

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 374 392**

51 Int. Cl.:

C12N 9/64

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03782702 .9**

96 Fecha de presentación: **23.12.2003**

97 Número de publicación de la solicitud: **1583825**

97 Fecha de publicación de la solicitud: **12.10.2005**

54 Título: **PROTEÍNAS DE METALOPROTEASA.**

30 Prioridad:
23.12.2002 GB 0230006

45 Fecha de publicación de la mención BOPI:
16.02.2012

45 Fecha de la publicación del folleto de la patente:
16.02.2012

73 Titular/es:
**ARES TRADING S.A.
ZONE INDUSTRIELLE DE L'OURIETTAZ
1170 AUBONNE, CH**

72 Inventor/es:
**FAGAN, Richard Joseph;
PHELPS, Christopher Benjamin;
POWER, Christine;
MITTER, Richard James;
BOSCHERT, Ursula y
CHVATCHKO, Yoïande**

74 Agente: **de Elzaburu Márquez, Alberto**

ES 2 374 392 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Proteínas de metaloproteasa

5 Esta invención se refiere a proteínas nuevas, denominadas INSP005a e INSP005b, identificadas en la presente memoria como proteínas secretadas, en particular miembros de la familia de metaloproteasas, y al uso de estas proteínas y secuencias de ácidos nucleicos de los genes codificantes en el diagnóstico, la prevención y el tratamiento de enfermedades.

Antecedentes

10 El proceso de descubrimiento de fármacos está experimentando en la actualidad una revolución fundamental, ya que la era de la genómica funcional ha alcanzado la mayoría de edad. La expresión "genómica funcional" se aplica a una aproximación que utiliza herramientas bioinformáticas para asignar una función a secuencias proteicas de interés. Tales herramientas se están haciendo cada vez más necesarias, ya que la velocidad de generación de datos de secuencias está dejando atrás rápidamente la capacidad de los laboratorios de investigación de asignar funciones a estas secuencias proteicas.

15 A medida que las herramientas bioinformáticas incrementan su potencia y su exactitud, estas herramientas están sustituyendo rápidamente a las técnicas convencionales de caracterización bioquímica. De hecho, las herramientas bioinformáticas avanzadas usadas en la identificación de la presente invención son capaces ya de producir resultados a los que se puede otorgar un grado elevado de confianza.

20 Diversas instituciones y organizaciones comerciales están examinando datos de secuencias a medida que aparecen, y se están realizando descubrimientos significativos continuamente. Sin embargo, sigue existiendo una necesidad continua de identificar y caracterizar más genes y los polipéptidos que codifican, como objetivos para la investigación y para el descubrimiento de fármacos.

Antecedentes de las proteínas secretadas

30 La capacidad de que las células produzcan y secreten proteínas extracelulares es fundamental para muchos procesos biológicos. Las células secretan enzimas, factores de crecimiento, proteínas de la matriz extracelular y moléculas de señalización. Esto se da a través de la fusión de una vesícula secretora con la membrana plasmática. En la mayoría de los casos, pero no en todos, las proteínas se dirigen al retículo endoplásmico y a las vesículas secretoras mediante un péptido señal. Los péptidos señal son secuencias que actúan en cis que afectan al transporte de las cadenas polipeptídicas desde el citoplasma hacia un compartimento asociado a la membrana, tal como una vesícula secretora. Los polipéptidos que se dirigen a las vesículas secretoras se secretan a la matriz extracelular o se retienen en la membrana plasmática. Los polipéptidos que se retienen en la membrana plasmática tendrán uno o más dominios transmembrana. Los ejemplos de proteínas secretadas que desempeñan un papel fundamental en el funcionamiento de una célula son las citocinas, hormonas, proteínas de la matriz extracelular (moléculas de adhesión), proteasas, y factores de crecimiento y diferenciación. A continuación se da una descripción de algunas de las propiedades de estas proteínas.

45 Las proteasas son enzimas que hidrolizan de manera irreversible enlaces amida en péptidos y proteínas. Las proteasas están ampliamente distribuidas, y están implicadas en muchos procesos biológicos diferentes, desde la activación de proteínas y péptidos hasta la degradación de proteínas. A pesar del hecho de que se ha demostrado que las proteasas están implicadas en muchas enfermedades diferentes, los fármacos dirigidos hacia las proteasas todavía son escasos en farmacia, aunque los inhibidores de la enzima convertidora de angiotensina (ACE) han estado entre los fármacos antihipertensivos más eficaces durante varios años. Las proteasas han recibido recientemente una publicidad considerable como objetivos terapéuticos valiosos tras la aprobación de los inhibidores de la proteasa de VIH.

