



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 329 689**

51 Int. Cl.:

C12N 9/64 (2006.01)

C12Q 1/68 (2006.01)

G01N 33/574 (2006.01)

C07K 16/40 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05019847 .2**

96 Fecha de presentación : **02.12.1999**

97 Número de publicación de la solicitud: **1642968**

97 Fecha de publicación de la solicitud: **05.04.2006**

54 Título: **Composiciones y procedimientos para el diagnóstico de tumores.**

30 Prioridad: **08.03.1999 PCT/US99/05028**
01.09.1999 PCT/US99/20111
29.10.1999 US 162506 P
30.11.1999 PCT/US99/28313
01.12.1999 PCT/US99/28634

45 Fecha de publicación de la mención BOPI:
30.11.2009

45 Fecha de la publicación del folleto de la patente:
30.11.2009

73 Titular/es: **Genentech, Inc.**
1 DNA Way
South San Francisco, California 94080-4990, US

72 Inventor/es: **Botstein, David;**
Goddard, Audrey;
Gurney, Austin L.;
Roy, Margaret Ann;
Watanabe, Colin K. y
Wood, William I.

74 Agente: **Ponti Sales, Adelaida**

ES 2 329 689 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

Composiciones y procedimientos para el diagnóstico de tumores.

5 **Campo de la invención**

La presente invención se refiere a composiciones y procedimientos para el diagnóstico y tratamiento de tumores.

Antecedentes de la invención

10

Los tumores malignos (cánceres) son la segunda causa más importante de muerte en los Estados Unidos, tras las enfermedades cardíacas (Boring *et al.*, CA Cancer J. Clin., 43: 7 [1993]).

15

El cáncer se caracteriza por un aumento del número de células anormales o neoplásicas derivadas de un tejido normal que proliferan para formar una masa tumoral, la invasión de tejidos adyacentes por estas células tumorales neoplásicas y la generación de células malignas que eventualmente se extienden a través de la sangre o el sistema linfático a ganglios linfáticos regionales y a puntos distantes (metástasis). En un estado canceroso, una célula prolifera en condiciones en las que las células normales no crecerían. El propio cáncer se manifiesta de muy diversas formas, caracterizadas por diferentes grados de invasividad y agresividad.

20

25

La alteración de la expresión génica está íntimamente relacionada con el crecimiento celular incontrolado y la desdiferenciación que son una característica común de todos los cánceres. Se ha descubierto que los genomas de ciertos tumores bien estudiados muestran una expresión reducida de genes recesivos, habitualmente denominados como genes de supresión tumoral, que normalmente funcionarían para impedir el crecimiento de células malignas, y/o la sobre-expresión de ciertos genes dominantes, tales como oncogenes, que actúan para promover el crecimiento canceroso. Cada uno de estos cambios genéticos parece ser responsable de la importación de algunos de los rasgos que, en conjunto, representan el fenotipo neoplásico completo (Hunter, Cell, 64: 1129 [1991] y Bishop, Cell, 64: 235-248 [1991]).

30

Un mecanismo bien conocido de sobre-expresión génica (por ejemplo, de oncogenes) en células cancerosas es la amplificación génica. Éste es un proceso donde en el cromosoma de la célula ancestral se producen múltiples copias de un gen particular. El proceso implica la replicación no programada de la región del cromosoma que comprende el gen, seguida de recombinación de los segmentos replicados de vuelta en el cromosoma (Alitalo *et al.*, Adv. Cancer Res., 47: 235-281 [1986]). Se cree que la sobre-expresión del gen es análoga a la amplificación génica, es decir, es proporcional al número de copias realizadas.

35

40

Se han identificado los proto-oncogenes que codifican factores de crecimiento y receptores de factores de crecimiento que juegan papeles importantes en la patogénesis de diversos cánceres humanos, incluyendo cáncer de mama. Por ejemplo, se ha descubierto que el gen Erb2 humano (erbB2, también conocido como her2, o c-erbB-2), que codifica un receptor de glucoproteína transmembrana de 185 kd (p185^{HER2}; HER2) relacionado con el receptor del factor de crecimiento epidérmico EGFR, se sobre-expresa en de aproximadamente el 25% al 30% de los cánceres de mama humanos (Slamon *et al.*, Science, 235: 177-182 [1987]; Slamon *et al.*, Science, 244: 707-712 [1989]).

45

50

55

Se ha descrito que la amplificación génica de un proto-oncogén es un suceso habitualmente implicado en las formas más malignas de cáncer, y podría actuar como factor pronóstico del resultado clínico (Schwab *et al.*, Genes Chromosomes Cancer, 1: 181-193 [1990]; Alitalo *et al.*, *supra*). Por lo tanto, la sobre-expresión de *erbB2* se contempla habitualmente como un factor pronóstico de un mal pronóstico, especialmente en pacientes con enfermedad primaria que implica ganglios linfáticos axilares (Slamon *et al.*, [1987] y [1989], *supra*; Ravdin y Chamness, Gene, 159: 19-27 [1995]; y Hynes y Stern, Biochim. Biophys. Acta, 1198: 165-184 [1994]), y se ha vinculado con la sensibilidad y/o resistencia a terapia hormonal y regímenes quimioterapéuticos, incluyendo CMF (ciclofosfamida, metotrexato y fluorouracilo) y antraciclinas (Baselga *et al.*, Oncology, 11 (3 Supl 1): 43-48 [1997]). Sin embargo, a pesar de la asociación de la sobre-expresión de *erbB2* con un mal pronóstico, las probabilidades de que pacientes positivos para HER2 respondan clínicamente al tratamiento con taxanos eran más de tres veces las de pacientes negativos para HER2 (ibídem). Un anticuerpo monoclonal anti-ErbB2 (anti-HER2) recombinante humanizado (una versión humanizada del anticuerpo anti-ErbB2 murino 4D5, denominado como rhuMAb HER2 o HerceptinTM) era clínicamente activo en pacientes con cánceres de mama metastáticos que sobre-expresan ErbB2 que habían recibido una exhaustiva terapia anti-cáncer anterior. (Baselga *et al.*, J. Clin. Oncol., 14: 737-744 [1996]).

60

En vista de lo anterior, existe un interés obvio en la identificación de nuevos procedimientos y composiciones que sean útiles para diagnosticar y tratar tumores, que se asocian con la amplificación génica.

Descripción resumida de la invención

65

A. Realizaciones

La presente invención se refiere a composiciones y procedimientos para el diagnóstico y tratamiento del crecimiento y proliferación de células neoplásicas en mamíferos, incluyendo seres humanos. La presente invención se basa

en la identificación de genes que se amplifican en el genoma de células tumorales. Se espera que dicha amplificación génica esté asociada con la sobre-expresión del producto génico y contribuya a la tumorigénesis. Por consiguiente, se cree que las proteínas codificadas por los genes amplificados son dianas útiles para el diagnóstico y/o tratamiento (incluyendo prevención) de algunos cánceres, y pueden actuar como factores pronóstico del pronóstico del tratamiento del tumor.

En una realización, la presente invención se refiere a un anticuerpo aislado que se une a un polipéptido designado en este documento como un polipéptido PRO para su utilización en diagnóstico. En un aspecto, el anticuerpo aislado se une específicamente a un polipéptido PRO. En otro aspecto, el anticuerpo induce la muerte de una célula que expresa un polipéptido PRO. A menudo, la célula que expresa el polipéptido PRO es una célula tumoral que sobre-expresa el polipéptido en comparación con una célula normal del mismo tipo tisular. En otro aspecto más, el anticuerpo es un anticuerpo monoclonal, que preferentemente tiene restos de la región determinante de la complementariedad (CDR) no humana y restos de la región marco (FR) humana. El anticuerpo puede estar marcado y estar inmovilizado en un soporte sólido. En otro aspecto más, el anticuerpo es un fragmento de anticuerpo, un anticuerpo de cadena sencilla o un anticuerpo humanizado que se une, preferentemente de forma específica, a un polipéptido PRO.

En otra realización, la invención se refiere a una composición de materia que comprende un anticuerpo que se une, preferentemente de forma específica, a un polipéptido PRO mezclado con un portador farmacéuticamente aceptable para su utilización en diagnóstico. En un aspecto, la composición de materia comprende una cantidad terapéuticamente eficaz del anticuerpo. En otro aspecto, la composición comprende un ingrediente activo adicional, que puede ser, por ejemplo, un anticuerpo adicional o un agente citotóxico o quimioterapéutico. Preferentemente, la composición es estéril.

En una realización adicional, la invención se refiere a moléculas de ácido nucleico aisladas que codifican anticuerpos anti-PRO, y vectores y células huésped recombinantes que comprenden dichas moléculas de ácido nucleico.

En otra realización adicional, la invención se refiere a un procedimiento para producir un anticuerpo anti-PRO, donde el procedimiento comprende cultivar una célula huésped transformada con una molécula de ácido nucleico que codifica el anticuerpo en condiciones suficientes para permitir la expresión del anticuerpo, y recuperar el anticuerpo del cultivo celular.

Aquí se describen antagonistas de un polipéptido PRO que inhiben una o más de las funciones o actividades biológicas y/o inmunológicas de un polipéptido PRO.

También se describe una molécula de ácido nucleico que hibrida a una molécula de ácido nucleico que codifica a un polipéptido PRO o al complemento del mismo. La molécula de ácido nucleico aislada es preferiblemente ADN, y la hibridación se produce preferentemente en condiciones de hibridación y lavado estrictas. Dichas moléculas de ácido nucleico pueden actuar como moléculas antisentido de los genes amplificados identificados en este documento, que, a su vez, pueden ser útiles en la modulación de la transcripción y/o traducción de los genes amplificados respectivos, o como cebadores antisentido en reacciones de amplificación. Además, dichas secuencias pueden usarse como parte de una ribozima y/o una secuencia de triple hélice que, a su vez, puede usarse en la regulación de los genes amplificados.

En otra realización, la invención proporciona un procedimiento para determinar la presencia de un polipéptido PRO en una muestra de la que se sospecha que contiene un polipéptido PRO, donde el procedimiento comprende exponer la muestra a un anticuerpo anti-PRO y determinar la unión del anticuerpo a un polipéptido PRO en la muestra. En otra realización, la invención proporciona un procedimiento para determinar la presencia de un polipéptido PRO en una célula, donde el procedimiento comprende exponer la célula a un anticuerpo anti-PRO y determinar la unión del anticuerpo a la célula.

En otra realización más, la presente invención se refiere a un procedimiento de diagnóstico de un tumor en un mamífero, que comprende detectar el nivel de expresión de un gen que codifica un polipéptido PRO (a) en una muestra de ensayo de células de tejido obtenidas del mamífero, y (b) en una muestra de control de células de tejido normal conocidas del mismo tipo celular, donde un nivel de expresión más alto en la muestra de ensayo en comparación con la muestra de control, es indicativo de la presencia del tumor en el mamífero del que se obtuvieron las células de tejido de ensayo.

En otra realización, la presente invención se refiere a un procedimiento de diagnóstico de un tumor en un mamífero, que comprende (a) poner en contacto un anticuerpo anti-PRO con una muestra de ensayo de células de tejido obtenidas del mamífero, y (b) detectar la formación de un complejo entre el anticuerpo anti-PRO y un polipéptido PRO en la muestra de ensayo, donde la formación de un complejo es indicativa de la presencia de un tumor en dicho mamífero. La detección puede ser cualitativa o cuantitativa, y puede realizarse en comparación con la supervisión de la formación de complejo en una muestra de control de células de tejido normal conocidas del mismo tipo celular. Una mayor cantidad de complejos formados en la muestra de ensayo indica la presencia del tumor en el mamífero del que se obtuvieron las células de tejido de ensayo. El anticuerpo preferentemente porta una marca detectable. La formación del complejo puede supervisarse, por ejemplo, mediante microscopía óptica, citometría de flujo, fluorimetría u otras técnicas conocidas en el sector.

La muestra de ensayo se obtiene habitualmente de un individuo del que se sospecha que tiene crecimiento o proliferación de células neoplásicas (por ejemplo, células cancerosas).

En otra realización, la presente invención se refiere a un kit de diagnóstico del cáncer que comprende un anticuerpo anti-PRO y un portador (por ejemplo, un tampón) en un envase adecuado. El kit contiene preferentemente instrucciones para usar el anticuerpo para detectar la presencia de un polipéptido PRO en una muestra de la que se sospecha que lo contiene.

B. Realizaciones adicionales

Aquí se describe una molécula de ácido nucleico aislada que comprende una secuencia de nucleótido que codifica un polipéptido PRO.

En un aspecto, la molécula de ácido nucleico aislada comprende una secuencia de nucleótido que tiene preferiblemente al menos el 80% de identidad en la secuencia, preferiblemente al menos el 81% de identidad en la secuencia, más preferiblemente al menos el 82% de identidad en la secuencia, incluso más preferiblemente al menos el 83% de identidad en la secuencia, incluso más preferiblemente al menos el 84% de identidad en la secuencia, incluso más preferiblemente al menos el 85% de identidad en la secuencia, incluso más preferiblemente al menos el 86% de identidad en la secuencia, incluso más preferiblemente al menos el 87% de identidad en la secuencia, incluso más preferiblemente al menos el 88% de identidad en la secuencia, incluso más preferiblemente al menos el 89% de identidad en la secuencia, incluso más preferiblemente al menos el 90% de identidad en la secuencia, incluso más preferiblemente al menos el 91% de identidad en la secuencia, incluso más preferiblemente al menos el 92% de identidad en la secuencia, incluso más preferiblemente al menos el 93% de identidad en la secuencia, incluso más preferiblemente al menos el 94% de identidad en la secuencia, incluso más preferiblemente al menos el 95% de identidad en la secuencia, incluso más preferiblemente al menos el 96% de identidad en la secuencia, incluso más preferiblemente al menos el 97% de identidad en la secuencia, incluso más preferiblemente al menos el 98% de identidad en la secuencia, incluso más preferiblemente al menos el 99% de identidad en la secuencia de (a) una molécula de ADN que codifica a un polipéptido PRO que tiene una secuencia de aminoácido de longitud completa como se describe en este documento, la secuencia codificante de un polipéptido PRO que carece del péptido señal como se describe en este documento, la secuencia codificante de un dominio extracelular de un polipéptido PRO transmembrana, con o sin el péptido señal, como se describe en este documento o la secuencia codificante de cualquier otro fragmento definido específicamente de la secuencia de aminoácidos de longitud completa como se describe en este documento, o (b) el complemento de la molécula de ADN de (a).

En otros aspectos, la molécula de ácido nucleico aislada comprende una secuencia de nucleótidos que tiene al menos aproximadamente el 80% de identidad en la secuencia, preferiblemente al menos el 81% de identidad en la secuencia, más preferiblemente al menos el 82% de identidad en la secuencia, incluso más preferiblemente al menos el 83% de identidad en la secuencia, incluso más preferiblemente al menos el 84% de identidad en la secuencia, incluso más preferiblemente al menos el 85% de identidad en la secuencia, incluso más preferiblemente al menos el 86% de identidad en la secuencia, incluso más preferiblemente al menos el 87% de identidad en la secuencia, incluso más preferiblemente al menos el 88% de identidad en la secuencia, incluso más preferiblemente al menos el 89% de identidad en la secuencia, incluso más preferiblemente al menos el 90% de identidad en la secuencia, incluso más preferiblemente al menos el 91% de identidad en la secuencia, incluso más preferiblemente al menos el 92% de identidad en la secuencia, incluso más preferiblemente al menos el 93% de identidad en la secuencia, incluso más preferiblemente al menos el 94% de identidad en la secuencia, incluso más preferiblemente al menos el 95% de identidad en la secuencia, incluso más preferiblemente al menos el 96% de identidad en la secuencia, incluso más preferiblemente al menos el 97% de identidad en la secuencia, incluso más preferiblemente al menos el 98% de identidad en la secuencia, incluso más preferiblemente al menos el 99% de identidad en la secuencia de (a) una molécula de ADN que comprende la secuencia codificante de un ADNc de polipéptido PRO de longitud completa como se describe en este documento, la secuencia codificante de un polipéptido PRO que carece del péptido señal como se describe en este documento, la secuencia codificante de un dominio extracelular de un polipéptido PRO transmembrana, con o sin el péptido señal, como se describe en este documento o la secuencia codificante de cualquier otro fragmento definido específicamente de la secuencia de aminoácidos de longitud completa como se describe en este documento, o (b) el complemento de la molécula de ADN de (a).

También se describe una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que tiene al menos aproximadamente el 80% de identidad en la secuencia, preferiblemente al menos el 81% de identidad en la secuencia, más preferiblemente al menos el 82% de identidad en la secuencia, incluso más preferiblemente al menos el 83% de identidad en la secuencia, incluso más preferiblemente al menos el 84% de identidad en la secuencia, incluso más preferiblemente al menos el 85% de identidad en la secuencia, incluso más preferiblemente al menos el 86% de identidad en la secuencia, incluso más preferiblemente al menos el 87% de identidad en la secuencia, incluso más preferiblemente al menos el 88% de identidad en la secuencia, incluso más preferiblemente al menos el 89% de identidad en la secuencia, incluso más preferiblemente al menos el 90% de identidad en la secuencia, incluso más preferiblemente al menos el 91% de identidad en la secuencia, incluso más preferiblemente al menos el 92% de identidad en la secuencia, incluso más preferiblemente al menos el 93% de identidad en la secuencia, incluso más preferiblemente al menos el 94% de identidad en la secuencia, incluso más preferiblemente al menos el 95% de identidad en la secuencia, incluso

más preferiblemente al menos el 96% de identidad en la secuencia, incluso más preferiblemente al menos el 97% de identidad en la secuencia, incluso más preferiblemente al menos el 98% de identidad en la secuencia, incluso más preferiblemente al menos el 99% de identidad en la secuencia de (a) una molécula de AND que codifica el mismo polipéptido maduro codificado por cualquiera de las proteínas humanas ADNc depositadas con el ATCC como aquí se describe, o (b) el complemento de la molécula de AND de (a).

También se describe una secuencia de nucleótidos que codifica un polipéptido que tiene eliminado el dominio transmembrana o que tiene inactivado el dominio transmembrana, o es complementario a dicha secuencia de nucleótidos codificante, donde el dominio o dominios transmembrana de dicho polipéptido se describen en este documento. Por lo tanto, se contemplan dominios extracelulares solubles de los polipéptidos PRO descritos en este documento. También se describe un polipéptido PRO aislado codificado por cualquiera de las secuencias de ácido nucleico aisladas identificadas anteriormente en este documento.

También se describen fragmentos de una secuencia codificante de polipéptido PRO, o el complemento de la misma, que pueden ser útiles como, por ejemplo, sondas de hibridación, para fragmentos codificantes de un polipéptido PRO que pueden codificar opcionalmente un polipéptido que comprende un punto de unión para un anticuerpo anti-PRO o como sondas de oligonucleótidos antisentido. Dichos fragmentos de ácido nucleico son habitualmente de al menos aproximadamente 20 nucleótidos de longitud, preferentemente de al menos aproximadamente 30 nucleótidos de longitud, más preferentemente de al menos aproximadamente 40 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 50 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 60 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 70 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 80 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 90 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 100 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 110 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 120 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 130 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 140 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 150 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 160 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 170 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 180 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 190 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 200 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 250 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 300 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 350 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 400 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 450 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 500 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 600 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 700 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 800 nucleótidos de longitud, aún más preferentemente de al menos aproximadamente 900 nucleótidos de longitud y aún más preferentemente de al menos aproximadamente 1000 nucleótidos de longitud, donde en este contexto el término "aproximadamente" significa la longitud de secuencia mencionada más o menos el 10% de esa longitud mencionada. Se observa que los nuevos fragmentos de una secuencia de nucleótidos que codifica un polipéptido PRO pueden determinarse de manera rutinaria alineando la secuencia de nucleótidos que codifica un polipéptido PRO con otras secuencias de nucleótidos conocidas usando cualquiera de varios programas de alineamiento de secuencias bien conocidos y determinando qué fragmento o fragmentos de secuencia de nucleótidos que codifica un polipéptido PRO son nuevos. Todas estas secuencias de nucleótidos que codifican un polipéptido PRO se contemplan en este documento. También se describen los fragmentos del polipéptido PRO codificados por estos fragmentos de molécula nucleotídica, preferentemente aquellos fragmentos de polipéptido PRO que comprenden un punto de unión para un PRO. También se describe un polipéptido PRO aislado codificado por cualquiera de las secuencias de ácido nucleico aisladas identificadas anteriormente en este documento.

También se describe un polipéptido PRO PRO2262, que comprende una secuencia de aminoácidos que tiene al menos aproximadamente el 80% de identidad en la secuencia, preferiblemente al menos el 81% de identidad en la secuencia, más preferiblemente al menos el 82% de identidad en la secuencia, incluso más preferiblemente al menos el 83% de identidad en la secuencia, incluso más preferiblemente al menos el 84% de identidad en la secuencia, incluso más preferiblemente al menos el 85% de identidad en la secuencia, incluso más preferiblemente al menos el 86% de identidad en la secuencia, incluso más preferiblemente al menos el 87% de identidad en la secuencia, incluso más preferiblemente al menos el 88% de identidad en la secuencia, incluso más preferiblemente al menos el 89% de identidad en la secuencia, incluso más preferiblemente al menos el 90% de identidad en la secuencia, incluso más preferiblemente al menos el 91% de identidad en la secuencia, incluso más preferiblemente al menos el 92% de identidad en la secuencia, incluso más preferiblemente al menos el 93% de identidad en la secuencia, incluso más preferiblemente al menos el 94% de identidad en la secuencia, incluso más preferiblemente al menos el 95% de identidad en la secuencia, incluso más preferiblemente al menos el 96% de identidad en la secuencia, incluso más preferiblemente al menos el 97% de identidad en la secuencia, incluso más preferiblemente al menos el 98% de identidad en la secuencia, incluso más preferiblemente al menos el 99% de identidad de secuencia a un polipéptido PRO que tiene una secuencia de aminoácido de longitud completa como se describe en este documento, la secuencia de aminoácido que carece del péptido señal como se describe en este documento, un dominio extracelular de un polipéptido PRO transmembrana,

con o sin el péptido señal, como se describe en este documento o cualquier otro fragmento definido específicamente de la secuencia de aminoácidos de longitud completa como se describe en este documento.

También se describe un PRO aislada que tiene al menos aproximadamente el 80% de identidad en la secuencia, preferiblemente al menos el 81% de identidad en la secuencia, más preferiblemente al menos el 82% de identidad en la secuencia, incluso más preferiblemente al menos el 83% de identidad en la secuencia, incluso más preferiblemente al menos el 84% de identidad en la secuencia, incluso más preferiblemente al menos el 85% de identidad en la secuencia, incluso más preferiblemente al menos el 86% de identidad en la secuencia, incluso más preferiblemente al menos el 87% de identidad en la secuencia, incluso más preferiblemente al menos el 88% de identidad en la secuencia, incluso más preferiblemente al menos el 89% de identidad en la secuencia, incluso más preferiblemente al menos el 90% de identidad en la secuencia, incluso más preferiblemente al menos el 91% de identidad en la secuencia, incluso más preferiblemente al menos el 92% de identidad en la secuencia, incluso más preferiblemente al menos el 93% de identidad en la secuencia, incluso más preferiblemente al menos el 94% de identidad en la secuencia, incluso más preferiblemente al menos el 95% de identidad en la secuencia, incluso más preferiblemente al menos el 96% de identidad en la secuencia, incluso más preferiblemente al menos el 97% de identidad en la secuencia, incluso más preferiblemente al menos el 98% de identidad en la secuencia, incluso más preferiblemente al menos el 99% de identidad en la secuencia de una secuencia de aminoácido codificada por cualquiera de los ADNc de proteína humana depositadas con el ATCC como aquí se describe.

También se describe un PRO aislado que tiene al menos aproximadamente una secuencia de aminoácidos que da una valoración de al menos 80% de positivos, más preferiblemente al menos un 82% de positivos, más preferiblemente al menos un 83% de positivos, más preferiblemente al menos un 84% de positivos, más preferiblemente al menos un 85% de positivos, más preferiblemente al menos un 86% de positivos, más preferiblemente al menos un 87% de positivos, más preferiblemente al menos un 88% de positivos, más preferiblemente al menos un 89% de positivos, más preferiblemente al menos un 90% de positivos, más preferiblemente al menos un 91% de positivos, más preferiblemente al menos un 92% de positivos, más preferiblemente al menos un 93% de positivos, más preferiblemente al menos un 94% de positivos, más preferiblemente al menos un 95% de positivos, más preferiblemente al menos un 96% de positivos, más preferiblemente al menos un 97% de positivos, más preferiblemente al menos un 98% de positivos, más preferiblemente al menos un 99% de positivos, al compararlo con la secuencia de aminoácidos de un polipéptido PRO que tiene la secuencia de aminoácido de longitud completa como aquí se describe, una secuencia de aminoácido que carece del péptido señal como se describe en este documento, un dominio extracelular de un polipéptido PRO transmembrana, con o sin el péptido señal, como se describe en este documento o cualquier otro fragmento definido específicamente de la secuencia de aminoácidos de longitud completa como se describe en este documento.

También se describe un polipéptido sin la secuencia de señal N-terminal y/o la metionina iniciadora y está codificada por una secuencia de nucleótidos que codifica la secuencia de aminoácido descrita anteriormente. Los procesos para producirla también se describen, donde aquellos procesos comprender cultivar una célula huésped que comprende un vector que comprende la molécula de ácido nucleico codificante apropiada bajo condiciones adecuadas para la expresión del polipéptido PRO y la recuperación del polipéptido PRO de los cultivos de células.

También se describe un polipéptido PRO que tiene eliminado el dominio transmembrana o que tiene inactivado el dominio transmembrana. Los procesos para producirlos son los mismos que los aquí descritos, en los que aquellos procesos comprenden el cultivo de una célula huésped que comprende un vector que comprende la molécula de ácido nucleico apropiada bajo condiciones adecuadas para la expresión del polipéptido PRO como aquí se define. El antagonista puede ser cualquier anticuerpo anti-PRO o una molécula pequeña.

También se describen los antagonistas de un polipéptido PRO como aquí se define. El antagonista puede ser un anticuerpo PRO o una molécula pequeña.

También se describe un procedimiento para identificar antagonistas en un polipéptido PRO que comprende poner en contacto el polipéptido PRO con una molécula candidata y supervisar una actividad biológica mediada por dicho polipéptido PRO. Preferiblemente, el polipéptido PRO es un polipéptido PRO nativo.

En todavía otra realización, la invención se refiere a una composición de material que comprende un anticuerpo anti-PRO para su utilización en diagnóstico, junto con un portador. Opcionalmente, el portador es un portador farmacéuticamente aceptable.

También se describe el uso de un polipéptido PRO, o un antagonista del mismo como se ha descrito anteriormente en este documento, o un anticuerpo anti-PRO, para la preparación de un medicamento útil en el tratamiento de una afección que es sensible al polipéptido PRO, un antagonista del mismo o un anticuerpo anti-PRO.

También se describen aquí vectores que comprenden ADN que codifica cualquiera de los polipéptidos descritos en este documento. También se proporcionan células huésped que comprenden cualquiera de dichos vectores. A modo de ejemplo, las células huésped pueden ser células CHO, células de *E. coli*, levadura, o de insecto infectadas por Baculovirus. Además se proporciona un proceso para producir cualquiera de los polipéptidos descritos en este documento y comprende cultivar células huésped en condiciones adecuadas para la expresión del polipéptido deseado y recuperar el polipéptido deseado del cultivo celular.

También se describen moléculas químéricas que comprenden cualquiera de los polipéptidos descritos en este documento fusionado a una secuencia de aminoácidos o de polipéptidos heteróloga. Los ejemplos de dichas moléculas químéricas comprenden cualquiera de los polipéptidos descritos en este documento fusionados a una secuencia de marca epitópica o una región Fc de una inmunoglobulina.

En otra realización, la invención proporciona un anticuerpo que se une específicamente a cualquiera de los polipéptidos descritos anterior o posteriormente. Opcionalmente, el anticuerpo es un anticuerpo monoclonal, anticuerpo humanizado, fragmento de anticuerpo o anticuerpo de cadena sencilla.

En otras realizaciones más, la invención proporciona sondas de oligonucleótidos útiles para aislar secuencias de nucleótidos genómicas y de ADNc o como sondas antisentido, donde estas sondas pueden obtenerse de cualquiera de las secuencias de nucleótidos descritas anterior o posteriormente.

Breve descripción de las figuras

La Figura 1 muestra la secuencia de nucleótidos (SEC ID N°: 26) de un ADNc que contiene una secuencia de nucleótidos que codifica PRO3567 de secuencia nativa, donde la secuencia de nucleótidos (SEC ID N°: 26) es un clon designado en este documento como DNA56049-2543. También se presentan en negrita y subrayadas las posiciones de los respectivos codones de inicio y de terminación.

La Figura 2 muestra la secuencia de aminoácidos (SEC ID N°: 27) de un polipéptido PRO3567 de secuencia nativa de la secuencia codificante de la SEC ID N°: 26 mostrada en la Figura 1.

Descripción detallada de la invención

I. Definiciones

Las frases “amplificación génica” y “duplicación génica” se usan de forma intercambiable y se refieren a un proceso mediante el cual múltiples copias de un gen o fragmento génico se forman en una célula o línea celular particular. La región duplicada (un tramo de ADN amplificado) a menudo se denomina como “amplicón”. Habitualmente, la cantidad de ARN mensajero (ARNm) producido, es decir, el nivel de expresión génica, también aumenta en proporción al número de copias realizadas del gen particular expresado.

“Tumor”, como se usa en este documento, se refiere a todo crecimiento y proliferación de células neoplásicas, ya sea maligno o benigno, y a todas las células y tejidos pre-cancerosos y cancerosos.

Los términos “cáncer” y “canceroso” se refieren a o describen la afección fisiológica en mamíferos que habitualmente se caracteriza por crecimiento celular no regulado. Los ejemplos de cáncer incluyen aunque sin limitación, carcinoma, linfoma, blastoma, sarcoma y leucemia. Ejemplos más particulares de dichos cánceres incluyen cáncer de mama, cáncer de próstata, cáncer de colon, cáncer de células escamosas, cáncer de pulmón microcítico, cáncer de pulmón macrocítico, cáncer gastrointestinal, cáncer pancreático, glioblastoma, cáncer de cuello de útero, cáncer de ovario, cáncer de hígado, cáncer de vejiga, hepatoma, cáncer colorrectal, carcinoma de endometrio, carcinoma de glándulas salivales, cáncer de riñón, cáncer de hígado, cáncer de vulva, cáncer de tiroides, carcinoma hepático y diversos tipos de cáncer de cabeza y cuello.

“Tratamiento” es una intervención realizada con la intención de impedir el desarrollo o alterar la patología de un trastorno. Por consiguiente, “tratamiento” se refiere a tratamiento terapéutico y medidas profilácticas o preventivas. Aquellos que necesitan tratamiento incluyen aquellos que ya padecen el trastorno así como aquellos en los que se quiere prevenir el trastorno. En el tratamiento de tumores (por ejemplo, cáncer), un agente terapéutico puede reducir directamente la patología de células tumorales, o hacer a las células tumorales más susceptibles al tratamiento mediante otros agentes terapéuticos, por ejemplo, radiación y/o quimioterapia.

La “patología” del cáncer incluye todos los fenómenos que ponen en peligro el bienestar del paciente. Esto incluye, sin limitación, crecimiento celular anormal o incontrolable, metástasis, interferencia en el funcionamiento normal de células adyacentes, liberación de citoquinas u otros productos de secreción a niveles anormales, supresión o agravamiento de la respuesta inflamatoria o inmunológica, etc.

“Mamífero” para fines de tratamiento se refiere a cualquier animal clasificado como mamífero, incluyendo seres humanos, animales domésticos y de granja, y animales de zoológico, para la práctica deportiva o de compañía, tales como perros, caballos, gatos, vacas, cerdos, ovejas, etc. Preferentemente, el mamífero es un ser humano.

“Portadores” como se usa en este documento incluye portadores, excipientes o estabilizantes farmacéuticamente aceptables que no son tóxicos para la célula o mamífero que se expone a ellos a las dosificaciones y concentraciones empleadas. A menudo el portador fisiológicamente aceptable es una solución acuosa con el pH regulado. Los ejemplos de portadores fisiológicamente aceptables incluyen tampones tales como fosfato, citrato y otros ácidos orgánicos; antioxidantes incluyendo ácido ascórbico; polipéptidos de bajo peso molecular (menos de aproximadamente 10 restos); proteínas, tales como albúmina de suero, gelatina o inmunoglobulinas; polímeros hidrofílicos tales como polivinilpirrolidona; aminoácidos tales como glicina, glutamina, asparagina, arginina o lisina; monosacáridos, disacáridos y

otros carbohidratos incluyendo glucosa, manosa o dextrinas; agentes quelantes tales como EDTA; alcoholes de azúcar tales como manitol o sorbitol; contraiones formadores de sales tales como sodio; y/o tensioactivos no iónicos tales como TWEENTM, polietilenglicol (PEG) y PLURONICSTM.

- 5 La administración “en combinación con” uno o más agentes terapéuticos adicionales incluye administración simultánea (concurrente) y consecutiva en cualquier orden.

La expresión “agente citotóxico” como se usa en este documento se refiere a una sustancia que inhibe o impide la función de células y/o causa la destrucción de células. La expresión pretende incluir isótopos radiactivos (por ejemplo, ¹³¹I, ¹²⁵I, ⁹⁰Y y ¹⁸⁶Re), agentes quimioterapéuticos, y toxinas tales como toxinas enzimáticamente activas de origen bacteriano, fúngico, vegetal o animal o fragmentos de las mismas.

Un “agente quimioterapéutico” es un compuesto químico útil en el tratamiento del cáncer. Los ejemplos de agentes quimioterapéuticos incluyen adriamicina, doxorrubicina, epirubicina, 5-fluorouracilo, citosina arabinósido (“Ara-C”), ciclofosfamida, tiotepa, busulfán, citoxina, taxoides, por ejemplo, paclitaxel (Taxol, Bristol-Myers Squibb Oncology, Princeton, NJ), y doxetaxel (Taxotere, Rhône-Poulenc Rorer, Antony, Rnace), toxotere, metotrexato, cisplatino, melfalán, vinblastina, bleomicina, etopósido, ifosfamida, mitomicina C, mitoxantrona, vincristina, vinorelbina, carboplatino, tenipósido, daunomicina, carminomicina, aminopterina, dactinomicina, mitomicinas, esperamicinas (véase la Patente de Estados Unidos N° 4.675.187), 5-FU, 6-tioguanina, 6-mercaptopurina, actinomicina D, VP-16, clorambucilo, melfalán, y otras mostazas de nitrógeno relacionadas. En esta definición también se incluyen agentes hormonales que actúan para regular o inhibir la acción hormonal sobre tumores tales como tamoxifeno y onapristona.

Un “agente inhibidor del crecimiento” cuando se usa en este documento se refiere a un compuesto o composición que inhibe el crecimiento de una célula, especialmente célula cancerosa que sobre-expresa cualquiera de los genes identificados en este documento, *in vitro* o *in vivo*. Por lo tanto, el agente inhibidor del crecimiento es uno que reduce significativamente el porcentaje de células que sobre-expresan dichos genes en fase S. Los ejemplos de agentes inhibidores del crecimiento incluyen agentes que bloquean la progresión del ciclo celular (en un punto diferente de la fase S), tales como agentes que inducen la detención en fase G I y detención en fase M. Los bloqueantes de fase M convencionales incluyen los alcaloides de Vinca (vincristina y vinblastina), taxol, e inhibidores de topo II tales como doxorrubicina, epirubicina, daunorrubicina, etopósido y bleomicina. Aquellos agentes que detienen G1 también extienden su acción a la detención en fase S, por ejemplo, agentes alquilantes de ADN tales como tamoxifeno, prednisona, dacarbazina, mecloretamina, cisplatino, metotrexato, 5-fluorouracilo y ara-C. Puede encontrarse información adicional en *The Molecular Basis of Cancer*, Mendelsohn and Israel, eds., Capítulo I, titulado “Cell cycle regulation, oncogens, and antineoplastic drugs” de Murakami *et al.*, (WB Saunders: Philadelphia, 1995), especialmente p. 13.

La “doxorrubicina” es un antibiótico de antraciclina. El nombre químico completo de la doxorrubicina es (8S-cis)-10-[(3-amino-2,3,6-tridesoxi- α -L-lixo-hexapiranosil)oxi]-7,8,9,10-tetrahidro-6,8,11-trihidroxi-8-(hidroxiacetil)-1-metoxi-5,12-naftacenediona.

La expresión “citoquina” es un término genérico para proteínas liberadas por una población celular que actúan sobre otra célula como mediadores intercelulares. Son ejemplos de dichas citoquinas linfoquinas, monoquinas, y hormonas polipeptídicas tradicionales. Entre las citoquinas se incluyen hormonas del crecimiento tales como hormona humana del crecimiento, N-metionil hormona humana del crecimiento, y hormona bovina del crecimiento; hormona paratiroidea; tiroxina; insulina; proinsulina; relaxina; prorrelaxina; hormonas de glucoproteínas tales como hormona estimulante del folículo (FSH), hormona estimulante de la tiroides (TSH) y hormona luteinizante (LH); factor de crecimiento hepático; factor de crecimiento de fibroblastos; prolactina; lactógeno placentario; factor de necrosis tumoral- α y - β ; sustancia inhibidora mulleriana; péptido asociado a gonadotropina de ratón; inhibina; activina; factor de crecimiento endotelial vascular; integrina; trombopoyetina (TPO); factores de crecimiento nerviosos tales como NGF- β ; factor de crecimiento de plaquetas; factores de crecimiento transformantes (TGF) tales como TGF- α y TGF- β ; factor de crecimiento similar a insulina I y II: eritropoyetina (EPO); factores osteoinductores; interferones tales como interferón - α , - β y - γ ; factores estimuladores de colonia (CSF) tales como macrófago-CSF (M-CSF); granulocito-macrófago-CSF (GM-CSF); y granulocito-CSF (G-CSF); interleuquinas (IL) tales como IL-1, IL-1a, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; un factor de necrosis tumoral tal como TNF- α o TNF- β ; y otros factores polipeptídicos incluyendo LIF y ligando kit (KL). Como se usa en este documento, el término citoquina incluye proteínas de fuentes naturales o de cultivo de células recombinantes y equivalentes biológicamente activos de las citoquinas de secuencia nativa.

El término “profármaco” como se usa en esta solicitud se refiere a una forma precursora o derivada de una sustancia farmacéuticamente activa que es menos citotóxica para células tumorales en comparación con el fármaco parental y es capaz de ser activado o convertido enzimáticamente en la forma parental más activa. Véase, por ejemplo, Wilman, “Prodrugs in Cancer Chemotherapy”, *Biochemical Society Transactions*, 14: 375-382, 615th Meeting, Belfast (1986), y Stella *et al.*, “Prodrugs: A Chemical Approach to Targeted Drug Delivery”, *Directed Drug Delivery*, Borchardt *et al.*, (ed.), p. 147-267, Humana Press (1985). Los profármacos de la presente invención incluyen, aunque sin limitación, profármacos que contienen fosfato, profármacos que contienen tiofosfato, profármacos que contienen sulfato, profármacos que contienen péptidos, profármacos modificados por D aminoácidos, profármacos glucosilados, profármacos que contienen β -lactama, profármacos que contienen fenoxiacetamida opcionalmente sustituida o profármacos que contienen fenilacetamida opcionalmente sustituida, 5-fluorocitosina y otros profármacos de 5-fluorouridina que pueden convertirse en el fármaco libre citotóxico más activo. Los ejemplos de fármacos citotóxicos que pueden deri-

vatizarse en forma de profármacos para uso en la presente invención incluyen, aunque sin limitación, aquellos agentes quimioterapéuticos descritos anteriormente.

Una “cantidad eficaz” de un polipéptido descrito en este documento o un antagonista del mismo, en referencia a la inhibición del crecimiento de células neoplásicas, crecimiento tumoral o crecimiento de células cancerosas, es una cantidad capaz de inhibir, hasta cierto punto, el crecimiento de células diana. La expresión incluye una cantidad capaz de inducir un efecto inhibidor del crecimiento, citostático y/o citotóxico y/o apoptosis de las células diana. Una “cantidad eficaz” de un antagonista de polipéptido PRO para fines de inhibir el crecimiento de células neoplásicas, crecimiento tumoral o crecimiento de células cancerosas, puede determinarse empíricamente y de manera rutinaria.

Una “cantidad terapéuticamente eficaz”, en referencia al tratamiento de tumores, se refiere a una cantidad capaz de inducir uno o más de los siguientes efectos: (1) inhibición, hasta cierto punto, del crecimiento tumoral, incluyendo, ralentizar y detener completamente el crecimiento; (2) reducción del número de células tumorales; (3) reducción del tamaño del tumor; (4) inhibición (es decir, reducción, ralentización o detención completa) de la infiltración de células tumorales en órganos periféricos; (5) inhibición (es decir, reducción, ralentización o detención completa) de la metástasis; (6) aumento de la respuesta inmune anti-tumoral, que puede, pero no debe, dar como resultado la regresión o el rechazo del tumor; y/o (7) alivio, hasta cierto punto, de uno o más síntomas asociados con el trastorno. Una “cantidad terapéuticamente eficaz” de un antagonista de polipéptido PRO para fines de tratamiento de tumores puede determinarse empíricamente y de manera rutinaria.

Una “cantidad inhibidora del crecimiento” de un antagonista de PRO es una cantidad capaz de inhibir el crecimiento de una célula, especialmente tumoral, por ejemplo, célula cancerosa, *in vitro* o *in vivo*. Una “cantidad inhibidora del crecimiento” de un antagonista de PRO para fines de inhibir el crecimiento de células neoplásicas puede determinarse empíricamente y de manera rutinaria.

Una “cantidad citotóxica” de un antagonista de PRO es una cantidad capaz de causar la destrucción de una célula, especialmente tumoral, por ejemplo, célula cancerosa, *in vitro* o *in vivo*. Una “cantidad citotóxica” de un antagonista de PRO para fines de inhibir el crecimiento de células neoplásicas puede determinarse empíricamente y de manera rutinaria.

Las expresiones polipéptido o proteína “PRO” cuando se usan en este documento abarcan polipéptidos PRO de secuencia nativa y variantes de polipéptido PRO (que se definen adicionalmente en este documento). El polipéptido PRO puede aislarse de diversas fuentes, tales como de tipos de tejido humano o de otra fuente, o prepararse mediante procedimientos recombinantes y/o sintéticos.

