólisis. Aunque se expresan eficazmente, la falta relativa de eficacia con respecto a secuencias naturales refleja las lagunas en el presente conocimiento sobre todas las variables que participan en la interacción de proteínas (Arndt y col. (2002) Structure 10: 1235-1248). Adicionalmente, el uso de secuencias modelo es problemático cuando el fin de la proteína de fusión producida es una proteína biológicamente funcional.

B. Repeticiones de héptada virales

Muchos virus producen una forma fusogénica de glicoproteínas de la envoltura viral. Entre los géneros o familias virales que presentan este tipo de proteínas de fusión están *Orthomyxovirus* (por ejemplo, virus de la gripe), *Filovirus* (virus del Ébola), *Betaretrovirus* (virus Mason-Pfizer del mono (VMMP)), *Gammaretrovirus* (virus de la leucemia murina de Friend (FRMLV); virus de la leucemia murina de Moloney (MoMLV)), *Deltaretrovirus* (virus de la leucemia de linfocitos T humana tipo 1 (HTLV-1)) y *Lentivirus* (virus de la inmunodeficiencia humana tipo 1 (VIH-1) y de la inmunodeficiencia simia (VIS)) (revisado en Cheynet y col. (2005) J Virol; 79: 5585-5593). En estos virus, la subunidad de transmembrana se produce como una proteína de fusión, codificada por el gen *env*. Este producto génico se escinde en dos proteínas, la proteína de la superficie que participa en el reconocimiento de receptores y la subunidad de transmembrana que ancla el complejo *env* entero a la membrana y participa en la entrada del virus por la fusión de membranas. Estas proteínas se caracterizan por la presencia de repeticiones de héptada dentro de la región TM que forman fuertes interacciones entre oligómeros de la proteína mediante la formación de una estructura de bobina en espiral con tres subunidades (Li y col. (1996) J Virol; 70: 1266-1270; Tucker y col. (1991) Virol; 185:710-720).

Se han determinado estructuras cristalinas de varias proteínas TM: MoMLV (Fass y col. (1996) Nat Struc Biol; 3:465-469); VIH-1 (Chan y col. (1997) Cell; 89: 263-273); HTLV-1 (Kobe y col. (1999) PNAS; 93:4319-4324). También se encontraron estructuras similares en el virus de la gripe (Wilson y col. (1981) Nature; 289: 366-373) y el virus del Ébola (Malashkevich y col. (1999) PNAS; 96:2662-2667) y se ha supuesto que esta estructura refleja un mecanismo común para el proceso de fusión y entrada viral (Chambers y col. J Gen Virol; 71: 3075-3080). Otros virus que incluyen repeticiones de héptada dentro de su genoma son *Cornavirus* (cornavirus asociado al síndrome respiratorio agudo grave (SARS-CoV)) (Bosch y col. J Virol 77: 8801-8811); virus del herpes (virus del herpes simple 1 (VHS-1)) (Gianni y col. (2005) J Virol 79:7042-7049) y citomegalovirus humano (CMV)) (Lopper y col. J Virol (2004) 78: 8333-8341); y *Paramyxovirus* (virus del sarampión) (Buckland y col. (1992) J Gen Virol 73:1703-1707).

La homología de secuencias en la región TM también ha permitido la identificación de retrovirus endógenos en bases de datos de secuencias. Se han realizado estas búsquedas y han identificado satisfactoriamente retrovirus endógenos en muchos genomas de organismos, que incluyen ser humano, rata y ratón. Un ejemplo muy conocido de una familia de retrovirus endógenos humanos (RVEH) es RVEH-W. Se ha mostrado que un sitio de esta familia, ERVWE1, codifica un marco de lectura abierto de *env* de longitud completa y produce una proteína también conocida como sincitina (Cheynet y col. (2005) J Virol; 79:5585-5593). Al igual que la proteína viral, esta proteína se produce y escinde en dos subunidades separadas, una subunidad de superficie gp50 y una subunidad transmembrana gp24. La subunidad gp24 incluye repeticiones de héptada y las subunidades se encuentran asociadas como homotrímeros. De forma interesante, esta proteína se produce naturalmente en la placenta y puede participar en interacciones célula-célula tales como la fusión de la placenta con la pared uterina.

Sigue existiendo la necesidad en la técnica de adaptar las secuencias de trimerización naturales a su uso en la producción de proteínas de fusión recombinantes biológicamente activas. Por consiguiente, la presente solicitud describe el cribado, descubrimiento y desarrollo de secuencias genéticas naturales apropiadas para la trimerización en la técnica de las proteínas recombinantes.

El documento WO 2004/033486 describe la producción de proteínas de fusión homotriméricas.

Chambers y col., J. General Virology, vol. 71(Parte 12), diciembre de 1990, páginas 3075-3080, informa de que las secuencias de repetición de héptada están localizadas adyacentes a regiones hidrófobas en varios tipos de glicoproteínas de fusión de virus.

Delwart y col., AIDS Research and Human Retroviruses, vol. 6(6) junio de 1990, páginas 703-706, informa de que las glicoproteínas de la envoltura retroviral contienen una repetición similar a cremallera de leucina.

Shugars y col., J. Virology, vol. 70(5) 1996, páginas 2982-2991, describen la caracterización biofísica de proteínas recombinantes que expresan el dominio similar a cremallera de leucina de la proteína transmembrana del VIH tipo 1 pg41.

RESUMEN DE LA INVENCION

La presente invención se refiere a un procedimiento de preparación de una proteína trimérica que comprende cultivar una célula huésped transformada o transfectada con un vector de expresión que codifica una proteína de fusión que comprende un dominio ZymoZipper (ZZ) como se define en las reivindicaciones y una proteína heteróloga. En una realización, la proteína heteróloga es una proteína de la membrana, la porción de la proteína heteróloga que está incluida en la proteína de fusión es el dominio extracelular de esa proteína, y la proteína de fusión resultante es soluble. En otra realización de la presente invención, el dominio ZZ se deriva de la subunidad transmembrana (TM) de una proteína de la envuelta del virus u otro gen que contiene repetición de héptada de un genoma de virus. En una realización, la proteína de la envuelta del virus es del virus Mason-Pfizer del mono (VMMP).

La proteína de fusión comprende una secuencia de adaptador como se define en las reivindicaciones. Una realización adicional es en la que el dominio ZZ se deriva de la subunidad transmembrana (TM) de un retrovirus endógeno, tal como un retrovirus endógeno de rata (RVER) o un retrovirus endógeno humano (RVEH) que incluye, pero no se limita a, ERVWE1.

5 Todavía otra realización de la presente invención es un procedimiento como se define en las reivindicaciones de preparación de una proteína hetero-trimérica soluble cultivando una primera célula huésped transformada o transfectada con un primer vector de expresión que codifica una proteína de fusión que comprende un primer dominio ZZ y una primera proteína heteróloga; cultivando una segunda célula huésped transformada o transfectada con un segundo vector de expresión que codifica un segundo dominio ZZ y una segunda proteína heteróloga, y cultivando una tercera célula huésped transformada o transfectada con un tercer vector de expresión que codifica un tercer dominio ZZ y una tercera proteína heteróloga. En esta realización, los tres dominios ZZ forman preferencialmente un hetero-trímero. Este cultivo puede producirse en la misma célula huésped. Los dominios ZZ pueden ser iguales o diferentes. La proteína de fusión comprende además una secuencia de adaptador. En una realización particular, las dos proteínas heterólogas usadas para formar la proteína hetero-trimérica son los dominios solubles de APRIL (un ligando de proliferación de linfocitos) y BLYS (estimulante de linfocitos B). La presente memoria descriptiva describe secuencias de ADN, vectores de expresión y células huésped transformadas usadas en el presente procedimiento y proteínas de fusión producidas por el presente procedimiento.

Estos y otros aspectos de la invención serán evidentes para aquellos expertos en la técnica tras la lectura de los detalles de la invención como se describen más completamente más adelante.

20 *BREVE DESCRIPCIÓN DE LOS DIBUJOS*

La FIG. 1. es una representación gráfica de la estructura de proteínas de bobina en espiral y la interacción entre residuos dentro de la bobina y los residuos entre bobinas.

La FIG. 2. representa en gráfica los resultados de las pruebas de estimulación de linfocitos B para construcciones de ZZ-APRIL representativas a diversas concentraciones. BLYS es un control positivo.

25 La FIG. 3. representa en gráfica los resultados del ensayo de proliferación de linfocitos B para construcciones de ZZ12-APRIL representativas con diversos adaptadores. BLYS es un control positivo.

La FIG. 4. representa en gráfica los resultados del bioensayo de TACI-Jurkat para construcciones de ZZ12-APRIL representativas con diversos adaptadores. BLYS es un control positivo.

30 La FIG. 5. representa en gráfica los resultados del ensayo de proliferación de linfocitos B para la construcción de ZZ12-6 APRIL y varios APRIL comercialmente disponibles. BLYS es un control positivo.

DESCRIPCIÓN DETALLADA DE LA INVENCION

La presente invención proporciona un procedimiento de preparación de una proteína trimérica cultivando una célula huésped transformada o transfectada con un vector de expresión que codifica una proteína de fusión que comprende un dominio ZymoZipper (ZZ) como se define en las reivindicaciones, una proteína heteróloga y una secuencia de adaptador como se define en las reivindicaciones. La invención se basa en el hallazgo de que secuencias de trimerización derivadas de la subunidad transmembrana TM de ciertos virus producen proteínas de fusión altamente bioactivas. Esta observación permitió el desarrollo de un procedimiento de producción de proteínas de fusión que puede usarse para producir proteínas homo- o hetero-triméricas que retienen su actividad biológica.

40 Antes de describir la presente invención, debe entenderse que la presente invención no se limita a las realizaciones particulares descritas, ya que tales pueden, por supuesto, variar. También debe entenderse que la terminología usada en el presente documento es solo con el fin de describir realizaciones particulares, y no pretende ser limitante, ya que el alcance de la presente invención solo se limitará por las reivindicaciones adjuntas.

45 A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que comúnmente es entendido por un experto habitual en la técnica a la que pertenece la presente invención. Aunque cualquier procedimiento y material similar o equivalente a aquellos descritos en el presente documento puede usarse en la práctica o prueba de la presente invención, ahora se describen los procedimientos y materiales preferidos.

50 Debe observarse que, como se usa en el presente documento y en las reivindicaciones adjuntas, las formas en singular "un", "una", "el" y "la" incluyen referentes plurales, a menos que el contexto dicte claramente de otro modo. Así, por ejemplo, referencia a "un polimorfismo" incluye una pluralidad de tales polimorfismos, referencia a "una molécula de ácido nucleico" incluye una pluralidad de tales moléculas de ácidos nucleicos, y referencia a "el procedimiento" incluye referencia a uno o más procedimientos, etapas de procedimiento y equivalentes de los mismos conocidos para aquellos expertos en

la materia, etc.

Las publicaciones tratadas en el presente documento se proporcionan únicamente para su divulgación antes de la fecha de presentación de la presente solicitud. Nada en el presente documento debe interpretarse como una admisión de que la presente invención no tiene derecho a anteceder tal publicación en virtud de invención previa. Además, las fechas de publicación proporcionadas pueden ser diferentes de las fechas de publicación reales que pueden necesitar ser confirmadas independientemente.

Definiciones

En la presente memoria descriptiva, el término "proteína de fusión" se usa en el presente documento para describir una proteína cuyas secuencias se derivan de al menos dos fuentes de genes diferentes. Las secuencias se manipulan genéticamente para transcribirse y traducirse en una proteína que comprende secuencias de al menos dos genes diferentes. Para la presente invención, una fuente de genes es una secuencia de repetición de héptada viral (conocida como el dominio ZymoZipper o ZZ) y la fuente o fuentes de genes adicionales son uno o más genes heterólogos. La proteína de fusión también comprende una secuencia de adaptador que generalmente se localizará entre el dominio ZZ y la secuencia de la proteína heteróloga, como se define en las reivindicaciones.

El término "heterólogo" se usa para describir un polinucleótido o proteína que no es codificado o expresado de modo natural con la secuencia de repetición de héptada viral del dominio ZZ dentro del genoma vírico. Si el dominio ZZ se deriva de un retrovirus endógeno humano, cualquier fuente de genes fuera de las proteínas que se cree que están codificadas por secuencias derivadas de virus se considera heteróloga. Una proteína heteróloga puede ser una proteína de longitud completa o un dominio particular de una proteína. Las proteínas heterólogas de la presente invención engloban tanto proteínas unidas a la membrana como proteínas solubles y dominios de las mismas.

Los términos "polinucleótido" y "molécula de ácido nucleico" se usan indistintamente en el presente documento para referirse a formas poliméricas de nucleótidos de cualquier longitud. Los polinucleótidos pueden contener desoxirribonucleótidos, ribonucleótidos y/o sus análogos. Los nucleótidos pueden tener cualquier estructura tridimensional, y pueden realizar cualquier función, conocida o desconocida. El término "polinucleótido" incluye moléculas monocatenarias, bicatenarias y de tres hélices. "Oligonucleótido" generalmente se refiere a polinucleótidos de entre aproximadamente 5 y aproximadamente 100 nucleótidos de ADN monocatenario o bicatenario. Sin embargo, para los fines de esta divulgación, no hay límite superior para longitud de un oligonucleótido. Los oligonucleótidos también se conocen como oligómeros u oligos y pueden aislarse de genes, o sintetizarse químicamente mediante procedimientos conocidos en la técnica.

Lo siguiente son realizaciones no limitantes de polinucleótidos: un gen o fragmento de gen, exones, intrones, ARNm, ARNt, ARNr, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. Una molécula de ácido nucleico también puede comprender moléculas de ácidos nucleicos modificadas tales como moléculas de ácidos nucleicos metiladas y análogos de moléculas de ácidos nucleicos. En la técnica se conocen análogos de purinas y pirimidinas. Los ácidos nucleicos pueden producirse naturalmente, por ejemplo, ADN o ARN, o pueden ser análogos sintéticos, como se conoce en la técnica. Tales análogos pueden ser preferidos para su uso como sondas debido a la superior estabilidad bajo condiciones de ensayo. Se ha mostrado que las modificaciones en la estructura nativa, que incluyen alteraciones en el esqueleto, azúcares o bases heterocíclicas, aumentan la estabilidad intracelular y afinidad de unión. Entre los cambios útiles en la química del esqueleto están fosforotioatos; fosforoditioatos, en los que ambos oxígenos de no unión por puente están sustituidos con azufre; fosforamiditos; alquilfosfotriésteres y boranofosfatos. Los derivados de fosfato acquiral incluyen 3'-O'-5'-S-fosforotioato, 3'-S-5'-O-fosforotioato, 3'-CH₂-5'-O-fosfonato y 3'-NH-5'-O-fosforoamidato. Los ácidos nucleicos peptídicos sustituyen el esqueleto del fosfodiéster de ribosa entero con un enlace peptídico.

También se usan modificaciones de azúcar para potenciar la estabilidad y afinidad. Puede usarse el α -anómero de desoxirribosa, en el que la base está invertida con respecto al β -anómero natural. El 2'-OH del azúcar ribosa puede alterarse para formar azúcares 2'-O-metilo o 2'-O-alilo, que proporcionan resistencia a la degradación sin comprender la afinidad.

La modificación de las bases heterocíclicas debe mantener el apropiado apareamiento de bases. Algunas sustituciones útiles incluyen desoxitimidina por desoxiuridina; desoxicitidina por 5-metil-2'-desoxicitidina y 5-bromo-2'-desoxicitidina. Se ha mostrado que la desoxitimidina y desoxicitidina aumentan la afinidad y actividad biológica cuando se sustituyen por 5-propinil-2'-desoxiuridina y 5-propinil-2'-desoxicitidina, respectivamente.

Los términos "polipéptido" y "proteína", usados indistintamente en el presente documento, se refieren a una forma polimérica de aminoácidos de cualquier longitud, que puede incluir aminoácidos codificados y no codificados, aminoácidos químicamente o bioquímicamente modificados o derivatizados, y polipéptidos que tienen esqueletos de péptido modificados. El término incluye proteínas de fusión que incluyen, pero no se limitan a, proteínas de fusión con una

secuencia de aminoácidos heteróloga, fusiones con secuencias conductoras heterólogas y homólogas, con o sin residuos de metionina en el extremo N; proteínas inmunológicamente elegidas como diana; y similares.

Un polinucleótido "sustancialmente aislado" o "aislado" es uno que está sustancialmente libre de las secuencias con las que está asociado en la naturaleza. Por sustancialmente libre se indica al menos el 50 %, preferentemente al menos el 70 %, más preferentemente al menos el 80 %, e incluso más preferentemente al menos el 90 % libre de los materiales con los que está asociado en la naturaleza. Como se usa en el presente documento, un polinucleótido "aislado" también se refiere a polinucleótidos recombinantes que, en virtud del origen o manipulación: (1) no están asociados a todo o una parte de un polinucleótido con el que está asociado en la naturaleza, (2) están ligados a un polinucleótido distinto de con el que está ligado en la naturaleza, o (3) no se produce en la naturaleza.

Las reacciones de hibridación pueden realizarse en condiciones de diferente "rigurosidad". Condiciones que aumentan la rigurosidad de una reacción de hibridación son ampliamente conocidas y publicadas en la materia. Véase, por ejemplo, Sambrook y col. (1989). Ejemplos de condiciones relevantes incluyen (en orden de rigurosidad creciente): temperaturas de incubación de 25 °C, 37 °C, 50 °C y 68 °C; concentraciones de tampón de 10×SSC, 6×SSC, 1×SSC, 0,1×SSC (en las que SSC es NaCl 0,15 M y tampón citrato 15 mM) y sus equivalentes usando otros sistemas tampón; concentraciones de formamida del 0 %, 25 %, 50 % y 75 %; tiempos de incubación de 5 minutos a 24 horas; 1, 2 o más etapas de lavado; tiempos de incubación de lavado de 1, 2 ó 15 minutos; y disoluciones de lavado de 6×SSC, 1×SSC, 0,1×SSC, o agua desionizada. Ejemplos de condiciones rigurosas son hibridación y lavado a 50 °C o superior y en 0,1×SSC (NaCl 9 mM / citrato de sodio 0,9 mM).

"T_m" es la temperatura en grados Celsius a la que el 50 % de un dúplex de polinucleótido constituido por hebras complementarias unidas por hidrógeno en dirección antiparalela por apareamiento de bases de Watson-Crick se disocia en sus hebras individuales en condiciones del experimento. T_m puede predecirse según una fórmula convencional, tal como:

en la que [X⁺] es la concentración de catión (normalmente ión sodio, Na⁺) en moles/l; (% de G/C) es el número de residuos G y C como porcentaje de residuos totales en el dúplex; (% de F) es el porcentaje de formamida en disolución (peso/vol); y L es el número de nucleótidos en cada hebra del dúplex.