50 Las proteasas se pueden dividir en grandes Familias. El término "Familia" se usa para describir un grupo de proteasas en que cada miembro muestra una relación evolutiva hacia al menos otro miembro, a lo largo de toda la secuencia o al menos en parte de la secuencia responsable de la actividad catalítica. El nombre de cada Familia refleja el tipo de actividad catalítica de las proteasas de la Familia. Así, las serín proteasas pertenecen a la familia S, las treonín proteasas pertenecen a la familia T, las aspartil proteasas pertenecen a la familia A, las cisteín proteasas pertenecen a la familia C y las metaloproteinasas pertenecen a la familia M. Las metaloproteasas y las Serín proteasas se hallan habitualmente en la matriz extracelular.

60 Metaloproteasas (Familia M):

Las metaloproteasas se pueden dividir en 2 grupos principales dependiendo de la presencia o ausencia de un motivo de unión de Zinc (HEXXH).

1.1 Presencia del motivo HEXXH (22 familias): Número de Prosite: PDOC00129Familias con miembros interesantes:

5 **M2:** Peptidil-dipeptidasa A (Enzima Conversora de Angiotensina I: ACE)

M13: Neprilisin (Encefalinasa A=endopeptidasa neutra=NEP), enzima conversora de endotelina (ECE)

10 **M10B:** Matrixina (Metaloproteasas de la Matriz=MMPs)

M12B: Reprolisina (ADAM-10; ADAM-17 = enzima conversora de TNF-alfa = TACE)/Desintegrina (otras proteasas ADAM). Las ADAMs son una familia de proteínas grande, expresadas de manera generalizada y reguladas durante el desarrollo con múltiples funciones potenciales en las interacciones célula-célula y célula-matriz. Entre ellas, TACE representa un nuevo objetivo emergente para la artritis.

15 **M41:** Esta familia contiene metaloproteasas dependientes de ATP: FtsH, proteínas del proteosoma.

Uno de los mayores grupos de metaloproteinasas terapéuticamente interesantes es la familia de Metaloproteinasas de la matriz (MMPs). Las metaloproteinasas de la matriz son una familia de enzimas que contienen zinc que son responsables de la remodelación de la matriz extracelular en todo el cuerpo. Se ha demostrado que están implicadas en el cáncer (incremento de la invasividad, efectos sobre los vasos sanguíneos nuevos), y en la artritis (implicación en la degradación del cartilago (Dahlberg, L., *et al.*, Arthritis Rheum. 2000 43(3):673-82) y también la conversión de TNF-alfa (Hanemaaijer, R., *et al.*, J Biol Chem. 1997 272(50):31504-9, Shlopov, B. V., *et al.*, Arthritis Rheum. 1997 40(11):2065-74)). De hecho, se ha demostrado que diferentes MMPs se sobreexpresan en enfermedades tales como artritis (Seitz, M., *et al.*, Rheumatology (Oxford). 2000 39(6):637-645, Yoshihara, Y., *et al.*, Ann Rheum Dis. 2000 59(6):455-61, Yamanaka, H., *et al.*, Lab Invest. 2000 80(5):677-87, Jovanovic, D. V., *et al.*, Arthritis Rheum. mayo de 2000; 43(5):1134-44, Ribbens, C., *et al.*, J Rheumatol. 2000 27(4):888-93) y el cáncer (Sakamoto, Y., *et al.*, Int J Oncol. 2000 17(2):237-43, Kerkela, E., *et al.*, J Invest Dermatol. 2000 114(6):1113-9, Fang, J., *et al.*, Proc Natl Acad Sci USA. 2000 97(8):3884-9, Sun, Y., *et al.*, J Biol Chem. 2000 275(15):11327-32, McCawley, L. J., *et al.*, Mol Med Today. 2000 6(4):149-56, Ara, T., *et al.*, J Pediatr Surg. 2000 35(3):432-7, Shigemasa, K., *et al.*, Med Oncol. 2000 17(1):52-8, Nakanishi, K., *et al.*, Hum Pathol. 2000 31(2):193-200, Dalberg, K., *et al.*, World J Surg. 2000 24(3):334-40). Se han propuesto los inhibidores de estas enzimas como agentes terapéuticos potenciales para el uso en el tratamiento del cáncer y la artritis. Más recientemente, se ha demostrado que los MMPs también pueden tener un papel en la liberación de receptores solubles de citocinas, factores de crecimiento y otros mediadores celulares, lo que sugiere que los inhibidores selectivos de MMPs pueden tener aplicaciones terapéuticas más amplias que las propuestas previamente.

Las MMPs se han dividido en 4 familias basándose en las homologías de las secuencias de aminoácidos de su estructura de dominios, aparte de la región catalítica.

40 Familia de dominio mínimo: matrilisina (PUMP-1, MMP-7) escinde proteoglicano, laminina y fibronectina.

Familia de dominio de hemopexina:

45 Colagenasas: capacidad única de escindir colágeno fibrilar. El papel de las colagenasas en la degradación del cartilago las hace objetivos atractivos para el tratamiento de la artritis reumatoide y osteoartritis.

