



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 274 977**

51 Int. Cl.:
C12N 15/55 (2006.01)
C12N 9/78 (2006.01)
C12Q 1/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **02732665 .1**
86 Fecha de presentación : **25.04.2002**
87 Número de publicación de la solicitud: **1387890**
87 Fecha de publicación de la solicitud: **11.02.2004**

54 Título: **Peptidilarginina deiminasa 6.**

30 Prioridad: **03.05.2001 EP 01201601**

45 Fecha de publicación de la mención BOPI:
01.06.2007

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

73 Titular/es: **N.V. Organon**
Kloosterstraat 6
5349 AB Oss, NL

72 Inventor/es: **Gossen, Jan, Albert y**
Van den Boogaart, Paul

74 Agente: **Ungría López, Javier**

ES 2 274 977 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

Peptidilarginina deiminasa 6.

5 La presente invención se refiere a polinucleótidos que codifican peptidilarginina deiminasa 6, a células transfectadas con estos polinucleótidos, a proteínas producidas por estas células así como a un método para producir estas proteínas y sus moduladores.

10 Las peptidilarginina deimasas (PAD) son una familia de enzimas de modificación post-traduccional que convierten peptidilarginina en citrulina de una manera dependiente de Ca^{2+} . La deiminación enzimática *in vitro* cambia las propiedades funcionales de diversas proteínas y altera sus estructuras secundaria y terciaria.

15 Hasta ahora, se han identificado cinco isoformas de PAD que muestran una amplia distribución de tejido. La PAD1 de ratón se detecta en la epidermis y útero (Rus'd, A.A. *et al.* 1990, Eur. J. Biochem. 259, 660-669); la PAD2 de ratón se expresa ampliamente en diversos tejidos tales como cerebro, pituitaria, médula espinal, glándulas salivales, páncreas, músculo esquelético, útero, bazo, estómago y timo (Takahara, H. *et al.* 1989, J. Biol. Chem. 264, 13361-13368); la PAD3 de ratón se expresa en la epidermis y folículos capilares (Terakawa, H. *et al.* 1991, J. Biochem (Tokio) 110, 661-666); la PAD4 (rata) es una enzima ubicua que se expresa en el páncreas, bazo, ovario, hígado, pulmón, estómago, riñón, útero, dermis, cerebro, corazón y epidermis (Yamakoshi, A *et al.* 1998, Biochim. Biophys. 20 Acta 1386, 227-232); finalmente se ha aislado la PAD5 humana como un nuevo miembro de la familia a partir de una línea celular de leucemia mieloide, pero no se ha determinado además su distribución en el tejido (Nakashima, K. *et al.* 1999, J. Biol. Chem. 274, 27786-27792).

25 Se conoce poco sobre las funciones fisiológicas de la PAD. En el cerebro, la proteína básica de la mielina es un sustrato natural y por lo tanto, la PAD juega un papel importante en el sistema nervioso central. Además, cuando la PAD se desregula, juega un papel en la etiología de la esclerosis múltiple (Mastronardi, F.G. *et al.* 1996, Clin. Invest. 97, 349-358). Parece que la PAD está implicada en la dermis en el procesamiento terminal de la filagrina, que de forma indirecta es importante para mantener la humedad en el estrato córneo superior (Senshu, T. *et al.* 1996, Biochem. Biophys. Res. Commun. 225, 712-719). De nuevo, la desregulación de esta PAD juega un papel en la etiología de la 30 artritis reumatoide (Girbal-Neuhauser, E. *et al.* 1999, J. Immunol. 162, 585-594). En los folículos capilares finalmente la solubilidad de la trichialina parece estar influenciada por la PAD; la función de esto permanece por determinar (Rogers, G.E. *et al.* 1997, J. Invest. Dermatol. 108, 700-707).

35 Los autores han descubierto una nueva PAD, llamada PAD6. El transcripto se ha encontrado en ovocitos de ratón. Su homóloga humana se describe también en este documento. Se descubrió que la proteína se expresaba exclusivamente en ovocitos/ovarios y testículos.

40 Los genes que se expresan específicamente en los gametos masculinos y/o femeninos pueden proporcionar nuevas dianas moleculares para la contracepción masculina y femenina. Para testículos, se ha descrito un gran número de secuencias de genes expresados únicamente en células germinales (Pawlak, A. *et al.* 1995 Genomics 26, 151-1588; Wolgemuth D.J. y Watrin F. 1991 Mamm Genome 1, 283-817). Por el contrario, sólo se han identificado unos pocos genes expresados específicamente en ovocitos de ese modo. La mayoría de las secuencias de genes específicos de gametos identificados tienen probablemente una función esencial debido a su expresión específica en gametos. Lo último se confirma con estudios que usan animales knockout que indican que la inactivación génica de genes específicos de 45 testículos y ovocitos produce generalmente infertilidad en el macho y/o hembra pero no produce patología adicional en otros órganos y tejidos (Dong, J. *et al.* 1996 Nature 383, 531-535; Nantel, F. *et al.* 1996, Nature 380, 159-162). Estos datos proporcionan evidencia adicional del papel específico y esencial de estos genes durante la gametogénesis. Esto subraya la importancia de la especificidad de tejido como un criterio de selección de dianas moleculares para la regulación de la fertilidad.

50 Estará claro que existe una gran necesidad de aclaración de los genes implicados en la regulación de la fertilidad para desenredar los diversos papeles que estos genes juegan en la infertilidad. Un mejor conocimiento de estos genes implicados en diferentes fases de la fertilidad femenina y masculina, por ejemplo en la gametogénesis y su regulación de la actividad y expresión, podrían ayudar a crear una mejor percepción de los trastornos de infertilidad. Esto podría 55 conducir eventualmente a la identificación de moduladores de la actividad para usar en protocolos terapéuticos *in vivo* o *in vitro*.

60 La presente invención proporciona tal gen. Más específicamente, la presente invención proporciona una secuencia polinucleotídica que codifica la peptidilarginina deiminasa 6 (PAD6). Preferiblemente, el polinucleótido es de origen mamífero, preferiblemente un ser humano. El ARN se expresa exclusivamente en órganos reproductores.

La secuencia polinucleotídica más preferida en la que codifica la SEC ID N°: 3.

65 La invención incluye la secuencia de ARNm humano completa como se indica en la SEC ID N°: 4 y la fase de lectura abierta correspondiente a la secuencia de nucleótidos 20-2077 de la SEC ID N°: 4. Esta secuencia codifica una proteína de 686 aminoácidos (SEC ID N°: 3). Para adaptar la variabilidad del codón, la invención incluye también secuencias polinucleotídicas que codifican las mismas secuencias de aminoácidos que las secuencias descritas en este documento. La información de secuencia como se proporciona en este documento no debe construirse de forma tan

ES 2 274 977 T3

estrecha como para requerir la exclusión de bases identificadas erróneamente. La secuencia específica descrita en este documento puede usarse fácilmente para aislar los genes completos de varias otras especies o variantes alélicas. La secuencia puede usarse por ejemplo para preparar sondas o como una fuente para preparar oligonucleótidos sintéticos para usar como cebadores en reacciones de amplificación de ADN que permitan el aislamiento e identificación de los genes completos de la variante. En particular, los polinucleótidos que hibridan en condiciones de lavado rigurosas con una sonda preparada por PCR en condiciones convencionales usando la SEC ID N°: 14 y la SEC ID N°: 15 con ADNc de origen mamífero como molde, preferiblemente de ser humano, son parte de la invención. Tal sonda (y su secuencia complementaria) se identifica por ejemplo por los nucleótidos 464-1052 de la SEC ID N°: 4.

La secuencia genética completa puede usarse en la preparación de moléculas de vector para la expresión de la proteína en células hospedadoras adecuadas.

De ese modo, en un aspecto, la presente invención proporciona polinucleótidos aislados que codifican la nueva proteína PAD6. Preferiblemente la PAD6 es de origen humano, pero también los ortólogos forman parte de la invención.

El ADN de acuerdo con la invención puede obtenerse a partir de ADNc. Los tejidos son preferiblemente de origen mamífero, más preferiblemente de origen humano. Preferiblemente, los ácidos ribonucleicos se aíslan de ovocitos o testículos. Como alternativa, la secuencia codificante puede ser ADN genómico, o prepararse usando técnicas de síntesis de ADN. El polinucleótido puede también estar en forma de ARN. Si el polinucleótido es ADN, puede estar en forma de cadena única o cadena doble. La cadena única puede ser la cadena codificante o la cadena no codificante (antisentido).

El ADN de acuerdo con la invención será muy útil para la expresión *in vivo* o *in vitro* de la nueva proteína de acuerdo con la invención en cantidades suficientes y en forma sustancialmente pura.

La presente invención refiere además a polinucleótidos que tienen ligeras variaciones o que tienen sitios polimórficos. Los polinucleótidos que tienen ligeras variaciones pueden codificar variantes de polipéptidos que mantienen la misma función o actividad biológica que la proteína natural, madura. Los sitios polimórficos son útiles para propósitos de diagnóstico.

En otro aspecto, la invención proporciona un método para aislar un polinucleótido que comprende las etapas de: a) hibridar un polinucleótido de acuerdo con la presente invención, o su complementario, en condiciones rigurosas frente ácidos nucleicos que son ADN (genómico), ARN o ADNc aislados preferiblemente de tejidos que expresan altamente el polinucleótido de interés y b) aislar dichos ácidos nucleicos por métodos conocidos por el especialista en la técnica. Los tejidos son preferiblemente de origen humano. Preferiblemente, los ácidos ribonucleicos se aíslan de ovocitos, ovarios o testículos. Las condiciones de hibridación son preferiblemente altamente rigurosas.

De acuerdo con la presente invención, el término “riguroso” significa condiciones de lavado de 1 x SSC, SDS al 0,1% a una temperatura de 65°C; condiciones altamente rigurosas se refieren a una reducción en SSC alrededor de 0,3 x SSC, más preferiblemente a 0,1 x SSC. Preferiblemente, los primeros dos lavados se realizan posteriormente dos veces cada uno durante 15-30 minutos. Si se necesita lavar en condiciones altamente rigurosas, se realiza un lavado adicional con 0,1 x SSC una vez durante 15 minutos. La hibridación puede realizarse por ejemplo durante toda la noche en tampón fosfato 0,5 M pH 7,5/SDS al 7% a 65°C.

Como una alternativa, el método de aislamiento del gen puede comprender metodología de amplificación génica usando cebadores derivados del ácido nucleico de acuerdo con la invención. Pueden obtenerse también ADNc completo combinando clones obtenidos por ejemplo por hibridación con por ejemplo clones RACE de ADNc.