Condiciones rigurosas para tanto hibridación de ADN/ADN como de ADN/ARN son como se describen por Sambrook y col. Molecular Cloning, A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989, incorporados en el presente documento por referencia. Por ejemplo, véase la página 7.52 de Sambrook y col.

El término "célula huésped" incluye una célula individual o cultivo celular que puede ser o ha sido un receptor de cualquier vector recombinante o polinucleótido aislado de la invención. Células huésped incluyen progenie de una única célula huésped, y la progenie puede no ser necesariamente completamente idéntica (en morfología o en complemento de ADN total) a la célula parental original debido a mutación y/o cambio natural, accidental o intencionado. Una célula huésped incluye células transfectadas o infectadas *in vivo* o *in vitro* con un vector recombinante o un polinucleótido de la invención. Una célula huésped que comprende un vector recombinante de la invención es un "célula huésped recombinante".

El término "secuencia señal secretora" indica una secuencia de ADN que codifica un polipéptido (un "péptido secretor") que, como componente de un polipéptido mayor, dirige al polipéptido mayor por una ruta de una célula en la que se sintetiza. El polipéptido mayor se escinde comúnmente para eliminar el péptido secretor durante el tránsito por la ruta secretora.

El término "marca de afinidad" se usa en el presente documento para indicar un segmento de polipéptido que puede unirse a un segundo polipéptido para proporcionar la purificación o detección del segundo polipéptido o proporcionar sitios para la unión del segundo polipéptido a un sustrato. En principio, cualquier péptido o proteína para el que está disponible un anticuerpo u otro agente de unión específico puede usarse como marca de afinidad. Las marcas de afinidad incluyen un tracto de poli-histidina, proteína A (Nilsson y col., EMBO J. 4:1075, 1985; Nilsson y col., Methods Enzymol. 198:3, 1991), glutatión S transferasa (Smith y Johnson, Gene 67:31, 1988), marca de afinidad Glu-Glu (Grussenmeyer y col., Proc. Natl. Acad. Sci. USA 82:7952-4, 1985), sustancia P, péptido Flag™ (Hopp y col., Biotechnology 6:1204-10, 1988), péptido de unión a estreptavidina, u otro epítope antigénico o dominio de unión. Véase, en general, Ford y col., Protein Expression and Purification 2: 95-107, 1991. Están disponibles ADN que codifican marcas de afinidad de proveedores comerciales (por ejemplo, Pharmacia Biotech, Piscataway, NJ).

Los términos "extremo amino" (extremo N) y "extremo carboxilo" (extremo C) se usan en el presente documento para indicar posiciones dentro de polipéptidos. Si el contexto lo permite, estos términos se usan con referencia a una secuencia o porción particular de un polipéptido para indicar proximidad o posición relativa. Por ejemplo, una cierta secuencia posicionada en el extremo carboxilo con respecto a una secuencia de referencia dentro de un polipéptido se localiza proximal al extremo carboxilo de la secuencia de referencia, pero no está necesariamente en el extremo carboxilo del polipéptido completo.

Como se usa en el presente documento, los términos “tratamiento”, “tratar” y similares se refieren a obtener un efecto farmacológico y/o fisiológico deseado. El efecto puede ser profiláctico en términos de prevenir completamente o parcialmente una enfermedad o síntoma de la misma, y/o pueden ser terapéuticos en términos de una cura parcial o completa de una enfermedad y/o efecto adverso atribuible a la enfermedad. “Tratamiento”, como se usa en el presente documento, cubre cualquier tratamiento de una enfermedad en un mamífero, particularmente en un ser humano, e incluye:

5 (a) prevenir que la enfermedad se produzca en un sujeto que puede tener predisposición a la enfermedad, pero que todavía no ha sido diagnosticado con que la tiene; (b) inhibir la enfermedad, es decir, detener su desarrollo; y (c) aliviar la enfermedad, es decir, causando regresión de la enfermedad.

Los términos “individuo”, “sujeto” y “paciente”, usados indistintamente en el presente documento, se refieren a un mamífero que incluye, pero no se limita a, murinos, simios, seres humanos, animales de granja mamíferos, animales para deportes mamíferos, y mascotas mamíferas.

10 El dominio ZymoZipper (ZZ)

El dominio ZymoZipper (ZZ)

La presente invención es un procedimiento de producción de una proteína trimérica que comprende una proteína de fusión como se define en las reivindicaciones que comprende un dominio ZZ y un dominio de proteína heterólogo. Los dominios ZZ se derivan de repeticiones de héptada virales como se encuentran en la subunidad TM de muchas especies de virus. Géneros o familias de virus representativas que tienen dominios TM apropiados u otros genes virales que incluyen repeticiones de héptada incluyen, pero no se limitan, a *Orthomyxovirus*, *Filovirus*, *Betaretrovirus*, *Gammaretrovirus*, *Deltaretrovirus*, *Lentivirus*, *Cornavirus* y *Paramyxovirus*. Un *Betaretrovirus* particular contemplado para la presente invención es el virus Mason-Pfizer del mono (VMMP). Otros virus particulares contemplados para la presente invención incluyen, pero no se limitan a, virus de la gripe, virus del Ébola, FRMLV, MoMLV, HTLV-1, VIH-1 y VIS. Las secuencias están seleccionadas por su capacidad esperada para formar estructura de proteínas de bobina en espiral, ya que esta estructura es importante para la capacidad para formar formas de proteínas multiméricas. Particularmente deseado para la presente invención es la capacidad de las proteínas de bobina en espiral para producir estructuras de proteína trimérica. La selección de secuencias apropiadas puede implicar el uso de software de comparación de secuencias, tal como el programa BLAST disponible del Centro nacional para información biotecnológica, como se describe más adelante en el Ejemplo 1.

15 20 25

Otra posible fuente de secuencias del dominio ZZ son subunidades TM de retrovirus endógenos. Éstos incluyen secuencias de retrovirus de rata, ratón o humanos endógenos (RVEH). Más allá de comprender las repeticiones de héptada virales que son características de la estructura de bobina en espiral, se ha encontrado que las proteínas de RVEH se expresan en tejidos humanos. Por ejemplo, la familia de RVEH-W incluye un sitio, ERVWE1, que produce una proteína conocida como sincitina. La ventaja de trabajar con tales secuencias como dominios ZZ es la antigenicidad reducida esperada que tendría una secuencia endógena cuando se introdujera en un sujeto humano.

30

El trabajar con otros tipos de secuencias de cremallera, por ejemplos, la cremallera de leucina, ha mostrado que un número limitado de sustituciones de aminoácidos conservativas (incluso en el residuo d) puede tolerarse frecuentemente en secuencias de cremallera sin la pérdida de la capacidad de las moléculas para multimerizarse (Landschulz y col., (1989), arriba). Así, cambios conservativos de la secuencia viral nativa por el dominio ZZ se contemplan dentro del alcance de la invención. La Tabla 1 muestra los cambios conservativos que se espera que sean tolerados por la estructura de cremallera.

35

Tabla 1

Sustituciones de aminoácidos conservativas

Básicos:	arginina
	lisina
	histidina
Ácidos:	ácido glutámico
	ácido aspártico
Polares:	glutamina
	asparagina
Hidrófobos:	leucina
	isoleucina

(continuación)

Sustituciones de aminoácidos conservativas

valina

metionina

Aromáticos: fenilalanina

triptófano

tirosina

Pequeños: glicina

alanina

serina

treonina

metionina

5 Si más de un proteína de fusión está siendo usada para producir proteínas hetero-multiméricas, por ejemplo, heterotrímeros, el dominio ZZ que se usa puede ser el mismo dominio para ambas proteínas de fusión o dominios ZZ diferentes, en tanto que los dominios tengan la capacidad para asociarse entre sí y formar proteínas multiméricas.

El dominio ZZ puede estar dispuesto en cualquiera del extremo N o C de la proteína heteróloga de interés, basándose en consideraciones de función (es decir, si la proteína heteróloga es una proteína de la membrana tipo I o tipo II) y facilidad de construcción de la construcción.

Proteínas heterólogas -- Proteínas de interés

10 Una proteína heteróloga de interés se selecciona principalmente basándose en un deseo de producir una versión multimérica, particularmente trimérica, de la proteína. Además, usando solo un dominio soluble de la proteína heteróloga, puede producirse en forma soluble una proteína transmembrana. De particular interés con la presente invención es la producción de proteínas biológicamente activas de interés. Una familia de proteínas que usa comúnmente trímeros para la actividad es la familia del factor de necrosis tumoral (TNF). Miembros de esta familia incluyen dos ligandos, APRIL y BLYS.

15 Se ha preparado APRIL humano soluble en células de mamífero en varias formas: con una marca Flag (Hahne y col. (1998) J Exp Med; 188:1185-1190), como una proteína de fusión con el dominio de TNF soluble ACRP30 (Rennert y col. (2000) J Exp Med; 192, 1677-1684) o fosfatasa alcalina secretada (Marsters y col. (2000) Curr Biol; 10: 785-788). Se han publicado dos residuos del extremo amino 105 y 110 para la preparación de la forma soluble de APRIL, y ambas formas incluyen el motivo de unión a heparina (residuos 109-114) (Ingold y col. (2005) J Exp Med; 201:1375-1383; Hendriks y col. (2005) Cell Death Differ; 12: 637-648).

20

A pesar del hecho de que la estructura cristalina de APRIL murino muestre una estructura homotrimérica (y se ha mostrado que APRIL humano soluble fusionado con el epítopo Flag también forma un homotrímero (datos sin publicar)), parece que APRIL requiere un componente de fusión trimérico con el fin de formar una molécula soluble activa (Rennert y col., arriba). Así, esta proteína se seleccionó como proteína heteróloga representativa a expresar usando el procedimiento de proteínas de fusión heterólogas de ZZ y producción y ensayos biológicos de su función se describen más adelante en los Ejemplos 1-4. Otra posible proteína que puede producirse usando el procedimiento de proteínas de fusión de ZZ es BLYS. Usando construcciones de proteínas de fusión de ZZ para ambos ligandos de TNF pueden producirse heterotrímeros de APRIL y BLYS.

25

Secuencias de adaptador, secuencias de marcas de afinidad y péptidos señal

30 Una proteína de interés puede ligarse directamente a otra proteína para formar una proteína de fusión; alternativamente, las proteínas pueden separarse una distancia suficiente para garantizar que las proteínas formen la estructura secundaria y terciaria apropiada necesaria para la actividad biológica. Secuencias de adaptador adecuadas adoptarán una conformación extendida flexible y no presentarán una propensión a desarrollar una estructura secundaria ordenada que pudiera interactuar con los dominios de función de las proteínas de fusión, y tendrán un mínimo carácter hidrófobo o cargado que también pudiera interferir con la función de dominios de fusión. Deben construirse secuencias de adaptador teniendo en cuenta la repetición de héptada viral, ya que puede no actuarse en el mejor interés de la producción de una

35

proteína biológicamente activa para estrechar fuertemente el extremo N o C de la secuencia heteróloga. Este asunto se examina experimentalmente en el Ejemplo 4, más adelante. Más allá de estas consideraciones, la longitud de la secuencia de adaptador puede variar sin afectar significativamente la actividad biológica de la proteína de fusión. Las secuencias de adaptador pueden usarse entre todos y cada uno de los componentes de la proteína de fusión (o construcción de expresión), que incluye marcas de afinidad y péptidos señal.

Otro componente de la proteína de fusión puede ser una marca de afinidad. Tales marcas no alteran la actividad biológica de proteínas de fusión, son altamente antigénicas y proporcionan un epítipo que puede unirse reversiblemente por una molécula de unión específica, tal como un anticuerpo monoclonal, permitiendo la rápida detección y purificación de una proteína de fusión expresada. Las marcas de afinidad también pueden transmitir resistencia a la degradación intracelular si las proteínas se producen en bacterias, como *E. coli*. Una marca de afinidad a modo de ejemplo es la marca FLAG (SEC ID N°: 58). Procedimientos de producción de proteínas de fusión usando esta marca de afinidad para la purificación se describen en la patente de EE.UU. n° 5.011.912.

Todavía otro componente de la proteína de fusión puede ser una secuencia señal o secuencia conductora. Estas secuencias se usan generalmente para permitir la secreción de la proteína de fusión de la célula huésped durante la expresión y también se conocen como una secuencia conductora, secuencia prepro o secuencia pre. La secuencia señal secretora puede ser la de la proteína heteróloga que se produce, si tiene una secuencia tal, o puede derivarse de otro proteína secretada (por ejemplo, t-PA) o sintetizarse *de novo*. La secuencia señal secretora está operativamente ligada a la secuencia de ADN de la proteína de fusión, es decir, la dos secuencias se unen en el marco de lectura correcto y se posicionan para dirigir el polipéptido recientemente sintetizado a la ruta secretora de la célula huésped. Las secuencias señal secretoras están comúnmente posicionadas 5' con respecto a la secuencia de ADN que codifica el polipéptido de interés, aunque ciertas secuencias señal pueden estar posicionadas en cualquier parte en la secuencia de ADN de interés (véase, por ejemplo, Welch y col., patente de EE.UU. n° 5.037.743; Holland y col., patente de EE.UU. n° 5.143.830). Un ejemplo de una secuencia conductora adecuada para la presente invención es un conductor optimizado del gen de activador de plasminógeno tisular humano (tPA) (SEC ID N°: 39 y 40).

Preparación de polinucleótidos que codifican proteínas de fusión heterólogas de ZZ

Las composiciones de ácido nucleico descritas en el presente documento se usan en la preparación de toda o una parte de las proteínas de fusión heterólogas de ZZ, como se ha descrito anteriormente. Los polinucleótidos objeto (incluyendo ADNc o el gen de longitud completa) pueden usarse para expresar un producto génico parcial o completo. Las construcciones que comprenden los polinucleótidos objeto pueden generarse sintéticamente. Alternativamente, el ensamblaje de una sola etapa de un gen y el plásmido entero de grandes números de oligodesoxirribonucleótidos se describe por, por ejemplo, Stemmer y col., *Gens* (Ámsterdam) (1995) 164(1):49-53. En este procedimiento se describe la PCR de ensamblaje (la síntesis de secuencias de ADN largas a partir de grandes números de oligodesoxirribonucleótidos (oligos)). El procedimiento se deriva de barajado de ADN (Stemmer, *Nature* (1994) 370:389-391) y no se basa en ADN ligasa, pero en su lugar se basa en ADN polimerasa para construir fragmentos de ADN cada vez más largos durante el procedimiento de ensamblaje. Construcciones de polinucleótidos apropiadas se purifican usando técnicas de ADN recombinante convencionales como se describen en, por ejemplo, Sambrook y col., *Molecular Cloning: A Laboratory Manual*, 2ª ed., (1989) Cold Spring Harbor Press, Cold Spring Harbor, N.Y., y bajo reglamentaciones actuales descritas en el Departamento de Estados Unidos de Salud y Servicios Humanos (HHS), Pautas del Instituto Nacional de Salud (NIH) para la investigación de ADN recombinante.

Moléculas de polinucleótido que comprenden una secuencia de polinucleótidos proporcionada en el presente documento se propagan disponiendo la molécula en un vector. Se usan vectores virales y no virales, incluyendo plásmidos. La elección del plásmido dependerá del tipo de célula en la que se desea la propagación y el fin de propagación. Ciertos vectores son útiles para amplificar y preparar grandes cantidades de la secuencia de ADN deseada. Otros vectores son adecuados para la expresión en células en cultivo. Todavía otros vectores son adecuados para la transferencia y expresión en células en un animal o persona completa. La elección del vector apropiado está perfectamente dentro de la habilidad de la técnica. Muchos de tales vectores están disponibles comercialmente. El polinucleótido de longitud parcial o completa se inserta en un vector normalmente por medio de unión de ADN ligasa a un sitio de enzima de restricción escindido en el vector. Alternativamente, la secuencia de nucleótidos deseada puede insertarse por recombinación homóloga *in vivo*. Normalmente, esto se lleva a cabo uniendo regiones de homología con el vector sobre los flancos de la secuencia de nucleótidos deseada. Las regiones de homología se añaden por ligación de oligonucleótidos, o por reacción en cadena de la polimerasa usando cebadores que comprenden, por ejemplo, tanto la región de homología como una parte de la secuencia de nucleótidos deseada.

Para la expresión puede emplearse un casete o sistema de expresión. El producto génico codificado por un polinucleótido de la invención se expresa en cualquier sistema de expresión conveniente que incluye, por ejemplo, sistemas bacterianos, de levadura, de insecto, anfibios y de mamífero. Vectores y células huésped adecuados se describen en la patente de EE.UU. n° 5.654.173. En el vector de expresión, un polinucleótido que codifica el polipéptido BLYS polimórfico se liga a una secuencia reguladora según convenga para obtener las propiedades de expresión deseadas. Éstas pueden incluir

promotores (unidos tanto en el extremo 5' de la hebra codificante como en el extremo 3' de la hebra no codificante), potenciadores, terminadores, operadores, represores e inductores. Los promotores pueden regularse o ser constitutivos. En algunas situaciones puede desearse usar promotores condicionalmente activos tales como promotores específicos de tejido o específicos del estadio de desarrollo. Éstos se ligan a la secuencia de nucleótidos deseada usando las técnicas descritas anteriormente para el enlace a vectores. Puede usarse cualquier técnica conocida en la técnica. En otras palabras, el vector de expresión proporcionará una región de iniciación transcripcional y traduccional, que puede ser inducible o constitutiva, estando la región codificante operativamente ligada bajo el control transcripcional de la región de iniciación de la transcripción, y una región de terminación transcripcional y traduccional. Estas regiones de control pueden ser nativas para el ADN que codifica la proteína de fusión heteróloga de ZZ, o pueden derivarse de fuentes exógenas.

5
10
15
Los vectores de expresión generalmente tienen sitios de restricción convenientes localizados cerca de la secuencia promotora para proporcionar la inserción de secuencias de ácidos nucleicos que codifican proteínas heterólogas. Puede estar presente un marcador de selección operativo en el huésped de expresión. Los vectores de expresión pueden usarse para la producción de proteínas de fusión, proporcionando el péptido de fusión exógeno funcionalidad adicional, es decir, elevada síntesis de proteínas, estabilidad, reactividad con antisueros definidos, un marcador de enzimas, por ejemplo, β -galactosidasa, etc.

