



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 278 392**

(51) Int. Cl.:

C12N 15/54 (2006.01)

C12N 9/12 (2006.01)

C12N 5/10 (2006.01)

C12N 15/62 (2006.01)

C07K 16/40 (2006.01)

G01N 33/50 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **97931459 .8**

(86) Fecha de presentación : **26.06.1997**

(87) Número de publicación de la solicitud: **0939821**

(87) Fecha de publicación de la solicitud: **08.09.1999**

(54) Título: **Fosfatidilinositol 3' quinasa regulada por una proteína G beta-gamma.**

(30) Prioridad: **27.06.1996 US 672211**

(45) Fecha de publicación de la mención BOPI:
01.08.2007

(45) Fecha de la publicación del folleto de la patente:
01.08.2007

(73) Titular/es: **ONYX PHARMACEUTICALS, Inc.**
2100 Powell Street, 12th Floor
Emeryville, California 94608, US

(72) Inventor/es: **Stephens, Len;**
Hawkins, Phillip, T. y
Brasemann, Sylvia

(74) Agente: **Zea Checa, Bernabé**

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

Fosfatidilinositol 3' quinasa regulada por una proteína G beta-gamma.

5 1. Introducción

La presente invención se refiere a una nueva subunidad de un enzima de fosfoinositida 3OH-quinasa regulada por proteína G aislada de células del linaje hematopoyético que están implicadas en las vías de transducción de la señal celular, y al uso de esta nueva subunidad en el tratamiento y al diagnóstico de enfermedades.

10 2. Antecedentes de la invención

Las fosfoinositidas 3OH-quinasas (PI3Ks) son una gran familia de enzimas capaces de fosforilar en la posición 3 como mínimo una de las fosfoinositidas celulares (Whitman *et al.*, 1988, Nature 332:644-646; Auger *et al.*, 1989, Cell 57:167-175). Las fosfoinositidas fosforiladas en la posición 3 se encuentran en todas las células de eucariotas superiores. Cada vez existen más pruebas de la implicación de PI3K y de un producto lipídico de esta enzima, fosfatidilinositol (3,4,5)-trifosfato (en donde en adelante "PtdIns(3,4,5)P₃"), como parte de un segundo sistema importante y nuevo de mensajero en la transducción de la señal celular. Los componentes de este nuevo sistema de señalización basado en PtdIns(3,4,5)P₃ son *independientes* de la vía de señalización caracterizada previamente basada en los fosfolípidos de inositol, en los que una fosfoinositidasa C (PIC) hidroliza PtdIns(4,5)P₂ para liberar los mensajeros secundarios estructuralmente diferentes inositol (1,4,5) - trifosfato (Ins(1,4,5)P₃) y diacilglicerol.

Los factores de crecimiento y los agonistas extracelulares seleccionados estimulan la actividad de PI3K intracelular y provocan la acumulación intracelular rápida y transitoria de PtdIns(3,4,5)P₃. Sorprendentemente, la estimulación de diferentes tipos de receptores de la superficie celular, incluyendo tirosina quinasas de receptores, receptores asociados con las tirosina quinasas no asociadas a receptores de la familia del *src*, factores de crecimiento de citoquinas, y receptores asociados a proteína G activan a miembros de la familia de PI3K. (Revisión en Stephens *et al.*, 1993, de Biochemica et Biophysica Acta, 1179:27-75). Por ejemplo, los receptores tirosina quinasa que, cuando se activan, producen un aumento en la acumulación de PtdIns(3,4,5)P₃ son el receptor de PDGF, el receptor de EGF, miembros de la familia de los receptores de FGF, el receptor de CSF-1, el receptor de la insulina, el receptor de IGF-1, y el receptor de NGF. Los receptores asociados con las tirosina quinasas no asociadas a receptores de la familia del *src*, que estimulan la acumulación de PtdIns(3,4,5)P₃ son el receptor IL-2, el receptor IL-3, el receptor de mIgM, el receptor de CD4, el receptor de CD2, y el receptor células de CD3/T. Además, el receptor IL-4 de citoquinas y el receptor de trombina asociado a proteína G, el receptor de ATP, y el receptor de fMLP estimulan la actividad de una PI3K, resultando en la acumulación posterior de PtdIns(3,4,5)P₃. De este modo, PtdIns(3,4,5)P₃ es un mensajero secundario en vías de señalización extremadamente diferentes.

Soporte en la hipótesis que la actividad de PI3K y la producción de PtdIns(3,4,5)P₃ es una vía fisiológica relevante de la transducción de la señal para estos receptores diferentes proviene, entre otros, de dos líneas diferentes de evidencias experimentales: inhibición de la actividad de PI3K mediante metabolitos fúngicos y observaciones de las asociaciones directas de proteínas. La wortmanina, un metabolito fúngico, inhibe de forma irreversible la actividad de PI3K uniéndose covalentemente al dominio catalítico de esta enzima. La inhibición de la actividad de PI3K por la wortmanina elimina la respuesta celular posterior del factor extracelular. Por ejemplo, los neutrófilos responden a la quimiocina fMet-Leu-Phe (fMLP) estimulando PI3K y sintetizando PtdIns(3,4,5)P₃. La síntesis se correlaciona con la activación de la explosión respiratoria implicada en la destrucción de neutrófilos de microorganismos invasores. El tratamiento de los neutrófilos con la wortmanina previene la respuesta de la explosión respiratoria inducida por fMLP. Thelen *et al.*, 1994, PNAS, USA 91:4960-4964. De hecho, estos experimentos con la wortmanina, así como otras pruebas experimentales, muestran que la actividad de PI3K en células del linaje hematopoyético, particularmente neutrófilos, monocitos, y otros tipos de leucocitos, está implicada en muchas de las respuestas inmunes de no memoria asociadas a la inflamación aguda y crónica.

Las enzimas PI3K interactúan directamente, y se pueden purificar conjuntamente con, las formas activadas de diferentes tirosina quinasas de receptores. Cuando están purificadas, las PI3K asociadas a tirosina quinasas de receptores consisten en heterodímeros de 170-200 kD (Otsu *et al.*, 1991, Cell 65:91-104, Pons *et al.*, 1995, Mol. Cell. Biol. 15:4453-4465, Inukai *et al.*, 1996, J. Biol. Chem. 271:5317-5320) y comprenden una subunidad catalítica y una subunidad adaptadora (o reguladora).

Dos homólogos diferentes de la subunidad catalítica, p110 α y p110 β han sido descritos y clonados. La subunidad catalítica, que se une irreversiblemente a wortmanina, se asocia fuertemente a uno u otro miembros de una pequeña familia de subunidades reguladoras altamente relacionadas, p55 α , p55P1K, p85 α y p85 β , para formar los heterodímeros de 170-200 kD. Las subunidades reguladoras conocidas contienen una gran colección de dominios de interacción proteína:proteína, incluyendo dos dominios SH2 (Cantley *et al.*, 1991, Cell 64:281-302).

Se cree que la presencia de los dominios SH2 es responsable de la unión y estimulación de los heterodímeros de PI3K a las tirosina quinasas de receptor activadas. Los receptores activados se fosforilan en residuos de tirosina claves en las secuencias locales consenso preferidas por los dominios SH2 encontrados en los adaptadores de PI3K de 55-87 kD (Songyang *et al.*, 1993, Cell 72:767-778). Una vez se une el heterodímero de PI3K, activa directamente la subunidad catalítica de PI3K (aunque este efecto es relativamente pequeño *in vitro*, Carpenter *et al.*, 1993, J. Biol.

Chem. 268:9478-9483, Backer *et al.*, 1992, EMBO J. 11: 3469-3479) y transloca el PI3K citosólico a una fuente de su sustrato fosfolipídico. La combinación de estos factores conduce a un aumento en la producción de PtdIns(3,4,5)P₃. Claramente, estas isoformas de PI3K (p100 α /p110 β /p55 α , p55PIK) parecen adaptadas estructuralmente para funcionar como transductores de señal “downstream” de las tirosina quinasas reguladas por receptores, de forma muy similar a la forma en la que la familia τ de PI-PLCs se encuentran reguladas por tirosina quinasas sensibles a receptor (Lee and Rhee, 1995, Current Biol. 7:193-189).

Sin embargo, la subfamilia p110/p85 de PI3Ks no parece estar implicada en la producción de PtdIns(3,4,5)P₃ que puede suceder como resultado de la activación de los receptores de la superficie celular que utilizan GTPasas heterotriméricas para transducir sus señales (*por ejemplo* fMLP, PAF, ATP, y trombina). Estos tipos de receptores de la superficie celular se han descrito principalmente en células de origen hematopoyético cuya activación está implicada en respuestas inflamatorias del sistema inmune. Pruebas recientes sugieren que una forma cromatográficamente diferente de PI3K sensible a wortmanina está presente en las células U937 y en los neutrófilos con una masa molecular nativa, relativa aproximadamente de aproximadamente 220 kD (Stephens *et al.*, 1994, Cell 77:83-93). Esta actividad de PI3K se puede estimular específicamente por subunidades G $\beta\gamma$, pero no por subunidades G α -GTP. Una actividad similar de PI3K también se ha descrito en una línea celular de Osteosarcoma (Morris *et al.*, 1995, Mol. Pharm. 48:532-539). Las plaquetas también contienen PI3K sensible a G $\beta\gamma$, aunque no está claro si pertenece a un miembro de la familia de PI3K de p85/p110 (Thomason *et al.*, 1994, J. Biol. Chem. 269:16525-16528). Parece probable que esta PI3K sensible a G $\beta\gamma$ mal caracterizada pudiera ser responsable de la producción de PtdIns(3,4,5)P₃ en respuesta a los agonistas como el ATP, fMLP etc.

Stoyanov *et al.*, (1995) han publicado recientemente la clonación y la expresión de una PI3K selectiva de PtdIns(4,5)P₂ sensible a wortmanina llamada p110 γ , a partir de una librería de ADNc de médula ósea humana. Se amplificó p110 γ por PCR usando cebadores diseñados para reconocer PI3Ks potenciales así como centros catalíticos de PtdIns4-quinasas. Es claramente diferente de p110 α y p110 β , ya que carece, por ejemplo, de un dominio de unión amino-terminal para un miembro de la familia adaptadora p85. Se especuló que p110 γ fuese la actividad de PI3K “downstream” de receptores heterotriméricos ligados a GTPasa en base a su sensibilidad a las subunidades G α -GTP y G $\beta\gamma$ *in vitro* y su expresión en células procedentes de mielóide. Sin embargo, esta hipótesis dejó varias preguntas sin resolver con respecto a la evidencia bioquímica anterior que indicaba que el PI3K sensible a G $\beta\gamma$ no se estaba estimulado por las subunidades de G α -GTP, y que ésta poseía una masa molecular mucho mayor de aproximadamente 220 kD.

Se ha sugerido que los efectos de las subunidades G $\beta\gamma$ en p110 γ están mediados a través de un dominio de homología “pleckstrin” (PH) NH₂-terminal putativo. Sin embargo, con la descripción de un incremento en el número de efectores regulados por G $\beta\gamma$, pruebas crecientes sugieren que los dominios de PH no representan un dominio de unión de G $\beta\gamma$ ampliamente utilizado. Un trabajo reciente, usando un grupo de péptidos relativamente pequeños basados en la secuencia de los dominios encontrados solamente en las adenilato ciclasas activadas por G $\beta\gamma$ (ACs 2 y 4) que bloquean específicamente la activación de G $\beta\gamma$ o la inhibición de diferentes efectores, ha sugerido que puede haber algunos argumentos para creer que las subunidades G $\beta\gamma$ contienen un dominio de activación de efector ampliamente utilizado. Además, regiones en diversos efectores que interaccionan con este dominio de activación de efectores muestran similitudes de secuencia significativas. Por lo tanto un motivo (Gln-X-X-Glu-Arg) dentro del dominio en AC2 destacado por estos estudios de péptidos también aparece en regiones de canales de potasio y β -ARKs y está implicado en la regulación por subunidades G $\beta\gamma$ (Chen *et al.* 1995, Science 268:1166-1169). Sin embargo, este motivo no se replica en todas las proteínas conocidas que son reguladas por subunidades G $\beta\gamma$, y por lo tanto el análisis de secuencia no puede predecir actualmente si una proteína es regulada por las subunidades G $\beta\gamma$.

Falta la identificación del mecanismo por el cual la actividad de PI3K es activada por agonistas celulares que transducen sus señales a través de receptores unidos a proteína G. Es importante tener en cuenta que la gran mayoría de los agonistas que activan la explosión respiratoria de neutrófilos implicada en la respuesta inflamatoria se unen a receptores asociados a proteína G en lugar de a tirosina quinasas de receptor. Por lo tanto, el mecanismo por el cual PI3K se encuentra regulado en respuesta a este tipo de quimiocinas es probable que sea muy diferente de la regulación por los factores del crecimiento cuya señalización es mediante tirosina quinasas. La presente invención está dirigida hacia la resolución de esta cuestión mediante la identificación, la purificación, y la clonación de una nueva y específica forma de PI3K que es activada por las subunidades $\beta\gamma$ de proteínas G triméricas.

3. Resumen de la invención

La presente invención se refiere al descubrimiento, la identificación, la purificación, y la clonación de los nucleótidos que codifican la subunidad p101 de PI3K regulada por proteína G trimérica, una nueva proteína que produce la acumulación del mensajero secundario PtdIns(3,4,5)P₃ en respuesta a la activación de receptores unidos a proteína G. Esta nueva PI3K regulada por proteína G comprende una subunidad catalítica, p120, y una subunidad reguladora, p101. La subunidad catalítica p120 comparte homología parcial en la secuencia de aminoácidos con las subunidades catalíticas p110 α , p110 β , y p101 γ de PI3K. Por otra parte p101 es una proteína totalmente nueva sin homología identificable con cualquier otra secuencia previamente disponible. En ausencia de la subunidad reguladora p101, las subunidades de proteína G activadas inducen un estímulo suave de la actividad catalítica por la subunidad catalítica p120. Sin embargo, en presencia de la subunidad p101, la actividad de PI3K de la subunidad catalítica se encuentra estimulada más de 100 veces por las proteínas G activadas.

Los ADNc que codifican tanto p101 y p120 de cerdo y p101 y p120 humanas se han clonado y secuenciado, y se describen en esta memoria. Los ADNc de p101 y p120 de cerdo codifican proteínas de aproximadamente 877 aminoácidos y de aproximadamente 1102 aminoácidos, respectivamente (Fig.2 y Fig. 4), mientras que los ADNc de p101 y p120 humanos codifican proteínas de aproximadamente 880 aminoácidos y de aproximadamente 1102 aminoácidos, respectivamente (Fig. 11 y Fig. 13). Aunque la secuencia de aminoácidos de las proteínas p120 son homólogas a las de otras subunidades catalíticas conocidas de PI3K, los ADNc de p120 descritos en esta memoria divergen de los ADNc que codifican las subunidades catalíticas conocidas de PI3K, particularmente en, por ejemplo, el extremo carboxilo (residuos aminoácidos del 1075 a 1102).

El transcrito de p101 se encuentra principalmente en células del linaje hematopoyético. p120, y otras proteínas de subunidades catalíticas de PI3K, tienen una distribución tisular y celular más amplia. Notablemente, la presencia de actividad de PI3K sensible a proteína G trimérica se ha encontrado solamente en un número limitado de células del linaje hematopoyético (*por ejemplo* neutrófilos, plaquetas, etc.). De este modo, la capacidad de activar las enzimas PI3K en respuesta a la estimulación de los receptores asociados a proteína G trimérica parece en gran parte dependiente de la presencia de la subunidad p101.

La invención engloba los siguientes nucleótidos, células hospedadoras que expresan estos nucleótidos, y productos de expresión de estos nucleótidos: (a) nucleótidos que codifican proteínas p101 de mamíferos, incluyendo la p101 de cerdo y la p101 humana, y los productos genéticos de p101, incluyendo los productos genéticos humanos y de cerdo; (b) nucleótidos que codifican porciones de p101 que corresponden a sus dominios funcionales, y los productos polipeptídicos especificados por estas secuencias nucleotídicas, incluyendo pero no limitado a los nucleótidos de p101 que codifican, por ejemplo, el dominio de interacción $G\beta\gamma$ de p101, o el dominio de asociación a la subunidad catalítica, o residuos aminoácidos de p101 de cerdo que se extienden de aproximadamente 1 a 160, 80-120, 161 a 263, 264 a 414, 415 a 565, 566 a 706, 707-832, y/o 833 a 877, o residuos aminoácidos de p101 humano que se extienden de aproximadamente 1 a 160, 80-120, 161 a 263, 264 a 416, 417 a 567, 568 a 709, 710-835, y/o 836 a 880; (c) los nucleótidos que codifican mutantes de p101 en los que todo o una parte de uno de los dominios se suprime o se altera, y los productos polipeptídicos especificados por estas secuencias de nucleótidos, incluyendo pero no limitado a los mutantes de p101 en que los nucleótidos que codifican el dominio de interacción de proteína G, o el dominio de asociación a la subunidad catalítica, o los residuos aminoácidos de aproximadamente 1 a 160, 80-120, 161 a 263, 264 a 414, 415 a 565, 566 a 706, 707-832, y/o 833 a 877 se suprimen; (d) nucleótidos que codifican proteínas de fusión que contienen la proteína p101 o uno de sus dominios (*por ejemplo*, el dominio de interacción de $G\beta\gamma$, o el dominio de asociación a la subunidad catalítica, o los dominios descritos por los residuos aminoácidos aproximadamente 1 a 160, 80-120, 161 a 263, 264 a 414, 415 a 565, 566 a 706, 707-832, y/o 833 a 877), fusionado a otro polipéptido.

También se describe en esta memoria a los agonistas y antagonistas de PI3K regulada por proteína G, incluyendo moléculas pequeñas, moléculas grandes, proteínas de p101 mutantes que compiten con p101 nativo, y anticuerpos, así como las secuencias de nucleótidos que se pueden utilizar para inhibir la expresión del gen de p101 (por ejemplo, moléculas antisentido y ribozimicas, y construcciones de sustitución de secuencias reguladoras o genes) o aumentar la expresión del gen de p101 (*por ejemplo*, construcciones de expresión que colocan el gen de p101 bajo del control de un sistema de promotor fuerte), y células y animales transgénicos que expresan un transgen de p101 o “*knock-out*” que no expresan el p101.

Además, la presente invención engloba métodos y composiciones para la evaluación del diagnóstico, la clasificación y el pronóstico de las enfermedades del sistema inmune, incluyendo la inflamación, y para la identificación de los sujetos que tienen predisposición a tales estados. Por ejemplo, las moléculas del ácido nucleico de p101 de la invención se pueden utilizar como sondas de hibridación diagnósticas o como cebadores para el diagnóstico por análisis de PCR para la identificación de mutaciones del gen de p101, variaciones alélicas y defectos reguladores en el gen de p101. La presente invención además proporciona los kits de diagnóstico para la práctica de estos métodos.

Además, la presente invención también se refiere a los métodos para el uso de productos genéticos de p101 para la identificación de compuestos que modulan, es decir, que actúan como agonistas o antagonistas, la expresión del gen de PI3K que es regulado por proteína G y/o la actividad del producto del gen de p101. Estos compuestos se pueden utilizar como agentes para controlar enfermedades del sistema inmune y, en particular, como agentes terapéuticos para el tratamiento de enfermedades inflamatorias tales como artritis, choque séptico, síndrome de dificultad respiratoria aguda en adultos (ARDS), pulmonía, asma, alergias, lesión por reperfusión, aterosclerosis, enfermedad de Alzheimer, y cáncer.

También se describen en esta memoria métodos y composiciones para el tratamiento de enfermedades inflamatorias tales como artritis, choque séptico, síndrome de dificultad respiratoria aguda en adultos (ARDS), pulmonía, asma, alergias, lesión por reperfusión, aterosclerosis, enfermedad de Alzheimer, y cáncer. Estos métodos y composiciones son capaces de modular el nivel de expresión del gen de p101 y/o el nivel de actividad del producto del gen de p101.

Esta invención se basa, en parte, en el sorprendente descubrimiento de una actividad de PI3K estimulada por proteína G en neutrófilos aislados; la purificación y la caracterización de esta actividad como una proteína heterodimérica de PI3K compuesta de una subunidad p101 y de una subunidad p120; la identificación y la clonación del ADNc de p101 a partir de librerías preparadas de mRNA de neutrófilos de cerdo, de mRNA de monocitos humanos, y de mRNA de leucocitos humanos; la identificación y la clonación del ADNc de p120 a partir de librerías preparadas de mRNA de neutrófilos de cerdo y de mRNA de leucocitos humano; caracterización de las nuevas secuencias; y estudios de las

proteínas aisladas p101 y p120 expresadas de forma recombinante aisladamente en células de insecto, células U937, y células Cos-7.

3.1 Definiciones

Tal como se usan en esta memoria, los términos siguientes, tanto si están utilizados en singular o en plural, tendrán los significados indicados:

PI3K regulada por proteína G: se refiere a una enzima PI3K cuya actividad se encuentra estimulada por proteínas G triméricas activadas tales como subunidades $G\beta\gamma$ y/o subunidades $G\alpha$ -GTP.

p101: significa la subunidad reguladora de la PI3K regulada por proteína G, también conocida como la subunidad adaptadora. p101 incluye las moléculas que son homólogas a p101 o que se unen a p120 y estimulan la actividad catalítica de PI3K en respuesta a la activación de las proteínas G triméricas. Las proteínas de fusión de p101 con un fragmento Glu N-terminal (específicamente MEEEFMPMPMEF) se refieren en esta memoria como proteínas de fusión de p101 (EE), mientras que las proteínas de fusión p101 que tienen un fragmento de epítipo myc N-terminal se refieren en esta memoria como proteínas de fusión de p101 myc.

Nucleótidos o secuencias de codificación de p101: significan secuencias de nucleótidos que codifican la proteína de la subunidad reguladora de p101, fragmentos polipeptídicos o peptídicos de la proteína p101, o proteínas de fusión p101. Las secuencias de nucleótidos de p101 engloban ADN, incluyendo el ADN genómico (*por ejemplo*, el gen de p101) o ADNc, o ARN.

p120: significa la subunidad catalítica de PI3K regulada por proteína G. Los fragmentos polipeptídicos o peptídicos de la proteína de p120 se refieren como los polipéptidos de p120 o péptidos de p120. Las fusiones de p120, o de los fragmentos polipeptídicos o peptídicos de p120 a una proteína no relacionada se refieren en esta memoria como proteínas de fusión de p120. Las proteínas de fusión de p120 (EE) tienen el fragmento Glu N-terminal MEEEFMPMEFSS. Los equivalentes funcionales de p120 se refieren a una proteína de la subunidad catalítica de PI3K que se une a la subunidad reguladora p101 con alta afinidad *in vivo* o *in vitro*. Otros equivalentes funcionales de p120 son subunidades catalíticas homólogas de una PI3K regulada por proteína G, tal como p117.

Nucleótidos o secuencias codificantes de p120: significan las secuencias de nucleótidos que codifican la proteína de la subunidad catalítica de p120, fragmentos polipeptídicos o peptídicos de la proteína de p120, o proteínas de fusión de p120. Las secuencias de nucleótidos de p120 engloban ADN, incluyendo ADN genómico (*por ejemplo*, el gen de p120) o ADNc, o ARN.

4. Descripción de las figuras

Figuras 1A, 1B, 1C, y 1D. Secuencia de nucleótidos de la subunidad adaptadora de p101 de cerdo. (SEC ID NO:1)

Figura 2. Secuencia de aminoácidos de p101 de cerdo deducida. (SEC ID NO:2) Los péptidos trípticos identificados por secuenciación proteica se subrayan con una línea gruesa.

Figuras 3A, 3B y 3C. Secuencia de nucleótidos de la subunidad catalítica de p120 de cerdo (SEC ID NO:3).

Figura 4. Secuencia de la subunidad catalítica de p120 deducida de cerdo. (SEC ID NO:4) Los péptidos trípticos identificados por secuenciación proteica se encuentran subrayados con una línea gruesa. La región de divergencia de la secuencia de aminoácidos publicada de p110 γ se subraya con una línea discontinua.

Figura 5. PI3K sensible a $G\beta\gamma$ en citosol de neutrófilo es diferente de PI3K de p85/p110. Alícuotas de cada una de las fracciones obtenidas de un perfil cromatográfico con Q-sefarsa fueron transferidas mediante Western ("Western blotting") y sondadas con los anticuerpos monoclonales $\alpha p85\alpha$ o $\alpha p85\beta$ y visualizadas con un anticuerpo secundario unido a HRP panel superior en la Figura 5), o incubadas con 100nM [^3H] -17-hidroxi-wortmanina, separadas por SDS-PAGE (acrilamida al 6% y medidas por fluorescencia (panel inferior en la Figura 5). La actividad de PI3K sensible a $G\beta\gamma$ se eluyó en las fracciones de 20 a 24.

Figura 6. El Pico B después de la purificación final con columna Mini Q contiene actividad de PI3K sensible a $G\beta\gamma$. El producto de purificación de la actividad final del Pico B se analizó después del aislamiento con columna Mini Q en presencia y ausencia de subunidades $G\beta\gamma$ ($\beta\gamma$), de un péptido fosforilado de tirosina (péptido YP), wortmanina, y/o subunidades $G\alpha_i$ -GDP (α_i -GDP).

Figura 7. Propiedades farmacológicas y reguladoras de p120 libre y de p101/p120 de PI3K expresada recombinante en células Sf9. Los ensayos contenían 7 nM de p101/p120 o 36 nM de p120 solo (concentraciones finales), que se incubaron con diferentes reactivos; NaF 10 mM y $AlCl_3$ (A/F) 30 μ M; 1 μ M de subunidades $G\beta\gamma$ ($\beta\gamma$); $G\alpha$ -GDP 2 μ M; wortmanina (W) 100 nM o 50 μ M de péptido fosforilado de tirosina (PY) durante un total de 15 minutos (a 0°C) antes de iniciar los ensayos añadiendo [γ ^{32}P]-ATP; se cuantificó el [^{32}P] incorporado a [^{32}P] -PtdIns(3,4,5)P $_3$. Los datos mostrados son la media (n=2).

Figura 8. p101 se puede asociar a actividad de PI3K estimulada por G $\beta\gamma$ en células U937. Las células U937 eran transfectadas transitoriamente con vectores de expresión de mamíferos que codificaban (EE)120 o (EE)101, co-transfectadas con vectores de expresión de mamíferos que codificaban myc101, myc120, o una proteína de fusión de myc de control irrelevante, según se indica. Se utilizó un total de 40 μ g de ADN de vector. Después de la co-transfección, las células se lisaron, se lavaron previamente y se inmunoprecipitaron con sefarsa de proteína G entrecruzada covalentemente al anticuerpo monoclonal α -(EE) tal como se describe en más detalle lo en los Ejemplos de esta memoria. Los inmunoprecipitados resultantes se lavaron y se evaluó la actividad de PI3K activada por G $\beta\gamma$ (1 μ M, concentración final). La actividad de PI3K detectada en los inmunoprecipitados de las células transfectadas con la proteína con fragmento (EE) irrelevante, con o sin G $\beta\gamma$ se restó de los datos mostrados (éstos son medias 1896 dpm y 2862 dpm en ausencia y presencia de G $\beta\gamma$ respectivamente). Transfecciones paralelas marcadas con [^{35} S]-metionina mostraron que la cantidad de [^{35} S]-p101 y [^{35} S]-p120 en los inmunoprecipitados se redujo un 40% cuando estaban co-transfectadas (datos no mostrados).

Figura 9. Regulación p101/p120 por G $\beta\gamma$ s *in vivo*. Las células Cos-7 se transfectaron con vectores de expresión de mamíferos que codificaban la subunidad G $\beta\gamma$ (" $\beta_1\gamma_2$ "), p101 ("101"), y/o p120 ("120"), tal como se ha indicado. Después de 48 horas de expresión transitoria, se realizó una transferencia Western ("Western Blot") a las células de cada transfección para determinar la "Expresión Relativa", o se evaluaron para determinar la acumulación de PtdIns(3,4,5)P $_3$ (barras sombreadas). Para realizar la transferencia Western, las muestras fueron sondadas con un anticuerpo monoclonal α -(myc) para cuantificar la expresión de las diferentes proteínas con fragmentos (myc). Los resultados se dan en relación a la expresión obtenida en presencia de los 4 vectores claves; los niveles absolutos de (myc)-p120 y (myc)-p101 fueron muy similares y aproximadamente 10 veces superiores a los de (myc)- γ_2 . Lotes paralelos de células fueron marcados con [^{32} P]-Pi; después de 90 minutos se extrajeron los lípidos, se desacetilaron y los grupos solubles en agua se separaron mediante HPLC de intercambio aniónico y se cuantificaron por conteo de centelleo líquido. Los datos mostrados son medias (n=2) +/- rangos. Los datos para [^{32} P]-PtdIns(3,4,5)P $_3$ son superiores al control ADN irrelevante (972 $\pm$ 41 dpm).

Figuras 10A y 10B. Secuencia de nucleótidos de la subunidad adaptadora de p101 humana. (SEC ID NO:11).

Figura 11. Secuencia de aminoácidos de p101 humana deducida (SEC ID NO: 12).

Figuras 12A y 12B. Secuencia de nucleótidos de la subunidad catalítica de p120 humana (SEC ID NO:13).

Figura 13. Secuencia de la subunidad catalítica de p120 humana deducida. (SEC ID NO:14).

5. Descripción detallada de la invención

La presente invención se refiere a la identificación, purificación, y clonación de una subunidad de p101 de una forma específica de PI3K que es activada por subunidades $\beta\gamma$ de proteínas G triméricas. La p101, descrita por primera vez en esta memoria, es una nueva subunidad de la PI3K regulada por proteína G. También se describe en esta memoria la identificación, clonación, y secuencia correcta de p120, la subunidad catalítica de la PI3K regulada por proteína G trimérica.

Las PI3K son enzimas que fosforilan fosfatidilinositoles en la posición 3d para generar la molécula de señalización intracelular PtdIns(3,4,5)P $_3$. Aunque se ha demostrado que las PI3K están inducidos en diferentes tipos celulares tras el estímulo de los receptores de tirosina quinasa, las PI3K que son activadas por la proteína G trimérica asociada a receptores solamente se han detectado en un número limitado de células derivadas del linaje hematopoyético tales como plaquetas, monocitos, y leucocitos. De manera interesante, la acumulación de PtdIns(3,4,5)P $_3$ en estas células se asocia a muchas de las respuestas inmunes implicadas en la inflamación aguda y crónica. Por ejemplo, los neutrófilos se activan por, entre otros, la quimiocina fMLP que es liberada en respuesta a la infección por microorganismos. Este agonista se une a un receptor asociado a una proteína G sensible a toxina pertúsica y activa la explosión respiratoria del neutrófilo, resultando en la producción de superóxido y citotoxicidad. La respuesta de explosión respiratoria se correlaciona con un rápido y transitorio aumento intracelular de actividad de PI3K.

Los estudios en neutrófilos demuestran que la wortmanina, que inhibe selectivamente el componente catalítico de PI3K, causa una parada en la producción de superóxido. Puesto que la gran mayoría de los agonistas que activan la explosión respiratoria se unen a receptores asociados a proteína G, la actividad de PI3K regulada por proteína G representa un candidato para una vía efectora dominante cuya inhibición reduzca los efectos celulares destructivos de la inflamación. Sin embargo, la wortmanina no es un inhibidor clínicamente apropiado para esta vía puesto que la wortmanina inhibe todas las actividades catalíticas de PI3K, incluyendo las implicadas en otras vías celulares tales como factores de crecimiento.

Se describe en esta memoria el descubrimiento que la PI3K regulada por proteína G está compuesta de dos subunidades: la subunidad catalítica (p110 γ , p117 o p120); y la subunidad reguladora p101. La capacidad de PI3K de responder a los receptores de la superficie celular que estimulan la liberación de las subunidades G $\beta\gamma$ de las proteínas G triméricas activadas es en gran parte dependiente de la presencia de la subunidad reguladora p101. Aunque las subunidades catalíticas presentan una pequeña estimulación (aproximadamente 1.7 veces) de la actividad de PI3K *in vitro* en presencia de las subunidades G $\beta\gamma$, la adición de proteínas p101 incrementa la estimulación de G $\beta\gamma$ de la actividad de PI3K en 100 veces. La neutralización de p101, la eliminación de p101, o la interferencia de la unión de p101 a sus

elementos de unión produce que las células sean incapaces de generar un gran incremento en los niveles de señalización intracelular $\text{PtdIns}(3,4,5)\text{P}_3$ en respuesta a las subunidades $G\beta\gamma$ activadas. De esta forma, la distribución limitada de la subunidad p101 en células procedentes de la médula ósea es el factor crítico en la especificidad de tipo celular de esta respuesta. Además, p101 no contiene ninguna homología significativa con ninguna otra secuencia identificada.

5 Este descubrimiento hace p101, y el complejo p101/p120, la diana de elección para inhibir, activar, o modular PI3K activada por proteína G con mínimos efectos no específicos.

La solicitud también describe el uso de nucleótidos de p101 y p120, proteínas y péptidos de p101 y p120, así como anticuerpos para p101 y p120 (que pueden, por ejemplo, actuar como agonistas o antagonistas de p101 o p120),

10 antagonistas que inhiben la actividad o expresión de PI3K activada por proteína G, o agonistas que activan la actividad de PI3K activada por proteína G o que incrementan su expresión en el tratamiento de los desórdenes de activación de células del linaje hematopoyético, incluyendo, pero no limitado a las respuestas inmunes asociadas a la inflamación aguda y crónica, en animales como los seres humanos. Por ejemplo, los antagonistas de las PI3K activadas por $G\beta\gamma$ son útiles en el tratamiento de la artritis, incluyendo artritis reumatoide, choque séptico, síndrome de dificultad

15 respiratoria aguda en adultos (ARDS), pulmonía, asma y de otras afecciones del pulmón, alergias, lesiones por reperfusión, aterosclerosis y otras enfermedades cardiovasculares, enfermedad de Alzheimer, y cáncer, para nombrar solo algunas enfermedades inflamatorias. Además, los nucleótidos de p101 y las subunidades reguladoras p101, así como los nucleótidos de p120 y las subunidades catalíticas p120, son útiles para la identificación de compuestos eficaces en el tratamiento de los desórdenes de activación de células del linaje hematopoyético que implican PI3K activadas por

20 proteína G.

Además, la invención engloba el uso de nucleótidos de p101, proteínas y péptidos de p101, así como los anticuerpos para p101 en el diagnóstico de los desórdenes de activación de células del linaje hematopoyético. El diagnóstico de una anomalía en la subunidad reguladora p101 en un paciente o una anomalía en la vía de trans-

25 ducción de señales de PI3K activada por proteína G, también permiten idear un tratamiento o régimen terapéutico apropiados.

En particular, la invención descrita en las secciones posteriores engloba la subunidad reguladora p101, polipéptidos o péptidos correspondientes a los dominios funcionales de la subunidad reguladora p101 (*por ejemplo* el dominio de asociación a la subunidad catalítica o el dominio de interacción con las proteínas G activadas), subunidades regula-

30 doras p101 mutadas, truncadas o suprimidas (*por ejemplo* una subunidad reguladora p101 con uno o más dominios funcionales o partes de los mismos suprimidos), proteínas de fusión de subunidades reguladoras p101 (*por ejemplo*, una subunidad reguladora p101 o un dominio funcional de la subunidad reguladora p101, fusionado a una proteína o péptido no relacionado tal como un fragmento de epítipo, es decir, el epítipo myc), secuencias de nucleótidos que codifican tales productos, y los sistemas de expresión en células hospedadoras que pueden producir estos productos de la subunidad reguladora p101.

35

Además, la presente solicitud describe las proteínas de subunidades catalíticas p120, polipéptidos, dominios funcionales de la subunidad p120 (*por ejemplo*, el dominio catalítico), proteínas de subunidades p120 mutadas, truncadas o suprimidas, proteínas de fusión de p120, secuencias de nucleótidos que codifican estos productos, y los sistemas de expresión en células hospedadoras que pueden producir estos productos de la subunidad catalítica p120.

40

La invención también engloba anticuerpos y anticuerpos anti-idiotípicos (incluyendo fragmentos Fab), antagonistas y agonistas de PI3K activada por proteína G, así como compuestos o construcciones nucleotídicas que inhiben la expresión del gen de p101 (inhibidores del factor de transcripción, moléculas antisentido y ribozimicas, o construcciones de sustitución de la secuencia reguladora del gen), o promueven la expresión de la subunidad reguladora p101 (*por ejemplo*, construcciones de expresión en los que secuencias codificadoras de p101 se encuentran asociadas de forma operativa a elementos de control de la expresión tales como promotores, promotores/potenciadores, etc.). La invención también se refiere a células hospedadoras y a animales modificados genéticamente que expresan la subunidad reguladora p101 humana (o mutantes de la misma) o que inhiben o suprimen ("*knock-out*") la expresión de la subunidad reguladora p101 endógena del animal.

45

50

Además, la invención engloba particularmente a antagonistas que previenen la asociación de las subunidades reguladoras p101 con sus elementos de unión, incluyendo p120 y otras proteínas de subunidades catalíticas de PI3K como p117 y p110 γ , así como proteínas de proteína G trimérica activadas, incluyendo las subunidades $G\beta\gamma$.

55

Las proteínas o péptidos de la subunidad reguladora p101, las proteínas de fusión de la subunidad reguladora p101, las secuencias de nucleótidos de p101, los anticuerpos, los antagonistas y agonistas pueden ser útiles en la detección de subunidades reguladoras p101 mutantes o subunidades reguladoras p101 que se expresan de forma inadecuada para el diagnóstico de enfermedades inmunes. Los péptidos o proteínas de las subunidades p101 y p120, y las proteínas de fusión de las subunidades p101 y p120, las secuencias de nucleótidos de p101 y p120, los sistemas de expresión en células hospedadoras, los anticuerpos, los antagonistas, los agonistas y las células y animales modificados genéticamente se pueden utilizar para hacer un cribaje de la efectividad de los fármacos en el tratamiento de estas enfermedades inmunes. El uso de células hospedadoras y/o de animales modificados genéticamente puede ofrecer una ventaja en que tales sistemas permiten no sólo la identificación de compuestos que se unen a la subunidad reguladora p101, sino que también permiten identificar compuestos que afectan la señal transducida por la subunidad reguladora p101 activada, específicamente, producción de la molécula de señalización intracelular $\text{PtdIns}(3,4,5)\text{P}_3$.

60

65

Finalmente, los productos proteicos de la subunidad reguladora p101 y los productos proteicos de fusión, los anticuerpos y anticuerpos anti-idiotípicos (incluyendo fragmentos Fab), los antagonistas o agonistas (incluyendo compuestos que modulan la transducción de señal que pueden actuar en dianas en las vías de transducción de señal posteriores de la subunidad reguladora p101) se pueden utilizar para la terapia de estas enfermedades. Por ejemplo, construcciones de nucleótidos que codifican subunidades reguladoras p101 funcionales, subunidades reguladoras p101 mutantes, así como moléculas antisentido y ribozimicas se pueden utilizar en aproximaciones de terapia génica para la modulación de la expresión y/o actividad de la subunidad reguladora p101 en el tratamiento de desórdenes de activación de células del linaje hematopoyético. De esta forma, la invención también engloba formulaciones y métodos farmacéuticos para tratar desórdenes de activación de células del linaje hematopoyético.

La invención se basa, en parte, en el sorprendente descubrimiento de nuevas enzimas PI3K en neutrófilos de cerdo. Como otras proteínas PI3K, estas enzimas fosforilan en la posición 3 a los sustratos PtdIns, PtdIns4P y PtdIns(4,5)P₂, y se inhibían completamente de wortmanina 100 nM. Sin embargo, a diferencia de los complejos proteicos p85/p101 de PI3K previamente descritos, la actividad de PI3K se encontraba estimulada más de 100 veces por la incubación con subunidades de Gβγ. La adición de subunidades Gα-GDP puede inhibir esta activación de Gβγ. Además, la actividad de PI3K no se encontraba estimulada por los péptidos de tirosina fosforilados. Cuando se purificaron, se identificaron dos complejos de proteína heterodiméricos diferentes: un complejo p117/p101 y un complejo p120/p101. La secuenciación de péptidos revela que la proteína p101 era idéntica en cada uno de los complejos. Las proteínas p120 y p117 eran homólogas a excepción del extremo amino (véanse los Ejemplos posteriores). Los ADNc de p101 y p120 de cerdo fueron clonados a continuación usando sondas degeneradas basadas en la secuencia peptídica para cribar una librería de expresión de los ADNc sintetizados a partir de mRNA de neutrófilos de cerdo. Se clonaron p101 y p120 humanos usando sondas provenientes de secuencias de ADNc de cerdo para cribar librerías de ADNc de monocitos y leucocitos humanos. Mientras que el análisis de la secuencia de las subunidades de cerdo y humanas reveló que las proteínas p120 son homólogas a las subunidades catalíticas previamente clonadas de PI3K (aunque divergen significativamente en el extremo carboxilo), las proteínas de p101 no están relacionadas con ninguna secuencia en las bases de datos.

La comparación de las secuencias aminoacídicas de los homólogos de cerdo y humano de p101 y p120 revela un alto grado de conservación entre las dos especies. Este alto grado de conservación entre dos especies de mamíferos sugiere que existe probablemente un grado similar de conservación entre especies de mamíferos en general. De la misma forma que se utilizaron las secuencias de cerdo para clonar sus homólogos humanos, un experto en la materia puede clonar los homólogos p101 y p120 de otras especies de mamíferos usando las secuencias humanas y/o de cerdo descritas y los métodos descritos en esta memoria.

Los experimentos descritos en esta memoria que expresaron proteínas de fusión de p101 y/o p120 en células de insecto demostraron que p101 se une firmemente a p120 en una estequiometría molar 1:1. p120 libre purificada presentaba actividad PI3K que era insensible a la presencia de subunidades Gα y de péptidos fosforilados en tirosinas, y únicamente estimulada de forma suave por subunidades Gβγ. Sin embargo, cuando está unida a p101, la actividad PI3K de p120 se encontraba estimulada 100 veces por la presencia de subunidades Gβγ. Cuando p101 se expresó sola en células de insecto, p101 no presentaba actividad PI3K.

Un resultado interesante ocurrió cuando p101 se expresó como una proteína de fusión con un fragmento de identificación en células U937 humanas. Cuando esta p101 de cerdo recombinante se inmunoprecipitó mediante el fragmento de identificación de la proteína de fusión, estos inmunoprecipitados contenían actividad PI3K regulada por proteína G. La co-expresión de p120 con p101 disminuyó ligeramente la cantidad de actividad PI3K que podía ser inmunoprecipitada. Estos resultados indicaron que las células U937 humanas contienen una subunidad catalítica de PI3K que se podía unir y ser activada por la subunidad reguladora p101 de cerdo. Estos resultados demuestran que junto con sus secuencias de aminoácidos altamente conservadas, que los homólogos de diversas especies de mamíferos se asemejan entre ellos funcional y estructuralmente.

Además, las transfecciones transitorias se realizaron en células Cos-7, que normalmente no estimulan la actividad PI3K en respuesta a proteínas G activadas, con construcciones que codifican p101, p120 y/o subunidades Gβγ. La transfección de una construcción que expresaba p120 solamente producía aumentos significativos en los niveles celulares de PtdIns(3,4,5)P₃ de una forma dependiente de Gβγ cuando se encontraba co-expresado en presencia de p101.

Diferentes aspectos de la invención se describen en mayor detalle en las secciones posteriores.

5.1 Genes de p101 y p120

La secuencia de ADNc (SEC ID NO:1) y la secuencia deducida de aminoácidos (SEC ID NO:2) de la subunidad reguladora p101 de cerdo se muestran en las Figs. 1 y 2, respectivamente. Una variación alélica relativamente común existe en el residuo aminoacídico 483 en la pauta de lectura abierta; una serina se puede sustituir por una glicina en esta posición.

La secuencia del ADNc (SEC ID NO:11) y la secuencia deducida de aminoácidos (SEC ID NO:12) de la subunidad reguladora p101 humana se muestran en las Figs. 10 y 11, respectivamente.

Se cree que la secuencia de nucleótidos que codifica los primeros 733 aminoácidos de la subunidad reguladora p101 de cerdo, que corresponde a los primeros 736 aminoácidos de la subunidad humana, es suficiente para unirse a la subunidad catalítica. Sin embargo, sin los 145 aminoácidos del extremo carboxilo, esta p101 truncada unida a p120 no estimula la actividad PI3K en respuesta a subunidades $G\beta\gamma$. Por lo tanto, se cree que el dominio de asociación a la subunidad catalítica está contenido dentro de los primeros 732 aminoácidos (735 para el ser humano), y los aminoácidos del extremo carboxilo 733 a 877 (736 a 880 para el ser humano) pueden estar implicados en la respuesta a las subunidades $G\beta\gamma$.

Otros dominios de p101 de cerdo se describen por los residuos aminoacídicos de aproximadamente 1 a 160, 80-120, 161 a 263, 264 a 414, 415 a 565, 566 a 706, 707-832, y/o 833 a 877. Los dominios correspondientes a p101 humana se describen por los residuos aminoacídicos de aproximadamente 1 a 160, 80-120, 161 a 263, 264 a 416, 417 a 567, 568 a 709, 710-835, y/o 836 a 880. Las secuencias de nucleótidos que codifican residuos aminoacídicos de aproximadamente 161 a 263 definen un dominio de homología de pleckstrin ("PH") (PROSITE:PS50003). Los dominios de PH pueden estar implicados en la unión de proteínas a subunidades $G\beta\gamma$ (Touhara *et al.*, 1994, J. Biol. Chem. 269:10217) y a fosfolípidos de membrana (véase Shaw, 1996, Bioessays 18:35-46). De este modo el dominio de PH de p101 puede estar implicado en estos dos acontecimientos.

Las secuencias de nucleótidos que codifican los aminoácidos 1 a 160 de p101 pueden ser responsables de la unión a la subunidad catalítica p120. Cuando se analiza la estructura secundaria de esta región de la proteína p101, se predice que contiene una estructura α -hélice/lámina- β alternante. Dentro de esta estructura se encuentra homología a un "dominio de WW" (Staub *et al.*, 1996, Structure 4:495-499 y TIBS 21:161-163). Este dominio de WW se puede unir a un dominio rico en prolina que se encuentra dentro del extremo N-terminal de la proteína p120 (residuos 310 a 315). De este modo, el dominio de WW de p101 puede estar implicado en mediar la interacción entre la subunidad reguladora y la subunidad catalítica.

La secuencia de ADNc (SEC ID NO:3) y la secuencia deducida de aminoácidos (SEC ID NO:4) de p120 de cerdo se muestran en las Figs. 3 y 4, respectivamente. La secuencia de ADNc (SEC ID NO:13) y la secuencia deducida de aminoácidos (SEC ID NO:14) de p120 humano se muestran en las Figs. 12 y 13, respectivamente. Un sitio de corte críptico de trombina está presente después de los primeros 40 residuos aminoacídicos aproximadamente. La proteína p120 truncada que carece de estos aproximadamente 40 residuos amino terminales todavía es capaz de unirse a p101, pero la actividad PI3K se encuentra reducida aproximadamente de un 20 a un 30%. Es decir, una proteína que tiene la secuencia de aminoácidos de la sección de SEC ID NO:4 (secuencia de aminoácidos deducida de p120 de cerdo) que se extiende desde aproximadamente el residuo aminoacídico 41 al residuo 1102 conserva la funcionalidad de p120 de longitud completa. Basándose en el alto grado de homología de secuencia entre p120 de cerdo y humana, una proteína que tengan la secuencia de aminoácidos de la sección de SEC ID NO:14 (secuencia de aminoácidos deducida de p120 humano) que se extiende desde aproximadamente el residuo aminoacídico 41 al residuo 1102 también conservará la funcionalidad de p120 de longitud completa. Aunque la proteína p120 era altamente homóloga a la proteína p110 γ previamente clonada descrita por Stoyanov *et al.*, (1995), el extremo C-terminal de p120 diverge de la proteína p110 γ descrita en el residuo aminoacídico 1075; de este modo, los últimos 28 residuos aminoacídicos de p120 no se han publicado en informes de ninguna proteína homóloga. Tal como se ha observado anteriormente, los nucleótidos que codifican una región rica en prolina, incluyendo los residuos 310 a 315 de p120, pueden estar implicados en la interacción entre p120 y p101. Además, los nucleótidos que codifican los residuos aminoacídicos de p120 desde aproximadamente de 173 a 302 definen un dominio de PH débil que puede ser un candidato de unión a membrana y/o interacción con la subunidad $G\beta\gamma$ del complejo p101/p120.

Los datos presentados en los siguientes ejemplos del trabajo, muestran que el ADNc de p120 codifica la subunidad catalítica de la PI3K activada por $G\beta\gamma$. El ADNc de p101 codifica una nueva proteína de la subunidad reguladora que se une a la subunidad p120. Este heterodímero fosforila en la posición 3 a PtdIns, PtdIns4P y PtdIns(4,5)P₂ en respuesta a la activación de los receptores asociados a la proteína G trimérica.

Las secuencias de nucleótidos de p101 de la invención incluyen: (a) la secuencia de ADN mostrada en la Fig. 1 o la Fig. 10 o contenida en el clon de ADNc pCMV3mycp 101 tal como se depositó en la American Type Culture Collection (ATCC) bajo número de acceso 97636; (b) la secuencia de nucleótidos que codifica las secuencias de aminoácidos mostrada en la Fig. 2 o la Fig. 11, o la secuencia de aminoácidos de la subunidad reguladora p101 codificada por el clon de ADNc pCMV3mycp 101 tal como se depositó en la ATCC; (c) cualquier secuencia de nucleótidos que hibrida con el complemento de la secuencia de ADN mostrada en la Fig. 1 o Fig. 10 o contenida en el clon de ADNc pCMV3mycp 101, tal como se depositó en la ATCC bajo condiciones altamente restrictivas, *por ejemplo*, hibridación a ADN unido a filtro en NaHPO₄ 0,5 M, dodecil sulfato sódico (SDS) al 7%, EDTA 1 mM a 65°C y lavado en 0,1xSSC/0,1% SDS a 68°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York, en p.2.10.3.) y que codifica un producto genético funcionalmente equivalente; y (d) cualquier secuencia de nucleótidos que se hibride al complemento de las secuencias de ADN que codifican la secuencia de aminoácidos mostrada en las Fig. 1 o Fig. 10 o contenido en el clon de ADNc pCMV3mycp 101, según se depositó en la ATCC, bajo condiciones menos restrictivas, tales como condiciones moderadamente restrictivas, *por ejemplo*, lavado en 0,2xSSC/0,1% SDS a 42°C (Ausubel *et al.*, 1989, anterior), y que todavía codifica un producto génico de p101 funcionalmente equivalente. Los equivalentes funcionales de la subunidad reguladora p101 incluyen la subunidad reguladora p101 que existe de forma natural presente en otras especies, y subunidades reguladoras mutantes de p101 que se encuentran de forma natural o diseñadas por ingeniería genética que conservan como mínimo algunas de las actividades funcionales de p101 (es decir, unión a la subunidad

catalíticas p120 o p117, estimulación de la actividad catalítica en respuesta a subunidades G β γ , y/o interacción con subunidades G β γ). La invención también incluye variantes degeneradas de las secuencias (a) a (d).