Un “PRO de secuencia nativa” comprende un polipéptido que tiene la misma secuencia de aminoácidos que el polipéptido PRO obtenido de la naturaleza. Dicho polipéptido PRO de secuencia nativa puede aislarse de la naturaleza o puede producirse mediante medios recombinantes y/o sintéticos. La expresión PRO de “secuencia nativa” abarca específicamente formas truncadas o secretadas de origen natural (por ejemplo, una secuencia de dominio extracelular), formas variantes de origen natural (por ejemplo, formas cortadas de forma alternativa) y variantes alélicas de origen natural de los polipéptidos PRO. En una realización aquí descrita, el polipéptido PRO de secuencia nativa es un polipéptido PRO de secuencia nativa maduro o de longitud completa como se muestra en la Figura 2 (SEC ID N°: 27) respectivamente. Además, aunque los polipéptidos PRO descritos en la Figura 2 (SEC ID N°: 27) respectivamente, han mostrado comenzar con el resto de metionina designado en su interior como posición de aminoácido 1, es concebible y posible que otro resto de metionina situado cadena arriba o cadena abajo de la posición de aminoácido 1 en la Figura 2 (SEC ID N°: 27) respectivamente, pueda emplearse como resto de aminoácido de inicio para el polipéptido PRO.

El “dominio extracelular” o “ECD” de un polipéptido descrito en este documento se refiere a una forma del polipéptido que está esencialmente libre de los dominios transmembrana y citoplasmático. Habitualmente, un ECD polipeptídico tendrá menos de aproximadamente el 1% de dichos dominios transmembrana y/o citoplasmático y preferentemente, tendrá menos de aproximadamente el 0,5% de dichos dominios. Se entenderá que cualesquier dominio o dominios transmembrana identificados para los polipéptidos de la presente invención se identifican según criterios empleados de forma rutinaria en el sector para identificar a ese tipo de dominio hidrofóbico. Los límites exactos de un dominio transmembrana pueden variar pero lo más probable es que en no más de aproximadamente aminoácidos en cualquier extremo del dominio según se identificó inicialmente y como se muestra en las figuras adjuntas. Como tal, en una realización de la presente invención, el dominio extracelular de un polipéptido de la presente invención comprende los aminoácidos 1 a X de la secuencia de aminoácidos madura, donde X es cualquier aminoácido en un radio de 5 aminoácidos en cualquier lado del límite del dominio extracelular/dominio transmembrana.

La ubicación aproximada de los “péptidos señal” de los diversos polipéptidos PRO descritos en este documento se muestra en las figuras adjuntas. Se observa, sin embargo, que el límite C-terminal de un péptido señal puede variar, pero lo más probable es que en no más de aproximadamente 5 aminoácidos en cualquier lado del límite C-terminal del péptido señal según se identificó inicialmente en este documento, donde el límite C-terminal del péptido señal puede identificarse según criterios empleados de forma rutinaria en el sector para identificar a ese tipo de elemento de secuencia de aminoácidos (por ejemplo, Nielsen *et al.*, Prot. Eng., 10: 1-6 (1997) y von Heinjc *et al.*, Nucl. Acids. Res., 14: 4683-4690 (1986)). Además, también se reconoce que, en algunos casos, la escisión de una secuencia señal de un polipéptido secretado no es completamente uniforme, dando como resultado más de una especie secretada. Estos polipéptidos maduros, donde el péptido señal se escinde en un radio de no más de aproximadamente 5 aminoácidos

en cualquier lado del límite C-terminal del péptido señal según se identificó en este documento, y los polinucleótidos que los codifican, están contemplados por la presente invención.

“Variante de polipéptido PRO” significa un polipéptido PRO activo como se define anterior o posteriormente que tiene al menos aproximadamente el 80% de identidad en la secuencia de aminoácidos con una secuencia de polipéptido PRO de secuencia nativa de longitud completa como se describe en este documento, una secuencia de polipéptido PRO que carece del péptido señal como se describe en este documento, un dominio extracelular de un polipéptido PRO, con o sin el péptido señal, como se describe en este documento o cualquier otro fragmento de una secuencia de polipéptido PRO de longitud completa como se describe en este documento. Dichas variantes de polipéptido incluyen, por ejemplo, polipéptidos PRO en los que se añaden o eliminan uno o más restos de aminoácidos, en el extremo N o C de la secuencia de aminoácidos nativa de longitud completa. Habitualmente, una variante de polipéptido PRO tendrá al menos aproximadamente el 80% de identidad en la secuencia de aminoácidos, preferentemente al menos aproximadamente el 81% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 82% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 83% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 84% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 85% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 86% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 87% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 88% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 89% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 90% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 91% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 92% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 93% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 94% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 95% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 96% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 97% de identidad en la secuencia de aminoácidos, más preferentemente al menos aproximadamente el 98% de identidad en la secuencia de aminoácidos y lo más preferentemente al menos aproximadamente el 99% de identidad en la secuencia de aminoácidos con una secuencia de polipéptido PRO de secuencia nativa de longitud completa como se describe en este documento, una secuencia de polipéptido PRO que carece del péptido señal como se describe en este documento, un dominio extracelular de un polipéptido PRO, con o sin el péptido señal, como se describe en este documento o cualquier otro fragmento de una secuencia de polipéptido PRO de longitud completa como se describe en este documento. Habitualmente, los polipéptidos PRO variantes tienen al menos aproximadamente 10 aminoácidos de longitud, a menudo al menos aproximadamente 20 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 30 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 40 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 50 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 60 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 70 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 80 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 90 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 100 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 150 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 200 aminoácidos de longitud, con mayor frecuencia al menos aproximadamente 300 aminoácidos de longitud, o más.

Como se muestra posteriormente, la Tabla I proporciona el código fuente completo para el programa informático de alineamiento de secuencias ALIGN-2. Este código fuente puede compilarse de forma rutinaria para su uso en un sistema operativo UNIX para proporcionar el programa informático de comparación de secuencias ALIGN-2.

Además, Las Tablas 2A-2B muestran ejemplificaciones hipotéticas para usar el procedimiento descrito posteriormente para determinar el % de identidad en la secuencia de aminoácidos (Tablas 2A-2B) y el % identidad en la secuencia de ácido nucleico (Tablas 2C-2D) usando el programa informático de comparación de secuencias ALIGN-2, donde “PRO” representa la secuencia de aminoácidos de un hipotético polipéptido PRO de interés, la “Proteína de Comparación” representa la secuencia de aminoácidos de un polipéptido con el cual el polipéptido “PRO” de interés se está comparando, “ADN de PRO” representa una hipotética secuencia de ácido nucleico que codifica PRO de interés, “ADN de Comparación” representa la secuencia de nucleótidos de una molécula de ácido nucleico con la cual la molécula de ácido nucleico “ADN de PRO” de interés se está comparando, “X”, “Y” y “Z” representan cada una diferentes restos de aminoácidos hipotéticos y “N”, “L” y “V” representan cada uno diferentes nucleótidos hipotéticos.

ES 2 329 689 T3

TABLA 1

```

5  /*
   *
   * C-C increased from 12 to 15
   * Z is average of EQ
10  * B is average of ND
   * match with stop is _M; stop-stop = 0; J (joker) match = 0
   */
   #define _M      -8      /* value of a match with a stop */

15  int  _day[26][26] = {
/*  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
20 /* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
25 /* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ {-1, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
30 /* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, 0, _M, _M, _M, _M, _M, _M, _M, _M, _M},
35 /* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
40 /* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
45 /* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

ES 2 329 689 T3

```

/*
 */
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX           4      /* save if there's at least MX-1 bases since last jmp */

#define DMAT         3      /* value of matching bases */
#define DMIS         0      /* penalty for mismatched bases */
#define DINS0        8      /* penalty for a gap */
#define DINS1        1      /* penalty per base */
#define PINS0        8      /* penalty for a gap */
#define PINS1        4      /* penalty per residue */

struct jmp {
    short          n[MAXJMP]; /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
}; /* limits seq to 2^16 -1 */

struct diag {
    int            score;      /* score at last jmp */
    long           offset;     /* offset of prev block */
    short          ijmp;       /* current jmp index */
    struct jmp      jp;        /* list of jmps */
};

struct path {
    int            spc;        /* number of leading spaces */
    short          n[JMPS];    /* size of jmp (gap) */
    int            x[JMPS];    /* loc of jmp (last elem before gap) */
};

char             *ofile;      /* output file name */
char             *namex[2];   /* seq names: getseqs() */
char             *prog;       /* prog name for err msgs */
char             *seqs[2];    /* seqs: getseqs() */
int              dmax;        /* best diag: nw() */
int              dmax0;       /* final diag */
int              dna;         /* set if dna: main() */
int              endgaps;     /* set if penalizing end gaps */
int              gapx, gapy;   /* total gaps in seqs */
int              len0, len1;   /* seq lens */
int              ngupx, ngapy; /* total size of gaps */
int              smax;        /* max score: nw() */
int              *xbm;        /* bitmap for matching */
long             offset;      /* current offset in jmp file */
struct diag      *dx;         /* holds diagonals */
struct path      pp[2];       /* holds path for seqs */

char             *calloc(), *malloc(), *index(), *strcpy();
char             *getseq0, *g_calloc();

```

ES 2 329 689 T3

```

/* Needleman-Wunsch alignment program
 *
 * usage: prog1 file1 file2
 * where file1 and file2 are two dna or two protein sequences.
 * The sequences can be in upper- or lower-case and may contain ambiguity
 * Any lines beginning with ';', '>' or '<' are ignored
 * Max file length is 65535 (limited by unsigned short x in the jmp struct)
 * A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 * Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static _dbval[26] = {
    1, 14, 2, 13, 0, 0, 4, 11, 0, 0, 12, 0, 3, 15, 0, 0, 0, 5, 6, 8, 7, 9, 0, 10, 0
};

static _pbval[26] = {
    1, 2 | (1 < < ('D'-'A')) | (1 < < ('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1 < < 10, 1 < < 11, 1 < < 12, 1 < < 13, 1 < < 14,
    1 < < 15, 1 < < 16, 1 < < 17, 1 < < 18, 1 < < 19, 1 < < 20, 1 < < 21, 1 < < 22,
    1 < < 23, 1 < < 24, 1 < < 25 | (1 < < ('E'-'A')) | (1 < < ('Q'-'A'))
};

main(ac, av)
int    ac;
char   *av[];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';', '>' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file 'align.out'\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0; /* 1 to penalize endgaps */
    ofile = "align.out"; /* output file */

    nw(); /* fill in the matrix, get the possible jmps */
    readjmps(); /* get the actual jmps */
    print(); /* print stats, alignment */

    cleanup(0); /* unlink any tmp files */
}

```

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()
{
    char      *px, *py;      /* seqs and ptrs */
    int       *ndely, *dely; /* keep track of dely */
    int       ndelx, delx;   /* keep track of delx */
    int       *tmp;         /* for swapping row0, row1 */
    int       mis;          /* score for each type */
    int       ins0, ins1;    /* insertion penalties */
    register  id;           /* diagonal index */
    register  ij;           /* jmp index */
    register  *col0, *col1; /* score for curr, last row */
    register  xx, yy;       /* index into seqs */

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0; /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
     */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

...nw

```

5   for (py = seqx[1], yy = 1; yy <= len1; py += yy++) {
      mis = col0[yy-1];
      if (dna)
          mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
      else
          mis += _day[*px-'A'][*py-'A'];

10      /* update penalty for del in x seq;
         * favor new del over ongoing del
         * ignore MAXGAP if weighting endgaps
         */
      if (endgaps || ndely[yy] < MAXGAP) {
          if (col0[yy] - ins0 >= dely[yy]) {
15              dely[yy] = col0[yy] - (ins0+ins1);
              ndely[yy] = 1;
          } else {
              dely[yy] -= ins1;
              ndely[yy]++;
          }
      } else {
          if (col0[yy] - (ins0+ins1) >= dely[yy]) {
20              dely[yy] = col0[yy] - (ins0+ins1);
              ndely[yy] = 1;
          } else
              ndely[yy]++;
      }

      /* update penalty for del in y seq;
         * favor new del over ongoing del
         */
30      if (endgaps || ndelx < MAXGAP) {
          if (col1[yy-1] - ins0 >= delx) {
              delx = col1[yy-1] - (ins0+ins1);
              ndelx = 1;
          } else {
35              delx -= ins1;
              ndelx++;
          }
      } else {
          if (col1[yy-1] - (ins0+ins1) >= delx) {
40              delx = col1[yy-1] - (ins0+ins1);
              ndelx = 1;
          } else
              ndelx++;
      }

45      /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */

```

50

55

60

65

...nw

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    coll[yy] = mis;
5 else if (delx >= dely[yy]) {
    coll[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
10 && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = ndelx;
        dx[id].jp.x[ij] = xx;
        dx[id].score = delx;
20     }
    else {
        coll[yy] = dely[yy];
        ij = dx[id].ijmp;

25 if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writejumps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
                offset += sizeof(struct jmp) + sizeof(offset);
            }
            dx[id].jp.n[ij] = -ndely[yy];
            dx[id].jp.x[ij] = xx;
            dx[id].score = dely[yy];
35     }
    if (xx == len0 && yy < len1) {
        /* last col
        */
        if (endgaps)
            coll[yy] -= ins0 + ins1*(len1-yy);
        if (coll[yy] > smax) {
            smax = coll[yy];
            dmax = id;
40     }
    }
}
if (endgaps && xx < len0)
    coll[yy-1] -= ins0 + ins1*(len0-xx);
if (coll[yy-1] > smax) {
    smax = coll[yy-1];
    dmax = id;
50 }
tmp = col0; col0 = coll; coll = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)coll);
}

```

ES 2 329 689 T3

```

/*
 *
 * print() -- only routine visible outside this module
 *
5  * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
10 * putline() -- put out a line (name, {num}, seq, {num}): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

15 #include "nw.h"

#define SPC      3
#define P_LINE 256 /* maximum output line */
#define P_SPC    3 /* space between name or num and seq */

20 extern _day[26][26];
int olen; /* set output line length */
FILE *fx; /* output file */

print() print
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

```

/*
 * trace back the best path, count matches
 */
static
5 getmat(lx, ly, firstgap, lastgap)
    int    lx, ly;          /* "core" (minus endgaps) */
    int    firstgap, lastgap; /* leading trailing overlap */
{
    int      nm, i0, i1, siz0, siz1;
    char     outx[32];
    double    pct;
    register  n0, n1;
    register char *p0, *p1;

10
    /* get total matches, score
    */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

20
    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
35
            if (xbrn[*p0-'A']&xbrn[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
40
        }
    }

    /* pct homology:
    * if penalizing endgaps, base is the shorter seq
    * else, knock off overhangs and take shorter core
    */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
50
    else
        lx = (lx < ly)? lx : ly;
    pct = 100.*((double)nm)/((double)lx);
    fprintf(fx, "\n");
    fprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
55
        nm, (nm == 1)? "" : "es", lx, pct);

```

getmat

ES 2 329 689 T3

...getmat

```

5      fprintf(fx, "< gaps in first sequence: %d", gapx);
      if (gapx) {
          (void) sprintf(ouix, " (%d %s%s)",
              ngapx, (dna)? "base": "residue", (ngapx == 1)? "" : "s");
          fprintf(fx, "%s", ouix);

      fprintf(fx, ", gaps in second sequence: %d", gapy);
10     if (gapy) {
          (void) sprintf(ouix, " (%d %s%s)",
              ngapy, (dna)? "base": "residue", (ngapy == 1)? "" : "s");
          fprintf(fx, "%s", ouix);
      }
      if (dna)
15         fprintf(fx,
            "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINSO, DINSI);
      else
          fprintf(fx,
20         "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            sinax, PINSO, PINSI);
      if (endgaps)
          fprintf(fx,
25         "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
            lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
      else
          fprintf(fx, "< endgaps not penalized\n");
    }

    static      nm;          /* matches in core -- for checking */
30    static      lmax;       /* lengths of stripped file names */
    static      ij[2];       /* jmp index for a path */
    static      nc[2];       /* number at start of current line */
    static      ni[2];       /* current elem number -- for gapping */
    static      siz[2];
35    static char  *ps[2];     /* ptr to current element */
    static char  *po[2];     /* ptr to next output char slot */
    static char  out[2][P_LINE]; /* output line */
    static char  star[P_LINE]; /* set by stars() */

    /*
40    * print alignment of described in struct path pp[]
    */
    static
    pr_align()
    {
45        int      nn;        /* char count */
        int      more;
        register  i;

        for (i = 0, lmax = 0; i < 2; i++) {
50            nn = stripname(namex[i]);
            if (nn > lmax)
                lmax = nn;

            nc[i] = 1;
            ni[i] = 1;
55            siz[i] = ij[i] = 0;
            ps[i] = seqx[i];
            po[i] = out[i];
        }
    }

```

pr_align

```

5      for (nn = nm = 0, more = 1; more;) {
        for (i = more = 0; i < 2; i++) {
            /*
            * do we have more of this sequence?
            */
            if (!*ps[i])
                continue;

            more++;

10         if (pp[i].spc) { /* leading space */
                *po[i]++ = ' ';
                pp[i].spc--;
            }
            else if (siz[i]) { /* in a gap */
                *po[i]++ = '-';
                siz[i]--;
            }
            else { /* we're putting a seq element
            */
                *po[i] = *ps[i];
                if (islower(*ps[i]))
                    *ps[i] = toupper(*ps[i]);
                po[i]++;
                ps[i]++;

25             /*
            * are we at next gap for this seq?
            */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
            * we need to merge all gaps
            * at this location
            */
                siz[i] = pp[i].n[ij[i] + 1];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i] + 1];

35             }
                ni[i]++;
            }
        }
        if (++nn == olen || !more && nn) {
            dumpblock();
            for (i = 0; i < 2; i++)
                po[i] = out[i];
            nn = 0;
        }
    }

    /*
    * dump a block of lines, including numbers, stars: pr_align()
    */
50     static
    dumpblock()
    {
        register i;

55         for (i = 0; i < 2; i++)
            *po[i]-- = '\0';
    }
}

```

...pr_align

dumpblock

...dumpblock

```

5      (void)putc('\n', fx);
      for (i = 0; i < 2; i++) {
          if (*out[i] && (*out[i] != ' ' || *(pof[i]) != ' ')) {
              if (i == 0)
                  nums(i);
              if (i == 0 && *out[1])
                  stars();
10             putline(i);
              if (i == 0 && *out[1])
                  fprintf(fx, star);
              if (i == 1)
                  nums(i);
15         }
    }

/*
   * put out a number line: dumpblock()
   */
static
nums(ix)                                nums
{
    int      ix;      /* index in out[] holding seq line */

    char      nline[P_LINE];
    register  i, j;
    register char *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax + P_SPC; i++, pn++)
        *pn = ' ';
30    for (i = nc[ix], py = out[ix]; *py: py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
35                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
40            }
            else
                *pn = ' ';
            i++;
        }
    }
45    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void)putc(*pn, fx);
    (void)putc('\n', fx);
50 }

/*
   * put out a line (name, {num}, seq, {num}): dumpblock()
   */
static
putline(ix)                                putline
{
    int      ix;

```

60

65

```

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void)putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void)putc(' ', fx);

/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
for (px = out[ix]; *px; px++)
    (void)putc(*px&0x7F, fx);
(void)putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumplock()
 */
static
stars()
{
    int          i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(p0[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(p1[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
}

/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
char *pn; /* file name (may be path) */
{
    register char *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;

    if (py)
        (void)strcpy(pn, py);
    return(strlen(pn));
}

```

...putline

stars

stripname

ES 2 329 689 T3

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq. set dna, len, maxlen
5  * g_alloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps. from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
10 #include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";    /* tmp file for jumps */
FILE    *fj;

int      cleanup();                    /* cleanup tmp file */
15 long   lseek();

/*
 * remove any tmp file if we blow
 */
20 cleanup(i)
    int    i;
{
    if (fj)
        (void) unlink(jname);
25     exit(i);
}

/*
 * read, return ptr to seq. set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
30 * seq in upper or lower case
 */
char    *
getseq(file, len)
35     char    *file;    /* file name */
    int      *len;    /* seq len */
{
    char        line[1024], *pseq;
    register char    *px, *py;
    int          natgc, tlen;
    FILE         *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
60
65

```

cleanup

getseq

...getseq

```

py = pseq + 4;
*len = tlen;
rewind(fp);

5      while (fgets(line, 1024, fp)) {
          if (*line == ';' || *line == '<' || *line == '>')
              continue;
          for (px = line; *px != '\n'; px++) {
10              if (isupper(*px))
                  *py++ = *px;
              else if (islower(*px))
                  *py++ = toupper(*px);
              if (index("ATGCU", *(py-1)))
15                  naigc++;
          }
      }
      *py++ = '\0';
      *py = '\0';
      (void) fclose(fp);
      dna = naigc > (tlen/3);
      return(pseq+4);
}

25 char *
g_alloc(msg, nx, sz)
    char *msg; /* program, calling routine */
    int nx, sz; /* number and size of elements */
{
    char *px, *calloc();

30      if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
          if (*msg) {
              fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
              exit(1);
          }
35      }
      return(px);
}

/*
40  * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
*/
readjmps()
    {
        int fd = -1;
        int siz, i0, i1;
        register i, j, xx;

        if (fj) {
            (void) fclose(fj);
            if ((fd = open(jname, O_RDONLY, 0)) < 0) {
20                fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                cleanup(1);
            }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
55            while (1) {
                for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                    ;
            }
        }
    }

```

g_alloc

readjmps

...readjumps

```

    if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;

        /* id = xx - yy + len1 - 1
        */
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
}
else
    break;
}

/* reverse the order of jumps
*/
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}