20
Pueden prepararse casetes de expresión que comprenden una región de iniciación de la transcripción, el gen o fragmento del mismo, y una región de terminación de la transcripción. Es de particular interés el uso de secuencias que permiten la expresión de epítopes o dominios funcionales, normalmente de al menos aproximadamente 8 aminoácidos de longitud, más normalmente de al menos aproximadamente 15 aminoácidos de longitud, a aproximadamente 25 aminoácidos, y hasta el marco de lectura abierto completo del gen. Después de la introducción del ADN, las células que contienen la construcción pueden seleccionarse por medio de un marcador de selección, expandirse las células y luego usarse para la expresión.

25
30
Las proteínas de fusión heterólogas de ZZ pueden expresarse en procariotas o eucariotas según formas convencionales que dependen del fin de la expresión. Para la producción a gran escala de la proteína, un organismo unicelular, tal como *E. coli*, *B. subtilis*, *S. cerevisiae*, células de insecto en combinación con vectores de baculovirus, o células de un organismo superior tal como vertebrados, particularmente mamíferos, por ejemplo, células COS 7, HEK 293, CHO, ovocitos de *Xenopus*, etc., puede usarse como las células huésped de expresión. En algunas situaciones se desea expresar una molécula de ácido nucleico BLYS polimórfica en células eucariotas, beneficiándose la proteína BLYS polimórfica del plegamiento nativo y las modificaciones post-traduccionales. También pueden sintetizarse péptidos pequeños en el laboratorio. Los polipéptidos que son subconjuntos de la secuencia de BLYS completa pueden usarse para identificar e investigar partes de la proteína importantes para la función.

Sistemas de expresión específicos de interés incluyen sistemas de expresión derivados de células bacterianas, levadura, de insecto y células de mamífero. Sistemas representativos de cada una de estas categorías se proporcionan más adelante:

35
Bacterias. Sistemas de expresión en bacterias incluyen aquellos descritos en Chang y col., *Nature* (1978) 275:615; Goeddel y col., *Nature* (1979) 281:544; Goeddel y col., *Nucleic Acids Res.* (1980) 8:4057; documento EP 0 036.776; patente de EE.UU. n° 4.551.433; DeBoer y col., *Proc. Natl. Acad. Sci. (USA)* (1983) 80:21-25; y Siebenlist y col., *Cell* (1980) 20:269.

40
45
Levadura. Sistemas de expresión en levadura incluyen aquellos descritos en Hinnen y col., *Proc. Natl. Acad. Sci. (USA)* (1978) 75:1929; Ito y col., *J. Bacteriol.* (1983) 153:163; Kurtz y col., *Mol. Cell. Biol.* (1986) 6:142; Kunze y col., *J. Basic Microbiol.* (1985)25:141; Gleeson y col., *J. Gen. Microbiol.* (1986) 132:3459; Roggenkamp y col., *Mol. Gen. Genet.* (1986) 202:302; Das y col., *J. Bacteriol.* (1984) 158:1165; De Louvencourt y col., *J. Bacteriol.* (1983) 154:737; Van den Berg y col., *Bio/Technology* (1990)8:135; Kunze y col., *J. Basic Microbiol.* (1985)25:141; Cregg y col., *Mol. Cell. Biol.* (1985) 5:3376; patentes de EE.UU. n° 4.837.148 y 4.929.555; Beach y Nurse, *Nature* (1981) 300:706; Davidow y col., *Curr. Genet.* (1985) 10:380; Gaillardin y col., *Curr. Genet.* (1985) 10:49; Ballance y col., *Biochem. Biophys. Res. Commun.* (1983) 112:284-289; Tilburn y col., *Gene* (1983) 26:205-221; Yelton y col., *Proc. Natl. Acad. Sci. (USA)* (1984) 81:1470-1474; Kelly y Hynes, *EMBO J.* (1985) 4:475479; documentos EP 0 244.234; y WO 91/00357.

50
55
Células de insecto. La expresión de genes heterólogos en insectos se lleva a cabo como se describe en la patente de EE.UU. n° 4.745.051; Friesen y col., "The Regulation of Baculovirus Gene Expression", en: *The Molecular Biology of Baculoviruses* (1986) (W. Doerfler, ed.); documentos EP 0 127.839; EP 0 155.476; y Vlak y col., *J. Gen. Virol.* (1988) 69:765-776; Miller y col., *Ann. Rev. Microbiol.* (1988) 42:177; Carbonell y col., *Gene* (1988) 73:409; Maeda y col., *Nature* (1985) 315:592-594; Lebacqz-Verheyden y col., *Mol. Cell. Biol.* (1988) 8:3129; Smith y col., *Proc. Natl. Acad. Sci. (USA)* (1985) 82:8844; Miyajima y col., *Gene* (1987) 58:273; y Martin y col., *DNA* (1988) 7:99. Numerosas cepas y variantes baculovirales y células huésped de insecto permisivas correspondientes de huéspedes se describen en Luckow y col., *Bio/Technology* (1988) 6:47-55, Miller y col., *Generic Engineering* (1986) 8:277-279, y Maeda y col., *Nature* (1985) 315:592-594.

Células de mamífero. La expresión en mamífero se lleva a cabo como se describe en Dijkema y col., EMBO J. (1985) 4:761, Gorman y col., Proc. Natl. Acad. Sci. (USA) (1982) 79:6777, Boshart y col., Cell (1985) 41:521 y patente de EE.UU. nº 4.399.216. Otras características de la expresión en mamífero se facilitan como se describe en Ham y Wallace, Meth. Enz. (1979) 58:44, Barnes y Sato, Anal. Biochem. (1980) 102:255, patentes de EE.UU. nº 4.767.704, 4.657.866, 4.927.762, 4.560.655, documentos WO 90/103430, WO 87/00195 y patente de EE.UU. nº RE 30.985.

Si cualquiera de las células huésped anteriores, u otras células huésped u organismos apropiados, se usan para replicar y/o expresar polinucleótidos o ácidos nucleicos, el ácido nucleico replicado resultante, ARN, proteína expresada o polipéptido se describe en el presente documento como un producto de la célula huésped u organismo. El producto se recupera por cualquier medio apropiado conocido en la técnica.

Una vez se identifica el gen correspondiente a un polinucleótido seleccionado, su expresión puede regularse en la célula para la que el gen es nativo. Por ejemplo, un gen endógeno de una célula puede regularse por una secuencia reguladora exógena insertada en el genoma de la célula en la localización suficiente para al menos potenciar la expresión del gen en la célula. La secuencia reguladora puede diseñarse para integrarse en el genoma mediante recombinación homóloga, como se desvela en las patentes de EE.UU. nº 5.641.670 y 5.733.761, o puede diseñarse para integrarse en el genoma mediante recombinación no homóloga, como se describe en el documento WO 99/15650.

Vectores y células huésped que comprenden polinucleótidos

La memoria descriptiva describe vectores recombinantes y células huésped que comprenden polinucleótidos. En general, los vectores recombinantes y las células huésped se aíslan; sin embargo, una célula huésped que comprende un polinucleótido puede ser parte de un animal genéticamente modificado.

Vectores recombinantes. La presente memoria descriptiva describe vectores recombinantes ("construcciones") que comprenden un polinucleótido. Los vectores recombinantes incluyen vectores usados para la propagación de un polinucleótido, y vectores de expresión. Vectores útiles para la introducción del polinucleótido incluyen plásmidos y vectores virales, por ejemplo, vectores basados en retrovirus, vectores de adenovirus, etc., que se mantienen transitoriamente o establemente en células de mamífero. Puede emplearse una amplia variedad de vectores para la transfección y/o integración del gen en el genoma de las células. Alternativamente puede emplearse micro-inyección, fusión o similares para la introducción de genes en una célula huésped adecuada.

Los vectores de expresión generalmente tienen sitios de restricción convenientes localizados próximos a la secuencia promotora para proporcionar la inserción de secuencias de ácidos nucleicos que codifican proteínas heterólogas. Puede estar presente un marcador de selección operativo en la expresión huésped. Los vectores de expresión pueden usarse para la producción de proteínas de fusión, proporcionando al péptido de fusión exógeno funcionalidad adicional, es decir, elevada síntesis de proteínas, estabilidad, reactividad con antisueros definidos, un marcador de enzimas, por ejemplo, β -galactosidasa, etc.

Pueden prepararse casetes de expresión que comprenden una región de iniciación de la transcripción, el gen o fragmento del mismo, y una región de terminación de la transcripción. Es de particular interés el uso de secuencias que permiten la expresión de epítopes o dominios funcionales, normalmente de al menos aproximadamente 8 aminoácidos de longitud, más normalmente de al menos aproximadamente 15 aminoácidos de longitud, de al menos aproximadamente 25 aminoácidos, de al menos aproximadamente 45 aminoácidos, y hasta el marco de lectura abierto completo del gen. Después de la introducción del ADN, las células que contienen la construcción pueden seleccionarse por medio de un marcador de selección, expandirse las células y luego usarse para la expresión.

Los casetes de expresión pueden introducirse en una variedad de vectores, por ejemplo, plásmido, BAC, YAC, bacteriófago tal como lambda, P1, M13, etc., virus de animal o planta y similares, caracterizándose normalmente los vectores por la capacidad para proporcionar selección de células que comprenden los vectores de expresión. Los vectores pueden proporcionar mantenimiento extracromosómico, particularmente como plásmidos o virus, o integración en el cromosoma huésped. Si se desea mantenimiento extracromosómico, se proporciona una secuencia de origen para la replicación del plásmido, que puede tener bajo o alto número de copias. Está disponible una amplia variedad de marcadores para la selección, particularmente aquellos que protegen contra toxinas, más particularmente contra antibióticos. El marcador particular que se elige está seleccionado según la naturaleza del huésped, pudiendo emplearse en algunos casos complementación con huéspedes auxotróficos. La introducción de la construcción de ADN puede usar cualquier procedimiento conveniente, por ejemplo, conjugación, transformación bacteriana, ADN precipitado con calcio, electroporación, fusión, transfección, infección con vectores virales, biolística, etc.

Células genéticamente modificadas. La presente memoria descriptiva describe células huésped, que pueden ser células huésped aisladas, que comprenden moléculas de ácidos nucleicos de BLYS polimórficas de la invención. Células huésped adecuadas incluyen procariotas tales como *E. coli*, *B. subtilis*, eucariotas, que incluyen células de insecto en combinación con vectores de baculovirus, células de levadura, tales como *Saccharomyces cerevisiae*, o células de un organismo superior tal como vertebrados, que incluyen anfibios (por ejemplo, ovocitos de *Xenopus laevis*), y mamíferos,

particularmente seres humanos, por ejemplo, células COS, células CHO, células HEK293 y similares, pueden usarse como células huésped. Las células huésped pueden usarse para los fines de propagar una molécula de ácido nucleico de BLYS polimórfica, para la producción de un polipéptido de BLYS polimórfico o procedimientos basados en células para identificar agentes que modulan un nivel de ARNm y/o proteína de BLYS y/o actividad biológica en una célula.

- 5 Pueden modificarse células y líneas celulares primarias o clonadas por la introducción de vectores que comprenden un ADN que codifica el (los) polimorfismo(s) de la proteína heteróloga de fusión de ZZ. La molécula de ácido nucleico de BLYS polimórfica aislada puede comprender una o más secuencias de variante, por ejemplo, un haplotipo de combinaciones que se producen comúnmente. Puede proporcionarse un panel de dos o más líneas celulares genéticamente modificadas, comprendiendo cada línea celular un polimorfismo de BLYS, para ensayos de sustrato y/o de expresión. El panel puede comprender además células genéticamente modificadas con otras secuencias genéticas, que incluyen polimorfismos, particularmente otras secuencias de interés para cribado farmacogenético, por ejemplo, otros genes/mutaciones de genes asociados a obesidad, varios de los cuales se conocen en la técnica.

- 10 Animales transgénicos. Los ácidos nucleicos objeto pueden usarse para generar animales no humanos genéticamente modificados o modificaciones de genes específicos de sitio en líneas celulares. El término "transgénico" pretende englobar animales genéticamente modificados que tienen la adición de ADN que codifica la proteína de fusión heteróloga de ZZ o que tiene un ADN exógeno que codifica la proteína de fusión heteróloga de ZZ que se transmite establemente en las células huésped. Los animales transgénicos pueden prepararse mediante recombinación homóloga. Alternativamente, una construcción de ácidos nucleicos se integra aleatoriamente en el genoma. Vectores para la integración estable incluyen plásmidos, retrovirus y otros virus animales, YAC y similares. De interés son mamíferos transgénicos, por ejemplo vacas, cerdos, cabras, caballos, etc., y particularmente roedores, por ejemplo, ratas, ratones, etc.

- 15 Las construcciones de ADN para la recombinación homóloga comprenderán al menos una parte del ADN que codifica la proteína de fusión heteróloga de ZZ e incluirán regiones de homología para el sitio diana. Convenientemente están incluidos marcadores para selección positiva y negativa. Los procedimientos para generar células que tienen modificaciones de genes elegidas como diana mediante recombinación homóloga son conocidos en la técnica. Para diversas técnicas para transfectar células de mamífero, véase Known y col. (1990) *Methods in Enzymology* 185:527-537.

- 20 Para células madre embrionarias (ES) puede emplearse una línea de células ES, o pueden obtenerse células ES recientes a partir de un huésped, por ejemplo ratón, rata, cobaya, etc. Tales células se cultivan sobre una capa alimentadora de fibroblastos apropiada o se cultivan en presencia de factor inhibidor de leucemia (LIF). Cuando las células ES se han transformado, pueden usarse para producir animales transgénicos. Después de la transformación, las células se siembran sobre una capa alimentadora en un medio apropiado. Pueden detectarse células que contienen la construcción empleando un medio selectivo. Después de un tiempo suficiente para que las colonias crezcan, se recogen y se analizan para la aparición de recombinación homóloga. Aquellas colonias que muestran recombinación homóloga pueden entonces usarse para la manipulación de embriones e inyección de blastocistos. Los blastocistos se obtienen de hembras superovuladas de 4 a 6 semanas de edad. Las células ES se tripsinan, y las células modificadas se inyectan en el blastocelo del blastocisto. Después de la inyección, los blastocistos se devuelven a cada trompa uterina de hembras seuspereñadas. Entonces se deja que las hembras lleguen a término y las camadas resultantes se criban para células mutantes que tienen la construcción. La progenie quimérica puede detectarse fácilmente proporcionando un fenotipo diferente del blastocisto y las células ES. Los animales quiméricos se criban para la presencia de ADN que codifica la proteína de fusión heteróloga de ZZ y los machos y hembras que tienen la modificación se aparean para producir progenie homocigótica. Los animales transgénicos pueden ser cualquier mamífero no humano, tal como animales de laboratorio, animales domésticos, etc. Los animales transgénicos pueden usarse para determinar el efecto de un fármaco candidato en un entorno *in vivo*.

Producción de proteínas hetero-triméricas usando construcciones de ZZ

- 45 La presente invención es un procedimiento de preparación de una proteína hetero-trimérica soluble como se define en las reivindicaciones, cultivando una primera célula huésped transformada o transfectada con un primer vector de expresión que codifica una primera proteína de fusión que comprende un primer dominio ZZ y una primera proteína heteróloga, cultivando una segunda célula huésped transformada o transfectada con un segundo vector de expresión que codifica un segundo dominio ZZ y una segunda proteína heteróloga, y cultivando una tercera célula huésped transformada o transfectada con un tercer vector de expresión que codifica un tercer dominio ZZ y una tercera proteína heteróloga. Con el fin de producir una proteína biológicamente funcional, los tres dominios ZZ forman preferencialmente un hetero-trímero. El cultivo también puede producirse en la misma célula huésped, si puede mantenerse la producción eficiente, y las proteínas hetero-triméricas se aíslan entonces del medio. Idealmente, las tres proteínas heterólogas se marcan diferencialmente con diversas secuencias de marca (es decir, marca de His, marca FLAG y marca de Glu-Glu) para permitir el análisis de la composición o purificación de las moléculas resultantes. Alternativamente, los tres componentes pueden producirse por separado y combinarse en relaciones intencionadas para producir las moléculas hetero-triméricas deseadas. Los dominios ZZ usados en la preparación de estas moléculas hetero-triméricas pueden ser iguales o diferentes. La(s) proteína(s) de fusión comprende(n) además una secuencia de adaptador. En una realización particular, las dos proteínas heterólogas

usadas para formar la proteína hetero-trimérica son los dominios solubles de APRIL (un ligando de proliferación de linfocitos) y BLYS (estimulante de linfocitos B).

Actividad biológica de las proteínas de fusión heterólogas de ZZ

5 La actividad biológica de las proteínas de fusión heterólogas de ZZ recombinantes está mediada por la unión de la proteína de fusión recombinante a una molécula afín, tal como un receptor. Una molécula afín se define como una molécula que une la proteína de fusión recombinante en una interacción no covalente basándose en la conformación apropiada de la proteína de fusión recombinante y la molécula afín. Por ejemplo, para una proteína de fusión recombinante que comprende una región extracelular de un receptor, la molécula afín comprende un ligando que une la región extracelular del receptor. En cambio, para una proteína de fusión soluble recombinante que comprende un ligando, la molécula afín comprende un receptor (o proteína de unión) que se une al ligando.

15 La unión de una proteína de fusión recombinante con una molécula afín es un marcador para actividad biológica. Tal actividad de unión puede determinarse, por ejemplo, por competición para unirse al dominio de unión de la molécula afín (es decir, ensayos de unión competitiva). Una configuración de un ensayo de unión competitiva para una proteína de fusión recombinante que comprende un ligando usa un receptor soluble radiomarcado y células intactas que expresan una forma nativa del ligando. Similarmente, un ensayo competitivo para una proteína de fusión recombinante que comprende un receptor usa un ligando soluble radiomarcado y células intactas que expresan una forma nativa del receptor. Un ensayo tal se describe en el Ejemplo 5. En lugar de células intactas que expresan una forma nativa de la molécula afín, podría sustituirse la molécula afín purificada unida a una fase sólida. Los ensayos de unión competitiva pueden realizarse usando metodología convencional. Pueden obtenerse resultados cualitativos o semi-cuantitativos por ensayos de unión en placa autorradiográfica competitiva, o citometría de flujo activada por fluorescencia, o pueden usarse representaciones de Scatchard para generar resultados cuantitativos.