La invención también incluye moléculas de ácidos nucleicos, preferentemente moléculas de ADN, que se hibridan a, y por lo tanto son complementos de las secuencias de nucleótidos (a) a (d), en el párrafo anterior. Tales condiciones de hibridación pueden ser altamente restrictivas menos restrictivas, tal como se ha descrito anteriormente. En los casos en que las moléculas de ácidos nucleicos son desoxiligonucleótidos ("oligos"), condiciones altamente restrictivas pueden ser, *por ejemplo*, lavado en 6xSSC/ pirofosfato de sodio 0,05% a 37°C (para oligos de 14 bases), 48°C (para oligos de 17 bases), 55°C (para oligos de 20 bases), y 60°C (para oligos de 23 bases). Estas moléculas de ácidos nucleicos pueden codificar o actuar como moléculas antisentido de p101, útiles, *por ejemplo*, en la regulación génica de p101 (para y/o como cebadores antisentido en reacciones de amplificación de secuencias de ácidos nucleicos del gen de p101). Con respecto a la regulación del gen de p101, estas técnicas se pueden utilizar para regular, *por ejemplo*, respuestas inmunes inflamatorias. Además, estas secuencias se pueden utilizar como parte de las secuencias de triple hélice y/o de ribozima, también útiles para la regulación génica de p101. Adicionalmente, estas moléculas se pueden utilizar como componentes de métodos de diagnóstico los cuales, *por ejemplo*, la presencia de un determinado alelo de p101 responsable de causar una respuesta inflamatoria, tal como artritis, pueda ser detectado.

Además de las secuencias de nucleótidos de p101 descritas anteriormente las secuencias génicas o el ADNc de p101 de longitud completa presentes en la misma especie y/o homólogos al gen de p101 presente en otras especies se pueden identificar y aislar fácilmente sin realizar una experimentación excesiva, mediante las técnicas de biología molecular bien conocidas en la técnica. Las pruebas experimentales descritas en esta memoria indican que las proteínas de p101 se conservan en diversas especies de mamíferos. *Por ejemplo*, p101 humana fue clonada a partir de librerías de monocitos y leucocitos usando sondas provenientes de la secuencia de ADNc de p101 de cerdo, y se observó que las secuencias deducidas de aminoácidos humanas y de cerdo eran altamente conservadas. La semejanza estructural y funcional de los dos homólogos se demuestra por los experimentos que muestran que cuando se expresó p101 de cerdo en las células U937, la p101 de cerdo se unía a un homólogo humano de la subunidad catalítica. Además, miembros de la familia de proteínas de PI3K reguladas por tirosina quinasas, *por ejemplo* p110/p85 PI3K, también están conservadas entre diversas especies de mamíferos. Por lo tanto, los homólogos de p101 y de p120 se pueden aislar a partir de varias células de mamíferos que se sabe o se sospecha que contienen PI3K regulada por proteína G trimérica, particularmente células de origen hematopoyético, y más particularmente plaquetas, monocitos, leucocitos, osteoclastos y neutrófilos.

La identificación de homólogos de p101 en especies relacionadas puede ser útil para desarrollar sistemas de modelos animales más relacionados con los seres humanos con objeto del descubrimiento de fármacos. *Por ejemplo*, las librerías de expresión de los ADNc sintetizados a partir del mRNA de neutrófilos proveniente del organismo de interés se pueden cribar usando la subunidad catalítica marcada proveniente de esa especie, *por ejemplo*, una proteína de fusión de la subunidad catalítica p120, p117 o p110 γ . Alternativamente, estas librerías de ADNc, o librerías de ADN genómico provenientes del organismo de interés se pueden cribar por hibridación usando los nucleótidos descritos en esta memoria como sondas de hibridación o de amplificación. Además, también se pueden identificar genes en otros locus genéticos dentro del genoma que codifica las proteínas que tienen una extensa homología con unos o más dominios del producto génico de p101 mediante técnicas similares. En el caso de librerías de ADNc, estas técnicas de cribaje pueden identificar clones provenientes de transcritos de corte y empalme ("*splicing*") alternativo en la misma o diferentes especies.

El cribaje puede ser por hibridación por filtro, usando filtros duplicados. La sonda marcada puede contener como mínimo 15-30 pares de bases de la secuencia de nucleótidos de p101 humana o de cerdo, tal como se muestra en las Figs. 1 y 10, respectivamente. Las condiciones de lavado de la hibridación usadas deben ser menos restrictivas cuando la librería de ADNc proviene de un organismo diferente del tipo de organismo del que proviene la secuencia marcada. Con respecto a la clonación de un homólogo de p101 de mamífero, usando sondas de p101 de cerdo y/o humanas, *por ejemplo*, la hibridación se puede, *por ejemplo*, realizar a 65°C hasta el día siguiente en tampón Church (SDS al 7%, NaHPO₄ 250 mM, EDTA 2 μ M, BSA al 1%). Los lavados se pueden realizar con 2xSSC, SDS al 0,1% a 65°C y a continuación en 0,1xSSC, SDS al 0,1% a 65°C.

Las condiciones menos restrictivas son bien conocidas por los expertos en la técnica, y variarán predeciblemente dependiendo de los organismos específicos de que proviene la librería y las secuencias marcadas. Para indicaciones con respecto a tales condiciones véanse, *por ejemplo*, Sambrook *et al.*, 1989, Molecular Cloning, A Laboratory Manual, Cold Springs Harbor Press, N.Y.; y Ausubel *et al.*, 1989, Current Protocols in Molecular Biology, Green Publishing Associates y Wiley Interscience, N.Y.

Alternativamente, la sonda de nucleótidos marcada de p101 se puede utilizar para cribar una librería genómica proveniente del organismo de interés, otra vez, usando condiciones apropiadamente restrictivas. La identificación y caracterización de los clones génicos humanos es útil para diseñar pruebas de diagnóstico y protocolos clínicos para tratar desórdenes de activación de células del linaje hematopoyético en pacientes humanos. *Por ejemplo*, las secuencias provenientes de regiones adyacentes a los límites de intron/exon del gen humano se pueden utilizar para diseñar cebadores para su uso en ensayos de amplificación para detectar mutaciones en exones, intrones, sitios de corte y empalme ("*splicing*") (*por ejemplo*, sitios de aceptor y/o donador del corte y empalme ("*splicing*")), etc., que se pueden utilizar en diagnóstico.

Además, se puede aislar un homólogo del gen de p101 a partir del ácido nucleico del organismo de interés realizando PCR usando dos grupos de cebadores oligonucleotídicos degenerados diseñados en base a secuencias de aminoácidos dentro del producto del gen de p101 descrito en esta memoria. La plantilla para la reacción puede ser ADNc obtenido por transcripción inversa del mRNA preparado, por ejemplo, a partir de tipos celulares o líneas celulares humanas o no humanas, tales como neutrófilos, que se sabe o se sospecha que expresan un alelo del gen de p101.

El producto de la PCR puede ser subclonado y secuenciado para asegurar de que las secuencias amplificadas representan las secuencias de un gen de p101. El fragmento de PCR se puede utilizar para aislar a continuación un clon del ADNc de longitud completa mediante diferentes métodos. Por ejemplo, el fragmento amplificado se puede marcar y utilizar para cribar una librería de ADNc, tal como una librería de ADNc de bacteriófagos. Alternativamente, el fragmento marcado se puede utilizar para aislar clones genómicos mediante el cribaje de una librería genómica.

La tecnología de PCR también se puede utilizar para aislar secuencias de ADNc de longitud completa. Por ejemplo, el ARN se puede aislar, siguiendo procedimientos estándares, a partir de una fuente celular apropiada (es decir, una que se conozca o se sospeche, que expresa el gen de p101, tal como por ejemplo, neutrófilos u otros tipos de leucocitos). Se puede realizar una reacción de transcripción inversa en el ARN usando un cebador de oligonucleótidos específico para el extremo 5' de un fragmento amplificado para cebar la síntesis de la primera cadena. A continuación, se le puede añadir una cola de guaninas al híbrido ARN/ADN resultante usando una reacción de transferasa terminal estándar, el híbrido se puede digerir con RNAasa H, y a continuación se puede cebar la síntesis de la segunda cadena con un cebador poly-c. De este modo, las secuencias de ADNc hacia arriba de la cadena ("*upstream*") del fragmento amplificado se pueden aislar fácilmente. Para una revisión de las estrategias de clonación que pueden usar, véase *por ejemplo*, Sambrock *et al.*, 1999, referencia anterior.

Las secuencias génicas de p101 se pueden utilizar adicionalmente para aislar alelos génicos de p101 mutantes. Estos alelos mutantes se pueden aislar de individuos que se sabe o se sospecha que tienen un genotipo que contribuye a los síntomas de los desordenes de la activación de células del linaje hematopoyético tales como inflamación. Los alelos mutantes y los productos alélicos mutantes se pueden usar a continuación en los sistemas diagnósticos y terapéuticos descritos a continuación. Adicionalmente, estas secuencias génicas de p101 se pueden usar para detectar defectos en la regulación del gen de p101 (*por ejemplo*, promotor o promotor/potenciador), que pueden afectar a la activación de células del linaje hematopoyético.

El ADNc de un gen de p101 mutante se puede aislar, por ejemplo, usando PCR, una técnica que es bien conocida por los expertos en la materia. En este caso, la primera cadena de ADNc se puede sintetizar por hibridación de un oligonucleótido oligo-dT a mRNA aislado de células que se sabe o se sospecha que se expresan en un individuo que supuestamente lleva el alelo de p101 mutante, y extendiendo la nueva cadena con transcriptasa inversa. La segunda cadena del ADNc se sintetiza a continuación usando un oligonucleótido que se hibrida específicamente al extremo 5' del gen normal. Usando estos dos cebadores, el producto se amplifica a continuación mediante PCR, se clona en un vector adecuado, y se somete a análisis de secuencia de ADN mediante métodos bien conocidos por los expertos en la materia. Comparando la secuencia de ADN del alelo de p101 mutante con el alelo de p101 normal, se puede determinar la mutación o mutaciones responsables de la pérdida o alteración de la función del producto del gen de p101 mutante.

Alternativamente, una librería genómica se puede construir usando ADN obtenido de un individuo que se sospecha o se conoce que tiene el alelo de p101 mutante, o una librería de ADNc se puede construir usando ARN de un tipo celular que se sabe o sospecha que expresa el alelo de p101 mutante. El gen de p101 normal o cualquier fragmento adecuado del mismo se puede marcar a continuación y utilizar como una sonda para identificar el correspondiente alelo de p101 mutante en estas librerías. Los clones que contienen las secuencias del gen de p101 mutado se pueden purificar a continuación y someter a análisis de secuencia según los métodos bien conocidos por los expertos en la materia.

Además, una librería de expresión se puede construir utilizando ADNc sintetizado, por ejemplo, de ARN aislado de un tipo celular que se sabe, o se sospecha, que expresa un alelo de p101 mutante en un individuo que se sospecha o es conocido para tener este alelo mutante. De este modo, los productos del gen generados por el tipo celular supuestamente mutante se pueden expresar y cribar usando técnicas estándares de cribaje con anticuerpos junto con anticuerpos contra el producto del gen de p101 normal, tal como se describe a continuación, en la sección 5.3. (Para las técnicas de cribaje, véase, por ejemplo, Harlow, E. and Lane, eds., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Col Spring Harbor). Además, el cribaje se puede lograr explorando con proteínas de fusión marcadas, tales como, por ejemplo, proteínas de fusión myc p120 o (EE)120. En los casos en los que una mutación en p101 tiene como resultado un producto de expresión del gen con la función alterada (*por ejemplo*, como resultado de una mutación sin sentido o de desplazamiento en la pauta de lectura), es probable que reaccione de forma cruzada un conjunto policlonal contra anticuerpos de la subunidad reguladora p101 una reacción cruzada con el producto génico de la subunidad reguladora p101 mutante. Los clones de la librería detectados mediante su reacción con estos anticuerpos marcados se pueden purificar y someter a análisis de secuencia según los métodos bien conocidos por los expertos en la materia.

La invención también engloba las secuencias de nucleótidos que codifican las subunidades reguladoras p101 mutantes, fragmentos peptídicos de la subunidad reguladora p101, subunidades reguladoras p101 truncadas, y proteínas

de fusión de la subunidad reguladora p101. Éstos incluyen, pero no se limitan a las secuencias de nucleótidos que codifican las subunidades reguladoras p101 mutantes descritas en la sección 5.2 posterior; polipéptidos o péptidos que corresponden a los dominios de unión a la subunidad G β γ o de unión catalítica de la subunidad reguladora p101 o porciones de estos dominios; subunidades reguladoras p101 truncadas en las que uno o dos de los dominios se suprime, o una subunidad reguladora p101 no funcional truncada. Los nucleótidos que codifican proteínas de fusión incluyen pero no se limitan a la subunidad reguladora p101 de longitud completa, la subunidad reguladora p101 truncada o fragmentos peptídicos de la subunidad reguladora p101 fusionada a una proteína o péptido sin relación, tal como por ejemplo, un fragmento de epítipo que ayuda en la purificación o la detección de la proteína de fusión resultante; o una enzima, proteína fluorescente, proteína luminiscente que se puede utilizar como marcador.

La invención también engloba (a) los vectores de ADN que contienen cualquiera de las secuencias que codifican las subunidades reguladoras p101 anteriores y/o sus complementos (es decir, antisentido); (b) Los vectores de expresión de ADN que contienen cualquiera de las secuencias que codifican la subunidad reguladora p101 operativamente asociadas a un elemento regulador que dirige la expresión de las secuencias codificantes; y (c) células hospedadoras modificadas genéticamente que contienen cualquiera de las secuencias que codifican las subunidades reguladoras p101 anteriores operativamente asociadas a un elemento regulador que dirige la expresión de las secuencias codificantes en la célula hospedadora. Tal como se usa en esta memoria, los elementos reguladores incluyen pero no se limitan a promotores inducibles y no-inducibles, potenciadores, operadores y otros elementos conocidos por los expertos en la materia que dirigen y regulan la expresión. Tales elementos reguladores incluyen pero no se limitan al promotor del baculovirus, al promotor del gen inmediato temprano de citomegalovirus hCMV, los promotores tempranos o tardíos del adenovirus SV40, el sistema *lac*, el sistema *trp*, el sistema *TAC*, el sistema *TRC*, regiones del operador principal del promotor del fago A, las regiones de control de la proteína de la cápside de fd, el promotor para la 3-fosfoglicerato quinasa, los promotores de la fosfatasa ácida y los promotores de los factores de unión α de la levadura.

Finalmente, la presente solicitud describe nucleótidos que codifican las proteínas de la subunidad p120 incluyendo el extremo carboxi terminal nuevamente descrito de esta subunidad catalítica, variantes de supresión de las proteínas de la subunidad p120, nucleótidos que hibridan a estos nucleótidos bajo condiciones altamente restrictivas y que codifican productos funcionalmente equivalentes, incluyendo el clon de ADNc pCMV3mycp 120 depositado en la ATCC bajo número de entrada 97637, variantes alélicas de p120 (*por ejemplo*, alelos mutantes o los alelos naturales tales como la variación alélica en el residuo aminoacídico 483), y nucleótidos de p120 equivalentes de diferentes organismos aislados como se ha descrito anteriormente para los nucleótidos de p101 de la invención. Además, también se describen en esta memoria vectores de expresión y células hospedadoras para la producción recombinante de p120.

5.2 Proteínas y polipéptidos de p101 y p120

La subunidad reguladora p101 y la subunidad catalítica p120, los fragmentos polipeptídicos y peptídicos, las formas mutadas, truncadas o suprimidas de la subunidad p101 y/o las proteínas de fusión de la subunidad reguladora p101 y la subunidad catalítica p120 y/o las proteínas de fusión de la subunidad catalítica p120, se pueden preparar para varios usos, incluyendo pero no limitado a la generación de anticuerpos, como reactivos en ensayos diagnósticos, la identificación de otros productos génicos celulares implicados en la regulación de la activación de células del linaje hematopoyético, como reactivos en los ensayos para cribar compuestos que se pueden utilizar en el tratamiento de los desórdenes de activación de células del linaje hematopoyético, y como reactivos farmacéuticos útiles en el tratamiento de los desórdenes de activación de células del linaje hematopoyético relacionados con PI3K activada por G β γ .

Las Figs. 2 y 11 muestran las secuencias de aminoácidos de la proteína de la subunidad reguladora p101 de cerdo y humana, respectivamente. Las Figs. 4 y 13 muestran la secuencia de aminoácidos de la proteína de la subunidad catalítica p120 de cerdo y humana, respectivamente. La línea discontinua en la Fig. 4 subraya la región de p120 que diverge de las secuencias publicadas de las subunidades catalíticas de PI3K p110 α , p110 β , y p110 γ .

La secuencia de la subunidad reguladora p101 empieza con una metionina en un contexto de secuencia de ADN consistente con un sitio de iniciación de traducción. La masa molecular predicha de las subunidades reguladoras p101 de cerdo y humanas es 97 kD.

Las secuencias aminoacídicas de la subunidad reguladora p101 de la invención incluyen las secuencias de aminoácidos mostradas en la Fig. 2 (SEC ID NO:2) o la Fig. 4 (SEC ID NO: 12), o la secuencia de aminoácidos codificada por el clon de ADNc pCMV3mycp 101, según lo depositado en la ATCC. Además, las subunidades reguladoras p101 de otras especies se encuentran englobadas por la invención. De hecho, cualquier proteína de la subunidad reguladora p101 codificada por las secuencias de nucleótidos de p101 descritas anteriormente, se encuentran dentro del ámbito de la invención.

Las secuencias aminoacídicas de la subunidad catalítica p120 de la invención incluyen el clon de ADNc pCMV3 mycp 120, según lo depositado en la ATCC. La presente solicitud también engloba las subunidades catalíticas p120 de otras especies, y las proteínas de p120 codificadas por las secuencias de nucleótidos de p120 descritas previamente en la sección anterior.

La invención también engloba proteínas que son funcionalmente equivalentes a la subunidad reguladora p101 codificada por las secuencias de nucleótidos descritas anteriormente, considerando varios criterios, incluyendo pero no limitado a la capacidad de unión a la subunidad catalítica, la afinidad de unión por la subunidad catalítica, la capacidad

de estimular la actividad de PI3K de la subunidad catalítica en respuesta a las proteínas G triméricas activadas, el efecto biológico resultante de la unión de la subunidad catalítica y la respuesta a la activación de las proteínas G triméricas, *por ejemplo*, transducción de señal, un cambio en el metabolismo celular (*por ejemplo*, generación de PtdIns(3,4,5) P₃) o cambio en el fenotipo cuando el equivalente de la subunidad reguladora de p101 se encuentra presente en un tipo celular apropiado (como la explosión de superóxido en neutrófilos). Estas proteínas de la subunidad reguladora p101 funcionalmente equivalentes incluyen pero no se limitan a adiciones o sustituciones de residuos aminoacídicos dentro de la secuencia de aminoácidos codificada por las secuencias de nucleótidos de p101 descritas anteriormente, pero que dan lugar a un cambio silencioso, produciendo de este modo un producto funcionalmente equivalente del gen. Las sustituciones de aminoácidos se pueden hacer en base a la semejanza en la polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad, y/o la naturaleza anfipática de los residuos implicados. Por ejemplo, los aminoácidos (hidrofóbicos) no polares incluyen alanina, leucina, isoleucina, valina, prolina, fenilalanina, triptófano, y metionina; los aminoácidos de polaridad neutra incluyen glicina, serina, treonina, cisteína, tirosina, asparagina, y glutamina; los aminoácidos (básicos) cargados positivamente incluyen arginina, lisina, e histidina; y los aminoácidos (ácidos) cargados negativamente incluyen ácido aspártico y ácido glutámico. De forma similar, la invención también engloba equivalentes funcionales de las proteínas de p120, como se ha descrito anteriormente.

Mientras que se pueden generar mutaciones al azar al ADN de p101 (usando técnicas de mutagénesis aleatoria bien conocidas por los expertos en la materia) y realizar pruebas de actividad a las subunidades reguladoras p101 mutantes resultantes, mutaciones dirigidas en la secuencia de codificación de p101 se pueden generar por ingeniería genética (usando técnicas de mutagénesis dirigida bien conocidas por los expertos en la materia) para generar subunidades reguladoras p101 mutantes con una función aumentada, *por ejemplo*, mayor afinidad de unión por la subunidad catalítica, y/o mayor capacidad de señalización; o función disminuida, *por ejemplo*, menor afinidad de unión por la subunidad catalítica, y/o menor capacidad de transducción de señal.

Por ejemplo, la secuencia de aminoácidos de p101 de cerdo se puede alinear con la subunidad reguladora p101 humana. Las subunidades reguladoras de p101 mutantes pueden ser diseñadas por ingeniería genética de modo que se mantengan las regiones de identidad inter-especie, mientras que los residuos variables son alterados, *por ejemplo*, mediante supresión o inserción de uno o varios residuos aminoacídicos o mediante sustitución de uno o más residuos aminoacídicos diferentes. Las alteraciones conservativas en las posiciones variables se pueden diseñar por ingeniería para producir una subunidad reguladora p101 mutante que conserve la función; *por ejemplo*, afinidad de unión a la subunidad catalítica o capacidad de transducción de la proteína G activada o ambas. Los cambios no-conservativos se pueden diseñar por ingeniería genética en estas posiciones variables para alterar la función, *por ejemplo*, afinidad de unión a la subunidad catalítica o capacidad de transducción de la señal o ambas. Alternativamente, donde se desea la alteración de la función, la supresión o las alteraciones no-conservativas de las regiones conservadas pueden ser diseñadas por ingeniería genética. Un experto en la materia puede probar fácilmente estas subunidades reguladoras mutadas o suprimidas de p101 para evaluar estas alteraciones en la función usando la descripción de esta memoria.

Se pueden realizar otras mutaciones a la secuencia de codificación de p101 para generar subunidades reguladoras p101 que son más adecuadas para la expresión, escalado, etc. en las células hospedadoras elegidas. Por ejemplo, el código triplete para cada aminoácido se puede modificar para adaptarse más al codón preferencial usado por la maquinaria de traducción de la célula hospedadora.

Péptidos que corresponden a uno o más dominios (o a una parte de un dominio) de la subunidad reguladora p101 (*por ejemplo*, el dominio de unión a p120, el dominio de interacción con la proteína G, o los dominios definidos por los residuos aminoacídicos desde aproximadamente 1 a 150, 151 a 300, 301 a 450, 451 a 600, 601 a 732 (cerdo) o 601 a 735 (humano), y 733 a 877 (cerdo) o 736 a 881 (humano)), subunidades reguladoras p101 truncadas o suprimidas (*por ejemplo*, subunidad reguladora p101 en la que se suprimen las partes de uno o más de los dominios anteriores) así como proteínas de fusión en las que la subunidad reguladora p101 de longitud completa, un péptido de la subunidad reguladora p101 o una subunidad reguladora p101 truncada, se encuentra fusionada a una proteína sin relación también se encuentra dentro del ámbito de la invención y se pueden diseñar en base los nucleótidos de p101 y las secuencias aminoacídicas de la subunidad reguladora p101 descritas en esta sección y anteriormente. Estas proteínas de fusión incluyen pero no se limitan a fusiones a un fragmento de epítipo (tal como se ejemplifica en los Ejemplos posteriores); o fusiones a una enzima, a una proteína fluorescente, o a una proteína luminiscente que proporcionan una función de marcador.

Mientras que los polipéptidos y los péptidos de la subunidad reguladora p101 se pueden sintetizar químicamente (*por ejemplo*, véase Creighton, 1983, Proteins: Structures and Molecular Principles, W.H. Freeman & Co., N.Y.), los grandes polipéptidos derivados de la subunidad reguladora p101 y la subunidad reguladora p101 de longitud completa se puede producir ventajosamente por tecnología del ADN recombinante usando técnicas bien conocidas en la técnica para expresar un ácido nucleico que contiene las secuencias codificantes y/o las secuencias del gen de p101. Estos métodos se pueden utilizar para construir vectores de expresión que contienen las secuencias de nucleótidos de p101 descritas anteriormente y señales de control de transcripción y traducción apropiadas. Estos métodos incluyen, por ejemplo, técnicas de ADN recombinante *in vitro*, técnicas sintéticas, y recombinación genética *in vivo*. Véanse, por ejemplo, las técnicas descritas en Sambrook *et al.*, 1989, referencia anterior, y Ausubel *et al.*, 1989, referencia anterior. Alternativamente, se puede sintetizar químicamente ARN capaz de codificar secuencias de nucleótidos de p101 sintetizadas usando, por ejemplo, sintetizadores. Véanse, por ejemplo, las técnicas descritas en "Oligonucleotide Synthesis", 1984, Gait, M.J. ed., IRL Press, Oxford, que se incorpora por referencia en esta memoria en su totalidad.

Se pueden utilizar varios sistemas de vectores de expresión en los hospedadores para expresar las secuencias de nucleótidos de p101 de la invención. Cuando el péptido o polipéptido de la subunidad reguladora de p101 es un derivado soluble, el péptido o el polipéptido se pueden recuperar del cultivo, es decir, de la célula hospedadora en los casos en los que el péptido o polipéptido de la subunidad reguladora p101 no se secreta, y de los medios de cultivo en los casos en los que el péptido o polipéptido de la subunidad reguladora p101 es secretado por las células. Sin embargo, estas células modificadas genéticamente se pueden utilizar en situaciones en las que es importante no sólo conservar las características estructurales y funcionales de la subunidad reguladora p101, sino también determinar la actividad biológica, *por ejemplo*, en ensayos de cribaje de fármacos.

Los sistemas de expresión que se pueden utilizar para los propósitos de la invención incluyen pero no se limitan a microorganismos tales como bacterias (*por ejemplo*, *E. coli*, *B. subtilis*) transformados con ADN recombinante de bacteriófagos, con ADN de plásmidos o con vectores de expresión de ADN de cósmidos que contienen secuencias de nucleótidos de p101; levaduras (*por ejemplo*, *Saccharomyces*, *Pichia*) transformadas con vectores de expresión de levaduras recombinantes que contienen las secuencias de nucleótidos de p101; sistemas de células de insectos infectadas con vectores de expresión de virus recombinantes (*por ejemplo*, baculovirus) que contiene las secuencias de p101; sistemas de células de plantas infectadas con vectores de expresión víricos recombinantes (*por ejemplo*, virus del mosaico de la coliflor, CaMV; virus del mosaico del tabaco, TMV) o transformados con vectores de expresión plasmídicos recombinantes (*por ejemplo*, plásmido Ti) que contienen secuencias de nucleótidos de p101; o sistemas de células de mamíferos (*por ejemplo*, COS, CHO, BHK, 293, 3T3, U937) que contienen construcciones de expresión recombinantes que contienen promotores provenientes del genoma de células de mamíferos (*por ejemplo*, promotor de metalotioneína) o de virus de mamíferos (*por ejemplo*, el promotor tardío de adenovirus, el promotor 7.5 K de virus vacunal).

En sistemas bacterianos, se pueden seleccionar de forma ventajosa varios vectores de expresión dependiendo del uso previsto del producto génico de p101 que es expresado. Por ejemplo, cuando se debe producir una gran cantidad de esta proteína, para la generación de composiciones farmacéuticas de proteína de la subunidad reguladora p101 o para producir anticuerpos para la proteína de la subunidad reguladora p101, *por ejemplo*, pueden ser deseables vectores que dirigen la expresión de altos niveles de los productos de la proteína de fusión y que se purifican fácilmente. Estos vectores incluyen, pero no se limitan, al vector de expresión de *E. coli* pUR278 (Ruther *et al.*, 1983, EMBO J. 2: 1791), en el que la secuencia de codificación de p101 puede estar ligada individualmente al vector en pauta con la región codificante de *lacZ* para producir una proteína de fusión; vectores pIN (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 264:5503-5509); y similares. Los vectores pGEX también se pueden utilizar para expresar polipéptidos extraños como proteínas de fusión con la glutatión S-transferasa (GST). Si la secuencia insertada codifica un polipéptido relativamente pequeño (menos de 25 kD), tales proteínas de fusión son generalmente solubles y se pueden purificar fácilmente a partir de las células lisadas por adsorción a bolitas de glutatión-agarosa seguido de elución en presencia de glutatión libre. Los vectores de pGEX se diseñan para incluir sitios de corte de trombina o de proteasa del factor Xa de modo que se pueda liberar el producto clonado del gen diana de la fracción de GST. Alternativamente, si la proteína de fusión resultante es insoluble y forma cuerpos de inclusión en la célula hospedadora, los cuerpos de inclusión se pueden purificar y la proteína recombinante solubilizar usando las técnicas bien conocidas a los expertos de la materia.

En un sistema de insecto, el virus de polihidrosis nuclear de *Autographa californica* (AcNPV) se puede utilizar como vector para expresar genes foráneos. (*por ejemplo*, véase Smith *et al.*, 1983, J. Virol. 46: 584; Smith, U.S. patent No. 4.215.051). En una realización específica descrita a continuación, las células de insecto Sf9 se infectan con vectores de baculovirus que expresan una construcción de p120 con un fragmento de 6xHIS o una construcción de p101 con un fragmento (EE).

En células hospedadoras de mamíferos, se pueden utilizar unos cuantos sistemas de expresión basados en virus. Las realizaciones específicas descritas más ampliamente debajo expresan secuencias de ADNc de p101 o p120 con cola usando un promotor CMV para expresar transitoriamente la proteína recombinante en las células U937 o en las células Cos-7. Alternativamente, los sistemas de vectores retrovirales bien conocidos en la materia se pueden utilizar para insertar la construcción de expresión recombinante en las células hospedadoras. Por ejemplo, los sistemas de vectores retrovirales para transducir las células hematopoyéticas han sido descritos en las solicitudes de PCT publicadas WO 96/09400 y WO 94/29438.

En los casos en que se utiliza un adenovirus como vector de expresión, la secuencia de nucleótidos de p101 de interés se puede ligar a un complejo de control de la transcripción/traducción de adenovirus, *por ejemplo*, el promotor tardío y la secuencia tripartida líder. Este gen quimérico se puede insertar a continuación en el genoma del adenovirus mediante recombinación *in vitro* o *in vivo*. La inserción en una región no esencial del genoma viral (*por ejemplo*, región EI o E3) dará lugar a un virus recombinante que es viable y capaz de expresar el producto génico de p101 en células hospedadoras infectadas (*por ejemplo*, véase Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA. 81:3655-3659). Las señales específicas de iniciación también pueden ser necesarias para la traducción eficiente de las secuencias de nucleótidos de p101 insertadas. Estas señales incluyen el codón de iniciación ATG y las secuencias adyacentes. En los casos en los que un ADNc o un gen de p101 entero, incluyendo su propio codón de iniciación y secuencias adyacentes, se inserta en el vector de expresión apropiado, puede que no sean necesarias señales de control de traducción adicionales. Sin embargo, en casos en que solamente una parte de la secuencia de codificación de p101 es insertada, se deben proporcionar las señales de control de traducción exógenas, incluyendo, quizás, el codón de iniciación ATG. Además, el codón de iniciación se debe encontrar en fase con la pauta de lectura de la secuencia de codificación deseada para

asegurar la traducción del inserto completo. Estas señales de control de la traducción exógenas y codones de iniciación pueden ser de diferentes orígenes, tanto naturales como sintéticos. La eficiencia de la expresión se puede aumentar mediante la inclusión de los elementos potenciadores de la transcripción, los terminadores de la transcripción, etc. apropiados. (Véase a Bittner *et al.*, 1987, *Methods in Enzymol.* 153:516-544).

Además, se puede seleccionar una cepa celular hospedadora que module la expresión de las secuencias insertadas, o modifique y procese el producto del gen de una forma específica deseada. Tales modificaciones (*por ejemplo*, glicosilación) y procesado (*por ejemplo*, escisión) de los productos proteicos pueden ser importantes para la función de la proteína. Diferentes células hospedadoras tienen mecanismos característicos y específicos para el proceso y la modificación post-traducción de proteínas y de productos genéticos. Se pueden elegir líneas celulares o sistemas hospedadores apropiados para asegurar la modificación y el procesamiento correctos de la proteína foránea expresada. Con este fin, se pueden utilizar las células hospedadoras de eucariotas que poseen la maquinaria celular para procesar apropiadamente el transcrito primario. Estas células hospedadoras de mamífero incluyen pero no se limitan a células CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, y U937.

Para la producción a largo plazo y de alto rendimiento de proteínas recombinantes, se prefiere la expresión estable. Por ejemplo, se pueden modificar por ingeniería genética las líneas celulares que expresan de forma estable las secuencias de p101 descritas anteriormente. En lugar de usar vectores de expresión que contienen orígenes de la replicación virales, las células hospedadoras se pueden transformar con ADN controlado por elementos de control de expresión adecuados (*por ejemplo*, promotor, secuencias potenciadoras, terminadores de la transcripción, sitios de poliadenilación, etc.), y un marcador seleccionable. Después de la introducción del ADN foráneo se deja que las células modificadas genéticamente crezcan durante 1-2 días en un medio enriquecido, y a continuación se cambian a un medio selectivo. El marcador seleccionable en el plásmido recombinante confiere resistencia a la selección y permite que las células integren de forma estable el plásmido en sus cromosomas y que crezcan para formar focos que a su vez se pueden clonar y ampliar en líneas celulares. Este método se puede utilizar ventajosamente para diseñar por ingeniería genética líneas celulares que expresan el producto génico de p101. Tales líneas celulares modificadas por ingeniería genética pueden ser particularmente útiles en el cribaje y la evaluación de los compuestos que afectan a la actividad endógena del producto del gen de p101.

Se pueden utilizar varios sistemas de selección, incluyendo pero no limitado a genes de timidina quinasa del virus herpes simplex (Wigler *et al.*, 1977, *Cell* 11:223), hipoxantina-guanina fosforibosiltransferasa (Szybalska & Szybalski, 1962, *Proc. Natl. Acad. Sci. USA.* 48:2026), y adenina fosforibosiltransferasa (Lowy *et al.*, 1980, *Cell* 22:817) que se pueden emplear en células tk⁻, hgp⁺ o apr⁺, respectivamente. También, se puede utilizar la resistencia a antimetabolitos como base de selección para los siguientes genes: dhfr, que confiere resistencia a metotrexato (Wigler *et al.*, 1980, *Natl. Acad. Sci. USA.* 77:3567; O'Hare *et al.*, 1981, *Proc. Natl. Acad. Sci. USA.* 78:1527); gpt, que confiere resistencia a ácido micofenólico (Mulligan & Berg, 1981, *Proc. Natl. Acad. Sci. USA.* 78:2072); neo, que confiere resistencia al aminoglicósido G-418 (Colberre-Garapin *et al.*, 1981, *J. Mol. Biol.* 150: 1); e hyg⁺, que confiere resistencia a higromicina (Santerre *et al.*, 1984, *Gene* 30: 147).

Los productos génicos de p101 se pueden expresar también en animales transgénicos. Se pueden usar animales de cualquier especie, incluyendo, pero no limitado a ratones, ratas, conejos, cobayas, cerdos, micro-cerdos, cabras, y primates no humanos, *por ejemplo*, babuinos, monos, y chimpancés para generar animales transgénicos de p101.

Cualquier técnica conocida en la materia se puede utilizar para introducir el transgen p101 en animales no humanos para producir las líneas fundadoras de los animales transgénicos no humanos. Estas técnicas incluyen, pero no se limitan a la microinyección pronuclear (Hoppe, P.C. and Wagner, T.E., 1989, U.S. Pat. No. 4.873.191); la transferencia genética medida por retrovirus en líneas germinales (Van der Putten *et al.*, 1985, *Proc. Natl. Acad. Sci., USA.* 82:6148-6152); clonación dirigida ("gen targeting") en células madre embrionarias no humanas (Thompson *et al.*, 1989, *Cell* 56:313-321); electroporación de embriones (Lo, 1983, *Mol Cell. Biol.* 3:1803-1814); y transferencia genética mediada por esperma (Lavitrano *et al.*, 1989, *Cell* 57:717-723); etc. Para una revisión de estas técnicas, véase Gordon, 1989, *Transgenic Animals*, *Intl. Rev. Cytol.* 115:171-229.

La presente invención describe animales transgénicos no humanos que llevan el transgen de p101 en todas sus células, así como animales que llevan el transgen en algunas, aunque no en todas sus células, es decir, animales mosaico. El transgen se puede integrar como un transgen único o en concatámeros, *por ejemplo*, tándems cabeza-cabeza o tándems cabeza-cola. El transgen también se puede introducir de forma selectiva y activar en un tipo celular particular siguiendo, por ejemplo, las descripciones de Lasko *et al.* (Lasko, M. *et al.*, 1992, *Proc. Natl. Acad. Sci. USA.* 89: 6232-6236). Las secuencias reguladoras requeridas para esta activación específica de tipo celular dependen del tipo celular de interés particular, y serán evidentes a los expertos en la materia. Cuando se desea que el transgen del gen de p101 se integre en el sitio cromosómico del gen de p101 endógeno, se prefiere clonación dirigida. En resumen, cuando se utiliza esta técnica, se diseñan vectores que contienen algunas secuencias de nucleótidos homólogas al gen de p101 endógeno con el fin de integrar, mediante la recombinación homóloga con secuencias cromosómicas, e interrumpir la función de la secuencia de nucleótidos del gen de p101 endógeno. De esta manera, la expresión del gen de p101 endógeno también se puede eliminar insertando secuencias no funcionales en el gen endógeno. El transgen también se puede introducir selectivamente en un tipo celular particular, inactivando de este modo el gen de p101 endógeno solamente en ese tipo celular, siguiendo, por ejemplo, las descripciones de Gu *et al.* (Gu *et al.*, 1994, *Science* 265: 103-106). Las secuencias reguladoras requeridas para esta inactivación específica de tipos celulares dependerán del tipo celular particular de interés, y serán evidentes a los expertos en la materia.

Una vez se hayan generado los animales transgénicos, la expresión del gen de p101 recombinante se puede evaluar utilizando técnicas estándares. El cribaje inicial se puede lograr mediante análisis de transferencia Southern (“*Southern blot*”) o técnicas de PCR para analizar tejidos animales y para analizar si se ha producido la integración del transgen. El nivel de expresión de mRNA del transgen en los tejidos de los animales transgénicos también se puede evaluar usando técnicas que incluyen pero no se limitan a análisis por transferencia Northern (“*Northern blot*”) de muestras de tipos celulares obtenidas del animal, análisis de hibridación *in situ* y RT-PCR. Muestras de tejido que expresa el gen de p101, también se pueden evaluar por inmunocitoquímica usando anticuerpos específicos para el producto transgénico de p101, como se describe a continuación.

5.3. Anticuerpos para las proteínas p101 y p120

Los anticuerpos que reconocen de manera específica uno o más epítomos de la subunidad reguladora p101, o epítomos de las variantes conservadas de la subunidad reguladora p101 o fragmentos del péptido de la subunidad reguladora p101 están englobados en la invención. También se describen en el presente documento anticuerpos que reconocen uno o más epítomos de la proteína p120, en particular, la nueva región carboxilo terminal. Dichos anticuerpos incluyen pero no se limitan a, anticuerpos policlonales, anticuerpos monoclonales (mAbs), anticuerpos humanizados o quiméricos, anticuerpos de cadena única, fragmentos Fab, fragmentos F(ab')₂, fragmentos producidos por una librería de expresión Fab, anticuerpos anti-idiotípicos (anti-Id) y fragmentos de unión a epítomo de cualquiera de los anteriores.

Los anticuerpos de la invención se pueden utilizar, por ejemplo, en la detección de la subunidad reguladora p101 o p120 en una muestra biológica y se puede utilizar, por lo tanto, como parte de una técnica de diagnóstico o pronóstico por lo que se pueden estudiar las cantidades anormales de estas proteínas en los pacientes. Dichos anticuerpos se pueden utilizar junto con, por ejemplo, esquemas de identificación de compuestos, según se describe más abajo, en la Sección 5.5, para la evaluación del efecto de los compuestos de estudio en la expresión y/o actividad de los productos de los genes p101 y p120. Adicionalmente, dichos anticuerpos se pueden utilizar junto con técnicas de terapia génica descritas más abajo, en la Sección 5.6 para, por ejemplo, evaluar las células que expresan la subunidad reguladora p101 normal y/o manipulada genéticamente antes de su introducción en el paciente. Dichos anticuerpos se pueden utilizar adicionalmente como método para la inhibición de la subunidad reguladora p101 anormal o la actividad p120. De esta manera, dichos anticuerpos se pueden utilizar como parte de métodos de tratamiento de una enfermedad inflamatoria.

Para la producción de anticuerpos, se pueden inmunizar varios animales hospedadores inyectando la subunidad reguladora p101, un péptido de la subunidad reguladora p101, polipéptidos truncados de la subunidad reguladora p101, equivalentes funcionales de la subunidad reguladora p101 o mutantes de la subunidad reguladora p101. Adicionalmente, los animales hospedadores se pueden inmunizar inyectando la subunidad catalítica p120 o péptidos de la subunidad p120. Dichos animales hospedadores pueden incluir pero no se limitan a conejos, ratones y ratas, por nombrar algunos. Se pueden utilizar varios adyuvantes para aumentar la respuesta inmune, dependiendo de la especie del hospedador, incluyendo pero no limitándose a adyuvante de Freund (completo o incompleto), geles minerales tales como el hidróxido de aluminio, sustancias activas de superficie como la lisolecitina, polioles plurónicos, polianiones, péptidos, emulsiones en aceite, hemocianina de *keyhole limpet*, dinitrofenol y adyuvantes humanos potencialmente útiles tales como BCG (bacilo de Calmette-Guerin) y *Corynebacterium parvum*. Los anticuerpos policlonales son poblaciones heterogéneas de moléculas de anticuerpo derivadas de sueros de animales inmunizados.

Los anticuerpos monoclonales, que son poblaciones homogéneas de anticuerpos para un antígeno concreto, se pueden obtener mediante cualquier técnica que proporcione la producción de moléculas de anticuerpo mediante líneas de células continuas en cultivo. Estas incluyen, pero no se limitan a, la técnica del hibridoma de Kohler y Milstein (1975, Nature 256:495-497; y patente US N° 4.376.110), la técnica del hibridoma de células B humanas (Kosbor *et al.*, 1983, Immunology Today 4:72, Cole *et al.*, 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030) y la técnica del hibridoma de EBV (Cole *et al.*, 1985, Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Dichos anticuerpos pueden ser de cualquier clase de inmunoglobulina incluyendo IgG, IgM, IgE, IgD y cualquier subclase de las mismas. El hibridoma que produce el mAb de esta invención se puede cultivar *in vitro* o *in vivo*. La producción de títulos elevados de mAb *in vivo* lo hace el método de producción actualmente preferido.

Además, se pueden utilizar técnicas desarrolladas para la producción de “anticuerpos quiméricos” (Morrison *et al.*, 1984, Proc. Natl. Acad. Sci. USA, 81:6851-6855; Neuberger *et al.*, 1984, Nature, 312:604-608; Takeda *et al.*, 1985, Nature, 314: 452-454) mediante el empalme de genes procedentes de una molécula de anticuerpo de ratón de especificidad de antígeno adecuada junto con genes de una molécula de anticuerpo humano de una actividad biológica adecuada. Un anticuerpo quimérico es una molécula en la que se derivan porciones procedentes de diferentes especies animales, tales como aquellas que tienen una región variable derivada de un mAb de cerdo y la región constante de una inmunoglobulina humana.

Alternativamente, se pueden adaptar técnicas descritas para la producción de anticuerpos de cadena sencilla (patente US No 4.946.778; Bird, 1988, Science 242: 423-426; Huston *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85: 5879-5883; y Ward *et al.*, 1989, Nature 334:544-546) para producir anticuerpos de cadena única contra productos del gen de p101. Los anticuerpos de cadena única están formados por la unión de fragmentos de cadena pesada y ligera de la región Fv mediante un puente aminoácido, resultando en un polipéptido de cadena única.

Se pueden generar fragmentos de anticuerpo que reconozcan epítomos específicos mediante técnicas conocidas. Por ejemplo, dichos fragmentos incluyen pero no se limitan a: los fragmentos de F'(ab)₂ que se pueden producir mediante

digestión con pepsina de la molécula de anticuerpo y los fragmentos de Fab que se pueden generar reduciendo los puentes disulfuro de los fragmentos F'(ab)₂. De manera alternativa, se pueden construir librerías de expresión de Fab (Huse *et al.*, 1989, Science, 246: 1275-1281) para permitir la identificación rápida y fácil de fragmentos monoclonales de Fab con la especificidad deseada.

Los anticuerpos para la subunidad reguladora p101 o la subunidad catalítica p120 pueden, a su vez, ser utilizados para generar anticuerpos anti-idiotipo que "mimetizen" la subunidad reguladora p101 o la subunidad p120, respectivamente, utilizando técnicas bien conocidas para aquellos expertos en la materia. (Véase, p.e., Greenspan y Bona, 1993, FASEB J 7(5):437-444; y Nissinoff, 1991, J. Immunol. 147(8):2429-2438). Por ejemplo, se pueden utilizar anticuerpos que se unan a la subunidad reguladora p101 e inhibir de manera competitiva la unión de la subunidad catalítica a la subunidad reguladora p101 para generar anti-idiotipos que "mimetizen" la subunidad reguladora p101 y, por lo tanto, unirse y neutralizar la subunidad catalítica. Dichos anti-idiotipos neutralizantes o fragmentos de Fab de dichos anti-idiotipos se pueden utilizar en regímenes terapéuticos para neutralizar la subunidad catalítica y reducir la inflamación.

5.4. Diagnóstico de enfermedades por activación de células hematopoyéticas

Se puede utilizar una variedad de procedimientos para el diagnóstico y la evaluación pronóstico de trastornos en la activación de células del linaje hematopoyético, incluyendo alteraciones inflamatorias, y para la identificación de sujetos que tengan una predisposición a dichos trastornos.

Estos procedimientos pueden utilizar, por ejemplo, reactivos tales como secuencias de nucleótidos de p101 y p120 descritas más arriba, y los anticuerpos de la subunidad reguladora p101 y p120, según se ha descrito en la sección 5.3. Específicamente, dichos reactivos se pueden utilizar, por ejemplo, para: (1) la detección de presencia de mutaciones en el gen de p101 y p120, o la detección de la sobreexpresión o subexpresión del ARNm de p101 o p120 relativo al estado de la alteración de la activación de células del linaje hematopoyético; (2) la detección de la abundancia en exceso o defecto del producto del gen p101 o p120 relativo al estado de alteración de la activación de células de linaje hematopoyético; y (3) la detección de perturbaciones o anomalías en la ruta de señalización de la transducción mediada por la subunidad reguladora p101 y la subunidad catalítica p120.

Los procedimientos descritos aquí se pueden llevar a cabo, por ejemplo, utilizando kits de diagnóstico pre-empaquetados que comprenden al menos una secuencia nucleotídica específica de p101 o p120 o un anticuerpo reactivo de la subunidad reguladora p101 o p120 descrito en el presente documento, que se puede utilizar convenientemente, p.e., en ensayos clínicos, para diagnosticar pacientes que exhiban anomalías de un trastorno de la activación de células del linaje hematopoyético.

Para la detección de mutaciones en p101 o p120, se puede utilizar cualquier célula nucleada como fuente de partida de ácido nucleico genómico. Para la detección de la expresión del gen de p101 o p120 o los productos de los genes de p101 o p120, se puede utilizar cualquier tipo de célula o tejido en el que el gen de p101 o p120 se exprese, tal como por ejemplo, células neutrófilas.

Las técnicas de detección basadas en ácido nucleico se describen más abajo, en la sección 5.4.1. Técnicas de detección de péptidos se describen más abajo, en la sección 5.4.2.

5.4.1. Detección del gen de p101 y transcritos

Las mutaciones dentro de los genes de p101 y p120 se pueden detectar utilizando una serie de técnicas. El ácido nucleico procedente de cualquier célula nucleada se puede utilizar como punto de partida para tales técnicas de ensayo, y se puede aislar de acuerdo con cualquier de los procedimientos de preparación de ácido nucleico estándar, los cuales son bien conocidos para aquellos expertos en la materia.

El ADN se puede utilizar en ensayos de hibridación o amplificación de muestras biológicas para detectar anomalías que impliquen la estructura del gen, incluyendo mutaciones puntuales, inserciones, deleciones y reorganizaciones cromosómicas. Dichos ensayos pueden incluir, pero no se limitan a, análisis Southern, análisis de polimorfismos conformacionales de cadena única (SSCP) y análisis de PCR.

Dichos procedimientos de diagnóstico para la detección de mutaciones específicas en el gen de p101 o p120 pueden implicar, por ejemplo, poner en contacto e incubar ácidos nucleicos que incluyen moléculas de ADN recombinantes, genes clonados o variantes degeneradas de las mismas, obtenidas de una muestra, p.e., derivadas de una muestra de paciente u otra fuente celular apropiada, con uno o más reactivos de ácido nucleico marcados, incluyendo moléculas de ADN recombinantes, genes clonados o variantes degeneradas de los mismos, como se ha descrito más arriba, en condiciones favorables para el emparejamiento específico de estos reactivos a sus secuencias complementarias dentro del gen de p101 o p120. Preferiblemente, las longitudes de estos reactivos de ácido nucleico son al menos de 15 a 30 nucleótidos. Después de la incubación, todos los ácidos nucleicos que no se han emparejado se eliminaron de la molécula híbrida de ácido nucleico. Se puede detectar, a continuación, la presencia de ácidos nucleicos que han hibridado, si existen tales moléculas. Utilizando dicho esquema de detección, se puede inmovilizar el ácido nucleico procedente del tipo celular o tejido de interés, por ejemplo, en un soporte sólido como una membrana, o una superficie de plástico como la de una placa de microtitulación o perlas de poliestireno. En este caso, después de la incubación, los reactivos de ácido nucleico marcados sin emparejar del tipo descrito más arriba son eliminados fácilmente. La

detección de los reactivos de ácido nucleico marcados de p101 o p120 emparejados se lleva a cabo utilizando técnicas estándar bien conocidas para aquellos expertos en la materia. Se pueden comparar las secuencias de los genes de p101 o p120 a las que se han emparejado los reactivos de ácido nucleico con el patrón de emparejamiento esperado para una secuencia del gen normal con el fin de determinar si está presente la mutación.

Procedimientos de diagnóstico alternativos para la detección de moléculas de ácido nucleico específicas del gen de p101 o p120, en muestras de pacientes u otras fuentes celulares adecuadas, pueden implicar su amplificación, p.e., mediante PCR (la realización experimental descrita por Mullis, K.B., 1987, patente U.S. No.4,683,202) seguida de la detección de las moléculas amplificadas utilizando técnicas bien conocidas para aquellos expertos en la materia. Las secuencias amplificadas resultantes se pueden comparar con aquellas que se esperarían si el ácido nucleico que se amplifica contuviera únicamente copias normales del gen p101 y p120 con el fin de determinar si existe mutación en el gen.

De manera adicional, se pueden llevar a cabo técnicas de genotipificación bien conocidas para identificar individuos que sean portadores de mutaciones en los genes de p101 o p120. Dichas técnicas incluyen por ejemplo, la utilización de polimorfismos de longitud de fragmento de restricción (RFLP) que implican variaciones de secuencia en uno de los sitios de reconocimiento para el enzima de restricción específico usado.

El nivel de expresión del gen p101 o p120 se puede determinar detectando y midiendo la transcripción de p101 o p120. Por ejemplo, el ARN de un tipo celular o tejido conocido o sospechoso de expresar el gen p101 o p120, tal como las células del linaje hematopoyético, especialmente las células mieloides y plaquetas, se pueden aislar y ensayar utilizando técnicas de hibridación o PCR como las descritas aquí más arriba. Las células aisladas se pueden derivar de un cultivo celular o de un paciente. El análisis de las células tomadas del cultivo puede ser una etapa necesaria en la evaluación de células a ser usadas como parte de una técnica de terapia génica basada en células o, alternativamente, para probar el efecto de compuestos sobre la expresión del gen p101 o p120. Dichos análisis pueden revelar tanto aspectos cuantitativos como cualitativos del patrón de expresión del gen p101 o p120, incluyendo la activación o inactivación de la expresión del gen p101 o p120.

5.4.2 Detección de los productos del gen de p101

Se pueden usar también anticuerpos dirigidos contra productos del gen de p101 o p120 natural o mutante o variantes conservadas o fragmentos peptídicos de los mismos, los cuales se han discutido más arriba, en la sección 5.3, como diagnóstico y pronóstico de trastornos por activación de células del linaje hematopoyético, según se describe en el presente documento. Dichos métodos de diagnóstico se pueden usar para detectar anomalías en el nivel de expresión del gen de p101 o p120 o anomalías en la estructura y/o localización temporal, tisular, celular o subcelular de la subunidad reguladora p101 y se puede llevar a cabo *in vivo* o *in vitro*, tal como, por ejemplo, en un tejido de biopsia.