```

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
5  writejmps(ix)                                writejmps
    int    ix;
    {
        char    *mkiemp();
10         if (!fj) {
            if (mkiemp(jname) < 0) {
                fprintf(stderr, "%s: can't mkiemp() %s\n", prog, jname);
                cleanup(1);
            }
            if ((fj = fopen(jname, "w")) == 0) {
15                 fprintf(stderr, "%s: can't write %s\n", prog, jname);
                exit(1);
            }
        }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
20         (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
    }
}

```

TABLA A

25	PRO	XXXXXXXXXXXXXXXXXX	(Longitud = 15 aminoácidos)
30	Proteína de Comparación	XXXXXXYYYYYYY	(Longitud = 12 aminoácidos)
35	% de identidad en la secuencia de aminoácidos = (el número de restos de aminoácidos que se emparejan de forma idéntica entre las dos secuencias de polipéptidos según se determinó mediante ALIGN-2) dividido entre (el número total de restos de aminoácidos del polipéptido PRO) = 5 dividido entre 15 = 33,3%		

TABLA 2B

50	PRO	XXXXXXXXXXXX	(Longitud = 10 aminoácidos)
55	Proteína de Comparación	XXXXXXYYYYYYZZYZ	(Longitud = 15 aminoácidos)
60	% de identidad en la secuencia de aminoácidos = (el número de restos de aminoácidos que se emparejan de forma idéntica entre las dos secuencias de polipéptidos según se determinó mediante ALIGN-2) dividido entre (el número total de restos de aminoácidos del polipéptido PRO) = 5 dividido entre 10 = 50%		

ES 2 329 689 T3

TABLA 2C

5	ADN de PRO	NNNNNNNNNNNNNNNN	(Longitud = 14 nucleótidos)
	ADN de Comparación	NNNNNNLLLLLLLLLLLL	(Longitud = 16 nucleótidos)
10			
	% identidad en la secuencia de ácido nucleico = (el		
	número de nucleótidos que se emparejan de forma idéntica		
15	entre las dos secuencias de ácido nucleico según se		
	determinó mediante ALIGN-2) dividido entre (el número		
	total de nucleótidos de la secuencia de ácido nucleico de		
20	ADN DE PRO) = 6 dividido entre 14 = 42,9%		

TABLA 2D

25	ADN de PRO	NNNNNNNNNNNNNN	(Longitud = 12 nucleótidos)
30	ADN de Comparación	NNNNLLLVV	(Longitud = 9 nucleótidos)
35	% identidad en la secuencia de ácido nucleico = (el		
	número de nucleótidos que se emparejan de forma idéntica		
	entre las dos secuencias de ácido nucleico según se		
40	determinó mediante ALIGN-2) dividido entre (el número		
	total de nucleótidos de la secuencia de ácido nucleico de		
45	ADN de PRO) = 4 dividido entre 12 = 33,3 %		

“Porcentaje (%) de identidad en la secuencia de aminoácidos” con respecto a las secuencias de polipéptido PRO identificadas en este documento se define como el porcentaje de restos de aminoácidos en una secuencia candidata que son idénticos a los restos de aminoácidos en una secuencia de PRO, después de alinear las secuencias e introducir huecos, si fuera necesario, para conseguir el máximo porcentaje de identidad en la secuencia, y sin considerar cualesquiera sustituciones conservativas como parte de la identidad en la secuencia. El alineamiento para fines de determinar el porcentaje de identidad en la secuencia de aminoácidos puede conseguirse de diversas maneras que están dentro del alcance de un experto en la materia, por ejemplo, usando software informático disponible públicamente tal como BLAST, BLAST-2, ALIGN, ALIGN-2 o Megalign (DNASTAR). Los expertos en la materia pueden determinar parámetros apropiados para medir el alineamiento, incluyendo cualesquiera algoritmos necesarios para conseguir el máximo alineamiento en toda la longitud de las secuencias comparadas. Para los fines de este documento, sin embargo, los valores de % de identidad en la secuencia de aminoácidos se obtienen como se describe posteriormente usando el programa informático de comparación de secuencias ALIGN-2, donde el código fuente completo para el programa ALIGN-2 se proporciona en la Tabla 1. EL programa informático de comparación de secuencias ALIGN-2 es propiedad de Genentech, Inc., y el código fuente mostrado en la Tabla 1 se ha presentado con documentación del usuario en la Oficina de los Derechos de Autor Estadounidense (U.S. Copyright Office), Washington D.C., 20559, donde está registrada con el N° de registro “U.S. Copyright Registration No. TXU510087”. El programa ALIGN-2 está disponible públicamente a través de Genentech, Inc., South San Francisco, California o puede compilarse a partir del código fuente que se proporciona en la Tabla I. El programa ALIGN-2 debe compilarse para su uso en un sistema operativo UNIX, preferentemente UNIX digital V4.0D. Todos los parámetros de comparación de secuencia los establece el programa ALIGN-2 y no varían.

ES 2 329 689 T3

Para los fines de este documento, el % de identidad en la secuencia de aminoácidos de una secuencia de aminoácidos dada A respecto a, con o contra una secuencia de aminoácidos dada B (que como alternativa puede expresarse como una secuencia de aminoácidos dada A que tiene o comprende cierto % de identidad en la secuencia de aminoácidos respecto a, con o contra una secuencia de aminoácidos dada B) se calcula de la siguiente manera:

$$100 \text{ veces la fracción } X/Y$$

donde X es el número de restos de aminoácidos valorado como emparejamientos idénticos por el programa de alineamiento de secuencias ALIGN-2 en el alineamiento que hace ese programa de A y B, y donde Y es el número total de restos de aminoácidos en B. Se entenderá que cuando la longitud de la secuencia de aminoácidos A no es igual a la longitud de la secuencia de aminoácidos B, el % de identidad en la secuencia de aminoácidos de A con respecto a B no será igual al % de identidad en la secuencia de aminoácidos de B con respecto a A. Como ejemplos de cálculos del % de identidad en la secuencia de aminoácidos, las Tablas 2A-2B demuestran cómo calcular el % de identidad en la secuencia de aminoácidos de la secuencia de aminoácidos designada "Proteína de Comparación" con respecto a la secuencia de aminoácidos designada "PRO".

A no ser que se afirme específicamente otra cosa, todos los valores de % de identidad en la secuencia de aminoácidos usados en este documento se obtienen como se ha descrito anteriormente usando el programa informático de comparación de secuencias ALIGN-2. Sin embargo, el % de identidad en la secuencia de aminoácidos también puede determinarse usando el programa de comparación de secuencias NCBI-BLAST2 (Altschul *et al.*, Nucleic Acids Res., 25: 3389-3402 (1997)). El programa de comparación de secuencias NCBI-BLAST2 puede descargarse de <http://www.ncbi.nlm.nih.gov>. NCBI-BLAST2 usa varios parámetros de búsqueda, donde todos estos parámetros de búsqueda se ajustan a valores por defecto incluyendo, por ejemplo, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0,01, constant for multi-pass - 25, dropoff for final gapped alignment = 25 y scoring matrix = BLOSUM62.

En situaciones en las que se emplea NCBI-BLAST2 para comparaciones de secuencias de aminoácidos, el % de identidad en la secuencia de aminoácidos de una secuencia de aminoácidos dada A respecto a, con o contra una secuencia de aminoácidos dada B (que como alternativa puede expresarse como una secuencia de aminoácidos dada A que tiene o comprende cierto % de identidad en la secuencia de aminoácidos respecto a, con o contra una secuencia de aminoácidos dada B) se calcula de la siguiente manera:

$$100 \text{ veces la fracción } X/Y$$

donde X es el número de restos de aminoácidos valorado como emparejamientos idénticos por el programa de alineamiento de secuencias NCBI-BLAST2 en el alineamiento que hace ese programa de A y B, y donde Y es el número total de restos de aminoácidos en B. Se entenderá que cuando la longitud de la secuencia de aminoácidos A no es igual a la longitud de la secuencia de aminoácidos B, el % de identidad en la secuencia de aminoácidos de A con respecto a B no será igual al % de identidad en la secuencia de aminoácidos de B con respecto a A.

Además, el % de identidad en la secuencia de aminoácidos también puede determinarse usando el programa informático WU-BLAST-2 (Altschul *et al.*, Methods in Enzymology, 266: 460-480 (1996)). La mayoría de los parámetros de búsqueda de WU-BLAST-2 se ajustan a los valores por defecto. Aquellos que no se ajustan a valores por defecto, es decir, los parámetros ajustables, se ajustan con los siguientes valores: overlap span = 1, overlap fraction = 0,125, word threshold (T) = 11, y scoring matrix = BLOSUM62. Para los fines de este documento, un valor de % de identidad en la secuencia de aminoácidos se determina dividiendo (a) el número de restos de aminoácidos con emparejamiento idéntico entre la secuencia de aminoácidos del polipéptido PRO de interés que tiene una secuencia obtenida del polipéptido PRO nativo y la secuencia de aminoácidos de comparación de interés (es decir, la secuencia con la que el polipéptido PRO de interés se está comparando que puede ser un polipéptido PRO variante) según se determinó mediante WU-BLAST-2 por (b) el número total de restos de aminoácidos del polipéptido PRO de interés. Por ejemplo, en la afirmación "un polipéptido que comprende una secuencia de aminoácidos A que tiene o teniendo al menos el 80% de identidad en la secuencia de aminoácidos con respecto a la secuencia de aminoácidos B", la secuencia de aminoácidos A es la secuencia de aminoácidos de comparación de interés y la secuencia de aminoácidos B es la secuencia de aminoácidos del polipéptido PRO de interés. "Polinucleótido PRO variante" o "secuencia de ácido nucleico PRO variante" significa una molécula de ácido nucleico que codifica un polipéptido PRO activo como se define posteriormente y que tiene al menos aproximadamente el 80% de identidad en la secuencia de ácido nucleico con una secuencia de ácido nucleico que codifica una secuencia de polipéptido PRO de secuencia nativa de longitud completa como se describe en este documento, una secuencia de polipéptido PRO de secuencia nativa de longitud completa que carece del péptido señal como se describe en este documento, un dominio extracelular de un polipéptido PRO, con o sin el péptido señal, como se describe en este documento o cualquier otro fragmento de una secuencia de polipéptido PRO de longitud completa como se describe en este documento. Habitualmente, un polinucleótido PRO variante tendrá al menos aproximadamente el 80% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 81% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 82% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 83% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 84% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 85% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 86% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 87% de identidad en la secuencia de ácido nucleico,

más preferentemente al menos aproximadamente el 88% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 89% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 90% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 91% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 92% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 93% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 94% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 95% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 96% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 97% de identidad en la secuencia de ácido nucleico, más preferentemente al menos aproximadamente el 98% de identidad en la secuencia de ácido nucleico y aún más preferentemente al menos aproximadamente el 99% de identidad en la secuencia de ácido nucleico con la secuencia de ácido nucleico que codifica una secuencia de polipéptido PRO de secuencia nativa de longitud completa como se describe en este documento, una secuencia de polipéptido PRO de secuencia nativa de longitud completa que carece del péptido señal como se describe en este documento, un dominio extracelular de un polipéptido PRO, con o sin la secuencia señal, como se describe en este documento o cualquier otro fragmento de una secuencia de polipéptido PRO de longitud completa como se describe en este documento. Las variantes no abarcan la secuencia de nucleótidos nativa.

Habitualmente, los polinucleótidos PRO variantes tienen al menos aproximadamente 30 nucleótidos de longitud, a menudo al menos aproximadamente 60 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 90 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 120 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 150 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 180 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 210 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 240 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 270 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 300 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 450 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 600 nucleótidos de longitud, con mayor frecuencia al menos aproximadamente 900 nucleótidos de longitud, o más.

“Porcentaje (%) de identidad en la secuencia de ácido nucleico” con respecto a las secuencias de ácido nucleico que codifican el polipéptido PRO identificadas en este documento se define como el porcentaje de nucleótidos en una secuencia candidata que son idénticos a los nucleótidos en una secuencia de ácido nucleico que codifica el polipéptido PRO, después de alinear las secuencias e introducir huecos, si fuera necesario, para conseguir el máximo porcentaje de identidad en la secuencia. El alineamiento para fines de determinar el porcentaje de identidad en la secuencia de ácido nucleico puede conseguirse de diversas maneras que están dentro del alcance de un experto en la materia, por ejemplo, usando software informático disponible públicamente tal como software BLAST, BLAST-2, ALIGN, MALIGN-2 o Megalign (DNASTAR). Los expertos en la materia pueden determinar parámetros apropiados para medir el alineamiento, incluyendo cualesquiera algoritmos necesarios para conseguir el máximo alineamiento en la longitud completa de las secuencias comparadas. Para los fines de este documento, sin embargo, los valores de % de identidad en la secuencia de ácido nucleico se obtienen como se describe posteriormente usando el programa informático de comparación de secuencias ALIGN-2, donde el código fuente completo para el programa ALIGN-2 se proporciona en la Tabla 1. El programa informático de comparación de secuencias ALIGN-2 es propiedad de Genentech, Inc., y el código fuente mostrado en la Tabla 1 se presentó con la documentación del usuario en la Oficina de los Derechos de Autor Estadounidense (U.S. Copyright Office), Washington D.C., 20559, donde se registró con el número “U.S. Copyright Registration No.” TXU510087. El programa ALIGN-2 está disponible públicamente a través de Genentech, Inc., South San Francisco, California o puede compilarse a partir del código fuente proporcionado en la Tabla 1. El programa ALIGN-2 debe compilarse para su uso en un sistema operativo UNIX, preferentemente UNIX digital V4.0D. Todos los parámetros de comparación de secuencias los establece el programa ALIGN-2 y no varían.

Para los fines de este documento, el % de identidad en la secuencia de ácido nucleico de una secuencia de ácido nucleico dada C respecto a, con o contra una secuencia de ácido nucleico dada D (que como alternativa puede expresarse como una secuencia de ácido nucleico dada C que tiene o comprende cierto % de identidad en la secuencia de ácido nucleico respecto a, con o contra una secuencia de ácido nucleico dada D) se calcula de la siguiente manera:

$$100 \text{ veces la fracción } W/Z$$

donde W es el número de nucleótidos valorado como emparejamientos idénticos por el programa de alineamiento de secuencias ALIGN-2 en el alineamiento que hace este programa de C y D, y donde Z es el número total de nucleótidos en D. Se entenderá que cuando la longitud de la secuencia de ácido nucleico C no es igual a la longitud de la secuencia de ácido nucleico D, el % de identidad en la secuencia de ácido nucleico de C con respecto a D no será igual al % de identidad en la secuencia de ácido nucleico de D con respecto a C. Como ejemplos de cálculos de % de identidad en la secuencia de ácido nucleico, las Tablas 2C-2D demuestran cómo calcular el % de identidad en la secuencia de ácido nucleico de la secuencia de ácido nucleico designada “ADN de Comparación” con respecto a la secuencia de ácido nucleico designada “ADN de PRO”.

A no ser que se afirme específicamente otra cosa, todos los valores de % de identidad en la secuencia de ácido nucleico usados en este documento se obtienen como se ha descrito anteriormente usando el programa informático de comparación de secuencias ALIGN-2. Sin embargo, el % de identidad en la secuencia de ácido nucleico también puede determinarse usando el programa de comparación de secuencias NCBI-BLAST2 (Altschul *et al.*, Nucleic

Acids Res., 25: 3389-3402 (1997)). El programa de comparación de secuencias NCBI-BLAST2 puede descargarse de <http://www.ncbi.nlm.nih.gov>. NCBI-BLAST2 usa varios parámetros de búsqueda, donde todos estos parámetros de búsqueda se ajustan a valores por defecto, por ejemplo, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0,01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 y scoring matrix = BLOSUM62.

En situaciones en las que se emplea NCBI-BLAST2 para comparaciones de secuencias, el % de identidad en la secuencia de ácido nucleico de una secuencia de ácido nucleico dada C respecto a, con o contra una secuencia de ácido nucleico dada D (que como alternativa puede expresarse como una secuencia de ácido nucleico dada C que tiene o comprende cierto % de identidad en la secuencia de ácido nucleico respecto a, con o contra una secuencia de ácido nucleico dada D) se calcula de la siguiente manera:

$$100 \text{ veces la fracción } W/Y$$

donde W es el número de nucleótidos valorado como emparejamientos idénticos por el programa de alineamiento de secuencias NCBI-BLAST2 en el alineamiento que hace ese programa de C y D, y donde Z es el número total de nucleótidos en D. Se entenderá que cuando la longitud de la secuencia de ácido nucleico C no es igual a la longitud de la secuencia de ácido nucleico D, el % de identidad en la secuencia de ácido nucleico de C con respecto a D no será igual al % de identidad en la secuencia de ácido nucleico de D con respecto a C.

Además, también pueden generarse valores de % de identidad en la secuencia de ácido nucleico usando el programa informático WU-BLAST-2 (Altschul *et al.*, Methods in Enzymology, 266: 460-480 (1996)). La mayoría de los parámetros de búsqueda de WU-BLAST-2 se ajustan a los valores por defecto. Aquellos no ajustados a valores por defecto, es decir, los parámetros ajustables, se ajustan con los siguientes valores: overlap span = 1, overlap fraction = 0,125, word threshold (T) = 11 y scoring matrix = BLOSUM62. Para los fines de este documento, un valor de % de identidad en la secuencia de ácido nucleico se determina dividiendo (a) el número de nucleótidos que se emparejan de forma idéntica entre la secuencia de ácido nucleico de la molécula de ácido nucleico que codifica un polipéptido PRO de interés que tiene una secuencia obtenida del ácido nucleico que codifica el polipéptido PRO de secuencia nativa y la molécula de ácido nucleico de comparación de interés (es decir, la secuencia con la que la molécula de ácido nucleico que codifica un polipéptido PRO de interés se está comparando, que puede ser un polinucleótido PRO variante) según se determinó mediante WU-BLAST.2 por (b) el número total de nucleótidos de la molécula de ácido nucleico que codifica un polipéptido PRO de interés. Por ejemplo, en la afirmación “una molécula de ácido nucleico aislada que comprende una secuencia de ácido nucleico A que tiene o teniendo al menos el 80% de identidad en la secuencia de ácido nucleico con respecto a la secuencia de ácido nucleico B”, la secuencia de ácido nucleico A es la molécula de ácido nucleico de comparación de interés y la secuencia de ácido nucleico B es la secuencia de ácido nucleico de la molécula de ácido nucleico que codifica un polipéptido PRO de interés.

En otras realizaciones, polinucleótidos PRO variantes son moléculas de ácido nucleico que codifican un polipéptido PRO activo y que son capaces de hibridar, preferentemente en condiciones de hibridación y de lavado astringentes, con secuencias de nucleótidos que codifican el polipéptido PRO de longitud completa mostrado en la Figura 2 (SEC ID N°: 27) los polipéptidos PRO variantes pueden ser aquellos que son codificados por un polinucleótido PRO variante.

El término “positivos”, en el contexto de las comparaciones de identidad de la secuencia de aminoácidos realizadas como se ha descrito anteriormente, incluye restos de aminoácidos en las secuencias comparadas que no solamente son idénticos, sino también aquellos que tienen propiedades similares. Restos de aminoácidos que marcan un valor positivo con respecto a un resto de aminoácido de interés son aquellos que son idénticos al resto de aminoácido de interés o son una sustitución preferida (como se define en la Tabla 3 posteriormente) del resto de aminoácido de interés.

Para los fines de este documento, el valor de % de positivos de una secuencia de aminoácidos dada A respecto a, con o contra una secuencia de aminoácidos dada B (que, como alternativa, puede expresarse como una secuencia de aminoácidos dada A que tiene o comprende cierto % de positivos respecto a, con o contra una secuencia de aminoácidos dada B) se calcula de la siguiente manera:

$$100 \text{ veces la fracción } X/Y$$

donde X es el número de restos de aminoácidos que marcan un valor positivo como se ha definido anteriormente mediante el programa de alineamiento de secuencias ALIGN-2 en el alineamiento que hace ese programa de A y B, y donde Y es el número total de restos de aminoácidos en B. Se entenderá que cuando la longitud de la secuencia de aminoácidos A no es igual a la longitud de la secuencia de aminoácidos B, el % de positivos de A con respecto a B no será igual al % de positivos de B con respecto a A.

“Aislado”, cuando se usa para describir los diversos polipéptidos descritos en este documento, significa un polipéptido que se ha identificado y separado y/o recuperado de un componente de su entorno natural. Preferentemente, el polipéptido aislado está libre de asociación con todos los componentes con los que se asocia de forma natural. Los componentes contaminantes de su entorno natural son materiales que habitualmente interferirían en usos de diagnóstico o terapéuticos para el polipéptido, y pueden incluir enzimas, hormonas y otros solutos proteínicos o no proteínicos. En realizaciones preferidas, el polipéptido se purificará (1) a un grado suficiente para obtener al menos 15 restos de secuencia de aminoácidos N-terminal o interna mediante el uso de secuenciador con taza giratoria, o

(2) a homogeneidad mediante SDS-PAGE en condiciones no reductoras o reductoras usando azul de Coomassie o, preferentemente, tinción de plata. El polipéptido aislado incluye polipéptido *in situ* dentro de células recombinantes, puesto que al menos un componente del entorno natural de PRO no estará presente. Habitualmente, sin embargo, el polipéptido aislado se preparará mediante al menos una etapa de purificación.

Una molécula de ácido nucleico “aislada” que codifica un polipéptido PRO o un ácido nucleico “aislado” que codifica un anticuerpo anti-PRO, es una molécula de ácido nucleico que se identifica y se separa de al menos una molécula de ácido nucleico contaminante con la que habitualmente está asociada en la fuente natural del ácido nucleico que codifica PRO o el ácido nucleico que codifica anti-PRO. Preferentemente, el ácido nucleico aislado está libre de asociación con todos los componentes con los que se asocia de forma natural. Una molécula de ácido nucleico que codifica PRO aislada o una molécula de ácido nucleico que codifica anti-PRO es diferente a la forma o emplazamiento en el que se encuentra en la naturaleza, las moléculas de ácido nucleico aisladas por lo tanto se distinguen de la molécula de ácido nucleico que codifica PRO o de la molécula de ácido nucleico que codifica anti-PRO como existe en células naturales. Sin embargo, una molécula de ácido nucleico aislada que codifica un polipéptido PRO o un anticuerpo anti-PRO incluye moléculas de ácido nucleico que codifican PRO y moléculas de ácido nucleico que codifican anti-PRO contenidas en células que habitualmente expresan polipéptidos PRO o expresan anticuerpos anti-PRO donde, por ejemplo, la molécula de ácido nucleico está en una ubicación cromosómica diferente de la de las células naturales.

La expresión “secuencias de control” se refiere a secuencias de ADN necesarias para la expresión de una secuencia codificante unida de forma operativa en un organismo huésped particular. Las secuencias de control que son adecuadas para procariotas, por ejemplo, incluyen un promotor, opcionalmente una secuencia operadora, y un punto de unión al ribosoma. Se sabe que las células eucariotas utilizan promotores, señales de poliadenilación y potenciadores.

Un ácido nucleico está “unido de forma operativa” cuando está situado en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el ADN para una presecuencia o líder de secreción está unido de forma operativa a ADN para un polipéptido si se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o potenciador está unido de forma operativa a una secuencia codificante si afecta a la transcripción de la secuencia; o un punto de unión al ribosoma está unido de forma operativa a una secuencia codificante si está situado para facilitar la traducción. Generalmente, “unido de forma operativa” significa que las secuencias de ADN que se están uniendo son contiguas, y, en el caso de un líder de secreción, contiguas y en fase de lectura. Sin embargo, no es necesario que los potenciadores sean contiguos. La unión se consigue mediante ligamiento en puntos de restricción convenientes. Si dichos puntos no existen, se usan los adaptadores o enlazadores de oligonucleótidos sintéticos según la práctica convencional.

El término “anticuerpo” se usa en el sentido más amplio y abarca específicamente, por ejemplo anticuerpos monoclonales anti-PRO únicos (incluyendo antagonistas, y anticuerpos de neutralización), composiciones de anticuerpos anti-PRO con especificidad poliepitópica, anticuerpos anti-PRO de cadena sencilla y fragmentos de anticuerpos anti-PRO (véase posteriormente). La expresión “anticuerpo monoclonal” como se usa en este documento se refiere a un anticuerpo obtenido de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que comprende la población son idénticos excepto por posibles mutaciones de origen natural que pueden estar presentes en cantidades menores.

La “astringencia” de reacciones de hibridación puede determinarla fácilmente un experto en la materia, y generalmente es un cálculo empírico que depende de la longitud de la sonda, temperatura de lavado y concentración de sal. En general, sondas más largas requieren temperaturas más altas para la hibridación apropiada, mientras que sondas más cortas requieren temperaturas más bajas. La hibridación generalmente depende de la capacidad del ADN desnaturizado para hibridar de nuevo cuando están presentes cadenas complementarias en un entorno por debajo de su temperatura de fusión. Cuanto mayor sea el grado homología deseada entre la sonda y la secuencia hibridable, mayor será la temperatura relativa que puede usarse. Como resultado, se deduce que temperaturas relativas más altas tenderían a hacer a las condiciones de la reacción más astringentes, mientras que las temperaturas más bajas lo harían en menor medida. Para detalles adicionales y explicación de la astringencia de reacciones de hibridación, véase Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

“Condiciones astringentes” o “condiciones del alta astringencia”, como se definen en este documento, pueden identificarse como aquellas que: (1) emplean baja fuerza iónica y alta temperatura de lavado, por ejemplo cloruro sódico 0,015 M/citrato sódico 0,0015 M/dodecil sulfato sódico al 0,1% a 50°C; (2) emplean durante la hibridación un agente desnaturizante, tal como formamida, por ejemplo, formamida al 50% (v/v) con albúmina de suero bovino al 0,1%/Ficoll al 0,1%/polivinilpirrolidona al 0,1%/tampón fosfato sódico 50 mM a pH 6,5 con cloruro sódico 750 mM, citrato sódico 75 mM a 42°C; o (3) emplean formamida al 50%, 5 x SSC (NaCl 0,75 M, citrato sódico 0,075 M), fosfato sódico 50 mM (pH 6,8), pirofosfato sódico al 0,1%, 5 x Solución de Denhardt, ADN de esperma de salmón sonificado (50 µg/ml), SDS al 0,1%, y sulfato de dextrano al 10% a 42°C, con lavados a 42°C en 0,2 x SSC (cloruro sódico/citrato sódico) y formamida al 50% a 55°C, seguido de un lavado a alta astringencia que consiste en 0,1 x SSC que contiene EDTA a 55°C.

“Condiciones moderadamente astringentes” pueden identificarse como se ha descrito por Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Nueva York: Cold Spring Harbor Press, 1989, e incluyen el uso de solución de lavado y condiciones de hibridación (por ejemplo, temperatura, fuerza iónica y % de SDS) menos astringentes que las

descritas anteriormente. Un ejemplo de condiciones moderadamente astringentes es la incubación durante una noche a 37°C en una solución que comprende: formamida al 20%, 5 x SSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7,6), 5 x Solución de Denhardt, sulfato de dextrano al 10% y 20 mg/ml de ADN de esperma de salmón cortado desnaturalizado, seguido por el lavado de los filtros en 1 x SSC a aproximadamente 35°C-50°C. El experto en la materia reconocerá cómo ajustar la temperatura, fuerza iónica, etc., según sea necesario para acomodar factores tales como longitud de la sonda y similares.

La expresión “marcado con epítipo” cuando se usa en este documento se refiere a un polipéptido quimérico que comprende un polipéptido PRO fusionado a un “polipéptido marca”. El polipéptido marca tiene suficientes restos para proporcionar un epítipo contra el cual puede prepararse un anticuerpo, aunque es lo bastante corto para no interferir en la actividad del polipéptido al que está fusionado. El polipéptido marca preferentemente también es bastante único de modo que el anticuerpo no reacciona sustancialmente de forma cruzada con otros epítipos. Los polipéptidos marca adecuados tienen generalmente al menos seis restos de aminoácidos y habitualmente entre aproximadamente 8 y 50 restos de aminoácidos (preferentemente, entre aproximadamente 10 y 20 restos de aminoácidos).

“Activo” o “actividad” para los fines de este documento se refiere a forma o formas de polipéptidos PRO que conservan una actividad/propiedad biológica y/o inmunológica de un polipéptido PRO nativo o de origen natural, donde actividad “biológica” se refiere a una función (inhibidora o estimuladora) causada por un polipéptido PRO nativo o de origen natural diferente de la capacidad de inducir la producción de un anticuerpo contra un epítipo antigénico poseído por un polipéptido PRO nativo o de origen natural y una actividad “inmunológica” se refiere a la capacidad de inducir la producción de un anticuerpo contra un epítipo antigénico poseído por un polipéptido PRO nativo o de origen natural.

“Actividad biológica” en el contexto de un anticuerpo u otra molécula antagonista que puede identificarse mediante los ensayos de cribado descritos en este documento (por ejemplo, una molécula pequeña orgánica o inorgánica, péptido etc.) se usa para referirse a la capacidad de dichas moléculas para unirse o complejarse con los polipéptidos codificados por los genes amplificados identificados en este documento, o interferir de otra forma en la interacción de los polipéptidos codificados con otras proteínas celulares o interferir de otra manera en la transcripción o traducción de un polipéptido PRO. Una actividad biológica preferida es la inhibición del crecimiento de una célula tumoral diana. Otra actividad biológica preferida es la actividad citotóxica que da como resultado la muerte de la célula tumoral diana.

La expresión “actividad biológica” en el contexto de un polipéptido PRO significa la capacidad de un polipéptido PRO de inducir crecimiento de células neoplásicas o crecimiento celular incontrolado.

La frase “actividad inmunológica” significa reactividad inmunológica cruzada con al menos un epítipo de un polipéptido PRO.

“Reactividad inmunológica cruzada” como se usa en este documento significa que el polipéptido candidato es capaz de inhibir de forma competitiva la actividad biológica cualitativa de un polipéptido PRO que tiene esta actividad con antisueros policlonales generados contra el polipéptido PRO activo conocido. Dichos antisueros se preparan de manera convencional inyectando a cabras o conejos, por ejemplo, por vía subcutánea con el análogo activo conocido en adyuvante completo de Freund, seguido de inyección de refuerzo intraperitoneal o subcutánea en adyuvante incompleto de Freund. La reactividad inmunológica cruzada preferentemente es “específica”, lo que significa que la afinidad de unión de la molécula inmunológicamente reactiva de forma cruzada (por ejemplo, anticuerpo) identificada, con el polipéptido PRO correspondiente es significativamente más alta (preferentemente de al menos aproximadamente 2 veces, más preferentemente al menos aproximadamente 4 veces, aún más preferentemente al menos aproximadamente 8 veces, lo más preferentemente de al menos aproximadamente 10 veces más alta) que la afinidad de unión de esa molécula con cualquier otro polipéptido nativo conocido.

El término “antagonista” se usa en el sentido más amplio, e incluye cualquier molécula que inhibe, bloquea o neutraliza parcial o completamente una actividad biológica de un polipéptido PRO nativo descrito en este documento o la transcripción o traducción del mismo. Las moléculas antagonistas adecuadas incluyen específicamente anticuerpos o fragmentos de anticuerpo antagonistas, fragmentos, péptidos, pequeñas moléculas orgánicas, ácidos nucleicos antisentido, etc., se incluyen procedimientos para identificar antagonistas de un polipéptido PRO con una molécula antagonista candidata y medir un cambio detectable en una o más actividades biológicas asociadas normalmente con el polipéptido PRO.

Una “molécula pequeña” se define en este documento como la que tiene un peso molecular por debajo de aproximadamente 500 Daltons.

“Anticuerpos” (Ab) e “inmunoglobulinas” (Ig) son glucoproteínas que tienen las mismas características estructurales. Aunque los anticuerpos muestran especificidad de unión a un antígeno específico, las inmunoglobulinas incluyen anticuerpos y otras moléculas similares a anticuerpos que carecen de especificidad antigénica. Los polipéptidos de este último tipo son producidos, por ejemplo, a niveles bajos por el sistema linfático y a niveles más altos por los mielomas. El término “anticuerpo” se usa en el sentido más amplio y abarca específicamente, sin limitación, anticuerpos monoclonales intactos, anticuerpos policlonales, anticuerpos multiespecíficos (por ejemplo, anticuerpos biespecíficos) formados a partir de al menos dos anticuerpos intactos, y fragmentos de anticuerpo siempre que muestren la actividad biológica deseada.

“Anticuerpos nativos” e “inmunoglobulinas nativas” son habitualmente glucoproteínas heterotetraméricas de aproximadamente 150.000 daltons, compuestas por dos cadenas ligeras (L) idénticas y dos cadenas pesadas (H) idénticas. Cada cadena ligera está unida a una cadena pesada mediante un puente disulfuro covalente, mientras que el número de enlaces disulfuro varía entre las cadenas pesadas de diferentes isotipos de inmunoglobulina. Cada cadena pesada y ligera también tiene puentes disulfuro intracatenarios separados regularmente. Cada cadena pesada tiene en un extremo un dominio variable (V_M) seguido por varios dominios constantes. Cada cadena ligera tiene un dominio variable en un extremo (V_L) y un dominio constante en su otro extremo; el dominio constante de la cadena ligera se alinea con el primer dominio constante de la cadena pesada, y el dominio variable de cadena ligera se alinea con el dominio variable de la cadena pesada. Se cree que restos de aminoácidos particulares forman un interfaz entre los dominios variable de cadena ligera y pesada.

El término “variable” se refiere al hecho de que algunas partes de los dominios variables difieren en gran medida en la secuencia entre anticuerpos y se usan en la unión y especificidad de cada anticuerpo particular por su antígeno particular. Sin embargo, la variabilidad no se distribuye uniformemente por todos los dominios variables de anticuerpos. Está concentrada en tres segmentos llamados regiones determinantes de la complementariedad (CDR) o regiones hipervariables en los dominios variables de cadena ligera y cadena pesada. Las partes más altamente conservadas de dominios variables se denominan las regiones marco (FR). Los dominios variables de cadenas pesadas y ligeras nativas comprenden cada uno cuatro regiones FR, adoptando en gran medida una configuración de lámina β , conectada mediante tres CDR, que forma bucles que conectan, y en algunos casos forman parte de, la estructura de la lámina β . Las CDR en cada cadena se mantienen juntas muy próximas mediante las regiones FR y, con las CDR de la otra cadena, contribuyen a la formación del punto de unión al antígeno de anticuerpos (véase Kabat *et al.*, NIH Publ. No. 91-3242, Vol. 1, páginas 647-669 (1991)). Los dominios constantes no están implicados directamente en la unión de un anticuerpo a un antígeno, sino que muestran varias funciones efectoras, tales como participación del anticuerpo en toxicidad celular dependiente de anticuerpos.

La expresión “región hipervariable” cuando se usa en este documento se refiere a los restos de aminoácidos de un anticuerpo que son responsables de la unión al antígeno. La región hipervariable comprende restos de aminoácidos de una “región determinante de la complementariedad” o “CDR” (es decir, restos 24-34 (L1), 50-56 (L2) y 89-97 (L3) en el dominio variable de la cadena ligera y 31-35 (H1), 50-65 (H2) y 95-102 (H3) en el dominio variable de la cadena pesada; Kabat *et al.*, Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institute of Health, Bethesda, MD. [1991]) y/o aquellos restos de un “bucle hipervariable” (es decir, restos 26-32 (L1), 50-52 (L2) y 91-96 (L3) en el dominio variable de la cadena ligera y 26-32 (H1), 53-55 (H2) y 96-101 (H3) en el dominio variable de la cadena pesada; Clothia y Lesk, J. Mol. Biol., 196: 901-917 [1987]). Restos “Marco” o “FR” son aquellos restos de dominio variable diferentes de los restos de región hipervariable como se definen en este documento.

“Fragmentos de anticuerpo” comprenden una parte de un anticuerpo intacto, preferentemente la región de unión al antígeno o variable del anticuerpo intacto. Los ejemplos de fragmentos de anticuerpo incluyen fragmentos Fab, Fab', F(ab')₂ y Fv; diacuerpos; anticuerpos lineales (Zapata *et al.*, Protein Eng., 8(10): 1057-1062); moléculas de anticuerpo de cadena sencilla; y anticuerpos multiespecíficos formados a partir de fragmentos de anticuerpo.

La digestión con papaína de anticuerpos produce dos fragmentos de unión al antígeno idénticos, llamados fragmentos “Fab”, cada uno con un único punto de unión al antígeno, y un fragmento “Fc” residual, cuyo nombre refleja su capacidad para cristalizar fácilmente. El tratamiento con pepsina produce un fragmento F(ab')₂ que tiene dos puntos de combinación con el antígeno y sigue siendo capaz de reticular el antígeno.

“Fv” es el mínimo fragmento de anticuerpo que contiene un punto de reconocimiento y unión al antígeno completo. Esta región consiste en un dímero de un dominio variable de cadena pesada y ligera en asociación fuerte no covalente. Es en esta configuración en la que las tres CDR de cada dominio variable interactúan para definir un punto de unión al antígeno en la superficie del dímero V_H - V_L . Colectivamente, las seis CDR otorgan al anticuerpo especificidad de unión al antígeno. Sin embargo, incluso un único dominio variable (o la mitad de un Fv que comprende solamente tres CDR específicas para un antígeno) tiene la capacidad de reconocer y unirse al antígeno, aunque con una menor afinidad que el punto de unión completo.

El fragmento Fab también contiene el dominio constante de la cadena ligera y el primer dominio constante (CH1) de la cadena pesada. Los fragmentos Fab difieren de los fragmentos Fab' por la adición de unos pocos restos en el extremo carboxilo del dominio CH de la cadena pesada incluyendo una o más cisteínas de la región bisagra del anticuerpo. Fab'-SH es la denominación en este documento para Fab' en el que el resto o restos cisteína de los dominios constantes tienen un grupo tiol libre. Los fragmentos de anticuerpo F(ab')₂ se produjeron originalmente como pares de fragmentos Fab' que tienen cisteínas bisagra entre ellos. También se conocen otros acoplamientos químicos de fragmentos de anticuerpo.

Las “cadenas ligeras” de anticuerpos (inmunoglobulinas) de cualquier especie de vertebrado pueden asignarse a uno de dos tipos claramente distintos, llamados (κ) y (λ), en base a las secuencias de aminoácidos de sus dominios constantes.

Dependiendo de la secuencia de aminoácidos del dominio constante de sus cadenas pesadas, las inmunoglobulinas pueden asignarse a diferentes clases. Hay cinco clases principales de inmunoglobulinas: IgA, IgD, IgE, IgG e IgM,

y varias de éstas pueden dividirse adicionalmente en subclases (isotipos), por ejemplo, IgG1, IgG2, IgG3, IgG4, IgA e IgA2. Los dominios constantes de cadena pesada que corresponden a las diferentes clases de inmunoglobulinas se denominan α , δ , ϵ , γ y μ , respectivamente. Las estructuras de subunidad y las configuraciones tridimensionales de diferentes clases de inmunoglobulinas se conocen bien.

La expresión “anticuerpo monoclonal” como se usa en este documento se refiere a un anticuerpo obtenido de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que comprende la población son idénticos excepto por posibles mutaciones de origen natural que pueden estar presentes en menores cantidades. Los anticuerpos monoclonales son altamente específicos, estando dirigidos contra un único punto antigénico. Además, al contrario que las preparaciones de anticuerpo (policlonal) convencionales que habitualmente incluyen diferentes anticuerpos dirigidos contra diferentes determinantes (epítopos), cada anticuerpo monoclonal está dirigido contra un único determinante en el antígeno. Además de su especificidad, los anticuerpos monoclonales son ventajosos ya que se sintetizan mediante el cultivo de hibridoma, sin contaminar por otras inmunoglobulinas. El modificador “monoclonal” indica el carácter del anticuerpo ya que se obtiene de una población de anticuerpos sustancialmente homogénea, y no debe interpretarse como que requiere la producción del anticuerpo mediante cualquier procedimiento particular. Por ejemplo, los anticuerpos monoclonales a usar según la presente invención pueden prepararse mediante el procedimiento de hibridoma descrito en primer lugar por Kohler *et al.*, Nature, 256: 495 [1975], o pueden prepararse mediante procedimientos de ADN recombinantes (véase, por ejemplo, Patente de Estados Unidos N° 4.816.567). Los “anticuerpos monoclonales” también pueden aislarse de bibliotecas de presentación en fagos usando las técnicas descritas en Clackson *et al.*, Nature, 352: 624-628 y Marks *et al.*, J. Mol. Biol., 222: 581-597 (1991), por ejemplo.

Los anticuerpos monoclonales en este documento incluyen específicamente anticuerpos “quiméricos” (inmunoglobulinas) en los que una parte de la cadena pesada y/o ligera es idéntica u homóloga a secuencias correspondientes en anticuerpos obtenidos de una especie particular o que pertenecen a una clase o subclase de anticuerpos particular, mientras que el resto de la cadena o cadenas es idéntico u homólogo a secuencias correspondientes en anticuerpos obtenidos de otra especie o que pertenecen a otra clase o subclase de anticuerpos, así como fragmentos de dichos anticuerpos, siempre que muestren la actividad biológica deseada (Patente de Estados Unidos N° 4.816.567; Morrison *et al.*, Proc. Natl. Acad. Sci. USA, 81: 6851-6855 [1984]).

Formas “humanizadas” de anticuerpos no humanos (por ejemplo, murinos) son inmunoglobulinas quiméricas, cadenas de inmunoglobulina o fragmentos de las mismas (tales como Fv, Fab, Fab', F(ab')₂ u otras subsecuencias de unión al antígeno de anticuerpos) que contienen una secuencia mínima obtenida de inmunoglobulina no humana. En su mayor parte, los anticuerpos humanizados son inmunoglobulinas humanas (anticuerpo receptor) en las que restos de una CDR del receptor se sustituyen por restos de una CDR de una especie no humana (anticuerpo donador) tal como ratón, rata o conejo que tiene la especificidad, afinidad y capacidad deseadas. En algunos casos, restos FR de Fv de la inmunoglobulina humana se sustituyen por restos no humanos correspondientes. Además, los anticuerpos humanizados pueden comprender restos que no se encuentran en el anticuerpo receptor ni en la CDR o secuencias marco importadas. Estas modificaciones se realizan para refinar y maximizar adicionalmente el rendimiento del anticuerpo. En general, el anticuerpo humanizado comprenderá sustancialmente todos de al menos uno, y habitualmente dos, dominios variables, en las que todas o sustancialmente todas las regiones CDR corresponden a las de una inmunoglobulina no humana y todas o sustancialmente todas las regiones FR son las de una secuencia de inmunoglobulina humana. El anticuerpo humanizado de forma óptima también comprenderá al menos una parte de una región constante de inmunoglobulina (Fc), habitualmente la de una inmunoglobulina humana. Para más detalles, véase, Jones *et al.*, Nature, 321: 522-525 (1986); Reichmann *et al.*, Nature, 332: 323-329 [1988]; y Presta, Curr. Op. Struct. Biol., 2: 593-596 (1992). El anticuerpo humanizado incluye un anticuerpo PRIMATIZED™ donde la región de unión al antígeno del anticuerpo se obtiene de un anticuerpo producido inmunizando Macacos Rhesus con el antígeno de interés.