25 La actividad biológica también puede medirse usando bioensayos que se conocen en la técnica, tal como un ensayo de proliferación celular. Bioensayos a modo de ejemplo se describen en los Ejemplos 5 y 6 en el presente documento. El tipo de ensayo de proliferación celular usado dependerá de la proteína de fusión soluble recombinante. Un bioensayo para una proteína de fusión soluble recombinante que en su forma nativa actúa sobre linfocitos T usará linfocitos T purificados obtenidos mediante procedimientos que se conocen en la técnica. Tales bioensayos incluyen ensayos de co-estimulación en los que los linfocitos T purificados se incuban en presencia de la proteína de fusión soluble recombinante y un nivel inferior al óptimo de un mitógeno tal como Con A o PHA. Similarmente, se usarán linfocitos B purificados para una proteína de fusión soluble recombinante que en su forma nativa actúa sobre linfocitos B. También pueden seleccionarse otros tipos de células basándose en el tipo de célula sobre el que actúa la forma nativa de la proteína de fusión soluble recombinante. La proliferación se determina midiendo la incorporación de una sustancia radiomarcada, tal como ³H-timidina, según procedimientos convencionales.

35 Todavía otro ensayo tipo para determinar la actividad biológica es la inducción de secreción de moléculas secundarias. Por ejemplo, ciertas proteínas inducen la secreción de citocinas por linfocitos T. Los linfocitos T se purifican y se estimulan con una proteína de fusión soluble recombinante en las condiciones requeridas para inducir la secreción de citocinas (por ejemplo, en presencia de un co-mitógeno). La inducción de la secreción de citocinas se determina por bioensayo, midiendo la proliferación de una línea celular dependiente de citocina. Similarmente, la inducción de la secreción de inmunoglobulinas se determina midiendo la cantidad de inmunoglobulina secretada por linfocitos B purificados estimulados con una proteína de fusión soluble recombinante que actúa sobre los linfocitos B en su forma nativa, usando un ensayo cuantitativo (o semi-cuantitativo) tal como un inmunoensayo enzimático.

Procedimientos de tratamiento

45 Para uso farmacéutico, las proteínas de fusión descritas en el presente documento se formulan para administración parenteral, particularmente intravenosa o subcutánea, según procedimientos convencionales. La administración intravenosa será mediante inyección en bolo o infusión durante un periodo típico de una a varias horas. En general, las formulaciones farmacéuticas incluirán una proteína de fusión heteróloga de ZZ en combinación con un vehículo farmacéuticamente aceptable tal como solución salina, solución salina tamponada, 5 % de dextrosa en agua o similares. Las formulaciones pueden incluir adicionalmente uno o más excipientes, conservantes, solubilizantes, agentes de tamponamiento, albúmina para prevenir la pérdida de proteína sobre las superficies del vial, etc. Los procedimientos de formulación son muy conocidos en la técnica y se desvelan, por ejemplo, en Remington's Pharmaceutical Sciences, Gennaro, ed., Mack Publishing Co., Easton PA, 1990, que se incorpora en el presente documento por referencia. Las dosis terapéuticas estarán generalmente en el intervalo de 0,1 a 100 µg/kg de peso de paciente por día, preferentemente 0,5-20 µg/kg por día, siendo la dosis exacta determinada por el profesional clínico según patrones aceptados, teniendo en cuenta la naturaleza y gravedad de la afección que va a tratarse, rasgos del paciente, etc. La determinación de la dosis está dentro del nivel del experto habitual en la técnica. Las proteínas pueden administrarse para tratamiento agudo, durante una semana o menos frecuentemente, durante un periodo de uno a tres días, o pueden usarse en tratamiento crónico, durante varios meses o años. En general, una cantidad terapéuticamente eficaz de proteína de fusión heteróloga

de ZZ es una cantidad suficiente para producir un cambio clínicamente significativo en los síntomas característicos de la falta de función de proteína heteróloga. Alternativamente, si la proteína de fusión heteróloga de ZZ es para actuar de antagonista, una cantidad terapéuticamente eficaz es la que produce un cambio clínicamente significativo en síntomas característicos de una sobreabundancia de la función de proteína heteróloga.

5 Eliminación enzimática del dominio ZZ

Dependiendo del uso previsto de la proteína de fusión heteróloga de ZZ, puede ser ventajoso eliminar el dominio ZZ después de la producción de la proteína, pero antes del uso final de la proteína. La eliminación de este dominio puede llevarse a cabo usando proteasas muy conocidas para un experto habitual. La secuencia diana específica para la escisión por diversas proteasas es muy conocida, y tales secuencias pueden estar presentes en la propia secuencia de ZymoZipper naturalmente o por ingeniería genética. Preferentemente, con el fin de evitar cambios en la función multimerizante del dominio ZZ, la secuencia diana puede encontrarse o manipularse en otra porción de la proteína de fusión, tal como la región de adaptador. Algunas enzimas apropiadas para este procedimiento de eliminación incluyen, pero no se limitan a tripsina, Lys C, Lys N, Arg C, Asp N, Glu C (condiciones de bicarbonato o fosfato), quimiotripsina, pepsina (a diversos intervalos de pH), proteinasa K y papaína. También pueden usarse combinaciones de estas enzimas, si fuera necesario, para optimizar la eliminación del dominio ZZ de la proteína de fusión final producida.

EJEMPLOS

Los siguientes ejemplos se presentan de manera que se provea a aquellos expertos habituales en la técnica de una divulgación y descripción completa de cómo preparar y usar la presente invención, y no pretenden limitar el alcance de lo que los inventores consideran como su invención ni pretenden representar que los experimentos más adelante son todos o los únicos experimentos realizados. Se han hecho esfuerzos para garantizar la precisión con respecto a los números usados (por ejemplo, cantidades, temperatura, etc.), pero debe contarse con algunos errores experimentales y desviaciones. A menos que se indique lo contrario, partes son partes en peso, peso molecular es peso molecular promedio en peso, temperatura es en grados Celsius, y presión es atmosférica o próxima a atmosférica.

EJEMPLO 1: SELECCIÓN DE SECUENCIAS DE ZZ CANDIDATAS

Se obtuvo una serie de secuencias de *env* retrovirales por la Herramienta básica de búsqueda de alineamientos locales (BLAST; Centro nacional para información biotecnológica (NCBI)) usando la subunidad TM de la envuelta de la leucemia murina de Moloney como secuencia de búsqueda. Este programa encuentra regiones de similitud local entre secuencias y calcula la significancia estadística de las coincidencias. BLAST puede usarse para deducir la función y relación evolucionaria entre secuencias, además de ayudar a identificar miembros de familias de genes. Las secuencias resultantes se dispusieron en un árbol genómico y se seleccionaron seis secuencias de diferentes agrupaciones dentro del conjunto para maximizar la diversidad de las secuencias de héptadas ensayadas. Se eligieron las secuencias de repetición de héptada mostradas en la Tabla 2 y se usaron para preparar proteínas de fusión con una proteína candidata.

Tabla 2

Componente de la héptada	fg	abcdefg	abcdefg	abcdefg	abcdefg	abcde
ZZ03 (SEC ID N°: 27)	KE	LREVEKS	IENLNKK	MEDGFLD	VLQNRRG	-----
ZZ07 (SEC ID N°: 28)	DD	LREVEKS	ISNLEKS	LTSLSEV	VLQNRRG	LDLLF
ZZ09 (SEC ID N°: 29)	DD	LGALEKS	VSALEKS	LTSLSEV	VLQNRRG	LDLLF
ZZ10 (SEC ID N°: 30)	AD	VQRLQQG	VDDLSDS	LSSLAEV	VLQNRRG	LDLLF

(continuación)

Componente de la héptada	fg	abcdefg	abcdefg	abcdefg	abcdefg	abcde
ZZ11 (SEC ID N°: 31)	NG	FRKMSQT	MLTIQKQ	IDSLAAV	VLQNRRG	LDVLT
ZZ12 (SEC ID N°: 32)	SD	VQAISSTI	QDLQDQ	VDSLAEV	VLQNRRG	LDLLT

Debido a las dificultades en la producción de APRIL biológicamente activo, esta proteína se seleccionó para producir una proteína de fusión. Las proteínas se manipularon para incluir un péptido señal, un marca Flag, la secuencia de ZZ, un adaptador y los aminoácidos 110-250 de APRIL.

5 EJEMPLO 2: CONSTRUCCIÓN DE VECTORES DE EXPRESIÓN DE ZZ

Cebadores para cada una de las construcciones de ZZ se organizaron del siguiente modo: dos mitades, delantera (D) y trasera (T), cada mitad compuesta por dos cebadores, 5' y 3'. Las mitades D y F se ensamblaron a partir de sus cebadores de 5' y 3' respectivos por PCR como se describe más adelante y se desvela en la Tabla 3, luego se ensamblaron alícuotas de D y T para cada ZZ en cada secuencia de ZZ completa por PCR de solapamiento con los cebadores zc49141 y zc49142. Las condiciones de PCR fueron del siguiente modo: 1 ciclo, 94 °C, 2 minutos; 35 ciclos, 94 °C, 30 segundos, seguido de 55 °C, 30 segundos, seguido de 72 °C, 30 segundos; 1 ciclo, 72 °C, 5 minutos. El diez por ciento de cada una de las mezclas de reacción de PCR se ejecutaron sobre un 1,8 % de gel de agarosa para análisis. La porción restante de los fragmentos de PCR se precipitó mediante la adición de 10 µl de acetato de Na 3 M y 250 µl de etanol, se centrifugó a 13.000 x g durante 30 minutos, el sedimento se aclaró una vez con 70 % de etanol y se dejó secar al aire. Los sedimentos se resuspendieron en 10 µl de agua destilada.

Estas PCR produjeron una serie de productos (228-246 pb) constituidos por cada secuencia de ZZ (ZZ3, ZZ7, ZZ9, ZZ10, ZZ11, ZZ12) con 40 pb de secuencia flanqueante en cualquier extremo para promover la recombinación homóloga en el sitio apropiado en el extremo amino de APRIL. Éstas se insertaron en un vector de expresión constituido por pZMP21 con APRIL [res n° 110-250], el conductor de tPA optimizado, y una secuencia que confiere propiedades trimerizantes. El vector pZMP21 es un vector de expresión de mamífero descrito en la publicación de patente n° US2003/023414 A1 y depositado en la Colección americana de cultivos tipo, 10801 University Boulevard, Manassas, VA 20110-2209 (designado ATCC n° PTA-5266). El vector APRIL se linealizó en un único sitio Bgl II en el sitio previsto de recombinación y la secuencia trimerizante se sustituyó por secuencias de ZZ usando recombinación homóloga en *Saccharomyces cerevisiae*.

La recombinación se realizó del siguiente modo: 100 µl de células de levadura competentes (*S. cerevisiae*) se combinaron con 10 µl de la mezcla de producto de PCR y el vector linealizado se transfirió luego a una cubera de electroporación de 0,2 cm. Las mezclas de levadura/ADN se electropulsaron a 0,75 kV (5 kV/cm), ∞ ohm ios, 25 µF. A cada cubeta se añadieron 600 µl de sorbitol 1,2 M y la levadura se sembró en dos alícuotas de 300 µl sobre dos placas URA-D y se incubaron a 30 °C. Esto se hizo para cada secuencia de ZZ y con vector linealizado solo. Después de aproximadamente 48 horas, las colonias se evaluaron contra el control de vector solo y los transformantes de levadura URA+ de placas positivas correspondientes a cada secuencia de ZZ se resuspendieron en 1 ml de H₂O y se centrifugaron brevemente para sedimentar las células de levadura. El sedimento de células se resuspendió en 1 ml de tampón Qiagen P1. Se añadieron quinientos microlitros de cada mezcla a un tubo Eppendorf que contenía 300 µl de perlas de vidrio lavadas con ácido, se removieron con vórtex dos o tres veces durante intervalos de 1 minuto, seguido de una centrifugación de 15 segundos en una centrífuga Eppendorf a velocidad máxima. Se transfirieron trescientos microlitros de lisado de levadura a un tubo fresco, y a partir de ese momento se siguió el procedimiento del fabricante para Qiagen miniprep. El eluato de las columnas QiaPrep se precipitó mediante la adición de 0,1 volúmenes de NaAc 3 M y 2,5 volúmenes de etanol absoluto y el sedimento de ADN se resuspendió en 10 µl de H₂O.

La transformación de células de *E. coli* electrocompetentes (DH10B, GibcoBRL) se hizo con 0,52 µl de preparación de ADN de levadura y 40 µl de células DH10B. Las células se electropulsaron a 1,7 kV, 25 µF y 400 ohmios. Tras la electroporación, 1 ml de SOC (2 % de Bacto Tryptone (Difco, Detroit, MI), 0,5 % de extracto de levadura (Difco), NaCl 10 mM, KCl 2,5 mM, MgCl₂ 10 mM, MgSO₄ 10 mM, glucosa 20 mM) se sembró en alícuotas de 250 µl sobre cuatro placas LB AMP (caldo LB (Lennox), 1,8 % de Bacto Agar (Difco), 100 mg/l de ampicilina). Cada construcción completada se sometió a análisis de secuencias y la construcción que contenía la secuencia correcta se seleccionó para ampliación por el kit QIAGEN Plasmid Mega.

45

Tabla 3: Secuencias de cebadores

oligo	nombre	Secuencia
zc49141	5' todo	AGCCAGGAAATCCATGCCGAGTTGAGACGCTTCCGTAGAG ATTATAAGGACGATGACGAT (SEC ID N°: 1)
zc49142	3' todo	GGTGGCGTTAATGGGAACCAGGTGCAGGACAGAGTGCTGC TTCTTCTGTTTCAGATCTCC (SEC ID N°: 2)
zc49317	ZZ3 F5'	GATTATAAGGACGATGACGATAAGGGCGGAGGTGGCTCAA AGGAGTTTAAACAAC (SEC ID N°: 3)
zc49363	ZZ3 F3'	CATCTTCTTGTTTCAGGTTCTCGATGCGGCGCTCCAGGTTGTT AAACTCCTTTGAGCC (SEC ID N°: 4)
zc49244	ZZ3 B5'	GAGAACCTGAACAAGAAGATGGAGGACGGTTTTCTGGACG TGTGGACCTAC (SEC ID N°: 5)
zc49310	ZZ3 B3'	CTTCTTCTGTTTCAGATCTCCCTCGGCGTTGTAGGTCCACAC GTCCAGAAA (SEC ID N°: 6)
zc49169	ZZ7 F5'	GATTATAAGGACGATGACGATAAGGGCGGAGGTGGCTCAG ATGATCTGCGCGAGGTG (SEC ID N°: 7)
zc49170	ZZ7 F3'	GGTCAGGGACTTCTCCAGATTGGAGATGGACTTCTCCACCT CGCGCAGATCATCTGA (SEC ID N°: 8)
zc49171	ZZ7 B5'	AATCTGGAGAAGTCCCTGACCTCCCTGTCCGAGGTGGTGCT CCAGAACCGCCGCGGC (SEC ID N°: 9)
zc49172	ZZ7 B3'	CTTCTTCTGTTTCAGATCTCCGAACAGCAGGTCCAGGCCGC GGCGGTTCTGGAGCAC (SEC ID N°: 10)
zc49223	ZZ9 F5'	GATTATAAGGACGATGACGATAAGGGCGGAGGTGGCTCAG ATGATCTGGGCGCCCTG (SEC ID N°: 11)
zc49224	ZZ9 F3'	GGTCAGGGACTTCTCCAGGGCGGACACGGATTTTTCCAGG GCGCCCAGATCATCTGA (SEC ID N°: 12)
zc49225	ZZ9 B5'	GCCCTGGAGAAGTCCCTGACCTCCCTTTCCGAGGTCGTCCT GCAGAACCGCCGCGGC (SEC ID N°: 13)
zc49226	ZZ9 B3'	CTTCTTCTGTTTCAGATCTCCAAAAGAAGGTCAAGGCCGC GGCGGTTCTGCAGGAC (SEC ID N°: 14)
zc49227	ZZ10 F5'	GATTATAAGGACGATGACGATAAGGGCGGAGGTGGCTCAG CCGACGTCCAGCGCCTC (SEC ID N°: 15)
zc49228	ZZ10 F3'	GGAAAGGGAGTCGGAAAGATCATCGACGCCCTGCTGGAGG CGCTGGACGTCGGC (SEC ID N°: 16)

(continuación)

oligo	nombre	Secuencia
zc49229	ZZ10 B5'	GATCTTTCGACTCCCTTTCCTCCCTCGCCGAGGTGGTGCT GCAGAATCGCCGCGGC (SEC ID N°: 17)
zc49230	ZZ10 B3'	CTTCTTCTGTTTCAGATCTCCGAAAAGAAGATCAAGGCCGC GGCGATTCTGCAGCAC (SEC ID N°: 18)
zc49231	ZZ11 F5'	GATTATAAGGACGATGACGATAAGGGCGGAGGTGGCTCAA ATGGCTTTCGCAAGATG (SEC ID N°: 19)
zc49232	ZZ11 F3'	ATCGATCTGTTTCTGGATCGTAAGCATGGTCTGGGACATCT TGCGAAAGCCATTTGA (SEC ID N°: 20)
zc49233	ZZ11 B5'	ACGATCCAGAAACAGATCGATTCCCTTGCCGCCGTGGTGCT TCAGAATCGCCGCGGC (SEC ID N°: 21)
zc49234	ZZ11 B3'	CTTCTTCTGTTTCAGATCTCCCGTGAGGACGTCGAGGCCGC GGCGATTCTGAAGCAC (SEC ID N°: 22)
zc49235	ZZ12 F5'	GATTATAAGGACGATGACGATAAGGGCGGAGGTGGCTCAT CCGACGTCCAGGCGATC (SEC ID N°: 23)
zc49236	ZZ12 F3'	GTCGACCTGGTCCTGAAGGTCCTGGATGGTGGAGGAGATC GCCTGGACGTCGGATGA (SEC ID N°: 24)
zc49237	ZZ12 B5'	GACCTTCAGGACCAGGTCGACTCCCTGGCGGAGGTGGTGC TGCAGAATCGCCGCGGC (SEC ID N°: 25)
zc49238	ZZ12 B3'	CTTCTTCTGTTTCAGATCTCCCGTAAGAAGGTCAAGGCCGC GGCGATTCTGCAGCAC (SEC ID N°: 26)

EJEMPLO 3: EXPRESIÓN DEL PANEL DE ZZ-APRIL EN CÉLULAS 293F Y CHO Y PURIFICACIÓN**5 Expresión transitoria en células 293F**

Las construcciones de ZZ-APRIL se expresaron primero transitoriamente en células 293F, así la proteína podría producirse para probar la actividad biológica. La expresión transitoria se logró usando un procedimiento de alta resolución que alcanza los siguientes parámetros de reactivo: relación de lípido/ADN de 1,33 mg/mg; ADN por transfección de 1500 µg; volumen de transfección total de 1500 ml; lípido por transfección de 2000 µl; número de transfecciones de 15; OptiMEM por transfección de 40 ml.