■ colagenasa de fibroblastos (colagenasa intersticial, MMP-1)

50 ■ colagenasa de neutrófilos (MMP-8)

■ colagenasa-3 (MMP-13)

Metaloelastasa: MME (MMP-12)

55 Stromelisina-1 (MMP-3), 2 (MMP-10) y 3 (MMP-11). MMP-11 se excreta en forma activa y su función podría ser activar otras MMPs.

60 Familia de dominio de fibronectina: degrada un gran número de sustratos de la matriz (gelatina, elastina, colágeno de tipo IV)

Gelatinasa A (MMP-2); además de su implicación en el cáncer (invasividad tumoral), se ha propuesto como un objetivo potencial para el descubrimiento de agentes antiplaquetarios, ya que puede desempeñar un papel importante en la activación de las plaquetas.

65

Gelatinasa B (MMP-9)

Familia de dominio transmembrana:

5 MT-1-MMP, MT-4-MMP, MMP-14, MMP-17

Muchos estudios relacionados con las diferentes especificidades de las MMPs y su implicación relativa en algunas enfermedades están en desarrollo.

10 1.2 Ausencia de motivos HEXXH (18 familias):

Familias con miembros interesantes:

15 **M24A:** Metionil aminopeptidasa, tipo 1 (que incluye MAP-1 procariótica y eucariótica) / número de Prosite: PDOC00575

M24C: Metionil aminopeptidasa, tipo 2 (que incluye MAP-2 eucariótica) / número de Prosite: PDOC00575

20 **Tabla 1. Resumen de metaloproteasas y su función**

Nombre de la proteasa	Número EC	Función biológica	Enfermedad asociada	Regulación
MMP-12	3.4.24.65	función de las MMPs; degradación de elastina; procesamiento de TNF-alfa; conversión de plasminógeno a angiotensina	implicación en trastornos pulmonares, enfisema, fibrosis quística	expresión aumentada en ciertas enfermedades cutáneas
MMP-2	3.4.24.24	función de las MMPs	cáncer	sobreexpresión en el cáncer colorrectal
ADAM-12	3.4.24	interacción célula-célula, célula-matriz		aumentada en varios carcinomas humanos
TACE	3.4.24.?	Procesamiento del TNF-alfa asociado a la membrana y otras moléculas asociadas a células	inflamación, artritis reumatoide, enfermedades neuroinmunológicas	aumentada en el cartílago afectado por artritis
ACE	3.4.15.1	producción de angiotensina II	hipertensión	
ECE-1	3.4.24.71	procesamiento del precursor de la endotelina vasoconstrictora	cardiovascular	
NEP	3.4.24.11	escinde neuropéptidos, hormonas y mediadores inmunitarios	cardiovascular, artritis (?)	
FtsH	?	secreción de proteínas, ensamblaje, degradación, ciclo celular, respuesta a agresiones	infecciones bacterianas	-
Desformilasa	3.5.1.31	elimina el grupo formilo de la posición N-terminal de proteínas recién sintetizadas	infecciones bacterianas	-
Proteasoma	3.4.99.46	degradación de proteínas, presentación de antígenos	cáncer	

Las metaloproteasas están implicadas en una amplia diversidad de áreas terapéuticas. Estas incluyen las enfermedades respiratorias (Segura-Valdez, L., *et al.*, Chest. 2000 117(3):684-94, Tanaka, H., *et al.*, J Allergy Clin Immunol. 2000 105(5):900-5, Hoshino, M., *et al.*, J Allergy Clin Immunol. 1999 104(2 Pt 1):356-63, Mautino, G., *et al.*, Am J Respir Crit Care Med. 1999 160(1):324-30, Dalal, S., *et al.*, Chest. 2000 117(5 Supl. 1):227S-8S, Ohnishi, K., *et al.*, Lab Invest. 1998 78(9):1077-87), cardiopatías (Taniyama, Y., *et al.*, Circulation. 2000 102(2):246-52, Hong, B. K., *et al.*, Yonsei Med J. 2000 41(1):82-8, Galis, Z. S., *et al.*, Proc Natl Acad Sci USA. 1995 92(2):402-6), infecciones bacterianas (Scozzafava, A., *et al.*, J Med Chem. 2000 43(9):1858-65, Vencill, C. F., *et al.*, Biochemistry. 1985 24(13):3149-57, Steinbrink, D. R., *et al.*, J Biol Chem. 1985 260(5):2771-6, Lopez-Boado, Y. S., *et al.*, J Cell Biol. 2000 148(6):1305-15, Chang, J. C., *et al.*, Thorax. 1996 51(3):306-11, Dammann, T., *et al.*, Mol. Microbiol. 6:2267-2278 (1992), Wassif, C., *et al.*, J. Bacteriol. 177 (20), 5790-5798 (1995), enfermedades oncológicas (Sakamoto, Y., *et al.*, Int J Oncol. 2000 17(2):237-43, Kerkela, E., *et al.*, J Invest Dermatol. 2000 114(6):1113-9, Fang, J., *et al.*, Proc Natl Acad Sci USA. 2000 97(8):3884-9, Sun, Y., *et al.*, J Biol Chem. 2000 275(15):11327-32, McCawley, L. J., *et al.*, Mol Med Today. 2000 6(4):149-56, Ara, T., *et al.*, J Pediatr Surg. 2000 35(3):432-7, Shigemasa, K., *et al.*, Med Oncol. 2000 17(1):52-8, Nakanishi, K., *et al.*, Hum Pathol. 2000 31(2):193-200, Dalberg, K., *et al.*, World J Surg. 2000 24(3):334-40), e inflamación (artritis reumatoide y osteoartritis (Ribbens, C., *et al.*, J Rheumatol. 2000 27(4):888-93,