También son parte de la invención porciones de las secuencias codificantes que codifican un polipéptido funcional, así como variaciones alélicas y de especie de las mismas. Algunas veces, se expresa un gen en un cierto tejido como una variante de empalme, produciendo un ARNm alterado en 5' ó 3' o la inclusión o exclusión de una o más secuencias de exones. Se espera que todas estas secuencias así como las proteínas codificadas por estas secuencias realicen la misma o similares funciones y forman parte de la invención.

La invención proporciona también peptidilarginina deiminasa 6 (PAD6). Preferiblemente, la proteína tiene una secuencia de aminoácidos de mamíferos, más preferiblemente una secuencia humana. Lo más preferido es la secuencia descrita en la SEC ID N°: 3. Puede obtenerse la expresión por introducción de moléculas de vector que comprenden un polinucleótido que codifica PAD6 dentro de células hospedadoras adecuadas. Las células pueden cultivarse y la proteína aislarse usando métodos conocidos por los especialistas en la técnica.

En otro aspecto más de la invención, se proporcionan equivalentes funcionales, es decir, polipéptidos que codifican actividades de la PAD6 y que comprenden esencialmente la misma secuencia SEC ID N°: 3 o partes de la misma que tienen variaciones en la secuencia mientras que mantienen todavía las características funcionales.

Las variaciones que pueden ocurrir en una secuencia pueden demostrarse por una diferencia o diferencias de aminoácidos en la secuencia total o por deleciones, sustituciones, inserciones, inversiones o adiciones de un aminoácido o aminoácidos en dicha secuencia. Se han descrito las sustituciones de aminoácidos que no se espera que

alteren esencialmente las actividades biológicas e inmunológicas. Los reemplazamientos de aminoácidos entre aminoácidos relacionados o reemplazamientos que han ocurrido frecuentemente en la evolución son, entre otros, Ser/Ala, Ser/Gly, Asp/Gly, Asp/Asn, Ile/Val (véase el documento de Dayhof, M.D., Atlas of protein sequence and structure, Nat. Biomed. Res. Found., Washington D. C., 1978, vol 5, supl. 3). Basándose en esta información, Lipman y Pearson desarrollaron un método para la comparación rápida y sensible de proteínas (Science, 1985 227, 1435-1441) y para determinar la semejanza funcional entre polipéptidos homólogos. Estará claro que también los polinucleótidos que codifican tales variantes son parte de la invención.

De ese modo, en otro aspecto de la invención, se proporcionan polipéptidos que comprenden la SEC ID N°: 3 pero también polipéptidos con una semejanza del 80%, preferiblemente del 90%, más preferiblemente del 95%.

Como se usa en este documento, el término semejanza se define en el documento NCBI-BLAST 2.0.10 [26 de agosto de 1999] (Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, y David J. Lipman (1997) "Grapped BLAST and PSI-BLAST: a new generation o protein database search programs", Nucleic Acids Res. 25, 3389-3402). El programa se usa para buscar alineamientos de secuencia usando ajustes por defecto. Para alineamientos de aminoácidos se usa la matriz BLOSUM62 por defecto y la semejanza se indica como el número de positivos. No se incluye filtración de complejidad composicional baja.

También se incluyen porciones de tales polipéptidos capaces todavía de conferir efectos biológicos. Especialmente, las porciones que son todavía capaces de convertir la arginina en citrulina forman parte de la invención. Tales proteínas o partes funcionales de las mismas pueden ser funcionales *per se*, por ejemplo en forma solubilizada o pueden unirse a otros polipéptidos (por ejemplo para dirigirlos a compartimientos subcelulares específicos, para aumentar su solubilidad o para facilitar su purificación), por medios biotecnológicos conocidos o por síntesis química, para obtener proteínas quiméricas.

Estará claro que también los polinucleótidos que codifican tales variantes están incluidos en la invención.

Pueden emplearse de forma útil una gran variedad de combinaciones de células hospedadoras y vehículos de clonación para clonar las secuencias de ácidos nucleicos que codifican el polipéptido de acuerdo con la invención.

Los vectores de expresión adecuados son por ejemplo plásmidos bacterianos o de levaduras, plásmidos de intervalo de hospedador amplio y vectores derivados de combinaciones de plásmidos y ADN de fagos o virus. Se incluyen también vectores derivados de ADN cromosómico. Además, puede estar presente un origen de replicación y/o un marcador de selección dominante en el vector de acuerdo con la invención. Los vectores de acuerdo con la invención son adecuados para transformar una célula hospedadora.

Los vehículos para usar en la expresión de la proteína o partes de la misma de la presente invención comprenderán además secuencias de control unidas funcionalmente a la secuencia de ácido nucleico que codifica la proteína. Tales secuencias de control comprenden generalmente una secuencia promotora y secuencias que regulan y/o potencian los niveles de expresión. Por supuesto que las secuencias de control y otras pueden variar dependiendo de la célula hospedadora seleccionada.

Los vectores de expresión recombinantes que comprenden el ADN de la invención así como células transfectadas con dicho ADN o dicho vector de expresión, de forma transitoria o estable, también forman parte de la presente invención.

Las células hospedadoras adecuadas de acuerdo con la invención son células hospedadoras bacterianas, levaduras y otras hospedadoras de hongos, plantas o animales tales como células de ovario de hámster chino, células de mono o células humanas. De ese modo, una célula hospedadora que comprenda el ADN o el vector de expresión de acuerdo con la invención está también dentro del alcance de la invención. Las células hospedadoras modificadas pueden cultivarse en medio con nutrientes convencionales que pueden modificarse por ejemplo por selección apropiada, amplificación o inducción de la transcripción. Las condiciones de cultivo tales como temperatura, pH, nutrientes, etc. las conocen bien los especialistas en la técnica.

Las técnicas para la preparación del ADN o del vector de acuerdo con la invención así como la transformación o transfección de una célula hospedadora con dicho ADN o vector son convencionales y se conocen bien en la técnica, véase por ejemplo el documento de Sambrook *et al.*, Molecular Cloning: A laboratory Manual. 2ª Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.

La proteína de acuerdo con la invención puede recuperarse y purificarse a partir de cultivos celulares recombinantes por métodos de purificación bioquímicos comunes (como los descritos en el documento Guide to Protein purification. Editado por Murray P. Deutscher (1990) Methods in Enzymology. Vol 182. Academic Press, inc. San Diego CA 92101. Harcourt Brace Jovanovich, Publishers) incluyendo precipitación con sulfato de amonio, extracción, cromatografía tal como cromatografía de interacción hidrófoba, cromatografía de intercambio catiónico o aniónico o cromatografía de afinidad y cromatografía líquida de alta resolución. Si es necesario, pueden incluirse también etapas de replegamiento de la proteína. Como alternativa, la proteína puede expresarse como una proteína de fusión que contiene ("señales") que pueden usarse para purificación por afinidad.

ES 2 274 977 T3

La regulación de la actividad de la proteína de acuerdo con la invención es útil *in vivo* para el control del reclutamiento folicular, pero también del crecimiento y maduración de los ovocitos y/o folículos. La inhibición de estos procesos *in vivo* puede usarse para retrasar (prematura) la menopausia y/o como un contraceptivo. Además, la proteína puede emplearse para la maduración *in vivo* del crecimiento de los folículos, por ejemplo a partir de tejido de ovarios congelado.

Los productos del gen de PAD de acuerdo con la presente invención pueden usarse para la identificación *in vivo* o *in vitro* de nuevos sustratos o análogos de los mismos. Para este propósito, pueden realizarse por ejemplo estudios de ensayo de la peptidilarginina deiminasa con células transformadas con ADN de acuerdo con la invención o un vector de expresión que comprende ADN de acuerdo con la invención, expresando dichas células los productos del gen de PAD6 de acuerdo con la invención. Como alternativa, pueden usarse también la proteína PAD6 por sí misma o los dominios de unión al sustrato de la misma en un ensayo para la identificación de sustratos o análogos funcionales.

Se conocen bien los métodos para determinar la actividad de la peptidil arginina deiminasa de productos del gen expresados en ensayos *in vitro* e *in vivo* para determinar la actividad biológica de los productos del gen. Véase por ejemplo el documento de Lamensa, F.E.W. y Moscarello, M.A. 1993 J. Neurochem. 61, 987-996. En este ensayo, se convierte arginina en forma de α -Nbenzoil-L-arginina etil éster (BAEE) en citrulina que puede medirse fácilmente después de la precipitación con ácido perclórico.

Otro ejemplo de determinación de la actividad enzimática de PAD6 hace uso de la inactivación de una proteína, por ejemplo el inhibidor de la tripsina de la soja (STI) (Takahara, H. *et al.* 1985, J. Biol. Chem 260, 8378-8383). Cuando una arginina esencial en STI se transforma en citrulina, no es capaz ya de inhibir la actividad proteolítica de la tripsina. Esto puede usarse como la base de un ensayo en dos etapas para la determinación de la actividad de la PAD. El ensayo consiste en dos etapas. En la primera reacción, PAD convierte la arginina (posición 63) de STI en una citrulina, inactivando el STI. En la segunda reacción, se añaden tripsina y un sustrato fluorescente y se mide la actividad de la tripsina.

Como alternativa, puede obtenerse la modulación de la actividad de PAD6 por regulación a la baja del nivel de expresión de la proteína, por ejemplo usando ácidos nucleicos anti-sentido por formación de triples hélices (Cooney *et al.*, 1988, Science 241, 456-459) o por unión al ARNm, o influenciando la estabilidad del ARNm o interacciones de la proteína con moléculas pequeñas. Esto por sí mismo puede conducir a la regulación de la fertilidad, es decir, contracepción o tratamiento de la infertilidad.

De ese modo, la presente invención proporciona un método para identificar compuestos que afectan a la función enzimática de la proteína de acuerdo con la invención. El método comprende las etapas de

a) contactar la proteína PAD6 con un sustrato que contiene arginina

b) contactar dicha mezcla con un compuesto de ensayo

c) medir la conversión de arginina en citrulina y

d) comparar dicha transformación con la actividad de la peptidilarginina deiminasa en ausencia de un compuesto de ensayo.

La transformación de arginina en citrulina puede medirse fácilmente, por ejemplo por métodos analíticos como HPLC, sensibilidad proteolítica del péptido alterada, cambio en las propiedades de actividad del péptido o reconocimiento del anticuerpo específico. Como sustrato pueden usarse péptidos o proteínas que comprendan arginina, pero también compuestos sintéticos como α -N-benzoil-L-arginina etil éster. Sin embargo, tienen que sustituirse los grupos amino y carboxilo o tienen que estar en forma de enlace peptídico.

Como alternativa, la presente invención proporciona un método para identificar compuestos que modulan la estabilidad del ARNm de PAD6 o los niveles de expresión de PAD6.