Fragmentos de anticuerpo “Fv de cadena sencilla” o “sFv” comprenden los dominios V_H y V_L de anticuerpo, donde estos dominios están presentes en una única cadena polipeptídica. Preferentemente, el polipéptido Fv comprende además un enlazador polipeptídico entre los dominios V_H y V_L que permite al sFv formar la estructura deseada para la unión al antígeno. Para una revisión de sFv véase Pluckthun en The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, Nueva York, p. 269-315 (1994).

El término “diacuerpos” se refiere a pequeños fragmentos de anticuerpo con dos puntos de unión al antígeno, fragmentos que comprenden un dominio variable de cadena pesada (V_H) conectado a un dominio variable de cadena ligera (V_L) en la misma cadena polipeptídica (V_H-V_L). Usando un enlazador que es demasiado corto para permitir el emparejamiento entre los dos dominios en la misma cadena, se fuerza a los dominios a emparejarse con los dominios complementarios de otra cadena y crear dos puntos de unión al antígeno. Los diacuerpos se describen más completamente en, por ejemplo, EP 404.097; WO 93/11161; y Hollinger *et al.*, Proc. Natl. Acad. Sci. USA, 90: 6444-6448 (1993).

Un anticuerpo “aislado” es uno que ha sido identificado y separado y/o recuperado de un componente de su entorno natural. Los componentes contaminantes de su entorno natural son materiales que interferirían en usos de diagnóstico y terapéuticos para el anticuerpo, y pueden incluir enzimas, hormonas y otros solutos proteínicos o no proteínicos. En realizaciones preferidas, el anticuerpo se purificará (1) a más del 95% en peso de anticuerpo según se determinó mediante el procedimiento de Lowry, y lo más preferentemente más del 99% en peso, (2) a un grado

suficiente para obtener al menos 15 restos de secuencia de aminoácidos N-terminal o interna mediante el uso de un secuenciador con taza giratoria, o (3) a homogeneidad mediante SDS-PAGE en condiciones reductoras o no reductoras usando azul de Coomassie o, preferentemente, tinción de plata. Un anticuerpo aislado incluye el anticuerpo *in situ* dentro de células recombinantes puesto que al menos un componente del entorno natural del anticuerpo no estará presente. Habitualmente, sin embargo, el anticuerpo aislado se preparará mediante al menos una etapa de purificación.

La palabra “marca” cuando se usa en este documento se refiere a un compuesto o composición detectable que se conjuga directa o indirectamente con el anticuerpo para generar un anticuerpo “marcado”. La marca puede ser detectable por sí misma (por ejemplo, marcas de radioisótopo o marcas fluorescentes) o, en el caso de una marca enzimática, puede catalizar la alteración química de un compuesto o composición sustrato que es detectable. Los radionúclidos que pueden servir como marcas detectables incluyen, por ejemplo, I-131, I-123, I-125, Y-90, Re-188, Re-186, At-211, Cu-67, Bi-212 y Pd-109. La marca también puede ser una entidad no detectable tal como una toxina.

Por “fase sólida” se entiende una matriz no acuosa a la que el anticuerpo de la presente invención puede adherirse. Los ejemplos de fases sólidas abarcados en este documento incluyen las formadas parcial o enteramente de vidrio (por ejemplo, vidrio de poro controlado), polisacáridos (por ejemplo, agarosa), poliacrilamidas, poliestireno, alcohol polivinílico y siliconas. En algunas realizaciones, dependiendo del contexto, la fase sólida puede comprender el pocillo de una placa de ensayo; en otras es una columna de purificación (por ejemplo, una columna de cromatografía de afinidad). Esta expresión también incluye una fase sólida discontinua de partículas individuales, tales como las descritas en la Patente de Estados Unidos N° 4.275.149.

Un “liposoma” es una pequeña vesícula compuesta por diversos tipos de lípidos, fosfolípidos y/o tensioactivo que es útil para la administración de un fármaco (tal como un polipéptido PRO o anticuerpo para éste y, opcionalmente, un agente quimioterapéutico) a un mamífero. Los componentes del liposoma se disponen habitualmente en una formación de bicapa, similar a la disposición lipídica de las membranas biológicas.

Como se usa en este documento, el término “inmunoadhesina” designa moléculas similares a anticuerpo que combinan la especificidad de unión de una proteína heteróloga (una “adhesina”) con las funciones efectoras de dominios constantes de inmunoglobulina. Estructuralmente, las inmunoadhesinas comprenden una fusión de una secuencia de aminoácidos con la especificidad de unión deseada que es diferente del punto de reconocimiento y unión al antígeno de un anticuerpo (es decir, es “heteróloga”), y una secuencia de dominio constante de inmunoglobulina. La parte de adhesina de una molécula de inmunoadhesina habitualmente es una secuencia de aminoácidos contigua que comprende al menos el punto de unión de receptores o un ligando. La secuencia de dominio constante de inmunoglobulina en la inmunoadhesina puede obtenerse de cualquier inmunoglobulina, tal como los subtipos IgG-1, IgG-2, IgG-3, o IgG-4, IgA (incluyendo IgA-1 e IgA-2), IgE, IgD o IgM.

II. Composiciones y Procedimientos

A. Polipéptidos PRO de Longitud Completa

En este documento se describen secuencias de nucleótidos que codifican polipéptidos denominadas en la presente solicitud como PRO. En particular, se describe ADNc que codifica polipéptidos, como se describe con más detalle en los ejemplos posteriormente. En la presente memoria descriptiva las proteínas codificadas por las secuencias de ácido nucleico descritas en este documento así como todos los homólogos y variantes adicionales nativos incluidos en la definición anterior de PRO se denominarán como “PRO” independientemente de su origen o modo de preparación.

Como se describe en los ejemplos posteriormente, se han depositados clones de ADNc en la ATCC. La secuencia de nucleótidos real de los clones puede determinarse fácilmente por el experto en la materia mediante secuenciación del clon depositado usando procedimientos rutinarios en el sector. Las secuencias de aminoácidos predichas pueden determinarse a partir de las secuencias de nucleótidos usando procedimientos rutinarios. Para los polipéptidos PRO y ácido nucleico codificante descritos en este documento, los Solicitantes identificaron lo que se cree que son los marcos de lectura que pueden identificarse mejor con la información de secuencia disponible en la actualidad.

B. Variantes de PRO

Además de los polipéptidos PRO de longitud completa de secuencia nativa descritos en este documento, se contempla que pueden prepararse variantes de PRO. Las variantes de PRO pueden prepararse introduciendo cambios de nucleótidos apropiados en el ADN de PRO y/o mediante síntesis del polipéptido PRO deseado. Los expertos en la materia entenderán que los cambios de aminoácidos pueden alterar procesos post-traduccionales del PRO tales como cambiar el número o posición de puntos de glucosilación o alterar las características de anclaje a la membrana.

Las variaciones en el PRO de secuencia nativa de longitud completa o en diversos dominios del PRO descrito en este documento, pueden realizarse, por ejemplo, usando cualquiera de las técnicas y directrices para mutaciones conservativas y no conservativas mostradas, por ejemplo, en la Patente de Estados Unidos N° 5.364.934. Las varia-

- ciones pueden ser una sustitución, eliminación o inserción de uno o más codones que codifican el PRO que da como resultado un cambio en la secuencia de aminoácidos del PRO en comparación con el PRO de secuencia nativa. Opcionalmente la variación es mediante sustitución de al menos un aminoácido por cualquier otro aminoácido en uno o más de los dominios del PRO. La guía para determinar que resto de aminoácido puede insertarse, sustituirse o eliminarse sin afectar de forma adversa a la actividad deseada puede encontrarse comparando la secuencia del PRO con la de moléculas de proteína homólogas conocidas y minimizando el número de cambios en la secuencia de aminoácidos realizados en regiones de alta homología. Las sustituciones de aminoácidos pueden ser el resultado de sustituir un aminoácido por otro aminoácido que tiene propiedades estructurales y/o químicas similares, tal como la sustitución de una leucina por una serina, es decir, sustituciones de aminoácidos conservativas. Las inserciones o eliminaciones pueden estar opcionalmente en el intervalo de aproximadamente 1 a 5 aminoácidos. La variación permitida puede determinarse realizando sistemáticamente inserciones, eliminaciones o substituciones de aminoácidos en la secuencia y ensayando las variantes resultantes para la actividad mostrada por la secuencia nativa de longitud completa o madura.
- En este documento se proporcionan fragmentos de polipéptido PRO. Dichos fragmentos pueden estar truncados en el extremo N o extremo C, o pueden carecer de restos internos, por ejemplo, en comparación con una proteína nativa de longitud completa. Algunos fragmentos carecen de restos de aminoácidos que no son esenciales para una actividad biológica deseada del polipéptido PRO.
- Pueden prepararse fragmentos de PRO mediante cualquiera de varias técnicas convencionales. Los fragmentos peptídicos deseados pueden sintetizarse químicamente. Un enfoque alternativo implica generar fragmentos de PRO mediante digestión enzimática, por ejemplo, tratando la proteína con una enzima que se sabe que escinde proteínas en puntos definidos mediante restos de aminoácidos particulares, o digiriendo el ADN con enzimas de restricción adecuadas y aislando el fragmento deseado. Otra técnica adecuada más implica aislar y amplificar un fragmento de ADN que codifica un fragmento polipeptídico deseado, mediante reacción en cadena de la polimerasa (PCR). Los oligonucleótidos que definen los extremos deseados del fragmento de ADN se emplean en los cebadores 5' y 3' en la PCR. Preferentemente, los fragmentos de polipéptido PRO comparten al menos una actividad biológica y/o inmunológica con el polipéptido PRO nativo.
- En realizaciones particulares, sustituciones conservativas de interés se muestran en la Tabla 3 bajo el título de sustituciones preferidas. Si dichas sustituciones dan como resultado un cambio de actividad biológica, entonces cambios más sustanciales, denominados sustituciones ejemplares en la Tabla 3, o como se describe con más detalle posteriormente en referencia a clases de aminoácidos, se introducen en los productos cribados.

(Tabla pasa a página siguiente)

ES 2 329 689 T3

TABLA 3

	Resto Original	Sustituciones ejemplares	Sustituciones preferidas
5	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
10	Asn (N)	Gln; his; lys; arg	gln
	Asp (D)	Glu	glu
	Cys (C)	Ser	ser
15	Gln (Q)	Asn	asn
	Glu (E)	Asp	asp
20	Gly (G)	Pro; ala	ala
	His (H)	Asn; gln; lys; arg	arg
	He (I)	leu; val; met; leu	
25		ala; phe;	
		norleucina	
	Leu (L)	norleucina; ile; ile	
30		val; met; ala; phe	
	Lys (K)	Arg; gln; asn	arg
	Met (M)	Leu; phe; ile	leu
35	Phe (F)	leu; val; ile; leu	
		ala; tyr	
40	Pro (P)	Ala	ala
	Ser (S)	Thr	thr
	Thr (T)	Ser	ser
45	Trp (W)	Tyr; phe	tyr
	Tyr (Y)	Trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; leu	
50		phe; ala;	
		norleucina	

55 Se consiguen modificaciones sustanciales en la función o identidad inmunológica del polipéptido seleccionando sustituciones que difieren significativamente en su efecto sobre el mantenimiento de (a) la estructura de la cadena principal del polipéptido en el área de la sustitución, por ejemplo, como una lámina o conformación helicoidal, (b) la carga o hidrofobicidad de la molécula en el punto diana, o (c) el volumen de la cadena lateral. Los restos de origen natural se dividen en grupos en base a propiedades comunes de la cadena lateral:

- 60 (1) hidrofóbicos: norleucina, met, ala, val, leu, ile;
- (2) hidrofílicos neutros: cys, ser, thr;
- 65 (3) ácidos: asp, glu;
- (4) básicos: asn, gln, his, lys, arg;

(5) restos que influyen en la orientación de la cadena: gly, pro; y

(6) aromáticos: trp, tyr, phe.

5 Las sustituciones no conservativas conllevarán cambiar un miembro de una de estas clases por uno de otra clase. Dichos restos sustituidos también pueden introducirse en los puntos de sustitución conservativa o, más preferentemente, en los restantes puntos (no conservados).

10 Las variaciones pueden realizarse usando procedimientos conocidos en el sector tales como mutagénesis mediada por oligonucleótidos (dirigida), barrido de alanina y mutagénesis por PCR. La mutagénesis dirigida [Carter *et al.*, Nucl. Acids Res., 1-3: 4331 (1986); Zoller *et al.*, Nucl. Acids Res., 10: 6481 (1987)], mutagénesis con casete [Wells *et al.*, Gene, 34: 315 (1985)], mutagénesis de selección por restricción [Wells *et al.*, Philos. Trans. R. Soc.-London SerA, 317: 415 (1986)] u otras técnicas conocidas pueden realizarse en el ADN clonado para producir el ADN de la variante de PRO.

15 El análisis de barrido de aminoácidos también puede emplearse para identificar uno o más aminoácidos a lo largo de una secuencia contigua. Entre los aminoácidos de barrido preferidos están aminoácidos neutros relativamente pequeños. Dichos aminoácidos incluyen alanina, glicina, serina y cisteína. La alanina es habitualmente un aminoácido de barrido preferido entre este grupo puesto que elimina la cadena lateral más allá del carbono beta y es menos probable que altere la conformación de la cadena principal de la variante [Cunningham y Wells, Science, 244: 1081-1085 (1989)]. La alanina también se prefiere habitualmente porque es el aminoácido más común. Además, se encuentra frecuentemente en posiciones escondidas y expuestas [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150: 1 (1976)]. Si la sustitución de alanina no produce cantidades adecuadas de variante, puede usarse un aminoácido isotérico.

25 C. Modificaciones de PRO

Las modificaciones covalentes de PRO se incluyen dentro del alcance de la presente invención. Un tipo de modificación covalente incluye hacer reaccionar restos de aminoácidos de un polipéptido PRO fijados como objetivo con un agente derivatizante orgánico que es capaz de reaccionar con cadenas laterales seleccionadas de los restos N- o C-terminales del PRO. La derivatización con agentes bifuncionales es útil, por ejemplo, para reticular PRO con una matriz o superficie de soporte insoluble en agua para su uso en un procedimiento para purificar anticuerpos anti-PRO, y viceversa. Los agentes reticulantes usados habitualmente incluyen, por ejemplo, 1,1-bis(diazoacetil)-2-feniletano, glutaraldehído, ésteres de N-hidroxisuccinimida, por ejemplo, ésteres con ácido 4-azidosalicílico, imidoésteres homobifuncionales, incluyendo ésteres de disuccinimidilo tales como 3,3'-ditiobis(succinimidilpropionato), maleimidias bifuncionales tales como bis-N-maleimido-1,8-octano y agentes tales como metil-3-[(p-azidofenil)ditio]propioimida-

40 Otras modificaciones incluyen desamidación de restos glutaminilo y asparaginilo a los correspondientes restos glutamilo y aspártico, respectivamente, hidroxilación de prolina y lisina, fosforilación de grupos hidroxilo de restos serilo o treonilo, metilación de los grupos α -amino de cadenas laterales de lisina, arginina e histidina [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, p. 79-86 (1983)], acetilación de la amina N-terminal, y amidación de cualquier grupo carboxilo C-terminal.

45 Otro tipo de modificación covalente del polipéptido PRO incluida dentro del alcance de la presente invención comprende alterar el patrón de glucosilación nativo del polipéptido. "Alterar el patrón de glucosilación nativo" se entiende para los fines de este documento que significa eliminar uno o más restos de carbohidrato descubiertos en PRO de secuencia nativa (retirando el punto de glucosilación subyacente o eliminando la glucosilación mediante medios químicos y/o enzimáticos), y/o añadiendo uno o más puntos de glucosilación que no están presentes en el PRO de secuencia nativa. Además, la frase incluye cambios cualitativos en la glucosilación de las proteínas nativas, que implica un cambio de la naturaleza y proporciones de los diversos restos de carbohidratos presentes.

50 La adición de puntos de glucosilación al polipéptido PRO puede conseguirse alterando la secuencia de aminoácidos. La alteración puede realizarse, por ejemplo, mediante la adición de, o sustitución por, uno o más restos de serina o treonina al PRO de secuencia nativa (para puntos de glucosilación enlazados a O). La secuencia de aminoácidos puede alterarse opcionalmente a través de cambios a nivel del ADN, particularmente mutando el ADN que codifica el polipéptido PRO en bases preseleccionadas de modo que se generen codones que se traducirán en los aminoácidos deseados.

60 Otro medio para aumentar el número de restos carbohidrato en el polipéptido PRO es mediante acoplamiento químico o enzimático de glucósidos al polipéptido. Dichos procedimientos se describen en el sector, por ejemplo, en WO 87/05330 publicado el 11 de septiembre de 1987, y en Aplin y Wriston, CRC Crit. Rev. Biochem., p. 259-306 (1981).

65 La retirada de restos carbohidrato presentes en el polipéptido PRO puede conseguirse química o enzimáticamente o mediante sustituciones mutacionales de codones que codifican restos de aminoácidos que sirven como dianas para glucosilación. En el sector se conocen técnicas de desglucosilación química y son descritas, por ejemplo, por Hakimuddin, *et al.*, Arch. Biochem. Biophys., 259: 52 (1987) y por Edge *et al.*, Anal. Biochem., 118: 131 (1981). La

escisión enzimática de restos carbohidrato en polipéptidos puede conseguirse mediante el uso de diversas endo- y exoglucosidasas según lo descrito por Thotakura *et al.*, Meth. Enzymol., 138: 350 (1987).

- Otro tipo de modificación covalente de PRO comprende enlazar el polipéptido PRO a uno de diversos polímeros no
 5 proteináceos, por ejemplo, polietilenglicol (PEG), polipropilenglicol, o polioxialquilenos, de la manera que se muestra en las Patentes de Estados Unidos N° 4.640.835; 4.496.689; 4.301.144; 4.670.417; 4.791.192 o 4.179.317. Los PRO de la presente invención también pueden modificarse de manera que formen una molécula quimérica que comprende PRO fusionado a otro polipéptido o secuencia de aminoácidos heteróloga.
- En una realización, dicha molécula quimérica comprende una fusión del PRO con un polipéptido marca que pro-
 10 porciona un epítipo al que puede unirse selectivamente un anticuerpo anti-marca. La marca epitópica se sitúa general- mente en el extremo amino o carboxilo del PRO. La presencia de dichas formas marcadas con epítipo del PRO puede detectarse usando un anticuerpo contra el polipéptido marca. Además, la provisión de la marca epitópica permite que el PRO se purifique fácilmente mediante purificación por afinidad usando un anticuerpo anti-marca u otro tipo de ma-
 15 triz de afinidad que se una a la marca epitópica. Diversos polipéptidos marca y sus respectivos anticuerpos se conocen bien en el sector. Los ejemplos incluyen marcas de poli-histidina (poli-His) o poli-histidina-glicina (poli-His-gly); el polipéptido marca HA de la gripe y su anticuerpo 12CA5 [Field *et al.*, Mol. Cell. Biol., 8: 2159-2165 (1988)]; la marca c-myc y los anticuerpos 8F9, 3C7, 6E10, G4, B7 y 9E10 para ésta [Evan *et al.*, Molecular and Cellular Biology, 5: 3610-3616 (1985)]; y la marca glucoproteína D de virus Herpes Simplex (gD) y su anticuerpo [Paborsky *et al.*, Protein
 20 Engineering 3(6): 547-553 (1990)]. Otros polipéptidos marca incluyen el péptido Flag (Hopp *et al.*, BioTechnology, 6: 1204-1210 (1988)]; el péptido epitópico KT3 [Martin *et al.*, Science, 255: 192-194 (1992)]; un péptido epitópico de α -tubulina [Skinner *et al.*, J. Biol. Chem., 266: 15163-15166 (1991)]; y la marca peptídica de la proteína del gen 10 de T7 [Lutz-Freyermuth *et al.*, Proc. Natl. Acad. Sci. USA, 87: 6393-6397 (1990)].
- En una realización alternativa, la molécula quimérica puede comprender una fusión del PRO con una inmunoglo-
 25 bulina o una región particular de una inmunoglobulina. Para una forma bivalente de la molécula quimérica (también denominada como una “inmunoadhesina”), dicha fusión podría ser con la región Fc de una molécula de IgG. Las fusiones de Ig preferentemente incluyen la sustitución de una forma soluble (dominio transmembrana eliminado o inactivado) de un polipéptido PRO en lugar de al menos una región variable dentro de una molécula de Ig. En una
 30 realización particularmente preferida, la fusión de inmunoglobulina incluye las regiones bisagra, CH2 y CH3, o las regiones bisagra, CH1, CH2 y CH3 de una molécula de IgG. Para la producción de fusiones de inmunoglobulina véase también la Patente de Estados Unidos N° 5.428.130 expedida el 27 de junio de 1995.

D. Preparación de Polipéptidos PRO

- La siguiente descripción se refiere principalmente a la producción de PRO cultivando células transformadas o
 35 transfectadas con un vector que contiene ácido nucleico de PRO. Por supuesto, se contempla que pueden emplearse procedimientos alternativos, que se conocen bien en el sector, para preparar PRO. Por ejemplo, la secuencia de PRO, o partes de la misma, puede producirse mediante síntesis peptídica directa usando técnicas en fase sólida [véase, por
 40 ejemplo, Stewart *et al.*, Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85: 2149-2154 (1963)]. La síntesis de proteínas *in vitro* puede realizarse usando técnicas manuales o mediante automatización. La síntesis automatizada puede conseguirse, por ejemplo, usando un sintetizador peptídico Applied Biosystems Peptide Synthesizer (Foster City, CA) usando las instrucciones del fabricante. Diversas partes del
 45 PRO pueden sintetizarse químicamente por separado y combinarse usando procedimientos químicos o enzimáticos para producir el PRO de longitud completa.

a. Aislamiento de ADN que Codifica un Polipéptido PRO

- El ADN que codifica PRO puede obtenerse de una biblioteca de ADNc preparada a partir de tejido que se cree
 50 que posee el ARNm de PRO y que lo expresa a un nivel detectable. Por consiguiente, el ADN de PRO humano puede obtenerse convenientemente de una biblioteca de ADNc preparada a partir de tejido humano, tal como se describe en los Ejemplos. El gen que codifica PRO también puede obtenerse de una biblioteca genómica o mediante síntesis de oligonucleótidos.
- Las bibliotecas pueden cribarse con sondas (tales como anticuerpos para el polipéptido PRO, u oligonucleótidos de
 55 al menos aproximadamente 20-80 bases) diseñadas para identificar el gen de interés o la proteína codificada por éste. El cribado de la biblioteca de ADNc o genómica con la sonda seleccionada puede realizarse usando procedimientos convencionales, tal como se describen en Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (Nueva York: Cold Spring Harbor Laboratory Press, 1989). Un medio alternativo para aislar el gen que codifica PRO es usar meto-
 60 dología de PCR [Sambrook *et al.*, *supra*; Dieffenbach *et al.*, PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

- Los Ejemplos posteriores describen técnicas para cribar una biblioteca de ADNc. Las secuencias de oligonucleóti-
 65 dos seleccionadas como sondas deben ser de suficiente longitud y lo suficientemente no ambiguas para que los falsos positivos se minimicen. El oligonucleótido preferentemente está marcado de modo que puede detectarse después de la hibridación con ADN en la biblioteca que se criba. Los procedimientos de marcado se conocen bien en el sector, e incluyen el uso de radiomarcas como ATP marcado con ^{32}P , biotilación o marcado con enzimas. Las condiciones de hibridación, incluyendo astringencia moderada y astringencia alta, se proporcionan en Sambrook *et al.*, *supra*.

Las secuencias identificadas en dichos procedimientos de cribado de bibliotecas pueden compararse y alinearse con otras secuencias conocidas depositadas y disponibles en bases de datos públicas tales como GenBank u otras bases de datos de secuencias privadas. La identidad en la secuencia (a nivel de aminoácidos o de nucleótidos) dentro de regiones definidas de la molécula o en toda la secuencia de longitud completa puede determinarse usando procedimientos conocidos en el sector y como se describen en este documento.

Puede obtenerse ácido nucleico que tiene secuencia codificante de proteínas cribando bibliotecas de ADNc o genómicas seleccionadas usando la secuencia de aminoácidos deducida descrita en este documento por primera vez, y, si fuera necesario, usando procedimientos de extensión con cebador convencionales como se describe en Sambrook *et al.*, *supra*, para detectar precursores e intermedios de procesamiento de ARNm que pueden no haberse transcrito de forma inversa a ADNc.

b. Selección y Transformación de Células Huésped

Las células huésped se transfectan o transforman con vectores de expresión o clonación descritos en este documento para la producción de PRO y se cultivan en medios nutrientes convencionales modificados según sea apropiado para inducir promotores, seleccionar transformantes o amplificar los genes que codifican las secuencias deseadas. Las condiciones de cultivo, tales como medios, temperatura, pH y similares, puede seleccionarlás el experto en la materia sin gran experimentación. En general, principios, protocolos y técnicas prácticas para maximizar la productividad de cultivos celulares pueden encontrarse en Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) y Sambrook *et al.*, *supra*.

Los procedimientos de transfección de células eucariotas y transformación de células procariotas son conocidos por el experto en la materia, por ejemplo, mediada por CaCl_2 , CaPO_4 , liposomas y electroporación. Dependiendo de la célula huésped usada, la transformación se realiza usando técnicas convencionales apropiadas para dichas células. El tratamiento con calcio que emplea cloruro cálcico, como se describe en Sambrook *et al.*, *supra*, o electroporación se usan generalmente para procariotas. La infección por *Agrobacterium tumefaciens* se usa para la transformación de ciertas células vegetales, según lo descrito por Shaw *et al.*, Gene, 23: 315 (1983) y WO 89/05859 publicado el 29 de junio de 1989. Para células de mamífero sin dichas paredes celulares, puede emplearse el procedimiento de precipitación con fosfato cálcico de Graham y van der Eb, Virology, 52: 456-457 (1978). Aspectos generales de transfecciones en sistemas de células huésped de mamífero se han descrito en la Patente de Estados Unidos N° 4.399.216. La transformación en levadura se realiza habitualmente según el procedimiento de Van Solingen *et al.*, J. Bact., 130: 946 (1977) y Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76: 3829(1979). Sin embargo, otros procedimientos para introducir ADN en células, tales como mediante microinyección nuclear, electroporación, fusión de protoplastos bacterianos con células intactas, o policonaciones, por ejemplo, polibreno, poliomitina, también pueden usarse. Para diversas técnicas para transformar células de mamífero, véase, Keown *et al.*, Methods in Enzymology, 185: 527-537 (1990) y Mansour *et al.*, Nature, 336: 348-352 (1988).

Células huésped adecuadas para clonar o expresar el ADN en los vectores en este documento incluyen células procariotas, de levadura o de eucariotas superiores. Los procariotas adecuados incluyen aunque sin limitación eubacterias, tales como organismos Gram-negativos o Gram-positivos, por ejemplo, Enterobacteriaceae tales como *E. coli*. Diversas cepas de *E. coli* están disponibles públicamente, tales como *E. coli* K12 cepa MM294 (ATCC 31.446); *E. coli* X1776 (ATCC 31.537); *E. coli* cepa W3110 (ATCC 27.325) y *E. coli* cepa K5 772 (ATCC 53.635). Otras células huésped procariotas adecuadas incluyen Enterobacteriaceae tales como *Escherichia*, por ejemplo, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, por ejemplo, *Salmonella typhimurium*, *Serratia*, por ejemplo, *Serratia marcescens* y *Shigella*, así como Bacilos tales como *B. subtilis* y *B. licheniformis* (por ejemplo, *B. licheniformis* 41P descrito en DD 266.710 publicado el 12 de abril de 1989), *Pseudomonas* tales como *P. aeruginosa*, y *Streptomyces*. Estos ejemplos son ilustrativos en lugar de limitantes. La cepa W3110 es un huésped o huésped parental particularmente preferido porque es una cepa huésped común para la fermentación de productos de ADN recombinante. Preferentemente, la célula huésped secreta cantidades mínimas de enzimas proteolíticas. Por ejemplo, la cepa W3110 puede modificarse para realizar una mutación genética en los genes que codifican proteínas endógenas para el huésped, con ejemplos de dichos huéspedes incluyendo *E. coli* W3110 cepa 1A2, que tiene el genotipo completo *tonA*; *E. coli* W3110 cepa 9E4, que tiene el genotipo completo *tonA ptr3*; *E. coli* W3110 cepa 27C7 (ATCC 55.244), que tiene el genotipo completo *tonA ptr3 phoA E15 (argF-lac) 169 degP ompT karf*; *E. coli* W3110 cepa 37D6, que tiene el genotipo completo *tonA ptr3 phoA E15 (argF-lac) 169 degP ompT rbs7 ilvG karf*; *E. coli* W3110 cepa 40B4, que es la cepa 37D6 con una mutación de eliminación de *degP* no resistente a kanamicina; y una cepa de *E. coli* que tiene proteasa periplásmica mutante descrita en la Patente de Estados Unidos N° 4.946.783 expedida el 7 de agosto de 1990. Como alternativa, son adecuados procedimientos de clonación *in vitro*, por ejemplo, PCR u otras reacciones de polimerasa de ácido nucleico.

Además de procariotas, microbios eucariotas tales como hongos filamentosos o levaduras son huéspedes de clonación o expresión adecuados para vectores que codifican PRO. *Saccharomyces cerevisiae* es un microorganismo huésped eucariota inferior usado habitualmente. Otros incluyen *Schizosaccharomyces pombe* (Beach y Nurse, Nature, 290: 140 [1981]; EP 139.383 publicado el 2 de mayo de 1985); Huéspedes de *Kluyveromyces* (Patente de Estados Unidos N° 4.943.529; Fleer *et al.*, Bio/Technology, 9: 968-975 (1991)) tales como, por ejemplo, *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt *et al.*, J. Bacteriol., 737 [1983]), *K. fragilis* (ATCC 12.424), *K. bulgaricus* (ATCC 16.045), *K. wickerhamii* (ATCC 24.178), *K. waltii* (ATCC 56.500), *K. drosophilae* (ATCC 36.906; Vanden Berg *et al.*, Bio/Technology, 8: 135 (1990)), *K. thermotolerans* y *K. marxianus*; *Yarrowia* (EP 402.226); *Pichia pastoris* (EP

183.070; Sreekrishna *et al.*, J. Basic Microbiol., 28: 265-278 [(1988)]; *Candida: Trichoderma reesia* (EP 244.234); *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76: 5259-5263 [1979]); *Schwanniomyces* tales como *Schwanniomyces occidentalis* (EP394.538 publicada el 31 de octubre 1990); y hongos filamentosos tales como, por ejemplo, *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 publicado el 10 de enero de 1991), y huéspedes de

5 *Aspergillus* tales como *A. nidularius* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112: 284-289 ; Tilburn *et al.*, Gene, 26: 205-221 [1983]; Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81: 1470-1474) y *A. niger* (Kelly y Hynes, EMBO J., 4: 475-479 [1985]). Las levaduras Metilotróficas son adecuadas en esta invención e incluyen, aunque sin limitación, levaduras capaces de crecer en metanol seleccionadas entre los géneros que consisten en *Hansenula*, *Candi-*

10 *da*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis* y *Rhodotorula*. Una lista de especies específicas que son ejemplares de esta clase de levaduras puede encontrarse en C. Anthony, The Biochemistry of Methylootrophs, 269 (1982).

Células huésped adecuadas para la expresión de PRO glucosilado se obtienen de organismos multicelulares. Los ejemplos de células de invertebrados incluyen células de insecto tales como *Drosophila* S2 y *Spodoptera* Sf9, así como células vegetales. Los ejemplos de líneas de células huésped de mamífero útiles incluyen células de ovario de hámster chino (CHO) y COS. Ejemplos más específicos incluyen línea CV1 de riñón de mono transformadas por SV40 (COS-7, ATCC CRL 1651); línea de riñón embrionario humano (células 293 ó 293 subclonadas para crecimiento en cultivo en suspensión, Graham *et al.*, J. Gen Virol., 36: 59 (1977)); células de ovario de hámster chino-DHFR(CHO), Urlaub y Chasin, Proc. Natl. Acad. Sci. USA, 77: 4216 (1980); células de sertoli de ratón (TM4, Mather, Biol. Reprod., 23: 243-251 (1980)); células pulmonares humanas (W138, ATCC CCL 75); células de hígado humanas (Hep G2, HB 8065); y tumor mamario de ratón (MMT 060562. ATCC CCL51). La selección de la célula huésped apropiada se considera dentro del alcance de un experto en la materia.

c. Selección y Uso de un Vector Replicable

25 El ácido nucleico (por ejemplo, ADNc o ADN genómico) que codifica PRO puede insertarse en un vector replicable para clonación (amplificación del ADN) o para expresión. Diversos vectores están disponibles públicamente. El vector puede estar, por ejemplo, en forma de un plásmido, cósmido, partícula viral o fago. La secuencia de ácido nucleico apropiada puede insertarse en el vector mediante diversos procedimientos. En general, el ADN se inserta en un punto o puntos apropiados de endonucleasa de restricción usando técnicas conocidas en el sector. Los componentes del vector

30 generalmente incluyen, aunque sin limitación, uno o más de una secuencia señal, un origen de replicación, uno o más genes marcadores, un elemento potenciador, un promotor y una secuencia de terminación de la transcripción. La construcción de vectores adecuados que contienen uno o más de estos componentes emplea técnicas de ligamiento convencionales que son conocidas por el experto en la materia.

35 El PRO puede producirse de forma recombinante no sólo directamente, sino también como polipéptido de fusión con un polipéptido heterólogo, que puede ser una secuencia señal u otro polipéptido que tiene un punto de escisión específico en el extremo N de la proteína o polipéptido maduro. En general, la secuencia señal puede ser un componente del vector, o puede ser una parte del ADN que codifica PRO que se inserta en el vector. La secuencia señal puede ser una secuencia señal procariota seleccionada, por ejemplo, entre el grupo de líderes de la fosfatasa alcalina, penicilina, lpp, o enterotoxina II termoestable. Para la secreción en levadura la secuencia señal puede ser, por ejemplo, el líder de invertasa de levadura, líder de factor alfa (incluyendo líderes de factor α de *Saccharomyces* y *Kluyveromyces*, el último descrito en la Patente de Estados Unidos N° 5.010.182), o líder de fosfatasa ácida, el líder de glucoamilasa de *C. albicans* (EP 362.179 publicada el 4 de abril de 1990), o la señal descrita en WO 90/13646 publicado el 15 de noviembre de 1990. En la expresión de células de mamífero, pueden usarse secuencias señal de mamífero para

45 dirigir la secreción de la proteína, tales como secuencias señal de polipéptidos secretados de la misma o de especies relacionadas, así como líderes de secreción virales.

Los vectores de expresión y de clonación contienen una secuencia de ácido nucleico que permite que el vector se replique en una o más células huésped seleccionadas. Dichas secuencias se conocen bien para diversas bacterias, levaduras y virus. El origen de replicación del plásmido pBR322 es adecuado para la mayoría de las bacterias Gram-negativas, el origen del plásmido 2μ es adecuado para levaduras, y diversos orígenes virales (SV40, polioma, adeno-virus, VSV o BPV) son útiles para vectores de clonación en células de mamífero.

Los vectores de expresión y clonación contendrán habitualmente un gen de selección, también denominado marcador seleccionable. Los genes de selección habituales codifican proteínas que (a) otorgan resistencia a antibióticos u otras toxinas, por ejemplo, ampicilina, neomicina, metotrexato o tetraciclina, (b) complementan deficiencias auxotróficas, o (c) suministran nutrientes críticos no disponibles de medios complejos, por ejemplo, el gen que codifica D-alanina racemasa para Bacilos.

60 Un ejemplo de marcadores seleccionables adecuados para células de mamífero son los que permiten la identificación de células competentes para captar el ácido nucleico que codifica PRO, tales como DHFR o timidina quinasa. Una célula huésped apropiada cuando se emplea DHFR de tipo salvaje es la línea celular de CHO deficiente en actividad de DHFR, preparada y propagada según se los descrito por Urlaub *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4216 (1980). Un gen de selección adecuado para uso en levadura es el gen *trpI* presente en el plásmido de levadura YRp7 [Stinchcomb *et al.*, Nature, 282: 39 (1979); Kingsman *et al.*, Gene, 7: 114 (1979); Tschemper *et al.*, Gene, 10: 157 (1980)]. El gen *trpI* proporciona un marcador de selección para una cepa mutante de levadura que carece de la capacidad de crecer en triptófano, por ejemplo, ATCC No. 44076 o PEP4-I (Jones, Genetics, 85: 12 (1977)).

Los vectores de expresión y clonación habitualmente contienen un promotor unido de forma operativa a la secuencia de ácido nucleico que codifica PRO para dirigir la síntesis de ARNm. Los promotores reconocidos por diversas células huésped potenciales se conocen bien. Los promotores adecuados para uso con huéspedes procariotas incluyen los sistemas promotores de β -lactamasa y lactosa [Chang *et al.*, Nature, 275: 615 (1978); Goeddel *et al.*, Nature, 281: 544 (1979)], fosfatasa alcalina, un sistema promotor de triptófano (trp) [Goeddel, Nucleic Acids Res., 8: 4057 (1980); EP 36.776] y promotores híbridos tales como el promotor tac (deBoer *et al.*, Proc. Natl. Acad. Sci. USA, 80: 21-25 (1983)). Los promotores para uso en sistemas bacterianos también contendrán una secuencia Shine-Dalgarno (S.D.) unida de forma operativa al ADN que codifica PRO.

Los ejemplos de secuencias promotoras adecuadas para uso con huéspedes de levadura incluyen los promotores para 3-fosfoglicerato quinasa (Hitzeman *et al.*, J. Biol. Chem., 255: 2073 (1980)) u otras enzimas glucolíticas [Hess *et al.*, J. Adv. Enzyme Reg., 7: 149 (1968); Holland, Biochemistry, 17: 4900 (1978)], tales como enolasa, gliceraldehído-3-fosfato deshidrogenasa, hexoquinasa, piruvato descarboxilasa, fosfofructoquinasa, glucosa-6-fosfato isomerasa, 3-fosfoglicerato mutasa, piruvato quinasa, triosafosfato isomerasa, fosfoglucosa isomerasa y glucoquinasa.

Otros promotores de levadura, que son promotores inducibles que tienen la ventaja adicional de transcripción controlada por condiciones de crecimiento, son las regiones promotoras para alcohol deshidrogenasa 2, isocitocromo C, fosfatasa ácida, enzimas degradativas asociadas con el metabolismo del nitrógeno, metalotioneína, gliceraldehído-3-fosfato deshidrogenasa, y enzimas responsables de la utilización de maltosa y galactosa. Los vectores y promotores adecuados para uso en la expresión en levadura se describen adicionalmente en EP 73.657.

La transcripción de PRO a partir de vectores en células huésped de mamífero está controlada, por ejemplo, por promotores obtenidos de los genomas de virus tales como virus de polioma, virus de la viruela aviar (UK 2.211.504 publicado el 5 de julio de 1989), adenovirus (tal como Adenovirus 2), virus del papiloma bovino, virus de sarcoma aviar, citomegalovirus, un retrovirus, virus de la hepatitis B y Virus de Simio 40 (SV40), a partir de promotores de mamífero heterólogos, por ejemplo, el promotor de actina o un promotor de inmunoglobulina, y a partir de promotores de choque térmico, siempre que dichos promotores sean compatibles con los sistemas de la célula huésped.

La transcripción de un ADN que codifica el PRO por eucariotas superiores puede aumentar insertando una secuencia potenciadora en el vector. Los potenciadores son elementos que actúan en posición cis del ADN, habitualmente de aproximadamente 10 a 300 pb, que actúan sobre un promotor para aumentar su transcripción. Actualmente se conocen muchas secuencias potenciadoras de genes de mamífero (globina, elastasa, albúmina, α -fetoproteína e insulina). Habitualmente, sin embargo, se usará un potenciador de un virus de célula eucariota. Los ejemplos incluyen el potenciador de SV40 en el lado tardío del origen de replicación (pb 100-270), el potenciador del promotor temprano de citomegalovirus, el potenciador de polioma en el lado tardío del origen de replicación y potenciadores de adenovirus. El potenciador puede cortarse y empalmarse en el vector en una posición 5' o 3' con respecto a la secuencia codificante de PRO, pero preferentemente se sitúa en un punto 5' respecto al promotor.

Los vectores de expresión usados en células huésped eucariotas (células de levadura, hongos, insecto, vegetales, animales, humanas, o nucleadas de otros organismos multicelulares) también contendrán las secuencias necesarias para la terminación de la transcripción y para estabilizar el ARNm. Dichas secuencias están disponibles habitualmente de las regiones no traducidas 5' y, ocasionalmente 3', de ADN o ADNc eucariotas o virales. Estas regiones contienen segmentos de nucleótidos transcritos como fragmentos poliadenilados en la parte no traducida del ARNm que codifica PRO.

Otros procedimientos, vectores, y células huésped más, adecuados para la adaptación a la síntesis de PRO en cultivo de células de vertebrados recombinantes se describen en Gething *et al.*, Nature, 293: 620-625 (1981); Mantei *et al.*, Nature, 281: 40-46 (1979); EP 117.060; y EP 117.058.

d. Detección de la Amplificación/Expresión Génica

La amplificación y/o expresión génica puede medirse en una muestra directamente, por ejemplo, mediante transferencia de Southern, transferencia de Northern convencional para cuantificar la transcripción de ARNm [Thomas, Proc. Natl. Acad. Sci. USA, 77: 5201-5205 (1980)], transferencia en mancha "dot blotting" (análisis de ADN), o hibridación *in situ*, usando una sonda marcada apropiadamente, en base a las secuencias proporcionadas en este documento. Como alternativa, pueden emplearse anticuerpos que pueden reconocer dobles cadenas específicas, incluyendo dobles cadenas de ADN, dobles cadenas de ARN y dobles cadenas de ADN-ARN o dobles cadenas de ADN-proteína. Los anticuerpos a su vez pueden marcarse y puede realizarse el ensayo en el que la doble cadena se une a una superficie, de modo que después de la formación de la doble cadena en la superficie, la presencia de anticuerpo unido a la doble cadena puede detectarse.

La expresión génica, como alternativa, puede medirse mediante procedimientos inmunológicos, tales como tinción inmunohistoquímica de secciones de células o tejidos y ensayo de cultivo celular o fluidos corporales, para cuantificar directamente la expresión del producto génico. Los anticuerpos útiles para la tinción inmunohistoquímica y/o ensayo de fluidos de muestra pueden ser monoclonales o policlonales, y pueden prepararse en cualquier mamífero. Convenientemente, los anticuerpos pueden prepararse contra un polipéptido PRO de secuencia nativa o contra un péptido sintético en base a las secuencias de ADN proporcionadas en este documento o contra una secuencia exógena fusionada a ADN de PRO y que codifica un epítipo de anticuerpo específico.

e. *Purificación de Polipéptido*

Pueden recuperarse formas de PRO de medios de cultivo o de lisados de células huésped. Si está unido a la membrana, puede liberarse de la membrana usando una solución detergente adecuada (por ejemplo, Triton-X 100) o mediante escisión enzimática. Las células empleadas en la expresión de PRO pueden romperse mediante diversos medios físicos o químicos, tales como ciclo de congelación-descongelación, sonicación, ruptura mecánica o agentes de lisis celular.

Puede desearse purificar PRO a partir de proteínas o polipéptidos de células recombinantes. Los siguientes procedimientos son ejemplares de procedimientos de purificación adecuados: mediante fraccionamiento en una columna de intercambio iónico; precipitación con etanol; HPLC de fase inversa; cromatografía en gel de sílice o en una resina de intercambio catiónico tal como DEAE; cromatografía de exclusión; SDS-PAGE, precipitación con sulfato amónico; filtración en gel usando, por ejemplo, Sephadex G-75; columnas de proteína A Sepharose para eliminar contaminantes tales como IgG; y columnas quelantes de metales para unir formas del PRO marcadas con epítopo. Pueden emplearse diversos procedimientos de purificación de proteínas y dichos procedimientos se conocen en el sector y se describen por ejemplo en Deutscher, *Methods in Enzymology*, 182 (1990); Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, Nueva York (1982). La etapa o etapas de purificación seleccionadas dependerán, por ejemplo, de la naturaleza del proceso de producción usado y del PRO particular producido.

E. *Amplificación de Genes que Codifican los Polipéptidos PRO en Tejidos y Líneas Celulares Tumorales*

La presente invención se basa en la identificación y caracterización de genes que se amplifican en algunas células cancerosas.