Kageyama, Y., *et al.*, Clin Rheumatol. 2000 19(1):14-20, Shlopov, B. V., *et al.*, Arthritis Rheum. enero de 2000; 43(1):195-205).

5 Las metaloproteasas también están implicadas en la fisiología y la patología de la reproducción sexual, y se han implicado en terapias asociadas a la modulación del estado del corion, la reacción de la zona, la formación de membranas de fecundación, la anticoncepción y la infertilidad (Shibata *et al.* (2000) J. Biol. Chem vol. 275, N° 12 pág 8349)

10 Por lo tanto, la identificación de metaloproteasas nuevas es de gran importancia en la comprensión creciente de las rutas subyacentes que conducen a ciertos estados patológicos en los que están implicadas estas proteínas, y en el desarrollo de terapias génicas o farmacológicas más eficaces para tratar estos trastornos.

LA INVENCIÓN

15 La invención se basa en el descubrimiento de que las proteínas INSP005a e INSP005b funcionan como moléculas proteasas secretadas, y además como moléculas proteasas secretadas de la familia de metaloproteasas. Preferiblemente, las proteínas INSP005a e INSP005b son miembros de la familia similar a coriolisina/astacina de metaloproteasas.

20 En una realización del primer aspecto de la invención, se proporciona un polipéptido que:

(i) comprende la secuencia de aminoácidos indicada en SEQ ID N°:14;

25 Preferiblemente, un polipéptido según esta realización consiste en la secuencia indicada en SEQ ID N°:14. El polipéptido que tiene la secuencia indicada en SEQ ID N°:14 se denomina más adelante en la presente memoria "el polipéptido INSP005a".

En una segunda realización del primer aspecto de la invención, se proporciona un polipéptido que:

30 (i) comprende la secuencia de aminoácidos indicada en SEQ ID N°:34 o SEQ ID N°:36;

35 Preferiblemente, un polipéptido según esta realización consiste en la secuencia indicada en SEQ ID N°:34 o SEQ ID N°:36. El polipéptido que tiene la secuencia indicada en SEQ ID N°:34 se denomina más adelante en la presente memoria "el polipéptido INSP005b".

Aunque el solicitante no desea limitarse a esta teoría, se postula que los primeros 23 aminoácidos del polipéptido INSP005b forman un péptido señal. La secuencia de nucleótidos que codifica el polipéptido maduro INSP005b postulado, y la secuencia de aminoácidos del polipéptido maduro INSP005b, se indican en SEQ ID N°:35 y SEQ ID N°:36, respectivamente. El polipéptido que tiene la secuencia indicada en SEQ ID N°:36 se denomina más adelante en la presente memoria "el polipéptido maduro INSP005b".

40 Un polipéptido según los aspectos anteriormente descritos de la invención funciona como una metaloproteasa. El término "metaloproteasa" se entiende bien en la técnica, y el experto podrá fácilmente determinar la actividad de metaloproteasa mediante el uso de uno de una diversidad de ensayos conocidos en la técnica. Por ejemplo, dos ensayos aplicados habitualmente son el ensayo cuantitativo de [³H] gelatina (Martin *et al.*, Kidney Int. 36, 790-801) y el ensayo de zimografía en gelatina (Herron G. S. *et al.*, J. Biol. Chem. 1986, 261, 2814-2818).

45 Más preferiblemente, un polipéptido según los aspectos anteriormente descritos de la invención es un miembro de la familia similar a coriolisina/astacina de metaloproteasas.

50 Se presentan pruebas en la sección de Ejemplos más adelante de que se descubrió que la administración de cADN de INSP005b (también denominado en la presente memoria IPAAA78836-2) en un modelo *in vivo* de hepatitis fulminante disminuía los niveles de TNF-alfa y m-IL-6 en suero y tenía un efecto significativo sobre la reducción de las transaminasas medidas en suero.