La presente invención proporciona de ese modo un método económico y rápido para detectar agentes terapéuticos para el control de la fertilidad relacionado con la actividad de PAD6. El método de acuerdo con la invención proporciona además la selección de agentes terapéuticos selectivos que discriminan entre peptidilarginina deiminasa diferentes conduciendo de ese modo a un agente terapéutico más eficaz y/o a disminuir los efectos secundarios. El método es especialmente adecuado para usarse en la detección de alta capacidad de numerosos compuestos diana potenciales.

Los compuestos que modulan la función de la peptidilarginina deiminasa 6 pueden emplearse en tratamientos terapéuticos modulando la PAD de la presente invención.

La invención proporciona también un método para la formulación de una composición farmacéutica que comprende mezclar los compuestos moduladores identificados con un vehículo farmacéuticamente aceptable.

Los vehículos farmacéuticamente aceptables los conocen bien los especialistas en la técnica e incluyen, por ejemplo, solución salina estéril, lactosa, sacarosa, fosfato de calcio, gelatina, dextrina, agar, pectina, aceite de cacahuete, aceite de oliva, aceite de sésamo y agua.

5 Además, las composiciones farmacéuticas pueden comprender uno o más estabilizadores tales como, por ejemplo, hidratos de carbono incluyendo sorbitol, manitol, almidón, sacarodextrina y glucosa, proteínas tales como albúmina o caseína, y tampones tales como fosfatos alcalinos. Los métodos para fabricar preparaciones y mezclas intravenosas se describen en el documento Remington's Pharmaceutical Sciences, pp. 1463-1497 (16ª ed. 1980, Mack Publ. Co of Easton, Pa, USA).

10 De ese modo, los compuestos moduladores identificados usando la peptidilarginina deiminasa de acuerdo con la invención son útiles en la preparación de un producto farmacéutico. El fármaco es para usarlo en el control de trastornos de la fertilidad.

15 Los siguientes ejemplos son ilustrativos para la invención y no deben interpretarse de ningún modo como limitantes del alcance de la invención.

Legendas de las figuras

20 Figura 1

Análisis por RT-PCR (30 ciclos) de la expresión de PAD6 de ratón en diversos tejidos de ratón (grupo superior). En el grupo inferior se muestran controles de GAPDH en ausencia y presencia de RT.

25 Figura 2

Análisis por ISH (hibridación *in situ*) usando el clon 1B11 como una sonda sobre ovarios de ratones jóvenes (7 días) y adultos.

30 S=folículo secundario A=folículo antral.

Figura 3

35 Análisis por RT-PCR (30 ciclos) de la expresión de PAD6 humana en diversos tejidos humanos (grupo superior). En el grupo inferior se muestran controles de GAPDH en ausencia y presencia de RT.

Figura 4

40 Transferencias de Northern (Clontech) de tejidos múltiples humanos hibridados con una sonda de PAD6h.

Figura 5

45 Medida de fluorescencia para determinar la actividad de PAD. Se preincubó STI (0,17 μg) en ausencia (cuadrado relleno) o presencia de 0,5 μg de GST-PAD6 (triángulo relleno), 1,0 μg de GST-PAD6 (triángulo abierto), o 0,1 μg de PAD de músculo de conejo (cuadrado abierto; N° de cat. Sigma P1584) respectivamente. Posteriormente, se añadió *N* α -benzoil-L-Arginina-7-amido-4-metilcumarina (100 μM ; N° de cat. Sigma B7260) y tripsina (0,25 μg), y se determinó la fluorescencia.

Ejemplos

50 Ejemplo 1

Preparación de clones de ADNc de ratón

55 *Generación de una biblioteca de ADNc de ovocitos*

Se aisló ARN total a partir de 2172 ovocitos de ratón desnudos, tratados *in vitro* durante 15 h con FF-MAS 50 μM , de acuerdo con el protocolo de aislamiento de ARN RNazol B™ (Campro scientific). Se añadió RNazol B directamente a los sedimentos de células congeladas que contenían aproximadamente 100 ovocitos cada uno. Los homogenizados se agruparon y se extrajeron con 0,1 volúmenes de cloroformo, se agitaron durante 15 segundos y se incubaron en hielo durante 10 minutos. Después de la centrifugación durante 15 minutos a 14000 rpm a 4°C, se recogió la fase acuosa. Se precipitó el ARN total añadiendo un volumen igual de isopropanol seguido por incubación durante una noche a 4°C. El ARN se centrifugó durante 45 minutos a 14000 rpm a 4°C, el sedimento se lavó una vez con 700 μl de etanol al 70% seguido por centrifugación a 14000 rpm a 4°C durante 30 minutos. El sedimento secado al aire se resuspendió finalmente en 7,5 μl de agua libre de ARNasa (Ambion). La cantidad total de ARN aislado usando este procedimiento se determinó usando el kit de cuantificación de ARN Ribogreen™ (Molecular Probes).

ES 2 274 977 T3

Para síntesis de ADNc, se usó el kit de construcción de bibliotecas de PCR SMART™ (Clontech). Se introdujeron las siguientes modificaciones. Se hibridó un cebador de oligodT(18) con el sitio de restricción EcoRI (Pharmacia) con el extremo 3' del ARNm y el oligo SMART™ extendido con un sitio de restricción EcoRI se hibridó con el extremo 5' del ARNm. La primera reacción de síntesis de la cadena de ADNc fue en un tampón de reacción que contenía Tris 50 mM (pH 8,3), KCl 75 mM, MgCl₂ 6 mM, DTT 2 mM, mezcla de NTPd 1 mM y 200 unidades de transcriptasa inversa Superscript II RNasa H (Gibco BRL) durante 1 hora a 42°C. Posteriormente se amplificó la primera cadena de ADNc por PCR usando un termociclador Perkin Elmer (9600). La PCR se realizó en un volumen total de 100 µl de tampón de reacción que contenía tampón PCR Klen Taq 1x (Clontech), mezcla de NTPd 0,2 mM (Clontech), cebador EcoRI-SMART 5' 0,2 mM, cebador NotI-EcoRI-dT(18) 0,2 mM (Pharmacia) y mezcla de polimerasa Taq Advantage Klen 1x (Clontech) comenzando con 1 minutos de desnaturalización a 95°C seguido por 28 ciclos de 15 segundos a 95°C y 5 minutos a 68°C.

Después de la purificación en una columna de centrifugación Ququick (QIAGEN), se digirió el ADNc con EcoRI (Pharmacia) a 37°C seguido por inactivación con calor a 70°C durante 10 minutos. El ADNc se purificó dos veces usando dos columnas de centrifugación Ququick posteriores y finalmente se resuspendió en 50 µl de Tris-CL 10 mM (pH 8,5). Se determinó la concentración de ADN midiendo la absorbancia a 260 nm usando un espectrofotómetro Genequant.

Fraccionamiento por tamaño del ADNc

Se fraccionó el ADNc por tamaño usando electroforesis en gel de agarosa y se extrajo de la matriz del gel usando el kit de extracción del gel de agarosa QiaexII (Qiagen). Se eluyó el ADN en 20 µl de H₂O, se purificó en una columna de centrifugación Ququick (Qiagen) y se eluyó en 50 µl de H₂O. Las muestras se precipitaron añadiendo 0,1 volúmenes de acetato sódico 3 M, 10 µg de glicógeno y 2,5 volúmenes de etanol (96% v/v) seguido por 1 h de incubación a -20°C. El ADNc fraccionado por tamaño se recogió por centrifugación a 14000 rpm durante 20 minutos a 4°C. El sedimento de ADN se lavó con etanol al 70% y se secó al aire antes de disolverse en MQ. Se determinó la concentración de ADN usando el kit de cuantificación de ADNds PicoGreen™ (Molecular Probes).

Después de la digestión con EcoRI, se ligaron 200 ng de ADNc de ovocitos dentro de 500 ng de brazos de fago λGT11 predigeridos y desfosforilados en un tampón que contenía Tris-Cl 50 mM pH 7,8, MgCl₂ 10 mM, dihidro-treitol 10 mM, ATP 1 mM y 750 unidades/ml de ligasa T4 (Pharmacia). Las reacciones se incubaron durante una noche a 16°C. La reacción de ligación completa se empaquetó finalmente en un extracto de empaquetado Max Plax™ (Epicentre) como se describe en la hoja de información del producto.

Ejemplo 2

Aislamiento y caracterización de PAD6 de ratón

Amplificación por PCR de clones de fagos

Se incubaron placas únicas durante al menos una hora en 100 µl de tampón de fago λ (Tris-HCl 10 mM pH 8,3, NaCl 100 mM y MgCl₂ 10 mM). De cada placa eluida se amplificaron por PCR 2,5 µl usando cebadores de λGT11 (SEC ID N°: 5 y SEC ID N°: 6). Las reacciones de pCR se realizaron en el PE9700 (modo 9600, Perkin Elmer), un ciclo de 5 min a 94°C, 30 ciclos de 30 segundos a 94°C, 30 segundos a 55°C y 3 min a 72°C, seguido de un ciclo de 5 min a 72°C. Los productos de PCR se analizaron por electroforesis en gel de agarosa y se seleccionaron por tamaño, pureza y concentración. Solamente se seleccionaron para secuenciación bandas únicas de 500 pb o más.

Análisis de la secuencia de ADN

Se analizaron 750 clones de la biblioteca de ADNc de ovocitos de ratón por secuenciación del ADN después de insertar la amplificación por PCR. El análisis de la secuencia se realizó usando el protocolo de reacción preparado de secuenciación del ADN Big Dye (Perkin Elmer) y las muestras se analizaron en el secuenciador de ADN automático ABI377 (Perkin Elmer). Las secuencias se compararon frente a varias bases de datos a.o.: bases de datos gb111rod, genpept, EMrodESTs59 y EMhumanESTs59 usando BLASTN o TBLASTN en un procedimiento automatizado y anotado en base a la homología con un gen o genes de funciones conocidas.

Identificación y caracterización de PAD6

Una de las secuencias obtenida muestra homología fuerte con la peptidil arginina deiminasa III. Basándose en búsquedas de homología se ha establecido que este clon, 1B11, codifica una nueva peptidilarginina deiminasa que se ha denominado PAD6.