El genoma de organismos procariotas y eucariotas está sujeto a dos requisitos aparentemente contradictorios. Uno es la conservación y la propagación de ADN como la información genética en su forma original, para garantizar una herencia estable a través de múltiples generaciones. Por otro lado, las células u organismos deben ser capaces de adaptarse a cambios medioambientales duraderos. Los mecanismos adaptativos pueden incluir modificaciones cualitativas o cuantitativas del material genético. Las modificaciones cualitativas incluyen mutaciones de ADN, en las que se alteran secuencias codificantes dando como resultado una proteína estructural y/o funcionalmente diferente. La amplificación génica es una modificación cuantitativa, con lo que el número real de secuencia codificante completa, es decir, un gen, aumenta, conduciendo a un mayor número de plantillas disponibles para transcripción, un mayor número de transcritos traducibles y, finalmente, a una mayor abundancia de la proteína codificada por el gen amplificado.

El fenómeno de amplificación génica y sus mecanismos subyacentes se han investigado *in vitro* en varios sistemas de cultivo procariotas y eucariotas. El ejemplo mejor caracterizado de amplificación génica implica el cultivo de células eucariotas en medio que contiene concentraciones variables del fármaco citotóxico metotrexato (MTX). MTX es un análogo del ácido fólico e interfiere en la síntesis de ADN bloqueando la enzima dihidrofolato reductasa (DHFR). Durante la exposición inicial a bajas concentraciones de MTX la mayoría de las células (> 99,9%) morirán. Una pequeña cantidad de células sobreviven, y son capaces de crecer en concentraciones en aumento de MTX produciendo grandes cantidades de ARN y proteína de DHFR. La base de esta sobreproducción es la amplificación del único gen de DHFR. Las copias adicionales del gen se encuentran como copias extracromosómicas en forma de pequeños, cromosomas supernumerarios (dobles diminutos) o como copias cromosómicas integradas.

La amplificación génica se encuentra más habitualmente en el desarrollo de resistencia a fármacos citotóxicos (antibióticos para bacterias y agentes quimioterapéuticos para células eucariotas) y transformación neoplásica. La transformación de una célula eucariota como suceso espontáneo o debido a agresión viral o química/medioambiental está asociada habitualmente con cambios en el material genético de esa célula. Uno de los cambios genéticos más comunes observados en tumores humanos son mutaciones de la proteína p53. p53 controla la transición de células de la fase estacionaria (G1) a la replicativa (S) e impide esta transición en presencia de daño al ADN. En otras palabras, una de las principales consecuencias de deshabilitar mutaciones de p53 es la acumulación y propagación de daño al ADN, es decir, cambios genéticos. Los tipos comunes de cambios genéticos en células neoplásicas son, además de mutaciones puntuales, amplificaciones y alteraciones estructurales groseras, tales como translocaciones.

La amplificación de secuencias de ADN puede indicar un requisito funcional específico como se ilustra en el sistema experimental de DHFR. Por lo tanto, la amplificación de algunos oncogenes en tumores apunta hacia un papel causante de estos genes en el proceso de transformación maligna y mantenimiento del fenotipo transformado. Esta hipótesis ha cobrado fuerza en estudios recientes. Por ejemplo, se descubrió que la proteína *bcl-2* se amplificaba en algunos tipos de linfoma no Hodgkin. Esta proteína inhibe la apoptosis y conduce a la progresiva acumulación de células neoplásicas. Se ha descubierto que los miembros de la familia génica de receptores del factor de crecimiento son amplificados en diversos tipos de cánceres lo que sugiere que la sobre-expresión de estos receptores puede hacer a las células neoplásicas menos susceptibles a cantidades limitantes de factor de crecimiento disponible. Los ejemplos incluyen la amplificación del receptor de andrógeno en cáncer de próstata recurrente durante terapia de privación de andrógenos y la amplificación del homólogo del receptor del factor de crecimiento ERB2 en cáncer de mama. Finalmente, los genes implicados en la señalización intracelular y el control de la progresión del ciclo celular pueden sufrir amplificación durante la transformación maligna. Esto se ilustra mediante la amplificación de los genes *bcl-1* y *ras* en diversos neoplasmas epiteliales y linfoides.

Estos primeros estudios ilustran la viabilidad de identificación de secuencias de ADN amplificadas en neoplasmas, puesto que este enfoque puede identificar genes importantes para la transformación maligna. El caso de ERB2 también demuestra la viabilidad desde un punto de vista terapéutico, puesto que las proteínas transformantes pueden representar nuevas y específicas dianas para terapia de tumores.

Pueden usarse varias técnicas diferentes para demostrar secuencias genómicas amplificadas. El análisis citogenético clásico de extensiones cromosómicas preparadas a partir de células cancerosas es adecuado para identificar alteraciones estructurales groseras, tales como translocaciones, eliminaciones e inversiones. Las regiones genómicas amplificadas solamente pueden visualizarse, si implican grandes regiones con grandes números de copias o están presentes como material extracromosómico. Aunque la citogenética era la primera técnica que demostró la asociación coherente de cambios cromosómicos específicos con neoplasmas particulares, ésta es inadecuada para la identificación y aislamiento de secuencias de ADN manejables. La técnica desarrollada más recientemente de hibridación genómica comparativa (CGH) ilustraba el fenómeno ampliamente extendido de amplificación genómica en neoplasmas. El ADN tumoral y normal hibridan simultáneamente en metafases de células normales y todo el genoma puede cribarse mediante análisis de imágenes para secuencias de ADN que están presentes en el tumor con una mayor frecuencia. (WO 93/18,186; Gray *et al.*, Radiation Res., 137: 275-289 (1994)). Como procedimiento de cribado, este tipo de análisis mostró gran número de amplicones recurrentes (un tramo de ADN amplificado) en diversos neoplasmas humanos. Aunque la CGH es más sensible que el análisis citogenético clásico en la identificación de tramos amplificados de ADN, no permite una rápida identificación y aislamiento de secuencias codificantes dentro del amplicón mediante técnicas de genética molecular convencionales.

Los procedimientos más sensibles para detectar amplificación génica son ensayos basados en la reacción en cadena de la polimerasa (PCR). Estos ensayos utilizan una cantidad muy pequeña de ADN tumoral como material de partida, son exquisitamente sensibles, proporcionan ADN que es susceptible a análisis adicionales, tales como secuenciación y son adecuados para análisis de rendimiento a volumen alto.

Los ensayos mencionados anteriormente no son mutuamente excluyentes, sino que frecuentemente se usan en combinación para identificar amplificaciones en neoplasmas. Aunque el análisis citogenético y la CGH representan procedimientos de cribado para supervisar todo el genoma para regiones amplificadas, los ensayos basados en PCR son más adecuados para la identificación final de secuencias codificantes, es decir, genes en regiones amplificadas.

Según la presente invención, dichos genes se han identificado mediante PCR cuantitativa (S. Gelmini *et al.*, Clin. Chem., 43: 752 [1997]), comparando ADN de diversos tumores primarios, incluyendo tumores o líneas celulares tumorales de mama, pulmón, colon, próstata, cerebro, hígado, riñón, páncreas, bazo, timo, testículos, ovario, útero, etc., con ADN reunido de donante sano. La PCR cuantitativa se realizó usando un instrumento TaqMan (ABI). Se diseñaron cebadores y sondas fluorogénicas específicas de genes en base a las secuencias codificantes de los ADN.

Las líneas celulares de carcinoma de pulmón humano incluyen A549 (SRCC768), Calu-1 (SRCC769), Calu-6 (SRCC770), H157 (SRCC771), H441 (SRCC772), H460 (SRCC773), SKMES-1 (SRCC774), SW900 (SRCC775), H522 (SRCC832) y H810 (SRCC833), todas disponibles de la ATCC. Las células tumorales primarias de pulmón humano habitualmente se obtienen de adenocarcinomas, carcinomas de células escamosas, carcinomas de células grandes, carcinomas no microcíticos, carcinomas microcíticos y carcinomas broncoalveolares, e incluyen, por ejemplo, SRCC724 (adenocarcinoma, abreviado como "AdenoCa")(LT1), SRCC725 (carcinoma de células escamosas, abreviado como "SqCCa")(LT1a), SRCC726 (adenocarcinoma)(LT2), SRCC727 (adenocarcinoma)(LT3), SRCC728 (adenocarcinoma)(LT4), SRCC729 (carcinoma de células escamosas)(LT6), SRCC730 (adeno/carcinoma de células escamosas)(LT7), SRCC731 (adenocarcinoma)(LT9), SRCC732 (carcinoma de células escamosas)(LT10), SRCC733 (carcinoma de células escamosas)(LT11), SRCC734 (adenocarcinoma)(LT12), SRCC735 (adeno/carcinoma de células escamosas)(LT13), SRCC736 (carcinoma de células escamosas)(LT15), SRCC737 (carcinoma de células escamosas)(LT16), SRCC738 (carcinoma de células escamosas)(LT17), SRCC739 (carcinoma de células escamosas)(LT18), SRCC740 (carcinoma de células escamosas)(LT19), SRCC741 (carcinoma de células pulmonares, abreviado como "LCCa")(LT21), SRCC811 (adenocarcinoma)(LT22), SRCC825 (adenocarcinoma)(LT8), SRCC886 (adenocarcinoma)(LT25), SRCC887 (carcinoma de células escamosas)(LT26), SRCC888 (adeno-carcinoma BAC)(LT27), SRCC889 (carcinoma de células escamosas)(LT28), SRCC890 (carcinoma de células escamosas)(LT29), SRCC891 (adenocarcinoma)(LT30), SRCC892 (carcinoma de células escamosas)(LT31), SRCC894 (adenocarcinoma)(LT33). También se incluyen tumores de pulmón humano denominados SRCC1125 [HF-000631], SRCC1127 [HF-000641], SRCC1129 [HF-000643], SRCC1133 [HF-000840], SRCC1135 [HF-000842], SRCC1227 [HF-001291], SRCC1229 [HF-001293], SRCC1230 [HF-001294], SRCC1231 [HF-001295], SRCC1232 [HF-001296], SRCC1233 [HF-001297], SRCC1235 [HF-001299] y SRCC1236 [HF-001300].

Las líneas celulares de cáncer de colon incluyen, por ejemplo, líneas celulares de ATCC SW480 (adenocarcinoma, SRCC776), SW620 (metástasis de ganglios linfáticos de adenocarcinoma de colon, SRCC777), Colo320 (carcinoma, SRCC778), HT29 (adenocarcinoma, SRCC779), HM7 (una variante de alta producción de mucina de la línea celular de la ATCC de adenocarcinoma de colon, SRCC780, obtenida del Dr. Robert Warren, UCSF), CaWi-Dr (adenocarcinoma, SRCC781), HCT116 (carcinoma, SRCC782), SKCO1 (adenocarcinoma, SRCC783), SW403 (adenocarcinoma, SRCC784), LS174T (carcinoma, SRCC785), Colo205 (carcinoma, SRCC828), HCT15 (carcinoma, SRCC829), HCC2998 (carcinoma, SRCC830) y KM 12 (carcinoma, SRCC831). Los tumores de colon primarios incluyen adenocarcinomas de colon llamados CT2 (SRCC742), CT3 (SRCC743), CT8 (SRCC744), CT10 (SRCC745), CT12 (SRCC746), CT14 (SRCC747), CT15 (SRCC748), CT16 (SRCC749), CT17 (SRCC750), CT1 (SRCC751),

CT4 (SRCC752), CT5 (SRCC753), CT6 (SRCC754), CT7 (SRCC755), CT9 (SRCC756), CT11 (SRCC757), CT18 (SRCC758), CT19 (adenocarcinoma, SRCC906), CT20 (adenocarcinoma, SRCC907), CT21 (adenocarcinoma, SRCC908), CT22 (adenocarcinoma, SRCC909), CT23 (adenocarcinoma, SRCC910), CT24 (adenocarcinoma, SRCC911), CT25 (adenocarcinoma, SRCC912), CT26 (adenocarcinoma, SRCC913), CT27 (adenocarcinoma, SRCC914), CT28 (adenocarcinoma, SRCC915), CT29 (adenocarcinoma, SRCC916), CT30 (adenocarcinoma, SRCC917), CT31 (adenocarcinoma, SRCC918), CT32 (adenocarcinoma, SRCC919), CT33 (adenocarcinoma, SRCC920), CT35 (adenocarcinoma, SRCC921) y CT36 (adenocarcinoma, SRCC922). También se incluyen centros tumorales de colon llamados SRCC1051 [HF-000499], SRCC1052 [HF-000539], SRCC1053 [HF-000575], SRCC1054 [HF-000698], SRCC1142 [HF-000762], SRCC1144 [HF-000789], SRCC1146 [HF-000795] y SRCC1148 [HF-000811].

Las líneas celulares de carcinoma de mama humano incluyen, por ejemplo, HBL100 (SRCC759), MB435s (SRCC760), T47D (SRCC761), MB468 (SRCC762), MB175 (SRCC763), MB361 (SRCC764), BT20 (SRCC765), MCF7 (SRCC766) y SKBR3 (SRCC767), y centros tumorales de mama humanos llamados SRCC1057 [HF-00054S]. También se incluyen tumores de mama humanos llamados SRCC1094, SRCC1095, SRCC1096, SRCC1097, SRCC1098, SRCC1099, SRCC1100, SRCC1101, y tumor de mama-met-pulmón-SN llamado SRCC893 [LT32].

Los centros tumorales de riñón humano incluyen SRCC989 [HF-000611] y SRCC1014 [HF-000613].

Los centros tumorales de testículo humanos incluyen SRCC1001 [HF-000733] y tumor marginal de testículo SRCC999 [HF-000716].

El tumor de paratiroides humano incluye SRCC1002 [HF-000831] y SRCC1003 [HF-000832].

F. Distribución Tisular

Los resultados de los ensayos de amplificación génica en este documento pueden verificarse mediante estudios adicionales, tales como, determinando la expresión del ARNm en diversos tejidos humanos.

Como se ha indicado anteriormente, la amplificación génica y/o la expresión génica en diversos tejidos puede medirse mediante transferencia de Southern convencional, transferencia de Northern para cuantificar la transcripción de ARNm (Thomas, Proc. Natl. Acad. Sci. USA, 77: 520 1-5205 [1980]), transferencia en mancha "dot blotting" (análisis de ADN), o hibridación *in situ*, usando una sonda marcada apropiadamente, en base a las secuencias proporcionadas en este documento. Como alternativa, pueden emplearse anticuerpos que pueden reconocer dobles cadenas específicas, incluyendo dobles cadenas de ADN, dobles cadenas de ARN y dobles cadenas híbridas de ADN-ARN o dobles cadenas de ADN-proteína.

La expresión génica en diversos tejidos, como alternativa, puede medirse mediante procedimientos inmunológicos, tales como tinción inmunohistoquímica de secciones de tejido y ensayo de cultivo celular o fluidos corporales, para cuantificar directamente la expresión de producto génico. Los anticuerpos útiles para tinción inmunohistoquímica y/o ensayo de fluidos de muestra pueden ser monoclonales o policlonales, y pueden prepararse en cualquier mamífero. Convenientemente, los anticuerpos pueden prepararse contra un polipéptido PRO de secuencia nativa o contra un péptido sintético en base a secuencias de ADN proporcionadas en este documento o contra secuencias exógenas fusionadas a secuencia de ADN de PRO y que codifican un epítipo de anticuerpo específico. Las técnicas generales para generar anticuerpos, y protocolos especiales para transferencia de Northern e hibridación *in situ* se proporcionan posteriormente en este documento.

G. Mapeo de Cromosomas

Si la amplificación de un gen dado es funcionalmente relevante, entonces ese gene debe amplificarse más que las regiones genómicas adyacentes que no son importantes para la supervivencia del tumor. Para ensayar esto, el gen puede mapearse para un cromosoma particular, por ejemplo, mediante análisis de radiación-híbrido. El nivel de amplificación se determina a continuación en la ubicación identificada, y en la región genómica adyacente. La amplificación selectiva o preferencial en la región genómica para la que el gen se ha mapeado es coherente con la posibilidad de que la amplificación génica observada promueva el crecimiento o la supervivencia tumoral. El mapeo de cromosomas incluye mapeo del marco y del epicentro. Para más detalles véase, por ejemplo, Stewart *et al.*, Genome Research. 7: 422-433 (1997).

H. Estudios de Unión a Anticuerpos

Los resultados del estudio de amplificación génica pueden verificarse adicionalmente mediante estudios de unión a anticuerpos, en los que se ensaya la capacidad de anticuerpos anti-PRO para inhibir la expresión de polipéptidos PRO en células tumorales (cancerosas). Los anticuerpos ejemplares incluyen anticuerpos policlonales, monoclonales, humanizados, biespecíficos y heteroconjugados, cuya preparación se describirá posteriormente en este documento.

Los estudios de unión a anticuerpos pueden realizarse en cualquier procedimiento de ensayo conocido, tal como ensayos de unión competitiva, ensayos sándwich directos e indirectos y ensayos de inmunoprecipitación. Zola, Monoclonal Antibodies: A Manual of Techniques, p. 147-158 (CRC Press, Inc., 1987).

Los ensayos de unión competitiva se basan en la capacidad de un patrón marcado para competir con el analito de la muestra de ensayo por la unión a una cantidad limitada de anticuerpo. La cantidad de proteína diana (codificada por un gen amplificado en una célula tumoral) en la muestra de ensayo es inversamente proporcional a la cantidad de patrón que queda unido a los anticuerpos. Para facilitar la determinación de la cantidad de patrón que queda unido, los anticuerpos preferentemente se insolubilizan antes o después de la competencia, de modo que el patrón y el analito que están unidos a los anticuerpos pueden separarse convenientemente del patrón y analito que permanecen sin unirse.

Los ensayos sándwich implican el uso de dos anticuerpos, cada uno capaz de unirse a una parte inmunogénica, o epítipo, diferente de la proteína a detectar. En un ensayo sándwich, el analito de la muestra de ensayo queda unido a un primer anticuerpo que se inmoviliza en un soporte sólido, y a continuación un segundo anticuerpo se une al analito, formando de este modo un complejo insoluble de tres partes. Véase, por ejemplo, la Patente de Estados Unidos N° 4.376.110. El propio segundo anticuerpo puede estar marcado con un resto detectable (ensayos sándwich directos) o puede medirse usando un anticuerpo anti-inmunoglobulina que está marcado con un resto detectable (ensayo de sándwich indirecto). Por ejemplo, un tipo de ensayo sándwich en un ensayo ELISA, en cuyo caso el resto detectable es una enzima.

Para inmunohistoquímica, la muestra tumoral puede ser fresca o congelada o puede incluirse en parafina y fijarse con un conservante tal como formalina, por ejemplo.

1. Ensayos Tumorales Basados en Células

Pueden usarse ensayos basados en células y modelos animales para tumores (por ejemplo, cánceres) para verificar los descubrimientos del ensayo de amplificación génica, y entender mejor la relación entre los genes identificados en este documento y el desarrollo y patogénesis del crecimiento de células neoplásicas. El papel de los productos génicos identificados en este documento en el desarrollo y patología de tumores o cáncer puede ensayarse usando células tumorales primarias o líneas celulares que se han identificado que amplifican los genes en este documento. Dichas células incluyen, por ejemplo, las células y líneas celulares de cáncer de mama, colon y pulmón enumeradas anteriormente.

En un enfoque diferente, células de un tipo celular que se sabe que está implicado en un tumor particular se transfectan con los ADNc en este documento, y se analiza la capacidad de estos ADNc para inducir un crecimiento excesivo. Las células adecuadas incluyen, por ejemplo, líneas celulares tumorales estables tales como, la línea celular B104-1-1 (línea celular NIH-3T3 estable transfectada con el proto-oncogén *neu*) y células NIH-3T3 transfectadas con *ras*, que pueden transfectarse con el gen deseado, y supervisarse su crecimiento tumoral. Dichas líneas celulares transfectadas pueden usarse a continuación para ensayar la capacidad de anticuerpos poli- o monoclonales o composiciones de anticuerpos para inhibir el crecimiento celular tumorigénico ejerciendo actividad citostática o citotóxica sobre el crecimiento de las células transformadas, o mediando en la citotoxicidad celular dependiente de anticuerpos (ADCC). Las células transfectadas con las secuencias codificantes de los genes identificados en este documento pueden usarse además para identificar fármacos candidatos para el tratamiento del cáncer.

Además, los cultivos primarios obtenidos de tumores en animales transgénicos (como se describe posteriormente) pueden usarse en los ensayos basados en células de este documento, aunque se prefieren líneas celulares estables. Las técnicas para obtener líneas celulares continuas de animales transgénicos se conocen bien en el sector (véase, por ejemplo, Small *et al.*, Mol. Cell. Biol., 5: 642-648 [1985]).

J. Modelos Animales

Pueden usarse diversos modelos animales bien conocidos para entender mejor el papel de los genes identificados en este documento en el desarrollo y patogénesis de tumores, y para ensayar la eficacia de agentes terapéuticos candidatos, incluyendo anticuerpos, y otros antagonistas de los polipéptidos nativos, incluyendo antagonistas de moléculas pequeñas. La naturaleza *in vivo* de dichos modelos los hace particularmente predictivos de respuestas en pacientes humanos. Los modelos animales de tumores y cánceres (por ejemplo, cáncer de mama, cáncer de colon, cáncer de próstata, cáncer de pulmón, etc.) incluyen animales no recombinantes y recombinantes (transgénicos). Los modelos animales no recombinantes incluyen, por ejemplo, modelos en roedores, por ejemplo, múridos. Dichos modelos pueden generarse introduciendo células tumorales en ratones singénicos usando técnicas convencionales, por ejemplo, inyección subcutánea, inyección en la vena caudal, implantación en el bazo, implantación intraperitoneal, implantación bajo la cápsula renal, o implantación de ortopina, por ejemplo, células de cáncer de colon implantadas en tejido colónico. (Véase, por ejemplo, Publicación PCT N°. WO 97/33551, publicada el 18 de septiembre de 1997).

Probablemente las especies animales usadas con mayor frecuencia en estudios oncológicos son ratones inmunodeficientes y, en particular, ratones desnudos. La observación de que los ratones desnudos con hipo/aplasia podían actuar de forma exitosa como huésped para xenoinjertos de tumores humanos condujo a su amplia extensión para este fin. El gen autosómico recesivo *nu* se introdujo en un muy gran número de distintas cepas oncogénicas de ratón desnudo, incluyendo, por ejemplo, ASW, A/He, AKR, BALB/c, B10-LP, C17, C3H, C57BL, C57, CBA, DBA, DDD, I/st, NC, NFR, NFS, NFS/N, NZB, NZC, NZW, P, RIII y SJL. Además, una gran variedad de otros animales con defectos inmunológicos heredados diferentes de los ratones desnudos se criaron y usaron como receptores de xenoinjertos tumorales. Para más detalles véase, por ejemplo, The Nude Mouse in Oncology Research, E. Boven y B. Winograd, eds., CRC Press, Inc., 1991.

Las células introducidas en dichos animales pueden obtenerse de líneas celulares tumorales/de cáncer conocidas, tales como, cualquiera de las líneas celulares tumorales enumeradas anteriormente, y, por ejemplo, la línea celular B104-1-1 (línea celular NIH-3T3 estable transfectada con el proto-oncogén *neu*); células NIH-3T3 transfectadas con *ras*; Caco-2 (ATCC HTB-37); una línea celular de adenocarcinoma de colon humano de grado II moderadamente bien diferenciado, HT-29 (ATCC HTB-38), o de tumores y cánceres. Las muestras de células tumorales o de cáncer pueden obtenerse de pacientes que se sometieron a cirugía, usando condiciones convencionales, que implican congelación y almacenamiento en nitrógeno líquido (Karmali *et al.*, Br. J. Cancer, 48: 689-696 [1983]).

Las células tumorales pueden introducirse en animales, tales como ratones desnudos, mediante diversos procedimientos. El espacio subcutáneo (s.c.) en ratones es muy adecuado para implantación de tumores. Los tumores pueden transplantarse s.c. como bloques sólidos, como biopsias de aguja mediante el uso de un trócar, o como suspensiones celulares. Para implantación como bloque sólido o con trócar, se introducen fragmentos de tejido tumoral de tamaño adecuado en el espacio s.c. Las suspensiones celulares se preparan de nuevo a partir de tumores primarios o líneas celulares tumorales estables, y se inyectan por vía subcutánea. Las células tumorales también pueden inyectarse como implantes subdérmicos. En esta ubicación, el inóculo se deposita entre la parte inferior del tejido conectivo dérmico y el tejido s.c. Boven y Winograd (1991), *supra*.

Pueden generarse modelos animales de cáncer de mama, por ejemplo, implantando células de neuroblastoma de rata (a partir de las cuales se aisló originalmente el oncogén *neu*), o células NIH-3T3 transformadas con *neu* en ratones desnudos, esencialmente según lo descrito por Drebin *et al.*, PNAS USA, 83: 9129-9133 (1986).

Análogamente, pueden generarse modelos animales de cáncer de colon subcultivando células de cáncer de colon en animales, por ejemplo, ratones desnudos, lo que conduce a la aparición de tumores en estos animales. Se ha descrito un modelo de trasplante ortotópico de cáncer de colon humano en ratones desnudos, por ejemplo, por Wang *et al.*, Cancer Research, 54: 4726-4728 (1994) y Too *et al.*, Cancer Research 55: 681-684 (1995). Este modelo se basa en el llamado "METAMOUSE" comercializado por AntiCancer, Inc., (San Diego, California).

Los tumores que surgen en animales pueden extirparse y cultivarse *in vitro*. Las células de los cultivos *in vitro* pueden subcultivarse a continuación en animales. Dichos tumores pueden servir como dianas para ensayos adicionales o cribado de fármacos. Como alternativa, los tumores resultantes del subcultivo pueden aislarse y el ARN de células pre-subcultivo y células aisladas después de una o más rondas de subcultivo analizarse para la expresión diferencial de genes de interés. Dichas técnicas de subcultivo pueden realizarse con cualesquiera líneas celulares tumorales o de cáncer conocidas.

Por ejemplo, Meth A, CMS4, CMS5, CMS21, y WEH1-164 son fibrosarcomas químicamente inducidos de ratones hembra BALB/c (DeLeo *et al.*, J. Exp. Med., 146: 720 [1977]), que proporcionan un sistema modelo altamente controlable para estudiar las actividades anti-tumorales de diversos agentes (Palladino *et al.*, J. Immunol., 138: 4023-4032 [1987]). En resumen, se propagan células tumorales *in vitro* en cultivo celular. Antes de la inyección en los animales, las líneas celulares se lavan y se suspenden en tampón, a una densidad celular de aproximadamente 10×10^6 a 10×10^7 células/ml. A continuación los animales se infectan por vía subcutánea con de 10 a 100 μ l de la suspensión celular, dejando de una a tres semanas para que aparezca un tumor.

Además, el carcinoma de pulmón de Lewis (3LL) de ratones, que es uno de los tumores experimentales más exhaustivamente estudiados, puede usarse como modelo tumoral de investigación. La eficacia en este modelo tumoral se correlacionaba con efectos beneficiosos en el tratamiento de pacientes humanos a los que se diagnosticó carcinoma de pulmón microcítico (SCCL). Este tumor puede introducirse en ratones normales después de la inyección de fragmentos tumorales de un ratón afectado o de células que se mantienen en cultivo (Zupi *et al.*, Br. J. Cancer, 41: supl. 4: 309 [1980]), y las pruebas indican que los tumores pueden iniciarse a partir de la inyección de incluso una única célula y que una gran proporción de células tumorales infectadas sobrevive. Para información adicional sobre este modelo tumoral véase, Zacharski, Haemostasis, 16: 300-320 [1986]).

Una manera de evaluar la eficacia de un compuesto de ensayo en un modelo animal sobre un tumor implantado es medir el tamaño del tumor antes y después del tratamiento. Tradicionalmente, el tamaño de tumores implantados se ha medido con un calibre pie de rey en dos o tres dimensiones. La medición limitada a dos dimensiones no refleja de forma precisa el tamaño del tumor, por lo tanto, habitualmente se convierte en el volumen correspondiente usando una fórmula matemática. Sin embargo, la medición del tamaño tumoral es muy imprecisa. Los efectos terapéuticos de un fármaco candidato pueden describirse mejor como retraso del crecimiento inducido por el tratamiento y retraso del crecimiento específico. Otra importante variable en la descripción del crecimiento del tumor es el tiempo de doblado del volumen tumoral. También están disponibles programas informáticos para el cálculo y descripción del crecimiento tumoral, tales como el programa descrito por Rygaard y Spang-Thomsen, Proc. 6th Int. Workshop on Immune-Deficient Animals, Wu y Sheng eds., Basel, 1989, 301. Se observa, sin embargo, que la necrosis y las respuestas inflamatorias después del tratamiento realmente pueden dar como resultado un aumento del tamaño tumoral, al menos inicialmente. Por lo tanto, estos cambios deben supervisarse cuidadosamente, mediante una combinación de un procedimiento morfométrico y análisis de citometría de flujo.

Pueden manipularse modelos animales recombinantes (transgénicos) introduciendo la parte codificante de los genes identificados en este documento en el genoma de animales de interés, usando técnicas convencionales para producir animales transgénicos. Los animales que pueden servir como diana para manipulación transgénica incluyen, sin

limitación, ratones, ratas, conejos, cobayas, ovejas, cabras, cerdos y primates no humanos, por ejemplo, babuinos, chimpancés y monos. Las técnicas conocidas en el sector para introducir un transgén en dichos animales incluyen minicroinyección pronucleica (Hoppe y Wanger, Patente de Estados Unidos N° 4.873.191); transferencia génica mediada por retrovirus a líneas germinales (por ejemplo, Van der Putten *et al.*, Proc. Natl. Acad. Sci. USA, 82: 6148-615 [1985]); dirección de genes en células madre embrionarias (Thompson *et al.*, Cell, 56: 313-321 [1989]); electroporación de embriones (Lo, Mol. Cell Biol., 3: 1803-1814 [1983]); transferencia de genes mediada por esperma (Lavitrano *et al.*, Cell, 57: 717-73 [1989]). Para una revisión, véase, por ejemplo, la Patente de Estados Unidos N° 4.736.866.

Para los fines de la presente invención, los animales transgénicos incluyen los que tienen el transgén solamente en parte de sus células (“animales mosaico”). El transgén puede integrarse como un único transgén, o en concatámeros, por ejemplo, tándems cabeza-a-cabeza o cabeza-a-cola. La introducción selectiva de un transgén en un tipo celular particular también es posible siguiendo, por ejemplo, la técnica de Lasko *et al.*, Proc. Natl. Acad. Sci. USA, 89: 6232-636 (1992).

La expresión del transgén en animales transgénicos puede supervisarse mediante técnicas convencionales. Por ejemplo, puede usarse análisis de transferencia de Southern o amplificación por PCR para verificar la integración del transgén. El nivel de expresión de ARNm puede analizarse a continuación usando técnicas tales como hibridación *in situ*, análisis de transferencia de Northern, PCR o inmunoquímica. Además se examina a los animales en busca de signos de desarrollo de tumores o cáncer.

Como alternativa, pueden construirse animales “knock out” que tienen un gen defectuoso o alterado que codifica un polipéptido PRO identificado en este documento, como resultado de recombinación homóloga entre el gen endógeno que codifica el polipéptido y ADN genómico alterado que codifica el mismo polipéptido introducido en una célula embrionaria del animal. Por ejemplo, puede usarse ADNc que codifica un polipéptido PRO para clonar ADN genómico que codifica ese polipéptido según técnicas establecidas. Una parte del ADN genómico que codifica un polipéptido PRO particular puede eliminarse o sustituirse por otro gen, tal como un gen que codifica un marcador seleccionable que puede usarse para supervisar la integración. Habitualmente, varias kilobases de ADN flanqueante inalterado (en los extremos 5' y 3') se incluyen en el vector [véase, por ejemplo, Thomas y Capecchi, Cell, 51: 503 (1987) para una descripción de vectores de recombinación homóloga. El vector se introduce en una línea de células madre embrionarias (por ejemplo, mediante electroporación) y se seleccionan células en las que el ADN introducido ha recombinado de forma homóloga con el ADN endógeno [véase, por ejemplo, Li *et al.*, Cell, 69: 915 (1992)]. Las células seleccionadas se inyectan a continuación en un blastocisto de un animal (por ejemplo, un ratón o una rata) para formar quimeras de agregación [véase, por ejemplo, Bradley, en Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E. J. Robertson, ed. (IRL, Oxford, 1987), p. 113-152]. Un embrión quimérico puede implantarse a continuación en un animal de acogida hembra pseudopreñada adecuado y el embrión puede llevarse a término para crear un animal “knock out”. La progenie que tiene el ADN recombinado de forma homóloga en sus células germinales puede identificarse mediante técnicas convencionales y usarse para criar animales en los que todas las células del animal contienen el ADN recombinado de forma homóloga. Los animales knockout pueden caracterizarse por ejemplo, por su capacidad para defenderse contra ciertas afecciones patológicas y por su desarrollo de afecciones patológicas debido a la ausencia del polipéptido PRO.

La eficacia de anticuerpos que se unen específicamente a los polipéptidos identificados en este documento y otros fármacos candidatos; puede ensayarse también en el tratamiento de tumores animales espontáneos. Una diana adecuada para dichos estudios es el carcinoma de células escamosas (SCC) oral felino. El SCC oral felino es un tumor maligno altamente invasivo que es el tumor oral más común en gatos, suponiendo más del 60% de los tumores orales descritos en esta especie. Este tumor raramente metastatiza a puntos distantes, aunque esta baja incidencia de metástasis puede ser meramente un reflejo de los cortos periodos de supervivencia para gatos con este tumor. Estos tumores habitualmente no son susceptibles a la cirugía, principalmente debido a la anatomía de la cavidad bucal felina. Actualmente, no existe un tratamiento eficaz para este tumor. Antes de enrolarse en el estudio, cada gato se somete a un examen clínico completo, biopsia, y a un barrido mediante tomografía computarizada (CT). Los gatos a los que se diagnosticó tumores de células escamosas orales sublinguales se excluyen del estudio. La lengua puede quedar paralizada como resultado de dicho tumor, e incluso aunque el tratamiento destruya el tumor, los animales pueden no ser capaces de alimentarse por sí mismos. Cada gato recibe tratamiento repetidamente, durante un largo periodo de tiempo. Se tomarán fotografías de los tumores a diario durante el periodo de tratamiento, y en cada posterior nueva comprobación. Después del tratamiento, cada gato se somete a otro barrido por CT. Los barridos por CT y los radiogramas torácicos se evalúan cada 8 semanas en lo sucesivo. Se evalúan los datos para las diferencias de supervivencia, respuesta y toxicidad en comparación con grupos de control. La respuesta positiva puede requerir pruebas de regresión tumoral, preferentemente con mejora de la calidad de vida y/o una mayor esperanza de vida.

Además, otros tumores animales espontáneos, tales como fibrosarcoma, adenocarcinoma, linfoma, condroma, leiomiomasarcoma de perros, gatos y babuinos también pueden ensayarse. De estos, el adenocarcinoma mamario en perros y gatos es un modelo preferido ya que su aspecto y comportamiento son muy similares a los de seres humanos. Sin embargo, el uso de este modelo está limitado porque este tipo de tumor raramente aparece en animales.

65 K. Ensayos de Cribado para Fármacos candidatos

Se han diseñado ensayos de cribado para fármacos candidatos para identificar compuestos que se unen a o completan con los polipéptidos codificados por los genes identificados en este documento, o que interfieren de otra manera

en la interacción de los polipéptidos codificados con otras proteínas celulares. Dichos ensayos de cribado incluirán ensayos susceptibles al cribado de alto rendimiento de bibliotecas químicas, lo que les hace particularmente adecuados para identificar fármacos candidatos de molécula pequeña. Las moléculas pequeñas contempladas incluyen compuestos sintéticos orgánicos o inorgánicos, incluyendo péptidos, preferentemente péptidos solubles, fusiones de (poli)péptido-inmunoglobulina, y, en particular, anticuerpos incluyendo, sin limitación, anticuerpos y fragmentos de anticuerpos poli- y monoclonales, anticuerpos de cadena sencilla, anticuerpos anti-idiotípicos, y versiones quiméricas o humanizadas de dichos anticuerpos o fragmentos, así como anticuerpos y fragmentos de anticuerpo humanos. Los ensayos pueden realizarse en diversos formatos, incluyendo ensayos de unión proteína-proteína, ensayos de cribado bioquímicos, inmunoensayos y ensayos basados en células, que están bien caracterizados en el sector.

Todos los ensayos tienen en común que requieren poner en contacto el fármaco candidato con un polipéptido codificado por un ácido nucleico identificado en este documento en condiciones y durante un periodo suficiente para permitir que estos dos componentes interactúen.

En los ensayos de unión, la interacción es la unión y el complejo formado puede aislarse o detectarse en la mezcla de reacción. En una realización particular, el polipéptido codificado por el gen identificado en este documento o el fármaco candidato se inmoviliza sobre una fase sólida, por ejemplo, en una placa de microtitulación, mediante uniones covalentes o no covalentes. La unión no covalente generalmente se consigue recubriendo la superficie sólida con una solución del polipéptido y secando. Como alternativa, un anticuerpo inmovilizado, por ejemplo, un anticuerpo monoclonal, específico para el polipéptido a inmovilizar puede usarse para anclarlo a una superficie sólida. El ensayo se realiza añadiendo el componente no inmovilizado, que puede estar marcado mediante una marca detectable, al componente inmovilizado, por ejemplo, La superficie recubierta que contiene el componente anclado. Cuando la restricción es completa, se eliminan los componentes que no reaccionaron, por ejemplo, lavando, y se detectan los complejos anclados en la superficie sólida. Cuando el componente originalmente no inmovilizado tiene una marca detectable, la detección de la marca inmovilizada en la superficie indica que se produjo la complejación. Cuando el componente originalmente no inmovilizado no tiene una marca, la complejación puede detectarse, por ejemplo, usando un anticuerpo marcado que se une específicamente al complejo inmovilizado.

Si el compuesto candidato interacciona con, pero no se une a un polipéptido PRO particular codificado por un gen identificado en este documento, su interacción con ese polipéptido puede ensayarse mediante procedimientos bien conocidos para detectar interacciones proteína-proteína. Dichos ensayos incluyen enfoques tradicionales, tales como, reticulación, co-inmunoprecipitación, y co-purificación a través de gradientes o columnas cromatográficas. Además, las interacciones proteína-proteína pueden supervisarse usando un sistema genético basado en levadura descrito por Fields y colaboradores (Fields y Song, *Nature*, 340: 245-246 (1989); Chien *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 9578-9582 (1991)] según lo descrito por Chevray y Nathans, *Proc. Natl. Acad. Sci. USA*, 89: 5789-5793 (1991)]. Muchos activadores transcripcionales, tales como GAL4 de levadura, consisten en dos dominios modulares físicamente separados, uno que actúa como el dominio de unión al ADN, mientras que el otro funciona como el dominio de activación de la transcripción. El sistema de expresión en levadura descrito en las anteriores publicaciones (generalmente denominado como el “sistema di-híbrido”) se aprovecha de esta propiedad, y emplea dos proteínas híbridas, una en la que la proteína diana se fusiona al dominio de unión al ADN de GAL4, y otra, en la que proteínas activadoras del candidato se fusionan al dominio de activación. La expresión de un gen informador GAL1-lacZ bajo el control de un promotor activado por GAL4 depende de la reconstitución de la actividad de GAL4 mediante la interacción proteína-proteína. Las colonias que contienen polipéptidos que interactúan se detectan con un sustrato cromogénico para β -galactosidasa. Un kit completo (MATCHMAKER™) para identificar interacciones proteína-proteína entre dos proteínas específicas usando la técnica di-híbrida está disponible en el mercado de Clontech. Este sistema también puede extenderse para mapear dominios de proteínas implicados en interacciones de proteínas específicas así como para localizar restos de aminoácidos que son cruciales para estas interacciones.

Los compuestos que interfieren en la interacción de un gen que codifica PRO identificado en este documento y otros componentes intra o extracelulares pueden ensayarse de la siguiente manera: habitualmente se prepara una mezcla de reacción que contiene el producto del gen amplificado y el componente intra o extracelular en condiciones y durante un periodo de tiempo que permiten la interacción y unión de los dos productos. Para ensayar la capacidad de un compuesto de ensayo para inhibir la unión, la reacción se realiza en ausencia y en presencia del compuesto de ensayo. Además, puede añadirse un placebo a una tercera mezcla de reacción, para que sirva como control positivo. La unión (formación de complejo) entre el compuesto de ensayo y el componente intra o extracelular presente en la mezcla se supervisa como se ha descrito anteriormente en este documento. La formación de un complejo en la reacción o reacciones de control pero no en la mezcla de reacción que contiene el compuesto de ensayo indica que el compuesto de ensayo interfiere en la interacción del compuesto de ensayo y su socio de reacción.

Para ensayar antagonistas, el polipéptido PRO puede añadirse a una célula junto con el compuesto a cribar para una actividad particular y la capacidad del compuesto para inhibir la actividad de interés en presencia del polipéptido PRO indica que el compuesto es un antagonista del polipéptido PRO. Como alternativa, pueden detectarse antagonistas combinando el polipéptido PRO y un antagonista potencial con receptores de polipéptido PRO unidos a membrana o receptores de PRO recombinantes en condiciones apropiadas para un ensayo de inhibición competitiva. El polipéptido PRO puede marcarse, tal como mediante radiactividad, de modo que el número de moléculas de polipéptido PRO unidas al receptor pueda usarse para determinar la eficacia del antagonista potencial. El gen que codifica el receptor puede identificarse mediante numerosos procedimientos conocidos por los expertos en la materia, por ejemplo, tapizado con ligando “*ligand panning*” y clasificación mediante FACS. Coligan *et al.*, *Current Protocols in Immun.*, 1(2): Capítulo

5 (1991). Preferentemente, se emplea clonación de expresión en la que se prepara ARN poliadenilado a partir de una célula sensible al polipéptido PRO y una biblioteca de ADNc creada a partir de este ARN se divide en reservas y se usa para transfectar células COS u otras células que no son sensibles al polipéptido PRO. Las células transfectadas que se cultivan en portaobjetos de vidrio se exponen a polipéptido PRO marcado. El polipéptido PRO puede marcarse
 5 mediante diversos medios incluyendo yodación o inclusión de un punto de reconocimiento para una proteína quinasa específica de un punto. Después de la fijación e incubación, los portaobjetos se someten a análisis autorradiográfico. Se identifican las reservas positivas y se preparan sub-reservas y se vuelven a transfectar usando un proceso de formación de sub-reservas y de re-cribado interactivo, que puede producir un único clon que codifica el supuesto receptor.

10 Como enfoque alternativo para la identificación del receptor, el polipéptido PRO marcado puede unirse mediante fotoafinidad con preparaciones de membrana o extracto celular que expresan la molécula receptora. El material reticulado se resuelve mediante PAGE y se expone a una película de rayos X. El complejo marcado que contiene el receptor puede escindirse, resolverse en fragmentos peptídicos y someterse a micro-secuenciación de proteínas. La secuencia de aminoácidos obtenida de la micro-secuenciación se usaría para diseñar un conjunto de sondas de oligonucleótidos
 15 degeneradas para cribar una biblioteca de ADNc para identificar el gen que codifica el supuesto receptor.

En otro ensayo para antagonistas, células de mamífero o una preparación de membrana que expresa el receptor se incubarían con polipéptido PRO marcado en presencia del compuesto candidato. La afinidad del compuesto para mejorar o bloquear esta interacción podía medirse a continuación.

20 Ejemplos más específicos de potenciales antagonistas incluyen un oligonucleótido que se une a las fusiones de inmunoglobulina con el polipéptido PRO, y, en particular, anticuerpos que incluyen, sin limitación, anticuerpos poli- y monoclonales y fragmentos de anticuerpo, anticuerpos de cadena sencilla, anticuerpos anti-idiotípicos y versiones quiméricas o humanizadas de dichos anticuerpos o fragmentos, así como anticuerpos y fragmentos de anticuerpo
 25 humanos. Como alternativa, un antagonista potencial puede ser una proteína estrechamente relacionada, por ejemplo, una forma mutada del polipéptido PRO que reconoce al receptor pero no ejerce ningún efecto, inhibiendo de este modo de forma competitiva la acción del polipéptido PRO.

Otro potencial antagonista del polipéptido PRO es una construcción de ARN o ADN antisentido preparada usando
 30 tecnología antisentido, donde, por ejemplo, una molécula de ARN o ADN antisentido actúa para bloquear directamente la traducción de ARNm hibridando con ARNm fijado como objetivo e impidiendo la traducción de proteínas. Puede usarse tecnología antisentido para controlar la expresión génica a través de formación de triple hélice o ADN o ARN antisentido, basándose ambos procedimientos en la unión de un polinucleótido a ADN o ARN. Por ejemplo, la parte codificante 5' de la secuencia de polinucleótidos, que codifica el polipéptido PRO maduro de este documento, se usa
 35 para diseñar un oligonucleótido de ARN antisentido de desde aproximadamente 10 a 40 pares de bases de longitud. Se diseña un oligonucleótido de ADN para que sea complementario a una región del gen implicado en la transcripción (triple hélice - véase, Lee *et al.*, Nucl. Acids Res., 6: 3073 (1979); Cooney *et al.*, Science, 241: 456 (1988); Dervan *et al.*, Science, 251: 1360 (1991)), impidiendo de este modo la transcripción y la producción del polipéptido PRO. El oligonucleótido de ARN antisentido hibrida con el ARNm *in vivo* y bloquea la traducción de la molécula de ARNm en