55 La disminución de los niveles de aspartato aminotransferasa (ASAT) y alanina aminotransferasa (ALAT) indicada podría ser debida a niveles disminuidos de TNF-alfa e IL-6. TNF-alfa es una citocina importante implicada en la lesión hepática tras una inyección de ConA. En este modelo de ratón de hepatitis, se produce TNF-alfa principalmente por los macrófagos hepáticos, las denominadas células de Kupfer. Se ha demostrado que los anticuerpos Anti-TNF-alfa confieren protección contra la enfermedad (Seino *et al.* 2001, Annals of surgery 234, 681). Por lo tanto, se considera que el polipéptido INSP005b y las proteínas funcionalmente equivalentes relacionadas serán útiles en el tratamiento de enfermedades hepáticas autoinmunitarias, virales o agudas, así como en insuficiencias hepáticas alcohólicas. También es probable que sean eficaces en el tratamiento de otras enfermedades inflamatorias.

Los polipéptidos INSP005a, polipéptidos INSP005b y los polipéptidos maduros INSP005b se denominan en la presente memoria "los polipéptidos INSP005".

5 En un segundo aspecto, la invención proporciona una molécula de ácido nucleico purificada que codifica un polipéptido del primer aspecto de la invención. Preferiblemente, la molécula de ácido nucleico purificada tiene la secuencia de ácido nucleico indicada en SEQ ID N°:13 (que codifica el polipéptido INSP005a), SEQ ID N°:33 (que codifica el polipéptido INSP005b) o SEQ ID N°:35 (que codifica el polipéptido maduro INSP005b).

10 En un aspecto adicional, la invención proporciona un vector, tal como un vector de expresión, que contiene una molécula de ácido nucleico del segundo aspecto de la invención. En una realización preferida de este aspecto de la invención, el vector es el vector PCR-TOPO-IPAAA78836-1 (véase la Figura 9 y SEQ ID N°:38). En una realización aún más preferida de este aspecto de la invención, el vector es el vector PCR-TOPO-IPAAA78836-2 (véase la Figura 12 y SEQ ID N°:39).

15 En un aspecto adicional, la invención proporciona una célula hospedadora transformada con un vector de la invención.

20 En un aspecto adicional, la invención proporciona un anticuerpo que se une específicamente, y que preferiblemente inhibe la actividad de metaloproteasa de un polipéptido del primer aspecto de la invención.

Se presentan pruebas en la sección de Ejemplos más adelante de que el polipéptido INSP005b se puede usar para prevenir o tratar enfermedades inflamatorias, enfermedades autoinmunitarias, enfermedades hepáticas virales o agudas o insuficiencia hepática alcohólica.

25 En un aspecto adicional, la invención proporciona un polipéptido del primer aspecto de la invención, o una molécula de ácido nucleico del segundo aspecto de la invención, o un vector de la invención, o una célula hospedadora de la invención, o un anticuerpo de la invención, para el uso en la terapia o el diagnóstico de enfermedades, y en el que dicha enfermedad es una enfermedad autoinmunitaria, enfermedad hepática viral o aguda, que incluye insuficiencia hepática alcohólica, o una enfermedad inflamatoria. Estas moléculas se pueden usar también en la fabricación de un medicamento para el tratamiento de tales enfermedades.

30 Se prefiere en particular que los restos de la invención se usen en la fabricación de un medicamento para el tratamiento de enfermedades inflamatorias, enfermedades autoinmunitarias, enfermedad hepática (que incluye enfermedad hepática viral o aguda) e insuficiencia hepática (que incluye insuficiencia hepática alcohólica).

35 En un aspecto adicional, la invención proporciona un método para diagnosticar una enfermedad en un paciente, que comprende determinar el nivel de expresión de un gen natural que codifica un polipéptido del primer aspecto de la invención o la actividad de un polipéptido del primer aspecto de la invención en un tejido de dicho paciente, y comparar dicho nivel de expresión o actividad con un nivel de control, en el que un nivel que es diferente de dicho nivel de control es indicativo de la enfermedad. Tal método se llevará a cabo *in vitro*. Se pueden usar métodos similares para monitorizar el tratamiento terapéutico de la enfermedad en un paciente, en los que la alteración del nivel de expresión o actividad de un polipéptido o molécula de ácido nucleico a lo largo del periodo de tiempo respecto de un nivel de control es indicativa de la regresión de la enfermedad.

40 Las enfermedades diagnosticadas mediante este método de la invención son una enfermedad inflamatoria, enfermedad autoinmunitaria, enfermedad hepática (que incluye enfermedad hepática viral o aguda) o insuficiencia hepática (que incluye insuficiencia hepática alcohólica).