El extremo 5' del ADNc de PAD6 de ratón pudo amplificarse a partir de una biblioteca de ADNc de ovarios de ratón. El ADNc de esta biblioteca se había clonado de forma direccional en los sitios NotI-SaII (5'-3') del vector pSPORT (Life Technologies). Este vector contiene las secuencias 5' del promotor delantero M13 y SP6 del sitio NotI que se ha usado en la PCR RACE 5' en combinación con dos cebadores inversos específicos de la PAD6. La primera PCR se realizó con el cebador M13F (SEC ID N°: 7) y el cebador inverso específico del gen (SEC ID N°: 8). Este producto de PCR se diluyó cincuenta veces y se usó un microlitro de esta dilución como molde en la PCR anidada con

ES 2 274 977 T3

el cebador SP6 (SEC ID N°: 9) y el cebador inverso específico del gen anidado (SEC ID N°: 10). Ambas reacciones de PCR se realizaron en un volumen total de 50 μ l de tampón de reacción que contenía tampón de PCR Klen Taq 1X (Clontech), mezcla de NTPd 0,2 mM (Clontech) y mezcla de polimerasa Taq Advantage Klen 1X (Clontech) comenzando con 5 minutos de desnaturalización a 94°C seguido por 30 ciclos de 30 segundos a 94°C, 30 segundos a 56°C, 3 minutos a 72°C con una extensión final de 5 minutos a 72°C.

Se clonaron bandas en los productos de la PCR anidada en el vector TA Topo PCR2,1 (Invitrogen), siguiendo la hoja de información del producto, y se secuenciaron. Se descubrió que un fragmento RACE 5' de 1800 pb completaba el clon PAD6 de ratón. La secuencia del ADNc de ratón de longitud completa se da en la SEC ID N°: 2.

Basándose en la información de secuencia obtenida, se diseñaron conjuntos de cebadores de PCR específicos del gen y se usaron en experimentos de RT-PCR para confirmar el perfil de expresión específico de tejido. Los datos obtenidos (Figura 1) confirman la expresión específica del ovocito/ovario-(y testículos) de la PAD6 de ratón. (se usaron la SEC ID N°: 8 y la SEC ID N°: 13 como cebadores).

Hibridación *in situ* (ISH)

Para estudiar adicionalmente la expresión de la PAD6 en las gónadas, se realizó hibridación *in situ* (ISH) sobre secciones de ovarios y testículos de ratón.

Se fijaron ovarios de ratones de 7 días y adultos en formalina tamponada al 4% durante 24 horas a temperatura ambiente. Los tejidos se embebieron en parafina. Se cortaron secciones de parafina (5 μ m), se montaron sobre portaobjetos de microscopio Superfrost plus, y se dejaron secar durante toda la noche a 37°C. Los portaobjetos se endurecieron a 60°C durante dos horas.

Las secciones de tejido se desparafinaron en xileno y se rehidrataron en concentraciones descendentes de etanol. Los portaobjetos se lavaron durante 20 min en agitación en HCl 0,2 mM, seguido de dos lavados en DEPC (di-etilpirocarbonato) tratado con Milli Q. Las secciones se trataron con proteinasa K (1 μ g/ml) en tampón de digestión (Tris 100 mM, EDTA 50 mM pH 8) durante 30 min a 37°C. La digestión se paró en glicina al 0,2% pre-enfriada en PBS durante 10 min de agitación a temperatura ambiente (TA). Los portaobjetos se acetilaron durante 25 min con anhídrido acético al 0,25% en tampón de trietanolamina 0,1 M, seguido por dos lavados en DEPC tratado con Milli Q. Las secciones se prehibridaron a temperatura de hibridación en una cámara húmeda con mezcla de prehibridación, que contenía formamida al 52%, Tris 21 mM, EDTA 1 mM, NaCl 0,33 M, sulfato de dextrano al 10%, solución de Denhardt 1x, 100 μ g/ml de ADN de esperma de salmón, 100 μ g/ml de ARNt y 250 μ g/ml de ARN total de levadura. Los portaobjetos se cubrieron con un cubreobjetos de vidrio. Después de dos horas, los cubreobjetos se reemplazaron por cubreobjetos que tenían 100 μ l de mezcla de sonda de hibridación, que contenía mezcla de prehibridación y las siguientes adiciones: DTT 0,1 mM, tiosulfato sódico al 0,1%, SDS al 0,1% y 200 ng/ml de sonda marcada con DIG.

Las sondas marcadas con DIG se generaron por transcripción *in vitro* a partir de un molde de ADN lineal, usando DIG-UTPd y ARN polimerasas dependientes de ADN (SP6 y T7). El sitio del promotor de cada ARN polimerasa se unió a las secuencias específicas del gen permitiendo la generación de un fragmento de pCR que contenía el sitio promotor SP6 en el sitio 5' y el promotor T7 en el sitio 3'. En general, se hicieron sondas de aproximadamente 250-500 nucleótidos situadas en el extremo 5' de la SEC ID N°: 2. Después de la transcripción *in vitro*, se analizó una pequeña cantidad de la sonda en un gel de agarosa al 1,5% para confirmar la transcripción *in vitro* exitosa. Las concentraciones de la sonda se estimaron aplicando puntualmente diluciones en serie (incluyendo DIG-ARN de control (100 ng/ μ l)) sobre una membrana Hybond N⁺ seguido por fragmentos Fab' de fosfatasa alcalina anti-DIG (anti-DIG-AP) e incubación con el sustrato de color NBT/BCIP.

La hibridación se realizó durante toda la noche (16 horas) en una cámara húmeda a 42°C o 50°C. Los portaobjetos se lavaron en SSC 2x, se agitaron durante 15 min, seguido por lavados en SSC 2x, SSC 1x y SSC 0,1x durante 15 minutos de agitación a temperatura de hibridación. Las secciones se digirieron con Ribonucleasa A (20 μ g/ml) en tampón de ARNasa (NaCl 0,6 M, Tris 20 mM, EDTA 10 mM) durante 1 hora a 37°C. Después de dos lavados (5 min de agitación a TA) en PBS pre-enfriado y un lavado en tampón 1 (ácido maleico 100 mM, NaCl 150 mM), las secciones se incubaron durante 30 min en solución de bloqueo (1 g/ml de reactivo de bloqueo en tampón 1). Después las secciones se incubaron con anti-DIG-AP, diluido 1:500 en solución de bloqueo, durante 1 hora a TA. Después de dos lavados en tampón 1 (15 min de agitación a TA), los portaobjetos se secaron con cuidado alrededor del tejido y las secciones se rodearon con un DAKO-pen. Las secciones se cubrieron con reactivo de desarrollo de color NBT/BCIP y se incubaron en una cámara húmeda a TA. Después de dos horas, las secciones se examinaron al microscopio. Si no se observó tinción o sólo tinción débil, se continuó la incubación durante toda la noche a 4°C y el día siguiente a 60°C. Finalmente, los portaobjetos se aclararon con agua y opcionalmente se tiñeron por contraste con hematoxilina de Mayer 1:5 durante tres segundos. Los portaobjetos se montaron en gelatina de glicerol Kaisers.

Como se muestra en la Figura 2, PAD6 se expresa en el ovario exclusivamente en ovocitos.

El ARNm de PAD6 tiene altos niveles de expresión en ovocitos de folículos primarios, secundarios y antrales, pero se expresa también en ovocitos de folículos primordiales. Basándose en los datos obtenidos hasta ahora, el nivel de expresión del ARNm de PAD6 disminuye en ovocitos de folículos antrales, sugiriendo que la función de PAD6 se requiere más probablemente durante las fases tempranas de la ovogénesis. Aunque los datos de RT-PCR revelan

ES 2 274 977 T3

la expresión en los testículos de PAD6, no se detectó expresión por encima del nivel de fondo de ARNm de PAD6 usando análisis de ISH, sugiriendo niveles bajos de expresión de PAD6 en los testículos.

Ejemplo 3

5

Aislamiento y caracterización de PAD6 humana

Una búsqueda BLAST usando el ADNc de PAD6 de ratón de longitud completa como un interrogante frente a las bases de datos EM63hsGeno(nuevo) identificó el homólogo humano de PAD6. Esta búsqueda identificó sólo la región C-terminal de la secuencia codificante de la PAD6 humana. Para extender la secuencia en dirección 5' se diseñaron cebadores y se realizó una PCR RACE 5' sobre ADNc Marathon Ready de ovarios humanos (Clontech) usando el manual de usuario de ADNc Marathon Ready™. La primera PCR se realizó en las siguientes condiciones: una desnaturalización de 30 segundos a 94°C, 5 ciclos de 5 segundos a 94°C y 3 minutos a 72°C, 5 ciclos de 5 segundos a 94°C y 3 minutos a 70°C y 25 ciclos de 5 segundos a 94°C y 3 minutos a 68°C. Una dilución de 50 veces de este primer producto de PCR sirvió como molde en la segunda, la reacción de PCR anidada usando las mismas condiciones de PCR. Se clonó una banda esperada de ~650 pb en el vector TA Topo PCR2.1 (Invitrogen) y se secuenció. Este clon contenía (por homología) las primeras 500 pares de bases 5' de la secuencia codificante de la PAD6 humana, completando de ese modo la secuencia codificante de la PAD6 humana.

20 Se seleccionaron cebadores de PCR para amplificar el ARN de ovarios humanos de ADNc de PAD6 humana de longitud completa. Para aislar ADNc de PAD6 se usaron los cebadores de la SEC ID N°: 11 y de la SEC ID N°: 12 sobre ADNc de ovarios Marathon Ready (Clontech). Las condiciones de PCR fueron: desnaturalización durante 5 minutos a 94°C seguido por 5 ciclos de 30 segundos a 94°C y 3 minutos a 68°C, 28 ciclos de 30 segundos a 94°C, 30 segundos a 62°C y 3 minutos a 72°C con una extensión final de 7 minutos a 72°C.

25

Los amplificados de longitud completa de tres reacciones de PCR independientes se clonaron dentro del vector PCR2.1 Topo (Invitrogen) y se secuenciaron para determinar las secuencias de nucleótidos consenso. Su secuencia se muestra en la SEC ID N°: 4.

30 Se diseñaron conjuntos de cebadores de PCR específicos del gen (SEC ID N°: 14 y SEC ID N°: 15) y se usaron en experimentos de RT-PCR para determinar el perfil de expresión de la PAD6 humana. La RT-PCR de ARN de testículos, útero, riñón, timo, hígado, cerebro, corazón, pulmón y bazo humanos reveló la expresión de PAD6 sólo en los testículos (Figura 3).

35 Se hibridaron múltiples transferencias de Northern de tejidos (Clontech) de tejidos humanos con el fragmento de PCR de la PAD6 humana (aproximadamente 590 pb; producto de PCR de los cebadores de la SEC ID N°: 14 y la SEC ID N°: 15 que se extienden entre los nucleótidos 464-1052 en la SEC ID N°: 4). Las sondas se marcaron con [³²P] CTPd y perlas Ready to Go Labellings (AP Biotech) de acuerdo con las instrucciones del fabricante usando un tiempo de incubación de 60 minutos a 37°C. Los NTPd no incorporados se retiraron en una columna de centrifugación de Sephadex G50 en un jeringa de 1 ml.