El tejido o tipo celular a ser analizado incluirá generalmente aquellas que son conocidas o sospechosas de contener células que expresan el gen de p101 o p120, tal como por ejemplo células neutrófilas que han infiltrado e inflamado el tejido. Procedimientos de aislamiento de proteínas utilizados en el presente documento pueden por ejemplo ser los descritos en Harlow y Lane (Harlow E. y Lane D., 1988, "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York), que se incorpora aquí por referencia en su totalidad. Las células aisladas se pueden derivar de un cultivo celular o de un paciente. El análisis de las células tomadas del cultivo puede ser una etapa necesaria en la valoración de células que podrían ser utilizadas como parte de una técnica de terapia génica basada en células o alternativamente para determinar el efecto de compuestos sobre la expresión del gen p101 o p120.

Por ejemplo, se pueden utilizar anticuerpos o fragmentos de anticuerpos tales como los descritos más arriba en la Sección 5.3, útiles en la presente invención para detectar de manera cuantitativa o cualitativa la presencia de los productos de los genes p101 o p120 o variantes conservadas o fragmentos peptídicos de los mismos. Esto se puede llevar a cabo, por ejemplo, mediante técnicas de inmunofluorescencia utilizando un anticuerpo marcado fluorescentemente (véase más abajo, esta sección) acoplado con microscopio óptico, citometría de flujo o detección fluorimétrica.

Los anticuerpos (o fragmentos de los mismos) o proteínas de fusión o conjugadas útiles en la presente invención pueden ser empleadas, adicionalmente, histológicamente, como ensayos de inmunofluorescencia, microscopía inmunoelectrónica o no inmunoensayos, para la detección *in situ* de productos de los genes p101 o p120 o variantes conservadas o fragmentos peptídicos de los mismos o para la unión de la subunidad catalítica (en el caso de la proteína de fusión de la subunidad catalítica marcada).

La detección *in situ* se puede llevar a cabo extrayendo una muestra histológica de un paciente y aplicando a la misma un anticuerpo marcado o una proteína de fusión de la presente invención. El anticuerpo (o fragmento) o proteína de fusión se aplica preferiblemente superponiendo el anticuerpo marcado (o fragmento) en una muestra biológica. Mediante el uso de este procedimiento es posible determinar no sólo la presencia del producto del gen p101 o p120, variantes conservadas o fragmentos de los mismos sino también su distribución en el tejido examinado. Utilizando la presente invención, aquellos con conocimiento general percibirán fácilmente que cualquiera de los numerosos procedimientos histológicos (como procedimientos de tinción) se pueden modificar con el fin de conseguir dicha detección *in situ*.

Los inmunoensayos y no inmunoensayos para productos del gen p101 o p120 o variantes conservadas o fragmentos peptídicos de los mismos comprenderán típicamente incubar una muestra, tal como un fluido biológico, un extracto tisular, células recogidas recientemente o lisados de células que han sido incubados en un cultivo celular, en presencia de un anticuerpo marcado de manera detectable capaz de identificar los productos de los genes p101 o p120 o variantes conservadas o fragmentos peptídicos de los mismos y detectar el anticuerpo unido mediante cualquiera de las numerosas técnicas conocidas en el estado de la técnica.

La muestra biológica se puede poner en contacto con e inmovilizarse sobre un soporte o portador de fase sólida tal como nitrocelulosa u otro soporte sólido que sea capaz de inmovilizar células, partículas celulares o proteínas solubles. El soporte se puede lavar a continuación con tampones adecuados seguido del tratamiento con el anticuerpo o proteína de fusión marcada de manera detectable de la subunidad reguladora p101 o de subunidad p120. El soporte de fase sólida se puede lavar a continuación con el tampón una segunda vez para eliminar el anticuerpo o proteína de fusión no unida. La cantidad de marcaje unido en el soporte sólido se puede detectar a continuación mediante métodos convencionales.

“Soporte de fase sólida o portador” se pretende que englobe cualquier soporte capaz de unirse a antígeno o anticuerpo. Soportes o portadores bien conocidos incluyen vidrio, poliestireno, polipropileno, polietileno, dextrano, nylon, amilasas, celulosas naturales o modificadas, poliacrilamidas, gabbros y magnetita. La naturaleza del portador puede ser hasta cierto punto soluble o insoluble para los propósitos de la presente invención. El material del soporte puede tener virtualmente cualquier posible configuración siempre y cuando la molécula acoplada sea capaz de unirse a un antígeno o anticuerpo. De esta manera, la configuración del soporte puede ser esférica, como en una perla o cilíndrica como en la superficie interior de un tubo de ensayo o la superficie externa de una barra. De manera alternativa, la superficie puede ser plana tal como una hoja, una tira de ensayo, etc. Soportes preferidos incluyen las perlas de poliestireno. Aquellos expertos en la materia conocerán muchos otros portadores adecuados para la unión de anticuerpo o antígeno o serán capaces de averiguarlo mediante experimentación rutinaria.

Se puede determinar la actividad de unión de un lote determinado del anticuerpo o proteína de fusión de la subunidad reguladora p101 o de la subunidad p120 de acuerdo con métodos bien conocidos. Aquellos expertos en la materia serán capaces de determinar las condiciones de ensayo operativas y óptimas para cada determinación utilizando experimentación rutinaria.

Con respecto a los anticuerpos, una de las maneras en las que el anticuerpo se puede marcar de manera detectable es mediante unión de éste con un enzima y utilizar en un inmunoensayo enzimático (EIA) (Voller, “The Enzyme Linked Immunosorbent Assay (ELISA)”, 1978, Diagnostic HorizonsTM 1-7 Microbiological Associates Quarterly Publication, Walkersville, MD); Voller *et al.*, 1978, J. Clin. Pathol. 31:507-520; Butler, 1981, Meth. Enzymol. 73:482-523; Maggio (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, FL.; Ishikawa *et al.*, (eds.), 1981, Enzyme Immunoassay, Kigaku Shoin, Tokyo). El enzima que se une al anticuerpo reaccionará con un sustrato adecuado, preferiblemente un sustrato cromogénico, de manera que produzca una unidad química que se pueda detectar, por ejemplo, mediante medios espectrofotométricos, fluorimétricos o visuales. Enzimas que se puedan usar para marcar de manera detectable el anticuerpo incluye, pero no se limita a, malato deshidrogenasa, nucleasa de *Staphylococcus*, delta-5-esteroide isomerasa, alcohol deshidrogenasa de levadura, alfa-glicerofosfato deshidrogenasa, triosafosfato isomerasa, peroxidasa de rábano, fosfatasa alcalina, asparaginasa, glucosa oxidasa, beta-galactosidasa, ribonucleasa, ureasa, catalasa, glucosa-6-fosfato deshidrogenasa, glucoamilasa y acetilcolinesterasa. La detección se puede llevar a cabo mediante métodos colorimétricos que utilicen un sustrato cromogénico del enzima. La detección también se puede llevar a cabo mediante comparación visual de la extensión de la reacción enzimática de un sustrato en comparación con estándares preparados de manera similar.

La detección también se puede llevar a cabo utilizando cualquiera de la variedad de otros inmunoensayos. Por ejemplo, mediante marcaje radioactivo de los anticuerpos o fragmentos de anticuerpo es posible detectar la subunidad reguladora p101 a través del uso de un radioinmunoensayo (RIA) (véase, por ejemplo, Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, Marzo, 1986, que se incorpora aquí por referencia). El isótopo radioactivo se puede detectar mediante dichos medios así como el uso de un contador gamma o un contador de centelleo o mediante autorradiografía.

También es posible marcar el anticuerpo con un compuesto fluorescente. Cuando el anticuerpo marcado fluorescentemente se expone a la luz de longitud de onda adecuada, su presencia se puede detectar a continuación debido a la fluorescencia. De entre los compuestos de marcaje fluorescente más comúnmente utilizados están el isotiocianato de fluoresceína, rodamina, ficoeritrina, ficocianina, aloficocianina y fluorescamina.

El anticuerpo también se puede marcar de manera detectable utilizando metales emisores de fluorescencia tales como el ¹²⁵Eu u otros de la serie de lantánidos. Estos metales se pueden unir al anticuerpo utilizando dichos grupos quelantes de metales como el ácido dietilentriaminopentacético (DTPA) o el ácido etilendiaminotetraacético (EDTA).

El anticuerpo también se puede marcar de manera detectable acoplándolo a un compuesto quimioluminiscente. La presencia del anticuerpo marcado con quimioluminiscencia se determina, a continuación, detectando la presencia de luminiscencia que se produce durante el curso de la reacción química. Ejemplos de compuestos de marcaje quimioluminiscentes particularmente útiles son el luminol, isoluminol, éster del acridinio teromático, imidazol, sal de acridinio y éster de oxalato.

Igualmente, se puede utilizar un compuesto bioluminiscente para marcar el anticuerpo de la presente invención. La bioluminiscencia es un tipo de quimioluminiscencia que se encuentra en sistemas biológicos en los que una proteína catalítica incrementa la eficacia de la reacción de quimioluminiscencia. La presencia de una proteína bioluminiscente se determina detectando la presencia de luminiscencia. Compuestos bioluminiscentes importantes para los propósitos de marcaje son la luciferina, luciferasa y aequorina.

5.5. Ensayos para la identificación de compuestos que modulan la expresión o la actividad de PI3K activada por la proteína G

Se diseñan los siguientes ensayos para identificar la interacción con (p.e. unirse a) la subunidad reguladora p101 o la subunidad catalítica p120, compuestos que interaccionan con (p.e. unirse a) proteínas intracelulares que interaccionan con la subunidad reguladora p101 y/o la subunidad catalítica p120, compuestos que interfieren en la interacción de la subunidad reguladora p101 con la subunidad catalítica p120 o con otras proteínas intracelulares implicadas en la transducción de señal mediada por PI3K estimulada por proteína G y a compuestos que modulan la actividad del gen de p101 y p120 (es decir, modulan el nivel de expresión del gen p101 o p120) o modulan el nivel de p101 o p120. Se pueden utilizar, adicionalmente, ensayos que identifiquen compuestos que se unan a secuencias reguladoras del gen p101 o p120 (p.e. secuencias de promotor) y que pueden modular la expresión del gen p101 o p120. Véase, p.e., Platt, K.A., 1994, J. Biol. Chem. 269:28558-28662.

Los compuestos que se pueden identificar de acuerdo con la invención incluyen pero no se limitan a, péptidos, anticuerpos y fragmentos de los mismos, prostaglandinas, lípidos y otros compuestos orgánicos (p.e. terpinas, peptidomiméticos) que se unen a la subunidad reguladora p101 o a la subunidad catalítica p120 y o bien mimetizan la actividad provocada por el ligando (i.e. agonistas) o inhiben la actividad provocada por el ligando natural (i.e., antagonistas); así como péptidos, anticuerpos o fragmentos de los mismos y otros compuestos orgánicos que mimetizan la subunidad reguladora p101 o la subunidad catalítica p120 (o una porción de la misma) y se une a y "neutraliza" el ligando natural.

Dichos compuesto pueden incluir, pero no se limitan a, péptidos tales como por ejemplo péptidos solubles, incluyendo pero no limitando a miembros de librerías aleatorias de péptidos (véase, p.e., Lam K. S. *et al.*, 1991, Nature 354:82-84; Houghten, R. *et al.*, 1991, Nature 354:84-86), péptidos de una librería molecular derivada de química combinatoria hechos de D- y/o L-aminoácidos, fosfopéptidos (incluyendo pero no limitando a miembros de librerías de fosfopéptidos dirigidos aleatorios o parcialmente degenerados; véase, p.e., Songyang, Z. *et al.*, Cell 72:767:778), anticuerpos (incluyendo, pero no limitándose a, anticuerpos policlonales, monoclonales, humanizados, anti-idiotípicos, quiméricos o anticuerpos de cadena única, y FAb, F'(ab)₂ y fragmentos de librerías de expresión de FAb y fragmentos de unión a epítipo de los mismos); y moléculas orgánicas e inorgánicas pequeñas.

Otros compuestos que se pueden identificar de acuerdo con la invención incluyen pero no se limitan a moléculas orgánicas pequeñas que son capaces de conseguir la entrada en una célula apropiada (p.e. en el neutrófilo) y afectan a la expresión del gen de p101 o p120 o a algún otro gen implicado en la ruta de transducción de señal de la subunidad reguladora p101 (p.e. interaccionando con la región reguladora o factores de transcripción implicados en la expresión del gen); o dichos compuestos que afectan a la actividad de la subunidad reguladora p101, p.e., inhibiendo o potenciando la unión de p101 a la subunidad catalítica de la PI3K o la unión de p101 a algún otro factor implicado en la ruta de transducción de señal de la subunidad catalítica p101, tal como por ejemplo Gβγ.

La modelación por ordenador y las tecnologías de búsqueda permiten la identificación de compuestos o la mejora de compuestos ya identificados, que pueden modular la actividad o expresión de la subunidad reguladora p101 o de la subunidad catalítica p120. Habiendo identificado dicho compuesto o composición, se identifican los sitios o regiones activos. Dichos sitios activos pueden ser, típicamente, sitios pareja de unión, tales como por ejemplo, los dominios de interacción de la subunidad catalítica p120 con la subunidad reguladora p101 en sí misma. El sitio activo se puede identificar utilizando procedimientos conocidos en la materia incluyendo, por ejemplo, secuencias aminoacídicas de péptidos, secuencias nucleotídicas de ácidos nucleicos o del estudio de complejos del compuesto o composición relevante con su ligando natural. En el último caso, se pueden utilizar métodos cristalográficos de rayos X para encontrar sitios activos encontrando dónde se encuentra el ligando complejoado sobre el factor.

A continuación, se determina la estructura geométrica tridimensional del sitio activo. Esto se puede realizar mediante procedimientos conocidos, incluyendo la cristalografía de rayos X la cual puede determinar una estructura molecular completa. Por otro lado, se puede utilizar la RMN en fase sólida o líquida para determinar ciertas distancias intramoleculares. Se puede utilizar cualquier otro procedimiento experimental de determinación estructural para obtener estructuras geométricas completas o parciales. Las estructuras geométricas se pueden determinar con un ligando complejoado, natural o artificial, que puede incrementar la precisión de la estructura del sitio activo determinada.

Si se determina una estructura incompleta o con insuficiente precisión, se pueden utilizar métodos basados en modelación numérica para completar la estructura o mejorar su precisión. Se puede utilizar cualquiera de los métodos de modelación, incluyendo modelos parametrizados específicos de biopolímeros concretos tal como proteínas o ácidos nucleicos, modelos dinámicos moleculares basados en movimientos moleculares por ordenador, modelos de mecánica estadística basados en conjuntos térmicos o modelos combinados. Para la mayoría de modelos, son necesarios campos moleculares de fuerza estándar que representan las fuerzas entre átomos y grupos constituyente y se pueden seleccionar de campos de fuerza conocidos en la química física. Las estructuras experimentales incompletas o con insuficiente

precisión pueden servir como restricciones de estructuras completas y más precisas por ordenador mediante estos métodos de modelación.

Finalmente, habiendo determinado la estructura del sitio activo, ya sea experimentalmente, mediante modelación o mediante combinación, se pueden identificar los compuestos candidatos de modulación mediante búsquedas en bases de datos que contienen compuestos junto con la información sobre su estructura molecular. Esta búsqueda solicita compuestos que tengan estructuras que coinciden con la estructura determinada del sitio activo y que interaccionan con los grupos que definen el sitio activo. Dicha búsqueda puede ser manual pero preferiblemente está asistida por ordenador. Estos compuestos encontrados a partir de esta búsqueda son compuestos potenciales de modulación de PI3K activada por proteína G.

Alternativamente, estos procedimientos se pueden utilizar para identificar compuestos de modulación mejorados a partir de un compuesto o ligando de modulación ya conocido. Se puede modificar la composición del compuesto conocido y determinar los efectos estructurales de la modificación utilizando los métodos de modelación experimentales y computacionales descritos más arriba aplicados a la nueva composición. La estructura modificada se compara a continuación con la estructura del sitio activo del compuesto para determinar si resulta en un ajuste o interacción mejorada. De esta manera se pueden evaluar rápidamente variaciones sistemáticas en la composición tal como variando los grupos laterales, para obtener compuestos con modulación modificada o ligandos con especificidad o actividad mejorada.

Otros métodos experimentales y de modelación por ordenador útiles en la identificación de compuestos de modulación basados en la identificación de los sitios activos de la subunidad catalítica p120, la subunidad reguladora p101 y los factores de transcripción y traducción relacionados serán evidentes para aquellos expertos en la materia.

Ejemplos de sistemas de modelación molecular son los programas CHARMM y QUANTA (Polygen Corporation, Waltham, MA). CHARMM lleva a cabo las funciones de minimización de energía y dinámica molecular. QUANTA lleva a cabo la construcción, modelación gráfica y análisis de la estructura molecular. QUANTA permite la construcción, modificación, visualización y análisis interactivo del comportamiento de unas moléculas con otras.

Una serie de publicaciones revisan la modelación por ordenador de fármacos que interaccionan con proteínas específicas, tal como Rotivinen *et al.*, 1988, Acta Pharmaceutica Fennica 97:159-166; Ripka, New Scientist 54-57 (16 de Junio, 1988); McKinaly and Rossmann, 1989, Annu. Rev. Pharmacol. Toxicol. 29:111-122; Perry y Davies, OSAR: Quantitative Structure-Activity Relationships in Drug Design pp. 189-193 (Alan R. Liss, Inc. 1989); Lewis y Dean, 1989, Proc. R. Soc. Lond. 236:125140 y 141-162; y con respecto a un receptor modelo para componentes de ácido nucleico, Askew *et al.*, 1989, J. Am. Chem. Soc. 111:1082-1090. Hay disponibles otros programas de ordenador que identifican y describen gráficamente productos químicos de compañías como BioDesign, Inc. (Pasadena, CA.), Allelix, Inc. (Mississauga, Ontario, Canada), y Hypercube, Inc. (Cambridge, Ontario). A pesar de que estos se diseñan de manera primaria para la aplicación a fármacos específicos de proteínas determinadas, estas se pueden adaptar para el diseño de fármacos específicos de regiones de ADN o ARN, una vez se identifica la región.

A pesar de que se ha descrito más arriba con referencia al diseño y generación de compuestos que pudieran afectar a la unión, uno también podría cribar en librerías de compuestos conocidos, incluyendo productos naturales o productos químicos sintéticos y materiales biológicamente activos, incluyendo proteínas, para compuestos que sean inhibidores o activadores.

Los compuestos identificados mediante ensayos tales como los descritos en el presente documento pueden ser útiles, por ejemplo, en la elaboración de la función biológica del producto del gen de p101 o p120 y en la mejora de los trastornos por activación de células del linaje hematopoyético. Ensayos para determinar la efectividad de compuestos identificados por ejemplo mediante técnicas tales como las descritas en la Sección 5.5.1 hasta 5.5.3 se discuten más abajo, en la Sección 5.5.4.

5.5.1. Ensayos *in vitro* de identificación de compuestos que se unen a la subunidad reguladora p101

Los sistemas *in vitro* se pueden diseñar para identificar compuestos capaces de interaccionar con (p.e. uniéndose a) la subunidad reguladora p101 o la subunidad catalítica p120. Los compuestos identificados pueden ser útiles, por ejemplo, en la modulación de la actividad de productos de los genes p101 y p120 naturales y/o mutantes; se pueden utilizar en cribajes para identificar compuestos que interrumpan las interacciones normales de subunidad reguladora p101/subunidad catalítica; o pueden interrumpir por sí mismas dichas interacciones.

El principio de los ensayos utilizados para identificar compuestos que se unen a la subunidad reguladora p101 implica la preparación de una mezcla de reacción de la subunidad reguladora p101 y el compuesto prueba en condiciones y durante un tiempo suficientes para dejar interaccionar y unirse a los dos componentes, formando de esta manera un complejo que se puede eliminar y/o detectar en la mezcla de reacción. La especie de la subunidad reguladora p101 utilizada puede variar dependiendo del objetivo del ensayo de cribado. Por ejemplo, cuando se buscan los agonistas del ligando natural, se puede utilizar la subunidad reguladora p101 de longitud completa o una proteína de fusión que contenga la subunidad reguladora p101 fusionada a una proteína o polipéptido que proporcione ventajas al sistema de ensayo (p.e., marcando, aislando el complejo resultante, etc.).

Los ensayos de cribado se pueden llevar a cabo de diferentes maneras. Por ejemplo, un procedimiento para llevar a cabo dicho ensayo implicaría anclar la proteína de la subunidad reguladora p101, polipéptido, péptido o proteína de fusión o la sustancia a ser ensayada sobre una fase sólida y detectar los complejos subunidad reguladora p101/compuesto de prueba fijados sobre la fase sólida al final de la reacción. En una realización de dicho método, la subunidad reguladora p101 reactiva se puede fijar a una superficie sólida y el compuesto prueba que no es fijado, puede ser marcado ya sea de manera directa o indirecta. En otra realización del procedimiento, una proteína de la subunidad reguladora p101 fijada a la fase sólida se compleja con la subunidad catalítica tal como p120. A continuación, el compuesto prueba se podría detectar por su capacidad a interrumpir la asociación del complejo p101/p120.

En la práctica, las placas de microtitulación se pueden utilizar convenientemente como la fase sólida. El componente anclado se puede inmovilizar mediante uniones no covalentes o covalentes. La unión no covalente se puede llevar a cabo mediante el simple recubrimiento de la superficie sólida con una solución de la proteína y secado. De manera alternativa, un anticuerpo inmovilizado, preferiblemente un anticuerpo monoclonal, específico de la proteína a ser inmovilizada se puede utilizar para anclar la proteína a la superficie sólida. Las superficies se pueden preparar por avanzado y guardarlas.

Con el fin de llevar a cabo el ensayo, el componente no inmovilizado se añade a la superficie recubierta que contiene el componente anclado. Después de que se complete la reacción, los componentes que no reaccionan se eliminan (p.e. mediante lavado) en condiciones tales que cualquiera de los complejos formados permanecerá inmovilizado en la superficie sólida. La detección de los complejos anclados en la superficie sólida se puede llevar a cabo de diferentes maneras. Cuando el componente previamente no inmovilizado se pre-marca, la detección del marcaje inmovilizado sobre la superficie indica que se forman los complejos; se puede utilizar un marcaje indirecto para detectar complejos anclados sobre la superficie, p.e. utilizando un anticuerpo específico marcado con un anticuerpo específico para el componente previamente no inmovilizado (el anticuerpo, a su vez, se puede marcar directamente o indirectamente con un anticuerpo anti-Ig marcado).

Alternativamente, se puede llevar a cabo una reacción en una fase líquida, los productos de reacción se separan de los componentes que no han reaccionado y se detectan los complejos, p.e. utilizando un anticuerpo inmovilizado específico de la proteína de la subunidad reguladora p101, polipéptido, péptido o proteína de fusión, o la proteína o proteína de fusión de la subunidad catalítica o el compuesto prueba para anclarlo a cualquiera de los complejos formados en solución, y un anticuerpo específico marcado para el otro componente del posible complejo con el fin de detectar complejos anclados.

5.5.2. Ensayos para proteínas intracelulares que interactúan con las proteínas p101 o p120

Se puede utilizar cualquier procedimiento adecuado para detectar interacciones proteína-proteína para identificar proteínas intracelulares que interaccionen con la subunidad reguladora p101 y/o la subunidad catalítica p120. De entre los procedimientos tradicionales que se pueden emplear están la co-inmunoprecipitación, reticulación y co-purificación a través de columnas de gradiente o cromatográficas de lisados celulares o proteínas obtenidas de lisados celulares y la subunidad reguladora p101 para identificar proteínas en el lisado que interaccionan con la subunidad reguladora p101. Para estos ensayos, el componente de la subunidad reguladora p101 utilizado puede ser una subunidad reguladora p101 de longitud completa, o un péptido truncado. De manera similar, el componente puede ser una subunidad catalítica p120 o un complejo de la subunidad reguladora p101 con la subunidad catalítica p120. Una vez aislada, dicha proteína intracelular se puede identificar y a su vez se puede usar en técnicas estándar para identificar proteínas con las que interactúe. Por ejemplo, se puede averiguar al menos una porción de la secuencia de aminoácidos de una proteína intracelular que interaccione con la subunidad reguladora p101, PI3K (complejo p101/p120) o la subunidad catalítica p120 utilizando técnicas bien conocidas para aquellos expertos en la materia, tales como mediante técnicas de degradación de Edman (véase, p.e., Creighton, 1983, "Proteins: Structures and Molecular Principles", W. H. Freeman & Co., N. Y. p. 34-49). La secuencia de aminoácidos obtenida se puede utilizar como guía para la generación de mezclas de oligonucleótidos que se pueden utilizar para la identificación de secuencias de genes que codifican dichas proteínas intracelulares. Se puede llevar a cabo el cribado, por ejemplo, mediante técnicas de hibridación estándar o PCR. Las técnicas para la generación de mezclas de oligonucleótidos y el cribado son bien conocidos (véase, p.e., Ausubel, *supra* y PCR protocols: A Guide to Methods and Applications, 1990, Innis, M. *et al.*, eds. Academia Press, Inc., Nueva York).

Adicionalmente, se pueden utilizar procedimientos que resultan en la identificación simultánea de genes que codifican las proteínas intracelulares que interaccionan con la subunidad reguladora p101 y/o la subunidad catalítica p120 y/o la PI3K. Estos procedimientos incluyen, por ejemplo, el sondeo de librerías de expresión, de manera similar a la técnica bien conocida del sondeo con anticuerpo de librerías λ 11, utilizando la proteína de la subunidad reguladora p101 marcada o un polipéptido, péptido o proteína de fusión de la subunidad reguladora p101, p.e., un polipéptido de la subunidad reguladora p101 o un dominio de la subunidad reguladora p101 fusionado a un marcador (p.e., un enzima, flúor, proteína luminiscente o colorante) o un dominio de Ig-Fc.

Se describe en detalle un procedimiento que detecta interacciones de proteínas *in vivo*, el sistema de doble híbrido, sólo a modo ilustrativo y no como limitativo. Una versión de este sistema ha sido descrita (Chien *et al.*, 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582) y está comercialmente disponible de Clontech (Palo Alto, CA).

Brevemente, utilizando dicho sistema, se construyen plásmidos que codifican dos proteínas híbridas: un plásmido que consiste en nucleótidos que codifican el dominio de unión a ADN de una proteína activadora de la transcripción

fusionada a una secuencia nucleotídica de p101 que codifica la subunidad reguladora p101, un polipéptido, péptido o proteína de fusión de la subunidad reguladora p101 y el otro plásmido consiste en nucleótidos que codifican el dominio de activación de la proteína activadora de la transcripción fusionado a un ADNc que codifica una proteína desconocida que ha sido recombinada en un plásmido como parte de una librería de ADNc. El plásmido de fusión del dominio de unión a ADN y la librería de ADNc se transforman en una cepa de levadura de *Saccharomyces cerevisiae* que contiene un gen reportero (P.e. HBS o lacZ) cuya región reguladora contiene el sitio de unión del activador de la transcripción. La proteína híbrida sola no puede activar la transcripción del gen reportero; el híbrido del dominio de unión a ADN no puede debido a que no proporciona una función de activación y el híbrido del dominio de activación no puede debido a que no puede localizar los sitios de unión del activador. La interacción de las dos proteínas híbridas reconstituye la proteína activadora funcional y resulta en la expresión del gen reportero que se detecta mediante un ensayo del producto del gen reportero.

El sistema de doble híbrido o la metodología relacionada se puede utilizar para identificar en librerías del dominio de activación para proteínas que interaccionen con el producto del gen de “cebo”. A modo de ejemplo, y no de limitación, se puede utilizar la subunidad reguladora p101 como producto del gen de cebo. Las secuencias genómicas totales o de ADNc se fusionan al ADN que codifica un dominio de activación. Se cotransforma esta librería y un plásmido que codifique un híbrido de un producto genético de cebo fusionado al dominio de unión a ADN, en una cepa de levadura reportera, y los transformantes resultantes se identifican por expresar el gen reportero. Por ejemplo, y no a modo limitativo, una secuencia del gen de p101 de cebo, tal como el marco de lectura abierto de p101 (o un dominio de p101), según se describe en Fig. 1 se puede clonar en un vector de manera que se fusione traduccionalmente al ADN codificante del dominio de unión a ADN de la proteína GALA. Estas colonias se purifican y se aíslan los plásmidos responsables de la expresión del gen reportero. La secuenciación de ADN se utiliza, a continuación, para identificar las proteínas codificadas por los plásmidos de las librerías.

Se puede preparar una librería de ADNc de la línea celular de la que se quiera detectar las proteínas que interaccionan con el producto del gen de p101 utilizando procedimientos practicados de manera rutinaria en el estado de la técnica. De acuerdo con el sistema concreto descrito en el presente documento, por ejemplo, se pueden insertar los fragmentos de ADNc en un vector de manera que se fusionen traduccionalmente al dominio de activación transcripcional de GAL4. Esta librería se puede co-transfectar con el plásmido de fusión de GAL-4-gen de p101 de cebo en una cepa de levadura que contiene un gen lacZ dirigido por un promotor que contiene la secuencia de activación de GAL4. Una proteína codificada por ADNc, fusionada al dominio de activación transcripcional de GAL4, que interacciona con el producto del gen de p101 de cebo reconstituirá una proteína GAL4 activa y, de esta manera, dirige la expresión del gen HIS3. Las colonias que expresen HIS3 se pueden detectar por su crecimiento en placas de Petri que contengan un medio basado en agar semisólido carente de histidina. El ADNc se puede purificar a continuación a partir de estas cepas y ser utilizado para producir y aislar la proteína que interacciona con el gen de p101 de cebo utilizando técnicas practicadas de manera rutinaria en el estado de la técnica.

5.5.3. Ensayos para compuestos que interfieren en la interacción de la subunidad reguladora p101/macromolécula intracelular

Las macromoléculas que interaccionan con la subunidad reguladora p101 son referidas a, para los propósitos de esta discusión, como “parejas de unión”. Estas parejas de unión están implicadas probablemente en la ruta de transducción de señal de la subunidad reguladora p101 y, por lo tanto, en la función de la subunidad reguladora p101 en la regulación de la activación de células de linaje hematopoyético. Parejas de unión conocidas son las subunidades catalíticas de la quinasas PI3K tal como p120, p117 y quizás algunas proteínas p110. Otras parejas de unión están activadas probablemente por proteínas G triméricas tales como las subunidades $G\beta\gamma$ y/o lípidos. Por lo tanto, es deseable identificar compuestos que interfieran en o interrumpan la interacción de dichas parejas de unión con p101 lo cual puede ser útil en la regulación de la actividad de la subunidad reguladora p101 y así controlar los trastornos en la activación de células del linaje hematopoyético asociado con la actividad de la subunidad reguladora p101.

El principio básico de los sistemas de ensayo utilizados para identificar compuestos que interfieren en la interacción entre la subunidad reguladora p101 y su pareja o parejas de unión implica la preparación de una mezcla de reacción que contiene la proteína, polipéptido, péptido o proteína de fusión de la subunidad reguladora p101, según se ha descrito en la sección 5.5.1 y 5.5.2 anteriores y la pareja de unión en condiciones y durante un tiempo suficiente para dejar a los dos que interaccionen y se unan, formando de esta manera un complejo. Con el fin de probar la actividad inhibidora de un compuesto, se prepara la mezcla de reacción en presencia y ausencia del compuesto de prueba. El compuesto de prueba se puede incluir inicialmente en la mezcla de reacción o se puede añadir en un tiempo subsiguiente a la adición del fragmento de la subunidad reguladora p101 y su pareja de unión. Las mezclas de reacción control se incuban sin el compuesto prueba o con un placebo. La formación de cualquier complejo entre el fragmento de la subunidad reguladora p101 y la pareja de unión se detecta a continuación. La formación de un complejo en la reacción control, pero no en la mezcla de reacción que contiene el compuesto de prueba, indica que el compuesto interfiere en la interacción de la subunidad reguladora p101 y la pareja de unión interactiva. Adicionalmente, la formación del complejo en mezclas de reacción que contienen el compuesto prueba y una proteína de la subunidad reguladora p101 también se puede comparar con la formación de complejos en las mezclas de reacción que contienen el compuesto prueba y una subunidad reguladora p101 mutante. Esta comparación puede ser importante en aquellos casos en los que sea deseable identificar compuestos que interrumpan las interacciones de las subunidades reguladoras mutantes pero no las de las subunidades normales.

El ensayo de compuestos que interfieren con la interacción de la subunidad reguladora p101 y las parejas de unión se puede llevar a cabo en un formato heterogéneo u homogéneo. Los ensayos heterogéneos implican el anclaje del producto del fragmento de la subunidad reguladora p101 o de la pareja de unión en una fase sólida y detectar los complejos anclados sobre la fase sólida al final de la reacción. En los ensayos homogéneos, la reacción entera se lleva a cabo en fase líquida. En cualquiera de las aproximaciones, el orden de adición de los reactivos puede variar para obtener diferente información sobre los compuestos a ser ensayados. Por ejemplo, los compuestos prueba que interfieren con la interacción por competición se pueden identificar llevando a cabo la reacción en presencia de una sustancia de prueba; i.e., añadiendo la sustancia de prueba a la mezcla de reacción antes de o simultáneamente con el fragmento de la subunidad reguladora p101 e interaccionar con la pareja de unión. Alternativamente, los compuestos de prueba que interrumpen complejos preformados, p.e., compuestos con constantes de unión elevadas que desplacen uno de los componentes del complejo, se puede probar añadiendo el compuesto prueba a la mezcla de reacción después de que se hayan formado los complejos.

En un sistema de ensayo heterogéneo, ya sea el fragmento de la subunidad reguladora p101 o la pareja de unión interactiva, se ancla en una superficie sólida mientras que la especie no anclada se marca, ya sea directa o indirectamente. En la práctica, se utilizan convenientemente placas de microtitulación. La especie anclada se puede inmovilizar mediante uniones covalentes o no covalentes. La unión no covalente se puede llevar a cabo simplemente recubriendo la superficie sólida con una solución del producto del gen p101 o p120 o la pareja de unión y el secado. De manera alternativa, se puede utilizar un anticuerpo inmovilizado específico para la especie a ser anclada para fijar la especie a la superficie sólida. Las superficies se pueden preparar por avanzado y guardarlas.

Con el fin de llevar a cabo el ensayo, la pareja de la especie inmovilizada se expone a la superficie recubierta con o sin el compuesto de prueba. Después de que se complete la reacción, los componentes que no han reaccionado se eliminan (p.e. mediante lavado) y los complejos formados permanecerán inmovilizados en la superficie sólida. La detección de los complejos anclados sobre la superficie sólida se puede llevar a cabo de diferentes maneras. Cuando la especie no inmovilizada se pre-marca. La detección del marcaje inmovilizado en la superficie indica que se formaron los complejos. Cuando la especie no inmovilizada no está pre-marcada se puede utilizar un marcaje indirecto para detectar complejos anclados a la superficie; p.e., utilizando un anticuerpo marcado específico de especies inicialmente no inmovilizadas (el anticuerpo, a su vez, se puede marcar directamente o indirectamente con un anticuerpo anti-Ig marcado). Dependiendo del orden de adición de los componentes de la reacción, se pueden detectar los compuestos en estudio que inhiban la formación del complejo o que interrumpen los complejos preformados.

Alternativamente, la reacción se puede llevar a cabo en una fase líquida en presencia o ausencia del compuesto prueba, los productos de reacción ser separados de los componentes sin reaccionar y detectar los complejos; p.e., utilizando un anticuerpo específico inmovilizado para uno de los componentes de unión para anclar cualquiera de los complejos formados en solución y un anticuerpo marcado específico de la otra pareja con el fin de detectar los complejos anclados. Otra vez, dependiendo del orden de adición de los reactivos a la fase líquida, se pueden identificar los compuestos prueba que inhiban el complejo o que interrumpen los complejos preformados.

En una realización alternativa de la invención, se puede utilizar un ensayo homogéneo. En esta aproximación, se prepara un complejo preformado del fragmento de la subunidad reguladora p101 y la pareja de unión interactiva, en el que la subunidad reguladora o su pareja de unión se encuentran marcadas pero la señal generada por el marcaje está apagada debido a la formación del complejo (véase, p.e., la patente US N° 4.109.496 de Rubenstein que utiliza esta aproximación para inmunoensayos). La adición de una sustancia prueba que compita con y desplaza una de las especies del complejo preformado resultará en la generación de una señal por encima del ruido de fondo. De esta manera, se pueden identificar las sustancias prueba que interrumpen la interacción subunidad reguladora p101/pareja de unión intracelular.

En una realización particular, se puede preparar una fusión de la subunidad reguladora p101 para la inmovilización. Por ejemplo, la subunidad reguladora p101 o un fragmento peptídico, p.e., correspondiente a CD, se puede fusionar a un gen de la glutatión-S-transferasa (GST) utilizando un vector de fusión, tal como pGEX-5X-1, de manera que su actividad de unión se mantenga en la proteína de fusión resultante. La pareja de unión interactiva se puede purificar y utilizar para generar un anticuerpo monoclonal, utilizando procedimientos practicados de manera rutinaria en la materia y descritos más arriba, en la Sección 5.3. Este anticuerpo se puede marcar con el isótopo radioactivo I¹²⁵, por ejemplo, mediante procedimientos practicados de manera rutinaria en la materia. En un ensayo heterogéneo, p.e., la proteína de fusión de GST-subunidad reguladora p101 se puede anclar a perlas de glutatión-agarosa. La pareja de unión interactiva se puede añadir, a continuación, en presencia o ausencia del compuesto de prueba de manera que deje que suceda la interacción y unión. Al final del periodo de reacción, se lava el material no unido y el anticuerpo monoclonal marcado se pueden añadir al sistema y dejar que se una a los componentes complejados. La interacción entre el producto del gen p101 o p120 y la pareja de unión interactiva se puede detectar midiendo la cantidad de radioactividad que permanece asociada a las partículas de glutatión-agarosa. Una inhibición exitosa de la interacción del compuesto de ensayo resultará en un descenso de la radioactividad.

Alternativamente, la proteína de fusión de GST-subunidad reguladora p101 y la pareja de unión interactiva se pueden mezclar juntos en un líquido en ausencia de las perlas de glutatión-agarosa sólida. El compuesto prueba se puede añadir bien durante o después de que se dejen interaccionar las especies. La mezcla se puede añadir a continuación a las perlas de glutatión-agarosa y el material no unido se lava. Otra vez, la extensión de la inhibición

de la interacción subunidad reguladora p101/pareja de unión se puede detectar añadiendo el anticuerpo marcado y midiendo la radioactividad asociada a las perlas.

En otra realización de la invención, estas mismas técnicas se puede utilizar usando fragmentos peptídicos que corresponden a los dominios de unión de la subunidad reguladora p101 y/o a la pareja interactiva o de unión (en casos en los que la pareja de unión sea una proteína), en lugar de una o las dos proteínas de longitud completa. Se pueden utilizar una serie de métodos practicados de manera rutinaria en la materia para identificar y aislar los sitios de unión. Estos procedimientos incluyen, pero no se limitan a, mutagénesis del gen codificante de una de las proteínas y la identificación por la interrupción de la unión en un ensayo de co-inmunoprecipitación. Se pueden detectar, a continuación, las mutaciones de compensación en el gen codificante de la segunda especie en el complejo. El análisis de secuencia en los genes que codifican las proteínas respectivas revelará las mutaciones que corresponden a la región de la proteína implicada en la unión interactiva. Alternativamente, se puede anclar una proteína a una superficie sólida utilizando procedimientos descritos más arriba y dejar que interaccionen con y se unan a su pareja de unión marcada la cual ha sido tratada con un enzima proteolítico, como tripsina. Después del lavado, un péptido marcado corto que comprende el dominio de unión puede permanecer asociado con el material sólido el cual ha sido aislado e identificado mediante secuenciación de aminoácidos. También, una vez se obtiene el gen codificante de una pareja de unión intracelular, se pueden manipular genéticamente segmentos del gen para que expresen fragmentos peptídicos de la proteína las cuales se pueden probar por su actividad de unión y purificar o sintetizar.

Por ejemplo, y no como limitación, un producto del gen p101 se puede anclar a un material sólido según se ha descrito más arriba, preparando una proteína de fusión de GST-subunidad reguladora p101 y dejando que se una a perlas de glutatión-agarosa. La pareja de unión interactiva se puede marcar con un isótopo radioactivo, tal como ^{35}S y escindir con un enzima proteolítico tal como tripsina. Los productos de escisión se pueden añadir, a continuación, a la proteína de fusión GST-p101 y dejar que se unan. Después de lavar los péptidos no unido, se puede eluir, purificar y analizar la secuencia de aminoácidos del material unido marcado que representa el dominio de unión de la pareja de unión intracelular mediante procedimientos bien conocidos. Los péptidos así identificados se pueden producir sintéticamente o fusionar a proteínas facilitadoras adecuadas utilizando la tecnología del ADN recombinante.

5.5.4 Ensayos para la determinación de compuestos que mejoran enfermedades inflamatorias

Los compuestos, incluyendo pero no limitando a los compuestos de unión identificados mediante técnicas de ensayo tales como las descritas más arriba, en las secciones 5.5.1 a 5.5.3, se pueden determinar por su capacidad para mejorar los síntomas del trastorno del sistema inmune, incluyendo la inflamación. Los ensayos descritos más arriba pueden identificar compuestos que afectan la actividad de la subunidad reguladora p101 (p.e. compuestos que se unen a la subunidad reguladora p101, inhiben la unión de los ligandos naturales y o bien activan la transducción de señal (agonistas) o bloquean la activación (antagonistas), y los compuestos que se unen a un ligando natural de la subunidad reguladora p101 y neutraliza la actividad del ligando); o compuestos que afectan la actividad del gen p101 o p120 (afectando la expresión del gen de p101 o p120, incluyendo moléculas, p.e., proteínas o moléculas orgánicas pequeñas que afectan o interfieren en acontecimientos de empalme de manera que se regula la expresión de la subunidad reguladora p101 de longitud completa o la forma truncada). Sin embargo, se debería indicar que los ensayos descritos en el presente documento también pueden identificar compuestos que modulan la transducción de señal de la subunidad reguladora p101 (p.e. compuestos que afectan a los acontecimientos de señalización "downstream", tales como inhibidores o potenciadores de actividades que participan en la transducción de la señal de $\text{PtdIns}(4,5)\text{P}_3$ que se genera mediante unión de la subunidad catalítica a la subunidad reguladora p101). La identificación y uso de dichos compuestos que afectan a otra etapa en la ruta de transducción de señal de la subunidad reguladora p101 en la que está implicado el gen p101 o p120 y/o el producto del gen p101 o p120 y que afecta la misma ruta puede modular el efecto de la subunidad reguladora p101 en el desarrollo de los trastornos de activación de células del linaje hematopoyético están dentro del ámbito de la invención. Dichos compuestos se pueden utilizar como parte de un método terapéutico para el tratamiento de trastornos de la activación de células del linaje hematopoyético.

La invención engloba los ensayos basados en células y basados en modelos animales para la identificación de compuestos que exhiben dicha capacidad de mejorar los síntomas de los trastornos de activación de células del linaje hematopoyético. Dichos sistemas de ensayo basados en células se pueden utilizar también como el estándar para determinar la pureza y potencia del ligando natural, subunidad catalítica, incluyendo la subunidad catalítica producida de manera recombinante o sintética o los mutantes de la subunidad catalítica.

Los sistemas basados en células se pueden usar para identificar compuestos que pueden actuar para mejorar los síntomas de trastornos de la activación de células del linaje hematopoyético. Dichos sistemas celulares pueden incluir, por ejemplo, células recombinantes o no recombinantes, tales como líneas celulares, que expresan el gen p101 y/o p120. Por ejemplo, se pueden utilizar células de leucocito o líneas celulares derivadas de células leucocitarias. Además, la expresión en células hospedadoras (p.e. células COS, células CHO, fibroblastos, células Sf9) modificadas genéticamente para que expresen un p101/p120 de PI3K funcional y respondan a la activación mediante subunidades $G\beta\gamma$ de ligando natural, p.e., se puede tomar como punto final del ensayo cuando se detecta un cambio químico o fenotípico, inducción de otro gen en la célula hospedadora, cambio en los niveles de mensajero intracelular (p.e. $\text{PtdIns}(3,4,5)\text{P}_3$, etc).

En la utilización de dichos sistemas celulares, las células se pueden exponer a un compuesto sospechoso de exhibir la capacidad de mejorar los síntomas de trastornos de activación de células del linaje hematopoyético, a una con-

centración suficiente y durante un tiempo suficiente como para estimular dicha mejora de los síntomas del trastorno de activación de células del linaje hematopoyético en las células expuestas. Después de la exposición, las células se pueden estudiar midiendo las alteraciones en la expresión del gen p101 o p120, p.e. determinando en lisados celulares los transcritos de ARNm de p101 o p120 (p.e., mediante análisis Northern) o la proteína de p101 o de p120 expresada en la célula; los compuestos que regulen o modulen la expresión del gen p101 o p120 son valiosos candidatos como agentes terapéuticos. Alternativamente, se examinan las células para determinar si uno o más de los fenotipos celulares similares a trastornos en la activación de células del linaje hematopoyético se ha alterado para parecerse a un fenotipo más natural o más normal o a un fenotipo más probable de producir una incidencia o severidad más baja de síntomas del trastorno. Todavía adicionalmente, se puede determinar la expresión y/o actividad de componentes de la ruta de transducción de la señal de la que la subunidad reguladora p101 es una parte o la actividad de la ruta de transducción de señal de la subunidad p101 en sí misma.

Por ejemplo, después de la exposición de las células, se puede determinar en los lisados celulares la presencia de niveles incrementados del segundo mensajero PtdIns(3,4,5)P₃, comparado con lisados derivados de células control no expuestas. La capacidad de un compuesto prueba de inhibir la producción de un segundo mensajero en estos sistemas de ensayo indica que el compuesto prueba inhibe la transducción de señal iniciada por la activación de la subunidad reguladora p101. Los lisados celulares se pueden ensayar fácilmente utilizando HPLC de intercambio aniónico. Alternativamente, los niveles de producción de superóxido u O₂ se pueden determinar mediante seguimiento de quimioluminiscencia de la oxidación del luminol catalizada por peroxidasa de rábano según se describe en Gimán *et al.*, 1987, Anal. Biochem. 165: 371-378. Finalmente, se puede determinar un cambio en la adhesión celular de células intactas usando técnicas bien conocidas para aquellos expertos en la materia.

Además, los sistemas de trastorno de activación de células del linaje hematopoyético basado en animales, los cuales pueden incluir, por ejemplo ratones, se pueden utilizar para identificar compuestos capaces de mejorar los síntomas relacionados con trastornos de activación de células del linaje hematopoyético. Dichos modelos animales se pueden utilizar como sistemas de ensayo para la identificación de drogas, productos farmacéuticos, terapias e intervenciones que pueden ser efectivas en el tratamiento de dichos trastornos. Por ejemplo, los modelos animales se pueden exponer a un compuesto, sospechoso de exhibir una capacidad de mejorar los síntomas del trastorno de activación de células del linaje hematopoyético, a una concentración suficiente y durante un tiempo suficiente como para provocar dicha mejora de los síntomas del trastorno de activación de células del linaje hematopoyético en los animales expuestos. La respuesta de los animales a la exposición se puede seguir determinando la inversión de trastornos asociados a trastornos por activación de células del linaje hematopoyético, tal como la inflamación. Con respecto a la intervención, cualquiera de los tratamientos que inviertan cualquier aspecto de los síntomas relacionados con un trastorno de la activación de células del linaje hematopoyético se debería de considerar como candidato para la intervención terapéutica de trastornos por activación de células del linaje hematopoyético. Las dosis de los agentes de prueba se pueden determinar derivando las curvas dosis-respuesta, según se discute más abajo.

5.6 El tratamiento de enfermedades asociadas con la estimulación de PI3K activada por proteína, incluyendo enfermedades inflamatorias

La invención también engloba procedimientos y composiciones para la modificación de la activación de células del linaje hematopoyético y el tratamiento de trastornos de la activación de células del linaje hematopoyético, incluyendo los trastornos inflamatorios. Por ejemplo, disminuyendo el nivel de expresión del gen p101 y/o la actividad del gen de la subunidad reguladora p101 y/o regulando negativamente la actividad de la ruta de la subunidad reguladora p101 (p.e. interfiriendo en la interacción de la subunidad reguladora p101 con la subunidad catalítica p120 o dirigiendo acontecimientos de señalización “downstream”), la respuesta de células leucocitarias a factores que activan receptores asociados a proteínas G triméricas, tales como citoquinas, se pueden reducir y los síntomas de enfermedades inflamatorias crónicas mejorar. A la inversa, la respuesta de células de leucocito a los receptores asociados a proteína G se pueden incrementar aumentando la actividad de la subunidad reguladora p101. Por ejemplo, dicho aumento puede servir para estimular la respuesta del sistema inmune a infecciones. Se discuten más abajo diferentes aproximaciones.

5.6.1 Inhibición de la expresión del adaptador p101 o actividad del adaptador p101 para reducir la actividad de pi3k activada por proteína G y reducir la inflamación

Se puede utilizar cualquier procedimiento que neutralice la subunidad catalítica o inhiba la expresión del gen p101 o p120 (ya sea transcripción o traducción) para reducir la respuesta inflamatoria. Dichas aproximaciones se pueden utilizar para tratar trastornos de respuesta inflamatoria tales como la artritis, incluyendo la artritis reumatoide, el choque séptico, síndrome de dificultad respiratoria aguda (ARDS), neumonía, asma y otros estados del pulmón, alergias, lesión de reperusión, aterosclerosis y otros trastornos cardiovasculares, enfermedad de Alzheimer y cáncer, por nombrar sólo unos pocos trastornos inflamatorios.

En una realización, la terapia inmune se puede diseñar para reducir el nivel de expresión del gen endógeno p101 o p120, p.e., utilizando aproximaciones antisentido o ribozimas para inhibir o prevenir la traducción de los transcritos de ARNm de p101 o p120; aproximaciones de triple hélice para inhibir la transcripción del gen p101 o p120; o recombinación homóloga dirigida para inactivar o “knock out” el gen de p101 o p120 o su promotor endógeno.

Las aproximaciones antisentido implican el diseño de oligonucleótidos (ya sea ADN o ARN) que son complementarios al ARNm de la subunidad reguladora p101 o p120. Los oligonucleótidos antisentido se unirán a los transcritos

de ARNm de p101 o p120 y evitarán la traducción. La complementariedad absoluta, aunque preferida, no es requerida. Una secuencia "complementaria" a una porción de ARN, tal y como se refiere en el presente documento, significa una secuencia que tiene suficiente complementariedad para ser capaz de hibridar con el ARN, formando un dúplex estable. En el caso de ácidos nucleicos de doble cadena antisentido, se puede analizar, de esta manera, una única cadena dúplex de ADN o se puede ensayar la formación del triplex. Generalmente, cuanto más largo sea el ácido nucleico que hibrida, más desemparejamientos de bases pueden haber en el ARN y todavía formar un dúplex estable (o triplex, según pueda ser el caso). Un experto en la materia puede averiguar un grado de tolerancia de desemparejamiento usando procedimientos estándar para determinar el punto de fusión del complejo hibridado.