10 El protocolo de transfección fue el siguiente: 1000 ml de células 293F se sembraron en un matraz de centrifuga de 3 l a 435 cm³/ml durante 72 h antes de la transfección. Las células se cultivaron en una estufa de incubación de cabina telefónica a 37 °C, 6 % de CO₂, a 95 rpm. Las células se diluyeron entonces a 136 cm³/ml en 1500 ml inmediatamente antes de la transfección. El complejo de ADN/lípido se preparó diluyendo 2 ml de lipofectamina 2000 en 20 ml de OptiMEM (Invitrogen (Gibco), Carlsbad, CA). A continuación, 1,5 mg de ADN de la construcción de ZZ se diluyeron en 20 ml de OptiMEM. Estas dos disoluciones se incuban durante 5 minutos a temperatura ambiente, luego se combinan y se incuban durante 20-30 minutos a temperatura ambiente, invirtiendo ocasionalmente. La mezcla de ADN/lípido se añade entonces al matriz de la centrifuga. Las células transfectadas se incuban en una estufa de incubación de cabina telefónica a 37 °C, 6 % de CO₂, a 75 rpm. El sobrenadante se recoge 96 horas después de la transfección.

Expresión en células CHO

1,5 mg de las construcciones de ZZ se digirieron cada una usando diversas enzimas de restricción a 37 °C durante tres horas y luego se precipitaron con IPA y se centrifugaron en un tubo de microcentrífuga de 1,5 ml. El sobrenadante se separó por decantación del sedimento, y el sedimento se lavó con 1 ml de 70 % de etanol y se dejó incubar durante 5 minutos a temperatura ambiente. El tubo se centrifugó en una microcentrífuga durante 10 minutos a 14.000 rpm y el sobrenadante se separó por decantación del sedimento. El sedimento se resuspendió entonces en 750 µl de medio PF-CHO en un entorno estéril, se dejó incubar a 60 °C durante 30 minutos y se dejó enfriar a temperatura ambiente. Células 5E6 APFDXB11 se centrifugaron en un tubo cónico de 15 ml para cada construcción de ZZ y se resuspendieron usando la disolución de medio de ADN. Las mezclas de ADN/célula se dispusieron en una cubeta de ranura de 0,4 cm y se sometieron a electroporación usando los siguientes parámetros: 950 µF, alta capacitancia, y 300 V. Entonces se sacó el contenido de las cubetas y se diluyó a 25 ml con medio PF-CHO y se dispuso en un matraz con agitación de 125 ml. El matraz se dispuso en una estufa de incubación sobre un agitador a 37 °C, 6 % de CO₂ y agitación a 120 rpm.

Las líneas celulares se sometieron a selección con nutriente, seguido de amplificación en metrotrexato 500 nM (MTX). La expresión se confirmó por transferencia Western, y la línea celular se amplió y siguió la purificación de proteínas.

Esquema de purificación para ZZ-APRIL-NF

Todos los procedimientos se realizaron a 4 °C, a menos que se indique lo contrario. Los medios acondicionados producidos a partir de tanto la expresión transitoria de células 293F como la expresión de células CHO se cargaron directamente sobre una columna de hidroxapatita 80 µM BIO-RAD Macro-Prep Ceramic. La columna se lavó con PO4 0,2 M, pH 6,8 y se eluyó con un gradiente de 10 VC de PO4 0,2 M, pH 6,8 a PO4 0,4 M, pH 6,8. La etapa de elución se realizó a temperatura ambiente.

Se realizaron geles NuPAGE (teñidos con Coomassie), Western y RP-HPLC para analizar la elución. Las fracciones de interés se reunieron y se purificaron adicionalmente sobre una columna de afinidad anti-FLAG M2 (Kodak). Para algunas proteínas de fusión, esta etapa se realiza mejor usando una columna de cationes. La columna se equilibró y se lavó con PBS, pH 7,3. La columna se eluyó entonces con PO4 10 mM, citrato 40 mM, AmSO₄ 100 mM, pH 3,0. Las fracciones se neutralizaron con Tris 2 M, pH 8,0.

El conjunto anti-FLAG (o conjunto de la columna de cationes) se concentró usando un dispositivo centrífugo Amicon Ultra, MWCO 30.000, y se cargó sobre una Superdex 200 equilibrada (GE Healthcare). El tampón usado fue PO4 50 mM, NaCl 250 mM, pH 6,8.

Las fracciones se reunieron basándose en los resultados de los geles NuPAGE (teñidos con Coomassie), Western y RP-HPLC.

Se realizaron Western y geles NuPAGE teñidos con Coomassie sobre el conjunto para confirmar la pureza y el contenido. Para análisis adicional, la proteína se sometió a AAA, secuenciación del extremo N y SEC-MALS, confirmando la identidad de la proteína como ZZ-APRIL del extremo N.

Las secuencias de ADN para ZZ7-APRIL, ZZ9-APRIL, ZZ10-APRIL, ZZ11-APRIL y ZZ12-APRIL son SEC ID N°: 43, 46, 49, 52 y 55, respectivamente. Las secuencias de proteína para ZZ3-APRIL, ZZ7-APRIL, ZZ9-APRIL, ZZ10-APRIL, ZZ11-APRIL y ZZ12-APRIL son SEC ID N°: 41, 44, 47, 50, 53 y 56, respectivamente. Las secuencias de ADN consenso para ZZ3-APRIL, ZZ7-APRIL, ZZ9-APRIL, ZZ10-APRIL, ZZ11-APRIL y ZZ12-APRIL son SEC ID N°: 42, 45, 48, 51, 54 y 57, respectivamente.

40 **EJEMPLO 4: ENSAYO DE PROLIFERACIÓN DE LINFOCITOS B USANDO BLYS**

Un vial que contiene 1×10^8 células mononucleares de sangre periférica (PBMC) aferesadas congeladas se descongeló rápidamente en baño de agua a 37 °C y se resuspendieron en 25 ml de medio de linfocitos B (medio RPMI-1640, 10 % de suero bovino fetal inactivado por calor, 5 % de L-glutamina, 5 % de pen/estrep) en un tubo de 50 ml. Las células se probaron para viabilidad usando azul de tripano (GIBCO BRL, Gaithersburg, MD). Los linfocitos B CD19+ se aislaron entonces por selección positiva usando microperlas recubiertas de anti-CD 19 (Miltenyi Biotech). Las células recubiertas se aislaron entonces sobre una columna MACS LS (Miltenyi Biotech). Los linfocitos B se resuspendieron a una concentración final de $1,6 \times 10^6$ células/ml en medio de linfocitos B y se sembraron a 100 µl/pocillo en una placa de fondo en U de 96 pocillos (Falcon, VWR, Seattle, WA).

Diversas ZZ-APRIL solubles como se han preparado en el Ejemplo 3 o BLYS se añadieron a las células en diluciones de 3 veces de 1000 ng/ml a 0 ng/ml. El volumen final fue 200 µl/pocillo.

Las células se incubaron entonces a 37 °C en una estufa de incubación humidificada durante 72 horas. Dieciséis horas

antes de la recogida se añadió 1 µCi de ³H-timidina a todos los pocillos. Las células se recogieron en una placa de filtro de 96 pocillos (UniFilter GF/C, Packard, Meriden, CT) en la que se recogieron usando un colector de células (Packard) y se recogieron según instrucciones del fabricante. Las placas se secaron a 55 °C durante 20-30 minutos y el fondo de los pocillos se tapó con un sellante de placas opaco. A cada pocillo se añadieron 0,25 ml de fluido de centelleo (Microscint-O, Packard) y la placa se leyó usando un contador de centelleo de microplacas TopCount (Packard).

Los resultados del ensayo de proliferación se informan en la Figura 3.

Este ensayo también se usó para comparar la función biológica de una construcción de ZZ-APRIL con diversos APRIL comercialmente disponibles (Alexis APRIL y R&D APRIL). Como puede apreciarse en la Figura 5, ZZ-APRIL imitó mejor el control positivo de BLyS que cualquiera de las fuentes comercialmente disponibles.

10 **EJEMPLO 5: BIOENSAYO DE CÉLULAS JURKAT TRANSFECTADAS CON TAC1**

El bioensayo *in vitro* de TAC1 usa una línea de células Jurkat (linfocito de linfocitos T agudo humano, KZ142, clon nº 24) que se ha transfectado con dos plásmidos. Primero, la línea celular se transfectó con un plásmido que contenía un gen indicador de luciferasa bajo el control del promotor NF-κβ/AP-1 y un gen de resistencia a neomicina. Se eligió un clon apropiado después de la selección con G418. Esta línea celular se transfectó entonces con un plásmido que contenía el ADNc de TAC1 de longitud completa bajo el control del promotor del CMV (TAC1/pZP7P) y un gen resistente a puromicina. Los clones se seleccionaron con puromicina y luego se eligió una línea celular apropiada para el ensayo evaluando la expresión de TAC1 por citometría de flujo usando anticuerpos monoclonales para TAC1.

El ensayo se basa en registrar la lectura de la expresión génica de luciferasa que es desencadenada por la unión del ligando de prueba (ZZ-APRIL) a TAC1 de la superficie celular producido a partir del ADNC de TAC1.

20 Las células de Jurkat transfectadas se propagaron en medio RPMI 1640 sin rojo de fenol (Rosewell Park Memorial Institute, Buffalo, NY) con 10 % de SBF añadido. Se añadió puromicina a 2 µg/ml como reactivo selectivo para la transfección. También se añadieron piruvato de sodio y L-glutamina a los medios al 1 % de volumen.

Se usaron el sustrato del sistema de ensayo de luciferasa Steady-Glo (Promega, Madison, WI #E2510) y tampón de ensayo según instrucciones del fabricante.

25 El ensayo se realizó resuspendiendo células de Jurkat transfectadas en medio a 1,6 x 10⁶ célula/ml. Las células se siembran sobre una placa de ensayo blanca con 50 µl por pocillo. Las muestras de ZZ-APRIL se llevaron a dilución apropiada y se dispusieron en una placa de 96 pocillos. Se añadieron diluciones a las células, a 50 µl por pocillo. La placa se incluyó 4 h en una estufa de incubación a 37 °C. Durante la incubación, el tampón Steady-GLO y el sustrato se equilibraron a temperatura ambiente. Después de la incubación de 4 h, la placa se dejó enfriar a temperatura ambiente durante 5 minutos. Se mezclaron tampón de ensayo y sustrato juntos y se añadieron a 100 µl por pocillo. Las placas se agitaron con vórtex a ajuste bajo durante 1 minuto para la mezcla, luego se incubaron a temperatura ambiente durante 10 minutos. La placa se leyó luego en un luminómetro con integración de 5 segundos.

Los resultados del bioensayo de células de Jurkat se informan en la Figura 2.

EJEMPLO 6: OPTIMIZACIÓN DE SECUENCIAS DE ADAPTADOR PARA ZZ12

35 Basándose en los resultados del ensayo de proliferación de linfocitos B y el bioensayo de células de Jurkat, la construcción de ZZ12 se seleccionó para optimización usando secuencias de adaptador. Se probaron diversas longitudes de adaptador en combinación con la construcción de ZZ12 para actividad biológica óptima usando los bioensayos descritos en los Ejemplos 4 y 5. La Tabla 4 desvela las secuencias de adaptador probadas. Los resultados del ensayo de proliferación de linfocitos B sobre construcciones de ZZ12-APRIL optimizadas para adaptador seleccionadas se informan en la Figura 3. Los resultados del bioensayo de TAC1-Jurkat de estas mismas construcciones se informan en la Figura 4.

Tabla 4

	abcde	fgabcd	
ZZ12.1 (SEC ID N°: 33)	LDLLT	GDL	--APRIL[110-250]
ZZ12.2 (SEC ID N°: 34)	LDLLT	GSGDL	--APRIL[110-250]

(continuación)

	abcde	fgabcd	
ZZ12.3 (SEC ID N°: 35)	LDLLT	L	--APRIL[110-250]
ZZ12.4 (SEC ID N°: 36)	LDLLT	GD	--APRIL[110-250]
ZZ12.5 (SEC ID N°: 37)	LDLLT	GERSL	--APRIL[110-250]
ZZ12.6 (SEC ID N°: 38)	LDLLT	RSGGGS	--APRIL[110-250]

EJEMPLO 7: ELIMINACIÓN ENZIMÁTICA DEL DOMINIO DE CREMALLERA

- 5 Con el fin de verificar que la presencia del dominio de cremallera del extremo amino no altera las propiedades funcionales y biológicas de APRIL, se usó tripsina para eliminar enzimáticamente el dominio ZZ12.6 de la proteína de fusión de ZymoZipper 12.6-April marcada con Flag en el extremo N (NF-ZZ12.6-APRIL). El homotrímero de APRIL es inesperadamente resistente a la escisión con tripsina. Adicionalmente, dos sitios de escisión con tripsina estrechamente separados están localizados en la unión entre el dominio de cremallera y el dominio de APRIL. Estos sitios son accesibles a tripsina y APRIL sin que la cremallera ZZ12.6 pueda obtenerse fácilmente usando el siguiente procedimiento:
- 10 Aproximadamente 24 mg de NF-ZZ12.6-APRIL a 1,2 mg/ml en NaPO₄ 50 mM, NaCl 250 mM, EDTA 2 mM a pH 7,2 se incubaron durante 30 minutos a 37 °C en presencia de una relación de masa 1:100 de tripsina (grado de secuenciación, porcina, Promega). La reacción se inactivó con la adición de 5 ml de 2,5 mg/ml de inhibidor de tripsina AEBSF (Roche) en una disolución de PBS similar.
- 15 APRIL escindido sin el dominio ZZ12.6 se separó de especies sin reaccionar usando cromatografía en heparina. La disolución de reacción se diluyó a NaCl 150 mM usando 17 ml de agua fría. La disolución diluida se cargó sobre una columna AF-Heparin 650-M de 3,4 ml (Toso Bioscience). El flujo continuo recogido se analizó por cromatografía en columna de heparina analítica y SDS-PAGE para confirmar que todas las especies moleculares que contenían APRIL se unieron a la columna de heparina preparativa. La columna se lavó con 10 volúmenes de columna de NaPO₄ 50 mM, NaCl 150 mM, 1 mg/ml de AEBSF a pH 7,0 y luego se eluyó usando un gradiente de 10 volúmenes de columna de NaCl 150 mM a NaCl 1 M en NaPO₄ 50 mM a pH 7,0, que contenía 1 mg/ml de AEBSF, seguido de una etapa de elución con NaCl 2 M. La elución de la proteína de la columna se monitorizó usando la absorbancia de la disolución a 280 nm (A280nm). Adicionalmente, fracciones que incluían las muestras de carga y de flujo continuo se analizaron por SDS-PAGE. Se eluyeron dos picos, el primero correspondiente a NF-ZZ12.6-APRIL intacto, mientras que el segundo pico de elución posterior se correspondió con APRIL sin el dominio ZZ12.6. La proteína en el segundo pico (conjunto de eluato de heparina) se procesó adicionalmente.
- 20
- 25 El conjunto de eluato de heparina que contenía APRIL sin el dominio ZZ12.6 se concentró a un volumen de < 3 ml usando una membrana de centrifuga Ultracel de 10 kD de MWCO (Millipore) y se purificó adicionalmente usando cromatografía de exclusión por tamaño sobre una columna Superdex 200 16/60 (GE Healthcare) con una elución isocrática usando NaPO₄ 50 mM a pH 7,2, NaCl 250 mM. Las fracciones de eluato se analizaron por SDS-PAGE. Se observó una pequeña cantidad de material agregado en la traza de A280 nm. Este material agregado no se incluyó en el eluato reunido (conjunto de eluato de Superdex 200).
- 30
- 35 El conjunto de eluato de Superdex 200 se purificó adicionalmente usando 1 m de resina de afinidad (resina de agarosa anti-FLAG M2: Sigma) con una incubación durante la noche a 4 °C. La resina de afinidad une el extremo amino de la marca de epítipo FLAG (NF) a ZZ12.6 en la molécula de NF-ZZ12.6-APRIL, y se usó para eliminar cualquier NF-ZZ12.6-APRIL sin escindir contaminante restante. APRIL sin el dominio ZZ-12.6 se separó de la resina de afinidad usando una unidad de filtración de 0,22µm (Millipore), se concentró a aproximadamente 1 mg/ml basándose en cálculos estimados de la concentración de heparina analítica usando una membrana centrifuga de 5 kD de MWCO (Millipore), se tomaron alícuotas y se guardó a -80 °C.
- 40 APRIL sin el dominio ZZ12.6 se caracterizó usando cromatografía de exclusión por tamaño analítica con detección por dispersión de la luz de ángulo múltiple y demostró la forma trimérica esperada. La secuenciación del extremo amino

verificó la escisión en dos de los posibles sitios de tripsina: R83 y K90, en una relación de aproximadamente 2:1. La afinidad de APRIL por Atacicept (TACI-Ig) como se mide usando un instrumento Biacore fue la misma para ambas formas moleculares de APRIL (con y sin ZZ12.6). Adicionalmente, ambas formas de APRIL mostraron actividad comparable en el bioensayo de TACI-Jurkat definido.

5 LISTADO DE SECUENCIAS

<110> Moore, Margaret D. Fox, Brian A.