40

Las transferencias se prehibridaron en Express hybmix (Clontech) durante al menos una hora a 65°C. Para la hibridación, se añadieron 4-8 x 10⁷ cpm de las sondas desnaturalizadas a la mezcla de prehibridación. Las transferencias se hibridaron a 65°C durante toda la noche y se lavaron una vez con SSC 2X, SDS 0,1% a temperatura ambiente, dos veces con SSC 1X, SDS al 0,1% a 65°C y una vez con SSC 0,1X, SDS al 0,1% a 65°C. Las transferencias hibridadas se analizaron con la cámara de fósforo STORM 840 (Molecular Dynamics), se escanearon a 200 micrómetros y se imprimieron con un intervalo de 0-50 después de la exposición de tres días a pantallas GP de fósforo de almacenamiento Kodak (Molecular dynamics).

50 En la Figura 4 puede verse una banda única de un ovario con una longitud estimada de aproximadamente 3 kB que aparece solo en ovarios. No pudo detectarse señal en testículos, lo más probable debido a que el nivel de expresión de PAD6 en testículos es demasiado bajo para detectarse en transferencias de Northern. Los análisis de hibridación *in situ* corroboran estos resultados: la expresión de PAD6 pudo detectarse en todos los tipos de folículos de ovarios humanos y de mono y en este aspecto son similares a los datos *in situ* en ratón. No se detectó expresión en testículos por hibridación *in situ* (datos no mostrados).

55

Ejemplo 4

Expresión de PAD6 humana y determinación de la actividad de PAD6

60

Clonación

Se clonó PAD6 humana de longitud completa dentro del vector de expresión bacteriano pGEX4T1 (AP Biotech) usando el kit de ligación de ADN Rapid (Boehringer). La construcción recombinante (pGEXhPAD6) se caracterizó por digestión con enzimas de restricción.

65

Se crecieron células de *E. coli* BL-21 transformadas con pGEXhPAD6 en medio YT 2 x a 25°C hasta una densidad celular de 1,0 a 650 nm. Después de la adición de 0,1 ml de isopropil-β-D-tiogalactopiranosido, se creció el cultivo

ES 2 274 977 T3

5 durante 5 horas adicionales a 25°C. Las células se centrifugaron y se resuspendieron en 0,1 volúmenes del volumen de cultivo original de Tris-HCl 20 mM, pH 7,6, EDTA 1 mM y se lisaron por sonicación en hielo. El sonicado se centrifugó a 15000 X g durante 30 min a 4°C (Sorvall, rotor SS34) y al sobrenadante se le añadió NaCl 1 M, 0,1% de Triton X-100 y perlas de glutatión-Sepharose 4B al 50% en PBS (Pharmacia Biotech, 1 ml a un equivalente de 250 ml del cultivo inicial), seguido por incubación a 4°C durante 60 minutos con agitación suave. Las perlas se lavaron después tres veces con 10 volúmenes de lecho de un tampón que contenía Tris-HCl 20 mM pH 7,6, EDTA 1 mM, Triton X-100 al 0,1% y NaCl 0,1 M a TA durante 5 minutos con agitación suave. La proteína de fusión PAD6h-gst recombinante se eluyó de las perlas en varias etapas con glutatión reducido de 10 a 100 mM en Tris-Cl 50 mM pH 8,0, NaCl 0,1 M y Triton X-100 al 0,1% a 4°C durante 30 minutos con agitación suave. Los eluidos se almacenaron con glicerol al 10% a -20°C para la determinación de la actividad enzimática. Se estimó que la pureza de la proteína era del 90% basándose en análisis de SDS PAGE.

Determinación de la actividad de la enzima PAD6

15 Se determinó la actividad de la PAD por la formación de citrulina en inhibidor de tripsina de soja (STI) como sustrato. Al contrario que la STI original, la STI citrulinada es incapaz de inhibir la actividad de la tripsina. Por lo tanto, un aumento en la actividad de la tripsina, detectada con un sustrato fluorescente de la tripsina, indica actividad de PAD.

20 Para la actividad de PAD, la mezcla de reacción consistía en HEPES 10 mM (pH 7,5), CaCl₂ 5 mM, DTT 2 mM, 0,17 µg de STI y una alícuota de la solución de la enzima purificada [GST-PAD6 o PAD disponible en el mercado (Sigma), derivada de músculo de conejo] en un volumen final de 20 µl. Después de la incubación de la mezcla de ensayo durante 30 minutos a 37°C, se añadieron posteriormente 10 µl del sustrato fluorescente N α -benzoil-L-Arginina-7-amido-4-metilcumarina [400 µM en HEPES 100 mM (pH 7,5), EDTA 50 mM] y 10 µl de solución de tripsina [0,25 µg en HEPES 100 mM (pH 7,5)]. Las medidas de fluorescencia (excitación a 360 nm, emisión a 460 nm) comenzaron directamente en un Victor V a temperatura ambiente, y continuaron durante una hora.

Pudo detectarse actividad de PAD6 como se puede ver en la Figura 5.

30

35

40

45

50

55

60

65

ES 2 274 977 T3

REIVINDICACIONES

5 1. Un polinucleótido que comprende una secuencia de nucleótidos que codifica la peptidilarginina deiminasa humana que se expresa exclusivamente en órganos reproductores y donde dicho polinucleótido hibrida en condiciones rigurosas con los nucleótidos 464-1052 de la SEC ID N°: 4.

2. Un polinucleótido de acuerdo con la reivindicación 1, codificando dicho polinucleótido la SEC ID N°: 3.

10 3. Un polinucleótido de acuerdo con la reivindicación 2, comprendiendo dicho polinucleótido los nucleótidos 20-2077 de la SEC ID N°: 4.

4. Un vector de expresión recombinante que comprende el polinucleótido de acuerdo con las reivindicaciones 1-3.

15 5. Un polipéptido que tiene actividad peptidilarginina deiminasa codificado por el polinucleótido de acuerdo con las reivindicaciones 1-3 o un vector de expresión de acuerdo con la reivindicación 4.

6. Una célula aislada transfectada con el polinucleótido de acuerdo con las reivindicaciones 1-3 o un vector de expresión de acuerdo con la reivindicación 4.

20 7. Una célula de acuerdo con la reivindicación 6 que es una célula transfectada estable que expresa el polipéptido de acuerdo con la reivindicación 5.

8. Un método para producir el polipéptido de la reivindicación 5 que comprende cultivar las células de la reivindicación 6 en condiciones en las que se produce dicha proteína y recuperar dicha proteína del cultivo.

25 9. Un método para identificar compuestos que afectan a la función enzimática de un polipéptido de acuerdo con la reivindicación 5, comprendiendo el método las etapas de

30 a) contactar dicho polipéptido con un sustrato que contiene arginina;

b) contactar dicha mezcla con un compuesto de ensayo;

c) medir la conversión de arginina en citrulina; y

35 d) comparar dicha transformación con la actividad de la peptidilarginina deiminasa en ausencia de un compuesto de ensayo.