Los oligonucleótidos que son complementarios al extremo 5' del mensajero, p.e., la secuencia 5' no traducida hasta e incluir el codón de inicio AUG, deberían de funcionar con más eficacia en la inhibición de la traducción. Sin embargo, las secuencias complementarias a las secuencias en 3' no traducidas de los ARNm han mostrado recientemente que también son efectivas en la inhibición de la traducción de ARNm. Véase, en general Wagner, R., 1994, Nature 372:333-335. De esta manera, los oligonucleótidos complementarios a regiones no codificantes no traducidas en 5' o 3' de p101 o p120 mostradas en la Fig. 1 y Fig. 3 se podrían utilizar en la aproximación antisentido para inhibir la traducción del ARNm endógeno de p101 o p120. La complementariedad de oligonucleótidos a la región 5' no traducida del ARNm debería incluir el complementario al codón de inicio AUG. Los oligonucleótidos antisentido complementarios a las regiones de ARNm codificantes son inhibidores de la traducción menos eficientes pero se podrían utilizar de acuerdo con la invención. Si se diseña para que hibride con 5' o 3' o la región codificante del ARNm de la subunidad reguladora p101, los ácidos nucleicos antisentido deberían ser al menos de 6 nucleótidos de longitud y son preferiblemente oligonucleótidos que oscilan entre 6 y aproximadamente 50 nucleótidos de longitud. En aspectos específicos el oligonucleótido es al menos de 10 nucleótidos, al menos de 17 nucleótidos, al menos de 25 nucleótidos o al menos de 50 nucleótidos.

Sin tener en cuenta la elección de la secuencia diana, se prefiere que se lleven a cabo primero estudios *in vitro* para cuantificar la capacidad del oligonucleótido antisentido para inhibir la expresión del gen. Se prefiere que estos estudios utilicen controles que distingan entre la inhibición del gen antisentido y los efectos biológicos no específicos de los oligonucleótidos. También se prefiere que estos estudios comparen niveles del ARN o proteína diana con la de un ARN o proteína de control interno. Adicionalmente, se tienen en cuenta que los resultados obtenidos utilizando el oligonucleótido antisentido se comparan con aquellos obtenidos utilizando un oligonucleótido control. Se prefiere que el oligonucleótido control sea de aproximadamente la misma longitud que el oligonucleótido en estudio y que la secuencia de nucleótidos del oligonucleótido difiera de la secuencia antisentido en no más de lo necesario para evitar la hibridación específica de la secuencia diana.

Los oligonucleótidos pueden ser de ADN o ARN o mezclas químicas o derivados o versiones modificadas de los mismos, de cadena única o cadena doble. El oligonucleótido se puede modificar en la unidad de la base, unidad de azúcar o esqueleto fosfato, por ejemplo, para mejorar la estabilidad de la molécula, la hibridación, etc. El oligonucleótido puede incluir otros grupos añadidos tales como péptidos (véase, p.e., para dirigir los receptores de células hospedadoras *in vivo*) o agentes que facilitan el transporte a través de la membrana celular (véase, p.e., Letsinger *et al.*, 1989, Proc. Natl. Acad. Sci. U.S.A., 86:6553-6556; Lemaitre *et al.*, 1987, Proc. Natl. Acad. Sci. 84:648-652; publicación de PCT N°. W088/09810, publicada el 15 de diciembre de 1988) o agentes de escisión provocados por hibridación. (Véase, p.e., Krol *et al.*, 1988, BioTechniques 6:958-976) o agentes intercalantes (véase, p.e., Zon., 1988, Pharma Res. 5:539-549). En este punto, el oligonucleótido se puede conjugar a otra molécula, p.e., a un péptido, agente de reticulación provocado por hibridación, agente de transporte, agente de escisión provocado por hibridación, etc.

El oligonucleótido antisentido puede comprender al menos una unidad base modificada que se selecciona del grupo que incluye pero no se limita a 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroximetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta D-manosilqueosina, 5'-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, metiléster del ácido uracil-5-oxiacético, ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w, y 2,6-diaminopurina.

El oligonucleótido antisentido puede comprender también al menos una unidad de azúcar modificada seleccionada del grupo que incluye pero no se limita a arabinosa, 2-fluoroarabinosa, xilulosa y hexosa.

En otra realización, el oligonucleótido antisentido comprende al menos un esqueleto de fosfato modificado del grupo que consiste en fosforotioato, fosforoditioato, fosforamidotioato, fosforamidoato, fósforodiamidato, metilfosfonato, alquilfosfotriéster y formacetal o análogo de los mismos.

En todavía otra realización, el oligonucleótido antisentido es un oligonucleótido α -anomérico. Un oligonucleótido α -anomérico forma híbridos de cadena doble con el ARN complementario, en el cual, al contrario a las unidades β habituales, las cadenas se organizan paralelas unas con las otras (Gautier *et al.*, 1987, Nucl. Acids Res. 15:6625-6641). El oligonucleótido es un 2'-O-metilribonucleótido (Inoue *et al.*, 1987, Nucl. Acids Res. 15:6131-6135) o un análogo químico de ADN/ARN (Inoue *et al.*, 1987, FEBS Lett. 215:327-330).

Los oligonucleótidos de la invención se pueden sintetizar mediante procedimientos estándar conocidos en la materia, p.e., utilizando un sintetizador de ADN automático (tales como los comercialmente disponibles de Biosearch, Applied Biosystems, etc). Como ejemplos, los oligonucleótidos de fosforotioato se pueden sintetizar mediante el procedimiento de Stein *et al.*, 1988, Nucl. Acids. Res. 16:3209. Los oligonucleótidos de metilfosfonato se pueden preparar utilizando soportes de polímero de vidrio de poro controlado (Sarin *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85: 7448-7451).

Si bien se deberían de utilizar oligonucleótidos antisentido complementarios a la secuencia de la región codificante de p101 o p120, son más preferidos aquellos complementarios a la región transcrita no traducida.

Las moléculas antisentido se deberían liberar a células que expresan la subunidad reguladora p101 *in vivo*, p.e., células de origen hematopoyético como plaquetas y neutrófilos y otros leucocitos. Se han desarrollado una serie de procedimientos para la liberación de ADN o ARN antisentido a células; p.e., se pueden inyectar moléculas antisentido directamente al sitio de derivación de tejido o célula o se pueden administrar de manera sistemática moléculas antisentido modificadas, diseñadas para dirigir las células deseadas (p.e. antisentido unido a péptidos o anticuerpos que se unen específicamente a receptores o antígenos expresados en la superficie de la célula diana).

Sin embargo, a menudo es difícil conseguir concentraciones intracelulares suficientes del antisentido como para suprimir la traducción de los ARNm endógenos. Por lo tanto, una aproximación preferida utiliza una construcción de ADN recombinante en la que el oligonucleótido antisentido se coloca bajo control de un promotor pol II o pol III fuerte. El uso de dicha construcción para transfectar células diana en el paciente resultará en la transcripción de cantidades suficientes de ARN de cadena única que formarán pares de bases complementarias con los transcritos endógenos de p101 o p120 y, de esta manera, se evita la traducción del ARNm de p101 o p120. Por ejemplo, se puede introducir un vector *in vivo* de manera que es asimilado por una célula y dirige la transcripción de un ARN antisentido. Dicho vector puede permanecer integrado en el episoma o llegar a estar integrado cromosómicamente, siempre y cuando se pueda transcribir para producir el ARN antisentido deseado. Dichos vectores se pueden construir mediante procedimientos de tecnología de ADN recombinante estándar en la materia. Los vectores pueden ser plasmídicos, víricos u otros conocidos en la materia, usados para la replicación y expresión en células de mamífero. La expresión de la secuencia codificante del ARN antisentido puede darse mediante cualquier promotor conocido por actuar en células de mamífero, preferiblemente células humanas. Dichos promotores pueden ser inducibles o constitutivos. Dichos promotores incluyen, pero no se limitan a: región del promotor temprano de SV40 (Berthoult y Chambon, 1981, Nature 290:304-310), el promotor contenido en la repetición larga terminal en 3' del virus del Sarcoma de Rous (Yamamoto *et al.*, 1980, Cell 22: 787-797), el promotor de la timidita quinasa del herpes (Wagner *et al.*, 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), las secuencias reguladoras del gen de la metalotioneína (Brinster *et al.*, 1982, Nature 296:39-42), etc. Cualquier tipo de plásmido, cósmido, YAC o vector vírico se puede usar para preparar la construcción de ADN recombinante que se puede introducir directamente en el tejido o sitio de derivación celular, p.e., la médula ósea. Alternativamente, se pueden utilizar vectores virales que infecten de manera selectiva el tejido deseado o tipo celular (p.e. virus que infectan células del linaje hematopoyético) en cuyo caso la administración se puede llevar a cabo por otra ruta (p.e. la sistemática).

Las moléculas de ribozima diseñadas para la escisión catalítica de los transcritos de ARNm de p101 o p120 se puede usar también para evitar la traducción del ARNm de p101 o p120 y la expresión de la subunidad reguladora p101. (Véase, p.e., la publicación internacional de la PCT WO90/11364, publicada el 4 de octubre de 1990; Sarver *et al.*, 1990, Science 247:1222-1225). Aunque se pueden utilizar para destruir los ARNm de p101 o p120, los ribozimas cortan el ARNm en secuencias de reconocimiento de lugar, se prefiere el uso de ribozimas de cabeza de martillo. Los ribozimas de cabeza de martillo cortan ARNm en sitios determinados por regiones flanqueantes que forman pares de bases complementarias con el ARNm diana. El único requisito es que el ARN diana tiene la siguiente secuencia de dos bases: 5'-UG-3'. La construcción y producción de ribozimas de cabeza de martillo es bien conocida en la materia y se describe en más detalle en Haseloff y Gerlach, 1988, Nature, 334:585-591. Existen cientos de sitios de corte de ribozimas de cabeza de martillo dentro de la secuencia de nucleótidos del ADNc de p101 o p120 (Fig. 3). Preferiblemente, el ribozima se manipula genéticamente de manera que el sitio de reconocimiento se localiza cerca del extremo 5' del ARNm de p101 o p120; es decir, se incrementa la eficacia y se minimiza la acumulación intracelular de transcritos de ARNm no funcionales.

Los ribozimas de la presente invención también incluyen endoribonucleasas de ARN (de aquí en adelante ribozimas "Tipo Cech") como la que se encuentra de manera natural en Tetrahymena Thermophila (conocida como el ARN de IVS o L-19 IVS) y que ha sido descrita extensamente por Thomas Cech y colaboradores (Zaug *et al.*, 1984, Science, 224:574-578; Zaug y Cech, 1986, Science, 231:470-475; Zaug *et al.*, 1986, Nature, 324:429-433; solicitud de patente internacional publicada con el n° WO 88/04300 por University Patents Inc.; Been y Cech, 1986, Cell, 47:207-216). Los ribozimas de tipo Cech tienen un sitio activo de ocho pares de bases que hibrida con una secuencia de ARN diana, de manera que tenga lugar la escisión del ARN diana. La invención engloba aquellos ribozimas de tipo Cech que dirigen secuencias de sitio activo de ocho pares de bases que están presentes en p101 o p120.

Como en la aproximación antisentido, los ribozimas pueden estar compuestos por oligonucleótidos modificados (p.e., para mejorar la estabilidad, dirección, etc.) y se debería de liberar a las células que expresan la subunidad reguladora p101 *in vivo*, p.e. neutrófilos. Un procedimiento preferido de liberación implica la utilización de una construcción de ADN "que codifica" el ribozima bajo el control de un promotor constitutivo fuerte pol III o pol II, de manera que

las células transfectadas produzcan cantidades suficientes del ribozima para destruir los mensajeros endógenos p101 o p120 e inhibir la traducción. Puesto que las moléculas antisentido diferentes de ribozimas son catalíticas se requiere una baja concentración intracelular para la eficacia.

5 La expresión del gen endógeno p101 o p120 se puede reducir también mediante la inactivación o “knock out” del gen p101 o p120 o su promotor utilizando recombinación homóloga dirigida. (P.e. véase Smithies *et al.*, 1985, Nature 317:230-234; Thomas y Capecchi, 1987, Cell 51:503-512; Thompson *et al.*, 1989, Cell 5:313-321; cada uno de los cuales incorporado por referencia en el presente documento en su totalidad). Por ejemplo, se puede utilizar una subunidad reguladora p101 mutante no funcional (o una secuencia de ADN completamente no relacionada) flanqueada por
10 ADN homólogo al gen endógeno p101 o p120 (ya sea las regiones codificantes o regiones reguladoras del gen p101 o p120), con o sin un marcador seleccionable y/o un marcador seleccionable negativo, para transfectar células que expresen la subunidad reguladora p101 *in vivo*. La inserción de una construcción de ADN, mediante recombinación homóloga dirigida, resulta en la inactivación del gen p101 o p120. Dichas aproximaciones son particularmente adecuadas en el campo de la agricultura donde las modificaciones a células ES (células madre embrionarias) se pueden
15 utilizar para generar un animal descendiente con la subunidad reguladora p101 inactiva (p.e. véase Thomas y Capecchi 1987 y Thompson 1989, *supra*). Sin embargo, esta aproximación se puede adaptar para el uso humano con la condición de que las construcciones de ADN recombinante se administren directamente o se dirijan al sitio requerido *in vivo* usando vectores víricos adecuados.

20 Alternativamente, la expresión endógena del gen p101 o p120 se puede reducir dirigiendo secuencias desoxiribonucleótidas complementarias a la región reguladora del gen p101 o p120 (es decir, el promotor de p101 o p120 y/o intensificadores) para formar estructuras helicoidales triples para evitar la transcripción del gen p101 o p120 en células diana del cuerpo. (Véase de manera general, Helene C. 1991, Anticancer Drug Des., 6(6):569-584; Helene, C. *et al.*, 1992, Ann. N. Y. Acad. Sci., 660:27-36; y Maher, L. J., 1992, Bioassays 14(12):807-815).

25 En todavía otra realización de la invención, la actividad de la subunidad reguladora p101 se puede reducir utilizando una aproximación de “negativo dominante” para interferir en la activación de la proteína G trimérica de PI3K. Con este fin, se pueden utilizar construcciones que codifican las subunidades reguladoras p101 defectivas en aproximaciones de terapia génica para disminuir la actividad de la subunidad reguladora p101 en células diana adecuadas.
30 Por ejemplo, se pueden introducir las secuencias nucleotídicas que dirigen la expresión en células hospedadoras de las subunidades reguladoras p101 en las que se elimina o muta el dominio de interacción $G\beta\gamma$ en células hematopoyéticas (ya sea mediante métodos de terapia génica *in vivo* o *ex vivo* descritos más arriba). De manera alternativa, las secuencias nucleotídicas que codifican sólo un dominio funcional de p101 se podrían usar como inhibidor de interacciones p101/p120 nativas. Alternativamente, la recombinación homóloga dirigida se puede utilizar para introducir
35 dichas deleciones o mutaciones en el gen endógeno p101 o p120 en la médula ósea del sujeto. Las células manipuladas genéticamente expresarán receptores no funcionales (es decir, una subunidad reguladora que es capaz de unirse a la subunidad catalítica, pero incapaz de estimular la actividad catalítica en respuesta a la activación de proteína G). Dichas células manipuladas genéticamente, como, neutrófilos u otras líneas de leucocitos, deberían demostrar una respuesta disminuida por la activación de receptores unidos a proteína G por quimioquinas extracelulares, resultando
40 en una reducción del fenotipo inflamatorio.

5.6.2. Restauración o incremento de la expresión o actividad de la subunidad reguladora p101 para promover la activación del sistema inmune

45 Con respecto a un incremento en el nivel normal de expresión del gen normal p101 o p120 y/o de la actividad del producto del gen de la subunidad reguladora p101, se pueden utilizar secuencias de ácido nucleico p101 o p120 para el tratamiento de trastornos de la activación de células del linaje hematopoyético, incluyendo respuesta inmunes reducidas por quimioquinas. Cuando la causa de la disfunción del sistema inmune es una subunidad reguladora p101 defectiva, se puede administrar el tratamiento, por ejemplo, en forma de terapia de sustitución génica. Específicamente,
50 se puede insertar una o más copias de un gen normal p101 o una porción del gen p101 que dirige la producción de un producto del gen p101 que exhibe función normal, en células adecuadas en un paciente o animal, utilizando vectores que incluyen, pero no se limitan a, adenovirus, virus adeno-asociado, retrovirus y vectores del virus del herpes, junto con otras partículas que introducen ADN en células, tal como liposomas.

55 Debido a que el gen p101 o p120 se expresa en células del linaje hematopoyético, incluyendo los neutrófilos y otros leucocitos, dichas técnicas de terapia de sustitución génica deberían ser capaces de liberar secuencias del gen p101 o p120 a estos tipos celulares en los pacientes. Alternativamente, la recombinación homóloga dirigida se puede utilizar para corregir el gen p101 o p120 endógeno defectivo del tipo celular adecuado; p.e., células de la médula ósea o neutrófilos y/u otros leucocitos. En animales, la recombinación homóloga dirigida se puede utilizar para corregir el defecto en células ES con el fin de generar un descendiente con un rasgo corregido.
60

Finalmente, los compuestos identificados en los ensayos descritos más arriba que estimulan, potencian o modifican la señal transducida por la subunidad reguladora p101 activada, p.e., mediante activación de las proteínas de señalización “*downstream*” en la cascada de la subunidad reguladora p101 y, de esta manera, pasando la subunidad reguladora p101 defectiva se pueden usar para conseguir la estimulación del sistema inmune. La formulación y modo
65 de administración dependerá de las propiedades físico-químicas del compuesto.

5.7 Preparaciones farmacéuticas y métodos de administración

Los compuestos que se determinan por tener un efecto sobre la expresión del gen p101 o p120 o de la actividad de la subunidad reguladora p101 o la interacción de p101 con cualquiera de sus parejas de unión incluye pero no se limita a la subunidad catalítica, se pueden administrar a un paciente a dosis terapéuticamente efectivas para tratar o mejorar los trastornos de la activación de células del linaje hematopoyético, incluyendo trastornos de respuesta inflamatoria como la artritis, incluyendo la artritis reumatoide, el choque séptico, el síndrome de dificultad respiratoria aguda (ARDS), neumonía, asma y otras patologías del pulmón, alergias, lesión por reperusión, aterosclerosis y otras enfermedades cardiovasculares, la enfermedad del Alzheimer y cáncer. Una dosis terapéuticamente efectiva se refiere a la cantidad de compuesto suficiente para que resulte en una mejora de los síntomas de los trastornos de la activación de células del linaje hematopoyético.

5.7.1 Dosis efectiva

La toxicidad y eficacia terapéutica de dichos compuestos se puede determinar mediante procedimientos farmacéuticos estándar en cultivos celulares o animales de experimentación, p.e., para determinar la LD50 (la dosis letal para el 50% de la población) y la ED50 (la dosis terapéuticamente efectiva en el 50% de la población). La relación de dosis entre efectos tóxicos y terapéuticos es el índice terapéutico y se puede expresar como la relación LD₅₀/ED₅₀. Son preferidos los compuestos que exhiben índices terapéuticos mayores. Si bien se pueden utilizar compuestos que exhiben efectos secundarios tóxicos, se debería de tener cuidado con el diseño de un sistema de liberación que dirija dichos compuestos al sitio del tejido afectado con el fin de minimizar el daño potencial en células no infectadas y, de esta manera, reducir los efectos secundarios.

Los datos obtenidos de los ensayos con cultivos celulares y estudios con animales se pueden utilizar en la formulación de un rango de dosis para uso en humanos. La dosis de dichos compuestos oscila preferiblemente en el rango de concentraciones circulantes que incluyen los ED50 con poca o ninguna toxicidad. Las dosis pueden variar en este rango dependiendo de la forma de dosificación empleada y de la ruta de administración utilizada. Para cualquiera de los compuestos usados en el procedimiento de la invención, la dosis terapéuticamente efectiva se puede estimar inicialmente a partir de ensayos con cultivos celulares. Se puede formular una dosis en modelos animales para conseguir un rango de concentración circulante en plasma que incluye el IC50 (es decir, la concentración del compuesto en ensayo que consigue una inhibición media máxima de síntomas) cuando se determina en el cultivo celular. Dicha información se puede utilizar para determinar de manera más precisa, dosis útiles en humanos. Los niveles en plasma se pueden medir, por ejemplo, mediante cromatografía líquida de alta resolución.

5.7.2. Formulaciones y uso

Las composiciones farmacéuticas para usar de acuerdo con la presente invención se puede formular de manera convencional utilizando uno o más portadores o excipientes fisiológicamente aceptables.

De esta manera, los compuestos y sus sales y solvatos fisiológicamente aceptables se pueden formular para la administración mediante inhalación o insuflación (ya sea a través de la boca o la nariz) o administración oral, bucal, parenteral o rectal.

Para la administración oral, las composiciones farmacéuticas pueden tomar la forma de, por ejemplo, tabletas o cápsulas preparadas mediante métodos convencionales con excipientes farmacéuticamente aceptables tales como agentes aglutinantes (p.e., almidón de maíz pregelatinizado, polivinilpirrolidona o hidroxipropilmetilcelulosa); rellenos (p.e. lactosa, celulosa microcristalina o hidrogenofosfato de calcio); lubricantes (p.e. estearato de magnesio, talco o sílice); desintegrantes (p.e. almidón de patata o glicolato de almidón sódico); o agentes humectantes (p.e. laurel sulfato sódico). Las tabletas se pueden recubrir mediante procedimientos bien conocidos en la materia. Las preparaciones líquidas para la administración oral pueden tomar la forma de, por ejemplo, soluciones, jarabes o suspensiones o se pueden presentar como producto seco para constituir con agua u otro vehículo antes de su uso. Dichas preparaciones líquidas se pueden preparar mediante medios convencionales con aditivos farmacéuticamente aceptables tales como agentes de suspensión (p.e. jarabe de sorbitol, derivados de celulosa o grasas comestibles hidrogenadas); agentes emulsionantes (p.e., lecitina o acacia); vehículos no acuosos (p.e. aceite de almendra, ésteres aceitosos, alcohol etílico o aceites vegetales fraccionados); y conservantes (p.e., metil- o propil-p-hidroxibenzoatos o ácido sórbico). Las preparaciones pueden contener también sales tampón, aromatizantes, colorantes y edulcorantes, según se considere apropiado.

Las preparaciones para la administración oral se pueden formular de manera adecuada para generar la liberación controlada del compuesto activo.

Para la administración bucal, las composiciones pueden tomar la forma de tabletas o pastillas formuladas de manera convencional.

Para la administración por inhalación, los compuestos para usar de acuerdo con la presente invención se liberan, de manera conveniente, en forma de una presentación en spray aerosol a partir de envases presurizados o un nebulizador, con el uso de un propulsor adecuado, p.e., diclorodifluorometano, triclorofluorometano, diclorotetrafluoroetano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado, la unidad de dosificación se puede determinar proporcionando una válvula para liberar una cantidad medida. Se pueden formular cápsulas y cartuchos

de, p.e., gelatina para usar en un inhalador o insuflador que contengan una mezcla en polvo del compuesto y una base en polvo adecuada como lactosa o almidón.

Los compuestos se pueden formular para la administración parenteral mediante inyección, p.e., mediante inyección de bolo o infusión continua. Las formulaciones para inyección se pueden presentar en una forma de unidad de dosis, p.e., en ampollas o en recipientes multi-dosis, con un conservante añadido. Las composiciones pueden tomar formas tales como suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos y pueden contener agentes de formulación tales como agentes de suspensión, estabilización y/o dispersión. Alternativamente, el ingrediente activo puede estar en forma de polvo para la constitución con un vehículo adecuado, p.e., agua estéril libre de pirógeno, antes de su uso.

Los compuestos se pueden formular también en composiciones rectales como supositorios o enemas de retención, p.e. conteniendo bases de supositorio convencionales tales como mantequilla de coco y otros glicéridos.

Además de las formulaciones descritas más arriba, los compuestos también se pueden formular como una preparación de tipo "depot". Dichas formulaciones de larga duración se pueden administrar mediante implantación (por ejemplo subcutánea o intracutánea) o mediante inyección intramuscular. De esta manera, por ejemplo, los compuestos se pueden formular con materiales poliméricos o hidrofóbicos adecuados (por ejemplo, como una emulsión en un aceite aceptable) o resinas de intercambio iónico, o como derivados poco solubles, por ejemplo, una sal poco soluble.

Las composiciones pueden ser presentadas, si se desea, en un dispositivo empaquetado o dispensador que puede contener una o más formas de unidad de dosificación que contienen el ingrediente activo. El paquete puede comprender, por ejemplo, una hoja de metal o de plástico, tal como un paquete blister. El paquete o dispositivo dispensador se puede acompañar de instrucciones para la administración.

Los siguientes ejemplos se presentan para ilustrar la presente invención y para ayudar a un experto en la materia en la preparación y uso de los mismos. Los ejemplos no tienen la intención, a no ser que se indique de otra manera, de limitar el ámbito de la descripción o la protección concedida por el contenido de la patente.

6. Ejemplo

Purificación y caracterización de las actividades PI3K porcinas activadas por Gβγ

6.1 Materiales y métodos

6.1.1 Ensayos de PI3K

Se diluyeron alícuotas purificadas de PI3K derivada de Sf9 o derivada de citosol neutrófilo de cerdo en NaCl 0,12 M enfriado en baño de hielo, 25 mM HEPES, 1 mM EGTA, 1 mM DTT, 1 mg·ml⁻¹ BSA, 1% de betaína, 0,02% p/v de Tween-20, pH 7,4, 0°C en la medida apropiada, a continuación se guardaron alícuotas de 5 μl en hielo hasta su uso. Si los ensayos de PI3K se realizaban sobre inmunoprecipitados de células U937 (ver los ejemplos siguientes), entonces el PI3K se inmovilizaba en 10 μl de partículas de proteína G Sefarosa empaquetada en una solución tampón enfriada en hielo como la definida previamente. Se añadieron 30 μl de una mezcla de fosfolípidos con o sin Gβγ y/o Gα (ya sea unidas a GDP o activadas) a las fracciones de 5 μl, o a los 10 μl de partículas, y se mezclaron. Después de 10 minutos en el hielo, se añadieron y se mezclaron 5-10 μl de tampón del último lavado, suplementado con MgCl₂, para obtener una concentración final de 3,5 mM en el volumen sobrante del ensayo. Seis minutos más tarde, se añadieron 5 μl de tampón del último lavado (para conseguir un volumen final de 50 μl) que contenía [γ³²P]-ATP (normalmente 10 μCi·muestra⁻¹, Amersham, PB10168) y 3,5 mM de MgCl₂, se mezclaron los tubos y se transfirieron a un baño de agua a 30°C. Los ensayos se pararon ("quenched") después de 15 minutos con 500 μl de cloroformo:metanol:H₂O (29:54:13,1 v/v/v). A continuación se añadieron 1 μl de ATP 100 mM, 103 μl de 2,4 HCl y 434 μl de cloroformo para crear una distribución "Folch" de disolvente en dos fases. Después del mezclado y del centrifugado, las fases inferiores se recogieron y se transfirieron a tubos limpios con 424 μl de "la fase superior" nueva (metanol:HCl 1 M:Cloroformo; 48:47:3, v/v/v). Después del mezclado y el centrifugado, la fase inferior se recogió y se transfirió a un tubo nuevo (durante la purificación del PI3K neutrófilo de cerdo, el producto lipídico [³²P] se cuantificó en este punto con un contador geiger), se secó al vacío, se desacidó y se resolvió mediante TLC en placas PEI (con HCl 0,6 M; PtdIns(3,4,5)P₃ tiene un R_f de 0,47) (Stephens *et al.*, 1994, Cell 77:83-93). En algunos experimentos el lípido seco se redisolvió en cloroformo: metanol 2:1 v/v, se depositó en una placa de TLC impregnada con oxalato de potasio y se redisolvió en una mezcla de cloroformo: acetona: metanol: ácido acético: H₂O (40: 15: 13: 12:8 v/v/v/v/v; Traynor-Kaplan *et al.*, 1988).

Las mezclas de lípido/subunidad de proteína G se prepararon como se describe a continuación: PtdIns(4,5)P₂ (que se preparó a partir de fracciones del fosfoinosítido de Folch mediante 2 ciclos de cromatografía en neomicina inmovilizada) y PtdEtn (Sigma) se secaron al vacío (suficiente para conseguir concentraciones finales para el ensayo de 50 μM y 0,5 mM, respectivamente). En algunos experimentos se incluyeron PtdIns4P (preparados de manera similar a PtdIns(4,5)P₂) y/o PtdIns (Sigma). El lípido seco se sonicó en un baño (a temperatura ambiente) en tampón del último lavado (arriba) suplementado con un 0,1% de colato de sodio. Después de enfriarla en hielo, se mezclaron porciones de este stock de lípido dispersado con una mezcla, totalizando un 1 μl·muestra⁻¹ de tampón de almacenamiento Gβγ (colato 1%; HEPES 50 mM, pH 7,5, 4°C, NaCl 0,1 M; 1 mM DTT; 1 mM DTT; 0,5 mM EDTA), Gβγ activo, o un volumen equivalente de Gβγ hervido a partir de 3-7 mg·ml⁻¹ de la stock de tampón de almacenamiento. En algunos

experimentos el 1 μ l incluiría subunidades $G\alpha$ o su tampón de almacenamiento, en tal caso las $G\beta\gamma$ se premezclaron con las subunidades $G\alpha$ (ambas unidas a GDP o activadas; ver a continuación) durante 10 minutos en hielo.

Las subunidades $G\alpha$ (una mezcla equimolar de $G\alpha$:o, i, i_2 y i_3 preparada como se ha descrito, y almacenadas en el mismo tampón que las subunidades $G\beta\gamma$ excepto porque se suplementó con GDP 10 μ M) se activaron mediante incubación en hielo con NaF 10 mM, $AlCl_3$ 30 μ M (A/F) durante 10 min; las muestras en las que se diluyeron estas subunidades también contenían A/F.

a) Purificación de Proteínas

Se recogió sangre de cerdo (42 lotes) directamente en 21 contenedores con anticoagulante. En los 40 minutos siguientes a la recolección, la sangre/anticoagulante se mezcló con un 3% (p/v) de polivinilpirrolidona (PVP, 360 kD) en solución salina fisiológica (4,2 de mezcla de sangre: 0,8 de PVP). Después de 35 minutos de reposo (en 25 contenedores) los eritrocitos se asentaron adecuadamente para permitir que el sobrenadante se pudiera extraer con un sifón y centrifugar (8 minutos, 1000 x g) en contenedores de 1 litro. Se recuperaron aproximadamente 28 litros de este sobrenadante primario. Las células sedimentadas se resuspendieron en solución salina de Hank de tal manera que el volumen final acumulado se dispuso en dos botellas de centrifugación de 1 litro. Las células se sedimentaron mediante centrifugación (8 minutos 1000 x g). Los sobrenadantes se aspiraron y el pellet celular se sometió a choque hipotónico (para lisar cualquier residuo de eritrocitos contaminantes) mediante la adición de 70 ml de H_2O enfriada en hielo. Después de 25-30 segundos de mezclado, se añadieron 77 ml de solución salina de Hank 10 x (sin calcio). Después de un lavado con la solución salina de Hank, las células se reunieron en una botella de centrifugado, se sedimentaron y se resuspendieron en 500 ml de agua/hielo; TRIS 40 mM, pH 7,5, 4°C; NaCl 0,12 M; $MgCl_2$ 2,5 mM. Se añadió di-isopropilfluorofosfato (concentración final de 0,5 mM), después de 5 minutos en hielo las células se sedimentaron (aproximadamente 80-90 ml de volumen compactado) y se resuspendieron en 300 ml de TRIS 40 mM helado, pH 7,5; NaCl 0,1 M; $MgCl_2$ 2,5 mM; EGTA 1 mM; EDTA 0,2 mM y antiproteasas I. La suspensión celular se sonicó (sonicador Heat Systems Probe, ajustado a 9,25, 4 x 15 segundos con 1 minuto de mezclado, en hielo, entre cada etapa) a continuación se centrifugó (2000 x g 10 min) para sedimentar las células y núcleos no lisados (menos del 5% de las células quedaron intactas). El sobrenadante se centrifugó (a 100.000 x g 60 min, 4°C) y los sobrenadantes se decantaron, se juntaron, se mezclaron con EDTA, betaína y DDT (concentraciones finales de 5 mM, 1% y 1 mM respectivamente) y finalmente se congelaron en nitrógeno líquido y se almacenaron a -80°C. Las fracciones citosólicas preparadas de esta manera normalmente tuvieron una concentración proteica de 8 a 9 mg·ml⁻¹ (aproximadamente 2,5 g de proteína por preparado). Una vez se hubo recogido el citosol equivalente a 750 litros de sangre (18 a 42 preparados, 9 x 10¹² células, 40 g de proteína) y se hubo almacenado a -80°C, se descongelaron en tres lotes separados 5 h entre ellos. A partir de este punto todos los procedimientos se llevaron a cabo a 2-4°C.

El citosol recién descongelado se suplementó con Tween-20 (0,05%, p/v, concentración final), se centrifugó (20.000 x g durante 30 min, 2°C) se filtró (nitrato de celulosa 5 μ m, 4,5 cm de diámetro; Whatman) se diluyó aproximadamente 2,5 x con tampón K (ver abajo) hasta una conductividad final de 200 μ S (a 4°C), a continuación se cargaron (12,5 ml·min⁻¹ con una bomba peristáltica) en una columna de 800 ml (5 cm de diámetro) de Q Sefarosa de retención baja equilibrada con tampón K. El volumen total de citosol diluido fue de aproximadamente 15,5 litros. Una vez cargada, la columna se lavó con 1 litro de tampón K y a continuación se eluyó utilizando 4,5 litros de NaCl 0,1 a 0,6 M en gradiente lineal en tampón K a 8 ml·min⁻¹. Se recogieron cincuenta fracciones de 1 ml. La conductividad y la Absorbancia (a 280 nm) del eluato de la columna se monitorizaron continuamente (y en todos los pasos siguientes). El tampón K tenía la siguiente composición: TRIS/H₁ 40 mM, EGTA 1 mM, EDTA 0,2 mM, betaína 1%; Tween 20 0,05% p/v, β -glicerofosfato 5 mM pH 7,5 a 4°C, β -mercaptoetanol 15 mM con 4 μ g·ml⁻¹ de cada uno de antepaína, leupeptina, bestatina, pepstatina A y aprotinina y PMSF 0,1 mM ("antiproteasas II"). Esta solución se filtró a 0,2 μ m, de la misma manera que las siguientes.

Después de identificarse las fracciones relevantes mediante ensayos PI3K, se juntaron y se cargaron inmediatamente (10 ml·min⁻¹) en una columna de 1,8 l (5 cm de diámetro) de Sephadex-G25-Fine, que se había pre-equilibrado previamente con 18 l de tampón L (solo los 2 últimos litros con antiproteasas II), (el tampón L contenía: β -glicerofosfato 5 mM, KCl 20 mM, Tween 20 0,05% p/v, betaína 1%, EDTA 0,1 mM, fosfato de potasio 10 mM a pH 7,0 y 4°C, β -mercaptoetanol 15 mM con antiproteasas II). La solución desalada ("desalted") de la sefarosa Q se cargó inmediatamente (5 ml·min⁻¹) en 80 ml de hidroxipatita (2,6 cm de diámetro; Macropep-ceramic, BioRad) equilibrada con 1 litro de tampón M (β -glicerofosfato 5 mM; 10 mM de fosfato de potasio a pH 7,0 y 4°C, Tween 20 0,05% p/v, betaína 1%, β -mercaptoetanol 15 mM) a un caudal de 10 ml·min⁻¹, y a continuación con 100 ml de tampón N, (comprendido por tampón M suplementado con EDTA 0,1 mM y antiproteasas I) inmediatamente antes de cargar la muestra. Después de la carga, la columna se lavó con 100 ml de tampón N y se eluyó con un gradiente lineal de 0,05 a 0,35 M de fosfato de potasio en tampón N (4 ml·min⁻¹). Se recogieron fracciones de 25 ml y se determinó la actividad PI3K estimulada por $G\beta\gamma$.

Las fracciones relevantes se juntaron (normalmente un total de 100 ml), se diluyeron 3 x con tampón O (hasta una conductividad de 250 PS, 4°C) y se cargaron (1,1 ml·min⁻¹) en sefarosa heparina HR (columna de 1,6 cm de diámetro que había sido pre-equilibrada con 150 ml de tampón O (ver abajo, a 2 ml·min⁻¹). Después de la carga, la columna se lavó con tampón O (30 ml) y se eluyó con 140 ml en un gradiente lineal de KCl 0,1-0,7 M en tampón O (caudal 1 ml·min⁻¹), el eluido se recogió en fracciones de 5 ml. El tampón O era: HEPES 20 mM, EGTA 1 mM, EDTA 0,2, Tween 20 0,05% p/v, butano 1,0%, β -glicerofosfato 1 mM a pH 7,2, 4°C, β -mercaptoetanol 15 mM, además de antiproteasas II.

La actividad PI3K activada por $G\beta\gamma$ se eluyó en sefrosa heparina HR en dos picos designados picos A y B. Ambos picos A y B se purificaron de nuevo mediante el uso secuencial de la misma combinación de columnas. El pico A se localizó a los 15 ml (KCl 0,4 M) y se diluyó 8 veces en tampón P (ver a continuación) hasta una conductividad final de 200 μ S, 4°C), el pico B se localizó a los 15 ml (0,6 M KCl) y se diluyó 10 x en tampón P (hasta una conductividad final de 200 μ S, 4°C). La dilución se llevó a cabo inmediatamente antes de la carga a 1 ml·min⁻¹ en una columna Mono Q 5/5 HR pre-equilibrada con 20 ml de tampón P. Después de la carga, la columna se lavó con 5 ml de tampón P. El eluato se recogió en fracciones de 0,5 ml. El tampón P contenía: Tris 10 mM, EGTA 1 mM, EDTA 0,2, Tween 20 0,05% p/v, betaína 1%, β -glicerofosfato 1 mM, pH 7,1, 4°C, β -mercaptoetanol 15 mM y antiproteasas II.

Se juntaron las fracciones relevantes de Mono Q (A) y (B) por separado (las dos con un volumen total de 3 ml), se concentraron con una unidad de ultrafiltración (límite de 50 kD, prelavado con tampón P) hasta 0,8 ml, se centrifugaron (10.000 x g durante 10 minutos, 0°C) y se cargaron (0,25 ml·min⁻¹) directamente en una columna de cromatografía de exclusión por tamaño de alto rendimiento (V 72 ml) se preequilibró con tampón Q (ver abajo; 2 litros sin antiproteasas II, a continuación 150 ml con antiproteasas II), se recogieron fracciones de 1,5 ml inmediatamente antes del V. El tampón Q contenía: KCl 0,17 M, betaína 1%, Tween 20 0,05% p/v, β -glicerofosfato 1 mM, EGTA 1 mM, EDTA 0,2 mM, fosfato de potasio 1,5 mM, HEPES 40 mM, pH 6,9 a 4°C, β -mercaptoetanol 15 mM.

Se juntaron las fracciones relevantes de Mono Q (A) y (B) por separado (las dos con un volumen total de 6 ml), se diluyeron con tampón R hasta 24 ml (Conductividad final 250 PS, 4°C) y se cargaron (0,8 ml·min⁻¹) en una columna de HPLC de intercambio de cationes acrílica (2,5 ml de volumen, BioRad) y se eluyó con un gradiente lineal de NaCl (0,1 a 0,5 M NaCl) en tampón S a 0,1 ml·min⁻¹, el eluato se recogió en fracciones de 1 ml. El tampón R contenía: betaína 1%, Tween 20 0,05% p/v, EGTA 1 mM, EDTA 0,2 mM, HEPES 20 mM, pH 6,8 a 4°C, β -mercaptoetanol 15 mM y antiproteasas II.

Se juntaron las fracciones relevantes (3 ml de (A), 2 ml de (B)), se diluyeron 7 x con tampón S (hasta una conductividad de 180 PS, 4°C) y se cargaron (0,15 ml·min⁻¹) en una columna Mini Q (0,24 ml, operada por un sistema Pharmacia Smart™). La columna se lavó con 1 ml de tampón S y se eluyó con un gradiente lineal de NaCl (0,1 a 0,5 M NaCl) en tampón S a 0,1 ml·min⁻¹, el eluato se recogió en fracciones de 75 μ l. El tampón S contenía: betaína 1%, Tween 20 0,05% p/v, EGTA 1 mM, EDTA 0,2 mM, β -glicerofosfato 2 mM, TRIS 10 mM, pH 7,7, 4°C, DTT 1 mM (sin antiproteasas).

Las concentraciones de proteína se estimaron de cuatro maneras durante la purificación: (a) con un tinte para proteínas (BioRad; sólo se utilizó en los lisados, en las fracciones citosólicas y de sefrosa Q); (b) mediante integración de las áreas de los picos de absorbancia a 280 nm (se calibró mediante el ensayo de tinte); (c) las proteínas de los filtros se tiñeron con Ponceau S y se compararon con la intensidad de tinte de las alícuotas definidas de un estándar inmovilizado de manera similar; y (d) proteínas resueltas mediante SDS-PAGE y se tiñeron con Coomassie R250 para compararlas con alícuotas de proteínas de recorrido conocido en el mismo gel.

La preparación final de PI3K (o el material en el primer paso de purificación) se incubó con 100 nM [³H]-17-hidroxi-wortmanina (17,7 Ci mmol⁻¹, Amersham, bajo pedido), se resolvió mediante SDS-PAGE, se tiñó con Coomassie Blue, y se fotografió. A continuación se detectó el [³H] por fluorescencia.

6.2 Resultados

El análisis del citosol de neutrófilos de cerdo mediante cromatografía de intercambio iónico mostró que presentaba una actividad PI3K activada por $G\beta\gamma$ de propiedades similares a las descritas en las células U937 y en las de osteosarcoma. Al utilizar [³H]-17-hidroxi-wortmanina como sonda se identificaron un doblete de proteínas de un tamaño aparente de 117 kD y 120 kD que se eluyeron en las fracciones que contenían la actividad PI3K activada por $G\beta\gamma$, y además, presentaron un 2-4% de los niveles de unión a [³H]-17-hidroxi-wortmanina de p110 α y/o p110 β (ver Fig. 5). Este pico de actividad PI3K activada por $G\beta\gamma$ se purificó adicionalmente (todas las figuras y las tablas detallan la purificación de las preparaciones de PI3K que fueron finalmente secuenciadas). Durante este procedimiento, se separó en dos picos (A y B) que mostraron masas moleculares aparentes, nativas y relativas de 220 kD. Una vez en estado esencialmente puro, tal como se comprobó mediante geles SDS-PAGE teñidos con Coomassie, fue evidente que ambas actividades co-migraron con dos proteínas: (A) con proteínas de 117 kD (que se habían unido específicamente a [³H]-17-hidroxi-wortmanina y por lo tanto se supuso que era la subunidad catalítica) y 101 kD; y (B) con proteínas de 120 kD (que también se enlazaron a la [³H]-17-hidroxi-wortmanina) y 101 kD. Este resultado indicó que las actividades PI3K eran heterodímeros p117/p101 y p120/p101 en su estado natural. En su forma final, los PI3Ks A y B fueron purificados aproximadamente 180.000 x y 380.000 x a partir de los lisados neutrófilos (1.000.000.000 x a partir de la sangre) con 5,5% de recuperación total de la actividad. La Tabla 1 define la recuperación de proteína y de actividad PI3K a través de cada paso.

TABLA 1

Purificación de PI3K activada por la subunidad $\beta\gamma$ de la proteína G de leucocito de cerdo				
Mezcla de actividad	Proteína total en las mezclas	Volumen de la Mezcla	Actividad Total en las mezclas	Orden de purificación
Citosol	40 g	151	100%	1
Sefarosa Q (desalada)	1,5 g (1,5 g)	40 ml (450 ml)	90% (124%)	24
Hidroxiapatita	162 mg	100 ml	125%	309
Heparina Pico A	19 mg	15 ml	46%	970
Sefarosa Pico B	12 mg	15 ml	53%	1769
Mezcla Mono Q A	5,4 mg	3 ml	16%	1187
B	1,4 mg	3 ml	15%	4291
Exclusión por A Tamaño (850 μ l cargados)	0,722 mg	6 ml	16%	8902
B	0,2 mg	6 ml	15,5%	51038
Intercambio Catiónico	0,13 mg	3 ml	6%	18489
A				
B	0,014 mg	2 ml	9,5%	271761
Mezcla Mini Q partiendo de A	15 μ g	0.225 ml	2,2%	58754
De B	10 μ g	0.225 ml	3,1%	174151

Estas cantidades de enriquecimiento son consecuentes con las cantidades de [^{3} H]-17-hidroxi-wortmanina que se unieron a estas proteínas en comparación con los miembros de la familia p85/p110. Todas estas proteínas se consideran de abundancia baja.

Las preparaciones purificadas de PI3K A y B fueron indistinguibles en base a sus actividades lipidoquinasa. Ambas preparaciones fueron (a) activadas por encima de 100 x mediante subunidades $G\beta\gamma$, de una manera sensible a $G\alpha$ -GDP, (b) completamente inhibidas mediante 100 nM de wortmanina (en los ensayos, con ATP 5 μ M), (c) insensible, tanto en presencia como ausencia de $G\beta\gamma$, hacia los péptidos tirosina fosforilados que activan la familia p85/p110 de PI3K del orden de cinco veces (ver Fig. 6), y (d) capaces de 3- fosforilar PtdIns, PtdIns4P y PtdIns(4,5)P₂ (la identidad del producto se ha establecido mediante la deacilación secuencial y la deglicerización seguida del análisis por HPLC por intercambio aniónico de los productos marcados en [^{32} P] solubles en agua con estándares internos marcados en [^{3} H]). Además, las preparaciones purificadas de PI3K A y B mostraron la K_m aparente más baja para el último sustrato (8 y 10 μ M para A y B, respectivamente) utilizando el g-fosfato de ATP como el donador de fosfato (no se muestran los resultados).

7. Ejemplo

Secuenciación de péptidos de PI3K porcina activada por Gβγ A y B

- 5 En este ejemplo, se analizaron ambas proteínas PI3K porcinas mediante secuenciación de aminoácidos. Los PI3K A y B purificados a partir del equivalente a 40 g de proteína citosólica se sometieron a una transferencia western sobre nitrocelulosa, se tiñeron con Ponceau S, las bandas correspondientes a las cuatro subunidades se escindieron, se trataron con tripsina y se procesaron para un análisis de secuencia interna de aminoácidos.

10 7.1 Métodos y materiales

- Construcción y secuenciación de péptidos Se sometió a alícuotas de proteína a transferencia western (en una cuba de transferencia húmeda) sobre nitrocelulosa (tamaño de poro 0,45 μ m BA85; Schleicher and Schuell). El tampón de transferencia contenía glicina 192 mM, TRIS 25 mM y metanol 10% v/v. El soporte de papel de filtro Whatman
15 No. 1 en la cara (-) del sándwich gel/filtro y el gel (1 mm de espesor) se bañaron (2-3 minutos) en el tampón de transferencia que contenía SDS 0,0005% (p/v) antes de ensamblar la unidad final de transferencia. La transferencia se llevo a cabo durante 16 h fijada a 35 V (0,25 a 0,35 Amp, a 5°C). Los filtros se tiñeron con Ponceau S 0,1% en ácido acético al 1% durante 1 min, a continuación se destiñeron durante 1 minuto en ácido acético 1%. En el filtro se recuperó aproximadamente el 85-90% de la proteína que se cargó en el gel. Las bandas de interés se escindieron de
20 la nitrocelulosa y se procesaron para un análisis de la secuencia interna de aminoácidos como se ha descrito (Tempst *et al.*, 1990, Electrophoresis 11:537-552), con modificaciones (Lui *et al.*, 1996). Resumidamente, el acoplamiento proteolítico *in situ* se llevó a cabo utilizando 0,5 μ g de tripsina (Promega, Madison, WI) en 25 μ l de NH_4HCO_3 (suplementado con Zwittergent 3-16 1%) a 37°C durante 2 horas. La mezcla de péptidos resultante se redujo y se S-alquiló con, respectivamente, 0,1% β -mercaptoetanol (BioRad, Richmond, CA) y 4-vinilpiridina 0,3% (Aldrich,
25 Milwaukee, WI), y se fraccionaron con HPLC de fase inversa. Se preparó un blanco enzimático en una tira de celulosa del mismo tamaño.

- Los solventes HPLC y la configuración del sistema fueron los descritos (Elicone *et al.*, 1994), excepto en que se utilizó una columna 2,1 mm 214 TP54 Vydac C4 (Separations Group, Hesperia, CA) con gradiente de elución a
30 un caudal de 100 μ l/min. La identificación de los péptidos que contienen Trp se llevó a cabo mediante un análisis manual de la relación de absorbancias a 297 y 277 nm, monitorizado en tiempo real utilizando una detector modelo 1000S diodarray detector Applied Biosystems (Foster City, CA) (Erdjument-Bromage *et al.*, 1994). Las fracciones se recogieron a mano, se mantuvieron en hielo durante el proceso y a continuación se almacenaron a -70°C antes del análisis.

- Los péptidos purificados se analizaron mediante la combinación de la degradación automatizada de Edman y de Espectrometría de Masas de Tiempo de Vuelo por Desorción/Ionización mediante Láser asistida por Matriz (en inglés, *matrix-assisted laser desorption/ionization time-of-flight mass spectrometry*, MALDI-TOF); los detalles de esta aproximación combinada, incluidas las rutinas de secuenciación postquímicas asistidas por masa, pueden encontrarse
40 fácilmente (Geromanos *et al.*, 1994, Techniques in Protein Chemistry V 143-150; Elicone *et al.*, 1994, J. Chromatogr. 676:121-137; Erdjument-Bromage *et al.*, 1994, Protein Sci. 3:2435-2446). Después del almacenamiento, las fracciones de la columna se suplementaron con TPA nuevo (hasta una concentración final del 10%) antes de cargarlas en los discos de secuenciación y disponerlas en el objetivo del espectrofotómetro de masas. El análisis de masas (en alícuotas al 2%) se llevó a cabo utilizando un modelo Voyager RP MALDI-TOF (Vestec/PerSeptive, Framingham, MA) en modo
45 lineal, con láser de nitrógeno de 337 nm de salida, un tubo de vuelo de 1,3 m y ácido α -ciano-4-hidroxycinámico (solución preparada suministrada por Linear Sci., Reno, NV) como matriz. Se utilizó un voltaje de aceleración de iones de 30 kV (voltaje del selector al 70%, voltaje del cable de guía al 0,1%) y un voltaje del multiplicador de -2,0 kV. La fluencia del Láser y el número de adquisiciones se ajustaron como se juzgó a partir de las desviaciones óptimas de un máximo específico, utilizando un osciloscopio digitalizador TDS 520 Tektronix (Beaverton, OR). Los espectros
50 Mz (masa-carga) se generaron a partir de los archivos de tiempo de vuelo utilizando un programa de análisis de datos por ordenador GRAMS (Galactic Ind., Salem, NH). Cada muestreo se analizó dos veces, en presencia y ausencia de dos calibradores (APID y P8930, 25 femtomoles cada uno), tal como está descrito (Geromanos *et al.*, 1994). La secuenciación química (en el 95% de la muestra) se llevó a cabo utilizando un instrumento modelo 477A de Applied Biosystems (AB). Los aminoácidos PHT liberados gradualmente se identificaron utilizando un sistema en línea (en inglés, *on-line*) de HPLC 120A (AB) equipado con una columna PTH C18 (AB) (2/1 x 220 mm; tamaño de partícula de 5 micrones). Los instrumentos y los procedimientos fueron optimizados para un análisis a nivel de femtomoles de aminoácido feniltiohidantoina como se ha descrito (Erdjument-Bromage *et al.*, 1994, Protein Sci. 3:2435-2446; Tempst *et al.*, 1994, Methods Companion Meth. Enzymol. 6:284-261).

- 60 Las masas isotópicas promedio del péptido se reunieron a partir de los residuos identificados (incluyendo los supuestos) utilizando un programa de análisis de datos por ordenador ProComp versión 1.2 (proporcionado por el Dr. P.C. Andrews, University of Michigan, Ann Arbor, MI).