45 Un método preferido para detectar los polipéptidos del primer aspecto de la invención comprende las etapas de: (a) poner en contacto un anticuerpo de la invención con una muestra biológica en condiciones adecuadas para la formación de un complejo anticuerpo-polipéptido; y (b) detectar dicho complejo.

50 Existen varios métodos diferentes según este aspecto de la invención, como sabrá el lector experto, tales como los métodos de hibridación de ácidos nucleicos con sondas cortas, análisis de mutaciones puntuales, amplificación mediante reacción en cadena de la polimerasa (PCR) y métodos mediante el uso de anticuerpos para detectar niveles anormales de proteínas. Se pueden usar métodos similares a corto plazo o a largo plazo para posibilitar la monitorización del tratamiento terapéutico de una enfermedad en un paciente.

55 En un aspecto adicional, la invención proporciona el uso de un polipéptido del primer aspecto de la invención como una metaloproteasa.

60 En un aspecto adicional, la invención proporciona una composición farmacéutica que comprende un polipéptido del primer aspecto de la invención, o una molécula de ácido nucleico del segundo aspecto de la invención, o un vector de la invención, o una célula hospedadora de la invención, o un anticuerpo de la invención, junto con un vehículo farmacéuticamente aceptable.

65

En un aspecto adicional, la presente invención proporciona un polipéptido del primer aspecto de la invención, o una molécula de ácido nucleico del segundo aspecto de la invención, o un vector de la invención, o una célula hospedadora de la invención, o un anticuerpo de la invención, para el uso en la fabricación de un medicamento para el diagnóstico o el tratamiento de una enfermedad inflamatoria, una enfermedad autoinmunitaria, una enfermedad hepática o insuficiencia hepática.

Para las enfermedades en las que la expresión de un gen natural que codifica un polipéptido del primer aspecto de la invención, o en las que la actividad de un polipéptido del primer aspecto de la invención, es menor en un paciente enfermo en comparación con el nivel de expresión o actividad en un paciente sano, el polipéptido, la molécula de ácido nucleico, el vector, la célula hospedadora, o el anticuerpo administrado al paciente debería ser un agonista. A la inversa, para las enfermedades en las que la expresión del gen natural o la actividad del polipéptido es mayor en un paciente enfermo en comparación con el nivel de expresión o la actividad en un paciente sano, el polipéptido, la molécula de ácido nucleico, el vector, la célula hospedadora, o el anticuerpo administrado al paciente debería ser un antagonista. Los ejemplos de tales antagonistas incluyen las moléculas inversas de ácidos nucleicos, ribozimas y ligandos, tales como los anticuerpos.

En un aspecto adicional, la invención proporciona animales no humanos transgénicos o con genes inactivados que se han transformado para expresar niveles mayores, menores o inexistentes de un polipéptido del primer aspecto de la invención. Tales animales transgénicos son modelos muy útiles para el estudio de la enfermedad, y se pueden usar también en regímenes de cribado para la identificación de compuestos que son eficaces en el tratamiento o el diagnóstico de tal enfermedad.

La enfermedad es una enfermedad inflamatoria, una enfermedad autoinmunitaria, una enfermedad hepática o insuficiencia hepática.

Más adelante se proporciona un resumen de las técnicas y procedimientos habituales que se pueden emplear para utilizar la invención. Se entenderá que esta invención no se limita a la metodología, protocolos, líneas celulares, vectores y reactivos particulares descritos.

En esta memoria descriptiva se usan las abreviaturas habituales para los nucleótidos y los aminoácidos.

La práctica de la presente invención empleará, a menos que se indique de otra manera, técnicas convencionales de biología molecular, microbiología, tecnología de ADN recombinante e inmunología, que están dentro de la experiencia de las personas que trabajan en la técnica.

Tales técnicas se explican completamente en la bibliografía. Los ejemplos de textos especialmente adecuados para la consulta incluyen los siguientes: Sambrook Molecular Cloning; A Laboratory Manual, Segunda Edición (1989); DNA Cloning, Volúmenes I y II (D. N. Glover ed. 1985); Oligonucleotide Synthesis (M. J. Gait ed. 1984); Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins eds. 1984); Transcription and Translation (B. D. Hames & S. J. Higgins eds. 1984); Animal Cell Culture (R. I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); la serie Methods in Enzymology (Academic Press, Inc.), especialmente los volúmenes 154 y 155, Gene Transfer Vectors for Mammalian Cells (J. H. Miller y M. P. Calos eds. 1987, Cold Spring Harbor Laboratory); Immunochemical Methods in Cell and Molecular Biology (Mayer y Walker, eds. 1987, Academic Press, Londres); Scopes, (1987) Protein Purification: Principles and Practice, Segunda Edición (Springer Verlag, N.Y.); y Handbook of Experimental Immunology, Volúmenes I-IV (D. M. Weir y C. C. Blackwell eds. 1986).