<120> POLIPÉPTIDOS TRIMERIZANTES

<130> 05-19PC

<150> 60/715.883

10 <151> 09/09/2005

<160> 58

<170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 60

15 <212> ADN

<213> Secuencia artificial

<220>

<223> Secuencia de cebador

<400> 1

20 agccaggaaa tccatgccga gttgagacgc ttccgtagag attataagga cgatgacgat 60

<210> 2

<211> 60

<212> ADN

<213> Secuencia artificial

25 <220>

<223> Secuencia de cebador

<400> 2

ggtggcgta atggaacca ggtgcaggac agagtgctgc ttctctgtt tcagatctcc 60

<210> 3

30 <211> 54

<212> ADN

<213> Secuencia artificial

<220>

<223> Secuencia de cebador

35 <400> 3

gattataagg acgatgacga taaggcgga ggtggctcaa aggagttaa caac 54

<210> 4

<211> 57
 <212> ADN
 <213> Secuencia artificial
 <220>
 5 <223> Secuencia de cebador
 <400> 4
 catcttcttg ttcaggttct cgatgcggcg ctccaggttg ttaaactcct ttgagcc 57
 <210> 5
 <211> 51
 10 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Secuencia de cebador
 <400> 5
 15 gagaacctga acaagaagat ggaggacggt ttctggacg tgtggaccta c 51
 <210> 6
 <211> 51
 <212> ADN
 <213> Secuencia artificial
 20 <220>
 <223> Secuencia de cebador
 <400> 6
 ctcttctgt ttcagatctc cctcggcgtt gtaggtccac acgtccagaa a 51
 <210> 7
 25 <211> 57
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Secuencia de cebador
 30 <400> 7
 gattataagg acgatgacga taagggcgga ggtggctcag atgatctgcg cgaggtg 57
 <210> 8
 <211> 57
 <212> ADN
 35 <213> Secuencia artificial

ES 2 415 890 T3

<220>
<223> Secuencia de cebador
<400> 8
ggtcaggac ttctccagat tggagatgga cttctccacc tcgcgagat catctga 57
5 <210> 9
<211> 57
<212> ADN
<213> Secuencia artificial
<220>
10 <223> Secuencia de cebador
<400> 9
aatctggaga agtccctgac ctccctgtcc gaggtgggtgc tccagaaccg ccgcggc 57
<210> 10
<211> 57
15 <212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia de cebador
<400> 10
20 cttcttctgt ttcagatctc cgaacagcag gtccaggccg cggcggttct ggagcac 57
<210> 11
<211> 57
<212> ADN
<213> Secuencia artificial
25 <220>
<223> Secuencia de cebador
<400> 11
gattataagg acgatgacga taagggcgga ggtggctcag atgatctggg cgcctg 57
<210> 12
30 <211> 57
<212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia de cebador
35 <400> 12

ES 2 415 890 T3

ggtcaggac ttctccaggg cggacacgga ttttccagg gcgcccagat catctga 57
<210> 13
<211> 57
<212> ADN
5 <213> Secuencia artificial
<220>
<223> Secuencia de cebador
<400> 13
gccttgaga agtcctgac ctcccttcc gaggtcgtcc tgcagaaccg ccgcggc 57
10 <210> 14
<211> 57
<212> ADN
<213> Secuencia artificial
<220>
15 <223> Secuencia de cebador
<400> 14
cttctctgt ttcatgctc caaaaagaag gtcaaggccg cggcggttct gcaggac 57
<210> 15
<211> 57
20 <212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia de cebador
<400> 15
25 gattataagg acgatgacga taagggcgga ggtggctcag ccgacgtcca gcgcctc 57
<210> 16
<211> 54
<212> ADN
<213> Secuencia artificial
30 <220>
<223> Secuencia de cebador
<400> 16
ggaaagggag tcgaaagat catcgacgcc ctgctggagg cgctggacgt cggc 54
<210> 17
35 <211> 57

ES 2 415 890 T3

<212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia de cebador
5 <400> 17
gatctttccg actcccttc ctcctcgcc gaggtggtgc tgcagaatcg ccgcggc 57
<210> 18
<211> 57
<212> ADN
10 <213> Secuencia artificial
<220>
<223> Secuencia de cebador
<400> 18
cttctctgt ttcagatctc cgaaaagaag atcaaggccg cggcgattct gcagcac 57
15 <210> 19
<211> 57
<212> ADN
<213> Secuencia artificial
<220>
20 <223> Secuencia de cebador
<400> 19
gattataagg acgatgacga taagggcgga ggtggctcaa atggcttccg caagatg 57
<210> 20
<211> 57
25 <212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia de cebador
<400> 20
30 atcgatctgt ttctggatcg taagcatggt ctgggacatc tgcgaaagc catttga 57
<210> 21
<211> 57
<212> ADN
<213> Secuencia artificial
35 <220>

ES 2 415 890 T3

<223> Secuencia de cebador
<400> 21
acgatccaga aacagatcga ttccctgcc gccgtggtgc ttcagaatcg ccgcggc 57
<210> 22
5 <211> 57
<212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia de cebador
10 <400> 22
cttctctgt ttcagatctc ccgtgaggac gtcgaggccg cggcgattct gaagcac 57
<210> 23
<211> 57
<212> ADN
15 <213> Secuencia artificial
<220>
<223> Secuencia de cebador
<400> 23
gattataagg acgatgacga taaggcgga ggtggctcat ccgacgtcca ggcgatc 57
20 <210> 24
<211> 57
<212> ADN
<213> Secuencia artificial
<220>
25 <223> Secuencia de cebador
<400> 24
gtcgacctgg tctgaaggt cctggatggt ggaggagatc gcctggacgt cggatga 57
<210> 25
<211> 57
30 <212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia de cebador
<400> 25
35 gacctcagg accaggtcga ctccctggcg gagtggtgc tgcagaatcg ccgcggc 57

ES 2 415 890 T3

<210> 26

<211> 57

<212> ADN

<213> Secuencia artificial

5 <220>

<223> Secuencia de cebador

<400> 26

cttcttctgt ttcagatctc ccgtaagaag gtcaaggccg cggcgattct gcagcac 57

<210> 27

10 <211> 30

<212> PRT

<213> gripe A

<400> 27

Lys	Glu	Leu	Arg	Glu	Val	Glu	Lys	Ser	Ile	Glu	Asn	Leu	Asn	Lys	Lys
1				5					10					15	
Met	Glu	Asp	Gly	Phe	Leu	Asp	Val	Leu	Gln	Asn	Arg	Arg	Gly		
			20					25					30		

15 <210> 28

<211> 35

<212> PRT

<213> Virus de la leucemia murina de Moloney

<400> 28

Asp	Asp	Leu	Arg	Glu	Val	Glu	Lys	Ser	Ile	Ser	Asn	Leu	Glu	Lys	Ser
1				5					10					15	
Leu	Thr	Ser	Leu	Ser	Glu	Val	Val	Leu	Gln	Asn	Arg	Arg	Gly	Leu	Asp
			20					25					30		
Leu	Leu	Phe													
		35													

20 <210> 29

<211> 35

<212> PRT

<213> Mus musculus

25 <400> 29

Asp	Asp	Leu	Gly	Ala	Leu	Glu	Lys	Ser	Val	Ser	Ala	Leu	Glu	Lys	Ser
1				5					10					15	
Leu	Thr	Ser	Leu	Ser	Glu	Val	Val	Leu	Gln	Asn	Arg	Arg	Gly	Leu	Asp
			20					25					30		
Leu	Leu	Phe													
		35													

<210> 30

<211> 35

<212> PRT

ES 2 415 890 T3

<213> Mus musculus

<400> 30

```

Ala Asp Val Gln Arg Leu Gln Gln Gly Val Asp Asp Leu Ser Asp Ser
 1      5      10      15
Leu Ser Ser Leu Ala Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp
      20      25      30
Leu Leu Phe
      35
    
```

<210> 31

5

<211> 35

<212> PRT

<213> Rat norvegicus

<400> 31

```

Asn Gly Phe Arg Lys Met Ser Gln Thr Met Leu Thr Ile Gln Lys Gln
 1      5      10      15
Ile Asp Ser Leu Ala Ala Val Val Leu Gln Asn Arg Arg Gly Leu Asp
      20      25      30
Val Leu Thr
      35
    
```

10

<210> 32

<211> 35

<212> PRT

<213> Virus Mason-Pfizer del mono

<400> 32

```

Ser Asp Val Gln Ala Ile Ser Ser Thr Ile Gln Asp Leu Gln Asp Gln
 1      5      10      15
Val Asp Ser Leu Ala Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp
      20      25      30
Leu Leu Thr
      35
    
```

15

<210> 33

<211> 8

<212> PRT

<213> Secuencia artificial

20

<220>

<223> Héptada final más adaptador

<400> 33

```

Leu Asp Leu Leu Thr Gly Asp Leu
 1      5
    
```

<210> 34

25

<211> 10

<212> PRT

<213> Secuencia artificial

<220>

ES 2 415 890 T3

<223> Héptada final más adaptador
 <400> 34
 Leu Asp Leu Leu Thr Gly Ser Gly Asp Leu
 1 5 10

5
 <210> 35
 <211> 6
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Héptada final más adaptador

10
 <400> 35
 Leu Asp Leu Leu Thr Leu
 1 5

<210> 36
 <211> 7
 <212> PRT

15
 <213> Secuencia artificial
 <220>
 <223> Héptada final más adaptador

<400> 36
 Leu Asp Leu Leu Thr Gly Asp
 1 5

20
 <210> 37
 <211> 10
 <212> PRT
 <213> Secuencia artificial
 <220>

25
 <223> Héptada final más adaptador

<400> 37
 Leu Asp Leu Leu Thr Gly Glu Arg Ser Leu
 1 5 10

<210> 38
 <211> 11

30
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Héptada final más adaptador

ES 2 415 890 T3

<400> 38

Leu Asp Leu Leu Thr Arg Ser Gly Gly Gly Ser
 1 5 10

<210> 39

<211> 105

5

<212> ADN

<213> Secuencia artificial

<220>

<223> conductor de tPA optimizado

<220>

10

<221> CDS

<222> (1)...(105)

<400> 39

atg gat gca atg aag aga ggg ctc tgc tgt gtg ctg ctg ctg tgt ggc 48
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15

gcc gtc ttc gtt tcg ctc agc cag gaa atc cat gcc gag ttg aga cgc 96
 Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg
 20 25 30

ttc cgt aga 105
 Phe Arg Arg
 35

<210> 40

15

<211> 35

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia conductora de tPA optimizada

20

<400> 40

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
 Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg
 20 25 30
 Phe Arg Arg
 35

<210> 41

<211> 222

<212> PRT

25

<213> Secuencia artificial

<220>

<223> ZZ3-APRIL

ES 2 415 890 T3

<400> 41

```

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

```

<210> 42

5 <211> 666

<212> ADN

<213> Secuencia artificial

<220>

<223> Consenso de ZZ3-APRIL

10 <220>

<221> misc_feature

<222> 9, 18, 21, 24, 33, 36, 39, 42, 48, 51, 54, 60, 63, 66, 69,

84, 90, 93, 96, 102, 105, 123, 129, 132, 144, 165, 171, 177, 183, 192, 222, 225, 228, 231, 234, 237, 243, 264, 267,
270, 276, 279, 282, 291, 294, 297, 309, 315, 318, 324, 336

15 <223> n = A,T,C o G

<220>

<221> misc_feature

<222> 339, 342, 345, 348, 351, 354, 357, 360, 366, 372, 378, 381,

384, 396, 399, 402, 408, 411, 417, 423, 426, 438, 441, 447, 453, 459, 462, 465, 468, 474, 480, 483, 492, 495, 501,
510, 513, 519, 522, 528, 534, 537, 546, 555, 558, 561, 564

<223> n = A,T,C o G

<220>

<221> misc_feature

ES 2 415 890 T3

<222> 573, 582, 591, 594, 597, 606, 609, 612, 615, 618, 624, 630,

633, 636, 642, 645, 651, 654, 660, 666

<223> n = A,T,C o G

<400> 42

```
atggaygcna tgaarmgngg nytntgytgy gtntyntytny tntgyggngc ngntntygtn 60
wsnytnwsnc argarathca ygcngarytn mgnmgnttym gnmgnaarga rttyaayaay 120
ytngarmgnm gnathgaraa yytnaayaar aaratggarg ayggnttyyt ngaygtntgg 180
acntayaayg cngargayta yaargaygay gaygayaarg gngngggngg nwsngggngay 240
ytnaarcara araarcarca ywsngtntytn cayytngtnc cnathaaygc nacnwsnaar 300
gaygaywsng aygtnacnga rgtnatgtgg carccngcny tnmgnmgngg nmgnngnytn 360
cargcncarg gntayggngt nmgnathcar gaygcngngg tntayytnyt ntaywsncar 420
gtnynttyc argaygtnac nttyacnatg ggncargtng tnwsnmngna rggncarggn 480
mgnccargara cnynttytg ntgyathmgn wsnatgccnw sncayccnga ymgngcntay 540
aaywsntgyt aywsngcngg ngnttycay ytncaycarg gngayathyt nwsngtnath 600
athccnmngg cnmngncnaa rytnaaytn wsncncayg gnacnttyt nggnttygtn 660
aarytn 666
```

5

<210> 43

<211> 684

<212> ADN

<213> Secuencia artificial

10

<220>

<223> ZZ7-APRIL

<220>

<221> CDS

<222> (1) ... (684)

15

<400> 43

ES 2 415 890 T3

```

atg gat gca atg aag aga ggg ctc tgc tgt gtg ctg ctg ctg tgt ggc 48
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

gcc gtc ttc gtt tgc ctc agc cag gaa atc cat gcc gag ttg aga cgc 96
Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg
20 25 30

ttc cgt aga gat tat aag gac gat gac gat aag ggc gga ggt ggc tca 144
Phe Arg Arg Asp Tyr Lys Asp Asp Asp Asp Lys Gly Gly Gly Gly Ser
35 40 45

gat gat ctg cgc gag gtg gag aag tcc atc tcc aat ctg gag aag tcc 192
Asp Asp Leu Arg Glu Val Glu Lys Ser Ile Ser Asn Leu Glu Lys Ser
50 55 60

ctg acc tcc ctg tcc gag gtg gtg ctc cag aac cgc cgc ggc ctg gac 240
Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp
65 70 75 80

ctg ctg ttc gga gat ctg aaa cag aag aag cag cac tct gtc ctg cac 288
Leu Leu Phe Gly Asp Leu Lys Gln Lys Lys Gln His Ser Val Leu His
85 90 95

ctg gtt ccc att aac gcc acc tcc aag gat gac tcc gat gtg aca gag 336
Leu Val Pro Ile Asn Ala Thr Ser Lys Asp Asp Ser Asp Val Thr Glu
100 105 110

gtg atg tgg caa cca gct ctt agg cgt ggg aga ggc cta cag gcc caa 384
Val Met Trp Gln Pro Ala Leu Arg Arg Gly Arg Gly Leu Gln Ala Gln
115 120 125

gga tat ggt gtc cga atc cag gat gct gga gtt tat ctg ctg tat agc 432
Gly Tyr Gly Val Arg Ile Gln Asp Ala Gly Val Tyr Leu Leu Tyr Ser
130 135 140

cag gtc ctg ttt caa gac gtg act ttc acc atg ggt cag gtg gtg tct 480
Gln Val Leu Phe Gln Asp Val Thr Phe Thr Met Gly Gln Val Val Ser
145 150 155 160

cga gaa ggc caa gga agg cag gag act cta ttc cga tgt ata aga agt 528
Arg Glu Gly Gln Gly Arg Gln Glu Thr Leu Phe Arg Cys Ile Arg Ser
165 170 175

atg ccc tcc cac ccg gac cgg gcc tac aac agc tgc tat agc gca ggt 576
Met Pro Ser His Pro Asp Arg Ala Tyr Asn Ser Cys Tyr Ser Ala Gly
180 185 190

gtc ttc cat tta cac caa ggg gat att ctg agt gtc ata att ccc cgg 624
Val Phe His Leu His Gln Gly Asp Ile Leu Ser Val Ile Ile Pro Arg
195 200 205

gca agg gcg aaa ctt aac ctc tct cca cat gga acc ttc ctg ggg ttt 672
Ala Arg Ala Lys Leu Asn Leu Ser Pro His Gly Thr Phe Leu Gly Phe
210 215 220

gtg aaa ctg taa 684
Val Lys Leu *
225

```

<210> 44

<211> 227

5

<212> PRT

<213> Secuencia artificial

<220>

<223> ZZ7-APRIL

<400> 44

```

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

```

5 <210> 45

<211> 672

<212> ADN

<213> Secuencia artificial

<220>

10 <223> Consenso de ZZ7-APRIL

<220>

<221> misc_feature

<222> 9, 18, 21, 24, 33, 36, 39, 42, 48,-51, 54, 60, 63, 66, 69,

15 84, 90, 93, 96, 102, 105, 132, 135, 138, 141, 144, 153, 156, 162, 171, 177, 183, 192, 195, 198, 201, 204, 207, 213, 216, 219, 228, 231, 234, 237, 243, 246, 252, 258, 279, 282

<223> n = A,T,C o G

<220>

<221> misc_feature

<222> 285, 291, 294, 297, 306, 309, 312, 324, 330, 333, 342, 345,

20 348, 351, 354, 357, 360, 363, 366, 372, 378, 384, 387, 390, 402, 405, 408, 414, 417, 423, 429, 432, 444, 447, 453, 459, 465, 468, 471, 474, 480, 486, 489, 498, 501, 507, 516

<223> n = A,T,C o G

ES 2 415 890 T3

<220>

<221> misc_feature

<222> 519, 525, 528, 534, 540, 543, 552, 561, 564, 567, 570, 579,

588, 597, 600, 603, 612, 615, 618, 621, 624, 630, 636, 639, 642, 648, 651, 657, 660, 666, 672

5

<223> n = A,T,C o G

<400> 45

```

atggaygcna tgaarmgngg nytntgytgy gtntyntny tntgyggngc ngntnttygtn 60
wsnytnwsnc argarathca ygcngarytn mgnmgnntym gnmngayta yaargaygay 120
gaygayaarg gnggnggngg nwsngaygay ytnmngarg tngaraarws nathwsnaay 180
ytngaraarw snytnacnws nytnwsngar gtngtntnyc araaymngmg nggnytngay 240
ytntynttyg gngayytnaa rcaraaraar carcaywsng tnytncaayt ngtnccnath 300
aaygcnacnw snaargayga ywsngaygtn acngarcarc cngcnytnmg nmngggnmgn 360
ggnytncarg cncarggnta yggngtnmgn athcargayg cngngtnta yytntntay 420
wscargtny tnttycarga ygtnacntty acnatgggnc argtngtnws nmngarggn 480
carggnmgnc argaracnyt nttymgntgy athmgnwsna tgccnwsnca yccngaymgn 540
gcntayaayw sntgytayws ngcngngtn ttycaytnc aycarggnga yathytnwsn 600
gtnathathc cnmngcnmg ngcnaarytn aaytnwsnc cncayggnac nttyytnggn 660
ttygtnaary tn 672

```

<210> 46

<211> 684

10

<212> ADN

<213> Secuencia artificial

<220>

<223> ZZ9-APRIL

<220>

15

<221> CDS

<222> (1)...(684)

<400> 46

```

atg gat gca atg aag aga ggg ctc tgc tgt gtg ctg ctg ctg tgt ggc 48
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

gcc gtc ttc gtt tcg ctc agc cag gaa atc cat gcc gag ttg aga cgc 96
Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg
20 25 30

ttc cgt aga gat tat aag gac gat gac gat aag ggc gga ggt ggc tca 144
Phe Arg Arg Asp Tyr Lys Asp Asp Asp Asp Lys Gly Gly Gly Gly Ser
35 40 45

gat gat ctg gcc gcc ctg gaa aaa tcc gtg tcc gcc ctg gag aag tcc 192
Asp Asp Leu Gly Ala Leu Glu Lys Ser Val Ser Ala Leu Glu Lys Ser