- Los servicios de microquímica del Babraham Institute prepararon un péptido tirosina doblemente fosforilado (18
65 residuos) basado en la secuencia que rodea a las tirosinas 740 y 751 en el PDGF β R.

7.2 Resultados

Los datos de la secuencia de péptidos resolvieron inmediatamente varias cuestiones en referentes a las relaciones entre estas proteínas. Las p101 derivadas de cada una de las PI3K A y B fueron idénticas y además se identificó una variante alélica relativamente común identificada como 483 en el ORF, (marcado en la Fig. 4) donde la serina fue reemplazada por glicina. p117 y p120 mostraron unos perfiles trípticos de HPLC prácticamente idénticos y todos los péptidos aparentemente comunes que se secuenciaron en ambas especies fueron idénticos con la excepción de un péptido bloqueado en el amino terminal de p117 (ver abajo). La información de la secuencia de péptidos se utilizó entonces para designar sondas para recuperar la secuencia de nucleótidos que codifican estas proteínas.

8. Ejemplo

Clonación de los ADNc que codifican el p120 y p101 de cerdo

Las sondas oligonucleotídicas degeneradas, basadas en la secuencia de un péptido del p120 de cerdo y un péptido del p101 de cerdo se utilizaron para cribar una librería de ADNc amplificada y cebada con oligo-dT (constituida a partir de ARN de cerdo neutrófilo poli A seleccionado). A continuación se describe la clonación y la caracterización del ADNc que codifica las dos proteínas p101 y p120.

8.1 Métodos y materiales

Se prepararon 0,7 mg de ARN total a partir de $4,2 \times 10^9$ neutrófilos cerdos (Chomczynski & Sacchi, 1987, Anal. Biochemistry 162:156-159). Este ARN se utilizó por Stratagene para la producción de ARNm PoliA-seleccionado a partir del cual prepararon librerías de ADN en lZAPII (aproximadamente $5,4 \times 10^{-6}$ y $3,2 \times 10^6$ p.f.u. primarios respectivamente) cebado con oligodT y cebado aleatoriamente. Las librerías amplificadas se construyeron a partir de aproximadamente 2×10^6 recombinantes originales y se utilizaron para el cribado de ADNc p120 y p101 mediante métodos comunes.

Se cribaron $2,5 \times 10^6$ placas derivadas de la librería cebada con oligodT utilizando un oligo marcado con [^{32} P] basado en la secuencia de péptidos de p120 [CA(T/C) GA(T/C) TT(T/C) ACI CA(A/G) CA(A/G) GTI CA(A/G) GTI AT(T/C/A) GA(T/C) ATG] (SEC ID NO:5). Se identificaron doce clones positivos aislados como plásmidos basados en BluescriptTM y se secuenciaron ambas cadenas DBA (utilizando la unidad de secuenciación automática ABI del Instituto Babraham). Los insertos de estos plásmidos representaron una serie de clones superpuestos con dos clones definiendo una ORF de cadena completa codificando toda la secuencia del péptido derivada de los sumarios trípticos de p117/p120 (Fig. 4).

Se cribaron $0,9 \times 10^6$ placas derivadas de la librería cebada con oligodT utilizando un oligo marcado con [^{32} P] basado en la secuencia de péptidos de p101 [GCITA(T/C)ATGGA(A/G)GA(T/C)ATIGA(A/G)GA] (SEC ID NO:6). Se identificó, aisló y se secuenció 1 clon positivo. El extremo 5' de este clon (D11) representó parte de la secuencia de uno de los péptidos trípticos, identificándolo así como un clon parcial. Se cribaron $0,6 \times 10^6$ placas derivadas de la librería cebada con oligodT utilizando un fragmento de restricción Cla-1 marcado con [^{32} P] derivado de D11. Se identificaron sesenta y seis clones positivos, 49 de los cuales se aislaron y algunos se secuenciaron parcialmente. Estos representaron una serie de clones superpuestos que contenían la secuencia de D11 aunque todos ellos fueron más pequeños que el mismo D11. Se cribaron $3,5 \times 10^6$ placas derivadas de la librería cebada aleatoriamente utilizando un fragmento de restricción Apa-1 marcado con [^{32} P] derivado de D11. Se identificaron noventa y ocho clones positivos. Estos clones se volvieron a cribar (en forma aislados primarios en placa) mediante una aproximación basada en PCR utilizando cebadores diseñados contra el vector BluescriptTM (tanto el cebador "sentido" como el "antisentido") y la secuencia interna de D11. Esto permitió identificar (independientemente de la orientación) la extensión N-terminal potencialmente más larga que codifica la secuencia p101. Se aislaron y se secuenciaron los 3 clones que dieron el fragmento de PCR más largo. Estos representaron clones superpuestos que, junto con D11, definieron un ORF de longitud completa codificando toda la secuencia peptídica derivada del sumario tríptico de p101 (Fig. 1 y Fig. 2)

8.2 Resultados

Dos de los clones definieron un ORF de longitud completa que codificaba todos los péptidos trípticos p117/p120 (ver Fig. 1-4). De tres potenciales metioninas de inicio, la central fue identificada como activa (a diferencia de la suposición de Stoyanov *et al.*, 1995, Science 269:690-693) basándose en la coincidencia precisa entre la masa medida de un péptido bloqueado en el N-terminal derivado de p120 y las masas teóricas de péptidos amino-terminales que derivarían de iniciar la traducción en cada una de las tres metioninas. Como tal, p120 tiene un tamaño teórico de 126 kD. La comparación de la masa del péptido bloqueado en el N-terminal preparado a partir de p117 con las regiones relevantes del extremo N terminal de p120 indica que no hay coincidencias exactas (que permitan el bloqueo amino-terminal normal). Por eso es improbable que p117 sea una forma modificada post-traduccionalmente de p120 o que se haya modificado proteolíticamente, tampoco es probable que sea resultado de la utilización de un segundo punto de inicio de la traducción en el mensaje de p120. Sin embargo, no se ha aislado un ADNc con un ORF codificando p117.

Las secuencias de proteína y de ADN que definen a p120 se utilizaron para buscar estructuras similares en bases de datos. Se identificaron las similitudes con todos los PI3Ks clonados previamente. Sin embargo, la secuencia fue prácticamente idéntica, teniendo en cuenta las diferencias entre especies, a p110 γ (Stoyanov *et al.*, 1995). La única

diferencia significativa entre nuestra secuencia y la de Stoyanov *et al.* se encuentra en el extremo terminal COOH. No se pudo confirmar, únicamente en base a la estructura primaria, la identificación de un dominio de homología COOH terminal con pleckstrina en p120.

- 5 Utilizando algunos fragmentos de solapamiento, derivados de las librerías de ADNc derivadas de neutrófilo de cerdo y preparadas tanto de manera aleatoria como con oligo-dT, se definió un ORF de longitud completa que codifica todas las secuencias peptídicas derivadas de p101. Se identificó un péptido bloqueado en su amino terminal derivado de p101; su masa fue exactamente equivalente a la predicha para una versión acetilada en el amino terminal de los primeros 12 residuos definidos por el punto de inicio predicho en el ORF descrito. La masa molecular relativa predicha de p101 es 97 kD. Aunque se buscaron estructuras y subestructuras similares en las bases de datos de secuencias de proteínas y ADN, no se encontraron coincidencias significativas.

9 Ejemplo

- 15 *Clonación de los ADNc que codifican el p120 y p101 de cerdo*

Los productos de PCR radiomarcados derivados de las secuencias de ADNc de cerdo p101 y p120 se utilizaron satisfactoriamente para cribar librerías de ADNc humano para sus homólogos humanos. Se describe a continuación la clonación y la caracterización de los ADNc que codifican los p101 y p120 humanos.

20 9.1 p101 humano

- Se cribó una librería de ADNc de monocito humano (Clontech, Palo Alto, CA; fuente de ARNm: línea celular U937, PMA tratado a 50 ng/ml durante 3,5 días) con un producto de PCR radiomarcado correspondiente a las pb 887-1287 de la secuencia de ADNc de cerdo p101. Fagos λ y11 que contenían la librería de ADNc humano se dispusieron en placas, como se describe en el Clontech Lambda Library Protocol Handbook, y se transfirieron a filtros de nylon. El ADN recuperado en los filtros se desnaturalizó mediante autoclave durante 1 minuto, después del cual los filtros se pre-hibridaron durante 2 horas a 42°C en formaldehído 50%, Denhardt 5x, SSC 5x, ADN de esperma de salmón 0,05 mg/ml, NaPO₄ 0,05 M a pH 6,8, pirofosfato de sodio 1 mM y SDS 1%. La sonda de PCR radiomarcada se desnaturalizó mediante ebullición durante 5 minutos, se coló en hielo durante 15 minutos, y entonces se añadió a los filtros en el tampón de hibridación a una concentración final de 1 millón cpm/ml. La hibridación se llevo a cabo durante una noche a 42°C con agitación constante. Los filtros se lavaron una vez en SSC 2x, SDS 1% a temperatura ambiente durante 20 minutos, seguido de un lavado en SSC 2x, 1% SDS a 42°C durante 20 minutos, y finalmente un lavado en SSC 0,2x, SDS 1% a 42°C durante 20 minutos. Después de la autoradiografía, los clones positivos se aislaron de acuerdo con el Clontech Library Handbook mediante el limpiado del área circundante e incluyendo el clon positivo y eluyendo el fago en un tampón de dilución de fagos. Los fagos se volvieron a disponer en placas a continuación con una densidad que permitiera el aislamiento de placas de un solo fago, se cultivaron durante la noche y se transfirieron a filtros de nylon. Los filtros se hibridaron con la misma sonda de PCR radioactiva que arriba, se lavaron, y se sometieron a autorradiografía. Se recogieron placas positivas, puras, y de un solo fago, se eluyó al fago en tampón de dilución de fagos y se volvieron a disponer en placas para proporcionar un sembrado unificado. El sembrado de placas con fagos purificados se eluyó en tampón de fagos y se aisló ADN de fago lambda según el protocolo del Clontech Library Protocol Handbook.

- El ADN de fago purificado lambda se escindió con el encima de restricción EcoRI, se separó en un gel de agarosa, se aisló el inserto EcoRI de ADN y se volvió a clonar en EcoRI Bluescript KS+ linealizado. Los plásmidos Bluescript que contenían los insertos de ADN se secuenciaron a continuación. Cinco de los clones de ADNc aislados independientemente contuvieron secuencias homólogas a la secuencia de ADNc de cerdo p101 entre el nucleótido 1137 y el codón de terminación en el nucleótido 3021. En relación a la secuencia de p101 de cerdo, los clones humanos aislados contenían 544 pb adicionales de secuencia 3' no traducida.

- Ya que la secuencia porcina es muy rica en GC en su extremo 5', la probabilidad de encontrar un clon de longitud completa en una librería cebada con oligo-dT es baja. Por esta razón se decidió utilizar una librería que estuviera cebada de manera aleatoria y con oligodT con el objetivo de incrementar la probabilidad de encontrar un clon que contuviera el extremo 5' rico en GC. Se cribó una librería de ADNc de leucocitos humanos 5'-Stretch Plus (Clontech; fuente de ARNm: leucocitos normales de la circulación periférica) utilizando el mismo producto de PCR radiomarcado de la secuencia p101 y los mismos métodos descritos arriba y se aislaron varios clones, uno de los cuales contenía el extremo 5' de la secuencia humana p101. Este clon incluyó apenas 300 pb cadena arriba de la secuencia de codificación y los primeros 1612 pb de la secuencia de codificación. A continuación se construyó un clon de ADNc de longitud completa que contenía la región de codificación entera de la secuencia de p101 humano mediante la fusión del extremo 5' del clon del leucocito (a partir del sitio HindIII a 77 pb cadena arriba del codón de inicio hasta un sitio SacI en la posición 1262) con el clon monocito (del sitio SacI hasta el sitio XbaI en la posición 3014, 305 pb secuencia abajo del codón de terminación). Se determinó la secuencia entera del clon de ADNc de longitud completa, y se dedujo la secuencia de aminoácidos codificada.

- 65 Puesto que este clon de longitud completa derivaba de dos clones diferentes de dos librerías diferentes, se comprobó la existencia del extremo 3' del clon p101 en la librería de leucocitos. Se demostró que el extremo 3' de p101, que se clonó a partir de la librería de monocito, estaba presente en las librerías de monocito y de leucocito utilizando cinco grupos diferentes de cebadores PCR.

9.2 p120 humano

El homólogo humano del p120 de cerdo se clonó mediante el cribado de una librería de ADNc de leucocito (Clontech; fuente de ARNm: leucocitos normales de la corriente sanguínea periférica) con un fragmento de PCR marcado radiactivamente de la secuencia porcina del pb 875 al 1315, según el protocolo descrito arriba para la clonación del p101 humano. Se aislaron independientemente dos clones que contenían el extremo 5' del p120 humano, correspondiente a los pb 6 a 2414 de la secuencia de la de cerdo tal como se ha descrito en Stephens *et al.* La secuencia humana contiene un sitio EcoRI en el pb 2408, y ya que la librería de ADNc se ha construido con conectores EcoRI, la secuencia de humano de longitud completa tuvo que estar contenida en dos clones separados, uno conteniendo el fragmento 5' de EcoRI, el otro conteniendo el fragmento 3' de EcoRI. Por lo tanto, la librería se cribó de nuevo bajo las mismas condiciones que arriba, con un fragmento de PCR marcado radiactivamente a partir del extremo 3' (pb 2801 a 3395) de la secuencia de p120 de cerdo. Se aislaron independientemente dos clones que contenían el extremo 3' de la secuencia desde la posición 2408 (secuencia humana) hasta el comienzo de la cola de poli(A) en la posición 5128. Se aislaron los insertos de los dos fagos conteniendo el fragmento 5' y el 3' de EcoRI y se fusionaron en el sitio EcoRI para proporcionar un clon de ADNc de cadena completa (1-5128) conteniendo el total de la región de codificación (84 a 3390) del homólogo humano p120. Se determinó la secuencia entera del clon de ADNc de longitud completa, y se dedujo la secuencia de aminoácidos codificada.

10. Ejemplo

Expresión de p120 y p101 cerdos en células de insecto

Se prepararon baculo-virus (rbv) clonales recombinantes que contenían o un amino terminal, ADNc 6 x p120 marcado con HIS (pAcHLT p120) o un amino terminal, ADNc p101 marcado con (EE) (pAcO-G1) p101) y se utilizaron para dirigir la producción de las proteínas descritas arriba en células de insecto (Sf9).

10.1 Métodos y materiales

10.1.1 Construcción de vectores de expresión

La utilización de PCR N-terminal y de las dianas de restricción apropiadas permitieron la manipulación de los ORF p120 y p101 de manera que pudieran ser insertados en el marco de lectura de varios vectores de expresión. En cada caso, el primer aminoácido codificado después de marcar el N-terminal fue la metionina de inicio. La región 3' sin traducir del ADNc p120 era la expresión dirigida del baculo-virus en las células Sf9 donde pAcHLT (que codifica un marcador 6 x His seguida de un sitio de enlace de trombina, Pharminogen; el ORFp120 se insertó en los sitios XhoI-EcoRI) y pAcO-GI (que codifica un marcador EE N-terminal, ONYX Pharmaceuticals; el ORF p101 se insertó en los sitios EcoRI-NotI). Todos los vectores se secuenciaron de forma N-terminal antes de su uso.

10.1.2 Transfecciones de la célula Sf9 y producción de proteínas recombinantes

Las células Sf9 se cultivaron en TNM FH con HI-FBS 11% en un matraz agitado y se mantuvieron entre 0,5 y 2 x 10⁶ células/ml⁻¹. Las células Sf9 se transfectaron con liposomas InsectinTM (Invitrogen) con ADN baculo-oro (*en inglés, baculo-gold*) (Pharminogen) y con los vectores baculo-víricos de transferencia relevantes tal como se recomienda (Invitrogen). Los baculo virus recombinantes resultantes se purificaron en placa y se amplificaron para obtener stocks virales de alta concentración. Las diluciones óptimas (para la producción de proteína) se determinaron para cada uno de los stocks de alta concentración. A los virus p101 pAcO-G1 se les permitió infectar células Sf9 adherentes durante 2,2 días a 27°C; a los virus p120 pAcHLT se les permitió infectar en un cultivo con agitación (normalmente) durante 1,9 días a 27°C. Las células se recogieron en KCl 0,41%; Sacarosa 2,66%; MgCl₂ 20 mM; NaH₂PO₄ 8 mM, pH 6,2, 25°C; se trataron con di-isopropilfluorofosfato (5 minutos en hielo), y se lavaron una vez en la solución de recogida. Las alícuotas de células compactadas por centrifugación se congelaron en N líquido, y se almacenaron a -80°C.

10.1.3 Purificación de proteínas derivadas de Sf9

El p120 de cerdo se purificó utilizando una columna de quelato metal-ión (Talon, Clontech). Los pellets celulares se descongelaron en NaCl 0,10 M; fosfato de sodio 50 mM, pH 8,0, 4°C; Tris-HCl 10 mM, pH 8,0, 4°C; 1 mM MgCl₂ y antiproteasas I (ver arriba) en una relación de 1 litro de cultivo infectado Sf9 en 50 ml de tampón de sonicado, se sonicaron (intervalos 4 x 15 segundos en hielo), y se centrifugaron (120.000 x g durante 40 minutos, 4°C). El sobrenadante se recogió y se juntó (el sobrenadante más esponjoso de la parte superior del tubo se recogió por separado, se filtró a 0,45 µm con un filtro con baja unión de proteínas y a continuación se juntó con el resto). La fracción citosólica se suplementó con Tween-20 y betaína (0,05%, p/v, y 1%, respectivamente) a continuación se bombearon en una columna de resina Talon equilibrada en tampón equivalente (1,2 ml de resina por litro de cultivo original de Sf9 infectada a una caudal lineal de 20 cm/hr⁻¹). Este y los pasos siguientes se llevaron a cabo a 4°C. La columna se lavó secuencialmente (con el mismo caudal) con 20 volúmenes de la columna, cada uno de ellos de: Tampón A, fosfato de sodio 50 mM, pH 8,0, 4°C; Tris/HCl 10 mM, pH 8,0, 4°C; NaCl 0,15 M; betaína 1%; Tween 20 0,05% p/v; tampón B, 1%, p/v, Triton X-100; 0,15 M NaCl; fosfato de sodio 50 mM, pH 8,0, 4°C; Tris 10 mM, pH 8,0, 4°C; betaína 1%, Tween-20 0,05%, p/v; buffer C, NaCl 0,15 M; fosfato de sodio 50 mM, pH 7,1, 4°C; betaína 1%; Tween-20 0,05%, p/v; tampón D; NaCl 0,15 M; Tris 30 mM, pH 7,5, 4°C, betaína 1%, Tween-20, 0,02%, p/v; etilenglicol 10%, v/v;

MgCl₂ 1 mM; y a continuación 8 volúmenes de columna de tampón E, comprendiendo tampón D suplementado con imidazole 10 mM (pH 7,5). Durante el lavado con tampón E, se recogieron fracciones de 2 ml. Finalmente, a la mitad de caudal utilizado previamente, se lavó la columna con tampón F, que comprendía tampón D suplementado con imidazole 70 mM (pH 7,5; concentración final). Se recogieron fracciones de 1 ml. Las fracciones se juntaron en función del espectro de Abs 280 nm (registrado continuamente) y se suplementó con DTT 1 mM y 1 mM EGTA (concentraciones finales) inmediatamente. Este proceso normalmente rindió 4 mg de p120 por litro de cultivo Sf9. La pureza del p120 preparado de esta manera normalmente fue mayor del 90%. La mezcla final de p120 se desaló a través de una columna de 15 cm de G-25 superfino equilibrado en tampón G, que estaba compuesto de tampón D suplementado con DTT 1 mM y EGTA 1 mM (concentraciones finales). El blanco de p120 utilizado en algunos experimentos se preparó totalmente en paralelo a las preparaciones normales de p120 excepto en las células de partida, que estaban infectadas con baculo virus natural, o no estaban infectadas. Las fracciones finales derivadas de estas preparaciones blanco se juntaron en un "volumen paralelo" ya que prácticamente no contenían proteína.

El marcador p120 6x HIS contenía un motivo de reconocimiento de escisión de trombina. La valoración cuidadosa con trombina y el análisis con geles SDS de poliacrilamida al 6% revelaron dos puntos de trombina, con similares sensibilidades a la trombina, ambos cercanos al amino terminal (ya que un anticuerpo aCOOH-terminal todavía precipitó el p120 escindido dos veces). Uno de los puntos está en el lugar esperado para el punto modificado del marcador; el otro estaba a aproximadamente 40 residuos del amino terminal (en una región sin secuencias de reconocimiento favorables a trombina). Bajo las condiciones optimizadas (trombina 2 U·ml⁻¹; p120 0,2 mg·ml⁻¹; 4 horas, 4°C, con EGTA 1 mM y MgCl₂ 1 mM) fue posible preparar preparaciones de p120 escindida con trombina que contenían el 15% de p120 sin escindir, el 50% escindido por el punto de trombina amino terminal real y el 35% con aproximadamente 40 residuos escindidos adicionales. En estos experimentos la acción de la trombina se terminó mediante la adición de ácido N-acetil D-phe Pro-2-amido-5 guanidino butano borónico 100 nM; un potente inhibidor de la trombina (Sigma). A través de todo el estudio, se utilizaron en paralelo preparaciones de p120 parcialmente escindidos con p120 totalmente sin escindir. Estaba claro que: (A) Los tres p120 se unieron a p101 con una afinidad muy similar, (B) p101 que se enlazó a la mezcla de p120 se activó por las Gβγ con una intensidad similar al p101 unido al p120 sin escindir y, (C) ningún p120, ni el escindido parcialmente ni el que no había sido escindido, respondió a las subunidades Gβγ en ausencia de p101 (aunque la escisión completa por trombina tendió a provocar la reducción de la actividad catalítica de p120 en un 20-30% y la activación aparente mediante Gβγ de 1,7 veces se incrementó en ausencia de p101 hasta aproximadamente 2,5 veces).

A partir de los pellets celulares de las células Sf9 congelados se purificó el (EE)-p101 de cerdo como se describe a continuación. Las células de 2 litros de cultivo Sf9 infectado se sonicaron en 50 ml de NaCl 0,12 M; MgCl₂ 1 mM; HEPES 25 mM, pH 7,4, 4°C; EGTA 1 mM con antiproteasas I como se describe arriba. Después de la centrifugación (120.000 x g, 4°C, 40 minutos), el sobrenadante se recogió (como se describe arriba), se suplementó con Triton X100 1% p/v, colato de sodio 0,4%, NaCl 0,4 M (concentraciones finales), y se mezclaron con 2 ml de proteína G sefarosa empaquetada entrelazada covalentemente a un anticuerpo monoclonal irrelevante α-(myc) (lavado en una solución equivalente). Después de 30 minutos mezclando a 4°C, las partículas se sedimentaron y se recogió el sobrenadante y se mezcló con 1 ml de proteína G sefarosa empaquetada entrelazada covalentemente a un anticuerpo monoclonal irrelevante α-(myc) (equilibrada en una solución equivalente). Después de 2 horas mezclando a 4°C, las partículas se lavaron como se describe a continuación para las inmunoprecipitaciones α-(EE) de la célula U937, excepto que (a) los lavados se llevaron a cabo en un tubo de centrifugación de 20 ml, y (b) las partículas se lavaron finalmente 3 x con tampón H, comprendido por tampón G con Triton X-100 1%, p/v. Las preparaciones del blanco p101 (p101C), que se utilizaron en alguno de los experimentos, se prepararon a partir de células infectadas o no infectadas por el baculo virus natural exactamente como se describe arriba.

Los heterodímeros p101/p120 formados *in vivo* mediante la coinfección de células Sf9 con ambas formas de los baculo virus recombinantes se purificaron como se describe para (EE)-p101 excepto porque los inmunoprecipitados se lavaron 4 x con tampón I, Triton X-100 1%, p/v; NaCl 0,15 M; HEPES 20 mM, pH 7,4, 4°C; EGTA 1 mM; 2 x con tampón J, que comprende tampón I suplementado con NaCl 0,4 M (concentración final), a continuación 3 x con tampón G antes de eluirse con 1 volumen de 150 μg·ml⁻¹ (concentración final) de (EE)-péptido en tampón G. El (EE)-péptido, EYMPTD acetilado en el amino terminal, tiene una alta afinidad por el anticuerpo monoclonal α-(EE). Después de incubar las partículas con (EE)-péptido, en hielo, durante 40 minutos, se recogió el sobrenadante. Se diluyeron las alícuotas de las proteínas eluidas y se ensayó su actividad PI3K (ver arriba) o se mezclaron con tampón de muestra SDS directamente.

En los experimentos de reconstitución *in vitro*, las preparaciones de p120 en tampón G suplementado con Triton X-100 1%, p/v (ahora equivalente a tampón H) se mezclaron con (EE)-p101 (relación molar de proteína 10:1) todavía unido a la matriz de proteína G, se mezcló durante 2 horas (siempre a 4°C), a continuación se lavaron y se eluyeron con (EE)-péptido tal como se ha descrito para la purificación del p101/p120-PI3K reconstituido *in vivo*, en el Ejemplo siguiente.

10.2 Resultados

Las purificaciones directas utilizando el marcador proporcionaron preparaciones purificadas de proteína comprobadas mediante tinción de Coomassie; sus tamaños aparentes fueron los esperados. Además, ambas se reconocieron correctamente en transferencia western mediante sueros policlonales de conejo con antipéptidos específicos, dirigidos al COOH terminal y dirigidos a la secuencia interna, demostrando su autenticidad.

p120 se unió fuertemente a p101 en una estequiometría 1:1 en moles, tanto (a) *in vitro*, cuando ambas proteínas habían sido purificadas por separado, mezcladas (con p101 todavía asociada a la matriz de proteína G utilizada para aislarla) a continuación lavadas a conciencia antes de eluir las con (EE)-péptido o, (b) *in vivo*, cuando las células Sf9 fueron coinfectadas con ambas formas de rbv y las proteínas se purificaron a través del marcador p101 (no se proporcionan datos).

El p120 purificado y libre fue un PI3K selectivo a $\text{PtdIns}(4,5)\text{P}_2$ y sensible a wortmanina. Las $G\beta\gamma$ tuvieron un efecto bimodal pequeño en la actividad PI3K p120 libre. Una mezcla equimolar (concentración final total de $2\text{ }\mu\text{M}$) de $G\alpha$ o, i1, i2 y i3 unidos a GDP y en presencia de fluoruro de aluminio no tuvieron efectos significantes sobre la actividad PI3K p120 libre (esta era una preparación de $G\alpha$ que, cuando se añadió un exceso molar de 1,5 veces en PI3K, pudo inhibir completamente los efectos de las $G\beta\gamma$). Los péptidos tirosina fosforilados capaces de activar los miembros de la familia p85/p110-PI3K tampoco tuvieron efecto apreciable sobre la actividad PI3K p120. Cuando p120 se unió a p101 (tanto *in vivo* como *in vitro*), la actividad PI3K se pudo activar por encima de las 100 x mediante las subunidades $G\beta\gamma$. Los péptidos tirosina fosforilados y el $G\alpha$ -GDP/fluoruro de aluminio no tuvieron efecto sobre la actividad p101/p120 PI3K basal o activada por $G\beta\gamma$. En la ausencia de las $G\beta\gamma$, la actividad específica de p120 en un complejo p101/p120 es menor que la actividad específica del p120 libre pero, en su presencia, se incrementa en más de 50x (ver Fig. 7).

La comparación de las actividades catalíticas específicas de PI3K derivado de neutrófilos de cerdo (B) y los heterodímeros p101/p120 derivados de Sf9, bajo condiciones idénticas de ensayo, estimulados por $G\beta\gamma$, mostró que el material derivado de Sf9 tenía una actividad 2 x mayor por mg de proteína. Este resultado no es inconsistente con el probable "age at assay" de una preparación PI3K derivada de neutrófilos circulantes mediante un protocolo de purificación de 4 días e indica que la masa del PI3K está correctamente plegado y que todas las modificaciones post-transcripcionales deben estar en su lugar correcto.

11. Ejemplo

Expresión de p101 y p120 de cerdo en células de mamífero

Se construyeron una familia de vectores de expresión para mamíferos que permitieron la expresión transitoria de formas amino terminales marcadas epitópicamente (tanto (myc) como (EE)) de p101 y p120 de cerdo. Este ejemplo describe la producción a partir de células de mamífero de las proteínas recombinantes de fusión p120 y p101 recombinantes purificadas, y el posterior análisis de sus propiedades.

11.1 Materiales y métodos

11.1.1 Cultivo celular

Las células U937 se cultivaron en RPMI 1640 con (HI)-FBS 10% v/v inactivado por calor y se diluyeron por 4 cada 2 días. Las células Cos-7 se cultivaron en DMEM 10% HI-FBS.

11.1.2 Construcción de vectores de expresión

La utilización de PCR N-terminal y de las dianas de restricción apropiadas permitió la manipulación de los ORF p120 y p101 de manera que pudieran ser insertados en el marco de lectura de varios vectores de expresión (el primer aminoácido codificado después del marcador N-terminal fue, en cada caso, la metionina de inicial). La resina 3' no traducida del ADNc p120 se utilizó en su totalidad y la del ADNc p101 se escindió en el punto BamHI (nucleótido 192, Fig. 1). Los vectores utilizados para la expresión dirigida por el citomegalovirus (CMV) en células de mamífero fueron (A) pCMV(EE) (que codifica un N-terminal MEEEEFMPMPMEF (SEC ID NO:7) o un marcador EE MEEEEFMPMEFSS (SEC ID NO:8) para la expresión de p101 o p120, respectivamente) y (B) pCMV (myc) (que codifica un N-terminal MEQKLISEEDLEF (SEC ID NO:9) o un marcador myc MEQKLISEEDLEFSS (SEC ID NO:10) para la expresión de p101 o p120, respectivamente). Todos los vectores se secuenciaron de manera N-terminal antes de su uso.

11.1.3 Protocolos de transfección de U937

Las células U937 con crecimiento exponencial (diluidas 12 horas antes) se lavaron 2 x con PBS y se resuspendieron en un medio estéril de electroporación (EM) que contenía: NaCl 30 mM, KCl 0,12 M, Na_2HPO_4 8,1 mM, KH_2PO_4 1,46 mM y MgCl_2 5 mM, a temperatura ambiente. El ADN circular plásmido se añadió en 1 x EM a las células para producir un volumen final de 0,5 ml que contenía $1,4 \times 10^7$ células y 40 μg de ADN total (normalmente llevado a 40 μg con un plásmido de expresión con el mismo promotor y que expresa una proteína irrelevante marcada de manera similar) y se transfirió a una cubeta (0,4 cm gap, BioRad). Después de 15 minutos de reposo a temperatura ambiente, las células se electroporificaron (1 pulso a 280V y 960 μF , con un pulsador BioRad Gene; las constantes de tiempo normalmente fueron 18 msec) a continuación se colocaron en hielo durante 8 minutos más antes de diluirse en 5 ml de RPMI, HI-FBS 10%. Después de reposar durante 5 minutos, para permitir que las células muertas se agruparan, las células se diluyeron con 35 ml de RPMI, HI-FBS 10%, suplementados con penicilina y estreptomycin, a continuación se añadieron TPA y ZnCl_2 (los cuales amplifican sustancialmente la expresión de los promotores en las células U937) hasta concentraciones finales de 5×10^{-8} M y 200 μM , respectivamente. Si se hubieran marcado las células

con metionina [³⁵S] (marcaje trans, ICN), el RPMI utilizado después de la electroporación hubiera sido RPMI libre de metionina y leucina y hubiera contenido NaHCO₃ 2 mM y HEPES 25 mM y 10% HI-FBS dializado y las células se resuspendieron en un volumen final de 10 ml con 20 µCi/ml metionina/leucina [³⁵S] (además de TPA y ZnCl₂, como arriba). Después de 12 horas (tanto para el caso con [³⁵S] y sin [³⁵S]) las suspensiones celulares se mezclaron con di-isopropilfluorofosfato (1 mM de concentración final), se dejó reposar 5 minutos, se recogieron por centrifugación, se lavaron 1x con PBS, y se lisaron para la inmunoprecipitación. Las proteínas-(EE) marcadas epitópicamente se inmunoprecipitaron a partir de lisados celulares de U937 como se describe a continuación. Las células de una de las electroporificaciones se lisaron en 1,25 ml de tampón lítico que contiene Triton X-100 1% p/v, HEPES 25 mM, pH 7,4, 4°C; EGTA 1 mM; MgCl₂ 1 mM; NaCl 0,15 M y PMSF 0,1 mM, leupeptina 10 µg·ml⁻¹, aprotinina 10 µg·ml⁻¹, antipaina 10 µg·ml⁻¹, pepstatina A 10 µg·ml⁻¹ y bestatina 10 µg·ml⁻¹ (de aquí en adelante referidas como antiproteasas I). Los lisados se centrifugaron (12,000 rpm, 0°C, 15 minutos). Los sobrenadantes resultantes se recogieron y se mezclaron con NaCl 4M (0,1 ml), colato de sodio (25 µl) 20% y proteína G sefarosa (volumen empaquetado 40 µl) unida covalentemente a un anticuerpo monoclonal irrelevante de ratón coincidente en isotipo ("isotype-matched") (aproximadamente 5-10 mg·ml⁻¹ de sefarosa). Los sobrenadantes se mezclaron totalmente durante 20 minutos, y a continuación se sedimentaron las partículas y se transfirió el sobrenadante a otro tubo con partículas de proteína G sefarosa (10 µl, empaquetadas) covalentemente reticuladas al anticuerpo monoclonal de ratón α-(EE) (aproximadamente 5-10 mg·ml⁻¹ de sefarosa). Después de 2 horas mezclándolos a 4°C, se lavaron los inmunoprecipitados 5 x con Triton X-100 1,0%, p/v; colato 0,4%; SDS 0,004%; NaCl 0,4 M; EGTA 1 mM; HEPES 25 mM, pH 7,4, a 0°C; 1 x con LiCl 0,5 M, Tris 0,1 M, pH 8,0, 4°C; 2 x con NaCl 0,12 M; EGTA 1 mM; HEPES 25 mM, pH 7,4, a 0°C. El último lavado con tampón se suplementó con DTT 1 mM (tampón del lavado final). Si los inmunoprecipitados se prepararon a partir de células marcadas con metionina [³⁵S], se omitió el último lavado y las partículas se hirvieron en tampón de muestra SDS. Aparte, se ensayó la actividad PI3K de las muestras tal como se describe a continuación.

11.1.4 Transfección de células COS-7

3 horas antes de la transfección, las células Cos-7 con crecimiento exponencial se tripsinizaron y se sembraron a una confluencia aproximada del 50-70%. En el momento de la transfección fueron tripsinizadas de nuevo, se diluyeron en FBS 10% DMEM, se contaron, se lavaron 2 x en PBS y se resuspendieron en EM (1 x 10⁷ por cubeta) se mezclaron con plásmido de ADN circular (40 µg totales por cubeta, comprende combinaciones de 10 µg de EXV-(EE)-β₁, 10 µg de EXV-(myc)-τ₂, 10 µg de pCMV-(myc)-p120 o 10-40 µg de un vector de expresión de mamífero), para dar un volumen final de 500 µl y a continuación se transfirió a una cubeta de electroporación (0,4 cm gap, BioRad). Después de 10 minutos a temperatura ambiente, las células se electroporificaron (250V pulso a 960 µF), se colocaron en hielo durante 8 minutos, y a continuación se diluyeron en DMEM FBS 10%. Se permitió a las células agregadas que se agruparan y se descartaron al formar alícuotas en los platos de 6 cm (4 por cada cubeta).

Después de 48 horas, los cuatro platos de cada tratamiento se lavaron en DMEM tamponado con HEPES que contenía NaHCO₃ 1 mM y BSA libre de ácidos grasos 0,2%. Después de 10 horas se recogió el producto de dos de los platos de replicación para transferencia western (con anticuerpo monoclonal α-(myc) como primer reactivo de detección, a-HRP de ratón como el segundo sistema de detección y cuantificación mediante ECL y escaneo de densidad). Se lavaron dos platos en DMEM libre de fosfatos, con NaHCO₃ 1 mM y BSA libre de ácidos grasos 0,2%, a continuación se incubaron durante 90 minutos más a 37°C con 300 µCi [³²P]-Pi por plato (en 4 ml). Se aspiró el medio y se añadió 1 ml de HCl 1 M enfriado en hielo. Se lisaron las células, se recogieron y se lavaron los platos con 1,33 ml de metanol. Los lavados con HCl y metanol se juntaron con 2,66 ml de cloroformo (para obtener una distribución de disolvente Folch de dos fases) se mezclaron y se centrifugaron. Las fases inferiores se recogieron y se mezclaron con 1,95 ml de la fase superior limpia (ver arriba) que contenía EDTA 0,5 mM y sulfato de tetrabutilamonio 1 mM. Después de mezclar y centrifugar, las fases inferiores se recogieron, se secaron, se deacilaron y se prepararon para el análisis por HPLC de intercambio aniónico (Stephens *et al.*, 1991, Nature. Lond. 351:33-39).

11.2 Resultados

Cuando se expresan transitoriamente (EE)-p101 y (EE)-p120 en las células U937, pueden inmunoprecipitarse a partir de las células marcadas con metionina [³⁵S] en cantidades aproximadamente iguales (teniendo en cuenta su cumplimiento relativo de metioninas; 8:25 respectivamente, datos no proporcionados). Los inmunoprecipitados α-(EE)-p101 lavados a conciencia contenían una actividad PI3K selectiva para PtdIns(4,5)P₂, sensible a wortmanina y sensible a Gβγ que no estaba presente en los controles utilizando un anticuerpo monoclonal cualquiera para la inmunoprecipitación, o un ADNc codificando una proteína cualquiera con el marcador (EE) que se expresó, a juzgar por el marcado con metionina [³⁵S], a niveles similares que los de (EE)-p101.

La activación mediante Gβγ se bloqueó mediante el preincubado con un exceso molar de dos veces de Gα-GDP. Además, la actividad PI3K de estos inmunoprecipitados p101 no fue sensible a Gα-GDP/fluoruro de aluminio. La cotransfección de (myc)-p120 con (EE)-p101 no incrementó la actividad PI3K recuperada específicamente en los inmunoprecipitados α-(EE). En realidad descendió, probablemente a causa de que la expresión de (EE)-p101 era relativamente inferior en presencia de los vectores de expresión (myc)-p120. Por el contrario, en las células transfectadas con (EE)-p120, los inmunoprecipitados α-(EE) prácticamente no presentaron actividad PI3K detectable ni en presencia ni en ausencia de Gβγ. Estas células contenían un número comparable de moles de p120 marcado con metionina [³⁵S] a los de las (EE)-p101 en los inmunoprecipitados α-(EE) de las células transfectadas con (EE)-p101. La cotransfección con (myc)-p101 resultó en un incremento sustancial específico de la actividad PI3K estimulada por Gβγ que pudo recuperarse, a pesar de un descenso en la expresión de (EE)-p120 (juzgado a partir del marcaje con

metionina [^{35}S] (ver Fig. 8). Estos datos indican que las células U937 (humanas) contienen una subunidad catalítica endógena PI3K que puede unirse a una p101 (porcina) expresada transitoriamente. Cuando se une a p101 (porcina) esta subunidad catalítica endógena muestra una regulación sustancial mediante $G\beta\gamma$ porque todo el p120 presente en los inmunoprecipitados se une a p101.

Por el contrario, en los inmunoprecipitados α -(EE) a partir de las células transfectadas con (EE)-p120, gran parte del p120 no está asociado con el p101 y por esta razón es relativamente inactivo (comparado con el unido a p101 y en presencia de $G\beta\gamma$). Sin embargo, esta actividad PI3K, cuando se ensaya en presencia de $G\beta\gamma$, se amplifica sustancialmente mediante la cotransfección con (myc)-p101. La explicación alternativa para estos datos es que el p120 se “desnaturaliza” (aunque continua soluble y con capacidad de ser inmunoprecipitado) a menos que se exprese en presencia de p101 - lo que es poco probable en vista de los datos obtenidos con las proteínas purificadas independientemente, derivadas de Sf9.

Para probar si la p101/p120 PI3K podría activarse mediante $G\beta\gamma$ y producir $\text{PtdIns}(3,4,5)\text{P}_3$ *in vivo*, se expresaron transitoriamente varias combinaciones de (myc)- γ_2 , (EE)- β_1 , (myc)-p101 y (myc)-p120 en las células Cos-7 y se midieron sus efectos en los niveles de fosoinositidos [^{35}P] en células 48 horas después de la transfección (ver Fig. 9). p120 sólo produjo incrementos significativos en $\text{PtdIns}(3,4,5)\text{P}_3$ y $\text{PtdIns}(3,4)\text{P}_2$ de manera dependiente de $\beta_1\gamma_2$ en presencia de p101. Este patrón de resultados no se puede explicar mediante los cambios en la expresión relativa de los diferentes ADNc cuando se introducen en combinaciones (ver Fig. 9).

12. Depósito de clones

Los microorganismos o clones siguientes se depositaron en la American Type Culture Collection (ATCC), Rockville, Maryland, en las fechas indicadas y se les asignó el número de acceso indicado:

Clon	Número de Acceso ATCC	Fecha del depósito
pCMV3mycp101	97636	6/27/96
pCMV3mycp120	97637	6/27/96

La presente invención no está limitada en el alcance de su protección por las realizaciones específicas descritas en el presente documento, las cuales se entienden como ilustraciones individuales de aspectos individuales de la invención, y por tanto los métodos y componentes funcionalmente equivalentes están dentro del alcance de la invención. Así, para los expertos en la materia, varias modificaciones de la invención se desprenderán de la descripción de las figuras que la acompañan. Tales modificaciones se suponen dentro del alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Molécula de ácido nucleico aislada que codifica una subunidad reguladora p101 de PI3K regulada por proteína G o un fragmento de la misma de al menos 30 nucleótidos, que tiene la secuencia de nucleótidos que:
 - a) codifica SEC ID NO: 2 o la secuencia de aminoácidos codificada por el ADNc contenido en el clon de ADNc pCMV3mycp101 tal y como está depositado en la ATCC con el número de acceso 97636;
 - b) codifica SEC ID NO: 12; o
 - c) hibrida en condiciones rigurosas con la secuencia nucleotídica de (a) o (b) o con sus complementarias.
2. Molécula de ácido nucleico aislado según la reivindicación 1, en la que la molécula de ácido nucleico hibrida en condiciones moderadamente rigurosas con la secuencia nucleotídica de (a) o (b) o con sus complementarias, y en la que la molécula de ácido nucleico codifica un producto del gen p101.
3. Molécula de ácido nucleico aislado según la reivindicación 1, en la que la molécula de ácido nucleico hibrida en condiciones altamente rigurosas con la secuencia nucleotídica de (a) o (b) o a sus complementarias.
4. Molécula de ácido nucleico aislado según la reivindicación 1, en la que la molécula de ácido nucleico codifica SEC ID NO: 2 o la secuencia de aminoácidos codificada por el ADNc contenido en el clon de ADNc pCMVmycp101 tal y como está depositada en ATCC con el número de acceso N° 97636.
5. Molécula de ácido nucleico aislado según la reivindicación 1, en la que la molécula de ácido nucleico codifica SEC ID NO: 12.
6. Molécula de ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos seleccionada del grupo que consiste en los residuos aminoácidos 1-160 de SEC ID NO: 2, los residuos aminoácidos 161-263 de SEC ID NO: 2, los residuos aminoácidos 1-732 de SEC ID NO: 2, los residuos aminoácidos 733-877 de SEC ID NO: 2, los residuos aminoácidos de 1-160 de SEC ID NO: 12, los residuos aminoácidos 161-263 de SEC ID NO: 12, los residuos aminoácidos 1-735 de SEC ID NO: 12 y los residuos aminoácidos 736-880 de SEC ID NO: 12.
7. Secuencia de nucleótidos aislada que codifica una proteína quimérica que comprende el polipéptido de cualquiera de las reivindicaciones 1-6, fusionado a un polipéptido heterólogo.
8. Secuencia de nucleótidos aislada según la reivindicación 7, en la que el polipéptido heterólogo es un marcador Glu o un marcador de epítipo myc.
9. Vector nucleotídico que contiene la secuencia de nucleótidos de cualquiera de las reivindicaciones 1-8.
10. Vector de expresión que contiene la secuencia de nucleótidos de cualquiera de las reivindicaciones 1-8, en asociación operativa con una secuencia de nucleótidos reguladora que controla la expresión de la secuencia de nucleótidos en una célula hospedadora.
11. Vector de expresión según la reivindicación 10, en el que dicha secuencia reguladora se selecciona del grupo que consiste en el promotor del gen inmediato temprano de citomegalovirus hCMV, el promotor temprano o tardío del adenovirus SV40, el promotor de baculovirus, el sistema lac, el sistema trp, el sistema TAC, el sistema TRC, las regiones principales del operador y promotor del fago λ , las regiones de control de la proteína de cubierta de fd, el promotor de la 3-fosfoglicerato quinasa, los promotores de la fosfatasa ácida y los promotores de los factores de acoplamiento a de levadura.
12. Célula hospedadora que ha sido manipulada genéticamente para contener el vector de cualquiera de las reivindicaciones 9-11.
13. Animal transgénico no humano que comprende la célula hospedadora de la reivindicación 12.
14. Célula hospedadora según la reivindicación 12, en la que la célula hospedadora ha sido manipulada genéticamente para contener el vector de la reivindicación 11.
15. Célula hospedadora según la reivindicación 14, en la que la célula hospedadora es una célula Sf9, una célula de ovario de hámster chino, una célula COS o una célula U937.
16. Célula hospedadora manipulada genéticamente en la que la secuencia del gen nativo para la subunidad p101 de PI3K regulada por proteína G se interrumpe de manera que se inhibe o evita la expresión del producto del gen nativo de P101, en donde la secuencia del gen nativo codifica la secuencia de nucleótidos de cualquiera de las reivindicaciones 1-5.

ES 2 278 392 T3

17. Subunidad reguladora aislada p101 de PI3K regulada por proteína G que tiene una secuencia de aminoácidos codificada por una secuencia de nucleótidos que:

(a) codifica SEC ID NO: 2 o la secuencia de aminoácidos codificada por el ADNc contenido en el clon de ADNc pCMV3mycp101 tal y como está depositado en ATCC con el número de acceso 97636;

(b) codifica SEC ID NO: 12; o

(c) hibrida en condiciones rigurosas con la secuencia nucleotídica de (a) o (b) o con sus complementarias, en la que dicha subunidad reguladora p101 está libre de la subunidad catalítica p120.

18. Subunidad reguladora p101 aislada de PI3K regulada por proteína G de acuerdo con la reivindicación 17, que tiene la secuencia de aminoácidos de SEC ID NO: 2, o la secuencia de aminoácidos codificada por el ADNc contenido en el clon de ADNc pCMV3mycp101 tal y como está depositada en ATCC con el número de acceso 97636.

19. Subunidad reguladora p101 aislada de PI3K regulada por proteína G de acuerdo con la reivindicación 17 que tiene la secuencia de aminoácidos SEC ID NO: 12.

20. Polipéptido que tiene una secuencia de aminoácidos seleccionada del grupo que consiste en los residuos aminoácidos 1-160 de SEC ID NO: 2, los residuos aminoácidos 161-263 de SEC ID NO:2, los residuos aminoácidos 1-732 de SEC ID NO: 2, los residuos aminoácidos 733-877 de SEC ID NO: 2, los residuos aminoácidos de 1-160 de SEC ID NO: 12, los residuos aminoácidos 161-263 de SEC ID NO: 12, los residuos aminoácidos 1-735 de SEC ID NO: 12 y los residuos aminoácidos 736-880 de SEC ID NO: 12.

21. Proteína quimérica que comprende el polipéptido de cualquiera de las reivindicaciones 17-20 fusionado a un polipéptido heterólogo.

22. Proteína quimérica de la reivindicación 21 en la que el polipéptido heterólogo es un marcador Glu o un marcador de epítipo myc.

23. Anticuerpo que se une inmunoespecíficamente a la subunidad reguladora p101 de PI3K regulada por proteína G de cualquiera de las reivindicaciones 17-20.

24. Método *in vitro* para el diagnóstico de una enfermedad de respuesta inflamatoria en un mamífero que comprende detectar una mutación en el gen de la subunidad reguladora p101 de PI3K regulada por proteína G contenido en el genoma del mamífero, en donde la subunidad reguladora p101 de PI3K regulada por proteína G natural expresa el polinucleótido de cualquiera de las reivindicaciones 1-5.

25. Procedimiento para la identificación de compuestos útiles en el tratamiento de enfermedades de respuesta inflamatoria, que comprende poner en contacto un compuesto con un cultivo de células hospedadoras de la reivindicación 12 que expresa el ácido nucleico de p101 y de PI3K regulada por proteína G de cualquiera de las reivindicaciones 1-5, y detectar un cambio en la expresión del ácido nucleico de p101 o un cambio en la actividad del producto del ácido nucleico de p101 expresado por la célula cultivada.

26. Procedimiento de la reivindicación 25 en el que la expresión del gen de p101 se detecta midiendo los transcritos de ARNm del ácido nucleico de p101.

27. Procedimiento de la reivindicación 25 en el que la expresión del ácido nucleico de p101 se detecta midiendo la proteína de la subunidad reguladora de p101.

28. Procedimiento para la identificación de compuestos útiles en el tratamiento de enfermedades de respuesta inflamatoria, que comprende poner en contacto un compuesto con la célula hospedadora cultivada de la reivindicación 12 que ha sido manipulada genéticamente para que exprese el ácido nucleico de p101 de la PI3K regulada por proteína G, y detectar un cambio en la producción del segundo mensajero PtdIns (3, 4, 5), un cambio en la adhesión celular o un cambio en la producción de O₂.

29. Procedimiento de la reivindicación 28 en el que se determinan los niveles de PtdIns (3, 4, 5) en la célula hospedadora utilizando HPLC de intercambio iónico.

30. Célula hospedadora que ha sido manipulada genéticamente para que contenga el vector de cualquiera de las reivindicaciones 9-11 y una molécula de ácido nucleico que codifica una subunidad catalítica p120 de PI3K regulada por proteína G, que tiene una secuencia de nucleótidos que:

(a) codifica SEC ID NO: 4 o la secuencia de aminoácidos codificada por el ADNc contenido en el clon de ADNc pCMV3mycp120 tal y como está depositado en ATCC con el número de acceso 97637;

(b) codifica SEC ID NO: 14; o

ES 2 278 392 T3

- (c) hibrida en condiciones rigurosas con al menos la secuencia de nucleótidos de (a) o (b) que codifica los últimos 28 aminoácidos de la subunidad p120, o con la complementaria de la secuencia de (a) o (b) que codifica los últimos 28 aminoácidos de la subunidad p120.