40

45

50

55

60

65

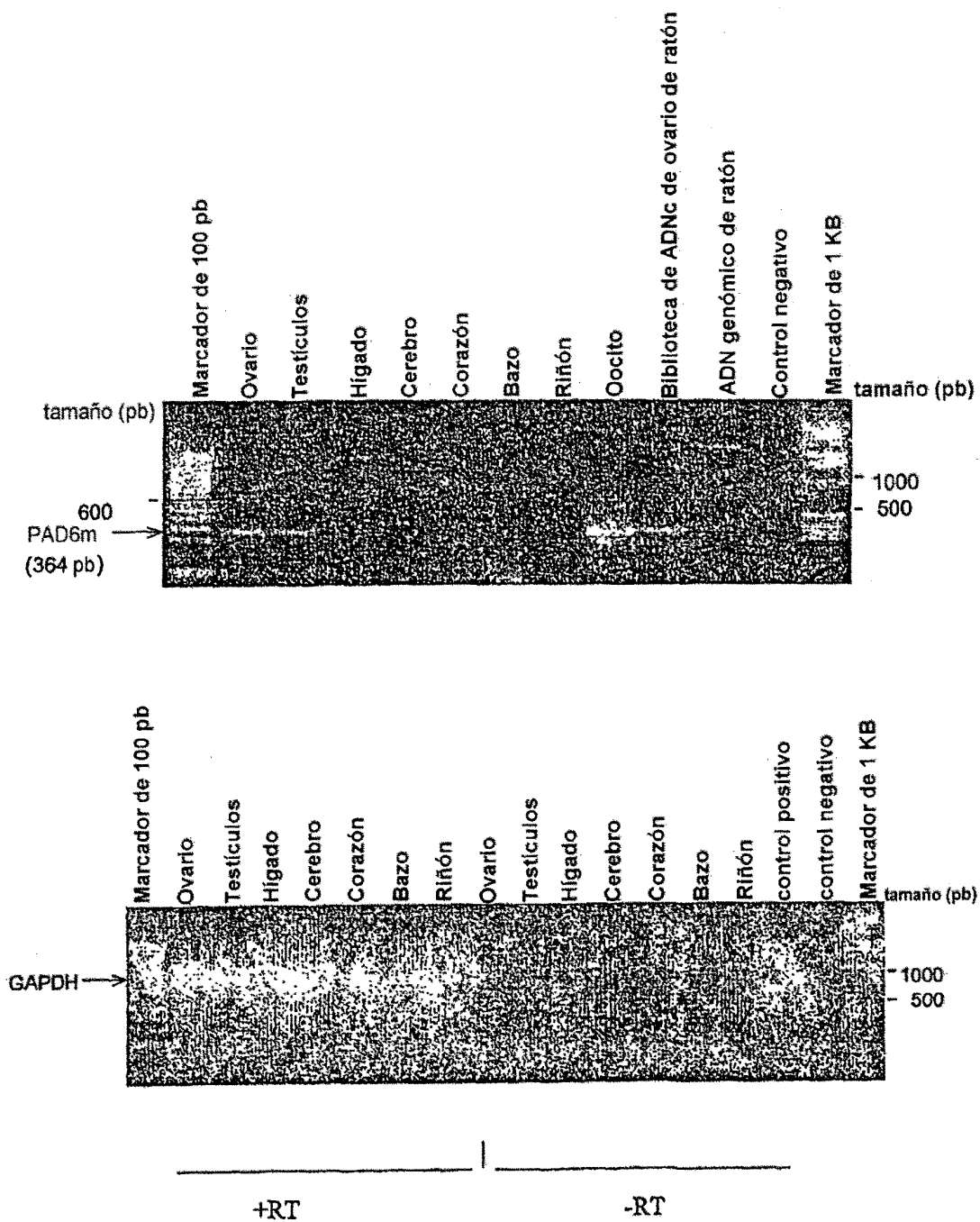


Figura 1

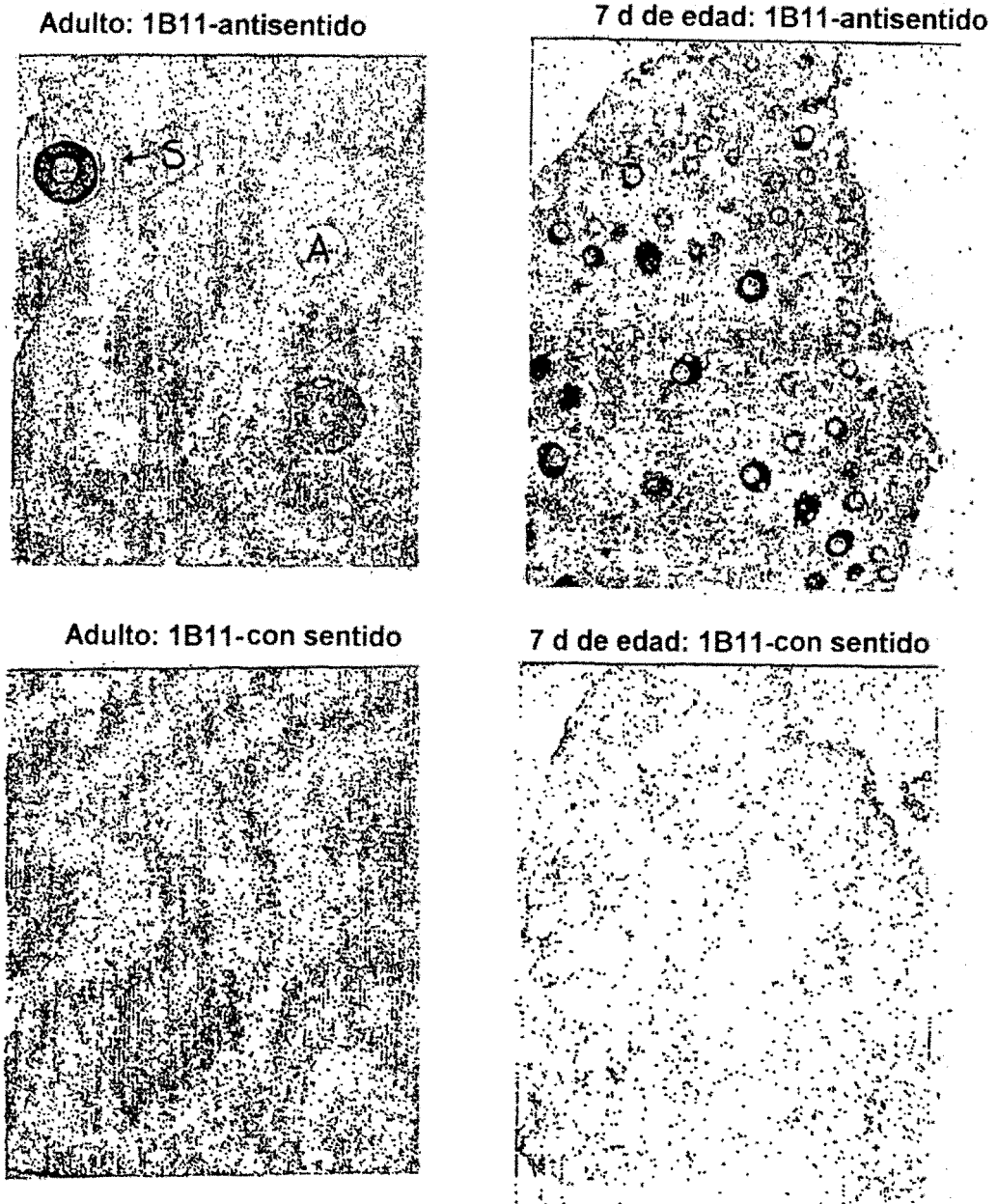


Figura 2

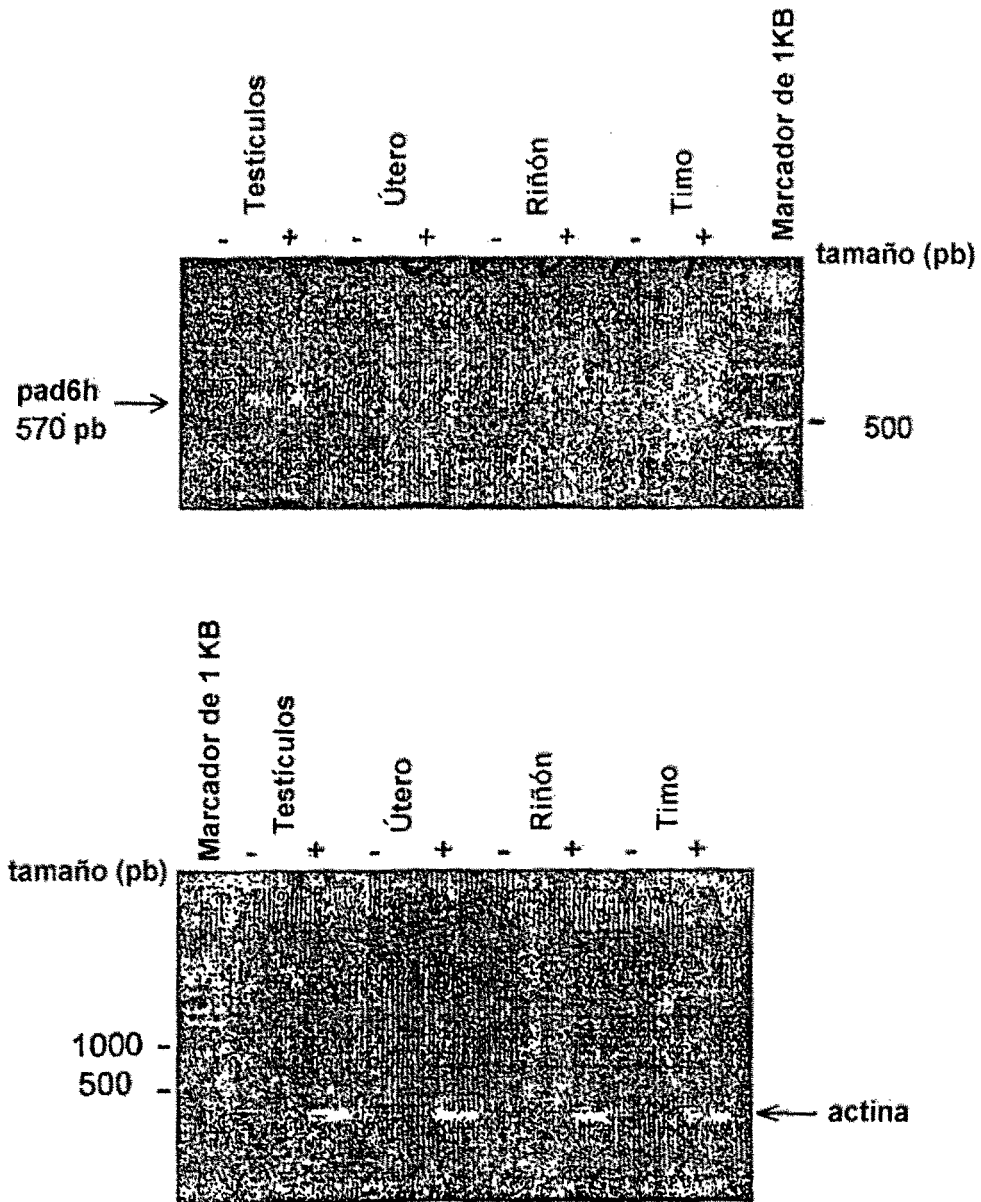


Figura 3

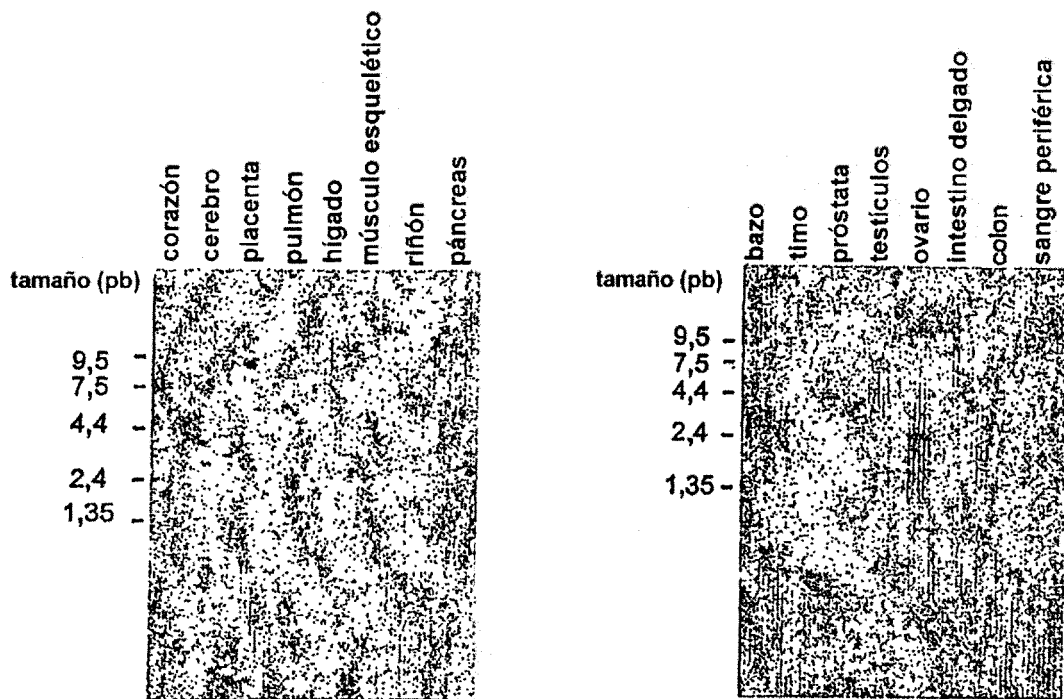


Figura 4

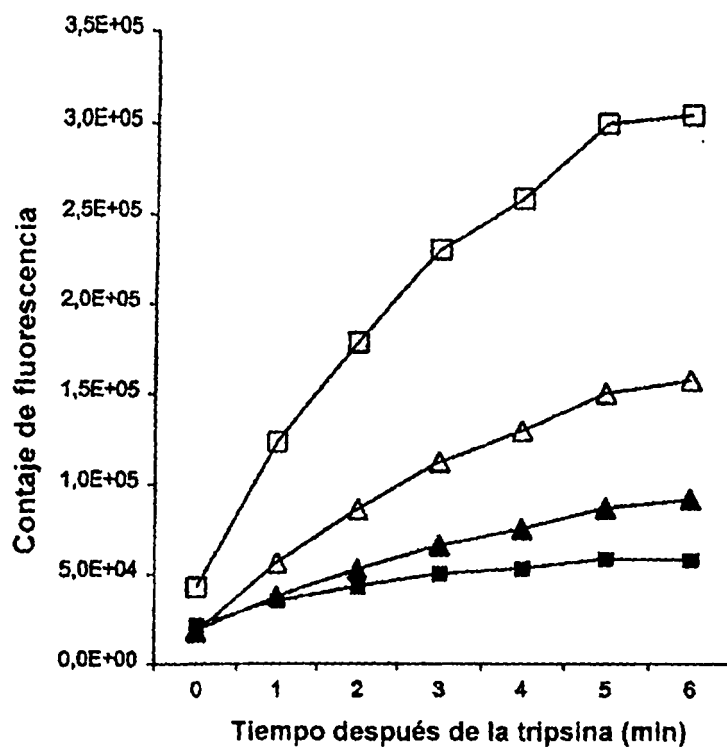


Figura 5

ES 2 274 977 T3

LISTA DE SECUENCIAS

<110> Akzo Nobel N. V.