```

ES 2 415 890 T3

50	55	60	
ctg acc tcc ctt tcc gag gtc gtc ctg cag aac cgc cgc ggc ctt gac			240
Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp			
65	70	75	80
ctt ctt ttt gga gat ctg aaa cag aag aag cag cac tct gtc ctg cac			288
Leu Leu Phe Gly Asp Leu Lys Gln Lys Lys Gln His Ser Val Leu His			
	85	90	95
ctg gtt ccc att aac gcc acc tcc aag gat gac tcc gat gtg aca gag			336
Leu Val Pro Ile Asn Ala Thr Ser Lys Asp Asp Ser Asp Val Thr Glu			
	100	105	110
gtg atg tgg caa cca gct ctt agg cgt ggg aga ggc cta cag gcc caa			384
Val Met Trp Gln Pro Ala Leu Arg Arg Gly Arg Gly Leu Gln Ala Gln			
	115	120	125
gga tat ggt gtc cga atc cag gat gct gga gtt tat ctg ctg tat agc			432
Gly Tyr Gly Val Arg Ile Gln Asp Ala Gly Val Tyr Leu Leu Tyr Ser			
	130	135	140
cag gtc ctg ttt caa gac gtg act ttc acc atg ggt cag gtg gtg tct			480
Gln Val Leu Phe Gln Asp Val Thr Phe Thr Met Gly Gln Val Val Ser			
	145	150	155
cga gaa ggc caa gga agg cag gag act cta ttc cga tgt ata aga agt			528
Arg Glu Gly Gln Gly Arg Gln Glu Thr Leu Phe Arg Cys Ile Arg Ser			
	165	170	175
atg ccc tcc cac ccg gac cgg gcc tac aac agc tgc tat agc gca ggt			576
Met Pro Ser His Pro Asp Arg Ala Tyr Asn Ser Cys Tyr Ser Ala Gly			
	180	185	190
gtc ttc cat tta cac caa ggg gat att ctg agt gtc ata att ccc cgg			624
Val Phe His Leu His Gln Gly Asp Ile Leu Ser Val Ile Ile Pro Arg			
	195	200	205
gca agg gcg aaa ctt aac ctc tct cca cat gga acc ttc ctg ggg ttt			672
Ala Arg Ala Lys Leu Asn Leu Ser Pro His Gly Thr Phe Leu Gly Phe			
	210	215	220
gtg aaa ctg taa			684
Val Lys Leu *			
225			

<210> 47

<211> 227

<212> PRT

5 <213> Secuencia artificial

<220>

<223> ZZ9-APRIL

<400> 47

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15
Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg
20 25 30

ES 2 415 890 T3

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

<210> 48

<211> 681

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Consenso de ZZ9-APRIL

<220>

<221> misc_feature

10 <222> 9, 18, 21, 24, 33, 36, 39, 42, 48, 51, 54, 60, 63, 66, 69,

84, 90, 93, 96, 102, 105, 132, 135, 138, 141, 144, 153, 156, 159, 162, 171, 174, 177, 180, 183, 192, 195, 198, 201, 204, 207, 213, 216, 219, 228, 231, 234, 237, 243, 246, 252

<223> n = A,T,C o G

<220>

15 <221> misc_feature

<222> 258, 279, 282, 285, 291, 294, 297, 306, 309, 312, 324, 330,

333, 339, 351, 354, 357, 360, 363, 366, 369, 372, 375, 381, 387, 393, 396, 399, 411, 414, 417, 423, 426, 432, 438, 441, 453, 456, 462, 468, 474, 477, 480, 483, 489, 495, 498

<223> n = A,T,C o G

20 <220>

<221> misc_feature

<222> 507, 510, 516, 525, 528, 534, 537, 543, 549, 552, 561, 570,

573, 576, 579, 588, 597, 606, 609, 612, 621, 624, 627, 630, 633, 639, 645, 648, 651, 657, 660, 666, 669, 675, 681

<223> n = A,T,C o G

<400> 48

```

atggaygcna tgaarmgngg nytntgytgy gtntyntytny tntgyggngc ngntnttygtn 60
wsnytnwsnc argarathca ygcngarytn mgnmgnttym gnmnggayta yaargaygay 120
gaygayaarg gngggngggg nwsngaygay ytngggngcny tngaraarws ngtnwsngcn 180
ytngaraarw snytnacnws nytnwsngar gtngtntytnc araaymgnmg nggnytngay 240
ytntynttyg gngayytnaa rcaraaraar carcaywsng tnytncayyt ngtnccnath 300
aaygcnacnw snaargayga ywsngaygtn acngargtna tgtggcarcc ngcnytnmgn 360
mgnggnmgng gnytncargc ncarggntay ggngtnmna thcargaygc nggngtntay 420
ytntyntayw sncargtnyt nttycargay gtnacnttya cnatgggnca rgtngtnwsn 480
mgngarggnc arggnmgnca rgaracnytn ttymngtgya thmgnwsnat gccnwsncay 540
ccngaymgng cntayaayws ntgytaywsn gcngnggtnt tycayytnc ycarggngay 600
athytnwsng tnathathcc nmngncnmgn gcnaarytna ayytnwsncc ncayggnacn 660
ttytnggnt tygtnaaryt n 681
    
```

<210> 49

<211> 684

5

<212> ADN

<213> Secuencia artificial

<220>

<223> ZZ10-APRIL

<220>

10

<221> CDS

<222> (1)...(684)

<400> 49

```

atg gat gca atg aag aga ggg ctc tgc tgt gtg ctg ctg ctg tgt ggc 48
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

gcc gtc ttc gtt tcg ctc agc cag gaa atc cat gcc gag ttg aga cgc 96
Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg
20 25 30

ttc cgt aga gat tat aag gac gat gac gat aag ggc gga ggt ggc tca 144
Phe Arg Arg Asp Tyr Lys Asp Asp Asp Asp Lys Gly Gly Gly Gly Ser
35 40 45

gcc gac gtc cag cgc ctc cag cag ggc gtc gat gat ctt tcc gac tcc 192
Ala Asp Val Gln Arg Leu Gln Gln Gly Val Asp Asp Leu Ser Asp Ser
50 55 60

ctt tcc tcc ctc gcc gag gtg gtg ctg cag aat cgc cgc ggc ctt gat 240
Leu Ser Ser Leu Ala Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp
65 70 75 80

ctt ctt ttc gga gat ctg aaa cag aag aag cag cac tct gtc ctg cac 288
Leu Leu Phe Gly Asp Leu Lys Gln Lys Lys Gln His Ser Val Leu His
85 90 95

ctg gtt ccc att aac gcc acc tcc aag gat gac tcc gat gtg aca gag 336
Leu Val Pro Ile Asn Ala Thr Ser Lys Asp Asp Ser Asp Val Thr Glu
100 105 110

gtg atg tgg caa cca gct ctt agg cgt ggg aga ggc cta cag gcc caa 384
Val Met Trp Gln Pro Ala Leu Arg Arg Gly Arg Gly Leu Gln Ala Gln
115 120 125

gga tat ggt gtc cga atc cag gat gct gga gtt tat ctg ctg tat agc 432
Gly Tyr Gly Val Arg Ile Gln Asp Ala Gly Val Tyr Leu Leu Tyr Ser
    
```

ES 2 415 890 T3

130	135	140	
cag gtc ctg ttt caa gac gtg act ttc acc atg ggt cag gtg gtg tct			480
Gln Val Leu Phe Gln Asp Val Thr Phe Thr Met Gly Gln Val Val Ser			
145	150	155	160
cga gaa ggc caa gga agg cag gag act cta ttc cga tgt ata aga agt			528
Arg Glu Gly Gln Gly Arg Gln Glu Thr Leu Phe Arg Cys Ile Arg Ser			
	165	170	175
atg ccc tcc cac ccg gac cgg gcc tac aac agc tgc tat agc gca ggt			576
Met Pro Ser His Pro Asp Arg Ala Tyr Asn Ser Cys Tyr Ser Ala Gly			
	180	185	190
gtc ttc cat tta cac caa ggg gat att ctg agt gtc ata att ccc cgg			624
Val Phe His Leu His Gln Gly Asp Ile Leu Ser Val Ile Ile Pro Arg			
	195	200	205
gca agg gcg aaa ctt aac ctc tct cca cat gga acc ttc ctg ggg ttt			672
Ala Arg Ala Lys Leu Asn Leu Ser Pro His Gly Thr Phe Leu Gly Phe			
	210	215	220
gtg aaa ctg taa			684
Val Lys Leu *			
225			

<210> 50

<211> 227

<212> PRT

5 <213> Secuencia artificial

<220>

<223> ZZ10-APRIL

<400> 50

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

ES 2 415 890 T3

Val	Phe	His	Leu	His	Gln	Gly	Asp	Ile	Leu	Ser	Val	Ile	Ile	Pro	Arg
		195					200					205			
Ala	Arg	Ala	Lys	Leu	Asn	Leu	Ser	Pro	His	Gly	Thr	Phe	Leu	Gly	Phe
	210					215					220				
Val	Lys	Leu													
225															

<210> 51

<211> 681

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Consenso de ZZ10-APRIL

<220>

<221> misc_feature

10 <222> 9, 18, 21, 24, 33, 36, 39, 42, 48, 51, 54, 60, 63, 66, 69,

84, 90, 93, 96, 102, 105, 132, 135, 138, 141, 144, 147, 153, 159, 162, 171, 174, 183, 186, 192, 195, 198, 201, 204, 207, 213, 216, 219, 228, 231, 234, 237, 243, 246, 252, 258

<223> n = A,T,C o G

<220>

15 <221> misc_feature

<222> 279, 282, 285, 291, 294, 297, 306, 309, 312, 324, 330, 333,

339, 351, 354, 357, 360, 363, 366, 369, 372, 375, 381, 387, 393, 396, 399, 411, 414, 417, 423, 426, 432, 438, 441, 453, 456, 462, 468, 474, 477, 480, 483, 489, 495, 498, 507

<223> n = A,T,C o G

20 <220>

<221> misc_feature

<222> 510, 516, 525, 528, 534, 537, 543, 549, 552, 561, 570, 573,

576, 579, 588, 597, 606, 609, 612, 621, 624, 627, 630, 633, 639, 645, 648, 651, 657, 660, 666, 669, 675, 681

<223> n = A,T,C o G

25 <400> 51

```

atggaygcna tgaarmgnngg nytnntgytgy gtnytnytny tntgyggngc ngtnnttygtn 60
wsnytnwsnc argarathca ygcngarytn mgnmgnntty mnmngayta yaargaygay 120
gaygayaarg gnggngngngg nwsngcngay gtncarmgny tncarcargg ngtnngaygay 180
ytnwsngayw snytnwsnws nytnngcngar gtngtnytnc araaymgnmg nggnytnngay 240
ytnytnnttyg gngayytnaa rcaraaraar carcaywsng tnytncayyt ngtnccnath 300
aaygcnacnw snaargayga ywsngaygtn acngargtna tgtggcarcc ngcnytnmgn 360
mngngnmngng gnytncargc ncarggntay ggngtnmna thcargaygc nggngntay 420
ytnytnntayw sncargtnyt nttycargay gtnacnttya cnatgggnca rgtngtnwsn 480
mngngarggnc arggnmgnca rgaracnytn tтымngtgya thmgnwsnat gccnwsncay 540
ccngaymgng cntayaayws ntgytaywsn gcngngntnt tycayytnc ycarggngay 600
athytnwsng tnathathcc nmngncnmgn gcnaarytna ayytnwsncc ncayggnacn 660
ttytnngnt tygtnaaryt n 681

```

<210> 52

<211> 684

<212> ADN

<213> Secuencia artificial

<220>

<223> ZZ11-APRIL

5

<220>

<221> CDS

<222> (1)...(684)

<400> 52

ES 2 415 890 T3

atg gat gca atg aag aga ggg ctc tgc tgt gtg ctg ctg ctg tgt ggc	48
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly	
1 5 10 15	
gcc gtc ttc gtt tcg ctc agc cag gaa atc cat gcc gag ttg aga cgc	96
Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg	
20 25 30	
ttc cgt aga gat tat aag gac gat gac gat aag ggc gga ggt ggc tca	144
Phe Arg Arg Asp Tyr Lys Asp Asp Asp Asp Lys Gly Gly Gly Ser	
35 40 45	
aat ggc ttt cgc aag atg tcc cag acc atg ctt acg atc cag aaa cag	192
Asn Gly Phe Arg Lys Met Ser Gln Thr Met Leu Thr Ile Gln Lys Gln	
50 55 60	
atc gat tcc ctt gcc gcc gtg gtg ctt cag aat cgc cgc ggc ctc gac	240
Ile Asp Ser Leu Ala Ala Val Val Leu Gln Asn Arg Arg Gly Leu Asp	
65 70 75 80	
gtc ctc acg gga gat ctg aaa cag aag aag cag cac tct gtc ctg cac	288
Val Leu Thr Gly Asp Leu Lys Gln Lys Lys Gln His Ser Val Leu His	
85 90 95	
ctg gtt ccc att aac gcc acc tcc aag gat gac tcc gat gtg aca gag	336
Leu Val Pro Ile Asn Ala Thr Ser Lys Asp Asp Ser Asp Val Thr Glu	
100 105 110	
gtg atg tgg caa cca gct ctt agg cgt ggg aga ggc cta cag gcc caa	384
Val Met Trp Gln Pro Ala Leu Arg Arg Gly Arg Gly Leu Gln Ala Gln	
115 120 125	
gga tat ggt gtc cga atc cag gat gct gga gtt tat ctg ctg tat agc	432
Gly Tyr Gly Val Arg Ile Gln Asp Ala Gly Val Tyr Leu Leu Tyr Ser	
130 135 140	
cag gtc ctg ttt caa gac gtg act ttc acc atg ggt cag gtg gtg tct	480
Gln Val Leu Phe Gln Asp Val Thr Phe Thr Met Gly Gln Val Val Ser	
145 150 155 160	
cga gaa ggc caa gga agg cag gag act cta ttc cga tgt ata aga agt	528
Arg Glu Gly Gln Gly Arg Gln Glu Thr Leu Phe Arg Cys Ile Arg Ser	
165 170 175	
atg ccc tcc cac ccg gac cgg gcc tac aac agc tgc tat agc gca ggt	576
Met Pro Ser His Pro Asp Arg Ala Tyr Asn Ser Cys Tyr Ser Ala Gly	
180 185 190	
gtc ttc cat tta cac caa ggg gat att ctg agt gtc ata att ccc cgg	624
Val Phe His Leu His Gln Gly Asp Ile Leu Ser Val Ile Ile Pro Arg	
195 200 205	
gca agg gcg aaa ctt aac ctc tct cca cat gga acc ttc ctg ggg ttt	672
Ala Arg Ala Lys Leu Asn Leu Ser Pro His Gly Thr Phe Leu Gly Phe	
210 215 220	
gtg aaa ctg taa	684
Val Lys Leu *	
225	

<210> 53

<211> 227

5 <212> PRT

<213> Secuencia artificial

<220>

<223> ZZ11-APRIL

<400> 53

```

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

```

5 <210> 54

<211> 681

<212> ADN

<213> Secuencia artificial

<220>

10 <223> Consenso de ZZ11-APRIL

<220>

<221> misc_feature

<222> 9, 18, 21, 24, 33, 36, 39, 42, 48, 51, 54, 60, 63, 66, 69,

15 84, 90, 93, 96, 102, 105, 132, 135, 138, 141, 144, 150, 156, 165, 171, 177, 180, 201, 204, 207, 210, 213, 216, 219,
228, 231, 234, 237, 243, 246, 249, 252, 258, 279, 282, 285

<223> n = A,T,C o G

<220>

<221> misc_feature

<222> 291, 294, 297, 306, 309, 312, 324, 330, 333, 339, 351, 354,

20 357, 360, 363, 366, 369, 372, 375, 381, 387, 393, 396, 399, 411, 414, 417, 423, 426, 432, 438, 441, 453, 456, 462,
468, 474, 477, 480, 483, 489, 495, 498, 507, 510, 516, 525

<223> n = A,T,C o G

ES 2 415 890 T3

<220>

<221> misc_feature

<222> 528, 534, 537, 543, 549, 552, 561, 570, 573, 576, 579, 588,

597, 606, 609, 612, 621, 624, 627, 630, 633, 639, 645, 648, 651, 657, 660, 666, 669, 675, 681

5

<223> n = A,T,C o G

<400> 54

```

atggaygcna tgaarmgnng nytnngytgy gtnytnytny tntgyggngc ngtnnttygtn 60
wsnytnwsnc argarathca ygcngarytn mgnmgnnttym gnmnggayta yaargaygay 120
gaygayaarg gngngngngg nwsnaayggn ttymgnaara tgwsncarac natgytnacn 180
athcaraarc arathgayws nytnngcngcn gtngtnytnc araaymgnmg nggnytnngay 240
gtnytnacng gngayytnaa rcaraaraar carcaywsng tnytncaayt ngtnccnath 300
aaygcnacnw snaargayga ywsngaygtn acngargtna tgtggcarcc ngcnytnmgn 360
mngngnmng gnytnchargc ncarggntay ggngtnmgna thcargaygc nggngntay 420
ytnyntayw sncargtnyt nttycargay gtnacnttya cnatggngca rgtngtnwsn 480
mngarggnc arggnmgna rgaracnytn ttymgntgya thmgnwsnat gccnwsncay 540
ccngaymng cntayaayws ntgytaywsn gcngngntnt tycayytnc ycarngngay 600
athytnwsng tnathathcc nmngncmgn gcnaarytna ayttnwsnc ncayggnaen 660
ttytnggnt tygtnaaryt n 681
    
```

<210> 55

<211> 684

10

<212> ADN

<213> Secuencia artificial

<220>

<223> ZZ12-APRIL

<220>

15

<221> CDS

<222> (1)...(684)

<400> 55

```

atg gat gca atg aag aga ggg ctc tgc tgt gtg ctg ctg ctg tgt ggc 48
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

gcc gtc ttc gtt tcg ctc agc cag gaa atc cat gcc gag ttg aga cgc 96
Ala Val Phe Val Ser Leu Ser Gln Glu Ile His Ala Glu Leu Arg Arg
20 25 30

ttc cgt aga gat tat aag gac gat gac gat aag ggc gga ggt ggc tca 144
Phe Arg Arg Asp Tyr Lys Asp Asp Asp Asp Lys Gly Gly Gly Gly Ser
35 40 45

tcc gac gtc cag gcg atc tcc tcc acc atc cag gac ctt cag gac cag 192
Ser Asp Val Gln Ala Ile Ser Ser Thr Ile Gln Asp Leu Gln Asp Gln
    
```