- 5 31. Animal transgénico no humano que comprende la célula hospedadora de la reivindicación 30.
32. Procedimiento para la preparación de una célula hospedadora que es capaz de expresar un producto del gen de p101 que comprende la manipulación genética de una célula para que contenga la secuencia de ácido nucleico de cualquiera de las reivindicaciones 1-11.
- 10 33. Procedimiento de producción de una proteína p101 que comprende proporcionar una célula hospedadora de acuerdo con la reivindicación 12 o reivindicación 30, y expresar la secuencia nucleotídica de p101 y opcionalmente la secuencia nucleotídica de p120 en dicha célula.
- 15 34. Procedimiento de producción de una proteína p101 que comprende:
- manipular genéticamente una célula para que contenga la secuencia de ácido nucleico de cualquiera de las reivindicaciones 1-11; y expresar la secuencia nucleotídica de p101 en dicha célula.
- 20 35. Procedimiento de la reivindicación 33 ó 34, que comprende además recuperar el polipéptido p101 y opcionalmente el polipéptido p120 de la célula hospedadora o del medio de cultivo.
36. Procedimiento de la reivindicación 35, en el que la célula hospedadora comprende el vector pGEX y en donde la recuperación del polipéptido p101 comprende liberar el polipéptido p10.
- 25
- 30
- 35
- 40
- 45
- 50
- 55
- 60
- 65

ES 2 278 392 T3

1	CCGTGCGCCC	CTCAAGACTA	ATGGACCCCC	GGCTCAGGAA	TCGCACGAGG
51	CAGGCTCACA	CCCGAGGCCC	ATGGAAGTTC	CCAGGCCAGG	GGTCAAGTTG
101	GAACCGAAGC	TGCTGCCAGC	TTACACCACA	GCCACAGCAA	CTTGGGATCT
151	GAGCTGCATC	TGTGACCTAC	ACCACAGCTC	ACGGCAATGC	TGGATTCCCA
201	ACACACCAAG	TGGGGCCAGG	GATCGAACCC	GCATCCTCTT	GGACAGTAGT
251	CAGATTCATT	ACCACTGAGC	CATTGACAGG	AACTCCAGGG	GCAGGGGGGA
301	GTCTCTTGTT	TTTGGCTCCT	CCCGACACCT	GGTGAAATGG	ACCAGCGCAG
351	GCACCCCTTT	CCAGTGGCTG	TCCCAGGCGA	TGACTCAGGA	TGCAGCCAGG
401	GGCCACGACG	TGCACGGAGG	ACCGCATCCA	GCACGCCCTG	GAGCGCTGCT
451	TGCACGGGCT	CAGCCTCAGC	CGCCGCTCCA	CCTCCTGGTC	AGCTGGGCTG
501	TGTCTAAACT	GTTGGAGCCT	GCAGGAGCTG	GTCAGCAGGG	ATCCGGGCCA
551	CTTCCTTATC	CTCCTGGAGC	AGATCCTGCA	GAAAACCCGA	GAGGTCCAAG
601	AGAAGGGCAC	CTATGACCTC	CTCGCGCCCC	TGGCCCTGCT	CTTCTATTCT
651	ACTGTCCTCT	GTACGCCACA	CTTCCCGCCA	GA CTCAGATC	TCCTTCTGAA
701	AGCAGCCAGA	ACCTACCACC	GATTCTTGAC	CTGGCCGGTT	CCGTACTGCA
751	GCATCTGCCA	GGA ACTGCTC	ACCTTCATCG	ATGCTGAGCT	GAAGGCCCCA
801	GGAATCTCCT	ACCAGCGACT	GGTGAGGGCG	GAGCAGGGCC	TGTCCACAAG
851	GAGTCACCGC	AGCTCCACCG	TCACGGTGCT	CTTGGTGAAC	CCCGTGGAGG
901	TGCAGGCTGA	GTTCTTGAC	GTGGCCGACA	AGCTGAGCAC	ACCAGGGCCC
951	TCGCCGCACA	GCGCCTACAT	CACCCTGCTC	CTGCATGCCT	TCCAGGCCAC
1001	CTTTGGGGCC	CACTGTGACC	TCTCTGGTCT	GCACCGCAGG	TTGCAGTCCA
1051	AGACCCTGGC	AGAGCTCGAG	GCCATCTTCA	CGGAGACAGC	CGAGGCACAG
1101	GAGCTGGCCT	CAGGCATCGG	GGATGCAGCT	GAGGCCCGGC	AGTGGCTCAG
1151	GACCAAGCTG	CAGGCGGTGG	GAGAGAAGGC	CGGCTTCCCT	GGTGTCTTAG

FIG.1A

ES 2 278 392 T3

1201	ACACCGCCAA	ACCTGGCAAG	CTCCGCACCA	TCCCCATCCC	GGTCGCCAGG
1251	TGCTACACCT	ACAGCTGGAA	CCAGGACAGC	TTCGACATCC	TGCAGGAAAT
1301	CCTGCTCAAG	GAGCAGGAGC	TGCTCCAGCC	AGAGATCCTG	GACGACGAGG
1351	AGGACGAGGA	CGAGGAGGAC	GAGGAAGAGG	ACTTGGACGC	CGACGGCCAC
1401	TGCGCGGAGA	GGGACTCCGT	GCTCTCCACC	GGCTCGGCGG	CCTCCCACGC
1451	CTCCACGCTG	TCCCTGGCCT	CGTCCCAGGC	CTCGGGGCCC	ACGCTCTCCC
1501	GCCAGTTGCT	GACCTCCTTC	GTCTCGGGCC	TCTCGGATGG	CGTGGACAGC
1551	GGCTACATGG	AGGACATCGA	GGAGAGCGCC	TACGAGCGGC	CCCGGAGGCC
1601	TGGCGGCCAC	GAGCGCCGGG	GCCACCGCCG	GCCCGGGCAG	AAGTTCAACA
1651	GGATCTATAA	ACTCTTCAAG	AGCACCAGCC	AGATGGTGCT	GCGGAGGGAC
1701	TCGCGCAGCC	TGGAGGGCAG	CCCGGACAGC	GGCCCGCCCC	TGCGTCGGGC
1751	CGGCAGCCTC	TGCAGCCCCC	TGGACAGCCC	GACCCTGCCC	CCGTCCCGGG
1801	CCCAGGGCTC	CCGCTCGCTG	CCCCAGCCCA	AGCTCAGCCC	CCAGCTGCCC
1851	GGCTGGCTCC	TGGCCCCCGC	CTCCCGCCAC	CAGCGCCGCC	GCCCCCTTCT
1901	GAGCGGGGAC	GAGGACCCCA	AGGCTTCCAC	GCTGCGTGTC	GTGGTCTTCG
1951	GCTCGGATCG	GATCTCGGGG	AAGGTGGTCC	GGGCTTACAG	CAACCTGCGG
2001	CGGCTGGAGA	ACAACCGTCC	TCTCCTCACA	CGGTTCTTCA	AGCTACAGTT
2051	CTTCTACGTG	CCTGTCAAGC	GGAGCCGTGG	GACAGGCACC	CCCACCAGCC
2101	CAGCCCCTCG	GAGCCAGACG	CCCCCCTCC	CCACAGACGC	CCCGAGGCAC
2151	CCGGGCCCTG	CAGAGCTGGG	CGCCGCCCCC	TGGGAGGAGA	GCACCAATGA
2201	CATCTCCCAC	TACCTCGGCA	TGCTCGACCC	CTGGTACGAG	CGAAACGTCC
2251	TGGGCCTCAT	GCACCTGCCT	CCTGAAGTCC	TGTGCCAGTC	CCTGAAGGCT
2301	GAGCCCCGGC	CCCTGGAGGG	CTCCCCTGCC	CAGCTGCCCA	TCCTGGCGGA
2351	CATGCTGCTC	TACTACTGCC	GCTTCGCTGC	CCGGCCGGTG	CTGCTGCAGG
2401	TCTATCAGAC	CGAGCTGACC	TTCATACCCG	GGGAGAAGAC	GACGGAGATC
2451	TTCATCCACT	CCCTGGAGCT	GGGCCACTCT	GCTGCCACAC	GTGCCATCAA
2501	GGCTTCGGGT	CCTGGCAGCA	AGCGGCTGGG	CATCGATGGT	GACCGGGAGG

Fig. 1B

ES 2 278 392 T3

2551	CCGTCCCTCT	AACACTACAG	ATAATTTACA	GCAAGGGGGC	CATCAGCGGC
2601	CGGAGTCGCT	GGAGCAACAT	GGAAAAGCTC	TGCACCTCTG	TCAACCTCAG
2651	CAAGGCCTGC	CGGCAGCAGG	AGGAGCTAGA	CTCCAGCACA	GAGGCCCTGA
2701	CGCTAAACCT	GACAGAAGTG	GTGAAAAGAC	AGACCCCTAA	ATCCAAGAAG
2751	GGCTTTAACC	AGATCAGCAC	CTCGCAGATC	AAAGTGGACA	AGGTGCAGAT
2801	CATCGGCTCT	AACAGCTGCC	CCTTTGCCGT	GTGTCTGGAC	CAGGACGAGA
2851	GGAAGATCCT	GCAGAGTGTC	ATCAGGTGCG	AGGTCTCGCC	CTGCTACAAG
2901	CCTGAGAAGA	GCAGCCTCTG	CCCCCACCC	CAGAGGCCCT	CCTACCCGCC
2951	AGCGCCGGCC	ACGCCCGACC	TCTGCTCCCT	GCTCTGCCTG	CCCATCATGA
3001	CTTTCAGCGG	AGCTCTGCCC	TAGCCGCCAC	CCTGCACCAG	CCTGGACAGG
3051	GAGCCGGGGG	GCAGCCTCCT	CGGAGCCCCC	TCCCCAGAAG	ACTGGCGGCT
3101	GAGAGGGTCG	TGCTCCCTGT	GGAGAACAGA	GGGGCCGTGT	ACTGGGTCAG
3151	GGTCCCCTG	TGGGCCCTGC	AGCAGCAAGA	GCGGGGGCTG	CTGGGGCCTC
3201	AGGGCTCTGT	TTGGGCGAGA	AGCAGGCATT	AGGGAGAGGG	GCCTGGCCCC
3251	ACGGCTCTCA	GCTTCCTCAC	GGTAGCGGAG	AGAGGGATGG	GTGAGCTTGA
3301	CCTCAAGGCC	CTGGCCTCCA	GTGGGGGTCC	AGGATCCTTT	CTGGAAGGAA
3351	GATCCCAAGG	CGCTGGTGCT	CTGGGGTGTG	GTGTTAGGGG	CTCCCCCCCC
3401	AGCCCTGGGC	CAGGGCCCCC	CCGTTACTTT	GTCAGAGACT	TGGGGATCCT
3451	GTGTCTGGAG	GGTCAAGTCC	CCCTCCCTGG	GGGTTCAAGC	AGTGAAGTA
3501	TGGTTGCGAC	TTTTCTGACG	TTGGTGCAAT	CCCCGCCCCC	ACCTCAACCC
3551	CCCCACAAAA	AAACCCCTTC	TCTCTTTCAA	GTTCCCTGGG	TCTTCTGTGA
3601	AACAGCACTA	ACACTTGACC	TGGCTGTGCC	AGCACTTGGA	ACAGATGCTC
3651	CCTGGATCGA	GAGCCTTGGG	AGACAGGACA	AGCTTAGGTT	CGGTGGTGGC
3701	TCAGTTACCT	TCTAGCGAAA	TGAGCAGAAG	GAGGTGAATT	GGCTCCTTCG
3751	AGGCTCCCCT	ACCTGGGCAC	TAAGATGGGG	GGAATAAGGC	CGCCTTAAAG
3801	GGTTGGGGTG	ATGTCGTCTG	CAAAGCGCCT	GGCCCAGTGG	CCGGCTGGTA
3851	GCAAGGTGCG	GCCTCACCTT	CTGGGCGTCG	ACTCCCTCGT	GTGGCGGGAG

Fig. 1 C

3901	GCTAAAAGGA	TGCCCTGCCC	CCGTGATGCT	GTCATTCCCT	CCTTCCAATT
3951	CACTGATGAG	GCAGGACCCA	GACTGAGGGG	GTGAGGGGCG	CACAGTTCTA
4001	CCTTTGAAGG	AGGAAGTGCC	TTGATCAGAG	TAAGAGGAGG	GTGGCCCAGG
4051	CGCCCCAAC	CGCCCCCTCC	TCCTCTCCCA	GGTTGGCCCC	TGTGCCTCCC
4101	ACTCCCATCT	CACTCCTTGG	GCTGGCGCAC	ATCACGGGCA	CAGTCCTCCA
4151	GCCCCACAGT	TCACTGGTAC	CATGGCCCCCT	GGGTCGGTTC	GCAGAGGATG
4201	GAGGATAAGA	CTTGCCTCGA	GAACCTGGGT	CTGATGGGGA	AACCGGGTGA
4251	TGGAAATGAT	TCCGGAAGAT	TAAACCTCC	CAGGTTCAAG	TGTCGGAGAA
4301	CCGCCCCAC	AACCGGACTA	GGTTGGTAGG	GAGAGGGCAG	GGCTTGGGCC
4351	CGGGATTTGG	ACTAGGAGAG	GCGGGGGGAG	GTAACCAGAG	AAgCAAGACA
4401	GTTGTATCCC	CGCAAAAGAC	CCTTCCCCGC	CCCTCCCCTC	CTGCTCTGGC
4451	TCCATCTGCT	TCAAAGGGTC	TGGGCTTTAG	GAGCCCGTGG	TGCCCAGCGC
4501	AGCGTACTCA	GGA CTGAGA	GACGCGGACC	GTGCCAGTTC	CCACCCTGTG
4551	CCACTCCAGG	CCCCAGGGAG	GGGTTTGCAA	TATACCCTCA	ACGTTTTTGT
4601	GTGTGTGGTA	AGGTCGTCCT	AGGACCCcAA	ATGGAATTTA	ACGTTATTGT
4651	CAAATAAAAC	TTGATTTGTC	TTGGAAAAAA	AAAAAAAAAA	AA

FIG. 1D

1 MQPGATTCTE DRIQHALERC LHGLSLSRRS TWSAGLCLN CWSLQELVSR
 51 DPGHFLILLE QILQKTREVO EKGTYDLLAP LALLFYSTVL CTPHFPPDSO
 101 LLLKAARTYH RFLTWVPVYC SICQELLTFI DAELKAPGIS YQRLVRAEQG
 151 LSTRSHRSST VTVLLVNPVE VQAEFLDVAD KLSTPGPSPH SAYITLLLHA
 201 FQATFGAHCD LSGLHRRLOS KTAELEAIF TETAEAQELA SGIGDAAEAR
 251 QWLRTKLQAV GEKAGFPGVL DTAKPGKLRT IPIPVARCYT YSWNQDSFDI
 301 LQEILLKEQE LLQPEILDDE EDEDEEDEEE DLDADGHCAE RDSVLSTGSA
 351 ASHASTLSLA SSQASGPTLS RQLLTSFVSG LSDGVDSGYM EDIEESAYER
 401 PRRPGGHERR GHRRPGQKFN RIYKLFKSTS QMLRRDSRS LEGSPDSGPP
 451 LRRAGSLCSP LDSPTLPPSR AQGSRSLPQP KLSPQLPGWL LAPASRHQRR
 501 RPFLSGDEDP KASTLRVVVF GSDRISGKVV RAYSNLRRLE NNRPLLTRFF
 551 KLQFFYVPVK RSRGTGTPTS PAPRSQTPPL PTDAPRHGPG AELGAAPWEE
 601 STNDISHYLG MLDPWYERNV LGLMHLPEV LCQSLKAEPR PLEGSPAQLP
 651 ILADMLLYYC RFAARPVLLQ VYQTELTFTT GEKTTEIFIH SLELGHSAAT
 701 RAIKASGPGS KRLGIDGDRE AVPLTLQIIY SKGAISGRSR WSNMEKLCTS
 751 VNLSKACRQQ EELDSSTEAL TLNLTEVVKR QTPKSKKGFN QISTSQIKVD
 801 KVQIIGNSNC PFAVCLDQDE RKILQSVIRC EVSPCYKPEK SSLCPPPQRP
 851 SYPPAPATPD LCSLLCLPIM TFGALP*

FIG.2

ES 2 278 392 T3

1	GGCACGAGGA	ATTTGTTTTG	TTTTCAGAAA	TTAAACAAAT	GATCCTTCAG
51	CATCATCACC	TCCGCTGCTT	TATCAGGTCG	CATAGGGCAT	GGAGCTGGAG
101	AACTATGAAC	AGCCCGTGGT	GCTGAGAGAG	GACAACCGCC	GCAGGCGTCG
151	GAGGATGAAG	CCGCGCAGCA	CGGCAGCCAG	CCTGTCCTCC	ATGGAGCTCA
201	TCCCCATCGA	GTTTGTTTTG	GCCACCAGCC	AGCGCAACAC	CAAGACCCCC
251	GAAACGGCAC	TGCTGCACGT	GGCCGGCCAC	GGCAATGTGG	AGAAGATGAA
301	GGCCCAGGTG	TTGTTGCGCG	CGCTGGAGAC	GAGCGTTTCT	TGGGACTTCT
351	ACCACCGGTT	CGGCCCCGAC	CACTTCCTCC	TGGTCTTCCA	GAAGAAGGGG
401	GAGTGGTACG	AGATCTATGA	CAAGTACCAG	GTGGTGCAGA	CCCTGGACTG
451	CCTGCGCTAC	TGGGAGGTGT	TGCACCGCAG	CCCCGGGCAG	ATCCACGTGG
501	TCCAGCGGCA	CGCGCCCTCG	GAGGAGACAT	TGGCCTTCCA	GCGCCAGCTC
551	AACGCCCTCA	TCGGCTACGA	CGTCACCGAC	GTCAGCAACG	TGCATGACGA
601	TGAGCTGGAG	TTCACGCGGC	GCCGCCTGGT	CACCCCGCGC	ATGGCCGAGG
651	TGGCCGGCCG	CGACCCCAAG	CTTTACGCCA	TGCACCCCTG	GGTGACATCC
701	AAGCCCCTCC	CTGAGTACCT	TCTGAAGAAG	ATCACTAACA	ACTGCGTCTT
751	CATCGTCATT	CACCGCAGCA	CCACCAGCCA	GACCATCAAG	GTCTCGGCCG
801	ATGACACCCC	AGGCACCATC	CTCCAGAGCT	TCTTTACCAA	GATGGCCAAG
851	AAGAAATCTC	TGATGGATAT	CCCTGAAAGC	CAGAACGAAC	GGGACTTTGT
901	GCTGCGCGTC	TGCGGCCGGG	ATGAGTACCT	GGTGGGTGAG	ACGCCCATCA
951	AAAATTTCCA	GTGGGTGAGG	CAGTGCCTCA	AGAATGGGGA	GGAGATTAC
1001	CTTGTGCTGG	AACTCCTCC	AGACCCAGCC	CTGGACGAGG	TGAGGAAGGA
1051	AGAGTGGCCG	CTGGTGGATG	ACTGCACGGG	AGTCACTGGC	TACCACGAGC
1101	AGCTGACCAT	CCACGGCAAG	GACCATGAAA	GTGTGTTTAC	CGTGTCCCTG
1151	TGGGACTGTG	ACCGCAAGTT	CAGGGTCAAA	ATCAGAGGCA	TTGATATCCC

FIG.3A

ES 2 278 392 T3

1201	TGTCCTGCCC	CGGACCGCTG	ACCTCACGGT	GTTTGTGGAG	GCAAACATCC
1251	AGTATGGGCA	GCAAGTCCTT	TGCCAAAGGA	GAACCAGCCC	CAAACCCTTC
1301	ACGGAGGAGG	TGCTCTGGAA	CGTGTGGCTT	GAGTTCAGTA	TTAAAAATCAA
1351	AGACTTACCC	AAAGGGGCTC	TGCTGAACCT	CCAGATCTAC	TGCGGCAAAG
1401	CTCCAGCACT	GTCTGGCAAG	ACCTCTGCAG	AGATGCCCAG	TCCCGAGTCC
1451	AAAGGCAAAG	CTCAGCTTCT	GTACTATGTC	AACCTATTGC	TGATAGACCA
1501	CCGCTTCCTC	CTGCGCCATG	GCGAGTATGT	GCTCCACATG	TGGCAGTTAT
1551	CCGGGAAGGG	GGAAGACCAA	GGGAGCTTCA	ATGCCGACAA	GCTCACGTGG
1601	GGAACCAACC	CGGACAAGGA	GGACTCAATG	TCCATCTCCA	TTCTTCTGGA
1651	CAATTACTGC	CACCCCATAG	CCTTGCCTAA	GCATCGGCCT	ACCCCTGACC
1701	CAGAAGGGGA	CCGGGTTCGG	GCAGAAATGC	CCAATCAGCT	TCGGAAGCAA
1751	CTGGAGGCAA	TCATAGCCAC	GGATCCGCTT	AACCCACTCA	CAGCTGAAGA
1801	CAAAGAACTG	CTCTGGCATT	TCAGATATGA	AAGCCTGAAG	GATCCCAAAG
1851	CGTATCCTAA	GCTCTTTAGC	TCGGTGAAAT	GGGGACAGCA	AGAAATTGTG
1901	GCCAAAACAT	ACCAATTATT	AGCCAAAAGG	GAGGTCTGGG	ATCAGAGTGC
1951	TTTGGATGTG	GGGTAAACCA	TGCAGCTCCT	GGACTGCAAC	TTCTCGGATG
2001	AAAACGTGAG	AGCCATTGCA	GTCCAGAAAC	TGGAGAGCTT	GGAGGATGAT
2051	GACGTGCTCC	ATTACCTGCT	CCAGCTGGTC	CAGGCTGTGA	AATTTGAACC
2101	ATACCATGAC	AGTGCCCTAG	CCAGATTTCT	GCTGAAGCGT	GGTTTAAGAA
2151	ACAAGAGAAT	TGGTCACTTC	TTGTTTTGGT	TCTTGAGAAG	TGAGATTGCC
2201	CAGTCTAGGC	ACTATCAGCA	GAGGTTTGCA	GTGATCCTGG	AAGCCTACCT
2251	GAGGGGCTGT	GGCACAGCCA	TGCTGCACGA	CTTCACCCAG	CAAGTCCAAG
2301	TAATTGACAT	GTTACAAAAA	GTCACCATTG	ACATTAAATC	GCTCTCTGCT
2351	GAAAAGTATG	ACGTCAGTTC	CCAAGTTATT	TCCCACTTA	AGCAAAAGCT
2401	TGAAAACCTA	CAGAATTTGA	ATCTCCCCCA	AAGCTTTAGA	GTTCCCTATG
2451	ATCCTGGACT	GAAAGCCGGG	GCACTGGTGA	TCGAAAATG	TAAAGTGATG
2501	GCCTCCAAGA	AGAAGCCCTT	GTGGCTTGAG	TTTAAATGTG	CCGATCCTAC

FIG.3B

ES 2 278 392 T3

2551	GGCTCTATCA	AATGAAACAA	TTGGAATTAT	CTTTAAACAC	GGTGACGATC
2601	TGCGCCAAGA	CATGCTTATT	TTACAGATTC	TACGAATCAT	GGAGTCCATT
2651	TGGGAGACCG	AATCTTTGGA	TCTGTGCCTC	CTGCCATATG	GCTGCATTTT
2701	AACTGGTGAC	AAAATAGGAA	TGATCGAGAT	CGTGAAGGAC	GCCACGACAA
2751	TCGCCAAAAT	TCAGCAAAGC	ACAGTGGGCA	ACACGGGTGC	CTTTAAAGAT
2801	GAAGTCCTGA	GTCACTGGCT	CAAAGAAAAA	TGCCCTATTG	AAGAAAAGTT
2851	TCAGGCAGCT	GTGGAGAGAT	TTGTTTATTC	CTGTGCCGGC	TACTGTGTGG
2901	CAACCTTTGT	TCTCGGAATA	GGCGACAGAC	ACAATGACAA	TATTATGATC
2951	TCAGAAACAG	GAAATCTATT	TCATATTGAT	TTCGGACACA	TTCTTGGGAA
3001	TTACAAAAGT	TTCCTGGGCA	TTAATAAAGA	GAGGGTGCCA	TTTGTGCTAA
3051	CCCCAGACTT	CCTGTTTGTG	ATGGGGACTT	CTGGAAAGAA	GACAAGTCTA
3101	CACTTCCAGA	AATTTTCAGGA	TGTCTGCGTC	AAGGCTTACC	TAGCCCTTCG
3151	TCATCACACA	AACCTACTGA	TCATCCTCTT	CTCCATGATG	CTGATGACAG
3201	GAATGCCCCA	GTTAACCAGC	AAAGAAGACA	TTGAATACAT	TCGGGATGCC
3251	CTCACAGTGG	GCAAAAAGTGA	GGAGGATGCT	AAAAAGTATT	TTCTGGATCA
3301	GATTGAAGTT	TGCAGAGACA	AAGGATGGAC	CGTGCAGTTT	AACTGGTTCT
3351	TACATCTTGT	TCTTGGCATC	AAACAAGGGG	AGAAGCATCC	CGCAT'AAAAC
3401	TTTGGGCCAA	GAGTTAAAC	CCAAGTTATT	GTCCTAATGC	TTTACGTCAG
3451	CAGGACAATC	ACCGAACTTG	ATGTCATGTA	GTGGGACATT	ATGAAAGCTG
3501	GCACTTGAGA	AATATAGCTC	TTCCCCTAAC	TGAACTCTTC	ACTGGAGAAA
3551	AACCTTGGCA	TGTTTAAGTA	ATGTTCAGTG	TTAGGCTTAT	TTGCATGTTT
3601	GTTTTTTCTC	ATGTGCCCCC	TCAGTCATGT	TGGAGACTGT	TCTAAATTTA
3651	AGTGGCCTAA	TGACCTCTGA	AGTTTCAACT	TTCTTGGTAC	TGAGTGCTTC
3701	TGAAATTCTT	TACAATAATT	GGTAACATCT	ATTGTCAGCT	GGGTATCCTC
3751	TCAATTTTGG	TTATCCTTGG	GTTTCTCAAA	CTCCTTACAG	GAAAAAAAAA
3801	AAAAAAAA				

FIG.3C

ES 2 278 392 T3

1 MELENYEQPV VLREDNRRRR RRMKPRSTAA SLSSMELIPI EFVLATSQRN
 51 TKTPETALLH VAGHGNVEKM KAQVLLRALE TSVSWDFYHR FGPDHFLLVF
 101 QKKGEWYEIY DKYQVQTLD CLRYWEVLHR SPGQIHVVQR HAPSEETLAF
 151 QRQLNALIGY DVTDVSNVHD DELEFTRRRL VTPRMAEVAG RDPKLYAWHP
 201 WVTSKPLPEY LLKKITNMCV FIVIHRSITS QTIKVSADDT PGTILQSFFT
 251 KMAKKKSLMD IPESQNERDF VLRVCGRDEY LVGETPIKNF QWVRQCLKNG
 301 EEIHLVLDTP PDPALDEVK EEWPLVDDCT GVTGYHEQLT IHGKDHESVF
 351 TVSLWDCDRK FRVKIRGIDI PVLPRADLT VFVEANIQYG QQVLCQRRTS
 401 PKPFTEEVW NWLEFSIKI KDLPKGALLN LQIYCGKAPA LSGKTSAMP
 451 SPESKGKAQL LYYVNLILLID HRFLLRHGEY VLHMWQLSGK GEDQGSFNAD
 501 KLTSGTNPOK EDSMSISILL DNYCHPIALP KHRPTDPEG DRVRAEMPNO
 551 LRKQLEAIIA TDPNLPLTAE DKELLWHFRY ESLKDPKAYP KLFSSVKWGO
 601 QEIVAKTYQL LAKREWDQS ALDVGLTMQL LDCNFSDENV RAIAVQKLES
 651 LEDDVLHYL LQLVQAVKFE PYHDSALARF LLKRGLRNKR IGHFLFWFLR
 701 SEIAQSRHYQ QRFVILEAY LRCGTAMLH DFTQQVQVID MLQKVTIDIK
 751 SLSAEKYDVS SQVISQLKOK LENLQNLNP QSFRVPYDPG LKAGALVIEK
 801 CKVMASKKKP LWLEFKCADP TALSNETIGI IFKHGDDLRO DMLILOILRI
 851 MESIWETESL DLCLLPYGGI STGDKIGMIE IVKDATTIAK IQOSTVGNTG
 901 AFKDEVLSHW LKEKCPIEEK FQAAVERFVY SCAGYCVATF VLGIGDRHND
 951 NIMISETGNL FHIDFGHILG NYKSFLGINK ERVPFVLTPD FLFWMTSGK
 1001 KTSLHFKQFQ DVCVKAYLAL RHHTNLLIIL FSMMLMTGMP QLTSKEDIY
 1051 IRDALTVGKS EEDAKKYFLD QIEVCROKGW TVQFNWFLHL VLGIKQGEKH
 1101 PA*

FIG.4

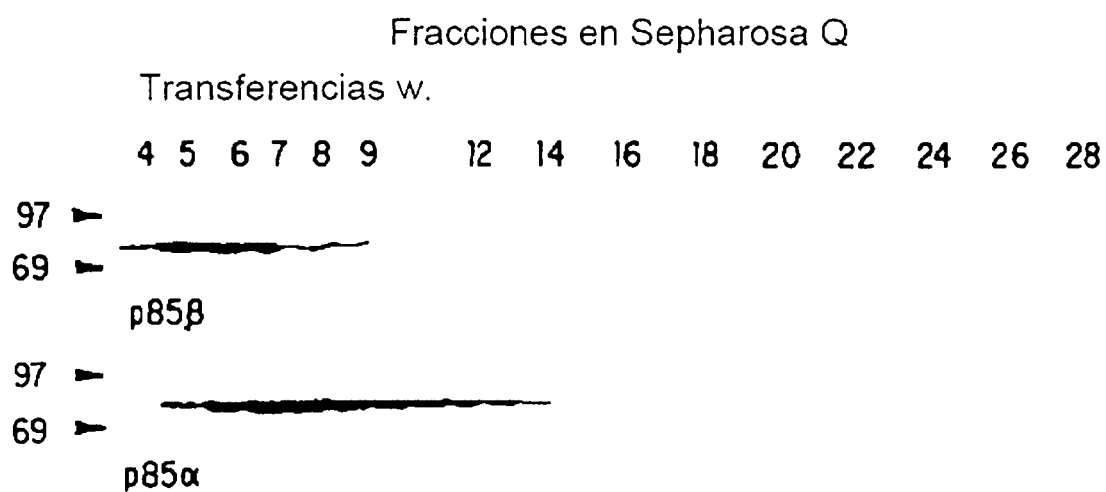


FIG. 5A

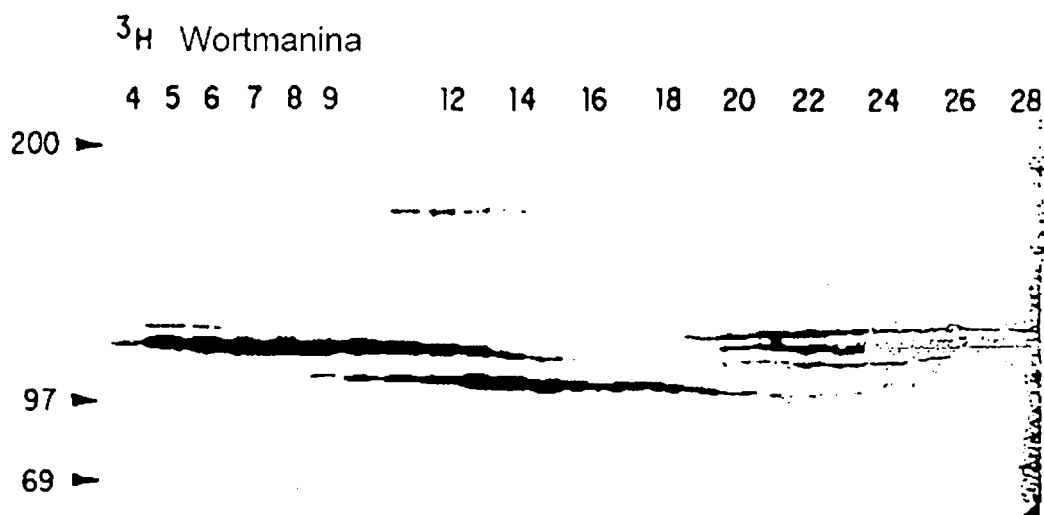


FIG. 5B

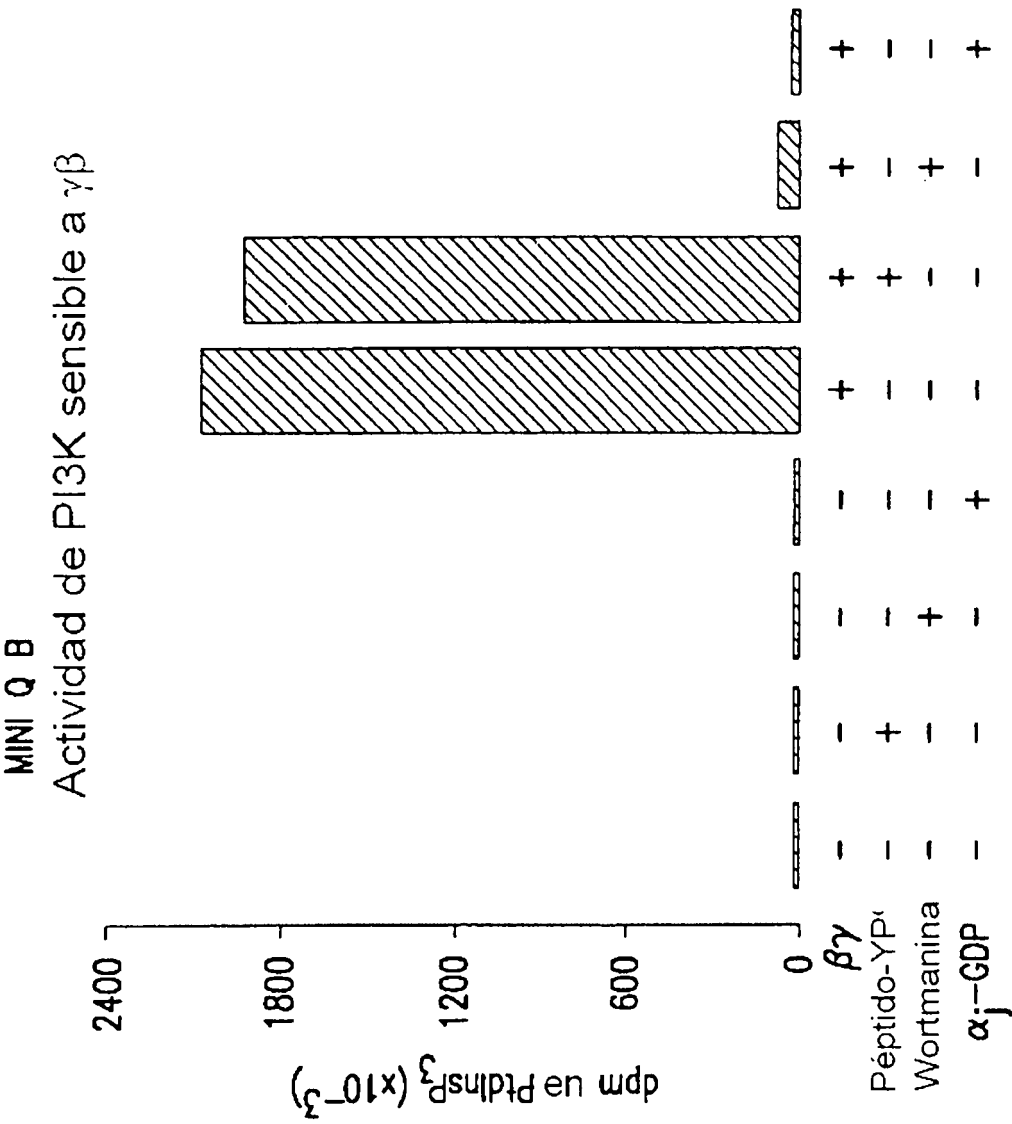


FIG.6

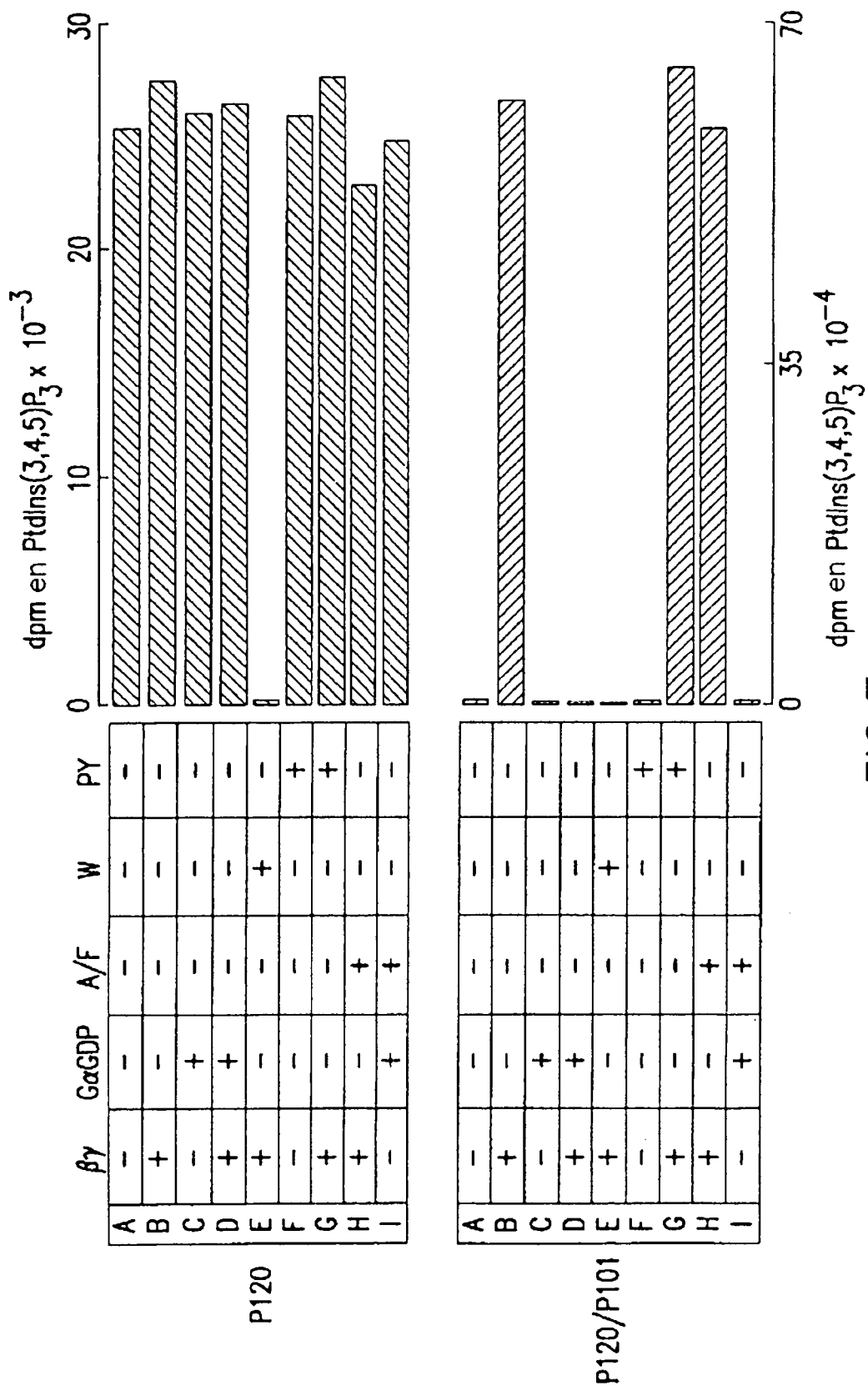


FIG.7

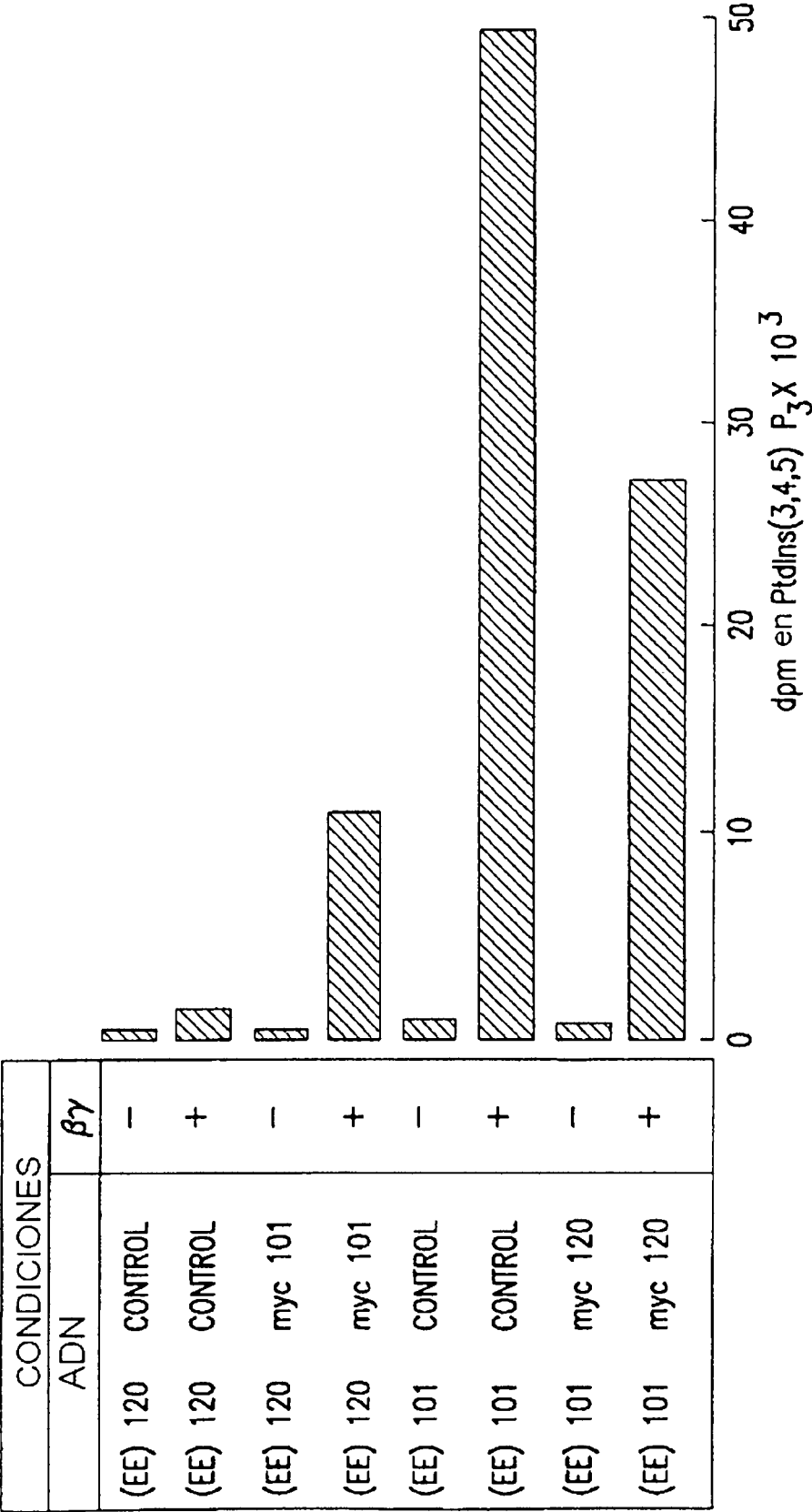


FIG.8

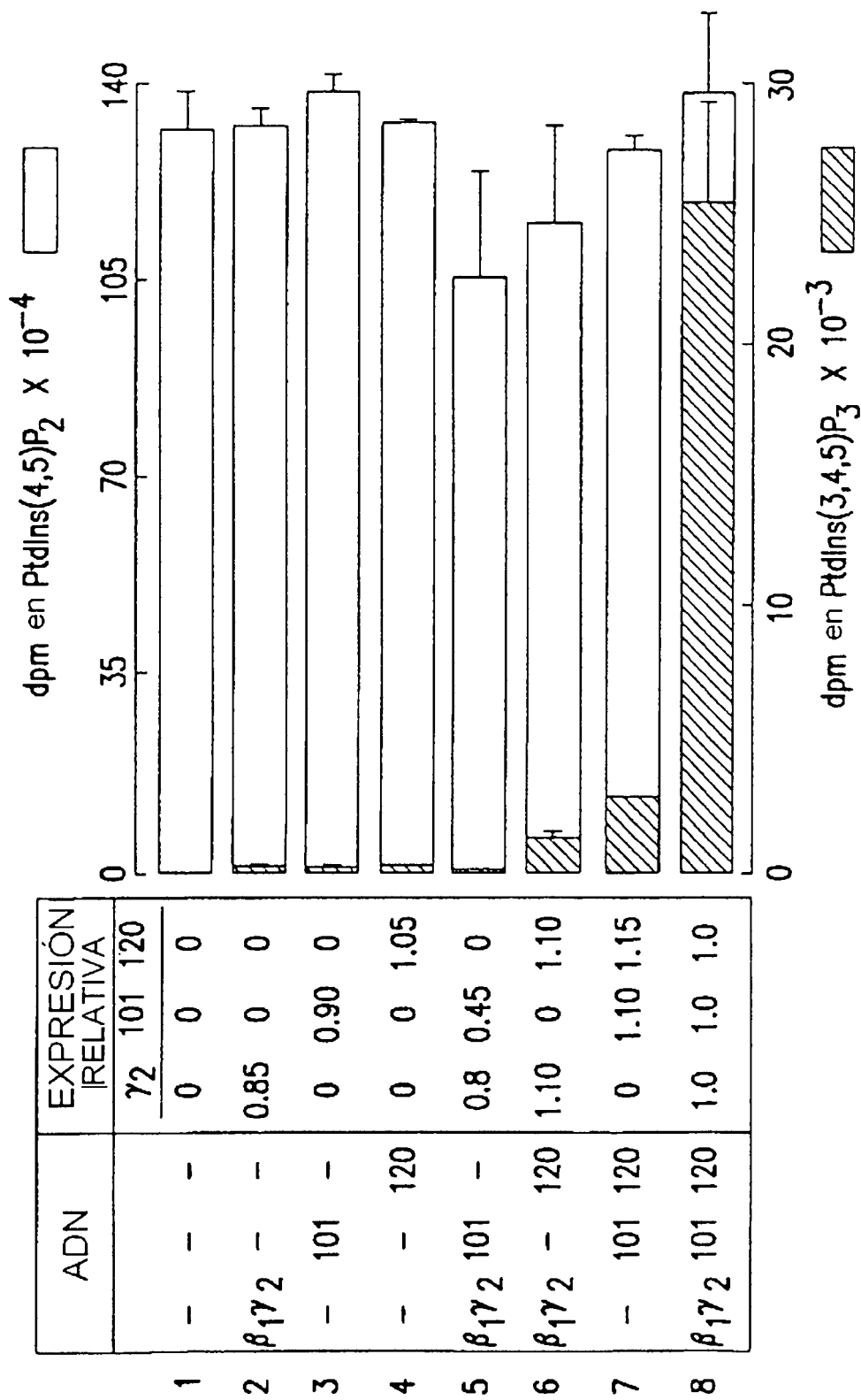


FIG.9

AAGCTTACTCAGGCGCATCTGCTACAGCATTTTATCTTCCCAAGAGCCCCATGAGGGGATGACCCAGGATGCAGCCAGGGG
 CCACGACATGCACGGAGGACCGCATCCAGCATGCCCTGGAAACGCTGCCATGGACTCAGCCTCAGCCGCCGCTCCACC
 TCCGGTCAGCTGGGCTGTGCTGAACCTGCTGGAGCTGCAGGAGCTGGTCAGCAGGACCCGGGCCACTTCTTTATCCT
 CCTTGAGCAGATCCTGCAGAAACCCGAGAGGTCCAGGAGAAGGCACCTACGACCTGCTACCCGCTGGCCCTGCTCT
 TCTATTCCACTGTTCTTTGTACACCACACTTCCCACCAGACTCGGATCTCCTTCTGAAGGCAGCCAGCACCACACCGG
 TTCTGACCTGGCCGTCTTCTTACTGCAGCATCTGCCAGGAGCTGCTCACCTTCTTGTGCTGAACCTCAAGGCCCCAGG
 GATCTCTACCAGAGACTGGTGAGGGCTGAGCAGGGCTGCCCATCAGGAGTCACCGAGCTCCACCGTCACCGTGTCTGC
 TGCTGAACCCAGTGGAAAGTGCAGGCCGAGTTCCTTGCTGTAGCCAATAAGCTGAGTACGCCGGACACTCGCCTCACAGT
 GCCTACACCACTGCTCCTGCACGCCCTCCAGGCCACCTTTGGGGCCCACTGTGACGTCCCGGGCTGCACCTGCAGGCT
 ACAGGCCAAGACCCCTGGCAGAGCTTGAGGACATCTTCACGGAGACCCGAGAGGCACAGGAGCTGGCATCTGGCATCGGGG
 ATGCTGCAGAGGCCCGGGTGGCTCAGGACCAAGCTGCAGGGGCTGGGAGAAAAAGCTGGCTTCCCTGGGGTGTAGAC
 ACTGCAAAACCCAGGGAAGCTCCACACCATCCCCATCCCTGTGCCAGGTGCTACACCTACAGCTGGAGCCAGGACAGCTT
 TGACATCCTGCAGGAAATCCTGCTCAAGGAACAGGAGCTACTCCAGCCAGGGATCCTGGGAGATGATGAAGAGGGAAG
 AGGAGGAGGAGGAGTGGAGGAGGACTTGGAAACTGACGGGCACTGTGCCGAGAGAGATTCCCTGCTCTCCACAGCTCT
 TTGGCGTCCCATGACTCCACCTTGTCCCCTTGCACTCTCCAGGCCCTCGGGGCCGGCCCTCTCGCGCCATCTGCTGACTTC
 CTTTGTCTCAGGCCCTCTCTGATGGCATGGACAGGGCTACGTGGAGGACAGCGAGGAGAGCTCCTCCGAGTGGCCTTGGA
 GCGGTGGCAGCCAGGAACGCCGAGGCCACCCAGGCCCTGGGCAGAACTTCAATCAGGATCTATAAATCTTCAAGAGCACC
 AGCCAGCTGGTACTGCGGAGGACTCTCGGAGCCTGGAGGGCAGCTCGGACACGGCCCTGCCCTGAGGCGGCGAGGAG
 CCTCTGCAGCCCCCTGGACGAACCAATACACCCCTTCCCGGCCCAAGGCTCCCGCTCCCTGCCCGAGCCCAACTCG
 GTACCCAGCTGCCCAGCTGGCTTCTGGCCCCCTGCTTACGCCCCCAGGGCCCGCCCCCTTCTGAGTGGAGATGAGGAT
 CCCAAGGCTTCCACGCTACGTGTTGTGGTCTTTGGCTCCGATCGGATTTTCAGGGAAGGTGGCTCGGGCGTACAGCAACCT
 TCGGCGGCTGGAGAACAAATCGCCCCACTCCTCACACGGTCTTCAAACTTCAGTTCCTACGTGCCCTGTGAAGCGAAGTC
 ATGGGACCAGCCCTGGTGCCTGTCCACCCCTCGGAGCCAGACGCCCTCACCCCGACAGACTCCCTAGGCAGCCAGC
 CCTGGAGAGCTGGGCACCAACCCATGGGAGGAGAGCAACATGACATCTCCCACCTACCTCGGCATGCTGGACCCCTGGTA
 TGAGCGCAATGTAATGAGGCTCATGACCTGCCCCCTGAAGTCTGTGCCAGCAGTCCCTGAAGGCTGAAGCCCGAGGCC
 TGGAGGGCTCCCCAACCCAGCTGCCCATCTTGGCTGACATGCTACTACTGCGGCTTTGCGGCCAGACCGGTGCTG