 40 el polipéptido PRO (antisentido-Okano, Neurochem., 56: 560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). Los oligonucleótidos descritos anteriormente también pueden suministrarse a células de modo que el ARN o ADN antisentido puedan expresarse *in vivo* para inhibir la producción del PRO. Cuando se usa ADN antisentido, se prefieren oligodesoxirribonucleótidos obtenidos del punto de inicio de la traducción, por ejemplo, entre aproximadamente -10 y +10 posiciones de la secuencia de nucleótidos del gen diana.

45 Las moléculas de ARN o ADN antisentido son generalmente de al menos aproximadamente 5 bases de longitud, aproximadamente 10 bases de longitud, aproximadamente 15 bases de longitud, aproximadamente 20 bases de longitud, aproximadamente 25 bases de longitud, aproximadamente 30 bases de longitud, aproximadamente 35 bases de longitud, aproximadamente 40 bases de longitud, aproximadamente 45 bases de longitud, aproximadamente 50 bases
 50 de longitud, aproximadamente 55 bases de longitud, aproximadamente 60 bases de longitud, aproximadamente 65 bases de longitud, aproximadamente 70 bases de longitud, aproximadamente 75 bases de longitud, aproximadamente 80 bases de longitud, aproximadamente 85 bases de longitud, aproximadamente 90 bases de longitud, aproximadamente 95 bases de longitud, aproximadamente 100 bases de longitud, o más.

55 Los potenciales antagonistas incluyen moléculas pequeñas que se unen al punto activo, al punto de unión al receptor, o factor de crecimiento u otro punto de unión relevante del polipéptido PRO, bloqueando de este modo la actividad biológica normal del polipéptido PRO. Los ejemplos de moléculas pequeñas incluyen, aunque sin limitación, péptidos pequeños o moléculas similares a péptidos, preferentemente péptidos solubles, y compuestos orgánicos o inorgánicos sintéticos no de péptido.

60 Las ribozimas son moléculas de ARN enzimáticas capaces de catalizar la escisión específica de ARN. Las ribozimas actúan mediante hibridación específica de secuencia con el ARN diana complementario, seguida de escisión endonucleolítica. Los puntos de escisión específicos de ribozima en un ARN diana potencial pueden identificarse mediante técnicas conocidas. Para más detalles véase, por ejemplo, Rossi, Current Biology, 4: 469-471 (1994), y
 65 Publicación PCT N°. WO 97/33551 (publicada el 18 de septiembre de 1997).

Las moléculas de ácido nucleico en formación de triple hélice usadas para inhibir la transcripción deben ser de cadena sencilla y estar compuestas por desoxinucleótidos. La composición de base de estos oligonucleótidos se diseña

de modo que promueva la formación de triple hélice mediante las reglas de emparejamiento de bases de Hoogsteen, que generalmente requieren tramos considerables de purinas o pirimidinas en una cadena de una cadena doble. Para más detalles véase, por ejemplo, la Publicación PCT N°. WO 97/33551, *supra*.

- 5 Estas pequeñas moléculas pueden ser identificadas mediante uno o más de los ensayos de cribado descritos anteriormente en este documento y/o mediante cualesquiera otras técnicas de cribado bien conocidas por los expertos en la materia.

L. Composiciones y Procedimientos para el Tratamiento de Tumores

- 10 Las composiciones útiles en el tratamiento de tumores asociados con la amplificación de los genes identificados en este documento incluyen, sin limitación, anticuerpos, pequeñas moléculas orgánicas e inorgánicas, péptidos, fosfo-péptidos, moléculas antisentido y de ribozima, moléculas de triple hélice, etc., que inhiben la expresión y/o actividad del producto génico diana.

- 15 Por ejemplo, las moléculas de ARN y ARN antisentido actúan para bloquear directamente la traducción de ARNm hibridando con ARNm fijado como objetivo e impidiendo la traducción de proteínas. Cuando se usa ADN antisentido, se prefieren oligodesoxirribonucleótidos obtenidos del punto de inicio de la traducción, por ejemplo, entre aproximadamente -10 y +10 posiciones de la secuencia de nucleótidos del gen diana.

- 20 Las ribozimas son moléculas de ARN enzimáticas capaces de catalizar la escisión específica de ARN. Las ribozimas actúan mediante hibridación específica de secuencia con el ARN diana complementario, seguida de escisión endonucleolítica. Los puntos de escisión específicos de ribozima en un ARN diana potencial pueden identificarse mediante técnicas conocidas. Para más detalles véase, por ejemplo, Rossi, Current Biology, 4: 469-471 (1994), y
25 Publicación PCT N°. WO 97/33551 (publicada el 18 de septiembre de 1997).

- 30 Las moléculas de ácido nucleico en formación de triple hélice usadas para inhibir la transcripción deben ser de cadena sencilla y estar compuestas por desoxinucleótidos. La composición de base de estos oligonucleótidos se diseña de modo que promueva la formación de triple hélice mediante las reglas de emparejamiento de bases de Hoogsteen, que generalmente requieren tramos considerables de purinas o pirimidinas en una cadena de una cadena doble. Para más detalles véase, por ejemplo, la Publicación PCT N°. WO 97/33551, *supra*.

- 35 Estas moléculas pueden ser identificadas mediante uno o más de los ensayos de cribado descritos anteriormente en este documento y/o mediante cualesquiera otras técnicas de cribado bien conocidas por los expertos en la materia.

M. Anticuerpos

- 40 Algunos de los más prometedores fármacos candidatos según la presente invención son anticuerpos y fragmentos de anticuerpo que pueden inhibir la producción o el producto génico de los genes amplificados identificados en este documento y/o reducir la actividad de los productos génicos.

1. Anticuerpos Policlonales

- 45 Los procedimientos para preparar anticuerpos policlonales son conocidos por el experto en la materia. Los anticuerpos policlonales pueden generarse en un mamífero, por ejemplo, mediante una o más inyecciones de un agente inmunizante y, si se desea, un adyuvante. Habitualmente, el agente inmunizante y/o adyuvante se inyectará en el mamífero mediante múltiples inyecciones subcutáneas o intraperitoneales. El agente inmunizante puede incluir el polipéptido PRO o una proteína de fusión del mismo. Puede ser útil conjugar el agente inmunizante con una proteína
50 que se sabe que es inmunógena en el mamífero que se inmuniza. Los ejemplos de dichas proteínas inmunogénicas incluyen aunque sin limitación hemocianina de lapa californiana, albúmina de suero, tiroglobulina bovina, e inhibidor de tripsina de soja. Los ejemplos de adyuvantes que pueden emplearse incluyen adyuvante completo de Freund y adyuvante MPL-TDM (Monofosforil Lípido A, dicorinomicolato de trehalosa sintético). El protocolo de inmunización puede ser seleccionado por un experto en la materia sin gran experimentación.

2. Anticuerpos Monoclonales

- Los anticuerpos anti-PRO pueden ser, como alternativa, anticuerpos monoclonales. Los anticuerpos monoclonales pueden prepararse usando procedimientos de hibridoma, tales como los descritos por Kohler y Milstein, Nature, 256: 495 (1975). En un procedimiento de hibridoma, un ratón, hámster u otro animal huésped apropiado, se inmuniza habitualmente con un agente inmunizante para generar linfocitos que producen o son capaces de producir anticuerpos que se unirán específicamente al agente inmunizante. Como alternativa, los linfocitos pueden inmunizarse *in vitro*.

- 65 El agente inmunizante habitualmente incluirá el polipéptido PRO, incluyendo fragmentos, o una proteína de fusión de dicha proteína o un fragmento de la misma. Generalmente, se usan linfocitos de la sangre periférica ("PBL") si se desean células de origen humano, o se usan células esplénicas o células de ganglios linfáticos si se desean fuentes de mamífero de origen no humano. Los linfocitos se fusionan a continuación con una línea celular inmortalizada usando un agente de fusión adecuado, tal como polietilenglicol, para formar una célula de hibridoma [Goding, Monoclonal

Antibodies: Principles and Practice, Academic Press, (1986) p. 59-103]. Las líneas celulares immortalizadas son habitualmente células de mamífero transformadas, particularmente células de mieloma de origen de roedor, bovino y humano. Habitualmente, se emplean líneas celulares de mieloma de rata o de ratón. Las células de hibridoma pueden cultivarse en un medio de cultivo adecuado que preferentemente contiene una o más sustancias que inhiben el crecimiento o la supervivencia de las células no fusionadas immortalizadas. Por ejemplo, si las células parentales carecen de la enzima hipoxantina guanina fosforribosil transferasa (HGPRT o HPRT), el medio de cultivo para los hibridomas habitualmente incluirá hipoxantina, aminopterina, y timidina ("medio HAT"), sustancias que impiden el crecimiento de células deficientes en HGPRT.

Las líneas celulares immortalizadas preferidas son aquellas que se fusionan eficazmente, soportan expresión estable a alto nivel de anticuerpo por las células productoras de anticuerpo seleccionadas, y son sensibles a un medio tal como medio HAT. Las líneas celulares immortalizadas más preferidas son líneas de mieloma múridas, que pueden obtenerse, por ejemplo, del Salk Institute Cell Distribution Center, San Diego, California y la American Type Culture Collection (ATCC), Manassas, Virginia. Las líneas celulares de mieloma humano y heteromieloma de ratón-humano también se han descrito para la producción de anticuerpos monoclonales humanos [Kozbor, J. Immunol., 133: 3001 (1984); Brodeur *et al.*, Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., Nueva York, (1987) p. 51-63].

En el medio de cultivo en el que las células de hibridoma se cultivan, puede ensayarse a continuación la presencia de anticuerpos monoclonales dirigidos contra PRO. Preferentemente, la especificidad de unión de anticuerpos monoclonales producidos por las células de hibridoma se determina mediante inmunoprecipitación o mediante un ensayo de unión *in vitro*, tal como radioinmunoensayo (RIA) o ensayo inmunoabsorbente ligado a enzima (ELISA). Dichas técnicas y ensayos se conocen en el sector. La afinidad de unión del anticuerpo monoclonal puede determinarse, por ejemplo, mediante el análisis de Scatchard de Munson y Pollard, Anal. Biochem., 107: 220 (1980).

Una vez que las células de hibridoma deseadas se han identificado, los clones pueden subclonarse mediante procedimientos de dilución limitante y cultivarse mediante procedimientos convencionales [Goding, *supra*]. Los medios de cultivo adecuados para este fin incluyen, por ejemplo, Medio Eagle Modificado por Dulbecco y medio RPMI-1640. Como alternativa, las células de hibridoma pueden cultivarse *in vivo* como ascitis en un mamífero.

Los anticuerpos monoclonales secretados por los subclones pueden aislarse o purificarse del medio de cultivo o fluido ascítico mediante procedimientos de purificación de inmunoglobulina convencionales tales como, por ejemplo, proteína A-Sepharose, cromatografía en hidroxipatita, electroforesis en gel, diálisis o cromatografía de afinidad.

Los anticuerpos monoclonales también pueden prepararse mediante procedimientos de ADN recombinantes, tales como los descritos en la Patente de Estados Unidos N°. 4.816.567. El ADN que codifica los anticuerpos monoclonales de la presente invención pueden aislarse y secuenciarse fácilmente usando procedimientos convencionales (por ejemplo, usando sondas de oligonucleótidos que son capaces de unirse específicamente a genes que codifican las cadenas pesada y ligera de anticuerpos múridos). Las células de hibridoma de la presente invención sirven como fuente preferida de dicho ADN. Una vez aislado, el ADN puede colocarse en vectores de expresión, que a continuación se transfectan en células huésped tales como células COS de simio, células de ovario de hámster chino (CHO), células de mieloma que, de otra manera no producen proteína inmunoglobulina, para obtener la síntesis de anticuerpos monoclonales en las células huésped recombinantes. El ADN también puede modificarse, por ejemplo, sustituyendo la secuencia que codifica dominios constantes de cadena pesada y ligera humanas en lugar de las secuencias múridas homólogas [Patente de Estados Unidos N° 4.816.567; Morrison *et al.*, *supra*] o uniendo covalentemente a la secuencia codificante de inmunoglobulina toda o parte de la secuencia codificante de un polipéptido no inmunoglobulina. Dicho polipéptido no inmunoglobulina puede sustituirse por los dominios constantes de un anticuerpo de la presente invención, o puede sustituirse por los dominios variables de un punto de combinación con el antígeno de un anticuerpo de la presente invención para crear un anticuerpo quimérico bivalente.

Los anticuerpos pueden ser anticuerpos monovalentes. Los procedimientos para preparar anticuerpos monovalentes se conocen bien en el sector. Por ejemplo, un procedimiento implica expresión recombinante de cadena ligera y cadena pesada modificada de inmunoglobulina. La cadena pesada está truncada generalmente en cualquier punto en la región Fc para evitar la reticulación de la cadena pesada. Como alternativa, los restos cisteína relevantes se sustituyen por otro resto de aminoácido o se eliminan para impedir la reticulación.

Los procedimientos *in vitro* también son adecuados para preparar anticuerpos monovalentes. La digestión de anticuerpos para producir fragmentos de los mismos, particularmente, fragmentos Fab, puede conseguirse usando técnicas rutinarias conocidas en el sector.

3. Anticuerpos Humanos y Humanizados

Los anticuerpos anti-PRO pueden comprender además anticuerpos humanizados o anticuerpos humanos. Las formas humanizadas de anticuerpos no humanos (por ejemplo, múridos) son inmunoglobulinas quiméricas, cadenas de inmunoglobulina o fragmentos de las mismas (tales como Fv, Fab, Fab', F(ab'), u otras sub-secuencias de anticuerpos de unión al antígeno) que contienen una mínima secuencia obtenida de inmunoglobulina no humana. Los anticuerpos humanizados incluyen inmunoglobulinas humanas (anticuerpo receptor) en las que restos de una región determinante de la complementariedad (CDR) del receptor se sustituyen por restos de una CDR de una especie no humana (anti-

cuerpo donador) tal como ratón, rata o conejo que tiene la especificidad, afinidad y capacidad deseadas. En algunos casos, restos marco Fv de la inmunoglobulina humana se sustituyen por restos no humanos correspondientes. Los anticuerpos humanizados también pueden comprender restos que se no encuentran ni en el anticuerpo receptor ni en la CDR o secuencias marco importadas. En general, el anticuerpo humanizado comprenderá sustancialmente todos de al menos uno, y habitualmente dos, dominios variables, en los que todas o sustancialmente todas las regiones CDR corresponden a las de una inmunoglobulina no humana y todas o sustancialmente todas las regiones FR son las de una secuencia consenso de inmunoglobulina humana. El anticuerpo humanizado óptimamente también comprenderá al menos una parte de una región constante de inmunoglobulina (Fc), habitualmente la de una inmunoglobulina humana [Jones *et al.*, Nature, 321: 522-525 (1986); Riechmann *et al.*, Nature, 332: 323-329 (1988); y Presta, Curr. Op. Struct. Biol., 2: 593-596 (1992)].

Los procedimientos para humanizar anticuerpos no humanos se conocen bien en el sector. Generalmente, un anticuerpo humanizado tiene uno o más restos de aminoácidos introducidos en él a partir de una fuente que no es humana. Estos restos de aminoácidos no humanos a menudo se denominan como restos "importados", que habitualmente se toman de un dominio variable "importado". La humanización puede realizarse esencialmente siguiendo el procedimiento de Winter y colaboradores [Jones *et al.*, Nature, 321: 522-525 (1986); Riechmann *et al.*, Nature, 332: 323-327 (1988); Verhoeyen *et al.*, Science, 239: 153A-1536 (1988)], sustituyendo CDR o secuencias de CDR de roedor por las secuencias correspondientes de un anticuerpo humano. Por consiguiente, dichos anticuerpos "humanizados" son anticuerpos quiméricos (Patente de Estados Unidos N° 4.816.567), donde sustancialmente menos de un dominio variable humano intacto ha sustituido a la secuencia correspondiente de una especie no humana. En la práctica, los anticuerpos humanizados son habitualmente anticuerpos humanos en los que algunos restos de CDR y posiblemente algunos restos de FR sustituyen a restos de puntos análogos en anticuerpos de roedor.

Los anticuerpos humanos también pueden producirse usando diversas técnicas conocidas en el sector, incluyendo bibliotecas de presentación en fagos [Hoogenboom y Winter, J. Mol. Biol., 227: 381 (1991); Marks *et al.*, J. Mol. Biol., 222: 581 (1991)]. Las técnicas de Cole *et al.*, y Boerner *et al.*, también están disponibles para la preparación de anticuerpos monoclonales humanos (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) y Boerner *et al.*, J. Immunol., 147(1): 86-95 (1991)]. Análogamente, pueden prepararse anticuerpos humanos introduciendo loci de inmunoglobulina humana en animales transgénicos, por ejemplo, ratones en los que los genes de inmunoglobulina endógenos se han inactivado parcial o totalmente. Después de la exposición, se observa la producción de anticuerpo humano, que recuerda estrechamente a la observada en seres humanos en todos los aspectos, incluyendo reordenación génica, ensamblaje y repertorio de anticuerpos. Este enfoque se describe, por ejemplo, en las Patentes de Estados Unidos N° 5.545.807; 5.545.806; 5.569.825; 5.625.126; 5.633.425; 5.661.016, y en las siguientes publicaciones científicas: Marks *et al.*, Bio/Technology, 10: 779-783 (1992); Lonberg *et al.*, Nature, 368: 856-859 (1994); Morrison, Nature, 368: 812-13 (1994); Fishwild *et al.*, Nature Biotechnology, 14: 845-51 (1996); Neuberger, Nature Biotechnology, 14: 826 (1996); Lonberg y Huszar, Intern. Rev. Immunol., 13: 65-93 (1995).

4. Terapia de Profármacos Mediada Por Enzima Dependiente de Anticuerpo (ADEPT)

Los anticuerpos de la presente invención también pueden usarse en ADEPT conjugando el anticuerpo a una enzima activadora de profármaco que convierte a un profármaco (por ejemplo, un agente quimioterapéutico de peptidilo, véase WO 81/01145) en un fármaco anti-cáncer activo. Véase, por ejemplo, WO 88/07378 y Patente de Estados Unidos N°. 4.975.278.

El componente enzimático del inmunoconjugado útil para ADEPT incluye cualquier enzima capaz de actuar sobre un profármaco de tal manera que lo convierta en su forma citotóxica, más activa.

Las enzimas que son útiles en el procedimiento de la presente invención incluyen, aunque sin limitación, glucosidasa, glucosa oxidasa, lisozima humana, glucuronidasa humana, fosfatasa alcalina útil para convertir profármacos que contienen fosfato en fármacos libres; arilsulfatasa útil para convertir profármacos que contienen sulfato en fármacos libres; citosina desaminasa útil para convertir 5-fluorocitosina no tóxica en el fármaco anti-cáncer 5-fluorouracilo; proteasas tales como proteasa de serrata, termolisina, subtilisina, carboxipeptidasas (por ejemplo, carboxipeptidasa G2 y carboxipeptidasa A) y catepsinas (tales como catepsinas B y L), que son útiles para convertir profármacos que contienen péptidos en fármacos libres: D-alanilcarboxipeptidasas, útiles para convertir profármacos que contienen sustituyentes de D-aminoácidos; enzimas que escinden carbohidratos tales como β -galactosidasa y neuraminidasa útiles para convertir profármacos glucosilados en fármacos libres; β -lactamasa útil para convertir fármacos derivatizados con β -lactamas en fármacos libres; y penicilina amidasas, tales como penicilina V amidasa o penicilina G amidasa, útiles para convertir fármacos derivatizados en sus nitrógenos de amina con grupos fenoxiacetilo o fenilacetilo, respectivamente, en fármacos libres. Como alternativa, pueden usarse anticuerpos con actividad enzimática, también conocidos en el sector como "abzimas" para convertir los profármacos de la presente invención en grupos activos libres (véase, por ejemplo, Massey, Nature, 328: 457-458 (1987)). Los conjugados anticuerpo-abzima pueden prepararse como se describe en este documento para el suministro de la abzima a una población de células tumorales.

Las enzimas de la presente invención pueden unirse covalentemente a los anticuerpos anti-PRO mediante técnicas bien conocidas en el sector tales como el uso de los agentes reticulantes heterobifuncionales descritos anteriormente. Como alternativa, pueden construirse proteínas de fusión que comprenden al menos la región de unión al antígeno del anticuerpo de la presente invención unida a al menos una parte funcionalmente activa de una enzima de la presente

invención, usando técnicas de ADN recombinantes bien conocidas en el sector (véase, por ejemplo, Neuberger *et al.*, Nature, 312: 604-608 (1984)).

5. Anticuerpos Biespecíficos

Los anticuerpos biespecíficos son anticuerpos monoclonales, preferentemente humanos o humanizados, que tienen especificidades de unión por al menos dos antígenos diferentes. En el presente caso, una de las especificidades de unión es por el PRO y la otra por cualquier otro antígeno, y preferentemente por una proteína o receptor o subunidad receptora de la superficie celular.

Los procedimientos para preparar anticuerpos biespecíficos se conocen en el sector. Tradicionalmente, la producción recombinante de anticuerpos biespecíficos se basa en la co-expresión de dos pares de cadena pesada/cadena ligera de inmunoglobulina, donde las dos cadenas pesadas tienen diferentes especificidades (Milstein y Cuello, Nature, 305: 537-539 [1983]). Debido a la selección aleatoria de cadenas pesadas y ligeras de inmunoglobulina, estos híbridos (cuadromas) producen una mezcla potencial de diez moléculas de anticuerpo diferentes, de las cuales solamente una tiene la estructura biespecífica correcta. La purificación de la molécula correcta se consigue habitualmente mediante etapas de cromatografía de afinidad. Procedimientos similares se describen en WO 93/08829, publicado el 13 de mayo de 1993, y en Traunecker *et al.*, EMBO J., 10: 3655-3659 (1991).

Los dominios variables de anticuerpos con las especificidades de unión deseadas (puntos de combinación anticuerpo-antígeno) pueden fusionarse a secuencias de dominio constante de inmunoglobulina. La fusión preferentemente es con un dominio constante de cadena pesada de inmunoglobulina, que comprende al menos parte de las regiones bisagra, CH2 y CH3. Se prefiere tener la primera región constante de cadena pesada (CH 1) que contiene el punto necesario para la unión a cadena ligera presente en al menos una de las fusiones. Los ADN que codifican las fusiones de cadena pesada de inmunoglobulina y, si se desea, la cadena ligera de inmunoglobulina, se insertan en vectores de expresión diferentes, y se co-transfectan en un organismo huésped adecuado. Para más detalles de la generación de anticuerpos biespecíficos véase, por ejemplo, Suresh *et al.*, Methods in Enzymology, 121: 210 (1986).

Según otro enfoque descrito en WO 96/27011, el interfaz entre un par de moléculas de anticuerpo puede diseñarse para maximizar el porcentaje de heterodímeros que se recuperan del cultivo de células recombinantes. El interfaz preferido comprende al menos una parte de la región CH3 de un dominio constante de anticuerpo. En este procedimiento, una o más pequeñas cadenas laterales de aminoácido del interfaz de la primera molécula de anticuerpo se sustituyen por cadenas laterales más grandes (por ejemplo, tirosina o triptófano). Se crean "cavidades" compensadoras de tamaño idéntico o similar a la cadena o cadenas laterales grandes en el interfaz de la segunda molécula de anticuerpo sustituyendo grandes cadenas laterales de aminoácidos por cadenas más pequeñas (por ejemplo, alanina o treonina). Esto proporciona un mecanismo para aumentar el rendimiento del heterodímero con respecto a otros productos finales no deseados tales como homodímeros.

Pueden prepararse anticuerpos biespecíficos como anticuerpos o fragmentos de anticuerpo de longitud completa (por ejemplo, anticuerpos biespecíficos $F(ab')_2$). Las técnicas para generar anticuerpos biespecíficos a partir de fragmentos de anticuerpo se han descrito en la bibliografía. Por ejemplo, pueden prepararse anticuerpos biespecíficos usando enlace químico. Brennan *et al.*, Science, 229: 81 (1985) describen un procedimiento en el que anticuerpos intactos se escinden proteolíticamente para generar fragmentos $F(ab')_2$. Estos fragmentos se reducen en presencia del agente complejante de ditioi arsenito sódico para estabilizar ditioles adyacentes e impedir la formación de puentes disulfuro intermoleculares. Los fragmentos Fab' generados se convierten a continuación en derivados de tionitrobenzoato (TNB). Uno de los derivados de Fab' -TNB se reconvierte a continuación en Fab' -tiol mediante reducción con mercaptoetilamina y se mezcla con una cantidad equimolar del otro derivado de Fab' -TNB para formar el anticuerpo biespecífico. Los anticuerpos biespecíficos producidos pueden usarse como agentes para la inmovilización selectiva de enzimas.

Los fragmentos Fab' pueden recuperarse directamente de *E. coli* y acoplarse químicamente para formar anticuerpos biespecíficos. Shalaby *et al.*, J. Exp. Med., 175: 217-225 (1992) describen la producción de una molécula $F(ab')_2$ de anticuerpo biespecífico completamente humanizado. Cada fragmento Fab' se secreta por separado a partir de *E. coli* y se sometió a acoplamiento químico dirigido *in vitro* para formar el anticuerpo biespecífico. El anticuerpo biespecífico formado de este modo era capaz de unirse a células que sobre-expresan el receptor ErbB2 y células T humanas normales, así como de desencadenar la actividad lítica de linfocitos citotóxicos humanos contra dianas tumorales de mama humanas.

También se han descrito diversas técnicas para preparar y aislar fragmentos de anticuerpo biespecíficos directamente a partir de cultivo de células recombinantes. Por ejemplo, se han producido anticuerpos biespecíficos usando cremalleras de leucina. Kostelny *et al.*, J. Immunol., 148(5): 1547-1553 (1992). Los péptidos de la cremallera de leucina de las proteínas Fos y Jun se enlazaron a las partes Fab' de dos anticuerpos diferentes mediante fusión génica. Los homodímeros de anticuerpo se redujeron en la región bisagra para formar monómeros y a continuación se oxidaron de nuevo para formar los heterodímeros de anticuerpo. Este procedimiento también puede utilizarse para la producción de homodímeros de anticuerpo. La tecnología de "diacuerpos" descrita por Hollinger *et al.*, Proc. Natl. Acad. Sci. USA, 90: 6444-6448 (1993) proporcionó un mecanismo alternativo para preparar fragmentos de anticuerpo biespecíficos. Los fragmentos comprenden un dominio variable de cadena pesada (V_H) conectados a un dominio variable de cadena ligera (V_L) mediante un enlazador que es demasiado corto para permitir el emparejamiento entre dos dominios de la

misma cadena. Por consiguiente, se obliga a los dominios V_H y V_L de un fragmento a emparejarse con los dominios V_L y V_H complementarios de otro fragmento, formando de este modo dos puntos de unión al antígeno. Otra estrategia para preparar fragmentos de anticuerpo biespecíficos mediante el uso de dímeros Fv de cadena sencilla (sFv) también se ha descrito. Véase, Gruber *et al.*, J. Immunol., 152: 5368 (1994).

Se contemplan anticuerpos con más de dos valencias. Por ejemplo, pueden prepararse anticuerpo triespecíficos. Tun *et al.*, J. Immunol., 147: 60 (1991).

Los anticuerpos biespecíficos ejemplares pueden unirse a dos epítomos diferentes en un polipéptido dado en este documento. Como alternativa, un brazo anti-polipéptido puede combinarse con un brazo que se une a una molécula inductora en un leucocito tal como una molécula receptora de células T (por ejemplo, CD2, CD3, CD28 o B7), o receptores de Fc para IgG (FcγR), tales como FcγRI (CD64), FcγRII (CD32) y FcγRIII (CD16) para enfocar mecanismos de defensa celular a la célula que expresa el polipéptido particular. También pueden usarse anticuerpos biespecíficos para localizar agentes citotóxicos para células que expresan un polipéptido particular. Estos anticuerpos poseen un brazo de unión al polipéptido y un brazo que se une a un agente citotóxico o un quelante de radionúclidos, tal como EOTUBE, DPTA, DOTA o TETA. Otro anticuerpo biespecífico de interés se une al polipéptido y se une además al factor tisular (TF).

6. Anticuerpos Heteroconjugados

Los anticuerpos heteroconjugados están compuestos por dos anticuerpos unidos covalentemente. Dichos anticuerpos se han propuesto, por ejemplo, para dirigir células del sistema inmune a células no deseadas [Patente de Estados Unidos N° 4.676.980], y para el tratamiento de infección por VIH [WO 91/00360; WO 92/200373; EP 03089]. Se contempla que los anticuerpos pueden prepararse *in vitro* usando procedimientos conocidos en química de proteínas sintéticas, incluyendo los que implican agentes reticulantes. Por ejemplo, pueden construirse inmunotoxinas usando una reacción de intercambio de disulfuro o formando un enlace tioéter. Los ejemplos de reactivos adecuados para este fin incluyen iminotiolato y metil-4-mercaptobutirimidato y los descritos, por ejemplo, en la Patente de Estados Unidos N° 4.676.980.

7. Manipulación de la Función Efectora

Puede ser deseable modificar el anticuerpo de la presente invención con respecto a la función efectora, para, por ejemplo, aumentar la eficacia del anticuerpo en el tratamiento del cáncer. Por ejemplo, pueden introducirse un resto o restos de cisteína en la región Fc, permitiendo de este modo la formación de puentes disulfuro intercatenarios en esta región. El anticuerpo homodimérico generado de este modo puede tener una mejor capacidad de internalización y/o mayor destrucción celular mediada por complemento y citotoxicidad celular dependiente de anticuerpo (ADCC). Véase, Caron *et al.*, J. Exp Med., 176: 1191-1195 (1992) y Shopes, J. Immunol., 148: 2918-2922 (1992). También pueden prepararse anticuerpos homodiméricos con mayor actividad anti-tumoral usando reticulantes heterobifuncionales como se describen en Wolff *et al.*, Cancer Research, 53: 2560-2565 (1993). Como alternativa, puede diseñarse un anticuerpo que tiene regiones Fc dobles y de este modo puede tener mayores capacidades de lisis del complemento y ADCC. Véase, Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).

8. Inmunoconjugados

La presente invención también se refiere a inmunoconjugados que comprenden un anticuerpo conjugado con un agente citotóxico tal como un agente quimioterapéutico, toxina (por ejemplo, una toxina enzimáticamente activa de origen bacteriano, fúngico, vegetal o animal, o fragmentos de la misma, o una toxina de molécula pequeña), o un isótopo radiactivo (es decir, un radioconjugado).

Los agentes quimioterapéuticos útiles en la generación de dichos inmunoconjugados se han descrito anteriormente. toxinas de proteínas enzimáticamente activas y fragmentos de las mismas que pueden usarse incluyen cadena A de difteria, fragmentos activos no de unión de toxina de difteria, toxina del cólera, toxina botulínica, cadena A de exotoxina (de *Pseudomonas aeruginosa*), cadena A de ricina, cadena A de abrina, cadena A de modicina, alfa-sarcina, proteínas de *Aleurites fordii*, proteínas de diantina, proteínas de *Phytolacca americana* (PAPI, PAPII, y PAP-S), inhibidor de *Momordica charantia*, curcina, crotina, inhibidor de *Saponaria officinalis*, gelonina, saporina, mitogelina, restrictocina, fenomicina, enomicina y los tricotecenos. Las toxinas de molécula pequeña incluyen, por ejemplo, caliqueamicinas, maitansinoides, palitoxina y CC1065. Diversos radionúclidos están disponibles para la producción de anticuerpos radioconjugados. Los ejemplos incluyen ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y y ^{186}Re .

Los conjugados del anticuerpo y agente citotóxico se preparan usando diversos agentes de acoplamiento de proteínas bifuncionales tales como N-succinimidil-3-(2-piridilditiol) propionato (SPDP), iminotiolano (IT), derivados bifuncionales de imidoésteres (tales como dimetil adipimidato HCL), ésteres activos (tales como disuccinimidil suberato), aldehídos (tales como glutaraldehído), compuestos bis-azido (tales como bis (p-azidobenzoyl) hexanodiamina), derivados de bis-diazonio (tales como bis-(p-diazoniobenzoyl)-etilendiamina), diisocianatos (tales como tolien 2,6-diisocianato) y compuestos de flúor bis-activos (tales como 1,5-difluoro-2,4-dinitrobenzeno). Por ejemplo, una inmunotoxina de ricina puede prepararse como se describe en Vitetta *et al.*, Science, 238: 1098 (1987). El ácido 1-isotiocianatobencil-3-metildietilen triaminedipentaacético marcado con carbono 14 (MX-DTPA) es un agente quelante ejemplar para la conjugación de un radionucleótido con el anticuerpo. Véase, WO94/11026.

En otra realización, el anticuerpo puede conjugarse con un “receptor” (tal como estreptavidina) para la utilización en la pre-fijación como objetivo de un tumor donde el conjugado anticuerpo-receptor se administra al paciente, seguido de la retirada del conjugado no unido de la circulación usando un agente de eliminación y a continuación administración de un “ligando” (por ejemplo, avidina) que está conjugado a un agente citotóxico (por ejemplo, un radionucleótido).

9. Inmunoliposomas

Los anticuerpos descritos en este documento también pueden formularse como inmunoliposomas. Los liposomas que contienen el anticuerpo se preparan mediante procedimientos conocidos en el sector, tales como se describen en Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); y Patentes de Estados Unidos N° 4.485.045 y 4.544.545. Los liposomas con mayor periodo de circulación se describen en la Patente de Estados Unidos N° 5.013.556.

Pueden generarse liposomas particularmente útiles mediante el procedimiento de evaporación en fase inversa con una composición lipídica que comprende fosfatidilcolina, colesterol y fosfatidiletanolamina derivatizada con PEG (PEG-PE). Los liposomas se extruyen a través de filtros de tamaño de poro definido para producir liposomas con el diámetro deseado. Los fragmentos Fab’ del anticuerpo de la presente invención pueden conjugarse con los liposomas como se describen en Martin *et al.*, J. Biol. Chem., 257: 286-288 (1982) mediante una reacción de intercambio de disulfuro. Un agente quimioterapéutico (tal como Doxorubicina) está contenido opcionalmente dentro del liposoma. Véase, Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

N. Composiciones Farmacéuticas

Los anticuerpos que se unen específicamente al producto de un gen amplificado identificado en este documento, así como otras moléculas identificadas mediante los ensayos de cribado descritos anteriormente en este documento, pueden administrarse para el tratamiento de tumores, incluyendo cánceres, en forma de composiciones farmacéuticas.

Si la proteína codificada por el gen amplificado es intracelular y se usan anticuerpos completos como inhibidores, se prefieren anticuerpos de internalización. Sin embargo, también pueden usarse lipofecciones o liposomas para suministrar el anticuerpo, o un fragmento de anticuerpo, a las células. Cuando se usan fragmentos de anticuerpo, se prefiere el fragmento inhibidor más pequeño que se une específicamente al dominio de unión de la proteína diana. Por ejemplo, en base a las secuencias de región variable de un anticuerpo, pueden diseñarse moléculas peptídicas que conservan la capacidad de unirse a la secuencia de la proteína diana. Dichos péptidos pueden sintetizarse químicamente y/o producirse mediante tecnología de ADN recombinante (véase, por ejemplo, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 [1993]).

Se preparan formulaciones terapéuticas del anticuerpo para almacenamiento mezclando el anticuerpo que tiene el grado de pureza deseado con portadores, excipientes o estabilizantes farmacéuticamente aceptables opcionales (Remington’s Pharmaceutical Sciences, 16th edition, Osol, A. ed. [1980]), en forma de formulaciones liofilizadas o soluciones acuosas. Los portadores, excipientes o estabilizantes aceptables no son tóxicos para los receptores a las dosificaciones y concentraciones empleadas, e incluyen tampones tales como fosfato, citrato, y otros ácidos orgánicos; antioxidantes incluyendo ácido ascórbico y metionina; conservantes (tales como cloruro de octadecildimetilbencilamonio; cloruro de hexametanio; cloruro de benzalconio, cloruro de bencetonio; fenol, alcohol butílico o bencílico; alquil parabenos tales como metil o propil parabeno; catecol; resorcinol; ciclohexanol; 3-pentanol; y m-cresol); polipéptidos de bajo peso molecular (menos de aproximadamente 10 restos); proteínas, tales como albúmina de suero, gelatina, o inmunoglobulinas; polímeros hidrofílicos tales como polivinilpirrolidona; aminoácidos tales como glicina, glutamina, asparagina, histidina, arginina o lisina; monosacáridos, disacáridos, y otros carbohidratos incluyendo glucosa, manosa o dextrinas; agentes quelantes tales como EDTA; azúcares tales como sacarosa, manitol, trehalosa o sorbitol; contraiones formadores de sales tales como sodio; complejos metálicos (por ejemplo, complejos de Zn-proteína); y/o tensioactivos no iónicos tales como TWEEN™, PLURONICS™ o polietilenglicol (PEG).

Los compuestos no de anticuerpo identificados mediante los ensayos de cribado de la presente invención pueden formularse de manera análoga, usando técnicas convencionales bien conocidas en el sector.

La formulación en este documento también puede contener más de un compuesto activo según sea necesario para la indicación particular que se está tratando, preferentemente aquellos con actividades complementarias que no se afectan de forma adversa entre sí. Como alternativa, o además, la composición puede comprender un agente citotóxico, citoquina o agente inhibidor del crecimiento. Dichas moléculas están idóneamente presentes en combinación en cantidades que son eficaces para el fin pretendido.

Los ingredientes activos también pueden estar atrapados en microcápsulas preparadas, por ejemplo, mediante técnicas de coacervación o mediante polimerización interfacial, por ejemplo, microcápsulas de hidroximetilcelulosa o gelatina y microcápsulas de poli-(metilmetacrilato), respectivamente, en sistemas de administración de fármaco coloidales (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nano-partículas y nanocápsulas) o en macroemulsiones. Dichas técnicas se describen en Remington’s Pharmaceutical Sciences, 16th edition, Osol, A. ed. (1980).

Las formulaciones que se usarán para administración *in vivo* deben ser estériles. Esto se consigue fácilmente mediante filtración a través de membranas de filtración estériles.

Pueden prepararse preparaciones de liberación sostenida. Los ejemplos adecuados de preparaciones de liberación sostenida incluyen matrices semipermeables de polímeros hidrofóbicos sólidos que contienen el anticuerpo, matrices que están en forma de artículos conformados, por ejemplo, películas o microcápsulas. Los ejemplos de matrices de liberación sostenida incluyen poliésteres, hidrogeles (por ejemplo, poli(2-hidroxietil-metacrilato), o alcohol polivinílico, polilacturos (Patente de Estados Unidos N° 3.773.919), copolímeros de ácido L-glutámico y etil-L-glutamato, elileno-acetato de vinilo no degradable, copolímeros de ácido láctico-ácido glicólico degradables tales como el LU-PRON DEPOT™ (microesferas inyectables compuestas por copolímero de ácido láctico-ácido glicólico y acetato de leuprolido), y ácido poli-D-(-)-3-hidroxibutírico. Aunque polímeros tales como elileno-acetato de vinilo y ácido láctico-ácido glicólico permiten la liberación de moléculas durante más de 100 días, algunos hidrogeles liberan proteínas durante periodos de tiempo más cortos. Cuando los anticuerpos encapsulados permanecen en el cuerpo durante un largo periodo, pueden desnaturalizarse o agregarse como resultado de la exposición a la humedad a 37°C, dando como resultado una pérdida de actividad biológica y posibles cambios en la inmunogenicidad. Pueden concebirse estrategias racionales para la estabilización dependiendo del mecanismo implicado. Por ejemplo, si se descubre que el mecanismo de agregación es la formación de puentes S-S intermoleculares a través de intercambio tio-disulfuro, puede conseguirse la estabilización modificando restos sulfhidrilo, liofilizando a partir de soluciones ácidas, controlando el contenido de humedad, usando aditivos apropiados y desarrollando composiciones de matriz polimérica específicas.

O. *Methods of Treatment*

Se contempla que los anticuerpos y otros compuestos anti-tumorales de la presente invención pueden usarse para tratar diversas afecciones, incluyendo las caracterizadas por sobre-expresión y/o activación de los genes amplificados identificados en este documento. Las afecciones o trastornos ejemplares a tratar con dichos anticuerpos y otros compuestos, incluyen, aunque sin limitación, moléculas pequeñas orgánicas e inorgánicas, péptidos, moléculas antisentido, etc., incluyen tumores benignos o malignos (por ejemplo, renal, de hígado, de riñón, de vejiga, de mama, gástrico, de ovario, colorrectal, de próstata, pancreático, de pulmón, de vulva, de tiroides, carcinomas hepáticos; sarcomas; glioblastomas; y diversos tumores de cabeza y cuello); leucemias y tumores linfoides; otros trastornos tales como neuronales, gliales, astrocíticos, hipotalámicos y otros trastornos glandulares, macrofágicos, epiteliales, estromáticos y blastocélicos; y trastornos inflamatorios, angiogénicos e inmunológicos.

Los agentes anti-tumorales por ejemplo, anticuerpos, se administran a un mamífero, preferentemente un ser humano, según procedimientos conocidos, tales como administración intravenosa en embolada o mediante infusión continua durante un periodo de tiempo, mediante vías intramuscular, intraperitoneal, intracerebroespinal, subcutánea, intra-articular, intrasinovial, intratecal, oral, tópica, o por inhalación. Se prefiere la administración intravenosa del anticuerpo.

Otros regímenes terapéuticos pueden combinarse con la administración de los agentes anti-cáncer, por ejemplo, anticuerpos de la presente invención. Por ejemplo, el paciente a tratar con dichos agentes anti-cáncer también puede recibir terapia de radiación. Como alternativa, o además, puede administrarse un agente quimioterapéutico al paciente. Los programas de preparación y dosificación para dichos agentes quimioterapéuticos pueden usarse según las instrucciones del fabricante o según determine empíricamente el experto en la materia. Los programas de preparación y dosificación para dicha quimioterapia también se describen en *Chemotherapy Service Ed.*, M.C. Perry, Williams & Wilkins, Baltimore, MD (1992). El agente quimioterapéutico puede preceder o seguir a la administración del agente anti-tumoral, por ejemplo, anticuerpo, o puede administrarse simultáneamente a éste. El anticuerpo puede combinarse con un compuesto anti-estrógeno tal como tamoxifeno o un anti-progesterona tal como onapristona (véase, EP 616812) en dosificaciones conocidas para dichas moléculas.

Puede ser deseable administrar también anticuerpos contra otros antígenos asociados a tumores, tales como anticuerpos que se unen a los ErbB2, EGFR, ErbB3, ErbB4, o factor vascular endotelial (VEGF). Como alternativa, o además, dos o más anticuerpos que se unen al mismo o a dos o más antígenos diferentes descritos en este documento pueden co-administrarse al paciente. En ocasiones, puede ser beneficioso administrar también una o más citoquinas al paciente. En una realización preferida, los anticuerpos de este documento se co-administran con un agente inhibidor del crecimiento. Por ejemplo, el agente inhibidor del crecimiento puede administrarse en primer lugar, seguido de un anticuerpo de la presente invención. Sin embargo, la administración simultánea o la administración del anticuerpo de la presente invención en primer lugar también se contemplan. Las dosificaciones adecuadas para el agente inhibidor del crecimiento son las que se usan actualmente y pueden rebajarse debido a la acción combinada (sinergia) del agente inhibidor del crecimiento y el anticuerpo de este documento.

Para la prevención o tratamiento de una enfermedad, la dosificación apropiada de un agente anti-tumoral, por ejemplo, un anticuerpo en este documento dependerá del tipo de enfermedad a tratar, como se ha definido anteriormente, la gravedad y el curso de la enfermedad, ya se administre el agente para fines preventivos o terapéuticos, la terapia previa, el historial clínico del paciente y la respuesta al agente, y la discreción del médico responsable. El agente se administra adecuadamente al paciente en una vez o a lo largo de una serie de tratamientos.

Por ejemplo, dependiendo del tipo y gravedad de la enfermedad, aproximadamente de 1 µg/kg a 15 mg/kg (por ejemplo, 0,1-20 mg/kg) de anticuerpo es una dosificación candidata inicial para administración al paciente, ya sea, por ejemplo, mediante una o más administraciones diferentes, o mediante infusión continua. Una dosificación diaria típica

puede variar de aproximadamente 1 $\mu\text{g/kg}$ a 100 mg/kg o más, dependiendo de los factores mencionados anteriormente. Para administraciones repetidas a lo largo de varios días o más, dependiendo de la afección, el tratamiento se prolonga hasta que se produce una supresión deseada de los síntomas de la enfermedad. Sin embargo, otros regímenes de dosificación pueden ser útiles. El progreso de esta terapia se supervisa fácilmente mediante técnicas y ensayos convencionales.

P. Artículos de Fabricación