5 <120> Peptidilarginina deiminasa 6
<130>

<140>

10 <141>

<160> 15

15 <170> PatentIn Ver. 2.1

<210> 1

<211> 682

20 <212> PRT

<213> *Mus musculus*

25 <400> 1

Met Ser Phe Gln Asn Ser Leu Ser Leu Ser Leu Val Asn Pro Thr His
1 5 10 15

30 Ala Leu Cys Met Val Gly Met Glu Ile Thr Leu Asp Ile Ser Lys Cys
20 25 30

35 Ala Pro Asp Lys Cys Lys Ser Phe Thr Ile Arg Gly Ser Pro Arg Ile
35 40 45

40 Leu Ile His Ile Ser Ser Ser Val Ile Ala Gly Lys Glu Asp Thr Val
50 55 60

Val Trp Arg Ser Met Asn His Pro Thr Val Ala Leu Val Arg Met Val
65 70 75 80

45 Ala Pro Ser Pro Thr Val Asp Glu Asp Lys Val Leu Val Ser Tyr Phe
85 90 95

50 Cys Pro Asp Gln Glu Val Pro Thr Ala Thr Ala Val Leu Phe Leu Thr
100 105 110

Gly Ile Glu Ile Ser Leu Glu Ala Asp Ile Tyr Arg Asp Gly Gln Leu
115 120 125

55 Asp Met Pro Ser Asp Lys Gln Ala Lys Lys Lys Trp Met Trp Gly Met
130 135 140

60 Asn Gly Trp Gly Ala Ile Leu Leu Val Asn Cys Ser Pro Asn Ala Val

65

ES 2 274 977 T3

	145				150					155				160		
5	Gly	Gln	Pro	Asp	Glu	Gln	Ser	Phe	Gln	Glu	Gly	Pro	Arg	Glu	Ile	Gln
					165					170				175		
10	Asn	Leu	Ser	Gln	Met	Asn	Val	Thr	Val	Glu	Gly	Pro	Thr	Ser	Ile	Leu
				180					185					190		
15	Gln	Asn	Tyr	Gln	Leu	Ile	Leu	His	Thr	Ser	Glu	Glu	Glu	Ala	Lys	Lys
			195					200						205		
20	Thr	Arg	Val	Tyr	Trp	Ser	Gln	Arg	Gly	Ser	Ser	Ala	Tyr	Glu	Leu	Val
	210						215					220				
25	Val	Gly	Pro	Asn	Lys	Pro	Val	Tyr	Leu	Leu	Pro	Thr	Phe	Glu	Asn	Arg
	225					230					235					240
30	Arg	Lys	Glu	Ala	Phe	Tyr	Val	Glu	Ala	Thr	Glu	Phe	Pro	Ser	Pro	Ser
				245						250					255	
35	Phe	Ser	Gly	Leu	Ile	Ser	Leu	Ser	Leu	Ser	Leu	Val	Glu	Lys	Ala	His
				260					265					270		
40	Asp	Glu	Cys	Ile	Pro	Glu	Ile	Pro	Leu	Tyr	Lys	Asp	Thr	Val	Met	Phe
			275					280					285			
45	Arg	Val	Ala	Pro	Tyr	Ile	Phe	Met	Pro	Ser	Thr	Gln	Met	Pro	Leu	Glu
	290						295					300				
50	Val	Tyr	Leu	Cys	Arg	Glu	Leu	Gln	Leu	Gln	Gly	Phe	Val	Asp	Ser	Val
	305					310					315					320
55	Thr	Lys	Leu	Ser	Glu	Lys	Ser	Lys	Val	Gln	Val	Val	Lys	Val	Tyr	Glu
				325						330					335	
60	Asp	Pro	Asn	Arg	Gln	Ser	Lys	Trp	Leu	Gln	Asp	Glu	Met	Ala	Phe	Cys
				340					345					350		
65	Tyr	Thr	Gln	Ala	Pro	His	Lys	Thr	Val	Ser	Leu	Ile	Leu	Asp	Thr	Pro
			355					360						365		
70	Arg	Val	Ser	Lys	Leu	Glu	Asp	Phe	Pro	Met	Lys	Tyr	Thr	Leu	Thr	Pro
	370						375					380				
75	Gly	Ser	Gly	Tyr	Leu	Ile	Arg	Gln	Thr	Glu	Asp	His	Arg	Val	Ala	Ser
	385					390					395					400
80	Leu	Asp	Ser	Ile	Gly	Asn	Leu	Met	Val	Ser	Pro	Pro	Val	Lys	Ala	Gln
				405						410					415	

ES 2 274 977 T3

5 Gly Lys Asp Tyr Pro Leu Gly Arg Val Leu Ile Gly Gly Ser Phe Tyr
420 425 430

Pro Ser Ser Glu Gly Arg Asp Met Asn Lys Gly Leu Arg Glu Phe Val
435 440 445

10 Tyr Ala Gln Gln Val Gln Ala Pro Val Glu Leu Phe Ser Asp Trp Leu
450 455 460

15 Met Thr Gly His Met Asp Gln Phe Met Cys Phe Val Pro Thr Asn Asp
465 470 475 480

Lys Asn Asn Asp Gln Lys Asp Phe Arg Leu Leu Leu Ala Ser Pro Ser
485 490 495

20 Ala Cys Phe Glu Leu Phe Glu Gln Lys Gln Lys Glu Gly Tyr Gly Asn
500 505 510

25 Val Thr Leu Phe Glu Asp Ile Gly Ala Glu Gln Leu Leu Ser Asn Gly
515 520 525

30 Arg Glu Ser Lys Thr Ile Ser Gln Ile Leu Ala Asp Lys Ser Phe Arg
530 535 540

Glu Gln Asn Thr Tyr Val Glu Lys Cys Ile Ser Leu Asn Arg Thr Leu
545 550 555 560

35 Leu Lys Thr Glu Leu Gly Leu Glu Asp Lys Asp Ile Ile Leu Ile Pro
565 570 575

40 Gln Leu Phe Cys Leu Glu Gln Leu Thr Asn Val Pro Ser Asn Gln Gln
580 585 590

45 Ser Thr Lys Leu Phe Ala Arg Pro Tyr Phe Pro Asp Met Leu Gln Ile
595 600 605

Ile Val Leu Gly Lys Asn Leu Gly Ile Pro Lys Pro Phe Gly Pro Lys
610 615 620

50 Ile Asn Gly Thr Cys Cys Leu Glu Glu Lys Val Cys Gly Leu Leu Glu
625 630 635 640

55 Pro Leu Gly Leu Lys Cys Thr Phe Ile Asp Asp Phe Asp Cys Tyr Leu
645 650 655

60 Ala Asn Ile Gly Asp Val Cys Ala Ser Ala Ile Ile Asn Arg Val Pro
660 665 670

Phe Ala Phe Lys Trp Trp Lys Met Thr Pro
675 680

65

ES 2 274 977 T3

<210> 2

<211> 2055

<212> ADN

5 <213> *Mus musculus*

<400>2

10 cagccatgtc ttttcagaac tcactcagcc tgtctctggt caatcccacc catgcoectct 60
 gcatggtagg catggaatc accttggaca tcagcaagtg tgcaccagac aagtgcaagt 120
 ctttcacccat ccgtggttcc cccaggatct tgatccacat ctctagctcc gtcacgctg 180
 15 gcaaagagga cactgtggtc tggaggtcaa tgaaccatcc cacagtggca ttggtgagga 240
 tgggtggcgcc cagccccact gtggatgaag acaaggtgct ggtctcctac ttctgtcctg 300
 accaagaagt ccccacggcc acagctgtgc tgtttctcac cggcatcgag atctccctgg 360
 aggcagacat ctatcgagat ggacaactgg acatgccaag tgataagcaa gctaagaaaa 420
 20 aatggatgtg gggatgaac ggctggggag ccatcctgct tgtgaattgt agccctaatt 480
 ctgtgggcca gcctgatgaa cagtcctttc aggagggccc cagagaaata cagaacctgt 540
 ctcatagtaa tgtaactgtg gagggccccca ccagcatcct acagaattac cagttgatcc 600
 tacatacctc cgaagaagag gcgaagaaga caagagtcta ctggtctcag agaggctcct 660
 ctgcgtatga actggtggtg ggacccaaca agcctgtcta tctcctgcct acctttgaga 720
 25 accgtaggaa agaggctttc tacgtagaag ccacggaatt cccatctccc agcttctcgg 780
 gcctgatctc cttgtcactc tccttagtag aaaaggetca cgacgagtgc atcccagaga 840
 ttccgctcta taaggataca gtgatgttcc ggggtggcacc ttatatcttc atgccagca 900
 30 ccagatgoc tctagaggtt tacctgtgca gggagctaca gctgcaaggc tttgtggact 960
 cagtgaccaa gctgagcgag aagagcaaaag tgcaggtggt aaaggtctat gaggacccca 1020
 accgccagag caagtggctc caggacgaga tggctttctg ctatactcag gctcctcaca 1080
 agacgggtgc attgatcctt gacaccccaa gggtttccaa gctggaagac tccccatga 1140
 35 aatacacact gacccctggc tctggctacc tgatccgaca aactgaggac caccgggtgg 1200
 ctagcctgga ttccatcggg aacctgatgg tatctccgcc tgtcaaggct cagggcaaaag 1260
 actaccctct agggagggtc ctcatgggtg gcagcttta ccccagctct gagggccggg 1320
 acatgaacaa gggcctgcca gaattcgtgt atgccagca ggtgcaggcc cctgtggaac 1380
 40 tcttctcgga ctggctgatg accggtcaca tggatcaatt catgtgcttt gtccctacca 1440
 atgataaaaa caacgaccag aaggacttcc gcctgctgct ggccagcccc agtgcctgct 1500
 ttgagctggt cgaacagaag cagaaggaag gctatgggaa cgtgaccctg tttgaagaca 1560
 ttggagcaga acagctcctt tctaattggga gggagagcaa aactatttcc caaatcctgg 1620
 45 ctgacaagag ttttcgagag cagaacacct atggtgagaa gtgtatcagc ctgaaccgca 1680
 cctcctgaa gacagaactg ggattggagg acaaggacat catcctgatc ccgagctct 1740
 tctgcctgga gcagctgacg aatgtcccct ccaaccagca gagcaccaa ctcttcgcca 1800
 ggccgtactt ccccgacatg ctgcagataa tcgtgttggg caagaacctt ggaatcccca 1860
 50 agccctttgg gcccaaatc aatggcacct gctgcctaga agagaaagtg tgtggattac 1920
 tggagccctt gggctcgaag tgcaocttca ttgatgattt tgactgctac ctggccaaca 1980
 taggggacgt ctgtgccagt gccatcataa acagggtgcc atttgcatc aagtgggtgga 2040
 agatgacccc ataaa 2055