FIG.10A

CTGCAACTCTATCAGACCGAGCTGACCTTCATCACTGGGAGAAAGCAGACAGAGATCTTCATCCACTCCTTGGAGCTGGG
 TCACCTCCGCTGCCACACGTGCCATCAAGGCGTCAGGTCTTGGCAGCAAGCGGCTGGGCATCGATGGCGACCGGGAGGCTG
 TTCCCTTAACACTACAGATTATTTACAGCAAGGGGGCCATCAGTGGACGAAGTCGCTGGAGCAACCTGGAGAAGGTCTGT
 ACCTCCGTGAACCTCAACAAGGCCTGCCGGAAGCAGGAGGAGCTGGATCCAGCATGGAGGCCCTGACGCTAAACCTGAC
 AGAAGTGGTGAAAGGCAGAACTCCAAATCCAAGAAGGGCTTTAACCCAGATTAGCACATGCGAGATCAAAGTGGACAAGG
 TGCAGATCATCGGCTCCAACAGCTGCCCTTTGCTGTGTGCTGGACCAAGGATGAGAGAAAGATCCTGCGAGAGTGTAGTC
 AGATGTGAGGTCTCACCGTGTCTAACGCCAGAGAAGAGCGACCTCTCCTCACCACCCAGACGCCCTCCTGACCTGCCGGC
 CCAGGCCGACCTGATCTCTGCTCCCCTCTCTGCTGCCCATCATGACTTTCAGTGGAGCTCTGCCCTAGTGTGGGCCCCA
 GCGCCAGACTGGACAGAAGCCCTGGGGCAACCTCCTCGGCCACCCCTCCAGGACAGTCCCCTCTCTGTGGAGAACTGAATG
 GCCCTGTGCAGAGCCATAGTCCCACTGTGGTCTCTGCAATGAGCAGGGGCTGGGAGTAGAGGGTTTCTGGGGCCCTCAGGG
 TTCTGGGAAAGCAACAGCTATCAGAGAGAGAAGGGCCAGACCCCATAGCCTCTTAGATTCTGCGCAGTAGAAGGAGAAGG
 ATGGGTAAATTGACCTCTGAAGTCCCAGCCATTAGCATGGTCTAGGATCCTTTCTAGAAGGAAGATCTGAGGCTCTGGT
 GCTCAGGGGGATGGCTTGGGCTTTCTCTCAACCTTGGCTGAGCCTACCCCTTACTTTGCCAAAGACTTGAGGACCCCTG
 TATGTCTGGAGTTCAGTCCCCTCCTCTGTGGGGCTCAGGTGATTGAAATGTGGATGAAACATTTCTCTACTTCAAGACCA
 CCTCTCCCTGCCAAACACCACACACATGGCATGCATGTACGCACATGCGCACCG

FIG. 10B

MQPGATTCTEDRIQHALLERCLHGLSLRRRSTWSAGLCLNCWSLQELVSRDPGHFLILLEQILQKTREVQKGTYDLLTP
 LALLFYSTVLCPTHFPPDSDLLLKAASTYHRFLTWPVPYCSICQELLTFIDAEKAPGISYQRLVRAEQGLPIRSHRSST
 VTVLLNPVEVQAEFLAVANKLSTPGHSPHSAYTTLLHAFQATFGAHCDVPGLHCRLOAKTLAELEDIFTETAQAELA
 SGIGDAAEARWLRTKLQAVGEKAGFPGVLDTAKPGKLHTIPIPVARCYTYSWSQDSFDILQEILLKEQELLQPGILGDD
 EEEEEEEVEEDLETDGHCAERDSSLSTSLASHDSTLSSASSQASGPALSRHLLTSFVSGLSGMDSGYVEDSEESS
 EWPWRRGSQERRGHRPQKFIRIYKLFKSTSQLVLRDRSRLEGSSTALPLRRAGSLCSPLDEPVSPPSRAQRSRLP
 QPKLGTQLPSWLLAPASRPQRRRPFLSGDEDPKASTLRVVVFGSDRISGKVARAYSNLRRLENNRPLLTRFFKLQFFYVP
 VKRSHGTSPGACPPRSQTPSPPTDSPRHASPGELGTTPEESTNDISHYLGMLDPWYERNVLGLMHLPPPEVLCQQSLKA
 EAQALEGSPTQLPILADMLLYYCRFAARPVLLQLYQTELTFITGEKTTEIFIHSELELGHSAATRAIKASGPGSKRLGIDG
 DREAVPLTLQIIYSKGAISSGRSRWSNLEKVCSTSVNLNKACRKQEELDSSMEALTLNLTEVVKRONSKSKGFNQISTSQI
 KVDKVQIIGNSNCPFAVCLDQDERKILQSVVRCEVSPCYKPEKSDLSSPPQTTPDLPAAQAPDLCSLCLPIMTFSGALP

FIG.11

TTGAAATTGTTTGTTCAAAATTAAACAAATGATCCTTCAGCATCATCGCCTCCGCTGCTTTATCAGGTCGCATAG
 GGATGGAGCTGGAGAACTATAACAGCCCGTGGTGTCTGAGAGAGGACAACCTGCCGAAGGCGCGGAGGATGAAGCGCGG
 CAGTGTGCGGCCAGCCTGTCTCTCATGGAGCTCATCCCATCGAGTTCGTGCTGCCACACAGCCAGCGCAAAATGCAAGA
 GCCCCGAACGGCGCTGCTGACGTGGCCGACGGCAACGTGGAGCAGATGAAGGCCAGGTGTTGGCTGCGAGCGCTG
 GAGACCAGCGTGGCGGGGACTTCTACCACGGCTGGGACCGCATCACTTCTCTGCTCTATCAGAAGAAGGGGCAGTGTG
 GTACGAGATCTACGACAAGTACCAGGTGGTGCAGACTCTGGACTGCTGCGCTACTTGAAGGCCACGACCCGGAGCCCGG
 GCCAGATCCACCTGGTGCAGCGGACCCGCCCTCCGAGGAGTCCCAAGCCTCCAGCGGACGCTACGGCGCTGATTGGC
 TATGACGTCACTGACGTACGCAACGTGCACGACGATGAGCTGGAGTTCACGGCGCGTGGCTTGGTGACCCCGCGCATGGC
 GGAGGTGGCCAGCCGACCCCAAGCTCTACGCCATGCACCCGTGGGTGACGTCCAAGCCCCCTCCGGAGTACCTGTGGA
 AGAAGATTGCCAACAACTGCATCTTCATCGTCAATCACCGCAGCACCAAGCCAGCCAGACCATTAAGGTCTACCCGACGAC
 ACCCCGGCGCCATCCTGCAGAGCTTCTTACCAGAATGGCCAGAAGAAATCTCTGATGGATATTCCTCGAAAGCCAAAG
 CGAACAGGATTTTGTGCTGGCGTCTGTGGCCGGATGAGTACCTGGTGGCGAAACGCCCATCAAAAACCTCCAGTGGG
 TGAGGCACTGCCTCAAGAACGGAGAAGAGATTACGTGGTACTGGACACGCTCCAGACCCGGCCCTAGACGAGGTGAGG
 AAGGAAGAGTGGCCGCTGGTGGACGACTGCACGGGAGTCAACGGCTACCATGAGCAGCTTACCATCCACGGCAAGGACCA
 CGAGAGTGTTCACCGTGTCCCTGTGGGACTGCGACCGCAAGTTCAGGGTCAAGATCAGAGGCAATTGATATCCCCGTCC
 TGCTCGGAACACCGACCTCACAGTTTGTAGAGGCAACATCCAGCATGGCAACAAGTCTTTGCCAAAGGAGAACC
 AGCCCCAAACCCCTTACAGAGGAGGTGCTGTGGAATGTGTGGCTTGAGTTCAGTATCAAAATCAAAGACTTGCCCCAAAGG
 GGCTCTACTGAACCTCCAGATCTACTGCGGTAAGCTCCAGCAGTGTCCAGCAAGGCTCTGCAGAGTCCCCCAGTCTTG
 AGTCCAAGGGCAAAGTTCGGCTTCTCTATTATGTGAACCTGCTGCTGATAGACCACCGTTTCTCTGCGCGGTGGAGAA
 TACGTCTCTCCACATGTGGCAGATATCTGGGAAGGGAGAAGACCAGGAAGCTTCAATGCTGACAAACTCAGTCTGCAAC
 TAACCCAGACAAGGAGAACTCAA TGTCCATCTCCATCTCTCTGGACAAATTAAGTCCACCCGATAGCCCTGCCTAAGCATC
 AGCCACCCCTGACCCGGAAAGGGACCGGGTTCGAGCAGAAATGCCCAACAGCTTCGCAAGCAATGGAGGGGATCATTA
 GCCACTGATCCACTTAACCCCTCACAGCAGAGGACAAAGAAATGCTCTGGCATTTTAGATACGAAGCCCTTAAGCACCC
 AAAAGCATATCCTAAGCTATTTAGTTCAGTGAAATGGGACAGCAAGAAATTTGGCCAAAACATACCAATTTGTGGCCA
 GAAGGGAAGTCTGGGATCAAAGTGTCTGGATTTGGGTAAACAAATGCAGTCTCTGGACTGCAACTTCTCAGATGAAAT
 GTAAGAGCCATTGCAGTTCAGAACTGGAGAGCTTGGAGGACGATGATGTTCTGCAATTAACCTTCTACAAATTTGTTCCAGG
 TGTGAAATTTGAACCATACCATGATAGCGCCCTTGCCAGATTTCTGCTGAAGCTGGTTTAAGAAACAAAAGAAATTTGGTC
 ACTTTTGTGTTTGGTCTTGAGAAGTGAGATAGCCCAAGTCCAGACACTATCAGCAGAGGTTGCTGTGATTTCTGGAAGCC

FIG. 12A

TATCTGAGGGGCTGTGGCAGGCCATGCTGCACGACTTTTACCCACAAGTCCAAGTAATCGAGATGTTACAAAAAGTCAC
 CCTTGATATTAAATCGCTCTCTGCTGAAAAAGTATGACGTGAGTCCCAAGTTATTTACAACTTAACAAAAAGCTTGAAA
 ACCTGCAGAAATCTCAACTCCCGAAAGCTTTAGAGTTCATATGATCTCGACTGAAAGCAGGAGCGCTGGCAATTTGAA
 AAATGTAAGTAATGGCTCCCAAGAAAAAACCACTATGGCTTGAGTTAAATGTCCGATCCCTACAGCCCTATCAAAATGA
 AACAAATTGGAAATATCTTTAAACATGGTGATGATCTGCGCCAAGACATGCTTATTTTACAGATTCACGAATCATGGAGT
 CTATTTGGGAGACTGAATCTTTGGATCTATGCCCTCCGACATATGGTTGCAATTTCAACTGGTGACAAAAATAGGAAATGATC
 GAGATTGTAAAGACGCCACGACAATTGCCAAAAATTCAGCAAGCACAGTGGGCAACACGGGAGCATTTAAAGATGAAGT
 CCTGAATCACTGGCTCAAVGAAAAATCCCTACTGAAGAAAGTTTCAGGCAGCAGTGGAGAGATTGTTTATTCCTGTG
 CAGGCTACTGTGTGGCAACCTTTGTCTTGGAAATAGGCGACAGACACAATGACAAATATATGATCACCGAGACAGGAAC
 CTATTTCATATTGACTTCGGGCACATTCCTGGGAATTACAAAAGTTTCTCGGCATTTAAATAAGAGAGAGTGCCTATTGT
 GCTAACCCCTGACTTCCTCTTTGTGATGGGAACCTCTGGAAAGAGACAAGCCACACTTCAGAAAATTCAGGACATCT
 GTGTTAAGGCTTATCTAGCCCTTCGTCAACACAAACCTACTGATCATCTCTGTTCTCCATGATGCTGATGACAGGAATG
 CCCAGTTAACAGCAAAGAAGACATTGAATATATCCGGGATGCCCTCACAGTGGGAAAAATGAGGAGGATGCTAAAAA
 GTATTTCTTGATCAGATCGAAGTTTGCAGAGACAAAGGATGGACTGTGCAGTTTAAATGGTTCTACATCTGTCTCTTG
 GCATCAAAACAAGGAGAGAAACATTCAGCCTAATACTTTAGGCTAGAAATCAAAAACAAGTTAGTGTCTATGGTTTAAAT
 AGCATAGCAATCATCGAACTTGGATTTCAAATGCAATAGACATTTGGAAGCTGGCAATTCAGAAGTATAGCTCTTTTCC
 TACCTGAACCTCTCCCTGGAGAAAAGATGTGGCATTTGCTGATTTGGTTAAGCAATGTCAGTGTCTAGGATTTATTG
 CAGTTTGGTTTTTCTCATTTGTCTGTGGCATTTGGAGAAATATCTTGGTTTAAACAGACTAATGACTTCCTTATTGTCC
 CTGATAATTTTGACTATCTTACTATTGAGTGTCTCTGGAAAATCTTGGAAATAATTGATGACATCTATTTTCATCTGGGTT
 TAGTCTCAATTTTGGTTATCTTTGTGTTCCCTCAAGCTCTTAAAGAAAAAGATGTAATCGTTGTAACTTTGTCTCATTC
 CTTAAATGATGCTTCCAAACATCTCCTTAGTGTCTGCAGGTGTTAGTGGTGTGCTAAAAGCAAGGAAAGCGAGTTAGTCT
 TTTCAAGTGTCTTTTGGCAATTCATTTTGTCAATGTAACCTGAGACACACAAACACAGCAGGAGAAATCTAAACCGTT
 GTGCCTTGACCTTCCTCTGCTGGTCTTGTTCCAGGGTTATGAATAAGAAAAATAGAGATGAGACTTTTGTGTCAACTC
 TGCCACAAAGAGTGAATATCTAGTATGATAGTCTTCTCCAGCATGGCAGCAGGAAGTAACACAGGGCCCTTT
 TTATGCCCTGACATTTCTCCCTTCCCTTCTCCCTGCTCCCTTTTTCATCAATTTGCAATGCTCCCACTCTTTACAGA
 CTTGTGAAATCTTCAAGAACACCTTTACTCTATAACTCAAAAAATAGTTGAAAAATTAATTCTCTCAAGGATTTIAGA
 ATCTTAGGTACTTATTTGTAAAGATGTTTAGTGACTTTTTTTTCAAGTATCTATAAAGGAGGCAGATCTTAGAAAAATG
 AATTAGTTTCCAAATGCCCTTAATTTTAAACTTTGGCCTGAACAGTTTTTCTTTTCTTAAATGGAAGAAGATATTAAATA

FIG. 12B

TCTTAAAAATATCCAAAGTTAGGAAGAACACTACTTGGCCTTATCCATTTCCATTTAAAGGACTTTTAAACTTTTGACACA
 GTCCTTCAGATTTCCTGAAATCCTTGAAATATCTTACTTTTAAAAATATTTTCACTCTGAAATATCTCGTTATTTATTG
 GAGGTATTGTTTAAACCTTAGATAGACCATTAAATATTATATAAAATATTTTGTAAATACCTGTAGYTAATACATTACATA
 GAAAAAACTATGTTAACAGTGCTCTGTTTAAGTAAATCAGATATAAATATAATACTTAAATTTTTTAAATTTAAAAATA
 GATACCTGTTTGACTTTGAGGTAGTCCAGACCTTTCTTTTTTTTTTTTAAATGTGTGCAAAAGCCCCAAAGGTTTC
 CTAAGCCTGGCTGCAAGAAGAAATCAACAGGGACACTTTTTTAAAAACACTCTTATCAGCCTGGGCCAACACAGTGAGACTC
 CATCTCTTAAAAAAAATTTAGCTGGGTATAGTGGTATGTGCCTGTAGTCCCAGGTACTCAGGAGGCTGASGCAGGAGGA
 TTGCCTGAGCCAGGAGGTGGAAACTGCAGAGAGTCATGATCATGTCTTACACTCCAGCTGGATACAGAGCGAGACC
 CTGTCTCAAAAAAAAAAAAAAAAAAAAAAGTCGACCCGAG

FIG.12C

MELENYKQPVVLRDNCRRRRRMKPRSAASLSSMELPIEFVLPTSQRCKKSPETALLHVAGHGNVEQMKAQVWLRalet
SVAADFYHRLGPHHFLLL YOKKGQWYEIYDKYQVVQTLDCRLRYwKATHRSPGQIHLVQRHPPSEESQAFQRQLTALIGYD
VTDVSNVHDDLEFTRRGLVTPRMAEVASRDPKLYAMHPWVTSKPLPEYLwKKIANNCIFIVIHRSSTTSQTIKVSPPDTP
GAILQSFFTKMAKKKSLMDIPESQSEQDFVLRVCGRDEYLVGETPIKNFQwVRHCLKNGEETHVVLDTPPDPALDEVrKE
EWPLVDDCTGVTGYHEQLTIHGKDHEsvFTVSLWDCORKFRVKIRGIDIPVLPRNTDLTVFVEANIQHGGQVLCQRRTSP
KPFTEEVLNwVwLEFSIKIKOLPKGALLNLQIYCGKAPALSSKASAESSESCKGVRLLYYVNLILLIDHRFLLRRGEYV
LHMWQISGKGEDQGSFNADKLTSATNPDKENSMSISILLDNYCHPIALPKHQTPDPEGDRVRAEMPnQLRKQLEAIAT
DPLNPLTAEDKELLwHFRYESLKHpkAYPKLFSSVKwGQQEIVAKTYQLLARREVWdQSALDVGLTMQLLDCNFSDENVR
AIAVQKLESLEDDDLHYLLQLVQAVKFEPYHDSALARFLLKRGLRNKRIGHFLFWFLRSEIAQSRHYQQRFavILEAYL
RGGGTAMLHdFTQQVQVIEMLQKVTLDIKSLSAEKYDVSSQVISQLKXKLENLQNSQLPESFRVPYDPGLKAGALAIeKc
KVMASKKKPLWLEFKCADPTALSNETIGIIFKHGDDLrQDMLILQILRIMESIwETESLDCLLPYGCISTGDKIGMIEI
VKDATTIAKIQQSTVGNTGAFKDEVLNHWLKEKSPTEEFQAAVERFVYSCAGYCVATFVLGIGDRHNDNIMITETGNLF
HIDFGHILGNyKSFLGINKERVFPFVLTPDFLwVMGTSGKkTSPHFQKFQDICVKAYALRHHTNLLIILFSMwLMTGMPQ
LTSKEDIeyIRDALTvGKNEEDAKKYFLDQIEVCROKGTvQFNwFLHLVLGIKQGEKHSA.

FIG.13

LISTA DE SECUENCIAS

(1) INFORMACIÓN GENERAL:

- 5 (i) SOLICITANTE: Stephens, Len
Hawkins, Phillip T.
- (ii) TÍTULO DE LA INVENCION: FOSFATIDILINOSITOL 3' QUINASA REGULADA POR UNA PRO-
TEÍNA G BETA-GAMMA
- 10 (iii) NÚMERO DE SECUENCIAS: 14
- (iv) DIRECCIÓN DE CORRESPONDENCIA:
- 15 (A) DIRECCIÓN: Pennie & Edmonds
(B) CALLE: 1155 Avenue of The Americas
(C) CIUDAD: Nueva York
(D) ESTADO: NY
20 (E) PAÍS: EE.UU.
(F) CÓDIGO POSTAL: 10036-2811
- (v) FORMATO LEGIBLE PARA ORDENADOR:
- 25 (A) MEDIO: Diskette
(B) ORDENADOR: Compatible IBM
(C) SISTEMA OPERATIVO: DOS
30 (D) PROGRAMA: FastSEQ Versión 2.0
- (vi) DATOS DE LA SOLICITUD ACTUAL:
- (A) NÚMERO DE SOLICITUD: PCT/US97/11219
(B) FECHA DE PRESENTACIÓN: 26-JUN-1997
35 (C) CLASIFICACIÓN:
- (vii) DATOS DE SOLICITUD ANTERIOR:
- (A) NÚMERO DE SOLICITUD: US 08/672,211
40 (B) FECHA DE PRESENTACIÓN: 27-JUN-1996
- (viii) INFORMACIÓN AGENTE/ABOGADO
- (A) NOMBRE: Stark, Jon R.
45 (B) NÚMERO DE REGISTRO: 30,111
(C) NÚMERO DE CASO/EXPEDIENTE: 8549-006-228
- (ix) INFORMACIÓN DE TELECOMUNICACIÓN:
- 50 (A) TELÉFONO: 650-493-4935
(B) TELEFAX: 650-493-5556
(C) TELEX: 66141 PENNIE

55 (2) INFORMACIÓN PARA LA SECUENCIA ID NO: 1

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 4692 pares de bases
60 (B) TIPO: Ácido nucléico
(C) CADENA: Única
(D) TOPOLOGÍA: Desconocida
- 65 (ii) TIPO MOLECULAR: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:1:

ES 2 278 392 T3

	CCGTGCGCCC	CTCAAGACTA	ATGGACCCCC	GGCTCAGGAA	TCGCACGAGG	CAGGCTCACA	60
	CCCGAGGCCC	ATGGAAGTTC	CCAGGCCAGG	GGTCAAGTTG	GAACCGAAGC	TGCTGCCAGC	120
5	TTACACCACA	GCCACAGCAA	CTTGGGATCT	GAGCTGCATC	TGTGACCTAC	ACCACAGCTC	180
	ACGGCAATGC	TGGATTCCCA	ACACACCAAG	TGGGGCCAGG	GATCGAACCC	GCATCCTCTT	240
	GGACAGTAGT	CAGATTCAAT	ACCACTGAGC	CATTGACAGG	AACTCCAGGG	GCAGGGGGGA	300
10	GTCTCTTGTT	TTTGGCTCCT	CCCGACACCT	GGTGAAATGG	ACCAGCGCAG	GCACCCCTTT	360
	CCAGTGGCTG	TCCCAGGCGA	TGACTCAGGA	TGCAGCCAGG	GGCCACGACG	TGCACGGAGG	420
	ACCGCATCCA	GCACGCCCTG	GAGCGCTGCT	TGCACGGGCT	CAGCCTCAGC	CGCCGCTCCA	480
15	CCTCCTGGTC	AGCTGGGCTG	TGCTAAACT	GTTGGAGCCT	GCAGGAGCTG	GTCAGCAGGG	540
	ATCCGGGCCA	CTTCCTTATC	CTCCTGGAGC	AGATCCTGCA	GAAAACCCGA	GAGGTCCAAG	600
	AGAAGGGCAC	CTATGACCTC	CTCGCGCCCC	TGGCCCTGCT	CTTCTATTCT	ACTGTCCTCT	660
20	GTACGCCACA	CTTCCCGCCA	GACTCAGATC	TCCTTCTGAA	AGCAGCCAGA	ACCTACCACC	720
	GATTCCTGAC	CTGGCCGGTT	CCGTACTGCA	GCATCTGCCA	GGAAC TGCTC	ACCTTCATCG	780
	ATGCTGAGCT	GAAGGCCCCA	GGAATCTCCT	ACCAGCGACT	GGTGAGGGCG	GAGCAGGGCC	840
25	TGTCCACAAG	GAGTCACCGC	AGCTCCACCG	TCACGGTGCT	CTTGSTGAAC	CCCGTGGAGG	900
	TGCAGGCTGA	GTTCCCTGAC	GTGGCCGACA	AGCTGAGCAC	ACCAGGGCCC	TCGCCGCACA	960
30	GCGCCTACAT	CACCCTGCTC	CTGCATGCCT	TCCAGGCCAC	CTTTGGGGCC	CACTGTGACC	1020
	TCTCTGGTCT	GCACCGCAGG	TTGCAGTCCA	AGACCCTGGC	AGAGCTCGAG	GCCATCTTCA	1080
	CGGAGACAGC	CGAGGCACAG	GAGCTGGCCT	CAGGCATCGG	GGATGCAGCT	GAGGCCCCGGC	1140
35	AGTGGCTCAG	GACCAAGCTG	CAGGCGGTGG	GAGAGAAGGC	CGGCTTCCCT	GGTGTCTTAG	1200
	ACACCGCCAA	ACCTGGCAAG	CTCCGCACCA	TCCCCATCCC	GGTCGCCAGG	TGCTACACCT	1260
	ACAGCTGGAA	CCAGGACAGC	TTCGACATCC	TGCAGGAAAT	CCTGCTCAAG	GAGCAGGAGC	1320
40	TGCTCCAGCC	AGAGATCCTG	GACGACGAGG	AGGACGAGGA	CGAGGAGGAC	GAGGAAGAGG	1380
	ACTTGACGC	CGACGGCCAC	TGCGCGGAGA	GGGACTCCGT	GCTCTCCACC	GGCTCGGCGG	1440
	CCTCCACGC	CTCCACGCTG	TCCCTGGCCT	CGTCCCAGGC	CTCGGGGCCC	ACGCTCTCCC	1500
45	GCCAGTTGCT	GACCTCCTTC	GTCTCGGGCC	TCTCGGATGG	CGTGGACAGC	GGCTACATGG	1560
	AGGACATCGA	GGAGAGCGCC	TACGAGCGGC	CCCGGAGGCC	TGGCGGCCAC	GAGCGCCGGG	1620
50	GCCACCGCCG	GCCCGGGCAG	AAGTTCAACA	GGATCTATAA	ACTCTTCAAG	AGCACCAGCC	1680
	AGATGGTGCT	GCGGAGGGAC	TCGCGCAGCC	TGGAGGGCAG	CCCGGACAGC	GGCCCGCCCC	1740
	TGCGTCGGGC	CGGCAGCCTC	TGCAGCCCCC	TGGACAGCCC	GACCCTGCCC	CCGTCCCCGGG	1800
55	CCCAGGGCTC	CCGCTCGCTG	CCCCAGCCCA	AGCTCAGCCC	CCAGCTGCCC	GGCTGGCTCC	1860
	TGGCCCCCGC	CTCCCGCCAC	CAGCGCCGCC	GCCCCCTCCT	GAGCGGGGAC	GAGGACCCCA	1920
	AGGCTTCCAC	GCTGCGTGTC	GTGGTCTTCG	GCTCGGATCG	GATCTCGGGG	AAGGTGGTCC	1980
60	GGGCTTACAG	CAACCTGCGG	CGGCTGGAGA	ACAACCGTCC	TCTCCTCACA	CGGTTCTTCA	2040
	AGCTACAGTT	CTTCTACGTG	CCTGTCAAGC	GGAGCCGTGG	GACAGGCACC	CCCACCAGCC	2100
65	CAGCCCCCTG	GAGCCAGACG	CCCCCCCCTC	CCACAGACGC	CCCGAGGCAC	CCGGGCCCTG	2160

ES 2 278 392 T3

	CAGAGCTGGG	CGCCGCCCCC	TGGGAGGAGA	GCACCAATGA	CATCTCCCAC	TACCTCGGCA	2220
	TGCTCGACCC	CTGGTACGAG	CGAAACGTCC	TGGGCCTCAT	GCACCTGCCT	CCTGAAGTCC	2280
5	TGTGCCAGTC	CCTGAAGGCT	GAGCCCCGGC	CCCTGGAGGG	CTCCCCTGCC	CAGCTGCCCA	2340
	TCCTGGCGGA	CATGCTGCTC	TACTACTGCC	GCTTCGCTGC	CCGGCCGGTG	CTGCTGCAGG	2400
	TCTATCAGAC	CGAGCTGACC	TTCATCACCG	GGGAGAAGAC	GACGGAGATC	TTCATCCACT	2460
10	CCCTGGAGCT	GGGCCACTCT	GCTGCCACAC	GTGCCATCAA	GGCTTCGGGT	CCTGGCAGCA	2520
	AGCGGCTGGG	CATCGATGGT	GACCGGGAGG	CCGTCCCTCT	AACACTACAG	ATAATTTACA	2580
15	GCAAGGGGGC	CATCAGCGGC	CGGAGTCGCT	GGAGCAACAT	GGAAAAGCTC	TGCACCTCTG	2640
	TCAACCTCAG	CAAGGCCTGC	CGGCAGCAGG	AGGAGCTAGA	CTCCAGCACA	GAGGCCCTGA	2700
	CGCTAAACCT	GACAGAAGTG	GTGAAAAGAC	AGACCCCTAA	ATCCAAGAAG	GGCTTTAACC	2760
20	AGATCAGCAC	CTCGCAGATC	AAAGTGGACA	AGGTGCAGAT	CATCGGCTCT	AACAGCTGCC	2820
	CCTTTGCCGT	GTGTCTGGAC	CAGGACGAGA	GGAAGATCCT	GCAGAGTGTC	ATCAGGTGCG	2880
	AGGTCTCGCC	CTGCTACAAG	CCTGAGAAGA	GCAGCCTCTG	CCCCCACC	CAGAGGCCCT	2940
25	CCTACCCGCC	AGCGCCGGCC	ACGCCCCGACC	TCTGCTCCCT	GCTCTGCCTG	CCCATCATGA	3000
	CTTTCAGCGG	AGCTCTGCCC	TAGCCGCCAC	CCTGCACCAG	CCTGGACAGG	GAGCCGGGGG	3060
30	GCAGCCTCCT	CGGAGCCCCC	TCCCCAGAAG	ACTGGCGGCT	GAGAGGGTCG	TGCTCCCTGT	3120
	GGAGAACAGA	GGGGCCGTGT	ACTGGGTCAG	GGTCCCCTG	TGGGCCCTGC	AGCAGCAAGA	3180
	GCGGGGGCTG	CTGGGGCCTC	AGGGCTCTGT	TTGGGCGAGA	AGCAGGCATT	AGGGAGAGGG	3240
35	GCCTGGCCCC	ACGGCTCTCA	GCTTCCTCAC	GGTAGCGGAG	AGAGGGATGG	GTGAGCTTGA	3300
	CCTCAAGGCC	CTGGCCTCCA	GTGGGGGTCC	AGGATCCTTT	CTGGAAGGAA	GATCCCAAGG	3360
40	CGCTGGTGCT	CTGGGGTGTC	GTGTTAGGGG	CTCCCCCCCC	AGCCCTGGGC	CAGGGCCCCC	3420
	CCGTTACTTT	GTCAGAGACT	TGGGGATCCT	GTGTCTGGAG	GGTCAAGTCC	CCCTCCCTGG	3480
	GGGTTCAAGC	AGTGGAAAGTA	TGGTTGCGAC	TTTTCTGACC	TTGGTGCAAT	CCCCGCCCCC	3540
45	ACCTCAACCC	CCCCACAAAA	AAACCCCTTC	TCTCTTTCAA	GTTCCCTGGG	TCTTCTGTGA	3600
	AACAGCACTA	ACACTTGACC	TGGCTGTGCC	AGCACTTGGA	ACAGATGCTC	CCTGGATCGA	3660
	GAGCCTTGGG	AGACAGGACA	AGCTTAGGTT	CGGTGGTGGC	TCAGTTACCT	TCTAGCGAAA	3720
50	TGAGCAGAAG	GAGGTGAATT	GGCTCCTTCG	AGGCTCCCCT	ACCTGGGCAC	TAAGATGGGG	3780
	GGAATAAGGC	CGCCTTAAAG	GTTTGGGGTG	ATGTCGTCTG	CAAAGCGCCT	GGCCCAGTGG	3840
55	CCGGCTGGTA	GCAAGGTGCG	GCCTCACCCCT	CTGGGCGTCG	ACTCCCTCGT	GTGGCGGGAG	3900
	GCTAAAAGGA	TGCCCTGCCC	CCGTGATGCT	GTCAATCCCT	CCTTCCAATT	CACTGATGAG	3960
	GCAGGACCCA	GACTGAGGGG	GTGAGGGGCG	CACAGTTCTA	CCTTTGAAGG	AGGAAGTGCC	4020
60	TTGATCAGAG	TAAGAGGAGG	GTGGCCCAGG	CGCCCCAAC	CGCCCCCTCC	TCCTCTCCCA	4080
	GGTTGGCCCC	TGTGCCCTCC	ACTCCCATCT	CACTCCTTGG	GCTGGCGCAC	ATCACGGGCA	4140
65	CAGTCCTCCA	GCCCCACAGT	TCACTGGTAC	CATGGCCCCCT	GGGTCGGTTC	GCAGAGGATG	4200

ES 2 278 392 T3

```

GAGGATAAGA CTTGCCTCGA GAACTTGGGT CTGATGGGGA AACCGGGTGA TGGAAATGAT 4260
TCCGGAAGAT TAAACCTCC CAGGTTCAAG TGTCGGAGAA CCGCCCCCAG AACCGGACTA 4320
GGTTGGTAGG GAGAGGGCAG GGCTTGGGCC CGGGATTGCG ACTAGGAGAG GCGGGGGGAG 4380
GTAACCAGAG AAGCAAGACA GTTGTATCCC CGCAAAAGAC CCTTCCCCGC CCCTCCCCCTC 4440
CTGCTCTGGC TCCATCTGCT TCAAAGGGTC TGGGCTTTAG GAGCCCGTGG TGCCCAGCGC 4500
AGCGTACTCA GGAATCGAGA GACGCGGACC GTGCCAGTTC CCACCCTGTG CCACTCCAGG 4560
CCCCAGGGAG GGGTTTGCAA TATACCTCA ACGTTTTTGT GTGTGTGGTA AGGTCGTCCT 4620
AGGACCCCAA ATGGAATTTA ACGTTATTGT CAAATAAAAC TTGATTGTGC TTGGAAAAAA 4680
AAAAAAAAAA AA 4692

```

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 877 aminoácidos
- (B) TIPO: Aminoácido
- (C) CADENA:
- (D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: Péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:2:

```

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

```

ES 2 278 392 T3

	Tyr	Ile	Thr	Leu	Leu	Leu	His	Ala	Phe	Gln	Ala	Thr	Phe	Gly	Ala	His
			195					200					205			
5	Cys	Asp	Leu	Ser	Gly	Leu	His	Arg	Arg	Leu	Gln	Ser	Lys	Thr	Leu	Ala
		210					215					220				
	Glu	Leu	Glu	Ala	Ile	Phe	Thr	Glu	Thr	Ala	Glu	Ala	Gln	Glu	Leu	Ala
	225					230					235					240
10	Ser	Gly	Ile	Gly	Asp	Ala	Ala	Glu	Ala	Arg	Gln	Trp	Leu	Arg	Thr	Lys
					245					250					255	
	Leu	Gln	Ala	Val	Gly	Glu	Lys	Ala	Gly	Phe	Pro	Gly	Val	Leu	Asp	Thr
15				260					265					270		
	Ala	Lys	Pro	Gly	Lys	Leu	Arg	Thr	Ile	Pro	Ile	Pro	Val	Ala	Arg	Cys
			275					280					285			
20	Tyr	Thr	Tyr	Ser	Trp	Asn	Gln	Asp	Ser	Phe	Asp	Ile	Leu	Gln	Glu	Ile
		290					295					300				
	Leu	Leu	Lys	Glu	Gln	Glu	Leu	Leu	Gln	Pro	Glu	Ile	Leu	Asp	Asp	Glu
	305					310					315					320
25	Glu	Asp	Glu	Asp	Glu	Glu	Asp	Glu	Glu	Glu	Asp	Leu	Asp	Ala	Asp	Gly
					325					330					335	
	His	Cys	Ala	Glu	Arg	Asp	Ser	Val	Leu	Ser	Thr	Gly	Ser	Ala	Ala	Ser
30				340					345					350		
	His	Ala	Ser	Thr	Leu	Ser	Leu	Ala	Ser	Ser	Gln	Ala	Ser	Gly	Pro	Thr
			355					360					365			
35	Leu	Ser	Arg	Gln	Leu	Leu	Thr	Ser	Phe	Val	Ser	Gly	Leu	Ser	Asp	Gly
		370					375					380				
	Val	Asp	Ser	Gly	Tyr	Met	Glu	Asp	Ile	Glu	Glu	Ser	Ala	Tyr	Glu	Arg
	385					390					395					400
40	Pro	Arg	Arg	Pro	Gly	Gly	His	Glu	Arg	Arg	Gly	His	Arg	Arg	Pro	Gly
					405					410					415	
	Gln	Lys	Phe	Asn	Arg	Ile	Tyr	Lys	Leu	Phe	Lys	Ser	Thr	Ser	Gln	Met
45				420					425					430		
	Val	Leu	Arg	Arg	Asp	Ser	Arg	Ser	Leu	Glu	Gly	Ser	Pro	Asp	Ser	Gly
			435					440					445			
50	Pro	Pro	Leu	Arg	Arg	Ala	Gly	Ser	Leu	Cys	Ser	Pro	Leu	Asp	Ser	Pro
		450					455					460				
	Thr	Leu	Pro	Pro	Ser	Arg	Ala	Gln	Gly	Ser	Arg	Ser	Leu	Pro	Gln	Pro
	465					470					475					480
55	Lys	Leu	Ser	Pro	Gln	Leu	Pro	Gly	Trp	Leu	Leu	Ala	Pro	Ala	Ser	Arg
					485					490					495	
	His	Gln	Arg	Arg	Arg	Pro	Phe	Leu	Ser	Gly	Asp	Glu	Asp	Pro	Lys	Ala
60				500					505					510		
	Ser	Thr	Leu	Arg	Val	Val	Val	Phe	Gly	Ser	Asp	Arg	Ile	Ser	Gly	Lys
			515					520					525			
65	Val	Val	Arg	Ala	Tyr	Ser	Asn	Leu	Arg	Arg	Leu	Glu	Asn	Asn	Arg	Pro
		530					535					540				

ES 2 278 392 T3

	Leu	Leu	Thr	Arg	Phe	Phe	Lys	Leu	Gln	Phe	Phe	Tyr	Val	Pro	Val	Lys
	545					550					555					560
5	Arg	Ser	Arg	Gly	Thr	Gly	Thr	Pro	Thr	Ser	Pro	Ala	Pro	Arg	Ser	Gln
					565					570					575	
	Thr	Pro	Pro	Leu	Pro	Thr	Asp	Ala	Pro	Arg	His	Pro	Gly	Pro	Ala	Glu
10				580					585					590		
	Leu	Gly	Ala	Ala	Pro	Trp	Glu	Glu	Ser	Thr	Asn	Asp	Ile	Ser	His	Tyr
			595					600					605			
15	Leu	Gly	Met	Leu	Asp	Pro	Trp	Tyr	Glu	Arg	Asn	Val	Leu	Gly	Leu	Met
		610					615					620				
	His	Leu	Pro	Pro	Glu	Val	Leu	Cys	Gln	Ser	Leu	Lys	Ala	Glu	Pro	Arg
20						630					635					640
	Pro	Leu	Glu	Gly	Ser	Pro	Ala	Gln	Leu	Pro	Ile	Leu	Ala	Asp	Met	Leu
					645					650					655	
25	Leu	Tyr	Tyr	Cys	Arg	Phe	Ala	Ala	Arg	Pro	Val	Leu	Leu	Gln	Val	Tyr
				660					665					670		
	Gln	Thr	Glu	Leu	Thr	Phe	Ile	Thr	Gly	Glu	Lys	Thr	Thr	Glu	Ile	Phe
30								680						685		
	Ile	His	Ser	Leu	Glu	Leu	Gly	His	Ser	Ala	Ala	Thr	Arg	Ala	Ile	Lys
		690					695					700				
35	Ala	Ser	Gly	Pro	Gly	Ser	Lys	Arg	Leu	Gly	Ile	Asp	Gly	Asp	Arg	Glu
						710					715					720
	Ala	Val	Pro	Leu	Thr	Leu	Gln	Ile	Ile	Tyr	Ser	Lys	Gly	Ala	Ile	Ser
					725					730					735	
40	Gly	Arg	Ser	Arg	Trp	Ser	Asn	Met	Glu	Lys	Leu	Cys	Thr	Ser	Val	Asn
				740					745					750		
	Leu	Ser	Lys	Ala	Cys	Arg	Gln	Gln	Glu	Glu	Leu	Asp	Ser	Ser	Thr	Glu
45			755					760					765			
	Ala	Leu	Thr	Leu	Asn	Leu	Thr	Glu	Val	Val	Lys	Arg	Gln	Thr	Pro	Lys
							775					780				
50	Ser	Lys	Lys	Gly	Phe	Asn	Gln	Ile	Ser	Thr	Ser	Gln	Ile	Lys	Val	Asp
						790					795					800
	Lys	Val	Gln	Ile	Ile	Gly	Ser	Asn	Ser	Cys	Pro	Phe	Ala	Val	Cys	Leu
55					805					810					815	
	Asp	Gln	Asp	Glu	Arg	Lys	Ile	Leu	Gln	Ser	Val	Ile	Arg	Cys	Glu	Val
				820					825					830		
60	Ser	Pro	Cys	Tyr	Lys	Pro	Glu	Lys	Ser	Ser	Leu	Cys	Pro	Pro	Pro	Gln
			835					840					845			
	Arg	Pro	Ser	Tyr	Pro	Pro	Ala	Pro	Ala	Thr	Pro	Asp	Leu	Cys	Ser	Leu
65							855					860				
	Leu	Cys	Leu	Pro	Ile	Met	Thr	Phe	Ser	Gly	Ala	Leu	Pro			
						870					875					

ES 2 278 392 T3

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 3808 pares de bases

(B) TIPO: Ácido nucleico

(C) CADENA: Única

(D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:3:

```

5      GGCACGAGGA ATTTGTTTTG TTTTCAGAAA TTAAACAAAT GATCCTTCAG CATCATCACC      60
      TCCGCTGCTT TATCAGGTCG CATAGGGCAT GGAGCTGGAG AACTATGAAC AGCCCGTGGT      120
      GCTGAGAGAG GACAACCGCC GCAGGCGTCG GAGGATGAAG CCGCGCAGCA CGGCAGCCAG      180
10     CCTGTCTCTC ATGGAGCTCA TCCCCATCGA GTTTGTTTTG GCCACCAGCC AGCGCAACAC      240
      CAAGACCCCC GAAACGGCAC TGCTGCACGT GGCCGGCCAC GGCAATGTGG AGAAGATGAA      300
      GGCCCAGGTG TTGTTGCGCG CGCTGGAGAC GAGCGTTTCT TGGGACTTCT ACCACCGGTT      360
25     CGGCCCCGAC CACTTCCTCC TGGTCTTCCA GAAGAAGGGG GAGTGGTACG AGATCTATGA      420
      CAAGTACCAG GTGGTGCGA CCCTGGACTG CCTGCGCTAC TGGGAGGTGT TGCACCGCAG      480
30     CCCCAGGAG ATCCACGTGG TCCAGCGGCA CGCGCCCTCG GAGGAGACAT TGGCCTTCCA      540
      GCGCCAGCTC AACGCCCTCA TCGGCTACGA CGTCACCGAC GTCAGCAACG TGCATGACGA      600
      TGAGCTGGAG TTCACGCGGC GCCGCTGGT CACCCCGCGC ATGGCCGAGG TGGCCGGCCG      660
35     CGACCCCAAG CTTTACGCCA TGCACCCCTG GGTGACATCC AAGCCCCTCC CTGAGTACCT      720
      TCTGAAGAAG ATCACTAACA ACTGCGTCTT CATCGTCATT CACCGCAGCA CCACCAGCCA      780
40     GACCATCAAG GTCTCGGCCG ATGACACCCC AGGCACCATC CTCCAGAGCT TCTTTACCAA      840
      GATGGCCAAG AAGAAATCTC TGATGGATAT CCCTGAAAGC CAGAACGAAC GGGACTTTGT      900
      GCTGCGCGTC TGCGGCCGGG ATGAGTACCT GGTGGGTGAG ACGCCCATCA AAAATTTCCA      960
45     GTGGGTGAGG CAGTGCCTCA AGAATGGGGA GGAGATTCAC CTTGTGCTGG AACTCCTCC      1020
      AGACCCAGCC CTGGACGAGG TGAGGAAGGA AGAGTGGCCG CTGGTGGATG ACTGCACGGG      1080
      AGTCACTGGC TACCACGAGC AGCTGACCAT CCACGGCAAG GACCATGAAA GTGTGTTTAC      1140
50     CGTGTCCCTG TGGGACTGTG ACCGCAAGTT CAGGGTCAAA ATCAGAGGCA TTGATATCCC      1200
      TGTCCTGCCC CGGACCGCTG ACCTCACGGT GTTTGTGGAG GCAAACATCC AGTATGGGCA      1260
55     GCAAGTCCTT TGCCAAAGGA GAACCAGCCC CAAACCCTTC ACGGAGGAGG TGCTCTGGAA      1320
      CGTGTGGCTT GAGTTCAGTA TTAAAATCAA AGACTTACCC AAAGGGGCTC TGCTGAACCT      1380
      CCAGATCTAC TGCGGCAAAG CTCCAGCACT GTCTGGCAAG ACCTCTGCAG AGATGCCCAG      1440
60     TCCCGAGTCC AAAGGCAAAG CTCAGCTTCT GACTATGTC AACCTATTGC TGATAGACCA      1500
      CCGCTTCCTC CTGCGCCATG GCGAGTATGT GCTCCACATG TGGCAGTTAT CCGGGAAGGG      1560
      GGAAGACCAA GGGAGCTTCA ATGCCGACAA GCTCACGTCG GGAACCAACC CGGACAAGGA      1620
65     GGAATCAATG TCCATCTCCA TTCTTCTGGA CAATTACTGC CACCCCATAG CCTTGCCTAA      1680

```

ES 2 278 392 T3

	GCATCGGCCT	ACCCCTGACC	CAGAAGGGGA	CCGGGTTCGG	GCAGAAATGC	CCAATCAGCT	1740
	TCGGAAGCAA	CTGGAGGCAA	TCATAGCCAC	GGATCCGCTT	AACCCACTCA	CAGCTGAAGA	1800
5	CAAAGAACTG	CTCTGGCATT	TCAGATATGA	AAGCCTGAAG	GATCCCAAAG	CGTATCCTAA	1860
	GCTCTTTAGC	TCGGTGAAAT	GGGGACAGCA	AGAAATTGTG	GCCAAAACAT	ACCAATTATT	1920
	AGCCAAAAGG	GAGGTCTGGG	ATCAGAGTGC	TTTGGATGTG	GGGTTAACCA	TGCAGCTCCT	1980
10	GGACTGCAAC	TTCTCGGATG	AAAACGTGAG	AGCCATTGCA	GTCCAGAAAC	TGGAGAGCTT	2040
	GGAGGATGAT	GACGTGCTCC	ATTACCTGCT	CCAGCTGGTC	CAGGCTGTGA	AATTTGAACC	2100
	ATACCATGAC	AGTGCCCTAG	CCAGATTTCT	GCTGAAGCGT	GGTTAAGAA	ACAAGAGAAT	2160
15	TGGTCACTTC	TTGTTTTGGT	TCTTGAGAAG	TGAGATTGCC	CAGTCTAGGC	ACTATCAGCA	2220
	GAGGTTTGCA	GTGATCCTGG	AAGCCTACCT	GAGGGGCTGT	GGCACAGCCA	TGCTGCACGA	2280
	CTTCACCCAG	CAAGTCCAAG	TAATTGACAT	GTTACAAAA	GTCACCATTG	ACATTAAATC	2340
20	GCTCTCTGCT	GAAAAGTATG	ACGTCAGTTC	CCAAGTTATT	TCCCAACTTA	AGCAAAAGCT	2400
	TGAAAACCTA	CAGAATTTGA	ATCTCCCCCA	AAGCTTTAGA	GTTCCCTATG	ATCCTGGACT	2460
25	GAAAGCCGGG	GCACTGGTGA	TCGAAAAATG	TAAAGTGATG	GCCTCCAAGA	AGAAGCCCCCT	2520
	GTGGCTTGAG	TTTAAATGTG	CCGATCCTAC	GGCTCTATCA	AATGAAACAA	TTGGAATTAT	2580
	CTTTAAACAC	GGTGACGATC	TGCGCCAAGA	CATGCTTATT	TTACAGATTC	TACGAATCAT	2640
30	GGAGTCCATT	TGGGAGACCG	AATCTTTGGA	TCTGTGCCTC	CTGCCATATG	GCTGCATTTTC	2700
	AACTGGTGAC	AAAATAGGAA	TGATCGAGAT	CGTGAAGGAC	GCCACGACAA	TCGCCAAAT	2760
	TCAGCAAAGC	ACAGTGGGCA	ACACGGGTGC	CTTTAAAGAT	GAAGTCCTGA	GTCAGTGGCT	2820
35	CAAAGAAAAA	TGCCCTATTG	AAGAAAAGTT	TCAGGCAGCT	GTGGAGAGAT	TTGTTTATTC	2880
	CTGTGCCGGC	TACTGTGTGG	CAACCTTTGT	TCTCGGAATA	GGCGACAGAC	ACAATGACAA	2940
40	TATTATGATC	TCAGAAACAG	GAAATCTATT	TCATATTGAT	TTCGGACACA	TTCTTGGGAA	3000
	TTACAAAAGT	TTCCTGGGCA	TTAATAAAGA	GAGGGTGCCA	TTTGTGCTAA	CCCCAGACTT	3060
	CCTGTTTGTG	ATGGGGACTT	CTGGAAAGAA	GACAAGTCTA	CACCTCCAGA	AATTTACAGG	3120
45	TGTCTGCGTC	AAGGCTTACC	TAGCCCTTCG	TCATCACACA	AACCTACTGA	TCATCCTCTT	3180
	CTCCATGATG	CTGATGACAG	GAATGCCCCA	GTTAACCAGC	AAAGAAGACA	TTGAATACAT	3240
	TCGGGATGCC	CTCACAGTGG	GCAAAAGTGA	GGAGGATGCT	AAAAAGTATT	TTCTGGATCA	3300
50	GATTGAAGTT	TGCAGAGACA	AAGGATGGAC	CGTGCACTTT	AACTGGTTCT	TACATCTTGT	3360
	TCTTGGCATC	AAACAAGGGG	AGAAGCATCC	CGCATAAAAC	TTTGGGCCAA	GAGTTAAAAC	3420
55	CCAAGTTATT	GTCCTAATGC	TTTACGTCAG	CAGGACAATC	ACCGAACTTG	ATGTCATGTA	3480
	GTGGGACATT	ATGAAAGCTG	GCACTTGAGA	AATATAGCTC	TTCCCTAAC	TGAACTCTTC	3540
	ACTGGAGAAA	AACCTTGGCA	TGTTTAAGTA	ATGTTCACTG	TTAGGCTTAT	TTGCATGTTT	3600
60	GTTTTTTCTC	ATGTGCCCCC	TCAGTCATGT	TGGAGACTGT	TCTAAATTTA	AGTGGCCTAA	3660
	TGACCTCTGA	AGTTTCAACT	TTCTTGGTAC	TGAGTGCTTC	TGAAATTCTT	TACAATAATT	3720
	GGTAACATCT	ATTGTCAGCT	GGGTATCCTC	TCAATTTTGG	TTATCCTTGG	GTTTCTCAAA	3780
65	CTCCTTACAG	GAAAAAAAAA	AAAAAAAAA				3808

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:4:

ES 2 278 392 T3

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1102 aminoácidos

(B) TIPO: Aminoácido

(C) CADENA:

(D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: Péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:4:

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

ES 2 278 392 T3

	Ser	Leu	Met	Asp	Ile	Pro	Glu	Ser	Gln	Asn	Glu	Arg	Asp	Phe	Val	Leu	
				260					265					270			
5	Arg	Val	Cys	Gly	Arg	Asp	Glu	Tyr	Leu	Val	Gly	Glu	Thr	Pro	Ile	Lys	
			275					280					285				
	Asn	Phe	Gln	Trp	Val	Arg	Gln	Cys	Leu	Lys	Asn	Gly	Glu	Glu	Ile	His	
10		290					295					300					
	Leu	Val	Leu	Asp	Thr	Pro	Pro	Asp	Pro	Ala	Leu	Asp	Glu	Val	Arg	Lys	
	305					310					315					320	
	Glu	Glu	Trp	Pro	Leu	Val	Asp	Asp	Cys	Thr	Gly	Val	Thr	Gly	Tyr	His	
15					325					330					335		
	Glu	Gln	Leu	Thr	Ile	His	Gly	Lys	Asp	His	Glu	Ser	Val	Phe	Thr	Val	
				340					345					350			
20	Ser	Leu	Trp	Asp	Cys	Asp	Arg	Lys	Phe	Arg	Val	Lys	Ile	Arg	Gly	Ile	
			355					360					365				
	Asp	Ile	Pro	Val	Leu	Pro	Arg	Thr	Ala	Asp	Leu	Thr	Val	Phe	Val	Glu	
25							375					380					
	Ala	Asn	Ile	Gln	Tyr	Gly	Gln	Gln	Val	Leu	Cys	Gln	Arg	Arg	Thr	Ser	
	385					390					395					400	
	Pro	Lys	Pro	Phe	Thr	Glu	Glu	Val	Leu	Trp	Asn	Val	Trp	Leu	Glu	Phe	
30					405					410					415		
	Ser	Ile	Lys	Ile	Lys	Asp	Leu	Pro	Lys	Gly	Ala	Leu	Leu	Asn	Leu	Gln	
				420					425					430			
35	Ile	Tyr	Cys	Gly	Lys	Ala	Pro	Ala	Leu	Ser	Gly	Lys	Thr	Ser	Ala	Glu	
			435					440					445				
	Met	Pro	Ser	Pro	Glu	Ser	Lys	Gly	Lys	Ala	Gln	Leu	Leu	Tyr	Tyr	Val	
40		450					455					460					
	Asn	Leu	Leu	Leu	Ile	Asp	His	Arg	Phe	Leu	Leu	Arg	His	Gly	Glu	Tyr	
	465				470						475					480	
45	Val	Leu	His	Met	Trp	Gln	Leu	Ser	Gly	Lys	Gly	Glu	Asp	Gln	Gly	Ser	
				485					490					495			
	Phe	Asn	Ala	Asp	Lys	Leu	Thr	Ser	Gly	Thr	Asn	Pro	Asp	Lys	Glu	Asp	
50				500					505					510			
	Ser	Met	Ser	Ile	Ser	Ile	Leu	Leu	Asp	Asn	Tyr	Cys	His	Pro	Ile	Ala	
			515					520					525				
55	Leu	Pro	Lys	His	Arg	Pro	Thr	Pro	Asp	Pro	Glu	Gly	Asp	Arg	Val	Arg	
		530					535					540					
	Ala	Glu	Met	Pro	Asn	Gln	Leu	Arg	Lys	Gln	Leu	Glu	Ala	Ile	Ile	Ala	
	545					550					555					560	
60	Thr	Asp	Pro	Leu	Asn	Pro	Leu	Thr	Ala	Glu	Asp	Lys	Glu	Leu	Leu	Trp	
				565					570					575			
	His	Phe	Arg	Tyr	Glu	Ser	Leu	Lys	Asp	Pro	Lys	Ala	Tyr	Pro	Lys	Leu	
65				580					585				590				
	Phe	Ser	Ser	Val	Lys	Trp	Gly	Gln	Gln	Glu	Ile	Val	Ala	Lys	Thr	Tyr	
			595					600					605				