En otra realización de la presente invención, se proporciona un artículo de fabricación que contiene materiales útiles para el diagnóstico o tratamiento de los trastornos descritos anteriormente. El artículo de fabricación comprende un recipiente y una etiqueta. Los recipientes adecuados incluyen, por ejemplo, frascos, viales, jeringas y tubos de ensayo. Los recipientes pueden estar formados de diversos materiales tales como vidrio o plástico. El recipiente contiene una composición que es eficaz para diagnosticar o tratar la afección y puede tener un puerto de acceso estéril (por ejemplo el recipiente puede ser una bolsa de solución intravenosa o un vial que tiene un tapón perforable mediante una aguja de inyección hipodérmica). El agente activo en la composición es habitualmente un agente anti-tumoral capaz de interferir en la actividad de un producto génico identificado en este documento, por ejemplo, un anticuerpo. La etiqueta en, o asociada con, el recipiente indica que la composición se usa para diagnosticar o tratar la afección de elección. El artículo de fabricación puede comprender además un segundo recipiente que comprende un tampón farmacéuticamente aceptable, tal como solución salina tamponada con fosfato, solución de Ringer y solución de dextrosa. Puede incluir además otros materiales deseables desde un punto de vista comercial y del usuario, incluyendo otros tampones, diluyentes, filtros, agujas, jeringas, e insertos en el envase con instrucciones de uso.

Q. Diagnóstico y Pronóstico de Tumores

Aunque las proteínas de la superficie celular, tales como receptores de crecimiento sobre-expresados en algunos tumores son excelentes dianas para fármacos candidatos o tratamiento de tumores (por ejemplo, cáncer), las mismas proteínas junto con proteínas secretadas codificadas por los genes amplificados en células tumorales son útiles además en el diagnóstico y pronóstico de tumores. Por ejemplo, pueden usarse anticuerpos dirigidos contra los productos proteicos de genes amplificados en células tumorales como diagnósticos o pronósticos de tumores.

Por ejemplo, pueden usarse anticuerpos, incluyendo fragmentos de anticuerpo, para detectar cualitativa o cuantitativamente la expresión de proteínas codificadas por los genes amplificados ("productos de genes marcadores"). El anticuerpo preferentemente está equipado con una marca detectable, por ejemplo, fluorescente, y la unión puede supervisarse mediante microscopía óptica, citometría de flujo, fluorimetría u otras técnicas conocidas en el sector. Estas técnicas son particularmente adecuadas, si el gen amplificado codifica una proteína de la superficie celular, por ejemplo, un factor de crecimiento. Dichos ensayos de unión se realizan esencialmente como se ha descrito en la sección 5 anteriormente.

La detección *in situ* de la unión de anticuerpos a los productos del gen marcador puede realizarse, por ejemplo, mediante inmunofluorescencia o microscopía inmunoelectrónica. Para este fin, se extrae una muestra histológica del paciente, y se aplica un anticuerpo marcado a ésta, preferentemente cubriendo con el anticuerpo una muestra biológica. Este procedimiento también permite determinar la distribución del producto del gen marcador en el tejido examinado. Será evidente para los expertos en la materia que una gran diversidad de procedimientos histológicos están disponibles fácilmente para la detección *in situ*.

Los siguientes ejemplos se ofrecen solamente con fines ilustrativos, y no pretenden limitar en absoluto el alcance de la presente invención.

50 Ejemplos

Los reactivos disponibles en el mercado mencionados en los ejemplos se usaron según las instrucciones del fabricante a menos que se indique otra cosa. La fuente de las células identificadas en los siguientes ejemplos, y en toda la memoria descriptiva, mediante el número de entrada ATCC es la American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209. Todos los depósitos originales mencionados en la presente solicitud se realizaron bajo las condiciones del Tratado de Budapest sobre reconocimiento internacional del depósito de microorganismos para los fines de procedimiento en materia de patentes y las regulaciones en virtud del mismo (Tratado de Budapest). Esto asegura el mantenimiento de un cultivo viable del depósito durante 30 años a partir de la fecha de depósito. El depósito estará disponible en la ATCC bajo los términos del Tratado de Budapest, y sometido a un acuerdo entre Genentech, Inc., y la ATCC, que asegura la disponibilidad permanente y sin restricciones de la progenie del cultivo del depósito al público después de la expedición de la Patente de Estados Unidos pertinente o después de la apertura al público de cualquier Solicitud de patente de Estados Unidos o extranjera, lo que suceda en primer lugar, y asegura la disponibilidad de la progenie a alguien designado para ello por el Comisionado de Patentes y Marcas de Estados Unidos según la disposición 35 USC § 122 las reglas del Comisionado al amparo de la misma (incluyendo la disposición 37 CDR § 1.14 con referencia particular a 886 OG 638).

A menos que se indique otra cosa, la presente invención usa procedimientos convencionales de tecnología de ADN recombinante, tales como las descritas anteriormente en este documento y en los siguientes libros de texto:

- Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press N.Y., 1989; Ausubel *et al.*, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y., 1989; Innis *et al.*, PCR Protocols: A Guide to Methods and Applications, Academic Press, Inc., N.Y., 1990; Harlow *et al.*, Antibodies: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, 1988; Gait, Oligonucleotide Synthesis, IRL Press, Oxford, 1984; R.I. Freshney, Animal Cell Culture, 1987; Coligan *et al.*, Current Protocols in Immunology, 1991.

Ejemplo 1

10 Cribado de Homología del Dominio Extracelular para Identificar Nuevos Polipéptidos y ADNc que los Codifica

Las secuencias de dominio extracelular (ECD) (incluyendo la secuencia señal de secreción, si hay alguna) de aproximadamente 950 proteínas secretadas conocidas de la base de datos pública Swiss-Prot se usaron para buscar bases de datos EST. Las bases de datos EST incluían bases de datos públicas (por ejemplo, GenBank), y bases de datos privadas (por ejemplo LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA). La búsqueda se realizó usando el programa informático BLAST o BLAST-2 (Altschul *et al.*, Methods in Enzymology, 266: 460-480 (1996)) como comparación de las secuencias proteicas ECD con una traducción de 6 marcos de las secuencias EST. Aquellas comparaciones con un valor BLAST de 70 (o en algunos casos 90) o mayor que no codificaban proteínas conocidas se agruparon y ensamblaron en secuencias de ADN consenso con el programa "phrap" (Phil Green, University of Washington, Seattle, Washington).

Usando este cribado de homología de dominio extracelular, se ensamblaron secuencias de ADN consenso con respecto a las demás secuencias EST identificadas usando phrap. Además, Las secuencias de ADN consenso obtenidas se extendieron a menudo (pero no siempre) usando ciclos repetidos de BLAST o BLAST-2 y phrap para extender la secuencia consenso tanto como fuera posible usando las fuentes de secuencias EST descritas anteriormente.

Partiendo de las secuencias consenso obtenidas como se ha descrito anteriormente, a continuación se sintetizaron oligonucleótidos y se usaron para identificar mediante PCR una biblioteca de ADNc que contenía la secuencia de interés y para su uso como sondas para aislar un clon de la secuencia codificante de longitud completa para un polipéptido PRO. Los cebadores de PCR directos e inversos generalmente varían entre 20 y 30 nucleótidos y a menudo se diseñan para dar un producto de PCR de aproximadamente 100-1000 pb de longitud. Las secuencias de sonda son habitualmente de 40-55 pb de longitud. En algunos casos, se sintetizan oligonucleótidos adicionales cuando la secuencia consenso es mayor de aproximadamente 1-1,5 kpb. Para cribar varias bibliotecas para un clon de longitud completa, se cribó ADN de las bibliotecas mediante amplificación por PCR, como Ausubel *et al.*, Current Protocols in Molecular Biology, con el par de cebadores de PCR. A continuación se usó una biblioteca positiva para aislar clones que codifican el gen de interés usando el oligonucleótido de sonda y uno de los pares de cebadores.

Las bibliotecas de ADNc usadas para aislar los clones de ADNc se construyeron mediante procedimientos convencionales usando reactivos disponibles en el mercado tales como los de Invitrogen, San Diego, CA. El ADNc se cebó con oligo dT que contenía un punto NotI, unido mediante adaptadores hemiquinasados blunt-a-SaII, escindidos con NotI, su tamaño se adapta apropiadamente mediante electroforesis en gel, y se clonan en una orientación definida en un vector de clonación adecuado (tal como pRKB o pRKD; pRK5B es un precursor de pRK5D que no contiene el punto Sfil; véase, Holmes *et al.*, Science, 253: 1278-1280 (1991)) en los puntos XhoI y NotI únicos.

Ejemplo 2

Aislamiento de Clones de ADNc Usando Análisis de Algoritmo de Señales

Se identificaron diversas secuencias de ácido nucleico que codifican polipéptidos aplicando un algoritmo de descubrimiento de secuencia señal privado desarrollado por Genentech, Inc., (South San Francisco, CA) sobre EST así como fragmentos de EST agrupados y ensamblados de bases de datos públicas (por ejemplo, GenBank) y/o privadas (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA). El algoritmo de secuencia señal computa un valor de señal de secreción basado en el carácter del los nucleótidos del ADN que rodean al primer y opcionalmente al segundo codón o codones de metionina (ATG) en el extremo 5' de la secuencia o fragmento de secuencia en consideración. Los nucleótidos que siguen al primer ATG deben codificar al menos 35 aminoácidos no ambiguos sin codones de terminación cualesquiera. Si el primer ATG tiene los aminoácidos necesarios, el segundo no se examina. Si ninguno cumple el requisito, la secuencia candidata no se valora. Para determinar si la secuencia EST contiene una auténtica secuencia señal, el ADN y las correspondientes secuencias de aminoácidos que rodean al codón ATG se valoran usando un conjunto de siete sensores (parámetros de evaluación) que se sabe que se asocian con señales de secreción. El uso de este algoritmo dio como resultado la identificación de numerosas secuencias de ácido nucleico que codifican polipéptidos.

Ejemplo 3

Aislamiento de clones de ADNc que codifican PRO3567 humano

Una secuencia de AND consenso se ensambló de forma relativa a otras secuencias EST utilizando phrap como se describe anteriormente en el ejemplo 1. Esta secuencia consenso se designa aquí como DNA52711. Basándose en la secuencia consenso ADN52711 y el clon ES de Merck número AA082750, clon EST de MERck número AA082750 se compró y se obtuvo y se secuenció. Toda la secuencia de nucleótido de la misma se muestra en la figura 1 (SEQ ID NO: 26) y aquí se designa como DNA56049-2543.

El clon DNA56049-2543 contiene un único marco abierto de lectura con un punto de inicio de la traducción evidente en las posiciones de nucleótidos 698-99 y que termina en el codón de terminación en las posiciones de nucleótidos 637-639 (Figura 1). El precursor polipeptídico predicho tiene 180 aminoácidos de longitud. El análisis de la secuencia PRO3567 de longitud completa en la figura 2 (SEQ ID N°: 27) demuestra la presencia de una variedad de dominios polipeptídicos importantes, donde las ubicaciones dadas para los dominios polipeptídicos importantes son aproximadas, tal y como se describe más abajo. El análisis de la secuencia PRO3567 de longitud completa mostrada en la Figura 2 demuestra la presencia de lo siguiente: un péptido señal de aproximadamente el aminoácido 1 a aproximadamente el aminoácido 25; un dominio transmembrana de aproximadamente el aminoácido 149 a aproximadamente el aminoácido 164; puntos de N-glicosilación de aproximadamente el aminoácido 141 a aproximadamente el aminoácido 145; un punto de N-miristoilación de aproximadamente el aminoácido 25 a aproximadamente el aminoácido 31; un sitio de unión de lípido lipoproteína de membrana procariota desde aproximadamente el aminoácido 16 a aproximadamente el aminoácido 27; una secuencia de unión celular desde aproximadamente 112 a aproximadamente el aminoácido 15; y una firma de proteína receptora Ton-B 1 desde aproximadamente el aminoácido 1 a aproximadamente el aminoácido 21. EL clon DNA56049-2543 se depositó en la ATCC el 9 de febrero de 1999 y se le asignó el número de depósito en ATCC. 203662. La proteína PRO3567 de longitud completa mostrada en la figura 2 tiene un peso molecular estimado de aproximadamente 20.313 daltons y un pI de aproximadamente 8,91.

Un análisis de la base de datos Dayhoff (versión 35.45 SwissProt 35), usando un análisis de alineamiento de secuencia WU-BLAST2 de la secuencia de longitud completa mostrada en la Figura 2 (SEC ID N°: 27), mostraba homología entre la secuencia de aminoácidos de PRO3567 y las siguientes secuencias de Dayhoff: SPC2CANFA, SPC2_CHICK, SPC2_CAEEL, AF057144_1, YD2B_SCHPO, S61639, SPC3_YEAST, AMU21992_1, EPU22004_1, y CMU20539_1.

Ejemplo 4

Amplificación génica

Este ejemplo muestra que los genes que codifican PRO se amplifican en el genoma de ciertos cánceres de pulmón, colon y/o de mama y/o líneas celulares humanas. La amplificación se asocia con la sobre-expresión del producto génico, lo que indica que los polipéptidos son dianas útiles para intervención terapéutica en algunos cánceres tales como cánceres de colon, pulmón, mama, paratiroides y otros. Los agentes terapéuticos pueden tomar la forma de antagonistas de polipéptidos PRO, por ejemplo, anticuerpos murinos-humanos quiméricos, humanizados o humanos contra un polipéptido PRO.

El material de partida para el cribado era ADN genómico aislado de diversos cánceres. El ADN se cuantifica de forma precisa, por ejemplo, mediante fluorometría. Como control negativo, se aisló ADN de las células de diez individuos normales sanos que se reunió y usó como controles de ensayo para la copia génica en individuos sanos (no se muestra). El ensayo de nucleasa 5' (por ejemplo, TaqMan™) y PCR cuantitativa en tiempo real (por ejemplo, ABI Prizm 7700 Sequence Detection System™ (Perkin Elmer, Applied Biosystems Division, Foster City, CA)), se usaron para descubrir genes potencialmente amplificados en algunos cánceres. Los resultados se usaron para determinar si el ADN que codifica PRO está sobre-representado en cualquiera de los cánceres de pulmón primario o cáncer de colon o líneas celulares cancerosas o líneas celulares de cáncer de mama que se cribaban. Los cánceres de pulmón primarios se obtuvieron de individuos con tumores del tipo y fase como se indica en la Tabla 4. Una explicación de las abreviaturas usadas para la designación de los tumores primarios enumerados en la Tabla 4 y los tumores primarios y líneas celulares mencionadas en todo este ejemplo se ha proporcionado anteriormente en este documento.

Los resultados del TaqMan™ se presentan en unidades delta (Δ) Ct. Una unidad corresponde a 1 ciclo de PCR o aproximadamente una amplificación de 2 veces con respecto a lo normal, dos unidades corresponden a 4 veces, 3 unidades a una amplificación de 8 veces y así sucesivamente. La cuantificación se obtuvo usando cebadores y una sonda fluorescente TaqMan™ obtenida del gen que codifica PRO, cuyas regiones es más probable que contengan secuencias de ácido nucleico únicas que es menos probable que tengan intrones retirados mediante corte, se prefieren para la derivación de cebador y sonda, por ejemplo, regiones 3' no traducidas. Las secuencias de los cebadores y sondas (directo, inverso y sonda) usadas para el análisis de amplificación génica de PRO eran las siguientes:

PR03567 (DNA56049-2543)

56049.tm.fl:

5 5'-GGAAATGGTCTCAAGGGAAA-3' (SEC ID N°: 83)

92264.tm.r1:

10 5'-TCACTTTGACCCTGTCTTGGAACGTC-3' (SEC ID N°: 84)

92264.tm.p1:

15 5'-GGTAGAATTCCAGCATTTGGTA-3' (SEQ ID N°: 85)

La reacción del ensayo de nucleasa 5' es una técnica basada en PCR fluorescente que hace uso de la actividad exonucleasa 5' de la enzima Taq ADN polimerasa para supervisar la amplificación en tiempo real. Se usan dos cebadores de oligonucleótidos para generar un amplicón típico de una reacción de PCR. Un tercer oligonucleótido, o sonda, se diseña para detectar una secuencia de nucleótidos situada entre los dos cebadores de PCR. La sonda no es extensible mediante enzima Taq ADN polimerasa, y se marca con un tinte fluorescente informador y un tinte fluorescente inactivador. Cualquier emisión inducida por láser del tinte informador es inactivada por el tinte inactivador cuando los dos se sitúan juntos y próximos puesto que están en la sonda. Durante la reacción de amplificación, la enzima Taq ADN polimerasa escinde la sonda de manera dependiente de una plantilla. Los fragmentos de sonda resultantes se disocian en solución, y la señal del tinte informador liberado está libre del efecto inactivador del segundo fluoróforo. Una molécula de tinte informador se libera por cada nueva molécula sintetizada, y la detección del tinte informador sin inactivar proporciona las bases para la interpretación cuantitativa de los datos.

El procedimiento de nucleasa 5' se realiza en un dispositivo de PCR cuantitativa en tiempo real como el ABI Prism 7700TM Sequence Detection. El sistema consiste en un termociclador, láser, dispositivo acoplado por carga (CCD), cámara y ordenador. El sistema amplifica las muestras en un formato de 96 pocillos en un termociclador. Durante la amplificación, la señal fluorescente inducida por láser se recoge en tiempo real a través de cables de fibra óptica para los 96 pocillos, y se detectan en el CCD. El sistema incluye software para controlar el instrumento y analizar los datos.

Los datos del ensayo de nucleasa 5' se expresan inicialmente como Ct, o el ciclo umbral. Esto se define como el ciclo en el que la señal informadora se acumula por encima del nivel de fluorescencia de fondo. Los valores de ΔCt se usan como medición cuantitativa del número relativo de copias de partida de una secuencia diana particular en una muestra de ácido nucleico cuando se compara resultados de ADN de cáncer con resultados de ADN humano normal.

(Tabla pasa a página siguiente)

ES 2 329 689 T3

La Tabla 4 describe la fase, fase T y fase N de diversos tumores primarios que se usaron para cribar el compuesto PRO de la presente invención.

TABLA 4

Perfiles de Tumores de Pulmón y Colon Primarios

Tumor Primario	Fase	Otra Fase	Fase de Dukes	Fase T	Fase N
Tumor de pulmón humano AdenoCa (SRCC724) [LT1]	IIA			T1	N1
Tumor de pulmón humano SqCCa (SRCC725) [LT1a]	IIB			T3	N0
Tumor de pulmón humano AdenoCa (SRCC726) [LT2]	IB			T2	N0
Tumor de pulmón humano AdenoCa (SRCC727) [LT3]	IIIA			T1	N2
Tumor de pulmón humano AdenoCa (SRCC728) [LT4]	IB			T2	N0
Tumor de pulmón humano SqCCa (SRCC729) [LT6]	IB			T2	N0
Tumor de pulmón humano Aden/SqCCa (SRCC730) [LT7]	IA			T1	N0
Tumor de pulmón humano AdenoCa (SRCC731) [LT9]	IB			T2	N0
Tumor de pulmón humano SqCCa (SRCC732) [LT10]	IIB			T2	N1

ES 2 329 689 T3

5	Tumor de pulmón humano SqCCa (SRCC733) [LT11]	IIA	T1	N1
10	Tumor de pulmón humano AdenoCa (SRCC734) [LT12]	IV	T2	N0
15	Tumor de pulmón humano AdenoSqCCa (SRCC735) [LT13]	IB	T2	N0
20	Tumor de pulmón humano SqCCa (SRCC736) [LT15]	IB	T2	N0
25	Tumor de pulmón humano SqCCa (SRCC737) [LT16]	IB	T2	N0
30	Tumor de pulmón humano SqCCa (SRCC738) [LT17]	IIB	T2	N1
35	Tumor de pulmón humano SqCCa (SRCC739) [LT18]	IB	T2	N0
40	Tumor de pulmón humano SqCCa (SRCC740) [LT19]	IB	T2	N0
45				
50				
55				
60				
65				

ES 2 329 689 T3

5	Tumor de pulmón humano LCCa (SRCC741) [LT21]	IIB			T3	N1
10	AdenoCa de pulmón humano (SRCC811) [LT22]	IA			T1	N0
15	AdenoCa de colon humano (SRCC742) [CT2]		M1	D	pT4	N0
20	AdenoCa de colon humano (SRCC743) [CT3]			B	pT3	N0
25	AdenoCa de colon humano (SRCC 744) [CT8]			B	T3	N0
30	AdenoCa de colon humano (SRCC745) [CT10]			A	pT2	N0
35	AdenoCa de colon humano (SRCC746) [CT12]		MO, R1	B	T3	N0
40	AdenoCa de colon humano (SRCC747) [CT14]		pMO, RO	B	pT3	pN0
45	AdenoCa de colon humano (SRCC748) [CT15]		M1, R2	D	T4	N2

ES 2 329 689 T3

5	AdenoCa de colon humano (SRCC749) [CT16]	pMO	B	pT3	pN0
10	AdenoCa de colon humano (SRCC750) [CT17]		C1	pT3	pN1
15	AdenoCa de colon humano (SRCC751) [CT1]	MO, R1	B	pT3	N0
20	AdenoCa de colon humano (SRCC752) [CT4]		B	pT3	M0
25	AdenoCa de colon humano (SRCC753) [CT5]	G2	C1	pT3	pN0
30	AdenoCa de colon humano (SRCC754) [CT6]	pMO, RO	B	pT3	pN0
35	AdenoCa de colon humano (SRCC755) [CT7]	G	A	pT2	pN0
40	AdenoCa de colon humano (SRCC756) [CT9]	G3	D	pT4	pN2
45	AdenoCa de colon humano (SRCC757) [CT11]		B	T3	N0
50	AdenoCa de colon humano (SRCC758) [CT18]	MO, RO	B	pT3	pN0

Preparación de ADN

Se preparó ADN a partir de líneas celulares cultivadas, tumores primarios, y sangre humana normal. El aislamiento se realizó usando un kit de purificación, conjunto de tampón y proteasa y todos de Quiagen, según las instrucciones del fabricante y la descripción posterior.

Lisis del Cultivo Celular

Las células se lavaron y tripsinizaron a una concentración de $7,5 \times 10^8$ por punta y se sedimentaron centrifugando a 1000 rpm durante 5 minutos a 4°C, seguido de lavado de nuevo con 1/2 volumen de PBS y centrifugado de nuevo. Los sedimentos se lavaron una tercera vez, las células suspendidas se recogieron y se lavaron 2 veces con PBS. A continuación, las células se suspendieron en 10 ml de PBS. El tampón C1 se equilibró a 4°C. Se diluyó proteasa # 19155 de Qiagen en 6,25 ml de ddH₂O (agua bidestilada) fría hasta una concentración final de 20 mg/ml y se equilibró a 4°C. Se prepararon 10 ml de Tampón G2 diluyendo la solución madre de ARNasa A de Qiagen (100 mg/ml) hasta una concentración final de 200 µg/ml.

A continuación se añadieron tampón C1 (10 ml, 4°C) y ddH₂O (40 ml, 4°C) a los 10 ml de suspensión celular, se mezclaron invirtiendo los tubos y se incubaron en hielo durante 10 minutos. Los núcleos celulares se sedimentaron centrifugando en un rotor de vasos basculantes de Beckman a 2500 rpm a 4°C durante 15 minutos. El sobrenadante se desechó y los núcleos se suspendieron con agitación en vórtice en 2 ml de Tampón C1 (a 4°C) y 6 ml de ddH₂O, seguido de un segundo centrifugado a 4°C a 2500 rpm durante 15 minutos. A continuación se resuspendieron los núcleos en el tampón residual usando 200 µl por punta. Se añadió tampón G2 (10 ml) a los núcleos suspendidos mientras se aplicaba agitación en vórtice suave. Una vez completada la adición de tampón, se aplicó agitación en vórtice vigorosa durante 30 segundos. Se añadió proteasa de Quiagen (200 µl, preparada como se ha indicado anteriormente) y se incubó a 50°C durante 60 minutos. La incubación y el centrifugado se repitieron hasta que los lisados eran transparentes (por ejemplo, incubando 30-60 minutos adicionales, sedimentando a 3000 x g durante 10 min., 4°C).

Preparación y lisis de una muestra de tumor humano sólido

Se pesaron muestras tumorales y se colocaron en tubos cónicos de 50 ml y se mantuvieron en hielo. El procesamiento estaba limitado a no más de 250 mg de tejido por preparación (1 punta/preparación). La solución de proteasa se preparó recientemente diluyendo en 6,25 ml de ddH₂O fría hasta una concentración final de 20 mg/ml y se almacenó a 4°C. Se preparó tampón G2 (20 ml) diluyendo ADNasa A hasta una concentración final de 200 mg/ml (a partir de 100 mg/ml de solución madre). El tejido tumoral se homogeneizó en 19 ml de tampón G2 durante 60 segundos usando la punta grande del polytron en una campana de flujo laminar para evitar la inhalación de aerosoles, y se mantuvo a temperatura ambiente. Entre muestras, se limpió el polytron centrifugando a 2 x 30 segundos cada vez en 2 l de ddH₂O, seguido de tampón G2 (50 ml). Si seguía estando presente tejido en la punta del generador, el aparato se desmontaba y se limpiaba.

Se añadió proteasa de Quiagen (preparada como se ha indicado anteriormente, 1,0 ml), seguida de agitación en vórtice e incubación a 50°C durante 3 horas. La incubación y el centrifugado se repitieron hasta que los lisados eran transparentes (por ejemplo, incubando 30-60 minutos adicionales, sedimentando a 3000 x g durante 10 min., 4°C).

Preparación y lisis de sangre humana

Se extrajo sangre de voluntarios sanos usando protocolos de agente infeccioso convencionales y se trató con citrato en muestras de 10 ml por punta. Se preparó de nuevo proteasa de Quiagen mediante dilución en 6,25 ml de ddH₂O fría hasta una concentración final de 20 mg/ml y se almacenó a 4°C. Se preparó tampón G2 diluyendo ARNasa A hasta una concentración final de 200 µg/ml a partir de 100 mg/ml de solución madre. La sangre (10 ml) se colocó en un tubo cónico de 50 ml y se añadieron 10 ml de tampón C1 y 30 ml de ddH₂O (ambos equilibrados previamente a 4°C), y los componentes se mezclaron invirtiendo los tubos y se mantuvieron en hielo durante 10 minutos. Los núcleos se sedimentaron con un rotor de vasos basculantes de Beckman a 2500 rpm, 4°C durante 15 minutos y el sobrenadante se desechó. Con una agitación en vórtice, los núcleos se suspendieron en 2 ml de tampón C1 (4°C) y 6 ml de ddH₂O (4°C). La agitación en vórtice se repitió hasta que el sedimento era blanco. A continuación se suspendieron los núcleos en el tampón residual usando una punta de 200 µl. Se añadió tampón G2 (10 ml) a los núcleos suspendidos mientras se agitaba en vórtice suavemente, seguido de agitación en vórtice vigorosa durante 30 segundos. Se añadió proteasas de Quiagen (200 µl) y se incubó a 50°C durante 60 minutos. La incubación y el centrifugado se repitieron hasta que los lisados eran transparentes (por ejemplo, incubando 30-60 minutos adicionales, sedimentando a 3000 x g durante 10 min., 4°C).

*Purificación de lisados depurados**(I) Aislamiento de ADN genómico*

Se equilibró ADN genómico (1 muestra por preparación de punta maxi) con 10 ml de tampón QBT. Se equilibró tampón de elución QF a 50°C. Las muestras se agitaban en vórtice durante 30 segundos, a continuación se cargaron en puntas equilibradas y se drenaron por gravedad. Las puntas se lavaron con 2x 15 ml de tampón QC. El ADN se eluyó en tubos Corex de 30 ml silanizados, autoclavados con 15 ml de tampón QF (50°C). Se añadió isopropanol (10,5

ml) a cada muestra, los tubos se cubrieron con parafina y se mezclaron mediante inversión repetida hasta que el ADN precipitaba. Las muestras se sedimentaron mediante centrifugado en el rotor SS-34 a 15.000 rpm durante 10 minutos a 4°C. La ubicación del sedimento se marcó, el sobrenadante se desechó, y se añadieron 10 ml de etanol al 70% (4°C). Las muestras se sedimentaron de nuevo mediante centrifugado en el rotor SS-34 a 10.000 rpm durante 10 minutos a 4°C. La ubicación del sedimento se marcó y el sobrenadante se desechó. A continuación se colocaron los tubos sobre un lado en una gradilla de secado y se secaron durante 10 minutos a 37°C, teniendo cuidado de no secar en exceso las muestras.

Después del secado, los sedimentos se disolvieron en 1,0 ml de TE (pH 8,5) y se colocaron a 50°C durante 1-2 horas. Las muestras se mantuvieron durante una noche a 4°C mientras continuaba la disolución. La solución de ADN se transfirió a continuación a tubos de 1,5 ml con una aguja del calibre 26 en una jeringa con tuberculina. La transferencia se repitió 5 veces para cortar el ADN. A continuación las muestras se colocaron a 50°C durante 1-2 horas.

(2) Cuantificación de ADN genómico y preparación para el ensayo de amplificación génica

Los niveles de ADN en cada tubo se cuantificaron mediante espectrofotometría A_{260}/A_{280} convencional en una dilución 1:20 (5 μ l de ADN + 95 μ l de ddH₂O) usando las cubetas de cuarzo de 0,1 ml en el espectrofotómetro DU640 de Beckman. Las proporciones de A_{260}/A_{280} estaban en el intervalo de 1,8-1,9. Cada muestra de ADN se diluyó a continuación adicionalmente a aproximadamente 200 ng/ml en TE (pH 8,3). Si el material original estaba muy concentrado (aproximadamente 700 ng/ μ l), el material se dejaba a 50°C durante varias horas hasta que se resuspendía.

La cuantificación fluorimétrica del ADN se realizó a continuación en el material diluido (20-600 ng/ml) usando las directrices del fabricante modificadas como posteriormente. Esto se conseguía permitiendo que un fluorómetro DyNA Quant 200 de Hoeffer se calentara durante aproximadamente 15 minutos. La solución de trabajo de tinte de Hoechst (#H33258, 10 μ l, preparada antes de 12 horas del uso) se diluyó en 100 ml de 1x tampón TNE. Una cubeta de 2 ml se llenó con la solución del fluorómetro, se colocó en la máquina, y la máquina se puso a cero. Se añadió pGEM 3Zf(+) (2 μ l, lote #360851026) a 2 ml de solución del fluorómetro y se calibró a 200 unidades. A continuación se ensayaron 2 μ l adicionales de ADN de pGEM 3Zf(+) y la lectura se confirmó a 400 ± 10 unidades. Cada muestra se leyó a continuación al menos por triplicado. Cuando se descubrió que 3 muestras estaban en un intervalo del 10% unas respecto a otras, se tomó su media y este valor se usó como el valor de cuantificación.

A continuación se usó la concentración determinada fluorométricamente para diluir cada muestra a 10 ng/ μ l en ddH₂O. Esto se realizó simultáneamente en todas las muestras de la plantilla para un único ensayo en placa TaqMan, y con material suficiente para realizar 500-1000 ensayos. Las muestras se ensayaron por triplicado con cebadores Taqman™ y sonda B-actina y GAPDH en una única placa con ADN humano normal y controles sin plantilla. Las muestras diluidas se usaron siempre que el valor CT del ADN humano normal restado del ADN de ensayo fuera ± 1 Ct. El ADN genómico diluido, cuantificado por lotes se almacenó en alícuotas de 1,0 ml a -80°C. Las alícuotas que posteriormente se usarían en el ensayo de amplificación génica se almacenaron a 4°C. Cada alícuota de 1 ml es bastante para 8-9 placas o 64 ensayos.

Ensayo de amplificación génica

Los compuestos PRO de la invención se cribaron en los siguientes tumores primarios y los valores de Δ Ct resultantes se presentan en la Tabla 5.

TABLA 5

Valores de Δ Ct en modelos de tumores primarios de pulmón y colón y de línea celular

Tumores Primarios o Líneas Celulares	PRO3567
Colo320	2,31
SW403	1,07
LS174T	1,17

PRO3567 (DNA56049-2543)

Los valores de Δ Ct para DNA56049-2543 en diversos tumores se presentan en la Tabla 5A. Un Δ Ct de > 1 se usaba habitualmente como el valor umbral para valoración de amplificación, puesto que éste representa una multiplicación por dos de la copia génica. La Tabla 5A indica que la amplificación significativa de ácido nucleico DNA56049-2543 que codifica PRO3567 se producía en líneas celulares de colon: Colo320, SW403 y LS 174T. Puesto que la amplificación de DNA56049-2543 se produce en diversos tumores, es muy probable que juegue un papel significativo

en el crecimiento o la formación tumoral. Como resultado, se esperaría que los antagonistas (por ejemplo, anticuerpos) dirigidos contra la proteína codificada por DNA56049-2543 (PRO3567) tuvieran utilidad en terapia de cáncer.

5 Ejemplo 5

Uso de PRO como sonda de hibridación

10 El siguiente procedimiento describe el uso de una secuencia de nucleótidos que codifica un polipéptido PRO como sonda de hibridación.

El ADN que comprende la secuencia codificante de un polipéptido "PRO" de longitud completa o maduro como se describe en este documento y/o fragmentos del mismo puede emplearse como sonda para cribar ADN homólogos (tales como los que codifican variantes de origen natural de PRO en bibliotecas de ADNc de tejido humano o bibliotecas genómicas de tejido humano).

La hibridación y el lavado de los filtros que contienen cualquier ADN de la biblioteca se realiza en las siguientes condiciones de alta astringencia. La hibridación de una sonda obtenida de PRO radiomarcada con los filtros se realiza en una solución de formamida al 50%, 5x SSC, SDS al 0,1%, pirofosfato sódico al 0,1%, fosfato sódico 50 mM, pH 6,8. 2x Solución de Denhardt, y sulfato de dextrano al 10% a 42°C durante 20 horas. El lavado de los filtros se realiza en una solución acuosa de SSC al 0,1 y SDS al 0,1% a 42°C.

Los ADN que tienen una identidad en la secuencia deseada con el ADN que codifica PRO de secuencia nativa de longitud completa pueden identificarse a continuación usando técnicas convencionales conocidas en el sector.

25

Ejemplo 6

Expresión de Polipéptidos PRO en E. coli

30

Este ejemplo ilustra la preparación de una forma no glucosilada de PRO mediante expresión recombinante en *E. coli*.

La secuencia de ADN que codifica el polipéptido PRO de interés se amplifica inicialmente usando cebadores de PCR seleccionados. Los cebadores deben contener puntos de enzimas de restricción que correspondan a los puntos de enzimas de restricción en el vector de expresión seleccionado. Pueden emplearse diversos vectores de expresión. Un ejemplo de un vector adecuado es pBR322 (obtenido de *E. coli*; véase Bolivar *et al.*, Gene, 2: 95 (1977)) que contiene genes para resistencia a ampicilina y tetraciclina. El vector se digiere con enzimas de restricción y se desfosforila. Las secuencias amplificadas por PCR se ligan a continuación en el vector. El vector incluirá preferentemente secuencias que codifican un gen de resistencia a antibióticos, un promotor trp, un líder poli-His (incluyendo los primeros seis codones STII, secuencia poli-His y punto de escisión de enteroquinasa), la región codificante de PRO, terminador de la transcripción lambda y un gen argU.

La mezcla de ligamiento se usa a continuación para transformar una cepa de *E. coli* seleccionada usando los procedimientos descritos en Sambrook *et al.*, *supra*. Los transformantes se identifican mediante su capacidad para crecer en placas LB y a continuación se seleccionan las colonias resistentes a antibióticos. El ADN plasmídico puede aislarse y confirmarse mediante análisis de restricción y secuenciación de ADN.

Pueden cultivarse clones seleccionados durante una noche en medio de cultivo líquido tal como caldo LB suplementado con antibióticos. El cultivo durante una noche puede usarse posteriormente para inocular un cultivo a mayor escala. Las células se cultivan a continuación hasta una densidad óptica deseada, durante lo cual se activa el promotor de expresión.

Después de cultivar las células durante varias horas más, las células pueden recogerse mediante centrifugado. El sedimento celular obtenido mediante centrifugado puede solubilizarse usando diversos agentes conocidos en el sector, y la proteína PRO solubilizada puede purificarse a continuación usando una columna quelante de metales en condiciones que permitan la unión fuerte de la proteína.

PRO puede expresarse con éxito en *E. coli* en una forma marcada con poli-His usando el siguiente procedimiento. El ADN que codifica PRO se amplifica inicialmente usando cebadores de PCR seleccionados. Los cebadores contienen puntos de enzimas de restricción que corresponden a los puntos de enzimas de restricción en el vector de expresión seleccionado, y otras secuencias útiles que posibilitan el inicio de la traducción eficaz y fiable, purificación rápida en una columna quelante de metales, y la retirada proteolítica con enteroquinasa. Las secuencias amplificadas por PCR, marcadas con poli-His se ligan a continuación en un vector de expresión, que se usó para transformar un huésped *E. coli* basado en la cepa 52 (W3110fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq)). Los transformantes se cultivan en primer lugar en LB que contiene 50 mg/ml de carbenicilina a 30°C con agitación hasta que se alcanzó una D.O. de 3-5 a 600 nm. Los cultivos se diluyen a continuación 50-100 veces en medio CRAP (preparado mezclando 3,57 g de (NH₄)₂SO₄, 0,71 g de citrato sódico-2H₂O, 1,07 g de KCl, 5,36 g de extracto de levadura Difco, 5,36 g de hicasa

SF de Sheffield en 500 ml de agua, así como MPOS 110 mM, pH 7,3, glucosa al 0,55% (p/v) y MgSO_4 7 mM) y se cultivan durante aproximadamente 20-30 horas a 30°C con agitación. Las muestras se retiran para verificar la expresión mediante análisis de SDS-PAGE, y el cultivo total se centrifugó para sedimentar las células. Los sedimentos celulares se congelan hasta la purificación y replegamiento.

La pasta de *E. coli* de fermentaciones de 0,5 a 1 l (sedimentos de 6-10 g) se resuspende en 10 volúmenes (p/v) en guanidina 7 M, Tris 20 mM, tampón a pH 8. Se añaden sulfito sódico sólido y tetratiónato sódico para preparar concentraciones finales de 0,1 M y 0,02 M, respectivamente, y la solución se agita durante una noche a 4°C. Esta etapa da como resultado una proteína desnaturalizada con todos los restos cisteína bloqueados mediante sulfitolización. La solución se centrifuga a 40.000 rpm en una Ultracentrífuga de Beckman durante 30 min. El sobrenadante se diluye con 3-5 volúmenes de tampón de columna de quelato metálico (guanidina 6 M, Tris 20 mM, pH 7,4) y se filtra a través de filtros de 0,22 micrómetros para purificarlo. El extracto purificado se carga en una columna de quelato metálico Ni 2+-NTA de 5 ml de Qiagen equilibrada en el tampón de columna de quelato metálico. La columna se lava con tampón adicional que contiene imidazol 50 mM (Calbiochem, grado Utról), pH 7,4. Las proteínas se eluyen con tampón que contiene imidazol 250 mM. Las fracciones que contienen la proteína deseada se reúnen y almacenan a 4°C. La concentración de proteína se estima mediante su absorbancia a 280 nm usando el coeficiente de extinción calculado en base a su secuencia de aminoácidos.

Las proteínas se repliegan diluyendo la muestra lentamente en tampón de replegamiento recién preparado que consiste en: Tris 20 mM, pH 8,6, NaCl 0,3 M, urea 2,5 M, cisteína 5 mM, glicina 20 mM y EDTA 1 mM. Los volúmenes de replegamiento se seleccionan de modo que la concentración final de proteína estaba entre 50 y 100 microgramos/ml. La solución de replegamiento se agita suavemente a 4°C durante 12-36 horas. La reacción de replegamiento se interrumpe mediante la adición de TFA hasta una concentración final del 0,4% (pH de aproximadamente 3). Antes de la purificación adicional de la proteína, la solución se filtra a través de un filtro de 0,22 micrómetros y se añade acetonitrilo hasta una concentración final del 2-10%. La proteína replegada se cromatografía en una columna Poros RI/H de fase inversa usando un tampón móvil de TFA al 0,1% con elución con un gradiente de acetonitrilo del 10 al 80%. Alícuotas de fracciones con absorbancia A_{280} se analizan en geles SDS de poliacrilamida y las fracciones que contienen proteína replegada homogénea se reúnen. Generalmente, las especies replegadas apropiadamente de la mayoría de las proteínas se eluyen a las concentraciones más bajas de acetonitrilo puesto que esas especies son las más compactas con sus interiores hidrofóbicos protegidos de la interacción con la resina de fase inversa. Las especies agregadas habitualmente se eluyen a concentraciones más altas de acetonitrilo. Además de resolver formas plegadas erróneamente de proteínas de la forma deseada, la etapa de fase inversa también retira endotoxina de las muestras.

Las fracciones que contienen las proteínas PRO plegadas deseadas se reúnen y el acetonitrilo se retira usando un suave chorro de nitrógeno dirigido a la solución. Las proteínas se formulan en Hepes 20 mM, pH 6,8 con cloruro sódico 0,14 M y manitol al 4% mediante diálisis o mediante filtración en gel usando resinas G25 Superfine (Pharmacia) equilibradas en el tampón de formulación y filtradas de forma estéril.

Ejemplo 7

Expresión de PRO en células de mamífero

Este ejemplo ilustra la preparación de una forma potencialmente glucosilada de PRO mediante expresión recombinante en células de mamífero.

El vector, pRK5 (véase EP 307.247, publicada el 15 de marzo de 1989), se emplea como vector de expresión. Opcionalmente, el ADN de PRO se liga en pRK5 con enzimas de restricción seleccionadas para permitir la inserción del ADN de PRO usando procedimientos de ligamiento tales como los descritos en Sambrook *et al.*, *supra*. El vector resultante se denomina pRK5-PRO.

En una realización, las células huésped seleccionadas pueden ser células 293. Células 293 humanas (ATCC CCL 1573) se cultivan hasta confluencia en placas de cultivo tisular en medio tal como DMEM suplementado con suero fetal de ternero y, opcionalmente, componentes nutrientes y/o antibióticos. Aproximadamente 10 µg de ADN de pRK5-PRO se mezclan con aproximadamente 1 µg de ADN que codifica el gen ARN de VA (Thimmappaya *et al.*, Cell, 31: 543 (1982)) y se disuelven en 500 µl de Tris-HCl 1 mM, EDTA 0,1 mM, CaCl_2 0,227 M. A esta mezcla se añaden, gota a gota, 500 µl de HEPES 50 mM (pH 7,35), NaCl 280 mM, NaPO_4 1,5 mM, y se deja que se forme un precipitado durante 10 minutos a 25°C. El precipitado se suspende y se añade a las células 293 y se deja sedimentar durante aproximadamente cuatro horas a 37°C. El medio de cultivo se retira mediante aspiración y se añaden 2 ml de glicerol al 20% en PBS durante 30 segundos. Las células 293 se lavan a continuación con medio libre de suero, se añade medio recién preparado y las células se incuban durante aproximadamente 5 días.

Aproximadamente 24 horas después de las transfecciones, el medio de cultivo se retira y se sustituye por medio de cultivo (en solitario) o medio de cultivo que contiene 200 µCi/ml de ^{35}S -cisteína y 200 µCi/ml de ^{35}S -metionina. Después de una incubación de 12 horas, el medio acondicionado se recoge, se concentra en un filtro rotatorio, y se carga en un gel de SDS al 15%. El gel procesado puede secarse y exponerse a una película durante un periodo de tiempo seleccionado para revelar la presencia del polipéptido PRO. Los cultivos que contienen células transfectadas pueden experimentar incubación adicional (en medio libre de suero) y el medio se ensaya en bioensayos seleccionados.

En una técnica alternativa, el ADN de PRO puede introducirse en células 293 transitoriamente usando el procedimiento de sulfato de dextrano descrito por Somparyrac *et al.*, Proc. Natl. Acad. Sci., 12: 7575 (1981). Las células 293 se cultivan hasta densidad máxima en un matraz de agitación y se añaden 700 µg de ADN de pRK5-PRO. Las células se concentran en primer lugar a partir del matraz de agitación mediante centrifugado y se lavan con PBS. El precipitado de ADN-dextrano se incuba en el sedimento celular durante cuatro horas. Las células se tratan con glicerol al 20% durante 90 segundos, se lavan con medio de cultivo tisular, y se reintroducen en el matraz de agitación que contiene medio de cultivo tisular, 5 µg/ml de insulina bovina y 0,1 µg/ml de transferrina bovina. Después de aproximadamente cuatro días, el medio acondicionado se centrifuga y se filtra para retirar células y restos celulares. La muestra que contiene PRO expresado puede concentrarse y purificarse a continuación mediante cualquier procedimiento seleccionado, tal como diálisis y/o cromatografía en columna.

En otra realización puede expresarse PRO en células CHO. El vector pRK5-PRO puede transfectarse en células CHO usando reactivos conocidos tales como CaPO₄ o DEAE-dextrano. Como se ha descrito anteriormente, los cultivos celulares pueden incubarse, y el medio sustituirse por medio de cultivo (en solitario) o medio que contiene una radiomarca tal como ³⁵S-metionina. Después de determinar la presencia del polipéptido PRO, el medio de cultivo puede sustituirse por medio libre de suero. Preferentemente, los cultivos se incuban durante aproximadamente 6 días, y a continuación se recoge el medio acondicionado. El medio que contiene el PRO expresado puede concentrarse y purificarse a continuación mediante cualquier procedimiento seleccionado.

También puede expresarse PRO marcado con epítipo en células huésped CHO. El PRO puede subclonarse a partir del vector pRK5. El inserto de subclón puede someterse a PCR para fusionarse en marco con una marca epitópica seleccionada tal como una marca poli-His en un vector de expresión en Baculovirus. El inserto de PRO marcado con poli-His puede subclonarse a continuación en un vector controlado por SV40 que contiene un marcador de selección tal como DHFR para la selección de clones estables. Finalmente, las células CHO pueden transfectarse (como se ha descrito anteriormente) con el vector controlado por SV40. El marcado puede realizarse, como se ha descrito anteriormente, para verificar la expresión. El medio de cultivo que contiene el PRO marcado con poli-His expresado puede concentrarse y purificarse a continuación mediante cualquier procedimiento seleccionado, tal como mediante cromatografía de afinidad de Ni²⁺-quelado. La expresión en células CHO y/o COS también puede conseguirse mediante un procedimiento de expresión transitoria.

PRO puede expresarse en células CHO mediante un procedimiento de expresión estable. La expresión estable en células CHO se realiza usando el siguiente procedimiento. Las proteínas se expresan como una construcción de IgG (inmunoadhesina), en la que las secuencias que codifican las formas solubles (por ejemplo, dominios extracelulares) de las respectivas proteínas se fusionaron a una secuencia de región constante de IgG1 que contenía los dominios bisagra, CH2 y CH2 y/o en una forma marcada con poli-His.

Después de la amplificación por PCR, los respectivos ADN se subclonan en un vector de expresión en CHO usando técnicas convencionales como se describe en Ausubel *et al.*, Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). Los vectores de expresión en CHO se construyen para que tengan puntos de restricción compatibles 5' y 3' del ADN de interés para permitir el traslado conveniente de ADNc. El vector usado para la expresión en células CHO es como se describe en Lucas *et al.*, Nucl. Acids Res., 24: 9 (1774-1779 (1996), y usa el promotor temprano de SV40/potenciador para regular la expresión del ADNc de interés y de dihidrofolato reductasa (DHFR). La expresión de DHFR permite la selección para el mantenimiento estable del plásmido después de la transfección.