<210> 3

60 <211> 686

<212> PRT

<213> *Homo sapiens*

65

ES 2 274 977 T3

<400> 3

5 Met Ser Phe Gln Ser Ile Ile His Leu Ser Leu Asp Ser Pro Val His
1 5 10 15

10 Ala Val Cys Val Leu Gly Thr Glu Ile Cys Leu Asp Leu Ser Gly Cys
20 25 30

15 Ala Pro Gln Lys Cys Gln Cys Phe Thr Ile His Gly Ser Gly Arg Val
35 40 45

20 Leu Ile Asp Val Ala Asn Thr Val Ile Ser Glu Lys Glu Asp Ala Thr
50 55 60

25 Ile Trp Trp Pro Leu Ser Asp Pro Thr Tyr Ala Thr Val Lys Met Thr
65 70 75 80

30 Ser Pro Ser Pro Ser Val Asp Ala Asp Lys Val Ser Val Thr Tyr Tyr
85 90 95

35 Gly Pro Asn Glu Asp Ala Pro Val Gly Thr Ala Val Leu Tyr Leu Thr
100 105 110

40 Gly Ile Glu Val Ser Leu Glu Val Asp Ile Tyr Arg Asn Gly Gln Val
115 120 125

45 Glu Met Ser Ser Asp Lys Gln Ala Lys Lys Lys Trp Ile Trp Gly Pro
130 135 140

50 Ser Gly Trp Gly Ala Ile Leu Leu Val Asn Cys Asn Pro Ala Asp Val
145 150 155 160

55 Gly Gln Gln Leu Glu Asp Lys Lys Thr Lys Lys Val Ile Phe Ser Glu
165 170 175

60 Glu Ile Thr Asn Leu Ser Gln Met Thr Leu Asn Val Gln Gly Pro Ser
180 185 190

65 Cys Ile Leu Lys Lys Tyr Arg Leu Val Leu His Thr Ser Lys Glu Glu
195 200 205

70 Ser Lys Lys Ala Arg Val Tyr Trp Pro Gln Lys Asp Asn Ser Ser Thr
210 215 220

75 Phe Glu Leu Val Leu Gly Pro Asp Gln His Ala Tyr Thr Leu Ala Leu
225 230 235 240

ES 2 274 977 T3

5 Ala Ser Pro Ser Ala Cys Tyr Lys Leu Phe Arg Glu Lys Gln Lys Glu
500 505 510

10 Gly Tyr Gly Asp Ala Leu Leu Phe Asp Glu Leu Arg Ala Asp Gln Leu
515 520 525

15 Leu Ser Asn Gly Arg Glu Ala Lys Thr Ile Asp Gln Leu Leu Ala Asp
530 535 540

20 Glu Ser Leu Lys Lys Gln Asn Glu Tyr Val Glu Lys Cys Ile His Leu
545 550 555 560

25 Asn Arg Asp Ile Leu Lys Thr Glu Leu Gly Leu Val Glu Gln Asp Ile
565 570 575

30 Ile Glu Ile Pro Gln Leu Phe Cys Leu Glu Lys Leu Thr Asn Ile Pro
580 585 590

35 Ser Asp Gln Gln Pro Lys Arg Ser Phe Ala Arg Pro Tyr Phe Pro Asp
595 600 605

40 Leu Leu Arg Met Ile Val Met Gly Lys Asn Leu Gly Ile Pro Lys Pro
610 615 620

45 Phe Gly Pro Gln Ile Lys Gly Thr Cys Cys Leu Glu Glu Lys Ile Cys
625 630 635 640

50 Cys Leu Leu Glu Pro Leu Gly Phe Lys Cys Thr Phe Ile Asn Asp Phe
645 650 655

55 Asp Cys Tyr Leu Thr Glu Val Gly Asp Ile Cys Ala Cys Ala Asn Ile
660 665 670

60 Arg Arg Val Pro Phe Ala Phe Lys Trp Trp Lys Met Val Pro
675 680 685

<210> 4

<211> 2092

55 <212> ADN

<213> *Homo sapiens*

60

65

ES 2 274 977 T3

<400> 4

5 tagcgtggag ggccgagcca tgtccttcca gagtatcate cacctgtccc tggacagccc 60
 tgtccatgcc gtttgtgtgt tgggcacaga aatctgcttg gatctcagcg ggtgtgcccc 120
 ccagaagtgc cagtgcttca ccatccatgg ctctgggagg gtcttgatcg atgtggccaa 180
 cacggtgatt tctgagaagg aggacgccac catctggtgg cccctgtctg atcccacgta 240
 10 cgccacagtg aagatgacat cgccagccc ttccgtggat gcggataagg tctcggtcac 300

atactatggg cccaacgagg atgccccgt gggcacagct gtgctgtacc tcactggcat 360
 tgaggtctct ctagaggtag acatctaccg caatgggcaa gttgagatgt caagtgacaa 420
 15 acaggctaag aaaaaatgga tctggggctc cagcggttgg ggtgccatcc tgcttgtaa 480
 ttgcaaccct gctgatgtgg gccagcaact tgaggacaag aaaaccaaga aagtgatctt 540
 ttcagaggaa ataacgaatc tgtccdagat gactctgaat gtccaaggcc ccagctgtat 600
 cttaaagaaa tatcggctag tcctccatac ctccaaggaa gagtcgaaga aggcgagagt 660
 20 ctactggccc caaaaagaca actccagtac ctttgagttg gtgctggggc ccgaccagca 720
 cgctataacc ttggccctcc tcgggaacca cttgaaggag actttctacg ttgaagctat 780
 agcattccca tctgccgaat tctcaggcct catctcctac tctgtgtccc tggaggagga 840
 gtotcaagac ccgtcaatte cagagactgt gctgtacaaa gacacggtgg tgttccgggt 900
 25 ggctccctgt gtcttcatte cctgtaccca ggtgcctctg gaggtttacc tgtgcaggga 960
 gctgcagctg cagggttttg tggacacagt gacgaagctg agtgagaaga gcaacagcca 1020
 ggtggcatct gtctatgagg accccaaccg cctgggcagg tggctccagg atgagatggc 1080
 cttctgctac acccaggctc cccacaagac aacgtccttg atcctcgaca cacctcaggc 1140
 30 cgccgatctc gatgagttcc ccatgaagta ctcactgagc cctggatttg gctacatgat 1200
 ccaggacact gaggaccata aagtggccag catggattcc attgggaacc tgatggtgtc 1260
 cccacctgtc aaggtccaag ggaaagagta cccgctgggc agagtccca ttggcagcag 1320
 cttttacccc agcgcagagg gccgggcat gagtaagacc ctccgagact tcctctatgc 1380
 35 ccagcaggtc caagcgccgg tggagctcta ctcagattgg ctaatgactg gccacgtgga 1440
 tgagttcatg tgcttcatec ccacagatga caagaatgag ggcaaaaagg gcttccctgct 1500
 gctcctggcc agccccagtg cctgctataa actgttccga gagaacaga aggaaggcta 1560
 40 tggcgacgct cttctgtttg atgagcttag agcagatcag ctctgtota atggaaggga 1620
 agccaaaacc atcgaccaac ttctggctga tgaaagcctg aagaagcaga atgaatacgt 1680
 ggagaagtgc attcacctga accgtgacat cctgaagacg gagctgggcc tgggtggaaca 1740
 ggacatcadc gagattcccc agctgttctg cttggagaag ctgactaaca tcccctctga 1800
 45 ccagcagccc aagaggtcct ttgcgagcc atacttccct gacctgttgc ggatgattgt 1860
 gatgggcaag aacctgggga tccccagcc ttttgggccc caaatcaagg ggacctgctg 1920
 cctggaagaa aagatttget gcttgctgga gccctgggc ttcaagtgca ccttcatcaa 1980
 50 tgactttgac tgttacctga cagaggtcgg agacatctgt gcctgtgcca acatccgccg 2040
 ggtgcccttt gccttcaaat ggtggaagat ggtacctag acccagggcc ta 2092

55 <210> 5

<211> 24

<212> ADN

<213> Secuencia artificial

60

<220>

<223> Descripción de secuencia artificial: cebador sintético

65 <400> 5

ttgacaccag accaactggt aatg

24

ES 2 274 977 T3

	<210> 6	
	<211> 24	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
10	<400> 6	
	ggtggcgacg actcctggag cccg	24
15	<210> 7	
	<211> 17	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
25	<400> 7	
	gttttcccag tcacgac	17
30	<210> 8	
	<211> 20	
	<212> ADN	
35	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
40	<400>8	
	cgtcagctgc tccagcagaa	20
45	<210> 9	
	<211> 19	
	<212> ADN	
50	<213> Secuencia artificial	
	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
55	<400> 9	
	gatttagtg acactatag	19
60	<210> 10	
	<211> 20	
	<212> ADN	
65	<213> Secuencia artificial	

ES 2 274 977 T3

	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
5	<400> 10	
	gtgcggttca ggctgataca	20
10	<210> 11	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
20	<400> 11	
	agcgtggagg gccgagccat g	21
25	<210> 12	
	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
35	<400> 12	
	agggcctggg tctaaggtac catc	24
40	<210> 13	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
45	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
50	<400> 13	
	ctgatgaccg gcacatggat	20
55	<210> 14	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
60	<220>	
	<223> Descripción de secuencia artificial: cebador sintético	
65	<400> 14	
	ggcggttggg gtcctcatag	20

ES 2 274 977 T3

<210> 15

<211> 20

<212> ADN

5 <213> Secuencia artificial

<220>

10 <223> Descripción de secuencia artificial: cebador sintético

<400> 15

15 gccatcctgc ttgtgaattg

20

20

25

30

35

40

45

50

55

60

65