ES 2 278 392 T3

	Gln	Leu	Leu	Ala	Lys	Arg	Glu	Val	Trp	Asp	Gln	Ser	Ala	Leu	Asp	Val
	610						615					620				
5	Gly	Leu	Thr	Met	Gln	Leu	Leu	Asp	Cys	Asn	Phe	Ser	Asp	Glu	Asn	Val
	625					630					635					640
	Arg	Ala	Ile	Ala	Val	Gln	Lys	Leu	Glu	Ser	Leu	Glu	Asp	Asp	Asp	Val
10					645					650					655	
	Leu	His	Tyr	Leu	Leu	Gln	Leu	Val	Gln	Ala	Val	Lys	Phe	Glu	Pro	Tyr
				660					665					670		
15	His	Asp	Ser	Ala	Leu	Ala	Arg	Phe	Leu	Leu	Lys	Arg	Gly	Leu	Arg	Asn
			675					680					685			
	Lys	Arg	Ile	Gly	His	Phe	Leu	Phe	Trp	Phe	Leu	Arg	Ser	Glu	Ile	Ala
		690					695					700				
20	Gln	Ser	Arg	His	Tyr	Gln	Gln	Arg	Phe	Ala	Val	Ile	Leu	Glu	Ala	Tyr
	705					710					715					720
	Leu	Arg	Gly	Cys	Gly	Thr	Ala	Met	Leu	His	Asp	Phe	Thr	Gln	Gln	Val
25					725					730					735	
	Gln	Val	Ile	Asp	Met	Leu	Gln	Lys	Val	Thr	Ile	Asp	Ile	Lys	Ser	Leu
				740					745					750		
30	Ser	Ala	Glu	Lys	Tyr	Asp	Val	Ser	Ser	Gln	Val	Ile	Ser	Gln	Leu	Lys
			755					760					765			
	Gln	Lys	Leu	Glu	Asn	Leu	Gln	Asn	Leu	Asn	Leu	Pro	Gln	Ser	Phe	Arg
		770					775					780				
35	Val	Pro	Tyr	Asp	Pro	Gly	Leu	Lys	Ala	Gly	Ala	Leu	Val	Ile	Glu	Lys
	785					790					795					800
	Cys	Lys	Val	Met	Ala	Ser	Lys	Lys	Lys	Pro	Leu	Trp	Leu	Glu	Phe	Lys
40					805					810					815	
	Cys	Ala	Asp	Pro	Thr	Ala	Leu	Ser	Asn	Glu	Thr	Ile	Gly	Ile	Ile	Phe
				820					825					830		
45	Lys	His	Gly	Asp	Asp	Leu	Arg	Gln	Asp	Met	Leu	Ile	Leu	Gln	Ile	Leu
			835					840					845			
	Arg	Ile	Met	Glu	Ser	Ile	Trp	Glu	Thr	Glu	Ser	Leu	Asp	Leu	Cys	Leu
50							855					860				
	Leu	Pro	Tyr	Gly	Cys	Ile	Ser	Thr	Gly	Asp	Lys	Ile	Gly	Met	Ile	Glu
	865					870					875					880
	Ile	Val	Lys	Asp	Ala	Thr	Thr	Ile	Ala	Lys	Ile	Gln	Gln	Ser	Thr	Val
55					885					890					895	
	Gly	Asn	Thr	Gly	Ala	Phe	Lys	Asp	Glu	Val	Leu	Ser	His	Trp	Leu	Lys
				900					905					910		
60	Glu	Lys	Cys	Pro	Ile	Glu	Glu	Lys	Phe	Gln	Ala	Ala	Val	Glu	Arg	Phe
			915					920					925			
	Val	Tyr	Ser	Cys	Ala	Gly	Tyr	Cys	Val	Ala	Thr	Phe	Val	Leu	Gly	Ile
		930					935					940				
65	Gly	Asp	Arg	His	Asn	Asp	Asn	Ile	Met	Ile	Ser	Glu	Thr	Gly	Asn	Leu
	945					950					955					960

ES 2 278 392 T3

	Phe His Ile Asp	Phe Gly His Ile	Leu Gly Asn Tyr	Lys Ser Phe Leu	
		965	970	975	
5	Gly Ile Asn Lys	Glu Arg Val Pro	Phe Val Leu Thr	Pro Asp Phe Leu	
	980	985	990		
	Phe Val Met Gly	Thr Ser Gly Lys	Lys Thr Ser Leu	His Phe Gln Lys	
10	995	1000	1005		
	Phe Gln Asp Val	Cys Val Lys Ala	Tyr Leu Ala Leu	Arg His His Thr	
	1010	1015	1020		
15	Asn Leu Leu Ile	Ile Leu Phe Ser	Met Met Leu Met	Thr Gly Met Pro	
	1025	1030	1035	1040	
	Gln Leu Thr Ser	Lys Glu Asp Ile	Glu Tyr Ile Arg	Asp Ala Leu Thr	
20	1045	1050	1055		
	Val Gly Lys Ser	Glu Glu Asp Ala	Lys Lys Tyr Phe	Leu Asp Gln Ile	
	1060	1065	1070		
25	Glu Val Cys Arg	Asp Lys Gly Trp	Thr Val Gln Phe	Asn Trp Phe Leu	
	1075	1080	1085		
	His Leu Val Leu	Gly Ile Lys Gln	Gly Glu Lys His	Pro Ala	
30	1090	1095	1100		

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- | | |
|----|---------------------------------|
| 35 | (A) LONGITUD: 36 pares de bases |
| | (B) TIPO: Ácido nucléico |
| | (C) CADENA: Única |
| 40 | (D) TOPOLOGÍA: Desconocida |

(ii) TIPO MOLECULAR: ADN genómico

(ix) CARACTERÍSTICAS:

- | | |
|----|--------------------------------------|
| 45 | (A) NOMBRE/CÓDIGO: misc_features |
| | (B) LOCALIZACIÓN: 12...0 |
| | (D) OTRA INFORMACIÓN: "N es inosina" |
| 50 | (A) NOMBRE/CÓDIGO: misc_features |
| | (B) LOCALIZACIÓN: 21...0 |
| | (D) OTRA INFORMACIÓN: "N es inosina" |
| 55 | (A) NOMBRE/CÓDIGO: misc_features |
| | (B) LOCALIZACIÓN: 27...0 |
| | (D) OTRA INFORMACIÓN: "N es inosina" |

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:5:

60	CAYGAYTTYA CNCARCARGT NCARGTNATH GAYATG
----	---

36

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- | | |
|----|---------------------------------|
| 65 | (A) LONGITUD: 23 pares de bases |
|----|---------------------------------|

- (B) TIPO: Ácido nucleico
- (C) CADENA: Única
- (D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: ADN genómico

(ix) CARACTERÍSTICAS:

- (A) NOMBRE/CÓDIGO: misc_features
- (B) LOCALIZACIÓN: 3...0
- (D) OTRA INFORMACIÓN: "N es inosina"
- (A) NOMBRE/CÓDIGO: misc_features
- (B) LOCALIZACIÓN: 18...0
- (D) OTRA INFORMACIÓN: "N es inosina"

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:6:

GCNTAYATGG ARGAYATNGA RGA

23

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 13 aminoácidos
- (B) TIPO: aminoácido
- (C) CADENA: Desconocida
- (D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:7:

Met Glu Glu Glu Glu Phe Met Pro Met Pro Met Glu Phe
1 5 10

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 13 aminoácidos
- (B) TIPO: aminoácido
- (C) CADENA: Desconocida
- (D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:8:

Met Glu Glu Glu Glu Phe Met Pro Met Glu Phe Ser Ser
1 5 10

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:9:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 13 aminoácidos
- (B) TIPO: aminoácido
- (C) CADENA: Desconocida
- (D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: péptido

ES 2 278 392 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:9:

Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Glu Phe
1 5 10

5

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:10:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

10

(A) LONGITUD: 15 aminoácidos

(B) TIPO: aminoácido

(C) CADENA: Desconocida

15

(D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:10:

20

Met Glu Gln Lys Leu Ile Ser Gu Glu Asp Leu glu Phe Ser Ser
1 5 10 15

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:11:

25

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 3255 pares de bases

(B) TIPO: Ácido nucleico

30

(C) CADENA: Desconocida

(D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: ADNc

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:11:

40

45

50

55

60

65

ES 2 278 392 T3

	AAGCTTACTC	AGCGCATCTG	CTACAGCATT	TTATCTTCCC	AAGAGCCCCA	TGAGGCGATG	60
	ACCCAGGATG	CAGCCAGGGG	CCACGACATG	CACGGAGGAC	CGCATCCAGC	ATGCCCTGGA	120
	ACGCTGCCTG	CATGGACTCA	GCCTCAGCCG	CCGCTCCACC	TCCTGGTCAG	CTGGGCTGTG	180
5	TCTGAACTGC	TGGAGCCTGC	AGGAGCTGGT	CAGCAGGGAC	CCGGGCCACT	TCCTTATCCT	240
	CCTTGAGCAG	ATCCTGCAGA	AGACCCGAGA	GGTCCAGGAG	AAGGGCACCT	ACGACCTGCT	300
	CACCCCGCTG	GCCCTGCTCT	TCTATTCCAC	TGTTCTTTGT	ACACCACACT	TCCCACCAGA	360
	CTCGGATCTC	CTTCTGAAGG	CAGCCAGCAC	CTACCACCGG	TTCCTGACCT	GGCCTGTTCC	420
	TTACTGCAGC	ATCTGCCAGG	AGCTGCTCAC	CTTCATTGAT	GCTGAACTCA	AGGCCCCAGG	480
10	GATCTCCTAC	CAGAGACTGG	TGAGGGCTGA	GCAGGGCCTG	CCCATCAGGA	GTCACCGCAG	540
	CTCCACCGTC	ACCGTGCTGC	TGCTGAACCC	AGTGGAAGTG	CAGGCCGAGT	TCCTTGCTGT	600
	AGCCAATAAG	CTGAGTACGC	CCGGACACTC	GCCTCACAGT	GCCTACACCA	CCCTGCTCCT	660
	GCACGCCTTC	CAGGCCACCT	TTGGGGCCCA	CTGTGACGTC	CCGGGCCTGC	ACTGCAGGCT	720
	ACAGGCCAAG	ACCCTGGCAG	AGCTTGAGGA	CATCTTCACG	GAGACCGCAG	AGGCACAGGA	780
	GCTGGCATCT	GGCATCGGGG	ATGCTGCAGA	GGCCCCGGCG	TGGCTCAGGA	CCAAGCTGCA	840
15	GGCGGTGGGA	GAAAAAGCTG	GCTTCCCTGG	GGTGTAGAC	ACTGCAAAAC	CAGGGAAGCT	900
	CCACACCATC	CCCATCCCTG	TCGCCAGGTG	CTACACCTAC	AGCTGGAGCC	AGGACAGCTT	960
	TGACATCCTG	CAGGAAATCC	TGCTCAAGGA	ACAGGAGCTA	CTCCAGCCAG	GGATCCTGGG	1020
	AGATGATGAA	GAGGAGGAAG	AGGAGGAGGA	GGAGGTGGAG	GAGGACTTGG	AAACTGACGG	1080
	GCACTGTGCC	GAGAGAGATT	CCCTGCTCTC	CACCAGCTCT	TTGGCGTCCC	ATGACTCCAC	1140
20	CTTGTCCCTT	GCATCCTCCC	AGGCCTCGGG	GCCGGCCCTC	TCGCGCCATC	TGCTGACTTC	1200
	CTTTGTCTCA	GGCCTCTCTG	ATGGCATGGA	CAGCGGTAC	GTGGAGGACA	GCGAGGAGAG	1260
	CTCCTCCGAG	TGGCCTTGGA	GGCGTGCCAG	CCAGGAACGC	CGAGGCCACC	GCAGGCCTGG	1320
	GCAGAAGTTC	ATCAGGATCT	ATAAACTCTT	CAAGAGCACC	AGCCAGCTGG	TACTGCGGAG	1380
	GGACTCTCGG	AGCCTGGAGG	GCAGCTCGGA	CACGGCCCTG	CCCCTGAGGC	GGGCAGGGAG	1440
25	CCTCTGCAGC	CCCCTGACG	AACCAGTATC	ACCCCTTCC	CGGGCCCAGC	GCTCCCGCTC	1500
	CCTGCCCCAG	CCCAAACCTG	GTACCCAGCT	GCCCAGCTGG	CTTCTGGCCC	CTGCTTCACG	1560
	CCCCCAGCGC	CGCCGCCCTT	TCCTGAGTGG	AGATGAGGAT	CCCAAGGCTT	CCACGCTACG	1620
	TGTTGTGGTC	TTTGGCTCCG	ATCGGATTTT	AGGGAAGGTG	GCTCGGGCGT	ACAGCAACCT	1680
	TCGGCGGGCT	GAGAACAATC	GCCCCACTCT	CACACGGTTC	TTCAAACCTT	AGTTCTTCTA	1740
30	CGTGCCCTGT	AAGCGAAGTC	ATGGGACCAG	CCCTGGTGCC	TGTCCACCCC	CTCGGAGCCA	1800
	GACGCCCTCA	CCCCGACAG	ACTCCCCTAG	GCACGCCAGC	CCTGGAGAGC	TGGGCACCAC	1860
	CCCATGGGAG	GAGAGCACC	ATGACATCTC	CCACTACCTC	GGCATGCTGG	ACCCCTGTGA	1920
	TGAGCGCAAT	GTAATGGGCC	TCATGCACCT	GCCCCCTGAA	GTCTGTGTGC	AGCAGTCCCT	1980
	GAAGGCTGAA	GCCCAGGCC	TGGAGGGCTC	CCCAACCCAG	CTGCCCATCC	TGGCTGACAT	2040
35	GCTACTCTAC	TACTGCCGCT	TTGCCGCCAG	ACCGGTGCTG	CTGCAACTCT	ATCAGACCGA	2100
	GCTGACCTTC	ATCACTGGGG	AGAAGACGAC	AGAGATCTTC	ATCCACTCCT	TGGAGCTGGG	2160
	TCACTCCGCT	GCCACACGTG	CCATCAAGGC	GTCAGGTCCT	GGCAGCAAGC	GGCTGGGCAT	2220
	CGATGGCGAC	CGGGAGGCTG	TTCTCTAAC	ACTACAGATT	ATTTACAGCA	AGGGGGCCAT	2280
	CAGTGGACGA	AGTCGCTGGA	GCAACCTGGA	GAAGGTCTGT	ACCTCCGTGA	ACCTCAACAA	2340
40	GGCCTGCCCG	AAGCAGGAGG	AGCTGGATTG	CAGCATGGAG	GCCCTGACGC	TAAACCTGAC	2400
	AGAAGTGGTG	AAAAGGCAGA	ACTCCAAATC	CAAGAAGGGC	TTTAACCAGA	TTAGCACATC	2460
	GCAGATCAAA	GTGGACAAGG	TGCAGATCAT	CGGCTCCAAC	AGCTGCCCTT	TTGCTGTGTG	2520
	CCTGGACCAG	GATGAGAGAA	AGATCCTGCA	GAGTGTAGTC	AGATGTGAGG	TCTCACCGTG	2580
	CTACAAGCCA	GAGAAGAGCG	ACCTCTCCTC	ACCACCCAG	ACGCCTCCTG	ACCTGCCGGC	2640
45	CCAGGCCGCA	CCTGATCTCT	GCTCCCTTCT	CTGCCTGCCC	ATCATGACTT	TCAGTGGAGC	2700
	TCTGCCCTAG	TGTGGGCCCA	GCGCCAGACT	GGACAGAAGC	CCTGGGGCAA	CCTCCTCGGC	2760
	CACCCCTCCA	GGACAGTCCC	TCTCTGTGGA	GAAGTGAATG	GCCCTGTGCA	GAGCCATAGT	2820
	CCCACTGTGG	GTCTTGCAAT	GAGCAGGGGC	TGGGAGTAGA	GGGTTTCTGG	GGCCTCAGGG	2880
50	TTCTGGGAAA	GCAACAGCTA	TCAGAGAGAG	AAGGGCCAGA	CCCCATAGCC	TCTTAGATTG	2940
	CTGGCAGTAG	AAGGAGAAGG	ATGGGTAAAT	TGACCTCTGA	AGTCCCTGAC	CATTAGCATG	3000
	GTCTAGGATG	CTTTCTAGAA	GGAAGATCTG	AGGCTCTGGT	GCTCAGGGGG	ATGGCTTGGG	3060
	CCTTTTCTCT	CAACCTTGCC	TGAGCCTACC	CCTTACTTTG	CCAAAGACTT	GAGGACCCTG	3120
	TATGTCTGGA	GTTCACTCCC	CTCCTCTGTG	GGGCTCAGGT	GATTGAAATG	TGGATGAAAC	3180
	ATTTCTCTAC	TTCAAGACCA	CCTCTCCCTG	CAAACACCAC	ACACACATGG	CATGCATGTA	3240
55	CGCACATGCG	CACCG					3255

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 880 aminoácidos
- (B) TIPO: aminoácido
- (C) CADENA: Desconocida
- (D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: proteína

ES 2 278 392 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:12:

5	Met	Gln	Pro	Gly	Ala	Thr	Thr	Cys	Thr	Glu	Asp	Arg	Ile	Gln	His	Ala
	1				5					10				15		
	Leu	Glu	Arg	Cys	Leu	His	Gly	Leu	Ser	Leu	Ser	Arg	Arg	Ser	Thr	Ser
				20					25					30		
	Trp	Ser	Ala	Gly	Leu	Cys	Leu	Asn	Cys	Trp	Ser	Leu	Gln	Glu	Leu	Val
			35					40					45			
10	Ser	Arg	Asp	Pro	Gly	His	Phe	Leu	Ile	Leu	Leu	Glu	Gln	Ile	Leu	Gln
	50						55					60				
	Lys	Thr	Arg	Glu	Val	Gln	Glu	Lys	Gly	Thr	Tyr	Asp	Leu	Leu	Thr	Pro
	65				70						75					80
15	Leu	Ala	Leu	Leu	Phe	Tyr	Ser	Thr	Val	Leu	Cys	Thr	Pro	His	Phe	Pro
					85					90					95	
	Pro	Asp	Ser	Asp	Leu	Leu	Leu	Lys	Ala	Ala	Ser	Thr	Tyr	His	Arg	Phe
				100					105					110		
20	Leu	Thr	Trp	Pro	Val	Pro	Tyr	Cys	Ser	Ile	Cys	Gln	Glu	Leu	Leu	Thr
			115					120					125			
	Phe	Ile	Asp	Ala	Glu	Leu	Lys	Ala	Pro	Gly	Ile	Ser	Tyr	Gln	Arg	Leu
	130						135					140				
	Val	Arg	Ala	Glu	Gln	Gly	Leu	Pro	Ile	Arg	Ser	His	Arg	Ser	Ser	Thr
25	145				150						155					160
	Val	Thr	Val	Leu	Leu	Leu	Asn	Pro	Val	Glu	Val	Gln	Ala	Glu	Phe	Leu
				165						170					175	
	Ala	Val	Ala	Asn	Lys	Leu	Ser	Thr	Pro	Gly	His	Ser	Pro	His	Ser	Ala
				180					185					190		
30	Tyr	Thr	Thr	Leu	Leu	Leu	His	Ala	Phe	Gln	Ala	Thr	Phe	Gly	Ala	His
			195					200					205			
	Cys	Asp	Val	Pro	Gly	Leu	His	Cys	Arg	Leu	Gln	Ala	Lys	Thr	Leu	Ala
	210						215					220				
35	Glu	Leu	Glu	Asp	Ile	Phe	Thr	Glu	Thr	Ala	Glu	Ala	Gln	Glu	Leu	Ala
	225				230					235						240
	Ser	Gly	Ile	Gly	Asp	Ala	Ala	Glu	Ala	Arg	Arg	Trp	Leu	Arg	Thr	Lys
				245						250				255		
40	Leu	Gln	Ala	Val	Gly	Glu	Lys	Ala	Gly	Phe	Pro	Gly	Val	Leu	Asp	Thr
				260					265					270		
	Ala	Lys	Pro	Gly	Lys	Leu	His	Thr	Ile	Pro	Ile	Pro	Val	Ala	Arg	Cys
			275					280					285			
	Tyr	Thr	Tyr	Ser	Trp	Ser	Gln	Asp	Ser	Phe	Asp	Ile	Leu	Gln	Glu	Ile
45	290						295					300				
	Leu	Leu	Lys	Glu	Gln	Glu	Leu	Leu	Gln	Pro	Gly	Ile	Leu	Gly	Asp	Asp
	305				310						315					320
	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Val	Glu	Glu	Asp	Leu	Glu	Thr
				325						330				335		
50	Asp	Gly	His	Cys	Ala	Glu	Arg	Asp	Ser	Leu	Leu	Ser	Thr	Ser	Ser	Leu
				340					345					350		

55

60

65

ES 2 278 392 T3

	Ala	Ser	His	Asp	Ser	Thr	Leu	Ser	Leu	Ala	Ser	Ser	Gln	Ala	Ser	Gly
			355					360					365			
5	Pro	Ala	Leu	Ser	Arg	His	Leu	Leu	Thr	Ser	Phe	Val	Ser	Gly	Leu	Ser
		370					375					380				
	Asp	Gly	Met	Asp	Ser	Gly	Tyr	Val	Glu	Asp	Ser	Glu	Glu	Ser	Ser	Ser
	385					390					395					400
	Glu	Trp	Pro	Trp	Arg	Arg	Gly	Ser	Gln	Glu	Arg	Arg	Gly	His	Arg	Arg
					405					410					415	
10	Pro	Gly	Gln	Lys	Phe	Ile	Arg	Ile	Tyr	Lys	Leu	Phe	Lys	Ser	Thr	Ser
				420					425						430	
	Gln	Leu	Val	Leu	Arg	Arg	Asp	Ser	Arg	Ser	Leu	Glu	Gly	Ser	Ser	Asp
			435					440					445			
15	Thr	Ala	Leu	Pro	Leu	Arg	Arg	Ala	Gly	Ser	Leu	Cys	Ser	Pro	Leu	Asp
		450					455					460				
	Glu	Pro	Val	Ser	Pro	Pro	Ser	Arg	Ala	Gln	Arg	Ser	Arg	Ser	Leu	Pro
	465					470					475					480
	Gln	Pro	Lys	Leu	Gly	Thr	Gln	Leu	Pro	Ser	Trp	Leu	Leu	Ala	Pro	Ala
					485					490					495	
20	Ser	Arg	Pro	Gln	Arg	Arg	Arg	Pro	Phe	Leu	Ser	Gly	Asp	Glu	Asp	Pro
				500					505					510		
	Lys	Ala	Ser	Thr	Leu	Arg	Val	Val	Val	Phe	Gly	Ser	Asp	Arg	Ile	Ser
		515					520						525			
25	Gly	Lys	Val	Ala	Arg	Ala	Tyr	Ser	Asn	Leu	Arg	Arg	Leu	Glu	Asn	Asn
		530					535					540				
	Arg	Pro	Leu	Leu	Thr	Arg	Phe	Phe	Lys	Leu	Gln	Phe	Phe	Tyr	Val	Pro
	545					550					555					560
	Val	Lys	Arg	Ser	His	Gly	Thr	Ser	Pro	Gly	Ala	Cys	Pro	Pro	Pro	Arg
					565					570					575	
30	Ser	Gln	Thr	Pro	Ser	Pro	Pro	Thr	Asp	Ser	Pro	Arg	His	Ala	Ser	Pro
				580					585					590		
	Gly	Glu	Leu	Gly	Thr	Thr	Pro	Trp	Glu	Glu	Ser	Thr	Asn	Asp	Ile	Ser
			595					600					605			
35	His	Tyr	Leu	Gly	Met	Leu	Asp	Pro	Trp	Tyr	Glu	Arg	Asn	Val	Leu	Gly
		610					615					620				
	Leu	Met	His	Leu	Pro	Pro	Glu	Val	Leu	Cys	Gln	Gln	Ser	Leu	Lys	Ala
	625					630					635					640
	Glu	Ala	Gln	Ala	Leu	Glu	Gly	Ser	Pro	Thr	Gln	Leu	Pro	Ile	Leu	Ala
					645					650					655	
40	Asp	Met	Leu	Leu	Tyr	Tyr	Cys	Arg	Phe	Ala	Ala	Arg	Pro	Val	Leu	Leu
					660				665					670		
	Gln	Leu	Tyr	Gln	Thr	Glu	Leu	Thr	Phe	Ile	Thr	Gly	Glu	Lys	Thr	Thr
			675				680						685			
45	Glu	Ile	Phe	Ile	His	Ser	Leu	Glu	Leu	Gly	His	Ser	Ala	Ala	Thr	Arg
		690					695					700				
	Ala	Ile	Lys	Ala	Ser	Gly	Pro	Gly	Ser	Lys	Arg	Leu	Gly	Ile	Asp	Gly
	705					710					715					720
	Asp	Arg	Glu	Ala	Val	Pro	Leu	Thr	Leu	Gln	Ile	Ile	Tyr	Ser	Lys	Gly
					725					730					735	
50	Ala	Ile	Ser	Gly	Arg	Ser	Arg	Trp	Ser	Asn	Leu	Glu	Lys	Val	Cys	Thr
				740					745					750		
	Ser	Val	Asn	Leu	Asn	Lys	Ala	Cys	Arg	Lys	Gln	Glu	Glu	Leu	Asp	Ser
			755					760					765			
55	Ser	Met	Glu	Ala	Leu	Thr	Leu	Asn	Leu	Thr	Glu	Val	Val	Lys	Arg	Gln
		770					775					780				
	Asn	Ser	Lys	Ser	Lys	Lys	Gly	Phe	Asn	Gln	Ile	Ser	Thr	Ser	Gln	Ile
	785					790					795					800
	Lys	Val	Asp	Lys	Val	Gln	Ile	Ile	Gly	Ser	Asn	Ser	Cys	Pro	Phe	Ala
					805					810					815	
60	Val	Cys	Leu	Asp	Gln	Asp	Glu	Arg	Lys	Ile	Leu	Gln	Ser	Val	Val	Arg
				820					825					830		
	Cys	Glu	Val	Ser	Pro	Cys	Tyr	Lys	Pro	Glu	Lys	Ser	Asp	Leu	Ser	Ser
			835					840					845			
65	Pro	Pro	Gln	Thr	Pro	Pro	Asp	Leu	Pro	Ala	Gln	Ala	Ala	Pro	Asp	Leu
		850					855					860				
	Cys	Ser	Leu	Leu	Cys	Leu	Pro	Ile	Met	Thr	Phe	Ser	Gly	Ala	Leu	Pro
	865					870					875					880

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:13:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 5162 pares de bases

(B) TIPO: Ácido nucléico

(C) CADENA: Desconocida

(D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:13:

15	TTGAATTTG	TTTTGTTTTC	AAAAATTAAA	CAAATGATCC	TTCAGCATCA	TCGCCTCCGC	60
	TGCTTTATCA	GGTCGCATAG	GGCATGGAGC	TGGAGAACTA	TAAACAGCCC	GTGGTGCTGA	120
	GAGAGGACAA	CTGCCGAAGG	CGCCGGAGGA	TGAAGCCGCG	CAGTGCTGCG	GCCAGCCTGT	180
	CCTCCATGGA	GCTCATCCCC	ATCGAGTTTCG	TGCTGCCAC	CAGCCAGCGC	AAATGCAAGA	240
	GCCCCGAAAC	GGCGCTGCTG	CACGTGGCCG	GCCACGGCAA	CGTGGAGCAG	ATGAAGGCCC	300
	AGGTGTGGCT	GCGAGCGCTG	GAGACCAGCG	TGGCGGCGGA	CTTCTACCAC	CGGCTGGGAC	360
20	CGCATCACTT	CCTCCTGCTC	TATCAGAAGA	AGGGGCAGTG	GTACGAGATC	TACGACAAGT	420
	ACCAGGTGGT	GCAGACTCTG	GACTGCCTGC	GCTACTGGAA	GGCCACGCAC	CGGAGCCCCG	480
	GCCAGATCCA	CCTGGTGCAG	CGGCACCCGC	CCTCCGAGGA	GTCCCAAGCC	TTCCAGCGGC	540
	AGCTCACGGC	GCTGATTGGC	TATGACGTCA	CTGACGTCAG	CAACGTGCAC	GACGATGAGC	600
	TGGAGTTCAC	GCGCCGTGGC	TTGGTGACCC	CGCGCATGGC	GGAGGTGGCC	AGCCGCGACC	660
25	CCAAGCTCTA	CGCCATGCAC	CCGTGGGTGA	CGTCCAAGCC	CCTCCCGGAG	TACCTGTGGA	720
	AGAAGATTGC	CAACAACCTG	ATCTTCATCG	TCATTACCCG	CAGCACCACC	AGCCAGACCA	780
	TTAAGGTCTC	ATCCGACGAC	ACCCCGGGCG	CCATCCTGCA	GAGCTTCTTC	ACCAAGATGG	840
	CCAAGAAGAA	ATCTCTGATG	GATATTCCCG	AAAGCCAAAG	CGAACAGGAT	TTTGTGCTGC	900
	GCGTCTGTGG	CCGGGATGAG	TACCTGGTGG	GCGAAACGCC	CATCAAAAAC	TTCCAGTGGG	960
30	TGAGGCACTG	CCTCAAGAAC	GGAGAAGAGA	TTACAGTGGT	ACTGGACACG	CCTCCAGACC	1020
	CGGCCCTAGA	CGAGGTGAGG	AAGGAAGAGT	GGCCGCTGGT	GGACGACTGC	ACGGGAGTCA	1080
	CCGGCTACCA	TGAGCAGCTT	ACCATCCACG	GCAAGGACCA	CGAGAGTGTG	TTACCCGTGT	1140
	CCCTGTGGGA	CTGCGACCGC	AAGTTCAGGG	TCAAGATCAG	AGGCATTGAT	ATCCCCGTTC	1200
	TGCCTCGGAA	CACCGACCTC	ACAGTTTTTG	TAGAGGCAAA	CATCCAGCAT	GGGCAACAAG	1260
35	TCCTTTGCCA	AAGGAGAACC	AGCCCCAAAC	CCTTCACAGA	GGAGGTGCTG	TGGAATGTGT	1320
	GGCTTGAGTT	CAGTATCAAA	ATCAAAGACT	TGCCCAAAGG	GGCTCTACTG	AACCTCCAGA	1380
	TCTACTGCGG	TAAAGCTCCA	GCACTGTCCA	GCAAGGCCTC	TGCAGAGTCC	CCCAGTTCTG	1440
	AGTCCAAGGG	CAAAGTTCGG	CTTCTCTATT	ATGTGAACCT	GCTGCTGATA	GACCACCGTT	1500
	TCCTCCTGCG	CCGTGGAGAA	TACGTCTCTC	ACATGTGGCA	GATATCTGGG	AAGGGAGAAG	1560
40	ACCAAGGAAG	CTTCAATGCT	GACAAACTCA	CGTCTGCAAC	TAACCCAGAC	AAGGAGAACT	1620
	CAATGTCCAT	CTCCATTCTT	CTGGACAATT	ACTGCCACCC	GATAGCCCTG	CCTAAGCATC	1680
	AGCCACCCCC	TGACCCGGAA	GGGGACCGGG	TTGAGCAGA	AATGCCCAAC	CAGCTTCGCA	1740
	AGCAATTGGA	GGCGATCATA	GCCACTGATC	CACTTAACCC	TCTCACAGCA	GAGGACAAAG	1800
	AAATGCTCTG	CCTTCTTAGA	TACGAAAGCC	TTAAGCACCC	AAAAGCATAT	CCTAAGCTAT	1860
45	TTAGTTCAGT	GAAATGGGGA	CAGCAAGAAA	TTGTGGCCAA	AACATACCAA	TTGTTGGCCA	1920
	GAAGGGAAGT	CTGGGATCAA	AGTGCTTTGG	ATGTTGGGTT	AACAATGCAG	CTCCTGGACT	1980
	GCAACTTCTC	AGATGAAAAT	GTAAGAGCCA	TTGCAGTTCA	GAAACTGGAG	AGCTTGGAGG	2040
	ACGATGATGT	TCTGCATTAC	CTTCTACAAT	TGGTCCAGGC	TGTGAAATTT	GAACCATAAC	2100
	ATGATAGCGT	CCTTGCCAGA	TTTCTGTGTA	AGCTGTGGTT	AAGAAACAAA	AGAATTGGTC	2160
	ACTTTTGTGT	TTGGTCTTTG	AGAAGTGAGA	TAGCCAGTCC	CAGACACTAT	CAGCAGAGGT	2220
50	TCGCTGTGAT	TCTGGAAGCC	TATCTGAGGG	GCTGTGGCAC	AGCCATGCTG	CACGACTTTA	2280
	CCCAACAAGT	CCAAGTAATC	GAGATGTTAC	AAAAAGTCAC	CCTTGATATT	AAATCGCTCT	2340
	CTGCTGAAAA	GTATGACGTC	AGTTCCCAAG	TTATTTTACA	ACTTAAACAA	AAGCTTGAAA	2400
	ACCTGCAGAA	TTCTCAACTC	CCCCGAAAGT	TTAGATTTCC	ATATGATCCT	GGACTGAAAG	2460
	CAGGAGCGCT	GGCAATTGAA	AAATGTAAAG	TAATGGCCTC	CAAGAAAAAA	CCACTATGGC	2520
55	TTGAGTTTAA	ATGTGCCGAT	CCTACAGCCC	TATCAAATGA	AACAATTGGA	ATTATCTTTA	2580
	AACATGGTGA	TGATCTGCGC	CAAGACATGC	TTATTTTACA	GATTCTACGA	ATCATGGAGT	2640
	CTATTTGGGA	GACTGAATCT	TTGGATCTAT	GCCTCCTGCC	ATATGGTTGC	ATTTCAACTG	2700
	GTGACAAAAA	AGGAATGATC	GAGATTGTGA	AAGACGCCAC	GACAATTGCC	AAAAATTCAGC	2760
	AAAGCACAGT	GGGCAACACG	GGAGCATTTA	AAGATGAAGT	CCTGAAATCAC	TGGCTCAAAG	2820
60	AAAAATCCCC	TACTGAAGAA	AAGTTTCAGG	CAGCAGTGGA	GAGATTGTGT	TATTCCTGTG	2880
	CAGGCTACTG	TGTGGCAACC	TTTGTTCCTG	GAATAGGCGA	CAGACACAAT	GACAAATATTA	2940
	TGATCACCBA	GACAGGAAAC	CTATTTTATA	TTGACTTCGG	GCACATTCTT	GGGAATTACA	3000
	AAAGTTTCCT	GGGCATTAAT	AAAGAGAGAG	TGCCATTTGT	GCTAACCCCT	GACTTCCTCT	3060
	TTGTGATGGG	AACCTCTGGA	AAGAAGACAA	CCCCACACTT	CCAGAAATTT	CAGGACATCT	3120
65	GTGTTAAGGC	TTATCTAGCC	CTTCGTCTATC	ACACAAACCT	ACTGATCATC	CTGTTCTCCA	3180
	TGATGCTGAT	GACAGGAATG	CCCCAGTTAA	CAAGCAAAGA	AGACATTGAA	TATATCCGGG	3240
	ATGCCCTCAC	AGTGGGGAAA	AATGAGGAGG	ATGCTAAAAA	GTATTTTCTT	GATCAGATCG	3300

ES 2 278 392 T3

	AAGTTTGCAG	AGACAAAGGA	TGGACTGTGC	AGTTTAATTG	GTTTCTACAT	CTTGTTCTTG	3360
	GCATCAAACA	AGGAGAGAAA	CATTACAGCT	AATACTTTAG	GCTAGAATCA	AAAACAAGTT	3420
	AGTGTTCAT	GGTTTAAATT	AGCATAGCAA	TCATCGAAGT	TGGATTTCAG	ATGCAATAGA	3480
	CATTGTGAAA	GCTGGCATT	CAGAAGTATA	GCTCTTTTCC	TACCTGAAGT	CTTCCCTGGA	3540
5	GAAAAGATGT	TGGCATTGCT	GATTGTTTGG	TTAAGCAATG	TCCAGTGCTA	GGATTATTTG	3600
	CAGGTTTGGT	TTTTTCTCAT	TTGTCTGTGG	CATTGGAGAA	TATTCTTGGT	TTAAACAGAC	3660
	TAATGACTTC	CTTATTGTCC	CTGATATTTT	GACTATCTTA	CTATTGAGTG	CTTCTGGAAA	3720
	TTCTTTGGAA	TAATTGATGA	CATCTATTTT	CATCTGGGTT	TAGTCTCAAT	TTTGGTTATC	3780
	TTTGTGTTCC	TCAAGCTCTT	TAAAGAAAAA	GATGTAATCG	TTGTAACCTT	TGTCTCATT	3840
	CTTAAATGAT	GCTTCCAAAC	ATCTCCTTAG	TGTCTGCAGG	TGTTAGTGGT	GTGCTAAAAG	3900
10	CAAGGAAAGC	GAGTTAGTCT	TTTCAGTGTC	TTTTGCAATT	CAATTCCTTT	GTCATGTATA	3960
	ACTGAGACAC	ACAAACACAG	CAGGAGAAAT	CTAAACCGTT	GTGCCTTGAC	CTTCTCTGCT	4020
	TGGTCTTGTT	CCAGGGTTAT	GAATATGAAA	AAATAGAGAT	GAGACTTTTT	GTGTCAACTC	4080
	TGTCCACAAG	AGTGAGTTAT	CTAGTATGAT	TAGTATAGCT	TTCTCCAGCA	TGGCAGCAGG	4140
	AAGTAACTAC	AGGGCCTCTT	TTATGCCTGA	CATTTCTTCC	CTTCTCTTTT	CCCTGCCTCC	4200
	CTTTTTCATC	AATTGCAATG	CTCCCAACA	TCTTTACAGA	CTTGTGAAAT	CTTCAAGAAC	4260
15	ACCTTTACTC	TATAACTCAA	AAATTAGTTG	AAAAATAATT	ACTTCTCAAG	GATTATTAGA	4320
	ATCTTTAGTA	CTTATTGTA	AAGATGTTTA	GTGACTTTTT	TTTCAAGTAT	CTATAAAGGA	4380
	GGCAGATTCT	AGAAAAATATG	AATTAGTTTC	CAAATGCCTT	AATTTTAAAC	TTTGGCCTGA	4440
	ACAGTTTTTT	CTTTTTCTTA	ATGGAAGAAG	ATATTTAATA	TCTTAAAAAT	ATTCCAAGTT	4500
	AGGAAGAACA	CTACTTGCCT	TATCCATTTT	CCATTTAAAG	GACTTTTAAA	CTTTGACACA	4560
20	GTCCTTCAGA	TTTCTTGAAA	ATCCTTGAAA	TATCTTACTT	TAAAAATATT	TTCATCTCTG	4620
	AAATATCTCG	TTATTTATTG	GAGGTATTGT	TTAACCTTAG	ATAGACCATT	AAATTATTTA	4680
	TAAAAATATT	TGTAATTACC	TGTAGYTAAT	ACATTACATA	GAAAAAACTA	TGTTAACAGT	4740
	GTCTCTGTTT	AAGTATAATC	AGATATAAAT	ATATAACTTA	ATTTTTTAAT	TTTAAAAATA	4800
	GATACCTGTT	GTAATTTGAG	GTAGTCCAGA	CCTTTTCTTT	TTTTTTTTTT	TTTTTAATGT	4860
25	GTGCAAAAGC	CCAAAGGTTT	CTAAGCCTGG	CTGCAAGAA	GAATCAACAG	GGACACTTTT	4920
	TAAAAACACT	CTTATCAGCC	TGGGCAACAC	AGTGAGACTC	CATCTCTTAA	AAAAAAAATT	4980
	AGCTGGGTAT	AGTGGTATGT	GCCTGTAGTC	CCAGGTACTC	AGGAGGCTGA	SGCAGGAGGA	5040
	TTGCCTGAGC	CCAGGAGGTG	GAAACTGCAG	AGAGTCATGA	TCATGTCCTT	ACACTCCAGC	5100
	CTGATAACA	GAGCGAGACC	CTGTCTCAAA	AAAAAAAATA	AAAAAAAATA	AAGTCGACCG	5160
	AG						5162

(2) INFORMACIÓN PARA LA SECUENCIA ID NO:14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1101 aminoácidos
- (B) TIPO: Aminoácido
- (C) CADENA: Desconocida
- (D) TOPOLOGÍA: Desconocida

(ii) TIPO MOLECULAR: Proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:14:

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

ES 2 278 392 T3

	Arg	Gly	Leu	Val	Thr	Pro	Arg	Met	Ala	Glu	Val	Ala	Ser	Arg	Asp	Pro
				180					185					190		
5	Lys	Leu	Tyr	Ala	Met	His	Pro	Trp	Val	Thr	Ser	Lys	Pro	Leu	Pro	Glu
			195					200					205			
	Tyr	Leu	Trp	Lys	Lys	Ile	Ala	Asn	Asn	Cys	Ile	Phe	Ile	Val	Ile	His
		210					215					220				
	Arg	Ser	Thr	Thr	Ser	Gln	Thr	Ile	Lys	Val	Ser	Pro	Asp	Asp	Thr	Pro
	225					230					235				240	
10	Gly	Ala	Ile	Leu	Gln	Ser	Phe	Phe	Thr	Lys	Met	Ala	Lys	Lys	Lys	Ser
					245					250					255	
	Leu	Met	Asp	Ile	Pro	Glu	Ser	Gln	Ser	Glu	Gln	Asp	Phe	Val	Leu	Arg
				260					265					270		
	Val	Cys	Gly	Arg	Asp	Glu	Tyr	Leu	Val	Gly	Glu	Thr	Pro	Ile	Lys	Asn
			275					280					285			
15	Phe	Gln	Trp	Val	Arg	His	Cys	Leu	Lys	Asn	Gly	Glu	Glu	Ile	His	Val
		290					295				300					
	Val	Leu	Asp	Thr	Pro	Pro	Asp	Pro	Ala	Leu	Asp	Glu	Val	Arg	Lys	Glu
	305					310					315					320
20	Glu	Trp	Pro	Leu	Val	Asp	Asp	Cys	Thr	Gly	Val	Thr	Gly	Tyr	His	Glu
					325					330					335	
	Gln	Leu	Thr	Ile	His	Gly	Lys	Asp	His	Glu	Ser	Val	Phe	Thr	Val	Ser
				340					345					350		
	Leu	Trp	Asp	Cys	Asp	Arg	Lys	Phe	Arg	Val	Lys	Ile	Arg	Gly	Ile	Asp
			355					360					365			
25	Ile	Pro	Val	Leu	Pro	Arg	Asn	Thr	Asp	Leu	Thr	Val	Phe	Val	Glu	Ala
		370					375					380				
	Asn	Ile	Gln	His	Gly	Gln	Gln	Val	Leu	Cys	Gln	Arg	Arg	Thr	Ser	Pro
	385					390					395					400
30	Lys	Pro	Phe	Thr	Glu	Glu	Val	Leu	Trp	Asn	Val	Trp	Leu	Glu	Phe	Ser
				405						410					415	
	Ile	Lys	Ile	Lys	Asp	Leu	Pro	Lys	Gly	Ala	Leu	Leu	Asn	Leu	Gln	Ile
				420					425					430		
	Tyr	Cys	Gly	Lys	Ala	Pro	Ala	Leu	Ser	Ser	Lys	Ala	Ser	Ala	Glu	Ser
			435					440					445			
35	Pro	Ser	Ser	Glu	Ser	Lys	Gly	Lys	Val	Arg	Leu	Leu	Tyr	Tyr	Val	Asn
		450					455					460				
	Leu	Leu	Leu	Ile	Asp	His	Arg	Phe	Leu	Leu	Arg	Arg	Gly	Glu	Tyr	Val
	465					470					475					480
40	Leu	His	Met	Trp	Gln	Ile	Ser	Gly	Lys	Gly	Glu	Asp	Gln	Gly	Ser	Phe
				485						490					495	
	Asn	Ala	Asp	Lys	Leu	Thr	Ser	Ala	Thr	Asn	Pro	Asp	Lys	Glu	Asn	Ser
			500						505					510		
	Met	Ser	Ile	Ser	Ile	Leu	Leu	Asp	Asn	Tyr	Cys	His	Pro	Ile	Ala	Leu
			515					520					525			
45	Pro	Lys	His	Gln	Pro	Thr	Pro	Asp	Pro	Glu	Gly	Asp	Arg	Val	Arg	Ala
		530					535					540				
	Glu	Met	Pro	Asn	Gln	Leu	Arg	Lys	Gln	Leu	Glu	Ala	Ile	Ile	Ala	Thr
	545					550					555					560
50	Asp	Pro	Leu	Asn	Pro	Leu	Thr	Ala	Glu	Asp	Lys	Glu	Leu	Leu	Trp	His
				565					570						575	
	Phe	Arg	Tyr	Glu	Ser	Leu	Lys	His	Pro	Lys	Ala	Tyr	Pro	Lys	Leu	Phe
			580						585					590		
	Ser	Ser	Val	Lys	Trp	Gly	Gln	Gln	Glu	Ile	Val	Ala	Lys	Thr	Tyr	Gln
			595					600					605			
55	Leu	Leu	Ala	Arg	Arg	Glu	Val	Trp	Asp	Gln	Ser	Ala	Leu	Asp	Val	Gly
		610					615					620				
	Leu	Thr	Met	Gln	Leu	Leu	Asp	Cys	Asn	Phe	Ser	Asp	Glu	Asn	Val	Arg
	625					630					635					640
60	Ala	Ile	Ala	Val	Gln	Lys	Leu	Glu	Ser	Leu	Glu	Asp	Asp	Asp	Val	Leu
				645						650					655	
	His	Tyr	Leu	Leu	Gln	Leu	Val	Gln	Ala	Val	Lys	Phe	Glu	Pro	Tyr	His
			660						665					670		
	Asp	Ser	Ala	Leu	Ala	Arg	Phe	Leu	Leu	Lys	Arg	Gly	Leu	Arg	Asn	Lys
			675					680					685			
65	Arg	Ile	Gly	His	Phe	Leu	Phe	Trp	Phe	Leu	Arg	Ser	Glu	Ile	Ala	Gln
		690					695					700				
	Ser	Arg	His	Tyr	Gln	Gln	Arg	Phe	Ala	Val	Ile	Leu	Glu	Ala	Tyr	Leu

ES 2 278 392 T3

	705				710				715				720
	Arg	Gly	Cys	Gly	Thr	Ala	Met	Leu	His	Asp	Phe	Thr	Gln
					725					730			Gln
5	Val	Ile	Glu	Met	Leu	Gln	Lys	Val	Thr	Leu	Asp	Ile	Lys
				740					745				Ser
	Ala	Glu	Lys	Tyr	Asp	Val	Ser	Ser	Gln	Val	Ile	Ser	Gln
			755					760				765	Leu
10	Lys	Leu	Glu	Asn	Leu	Gln	Asn	Ser	Gln	Leu	Pro	Glu	Ser
		770					775					780	Phe
	Pro	Tyr	Asp	Pro	Gly	Leu	Lys	Ala	Gly	Ala	Leu	Ala	Ile
		785				790				795			Glu
15	Lys	Val	Met	Ala	Ser	Lys	Lys	Lys	Pro	Leu	Trp	Leu	Glu
				805						810			Phe
	Ala	Asp	Pro	Thr	Ala	Leu	Ser	Asn	Glu	Thr	Ile	Gly	Ile
				820					825				Ile
	His	Gly	Asp	Asp	Leu	Arg	Gln	Asp	Met	Leu	Ile	Leu	Gln
			835					840					Ile
20	Ile	Met	Glu	Ser	Ile	Trp	Glu	Thr	Glu	Ser	Leu	Asp	Leu
		850					855					860	Cys
	Pro	Tyr	Gly	Cys	Ile	Ser	Thr	Gly	Asp	Lys	Ile	Gly	Met
		865				870					875		Ile
25	Val	Lys	Asp	Ala	Thr	Thr	Ile	Ala	Lys	Ile	Gln	Gln	Ser
				885						890			Thr
	Asn	Thr	Gly	Ala	Phe	Lys	Asp	Glu	Val	Leu	Asn	His	Trp
				900					905				Leu
30	Lys	Ser	Pro	Thr	Glu	Glu	Lys	Phe	Gln	Ala	Ala	Val	Glu
			915					920					Arg
	Tyr	Ser	Cys	Ala	Gly	Tyr	Cys	Val	Ala	Thr	Phe	Val	Leu
		930					935				940		Gly
	Asp	Arg	His	Asn	Asp	Asn	Ile	Met	Ile	Thr	Glu	Thr	Gly
35						950					955		Asn
	His	Ile	Asp	Phe	Gly	His	Ile	Leu	Gly	Asn	Tyr	Lys	Ser
				965						970			Phe
	Ile	Asn	Lys	Glu	Arg	Val	Pro	Phe	Val	Leu	Thr	Pro	Asp
				980					985				Phe
40	Val	Met	Gly	Thr	Ser	Gly	Lys	Lys	Thr	Ser	Pro	His	Phe
			995					1000					Gln
	Gln	Asp	Ile	Cys	Val	Lys	Ala	Tyr	Leu	Ala	Leu	Arg	His
		1010					1015				1020		His
45	Leu	Leu	Ile	Ile	Leu	Phe	Ser	Met	Met	Leu	Met	Thr	Gly
		025				1030				1035			Met
	Leu	Thr	Ser	Lys	Glu	Asp	Ile	Glu	Tyr	Ile	Arg	Asp	Ala
				1045					1050				Leu
50	Gly	Lys	Asn	Glu	Glu	Asp	Ala	Lys	Lys	Tyr	Phe	Leu	Asp
				1060					1065				Gln
	Val	Cys	Arg	Asp	Lys	Gly	Trp	Thr	Val	Gln	Phe	Asn	Trp
		1075						1080				1085	Phe
	Leu	Val	Leu	Gly	Ile	Lys	Gln	Gly	Glu	Lys	His	Ser	Ala
55		1090					1095					1100	
60													
65													