Doce microgramos del ADN plasmídico deseado se introducen en aproximadamente 10 millones de células CHO usando reactivos de transfección disponibles en el mercado Superfect® (Quiagen), Dosper® o Eugene® (Boehringer Mannheim). Las células se cultivan como se describe en Lucas *et al.*, *supra*. Aproximadamente 3 x 10⁻⁷ células se congelan en una ampolla para un crecimiento y producción adicionales como se describe posteriormente.

Las ampollas que contienen el ADN plasmídico se descongelan colocándolas en un baño de agua y se mezclan agitando en vórtice. El contenido se pipetea en un tubo de centrifuga que contiene 10 ml de medio y se centrifuga a 1000 rpm durante 5 minutos. El sobrenadante se aspira y las células se resuspenden en 10 ml de medio selectivo (PS20 filtrado a 0,2 µm con el 5% de suero fetal bovino diafiltrado a 0,2 µm). Las células se dividen a continuación en alícuotas en un tubo de centrifuga de 100 ml que contiene 90 ml de medio selectivo. Después de 1-2 días, las células se transfectan a un tubo de centrifuga de 250 ml cargado con 150 ml de medio de cultivo selectivo y se incuban a 37°C. Después de otros 2-3 días, se siembran tubos de centrifuga de 250 ml, 500 ml y 2000 ml con 3 x 10⁵ células/ml. El medio con las células se intercambia con medio recién preparado mediante centrifugado y resuspensión en medio de producción. Puede emplearse cualquier medio adecuado para CHO, un tubo de centrifuga de producción de 3 l se siembra a 1,2 x 10⁶ células/ml. El día 0, se determinan el número de células y el pH. El día 1, se muestrea el tubo de centrifuga y se inicia el burbujeo con aire filtrado. El día 2, se muestrea el tubo de centrifuga, la temperatura se cambia a 33°C, y se añaden 30 ml de 500 g/l de glucosa y 0,6 ml de antiespuma al 10% (por ejemplo, emulsión de polidimetilsiloxano al 35%, Dow Corning 365 grado médico). En toda la producción, el pH se ajusta según sea necesario para mantenerlo a aproximadamente 7,2. Después de 10 días, o hasta que la viabilidad caiga por debajo del 70%, el cultivo celular se recoge mediante centrifugado y se filtra a través de un filtro de 0,22 µm. El filtrado se almacena a 4°C o se carga inmediatamente en columnas para la purificación.

Para las construcciones marcadas con poli-His, las proteínas se purifican usando una columna Ni²⁺-NTA (Qiagen). Antes de la purificación, se añade imidazol al medio acondicionado hasta una concentración de 5 mM. El medio acondicionado se bombea a una columna Ni²⁺-NTA de 6 ml equilibrada en Hepes 20 mM, pH 7,4, tampón que contiene NaCl 0,3 M e imidazol 5 mM a un caudal de 4-5 ml/min. a 4°C. Después de la carga, la columna se lava con tampón de equilibrado adicional y la proteína se eluye con tampón de equilibrado que contiene imidazol 0,25 M. La proteína altamente purificada se desala posteriormente en un tampón de almacenamiento que contiene Hepes 10 mM, NaCl 0,14 M y manitol al 4%, pH 6,8, con una columna G25 Superfine (Pharmacia) de 25 ml y se almacena a -80°C.

Las construcciones de inmunoadhesina (que contienen Fc) se purifican del medio acondicionado de la siguiente manera. El medio acondicionado se bombea en una columna Proteína A de 5 ml (Pharmacia) que se había equilibrado en tampón fosfato de Na 20 mM, pH 6,8. Después de la carga, la columna se lava exhaustivamente con tampón de equilibrado antes de la elución con ácido cítrico 100 mM, pH 3,5. La proteína eluida se neutraliza inmediatamente recogiendo fracciones de 1 ml en tubos que contienen 275 µl de tampón Tris 1 M, pH 9. La proteína altamente purificada se desala posteriormente en tampón de almacenamiento como se ha descrito anteriormente para las proteínas marcadas con poli-His. La homogeneidad se evalúa mediante geles de SDS poliacrilamida y mediante secuenciación de aminoácido N-terminal mediante degradación de Edman.

Ejemplo 8

Expresión de PRO en Levadura

El siguiente procedimiento describe la expresión recombinante de PRO en levaduras.

En primer lugar, se construyen vectores de expresión en levadura para la producción o secreción intracelular de PRO a partir del promotor ADH2/GAPDH. El ADN que codifica PRO y el promotor se inserta en puntos de enzimas de restricción adecuados en el plásmido seleccionado para dirigir la expresión intracelular de PRO. Para la secreción, puede clonarse ADN que codifica PRO en el plásmido seleccionado, junto con ADN que codifica el promotor ADH2/GAPDH, un péptido señal de PRO nativo u otro péptido señal de mamífero, o, por ejemplo, un factor alfa o señal secretora de invertasa/secuencia líder de levadura, y secuencias enlazadoras (si fuera necesario) para la expresión de PRO.

Células de levadura, tales como la cepa de levadura AB110, pueden transformarse a continuación con los plásmidos de expresión descritos anteriormente y cultivarse en medios de fermentación seleccionados. Los sobrenadantes de levadura transformados pueden analizarse mediante precipitación con ácido tricloroacético al 10% y separación mediante SDS-PAGE, seguida de tinción de los geles mediante tinción con Azul de Coomassie.

Posteriormente el PRO recombinante puede aislarse y purificarse retirando las células de levadura del medio de fermentación mediante centrifugado y a continuación concentrando el medio usando filtros de cartucho seleccionados. El concentrado que contiene PRO puede purificarse adicionalmente usando resinas de cromatografía en columna seleccionadas.

Ejemplo 9

Expresión de PRO en Células de Insecto infectadas por Baculovirus

El siguiente procedimiento describe la expresión recombinante en células de insecto infectadas por Baculovirus.

La secuencia que codifica PRO se fusiona cadena arriba de una marca epitópica contenida dentro de un vector de expresión en baculovirus. Dichas marcas epitópicas incluyen marcas poli-His y marcas de inmunoglobulina (como regiones Fc de IgG). Pueden emplearse diversos plásmidos, incluyendo plásmidos obtenidos de plásmidos disponibles en el mercado tales como pVL1393 (Novagen). En resumen, la secuencia que codifica PRO o la parte deseada de la secuencia codificante de PRO [tal como la secuencia que codifica el dominio extracelular de una proteína transmembrana o la secuencia que codifica la proteína madura si la proteína es extracelular] se amplifica mediante PCR con cebadores complementarios a las regiones 5' y 3'. El cebador 5' puede incorporar puntos de enzimas de restricción flanqueantes (seleccionados). El producto se digiere a continuación con aquellas enzimas de restricción seleccionadas y se subclona en el vector de expresión.

El baculovirus recombinante se genera co-transfectando el plásmido anterior y ADN de virus BaculoGold™ (Pharmingen) en células de *Spodoptera frugiperda* ("Sf9") (ATCC CRL 1711) usando lipofectina (disponible en el mercado de GIBCO-BRL). Después de 4-5 días de incubación a 28°C, los virus liberados se recogen y se usan para ampliificaciones adicionales. La infección vírica y la expresión de proteína se realizan según lo descrito por O'Reilly *et al.*, Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

El PRO marcado con poli-His expresado puede purificarse a continuación, por ejemplo, mediante cromatografía de afinidad por Ni²⁺-quelato de la siguiente manera. Los extractos se preparan a partir de células Sf9 infectadas por virus recombinantes según lo descrito por Rupert *et al.*, Nature, 362: 175-179(1993). En resumen, células Sf9 se lavan,

se resuspenden en tampón de sonicación (25 ml de Hepes, pH 7,9; MgCl₂ 12,5 mM; EDTA 0,1 mM; glicerol al 10%; NP-40 al 0,1%; KCl 0,4 M), y se sonicán dos veces durante 20 segundos en hielo. Los sonicados se eliminan mediante centrifugado, y el sobrenadante se diluye 50 veces en tampón de carga (fosfato 50 mM, NaCl 300 mM, glicerol al 10%, pH 7,8) y se filtran a través de un filtro de 0,45 μ m. Una columna de agarosa Ni²⁺-NTA (disponible en el mercado de Qiagen) se prepara con un volumen de lecho de 5 ml, se lava con 25 ml de agua y se equilibra con 25 ml de tampón de carga. El extracto de células filtrado se carga en la columna a 0,5 ml por minuto. La columna se lava hasta el valor de referencia A₂₈₀ con tampón de carga, momento en que se inicia la recogida de fracciones. A continuación, la columna se lava con un tampón de lavado secundario (fosfato 50 mM; NaCl 300 mM, glicerol al 10%, pH 6,0), que eluye proteína unida de forma no específica. Después de alcanzar el valor de referencia A₂₈₀ de nuevo, la columna se desarrolla con un gradiente de imidazol de 0 a 500 mM en el tampón de lavado secundario. Se recogen fracciones de un ml y se analizan mediante SDS-PAGE y tinción de plata o transferencia de Western con Ni²⁺-NTA conjugado a fosfatasa alcalina (Qiagen). Las fracciones que contienen el PRO marcado con His₁₀ eluido respectivamente, se reúnen y dializan contra tampón de carga.

Como alternativa, la purificación del PRO marcado con IgG (o marcado con Fc) puede realizarse usando técnicas de cromatografía conocidas, incluyendo por ejemplo, cromatografía en columna de Proteína A o Proteína G.

Aunque la expresión realmente se realiza a una escala de 0,5-2 l, su escala puede aumentarse fácilmente para preparaciones más grandes (por ejemplo, 8 l). Las proteínas se expresan como una construcción de IgG (inmunoadhesina), en la que la región extracelular de la proteína se fusiona a una secuencia de región constante de IgG1 que contiene los dominios bisagra, CH2 y CH3 y/o en formas marcadas con poli-His.

Después de la amplificación por PCR, las respectivas secuencias codificantes se subclonan en un vector de expresión en Baculovirus (pb.PH.IgG para fusiones de IgG y pb.PH.His.c para proteínas marcadas con poli-His), y el vector y el ADN de baculovirus Baculogold[®] (Pharmingen) se co-transfectan en 10⁵ células de *Spodoptera frugiperda* ("Sf9") (ATCC CRL 1711), usando Lipofectina (Gibco BRL). pb.PH.IgG y pb.PH.His son modificaciones del vector de expresión en baculovirus disponible en el mercado pVL 1393 (Pharmingen), con regiones polienlazadoras modificadas para incluir las secuencias marca His o Fc. Las células se cultivan en medio TNM-FH de Hink suplementado con FBS al 10% (Hyclone). Las células se incuban durante 5 días a 28°C. El sobrenadante se recoge y posteriormente se usa para la primera amplificación viral infectando células Sf9 en medio TNM-FH de Hink suplementado con FBS al 10% a una multiplicidad de infección (MOI) aproximada de 10. Las células se incuban durante 3 días a 28°C. El sobrenadante se recoge y la expresión de las construcciones en el vector de expresión en baculovirus se determina mediante la unión por lotes de 1 ml de sobrenadante a 25 ml de perlas de Ni²⁺-NTA (QIAGEN) para proteínas marcadas con histidina o perlas de Proteína A Sepharose CL-4B (Pharmacia) para proteínas marcadas con IgG seguido de análisis de SDS-PAGE en comparación con una concentración conocida de patrón proteico mediante tinción con azul de Coomassie.

El primer sobrenadante de amplificación viral se usa para infectar un cultivo en tubo de centrífuga (500 ml) de células Sf9 cultivadas en medio ESF-921 (Expression Systems LLC) a una MOI aproximada de 0,1. Las células se incuban durante 3 días a 28°C. El sobrenadante se recoge y se filtra. Se repiten análisis de unión por lotes y SDS-PAGE, según sea necesario, hasta que se confirme la expresión del cultivo del tubo de centrífuga.

El medio acondicionado de las células transfectadas (0,5 a 3 l) se recoge mediante centrifugado para retirar las células y se filtra a través de filtros de 0,22 micrómetros. Para las construcciones marcadas con poli-His, la construcción de proteína se purifica usando una columna de Ni²⁺-NTA (Qiagen). Antes de la purificación, se añade imidazol al medio acondicionado hasta una concentración de 5 mM. El medio acondicionado se bombea en una columna de Ni²⁺-NTA de 6 ml equilibrada en Hepes 20 mM, pH 7,4, tampón que contiene NaCl 0,3 M e imidazol 5 mM a un caudal de 4-5 ml/min. a 4°C. Después de la carga, la columna se lava con tampón de equilibrado adicional y la proteína se eluye con tampón de equilibrado que contiene imidazol 0,25 M. La proteína altamente purificada se desala posteriormente en un tampón de almacenamiento que contiene Hepes 10 mM, NaCl 0,14 M y manitol al 4%, pH 6,8, con una columna G25 Superfine (Pharmacia) de 25 ml y se almacena a -80°C.

Las construcciones de inmunoadhesina (que contienen Fc) de proteínas se purifican a partir del medio acondicionado de la siguiente manera, el medio acondicionado se bombea en una columna de Proteína A (Pharmacia) de 5 ml que se había equilibrado en tampón fosfato de Na 20 mM, pH 6,8. Después de la carga, la columna se lava exhaustivamente con tampón de equilibrado antes de la elución con ácido cítrico 100 mM, pH 3,5. La proteína eluida se neutraliza inmediatamente recogiendo fracciones de 1 ml en tubos que contienen 275 ml de tampón Tris 1 M, pH 9. La proteína altamente purificada se desala posteriormente en tampón de almacenamiento como se ha descrito anteriormente para las proteínas marcadas con poli-His. La homogeneidad de las proteínas se verifica mediante electroforesis en gel de SDS poliacrilamida (PEG) y secuenciación del aminoácido N-terminal mediante degradación de Edman.

Como alternativa, puede usarse un procedimiento de baculovirus modificado que incorpora células High 5. En este procedimiento, el ADN que codifica la secuencia deseada se amplifica con sistemas adecuados, tales como Pfu (Stratagene), o se fusiona cadena arriba (5') de una marca epitópica contenida en un vector de expresión en baculovirus. Dichas marcas epitópicas incluyen marcas poli-His y marcas de inmunoglobulina (como regiones Fc de IgG). Pueden emplearse diversos plásmidos, incluyendo plásmido obtenidos de plásmidos disponibles en el mercado tales como pIE1-1 (Novagen). Los vectores pIE1-1 y pIE1-2 se diseñan para la expresión constitutiva de proteínas recombinantes de los promotores del de baculovirus en células de insecto transformadas de forma estable. Los plásmidos solamente

difieren en la orientación de los múltiples puntos de clonación y contienen todas las secuencias promotoras que se sabe que son importantes para la expresión génica mediada por el *in* en células de insecto no afectadas así como el elemento potenciador *hr5*. *pIE1-1* y *pIE1-2* incluyen el punto de inicio de la traducción y pueden usarse para producir proteínas de fusión. En resumen, la secuencia deseada o la parte deseada de la secuencia (tal como la secuencia que codifica el dominio extracelular de una proteína transmembrana) se amplifica por PCR con cebadores complementarios a las regiones 5' y 3'. El cebador 5' puede incorporar puntos de enzimas de restricción flanqueantes (seleccionados). El producto se digiere a continuación con aquellas enzimas de restricción seleccionadas y se subclona en el vector de expresión. Por ejemplo, los derivados de *pIE1-1* pueden incluir la región Fc de IgG humana (*pb.PH.IgG*) o una marca de 8 histidinas (*pb.PH.His*) cadena abajo (3' de) la secuencia deseada. Preferentemente, la construcción del vector se realiza para la confirmación.

Células High 5 se cultivan hasta una confluencia del 50% en las condiciones de 27°C, sin CO₂, sin pen./estrep. Para cada placa de 150 mm, 30 µg de vector basado en *pIE* que contiene la secuencia se mezclan con 1 ml de medio Ex-Cell (Medio: Ex-Cell 401 + 1/100 L-Glu JRH Biosciences #14401-78P (nota: este medio es fotosensible)), y en un tubo diferente, 100 µl de CellFectin (CellFECTIN (GibcoBRL #10362-010) (agitado en vórtice para mezclarlo)) se mezcla con 1 ml de medio Ex-Cell. Las dos soluciones se combinan y se dejan incubar a temperatura ambiente durante 15 minutos. Se añaden 8 ml de medio Ex-Cell a los 2 ml de la mezcla de ADN/CellFECTIN y esto se deposita en capas en células High 5 que se habían lavado con medio Ex-Cell. La placa se incuba a continuación en la oscuridad durante 1 hora a temperatura ambiente. La mezcla de ADN/CellFECTIN se aspira a continuación, y las células se lavan una vez con Ex-Cell para retirar el exceso de CellFECTIN, se añaden 30 ml de medio Ex-Cell recién preparado y las células se incuban durante 3 días a 28°C. El sobrenadante se recoge y la expresión de la secuencia en el vector de expresión en baculovirus se determina mediante unión por lotes de 1 ml de sobrenadante a 25 ml de perlas de Ni²⁺-NTA (QIAGEN) para proteínas marcadas con histidina o perlas de Proteína A Sepharose CL-4B (Pharmacia) para proteínas marcadas con IgG seguido de análisis por SDSPAGE en comparación con una concentración conocida de patrón proteico mediante tinción con azul de Coomassie.

El medio acondicionado a partir de las células transfectadas (de 0,5 a 3 l) se recoge mediante centrifugado para retirar las células y se filtra a través de filtros de 0,22 micrómetros. Para las construcciones marcadas con poli-His, la proteína que comprende la secuencia se purifica usando una columna Ni²⁺-NTA (Qiagen). Antes de la purificación, se añade imidazol al medio acondicionado hasta una concentración de 5 mM. El medio acondicionado se bombea en una columna Ni²⁺-NTA de 6 ml equilibrada en Hepes 20 mM, pH 7,4, tampón que contiene NaCl 0,3 M e imidazol 5 mM a un caudal de 4-5 ml/min. a 48°C. Después de la carga, la columna se lava con tampón de equilibrio adicional y la proteína se eluye con tampón de equilibrio que contiene imidazol 0,25 M. La proteína altamente purificada se desala posteriormente a continuación en un tampón de almacenamiento que contiene Hepes 10 mM, NaCl 0,14 M y manitol al 4%, pH 6,8, con una columna G25 Superfine (Pharmacia) de 25 ml y se almacena a -80°C.

Construcciones de inmunoadhesina (que contienen Fc) de proteínas se purifican del medio acondicionado de la siguiente manera. El medio acondicionado se bombea en una columna de Proteína A de 5 ml (Pharmacia) que se ha equilibrado en tampón fosfato de Na 20 mM, pH 6,8. Después de la carga, la columna se lava exhaustivamente con tampón de equilibrio antes de la elución con ácido cítrico 100 mM, pH 3,5. La proteína eluida se neutraliza inmediatamente recogiendo fracciones de 1 ml en tubos que contienen 275 ml de tampón Tris 1 M, pH 9. La proteína altamente purificada se desala posteriormente en tampón de almacenamiento como se ha descrito anteriormente para las proteínas marcadas con poli-His. La homogeneidad de la secuencia se evalúa mediante geles de SDS poliacrilamida y mediante secuenciación del aminoácido N-terminal mediante degradación de Edman y otros procedimientos analíticos según se desee o sea necesario.

PRO381, PRO1410 y PRO4354 se expresaron con éxito mediante el procedimiento en baculovirus modificado anterior que incorpora células High 5.

Ejemplo 10

Preparación de Anticuerpos que se unen a PRO

Este ejemplo la ilustra preparación de anticuerpos monoclonales que pueden unirse específicamente a PRO.

En el sector se conocen técnicas para producir los anticuerpos monoclonales y se describen, por ejemplo, en Goding, *supra*. Los inmunógenos que pueden emplearse incluyen proteínas de fusión a PRO purificadas que contienen PRO y células que expresan PRO recombinante en la superficie celular. La selección del inmunógeno puede realizarla el experto en la materia sin gran experimentación.

Se inmunizan ratones, tales como Balb/c, con el inmunógeno PRO emulsionado en adyuvante completo de Freund y se inyectan por vía subcutánea o por vía intraperitoneal en una cantidad de 1-100 microgramos. Como alternativa, el inmunógeno se emulsiona en adyuvante MPL-TDM (Ribi Immunochemical Research, Hamilton, MT) y se inyecta en las almohadillas plantares de la pata trasera del animal. Los ratones inmunizados reciben a continuación un refuerzo de 10 a 12 días después con inmunógeno adicional emulsionado en el adyuvante seleccionado. En adelante, durante varias semanas, Los ratones también pueden recibir refuerzos con inyecciones de inmunización adicionales. Pueden

obtenerse periódicamente muestras de suero de los ratones mediante extracción de sangre retro-orbital para ensayarla en ensayos ELISA para detectar anticuerpos anti-PRO.

Una vez que se ha detectado un valor cuantitativo de anticuerpos adecuado, a los animales "positivos" para anticuerpos se les puede inyectar una inyección intravenosa final de PRO. Tres o cuatro días después, los ratones se sacrifican y se recogen las células esplénicas. Las células esplénicas se fusionan a continuación (usando polietilenglicol al 35%) con una línea celular de mieloma múrida tal como P3X63AgU.1, disponible de ATCC. No. CRL 1597. Las fusiones generan células de hibridoma que a continuación pueden colocarse en placas de cultivo tisular de 96 pocillos que contienen medio HAT (hipoxantina, aminopterina y timidina) para inhibir la proliferación de células no fusionadas, híbridos de mieloma, e híbridos de células esplénicas.

Las células de hibridoma se cribarán en un ELISA para la reactividad contra PRO. La determinación de células de hibridoma "positivas" que secretan los anticuerpos monoclonales deseados contra PRO está dentro del alcance de un experto en la materia.

Las células de hibridoma positivas pueden inyectarse por vía intraperitoneal en ratones Balb/c singénicos para producir ascitis que contienen los anticuerpos monoclonales anti-PRO. Como alternativa, las células de hibridoma pueden cultivarse en matraces de cultivo tisular o frascos rotatorios. La purificación de los anticuerpos monoclonales producidos en los ascitis puede conseguirse usando precipitación con sulfato de amonio, seguida por cromatografía de exclusión en gel. Como alternativa, puede emplearse cromatografía de afinidad en base a anticuerpos para proteína A o proteína G.

Depósito de Material

Los siguientes materiales se depositaron en la American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC):

Material	Número de depósito de ATCC.:	Fecha del depósito
DNA56049-2543	203662	9 de febrero de 1999

Estos depósitos se realizaron bajo las condiciones del Tratado de Budapest sobre reconocimiento internacional del depósito de microorganismos para los fines de procedimiento en materia de patentes y las regulaciones en virtud del mismo (Tratado de Budapest). Esto asegura el mantenimiento de un cultivo viable del depósito durante 30 años a partir de la fecha de depósito. El depósito estará disponible en la ATCC bajo los términos del Tratado de Budapest, y sometido a un acuerdo entre Genentech, Inc., y la ATCC, que asegura la disponibilidad permanente y sin restricciones de la progenie del cultivo del depósito al público después de la expedición de la Patente de Estados Unidos pertinente o después de la apertura al público de cualquier Solicitud de patente de Estados Unidos o extranjera, lo que suceda en primer lugar, y asegura la disponibilidad de la progenie a alguien designado para ello por el Comisionado de Patentes y Marcas de Estados Unidos según la disposición 35 USC § 122 las reglas del Comisionado al amparo de la misma (incluyendo la disposición 37 CDR § 1.14 con referencia particular a 886 OG 638).

El cesionario de la presente solicitud acordó que si un cultivo de los materiales en depósito muriera o se perdiera o destruyera mientras se cultiva en condiciones adecuadas, los materiales serán inmediatamente sustituidos tras recibir la notificación por otro cultivo igual. La disponibilidad del material depositado no debe interpretarse como una licencia para poner en práctica la la presente invención contraviniendo los derechos concedidos bajo la autoridad de cualquier gobierno según sus leyes de patentes.

Se considera que la anterior memoria descriptiva escrita es suficiente para permitir a un experto en la materia poner en práctica la presente invención. El alcance de la presente invención no debe limitarse por la construcción depositada, puesto que la realización depositada está concebida como una mera ilustración de ciertos aspectos de la presente invención. El depósito de material de este documento no constituye una admisión de que la descripción escrita contenida en este documento es inadecuada para permitir la puesta en práctica de cualquier aspecto de la presente invención, incluyendo su mejor modo, ni debe interpretarse como limitante del alcance de las reivindicaciones a las ilustraciones específicas que representa. De hecho, diversas modificaciones de la presente invención además de las que se muestran y describen en este documento serán evidentes para los expertos en la materia a partir de la anterior descripción.

Referencias citadas en la descripción

Esta lista de referencias citadas por el solicitante está prevista únicamente para ayudar al lector y no forma parte del documento de patente europea. Aunque se ha puesto el máximo cuidado en su realización, no se pueden excluir errores u omisiones y la OEP declina cualquier responsabilidad en este respecto.

Documentos de patente citados en la descripción

- US 4675187 A [0050]
- 10 • US 4816567 A [0106] [0107] [0235] [0235] [0239]
- EP 404097 A [0110]
- 15 • WO 9311161 A [0110]
- US 4273149 A [0113]
- US 5364934 A [0119]
- 20 • WO 8705330 A [0131]
- US 4640835 A [0133]
- 25 • US 4496689 A [0133]
- US 4301144 A [0133]
- US 4670417 A [0133]
- 30 • US 4791192 A [0133]
- US 4179337 A [0133]
- 35 • US 5428130 A [0135]
- WO 8905859 A [0143]
- US 4399216 A [0143]
- 40 • DD 266710 [0144]
- US 4946783 A [0144]
- 45 • EP 139383 A [0145]
- US 4943529 A [0145]
- EP 402226 A [0145]
- 50 • EP 183070 A [0145]
- EP 244234 A [0145]
- 55 • EP 394538 A [0145]
- WO 9100357 A [0145]
- US 5010182 A [0148]
- 60 • EP 362179 A [0148]
- WO 9013646 A [0148]
- 65 • EP 36776 A [0152]
- EP 73657 A [0154]

- GB 2211504 A [0155]
- EP 117060 A [0158]
- 5 • EP 117058 A [0158]
- WO 9318186 A [0169]
- US 4376110 A [0186]
- 10 • WO 9733551 A [0191] [0219] [0220] [0224] [0225]
- US 4873191 A [0201]
- 15 • US 4736866 A [0201]
- US 5545807 A [0240]
- US 5545806 A [0240]
- 20 • US 5569825 A [0240]
- US 5625126 A [0240]
- 25 • US 5633425 A [0240]
- US 5661016 A [0240]
- WO 8101145 A [0241]
- 30 • WO 8807378 A [0241]
- US 4975278 A [0241]
- 35 • WO 9308829 A [0246]
- WO 9627011 A [0248]
- US 4676980 A [0254] [0254]
- 40 • WO 9100360 A [0254]
- WO 92200373 A [0254]
- 45 • EP 03089 A [0254]
- WO 9411026 A [0258]
- US 4485045 A [0260]
- 50 • US 4544545 A [0260]
- US 5013556 A [0260]
- 55 • US 3773919 A [0269]
- EP 616812 A [0272]
- EP 307247 A [0326].

Documentos no procedentes de patentes citados en la descripción

- **Boring** *et al.* CA Cancel *J. Clin.*, 1993, vol. 43, 7 [0002]
- 65 • **Hunter.** *Cell*, 1991, vol. 64, 1129 [0004]
- **Bishop.** *Cell*, 1991, vol. 64, 233-248 [0004]

ES 2 329 689 T3

- **Alitalo et al.** *Adv. Cancer Res.*, 1986, vol. 47, 235-281 [0005]
- **Slamon et al.** *Science*, 1987, vol. 235, 177-182 [0006]
- 5 • **Slamon et al.** *Science*, 1989, vol. 244, 707-712 [0006]
- **Schwab et al.** *Genes Chromosomes Cancer*, 1990, vol. 1, 181-193 [0007]
- **Ravdin; Chamness.** *Gene*, 1995, vol. 159, 19-27 [0007]
- 10 • **Hynes; Stern.** *Biochim. Biophys. Acta*, 1994, vol. 1198, 165-184 [0007]
- **Baselga et al.** *Oncology*, 1997, vol. 11 (3), 43-48 [0007]
- 15 • **Baselga et al.** *J. Clin. Oncol.*, 1996, vol. 14, 737-744 [0007]
- Cell cycle regulation, oncogens, and antineoplastic drugs. **Murakami et al.** *The Molecular Basis of Cancer*. WB Saunders, 1995, 13 [0051]
- 20 • **Wilman.** Prodrugs in Cancer Chemotherapy. *Biochemical Society Transactions*, 1986, vol. 14, 375-382 [0054]
- Prodrugs: A Chemical Approach to Targeted Drug Delivery. **Stella et al.** *Directed Drug Delivery*. Humana Press, 1985, 147-267 [0054]
- 25 • **Nielsen et al.** *Prot. Eng.*, 1997, vol. 10, 1-6 [0062]
- **Hcinjc et al.** *Nucl. Acids. Res.*, 1986, vol. 14, 4683-4690 [0062]
- **Altschul et al.** *Nucleic Acids Res.*, 1997, vol. 25, 3389-3402 [0068] [0074]
- 30 • **Altschul et al.** *Methods in Enzymology*, 1996, vol. 266, 460-480 [0070] [0076] [0283]
- **Ausubel et al.** *Current Protocols in Molecular Biology*. Wiley Interscience Publishers, 1995 [0085]
- 35 • **Sambrook et al.** *Molecular Clonin: A Laboratory Manual*. Cold Spring Harbor Press, 1989 [0087]
- **Kabat et al.** NIH Publ. No.91-3242, 1991, vol. 1, 647-669 [0098]
- **Kabat et al.** *Sequences of Proteins of Immunological Interest*. National Institute of Health, 1991 [0090]
- 40 • **Clothia; Lesk.** *J. Mol. Biol.*, 1987, vol. 196, 901-917 [0099]
- **Zapata et al.** *Protein Eneg.*, 1995, vol. 8 (10), 1057-1062 [0100]
- 45 • **Kohler et al.** *Nature*, 1975, vol. 256, 495 [0106]
- **Clackson et al.** *Nature*, 1991, vol. 352, 624-628 [0106]
- **Marks et al.** *J. Mol. Biol.*, 1991, vol. 222, 58-597 [0106]
- 50 • **Morrison et al.** *Proc. Natl. Acad. Sci. USA*, 1984, vol. 81, 6851-6855 [0107]
- **Jones et al.** *Nature*, 1986, vol. 321, 522-525 [0108] [0238] [0239]
- 55 • **Reichmann.** *Nature*, 1988, vol. 332, 323-329 [0108]
- **Presta.** *Curr. Op. Struct. Biol.*, 1992, vol. 2, 593-596 [0108] [0238]
- **Pluckthun.** *The Pharmacology of Monoclonal Antibodies*. Springer-Verlag, 1994, vol. 113, 269-315 [0109]
- 60 • **Hollinger et al.** *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 6444-6448 [0110]
- **Carter.** *Nucl. Acids Res.*, 1986, vol. 13, 4331 [0125]
- 65 • **Zoller et al.** *Nucl. Acids Res.*, 1987, vol. 10, 6487 [0125]
- **Wells et al.** *Gene*, 1985, vol. 34, 315 [0125]

ES 2 329 689 T3

- **Wells** *et al.* Philos. Trans. R. Soc. London SerA., 1986, vol. 317, 415 [0125]
- **Cunningham; Wells.** *Science*, 1989, vol. 244, 1081-1085 [0126]
- 5 • **Creighton.** The Proteins. W.H. *Freeman & Co*, [0126]
- **Chothia.** *J. Mol. Biol.*, 1976, vol. 150, 1 [0126]
- **Aplin; Wriston.** CRC Crit. Rev. *Biochem.*, 1981, 259-306 [0131]
- 10 • **Hakimuddin** *et al.* Arch. *Biochem. Biophys.*, 1987, vol. 259, 52 [0132]
- **Edge** *et al.* Anal. *Biochem.*, 1981, vol. 118, 131 [0132]
- 15 • **Thotakura** *et al.* Meth. *Enzymol.*, 1987, vol. 138, 350 [0132]
- **Field** *et al.* Mol. Cell. *Biol.*, 1988, vol. 8, 2159-2165 [0134]
- **Evan** *et al.* *Molecular and Cellular Biology*, 1985, vol. 5, 3610-3616 [0134]
- 20 • **Paborsky** *et al.* *Protein Engineering*, 1990, vol. 3 (6), 547-553 [0134]
- **Hopp** *et al.* *BioTechnology*, 1988, vol. 6, 1204-1210 [0134]
- 25 • **Martin** *et al.* *Science*, 1992, vol. 255, 192-194 [0134]
- **Skinner** *et al.* *J. Biol. Chem.*, 1991, vol. 266, 15163-15166 [0134]
- **Lutz-Freyemuth.** *Proc. Natl. Acad. Sci. USA*, 1990, vol. 87, 6393-6397 [0134]
- 30 • **Stewart** *et al.* Solid-Phase Peptide Synthesis. W.H. *Freeman Co*, 1969 [0136]
- **Merrifield.** *J. Am. Chem. Soc.*, 1963, vol. 85, 2149-2154 [0136]
- 35 • **Sambrook** *et al.* Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press*, 1989 [0138]
- **Dieffenbach** *et al.* PCR Primer: A Laboratory Manual. *Cold Spring Harbor Laboratory Press*, 1995 [0138]
- Mammalian Cell Biotechnology: a Practical Approach. IRL *Press*, 1991 [0142]
- 40 • **Shaw** *et al.* *Gene*, 1983, vol. 23, 315 [0143]
- **Graham; van der Eb.** *Virology*, 1978, vol. 52, 456-457 [0143]
- 45 • **Van Solingen** *et al.* *J. Bact.*, 1977, vol. 130, 946 [0143]
- **Hsiao** *et al.* *Proc. Natl. Acad. Sci. (USA)*, 1979, vol. 76, 3829 [0143]
- **Keown** *et al.* *Methods in Enzymology*, 1990, vol. 185, 527-537 [0143]
- 50 • **Mansour** *et al.* *Nature*, 1988, vol. 336, 348-352 [0143]
- **Beach; Nurse.** *Nature*, 1981, vol. 290, 140 [0145]
- 55 • **Fleer** *et al.* *Bio/Technology*, 1991, vol. 9, 968-975 [0145]
- **Louvencourt** *et al.* *J. Bacteriol.*, 1983, vol. 737 [0145]
- **Vanden Berg** *et al.* *Biol Technology*, 1990, vol. 8, 135 [0145]
- 60 • **Sreekrishna** *et al.* *J. Basic Microbiol.*, 1988, vol. 28, 265-278 [0145]
- **Case** *et al.* *Proc. Natl. Acad. Sci. USA*, 1979, vol. 76, 5259-5263 [0145]
- 65 • **Ballance** *et al.* *Biochem. Biophys. Res. Commun.*, 1983, vol. 112, 284-289 [0145]
- **Tilburn** *et al.* *Gene*, 1983, vol. 26, 205-221 [0145]

- **Yelton et al.** *Proc. Natl. Acad. Sci. USA*, 1984, vol. 81, 1470-1474 [0145]
- **Kelly; Hynes.** *EMBO J.*, 1985, vol. 4, 475-479 [0145]
- 5 • **C. Anthony.** *The Biochemistry of Methylotrophs*, 1982, vol. 269 [0145]
- **Graham et al.** *J. Gen Virol.*, 1977, vol. 36, 59 [0146]
- **Urlaub; Chasin.** *Proc. Natl. Acad. Sci. USA*, 1980, vol. 77, 4216 [0146]
- 10 • **TM4, Mather.** *Biol. Reprod.*, 1980, vol. 23, 243-251
- **Urlaub et al.** *Proc. Natl. Acad. Sci. USA*, 1980, vol. 77, 4216 [0151]
- 15 • **Stinchcomb et al.** *Nature*, 1979, vol. 282, 39 [0151]
- **Kingsman et al.** *Gene*, 1979, vol. 7, 141 [0151]
- **Tschemper et al.** *Gene*, 1980, vol. 10, 157 [0151]
- 20 • **Jones.** *Genetics*, 1977, vol. 85, 12 [0151]
- **Chang et al.** *Nature*, 1978, vol. 275, 615 [0152]
- 25 • **Goeddel et al.** *Nature*, 1979, vol. 281, 544 [0152]
- **Goeddel.** *Nucleic Acids Res.*, 1980, vol. 8, 4057 [0152]
- **deBoer et al.** *Proc. Natl. Acad. Sci. USA*, 1983, vol. 80, 21-25 [0152]
- 30 • **Hitzeman et al.** *J. Biol. Chem.*, 1980, vol. 255, 2073 [0153]
- **Hess et al.** *J. Adv. Enzyme Reg.*, 1968, vol. 7, 149 [0153]
- 35 • **Holland.** *Biochemistry*, 1978, vol. 17, 4900 [0153]
- **Gething et al.** *Nature*, 1981, vol. 293, 620-625 [0158]
- **Mantei et al.** *Nature*, 1979, vol. 281, 40-46 [0158]
- 40 • **Thomas.** *Proc. Natl. Acad. Sci. USA*, 1980, vol. 77, 5201-5205 [0159] [0180]
- **Deutscher.** *Methods in Enzymology*, 1990, vol. 182 [0162]
- 45 • **Scopes.** *Protein Purification: Principles and Practice.* *Springer-Verlag*, 1982 [0162]
- **Gray et al.** *Radiation Res.*, 1994, vol. 137, 275-289 [0169]
- **S. Gelmini et al.** *Clin. Chem.*, 1997, vol. 43, 752 [0172]
- 50 • **Stewart et al.** *Genome Research*, 1997, vol. 7, 422-433 [0182]
- **Zola.** *Monoclonal Antibodies: A Manual of Techniques.* *CRC Press, Inc.*, 1987, 147-158 [0184]
- 55 • **Small et al.** *Mol. Cell. Biol.*, 1985, vol. 5, 642-648 [0190]
- **The Nude Mouse in Oncology Research.** *CRC Press, Inc.*, 1991 [0192]
- **Karmali et al.** *Br. J. Cancer*, 1983, vol. 48, 689-696 [0193]
- 60 • **Drebin et al.** *PNAS USA*, 1986, vol. 83, 9129-9133 [0195]
- **Wang et al.** *Cancer Research*, 1994, vol. 54, 4726-4728 [0196]
- 65 • **Too et al.** *Cancer Research*, 1995, vol. 55, 681-684 [0196]
- **DeLeo et al.** *J. Exp. Med.*, 1977, vol. 146, 720 [0198]

- **Palladino et al.** *J. Immunol.*, 1987, vol. 138, 4023-4032 [0198]
- **Zupi et al.** *Br. J. Cancer.*, 1980, vol. 41 (4), 309 [0199]
- 5 • **Zacharski.** *Haemostasis*, 1986, vol. 16, 300-320 [0199]
- **Rygaard; Spang-Thomsen.** *Proc. 6th Int. Workshop on Immune-Deficient Animals.* 1989, 301 [0200]
- **Van der Putten et al.** *Proc. Natl. Acad. Sci. USA*, 1985, vol. 82, 6148-615 [0201]
- 10 • **Thompson et al.** *Cell*, 1989, vol. 56, 313-321 [0201]
- **Lo.** *Mol. Cell Biol.*, 1983, vol. 3, 1803-1814 [0201]
- 15 • **Lavitrano et al.** *Cell*, 1989, vol. 57, 717-73 [0201]
- **Lasko et al.** *Proc. Natl. Acad. Sci. USA*, 1992, vol. 89, 6232-636 [0202]
- **Thomas; Capecchi.** *Cell*, 1987, vol. 51, 503 [0204]
- 20 • **Li et al.** *Cell*, 1992, vol. 69, 915 [0204]
- **Bradley.** *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach.* *IRL*, 1987, 113-152 [0204]
- 25 • **Fields; Song.** *Nature*, 1989, vol. 340, 245-246 [0210]
- **Chien et al.** *Proc. Natl. Acad. Sci. USA*, 1991, vol. 88, 9578-9582 [0210]
- **Chevray; Nathans.** *Proc. Natl. Acad. Sci. USA*, 1991, vol. 89, 5789-5793 [0210]
- 30 • **Coligan et al.** *Current Protocols in Immun.* 1991, vol. 1 [0212]
- **Lee et al.** *Nucl. Acids Res.*, 1979, vol. 6, 3073 [0216]
- 35 • **Cooney et al.** *Science*, 1988, vol. 241, 456 [0216]
- **Dervan et al.** *Science*, 1991, vol. 251, 1360 [0216]
- **Okano.** *Neurochem.*, 1991, vol. 56, 560 [0216]
- 40 • **Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression.** *CRC Press*, 1988 [0216]
- **Rossi.** *Current Biology*, 1994, vol. 4, 469-471 [0219]
- 45 • **Kohler; Milstein.** *Nature*, 1975, vol. 256, 495 [0229]
- **Goding.** *Monoclonal Antibodies: Principles and Practice.* *Academic Press*, 1986, 59-103 [0230]
- **Kozbor.** *J. Immunol.*, 1984, vol. 133, 3001 [0231]
- 50 • **Brodeur et al.** *Monoclonal Antibody Production Techniques and Applications.* *Marcel Dekker, Inc.*, 1987, 51-63 [0231]
- **Munson; Pollard.** *Anal. Biochem.*, 1980, vol. 107, 220 [0232]
- 55 • **Riechmann.** *Nature*, 1988, vol. 332, 323-329 [0238]
- **Riechmann et al.** *Nature*, 1988, vol. 332, 323-327 [0239]
- 60 • **Verhoeven et al.** *Science*, 1988, vol. 239, 1534-1536 [0239]
- **Hoogenboom; Winter.** *J. Mol. Biol.*, 1991, vol. 227, 381 [0240]
- **Marks et al.** *J. Mol. Biol.*, 1991, vol. 222, 581 [0240]
- 65 • **Cole et al.** *Monoclonal Antibodies and Cancer Therapy.* *Alan R. Liss*, 1985, 77 [0240]
- **Boerner et al.** *J. Immunol.*, 1991, vol. 147 (1), 86-95 [0240]

ES 2 329 689 T3

- **Marks et al.** *Bio/Technology*, 1992, vol. 10, 779-783 [0240]
- **Lonberg et al.** *Nature*, 1994, vol. 368, 856-859 [0240]
- 5 • **Morrison.** *Nature*, 1994, vol. 368, 812-13 [0240]
- **Fishwild et al.** *Nature Biotechnology*, 1996, vol. 14, 945-51 [0240]
- **Neuberger.** *Nature Biotechnology*, 1996, vol. 14, 826 [0240]
- 10 • **Lonberg; Huszar.** *Intern. Rev. Immunol.*, 1995, vol. 13, 65-93 [0240]
- **Massey.** *Nature*, 1987, vol. 328, 457-459 [0243]
- 15 • **Neuberger et al.** *Nature*, 1984, vol. 312, 604-608 [0244]
- **Milstein; Cuello.** *Nature*, 1983, vol. 305, 537-539 [0246]
- **Traunecker et al.** *EMBO J.*, 1991, vol. 10, 3655-3659 [0246]
- 20 • **Suresh et al.** *Methods in Enzymology*, 1986, vol. 121, 210 [0247]
- **Brennan et al.** *Science*, 1985, vol. 229, 81 [0249]
- 25 • **Shalaby et al.** *J. Exp. Med.*, 1992, vol. 175, 217-225 [0250]
- **Kostelny et al.** *J. Immunol.*, 1992, vol. 148 (5), 1547-1553 [0251]
- **Hollinger et al.** *Proc. Natl. Acad. Sci. USA.*, 1993, vol. 90, 6444-6448 [0251]
- 30 • **Gruber et al.** *J. Immunol.*, 1994, vol. 152, 5368 [0251]
- **Tutt et al.** *J. Immunol.*, 1991, vol. 147, 60 [0252]
- 35 • **Caron et al.** *J. Exp Med.*, 1992, vol. 176, 1191-1195 [0255]
- **Shopes.** *J. Immunol.*, 1992, vol. 148, 2918-2922 [0255]
- **Wolff et al.** *Cancer Research*, 1993, vol. 53, 2560-2565 [0255]
- 40 • **Stevenson et al.** *Anti-Cancer Drug Design*, 1989, vol. 3, 219-230 [0255]
- **Vitetta et al.** *Science*, 1987, vol. 238, 1098 [0258]
- 45 • **Epstein et al.** *Proc. Natl. Acad. Sci. USA.*, 1985, vol. 82, 3688 [0260]
- **Hwang et al.** *Proc. Natl. Acad. Sci. USA.*, 1980, vol. 77, 4030 [0260]
- **Martin et al.** *J. Biol. Chem.*, 1982, vol. 257, 286-288 [0261]
- 50 • **Gabizon et al.** *J. National Cancer Inst.*, 1989, vol. 81 (19), 1484 [0261]
- **Marasco et al.** *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 7889-7893 [0263]
- 55 • Remington's *Pharmaceutical Sciences*. 1980 [0264] [0267]
- Chemotherapy Service. *Williams & Wilkins*, 1992 [0272]
- **Sambrook et al.** *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press*, 1989 [0282]
- 60 • **Ausubel et al.** *Current Protocols in Molecular Biology. Green Publishing Associates and Wiley Interscience*, 1989 [0282]
- **Innis et al.** *PCR Protocols: A Guide to Methods and Applications. Academic Press, Inc.*, 1990 [0282]
- 65 • **Harlow et al.** *Antibodies: A Laboratory Manual. Cold Spring Harbor Press*, 1988 [0282]
- **Gait.** *Oligonucleotide Synthesis. IRL Press*, 1984 [0282]

ES 2 329 689 T3

- **R.I. Freshney.** *Animal Cell Culture*, 1987 [0282]
- **Coligan et al.** *Current Protocols in Immunology*, 1991 [0282]
- 5 • **Ausubel et al.** *Current Protocols in Molecular Biology* [0285]
- **Holmes et al.** *Science*, 1991, vol. 253, 1279-1280 [0286]
- **Bolivar.** *Gene*, 1977, vol. 2, 95 [0317]
- 10 • **Thimmappaya et al.** *Cell.*, 1982, 31-543 [0327]
- **Somparyrac et al.** *Proc. Natl. Acad. Sci.*, 1981, vol. 12, 7575 [0329]
- 15 • **Ausubel et al.** *Current Protocols of Molecular Biology. John Wiley and Sons*, 1997 [0333]
- **Lucas et al.** *Nucl. Acids Res.*, 1996, vol. 24 (9), 1774-1779 [0333]
- 20 • **O'Reilley et al.** *Baculovirus expression vectors: A Laboratory Manual. Oxford University Press*, 1994 [0344]
- **Rupert et al.** *Nature*, 1993, vol. 362, 175-179 [0345].

REIVINDICACIONES

1. Anticuerpo aislado que se une a un polipéptido PRO3567, para su uso en un procedimiento de diagnóstico de

a) que tiene al menos el 99% de identidad en la secuencia de aminoácidos mostrada en la Figura 2 (SEC ID N°: 27);

b) que da una valoración al menos el 99% de positivos cuando se compara con una secuencia de aminoácidos seleccionada mostrada en la Figura 2 (SEC ID N°: 27);

c) que tiene al menos el 99% de identidad en la secuencia de aminoácidos con una secuencia de aminoácidos codificada por la secuencia codificante de longitud completa del ADN depositado con el número de entrada en la ATCC 203662;

d) que tiene al menos el 99% de identidad en la secuencia de aminoácidos con el polipéptido mostrado en la Figura 2 (SEC ID N°: 27), que carece de su péptido señal asociado;

e) que tiene al menos el 99% de identidad en la secuencia de aminoácidos con un dominio extracelular del polipéptido mostrado en la Figura 2 (SEC ID N°: 27), con su péptido señal asociado; o

f) que tiene al menos el 99% de identidad en la secuencia de aminoácidos con un dominio extracelular del polipéptido mostrado en la Figura 2 (SEC ID N°: 27), que carece de su péptido señal asociado.

2. Anticuerpo de la reivindicación 1 que es un anticuerpo monoclonal.

3. Anticuerpo de la reivindicación 2 que comprende una región determinante de la complementariedad (CDR) no humana o una región marco (FR) humana.

4. Anticuerpo de cualquiera de las reivindicaciones 1 a 3 que está marcado.

5. Anticuerpo de cualquiera de las reivindicaciones 1 a 4 que es un fragmento de anticuerpo o un anticuerpo de cadena sencilla.

6. Procedimiento para determinar la presencia de un polipéptido PRO3567 como se define en la reivindicación 1 en una muestra sospechosa de contener dicho polipéptido, comprendiendo dicho procedimiento la exposición de la muestra a un anticuerpo anti-PRO3567 y determinar la unión de dicho anticuerpo a dicho polipéptido en dicha muestra, en el que dicha muestra comprende una célula de cáncer de colon sospechosa de contener un polipéptido PRO3567.

7. Procedimiento para diagnosticar el tumor de colon en un mamífero, comprendiendo dicho procedimiento el nivel de expresión de un gen que codifica a un polipéptido PRO3567 como se define en la reivindicación 1 (i) en una muestra de ensayo de tejidos celulares obtenidos a partir del mamífero, y (ii) en una muestra de control de células de tejido normal conocidas del mismo tipo celular, en el que un nivel de expresión mayor en la muestra de ensayo, en comparación con la muestra de control, es indicativo de la presencia de células de tejido celular en el mamífero, a partir de las que se obtuvieron las células del tejido de ensayo.

8. Procedimiento para diagnosticar tumor de colon en un mamífero, comprendiendo dicho procedimiento

(i) poner en contacto un anticuerpo que se une al polipéptido PRO32567 como se define en la reivindicación 1 con (a) una muestra de ensayo de células de tejido obtenidas del mamífero, y (b) una muestra de control de células de tejido normal conocidas del mismo tipo celular; y

(ii) detectar la formación de un complejo entre dicho anticuerpo y un polipéptido PRO3567 como se define en la reivindicación 1 en cada muestra, en el que la formación de una cantidad mayor de complejos en la muestra de ensayo en comparación con la muestra normal es indicativo de la presencia de un tumor en un mamífero.

9. Procedimiento de la reivindicación 8, en el que dicho anticuerpo está marcado de forma detectable.

10. Procedimiento de la reivindicación 9, en el que dicha muestra de ensayo de células de tejido se obtiene de un individuo sospechoso de tener crecimiento o proliferación celular neoplástico.

11. Procedimiento según la reivindicación 7, donde el nivel de expresión de dicho gen se determina utilizando una molécula de ácido nucleico aislada que hibrida con una secuencia de ácido nucleico que codifica a un polipéptido PRO3567 como se describe en la reivindicación 1, o el complemento del mismo.

12. Una molécula de ácido nucleico aislado que codifica el anticuerpo de cualquiera de las reivindicaciones 1 a 5.

ES 2 329 689 T3

13. Un vector que comprende la molécula de ácido nucleico de la reivindicación 12.

14. Una célula huésped que comprende el vector de la reivindicación 13.

5 15. Procedimiento para producir un anticuerpo según la reivindicación 1, comprendiendo dicho procedimiento el cultivo de la célula huésped de la reivindicación 14 bajo condiciones suficientes para permitir la expresión de dicho anticuerpo y la recuperación de dicho anticuerpo del cultivo celular.

10

15

20

25

30

35

40

45

50

55

60

65

FIGURE 1

GCACCGCAGACGGCGCGGATCGCAGGGAGCCGGTCCGCCGCCGGAACGGGAGCCTGGGTGTGCGTGTGGAGTCCG
GACTCGTGGGAGACGATCGCGATGAACACGGTGCTGTGCGGGCGAACTCACTGTTGCGCTTCTCGCTGAGCGTG
ATGGCGGCGCTCACCTTCGGCTGCTTCATCACCACCGCCTTCAAAGACAGGAGCGTCCCGGTGCGGCTGCACGTC
TCGCGGATCATGCTAAAAATGTAGAAGATTTCACTGGACCTAGAGAAAGAAGTGATCTGGGATTTATCACATTT
GATATAACTGCTGATCTAGAGAATATATTTGATTGGAATGTTAAGCAGTTGTTTCTTTATTTATCAGCAGAATAT
TCAACAAAAAATAATGCTCTGAACCAAGTTGTCCTATGGGACAAGATTGTTTGTAGAGGTGATAATCCGAAGCTG
CTGCTGAAAGATATGAAAACAAAATATTTTCTTTGACGATGGAAATGGTCTCAAGGGAAACAGGAATGTCAC
TTGACCCCTGTCTTGGAACGTCGTACCAAATGCTGGAATTCTACCTCTTGTGACAGGATCAGGACACGTATCTGTC
CCATTTCAGATACATATGAAATAACGAAGAGTTATTAATTATTCTGAATTTGAAACAAAA

FIGURA 2

Péptido señal:	Aminoácidos 1-25
Dominio transmembrana:	Aminoácidos 149-164
Sitio de N-glicosilación:	Aminoácidos 141-145
Sitios de N-miristoilación:	Aminoácidos 25-31; 135-141
Sitio de unión de lípido lipoproteína de membrana procariota:	Aminoácidos 16-27
Secuencia de unión celular:	Aminoácidos 112-115
Firma de proteína receptora Ton-B 1:	Aminoácidos 1-21

**MNTVLSRANSLFAFSLSVMAALTFGCFITTAFKDRSVFVRLHVSRIMLKNVEDFTGPRERSDLGFITFDITADLE
NIFDWNVKQLFLYLSAEYSTKNNALNQVVLWDKIVLRGDNPKLLLKDMKTKYFFFDDGNGLKGNRNVTLTSLWNV
VPNAGILPLVTGSGHVSVPFPDTYEITKSY**