ES 2 415 890 T3

50	55	60	
gtc gac tcc ctg gcg gag gtg gtg ctg cag aat cgc cgc ggc ctt gac			240
Val Asp Ser Leu Ala Glu Val Val Leu Gln Asn Arg Arg Gly Leu Asp			
65	70	75	80
ctt ctt acg gga gat ctg aaa cag aag aag cag cac tct gtc ctg cac			288
Leu Leu Thr Gly Asp Leu Lys Gln Lys Lys Gln His Ser Val Leu His			
	85	90	95
ctg gtt ccc att aac gcc acc tcc aag gat gac tcc gat gtg aca gag			336
Leu Val Pro Ile Asn Ala Thr Ser Lys Asp Asp Ser Asp Val Thr Glu			
	100	105	110
gtg atg tgg caa cca gct ctt agg cgt ggg aga ggc cta cag gcc caa			384
Val Met Trp Gln Pro Ala Leu Arg Arg Gly Arg Gly Leu Gln Ala Gln			
	115	120	125
gga tat ggt gtc cga atc cag gat gct gga gtt tat ctg ctg tat agc			432
Gly Tyr Gly Val Arg Ile Gln Asp Ala Gly Val Tyr Leu Leu Tyr Ser			
	130	135	140
cag gtc ctg ttt caa gac gtg act ttc acc atg ggt cag gtg gtg tct			480
Gln Val Leu Phe Gln Asp Val Thr Phe Thr Met Gly Gln Val Val Ser			
	145	150	155
cga gaa ggc caa gga agg cag gag act cta ttc cga tgt ata aga agt			528
Arg Glu Gly Gln Gly Arg Gln Glu Thr Leu Phe Arg Cys Ile Arg Ser			
	165	170	175
atg ccc tcc cac ccg gac cgg gcc tac aac agc tgc tat agc gca ggt			576
Met Pro Ser His Pro Asp Arg Ala Tyr Asn Ser Cys Tyr Ser Ala Gly			
	180	185	190
gtc ttc cat tta cac caa ggg gat att ctg agt gtc ata att ccc cgg			624
Val Phe His Leu His Gln Gly Asp Ile Leu Ser Val Ile Ile Pro Arg			
	195	200	205
gca agg gcg aaa ctt aac ctc tct cca cat gga acc ttc ctg ggg ttt			672
Ala Arg Ala Lys Leu Asn Leu Ser Pro His Gly Thr Phe Leu Gly Phe			
	210	215	220
gtg aaa ctg taa			684
Val Lys Leu			
225			

<210> 56

<211> 227

<212> PRT

5 <213> Secuencia artificial

<220>

<223> ZZ12-APRIL

<400> 56

Met	Asp	Ala	Met	Lys	Arg	Gly	Leu	Cys	Cys	Val	Leu	Leu	Leu	Cys	Gly
1			5					10						15	
Ala	Val	Phe	Val	Ser	Leu	Ser	Gln	Glu	Ile	His	Ala	Glu	Leu	Arg	Arg
			20					25						30	

ES 2 415 890 T3

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

<210> 57

<211> 681

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Consenso de ZZ12-APRIL

<220>

<221> misc_feature

10 <222> 9, 18, 21, 24, 33, 36, 39, 42, 48, 51, 54, 60, 63, 66, 69,

84, 90, 93, 96, 102, 105, 132, 135, 138, 141, 144, 147, 153, 159, 165, 168, 171, 183, 195, 201, 204, 207, 213, 216, 219, 228, 231, 234, 237, 243, 246, 249, 252, 258, 279, 282

<223> n = A,T,C o G

<220>

15 <221> misc_feature

<222> 285, 291, 294, 297, 306, 309, 312, 324, 330, 333, 339, 351,

354, 357, 360, 363, 366, 369, 372, 375, 381, 387, 393, 396, 399, 411, 414, 417, 423, 426, 432, 438, 441, 453, 456, 462, 468, 474, 477, 480, 483, 489, 495, 498, 507, 510, 516

<223> n = A,T,C o G

20 <220>

<221> misc_feature

<222> 525, 528, 534, 537, 543, 549, 552, 561, 570, 573, 576, 579,

588, 597, 606, 609, 612, 621, 624, 627, 630, 633, 639, 645, 648, 651, 657, 660, 666, 669, 675, 681

<223> n = A,T,C o G

<400> 57

```

atggaygcna tgaarmgngg nytntgytgy gtnytnytny tntgyggngc ngtnnttygtn 60
wsnytnwsnc argarathca ygcngarytn mgnmgnttym gnmnggayta yaargaygay 120
gaygayaarg gnggnggngg nwsnwsngay gtncargcna thwsnwsnac nathcargay 180
ytncargayc argtngayws nytngcngar gtngtnytnc araaymgng nggnytngay 240
ytnytnacng gngayytnaa rcaraaraar carcaywsng tnytncayyt ngtnccnath 300
aaygcnacnw snaargayga ywsngaygtn acngargtna tgtggcarcc ngcnytnmgn 360
mgnggnmgng gnytncargc ncarggntay ggngtnmgna thcargaygc nggngtntay 420
ytnytnayw sncargtnyt nttycargay gtnacnttya cnatggngca rgtngtnwsn 480
mgngarggnc arggnmgnc rgaracnytn ttymngtgya thmgwnsnat gccnwsncay 540
ccngaymgng cntayaayws ntgytaywsn gcnggngtnt tycayytncaycarggngay 600
athytnwsng tnathathcc nmngncnmgn gcnaarytna ayytnwsncc ncaygggnacn 660
ttyytnngnt tygtnaaryt n 681
    
```

<210> 58

<211> 8

5

<212> PRT

<213> Secuencia artificial

<220>

<223> Marca FLAG

<400> 58

```

Asp Tyr Lys Asp Asp Asp Asp Lys
1           5
    
```

10

REIVINDICACIONES

1. Un procedimiento de preparación de una proteína trimérica que comprende cultivar una célula huésped transformada o transfectada con un vector de expresión que codifica una proteína de fusión que comprende:
 - (i) un dominio ZymoZipper (ZZ) en el que el dominio ZymoZipper (ZZ) comprende una secuencia de repetición de héptada viral;
 - (ii) una proteína heteróloga; y
 - (iii) una secuencia de adaptador, en el que el adaptador comprende una secuencia seleccionada del grupo constituido por la SEC ID N°: 33, SEC ID N°: 34, SEC ID N°: 35, SEC ID N°: 36, SEC ID N°: 37 y SEC ID N°: 38.
2. El procedimiento de la reivindicación 1, en el que la proteína heteróloga comprende el dominio extracelular de dicha proteína.
3. El procedimiento de la reivindicación 1, en el que dicha proteína de fusión es soluble.
4. El procedimiento de la reivindicación 1, en el que el dominio ZZ se deriva de la subunidad transmembrana (TM) de un virus clasificado dentro del género o de la familia seleccionados del grupo constituido por *Orthomyxovirus*, *Filovirus*, *Betaretrovirus*, *Gammaretrovirus*, *Deltaretrovirus*, *Lentivirus*, *Cornavirus* y *Paramyxovirus*, preferentemente en el que el virus es gripe A o un *Betaretrovirus* tal como el virus Mason-Pfizer del mono (VMMP).
5. El procedimiento de cualquier reivindicación anterior, en el que el dominio ZZ comprende una secuencia seleccionada del grupo constituido por la SEC ID N°: 27, SEC ID N°: 28, SEC ID N°: 29, SEC ID N°: 30, SEC ID N°: 31 y SEC ID N°: 32.
6. El procedimiento de cualquier reivindicación anterior, en el que el adaptador comprende la SEC ID N°: 35 y el dominio ZZ comprende SEC ID N°: 32.
7. El procedimiento de la reivindicación 1, en el que dicho dominio ZZ se deriva de la subunidad transmembrana (TM) de un retrovirus endógeno humano (RVEH), preferentemente en el que dicho RVEH es ERVWE1.
8. El procedimiento de cualquier reivindicación anterior que comprende además eliminar el dominio ZZ de la proteína trimérica usando digestión con proteasa, preferentemente en el que dicha proteasa es tripsina.
9. Un procedimiento de preparación de una proteína hetero-trimérica soluble:
 - cultivando una primera célula huésped transformada o transfectada con un primer vector de expresión que codifica una primera proteína de fusión que comprende un primer dominio ZZ y una primera proteína heteróloga,
 - cultivando una segunda célula huésped transformada o transfectada con un segundo vector de expresión que codifica una segunda proteína de fusión que comprende un segundo dominio ZZ y una segunda proteína heteróloga, y
 - cultivando una tercera célula huésped transformada o transfectada con un tercer vector de expresión que codifica una tercera proteína de fusión que comprende un tercer dominio ZZ y una tercera proteína heteróloga;
 - en el que dicha primera, segunda y tercera proteínas de fusión comprenden además una secuencia de adaptador, en el que el adaptador comprende una secuencia seleccionada del grupo constituido por la SEC ID N°: 33, SEC ID N°: 34, SEC ID N°: 35, SEC ID N°: 36, SEC ID N°: 37 y SEC ID N°: 38;
 - en el que el primer, segundo y tercer dominios ZZ comprenden una secuencia de repetición de héptada viral y en el que el primer, segundo y tercer dominios ZZ forman preferencialmente un hetero-trímero;
 - y combinando la primera, segunda y tercera proteínas de fusión en condiciones que promuevan un hetero-trímero.
10. El procedimiento de la reivindicación 9, en el que dicha primera y segunda células huésped son la misma célula.
11. El procedimiento de la reivindicación 9, en el que dicha primera y segunda proteínas heterólogas son las mismas.
12. El procedimiento de la reivindicación 11, en el que dicha primera y segunda proteínas heterólogas son el dominio extracelular de BLYS y dicha tercera proteína heteróloga es el dominio extracelular de APRIL.
13. El procedimiento de la reivindicación 9 que comprende además eliminar el dominio ZZ de la proteína hetero-trimérica usando digestión con proteasa, preferentemente en el que dicha proteasa es tripsina.

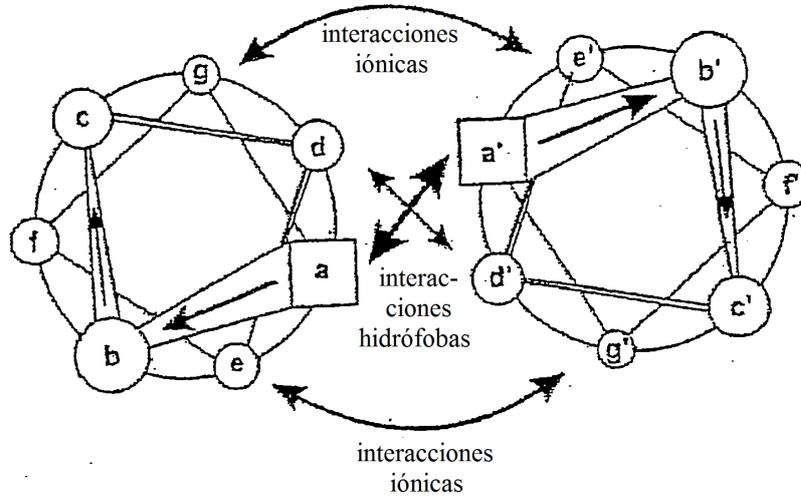


FIG. 1

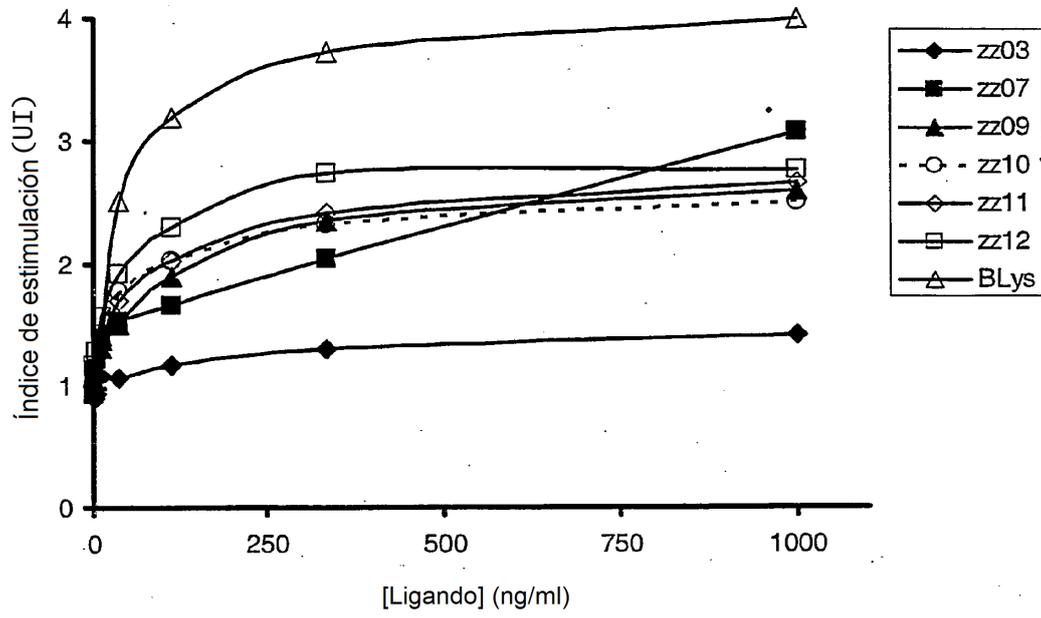


FIG. 2

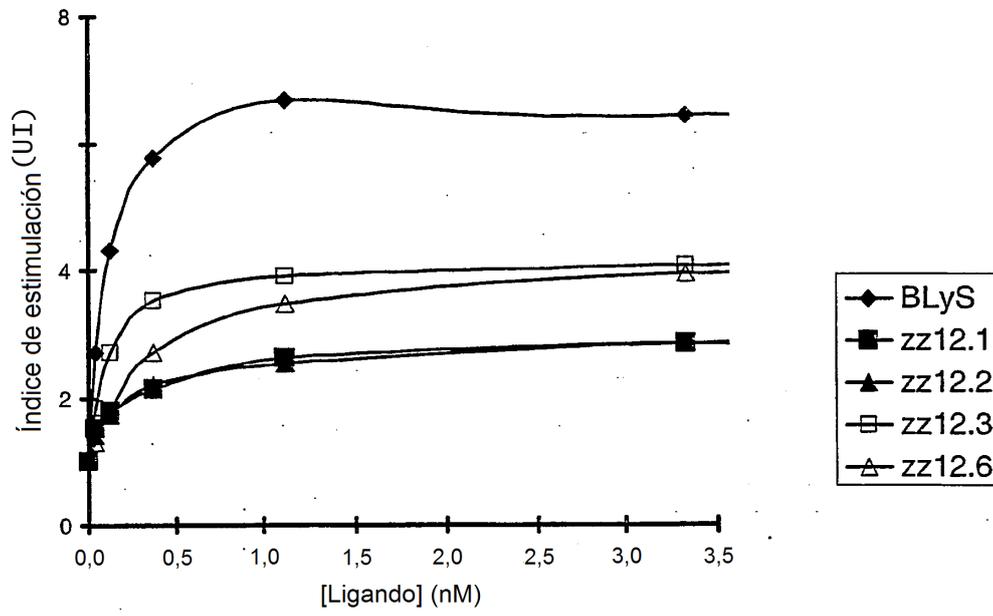


FIG. 3

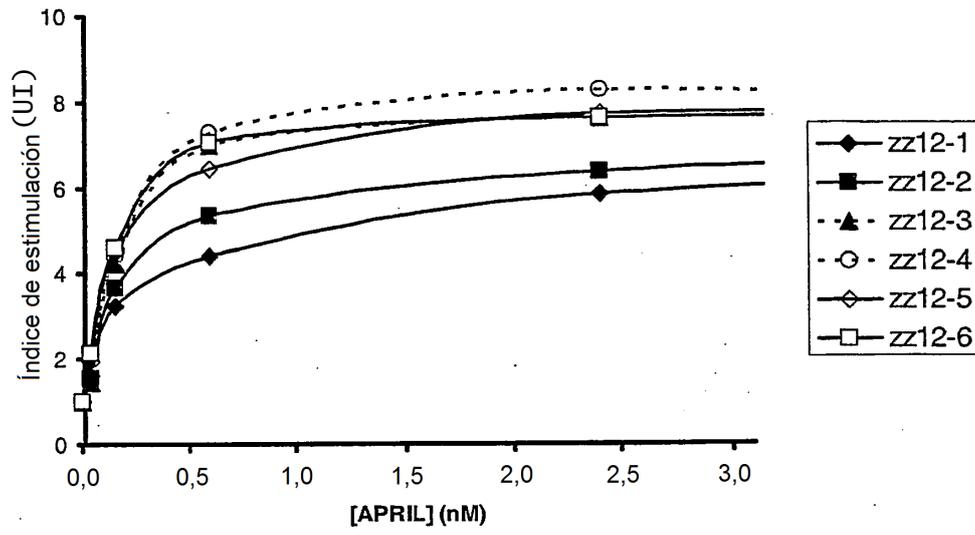


FIG. 4

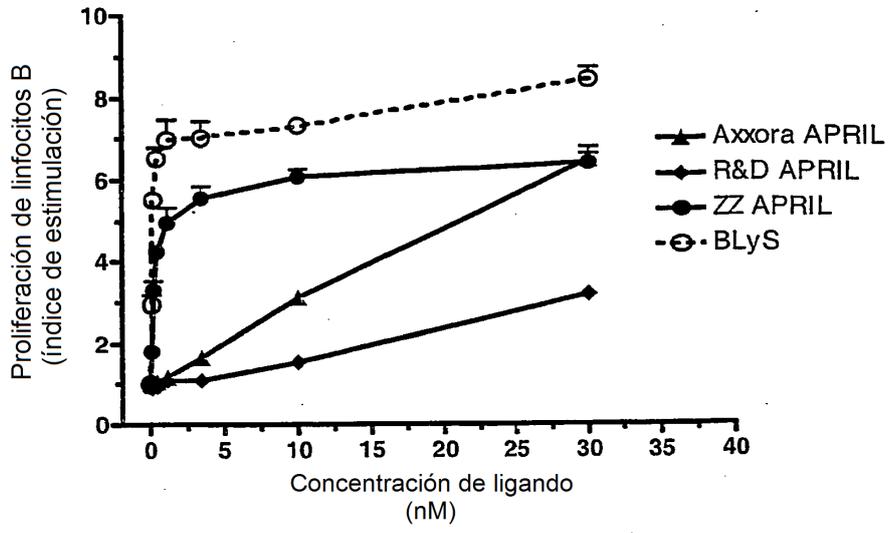


FIG. 5