Tal como se usa en la presente memoria, el término "polipéptido" incluye cualquier péptido o proteína que comprende dos o más aminoácidos unidos entre sí mediante enlaces peptídicos o enlaces peptídicos modificados, es decir, isómeros peptídicos. Este término se refiere tanto a las cadenas cortas (péptidos y oligopéptidos) como a las cadenas largas (proteínas).

Como se describió anteriormente, los polipéptidos de la presente invención pueden estar en forma de una proteína madura o pueden ser una pre-, pro- o prepro-proteína que se puede activar mediante la escisión de la porción pre-, pro- o prepro- para producir un polipéptido maduro activo. En tales polipéptidos, la pre-, pro- o prepro-secuencia puede ser una secuencia líder o secretora, o puede ser una secuencia que se emplea para la purificación de la secuencia polipeptídica madura.

El polipéptido del primer aspecto de la invención puede formar parte de una proteína de fusión. Por ejemplo, a menudo es ventajoso incluir una o más secuencias de aminoácidos adicionales que pueden contener secuencias secretoras o líder, pro-secuencias, secuencias que ayudan en la purificación, o secuencias que confieren una estabilidad superior a la proteína, por ejemplo durante la producción recombinante. De manera alternativa o adicional, el polipéptido maduro se puede fusionar a otro compuesto, tal como un compuesto para incrementar la semivida del polipéptido (por ejemplo, polietileno glicol).

Los polipéptidos pueden contener aminoácidos distintos de los 20 aminoácidos codificados genéticamente, modificados mediante procesos naturales, tales como mediante procesamiento postraduccion, o mediante técnicas de modificación química que se conocen bien en la técnica. Entre las modificaciones conocidas que pueden estar presentes habitualmente en los polipéptidos de la presente invención están la glicosilación, la unión de lípidos, sulfatación, gamma-carboxilación, por ejemplo de residuos de ácido glutámico, hidroxilación y ADP-ribosilación. Otras modificaciones potenciales incluyen la acetilación, acilación, amidación, unión covalente de flavina, unión covalente de un resto hemo, unión covalente de un nucleótido o derivado de nucleótido, unión covalente de un derivado de lípido, unión covalente de fosfatidilinositol, reticulación, ciclación, formación de puentes disulfuro, desmetilación, formación de entrecruzamientos covalentes, formación de cisteína, formación de piroglutamato, formilación, formación de anclajes de GPI, yodación, metilación, miristoilación, oxidación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenilación, adición de aminoácidos mediada por ARN de transferencia a proteínas, tal como arginilación, y ubiquitinación.

Las modificaciones se pueden dar en cualquier parte de un polipéptido, que incluye el esqueleto peptídico, las cadenas laterales de los aminoácidos y los extremos amino o carboxilo. De hecho, el bloqueo del extremo amino o carboxilo de un polipéptido, o ambos, mediante una modificación covalente es habitual en los polipéptidos naturales y sintéticos, y tales modificaciones pueden estar presentes en los polipéptidos de la presente invención.

Las modificaciones que se dan en un polipéptido serán a menudo una función de cómo se produce el polipéptido. Para los polipéptidos que se producen de manera recombinante, la naturaleza y el grado de las modificaciones estarán determinados en gran parte por la capacidad de modificación postraduccion de la célula hospedadora particular y las señales de modificación que estén presentes en la secuencia de aminoácidos del polipéptido en cuestión. Por ejemplo, los patrones de glicosilación varían entre los diferentes tipos de células hospedadoras.

Los polipéptidos de la presente invención se pueden preparar de cualquier manera adecuada. Tales polipéptidos incluyen, cuando el polipéptido es un polipéptido que se da de manera natural, polipéptidos que se dan de manera natural aislados (por ejemplo, purificados a partir de un cultivo celular) y también polipéptidos producidos de manera recombinante (que incluyen las proteínas de fusión), polipéptidos producidos de manera sintética o polipéptidos que se producen mediante una combinación de estos métodos. El término "aislado" no indica el método mediante el cual se obtiene el polipéptido, o el nivel de pureza de la preparación. Así, tales especies aisladas se pueden producir de manera recombinante, aislarlas directamente de la célula o tejido de interés, o producirlas de manera sintética basándose en las secuencias determinadas.

La "identidad" indica que en cualquier posición particular de las secuencias alineadas, el residuo de aminoácido es idéntico entre las secuencias. La "similitud" indica que, en cualquier posición particular de las secuencias alineadas, el residuo de aminoácido es de un tipo similar entre las secuencias. Los grados de identidad y de similitud se pueden calcular fácilmente (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, Nueva York, 1988; Biocomputing. Informatics and Genome Projects, Smith, D. W., ed., Academic Press, Nueva York, 1993; Computer Analysis of Sequence Data, Parte 1, Griffin, A. M., y Griffin, H. G., eds., Humana Press, Nueva Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. y Devereux, J., eds., M Stockton Press, Nueva York, 1991).

Los polipéptidos de la presente invención se pueden usar para generar anticuerpos policlonales o monoclonales, que son inmunoespecíficos para los polipéptidos. Tales anticuerpos se pueden emplear para aislar o identificar clones que expresan los polipéptidos de la invención o para purificar los polipéptidos mediante cromatografía de afinidad. Los anticuerpos también se pueden emplear como herramientas diagnósticas o terapéuticas, entre otras aplicaciones, como será evidente para el lector experto.

El término "inmunoespecífico" significa que los anticuerpos tienen una afinidad sustancialmente mayor por los polipéptidos de la invención que su afinidad hacia otros polipéptidos relacionados de la técnica anterior. Tal como se usa en la presente memoria, el término "anticuerpo" se refiere a moléculas intactas, así como a los fragmentos de las mismas, tales como Fab, F(ab')₂ y Fv, que son capaces de unirse al determinante antigénico en cuestión. Tales anticuerpos se unen así a los polipéptidos del primer aspecto de la invención.

"Afinidad sustancialmente mayor" significa que existe un incremento medible en la afinidad hacia un polipéptido de la invención en comparación con la afinidad hacia las proteínas secretadas conocidas.

Preferiblemente, la afinidad es al menos 1,5 veces, 2 veces, 5 veces, 10 veces, 100 veces, 10³ veces, 10⁴ veces, 10⁵ veces o 10⁶ veces mayor hacia un polipéptido de la invención que hacia las proteínas secretadas conocidas.

Si se desean anticuerpos policlonales, se puede inmunizar a un mamífero seleccionado, tal como un ratón, conejo, cabra o caballo, con un polipéptido del primer aspecto de la invención. El polipéptido usado para inmunizar al animal se puede obtener mediante técnicas de ADN recombinante, o se puede sintetizar químicamente. Si se desea, el polipéptido se puede conjugarse a una proteína portadora. Las proteínas portadoras usadas habitualmente a las que se pueden acoplar químicamente los polipéptidos incluyen albúmina de suero bovino, tiroglobulina y hemocianina de

lapa californiana. El polipéptido acoplado se usa después para inmunizar al animal. El suero del animal inmunizado se recoge y se trata según procedimientos conocidos, por ejemplo mediante cromatografía de inmunoafinidad.

5 Un experto en la técnica también puede producir fácilmente anticuerpos monoclonales hacia los polipéptidos del primer aspecto de la invención. Se conoce muy bien la metodología general para producir anticuerpos monoclonales mediante el uso de la técnica de hibridomas (véase, por ejemplo, Kohler, G. y Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, 77-96 en *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985).

10 Se pueden cribar paneles de anticuerpos monoclonales producidos hacia los polipéptidos del primer aspecto de la invención en función de diversas propiedades, es decir, en función del isotipo, epítipo, afinidad, etc. Los anticuerpos monoclonales son especialmente útiles en la purificación de los polipéptidos individuales hacia los que se dirigen. De manera alternativa, se pueden aislar los genes que codifican los anticuerpos monoclonales de interés a partir de hibridomas, por ejemplo mediante métodos de PCR conocidos en la técnica, y clonarlos y expresarlos en vectores adecuados.

15 También pueden ser útiles los anticuerpos quiméricos, en los que se unen o fusionan regiones variables no humanas a regiones constantes humanas (véase, por ejemplo, Liu *et al.*, *Proc. Natl. Acad. Sci. USA*, 84, 3439 (1987)).

20 El anticuerpo se puede modificar para hacerlo menos inmunógeno en un individuo, por ejemplo mediante humanización (véase Jones *et al.*, *Nature*, 321, 522 (1986); Verhoeyen *et al.*, *Science*, 239, 1534 (1988); Kabat *et al.*, *J. Immunol.*, 147, 1709 (1991); Queen *et al.*, *Proc. Natl Acad. Sci. USA*, 86, 10029 (1989); Gorman *et al.*, *Proc. Natl Acad. Sci. USA*, 88, 34181 (1991); y Hodgson *et al.*, *Bio/Technology*, 9, 421 (1991)). La expresión "anticuerpo humanizado", tal como se usa en la presente memoria, se refiere a moléculas de anticuerpo en las que los aminoácidos de la CDR y otros aminoácidos seleccionados de los dominios variables de las cadenas pesadas y/o ligeras de un anticuerpo donante no humano se han sustituido por los aminoácidos equivalentes de un anticuerpo humano. El anticuerpo humanizado, así, se parece mucho a un anticuerpo humano, pero tiene la capacidad de unión del anticuerpo donante.

30 En una alternativa adicional, el anticuerpo puede ser un anticuerpo "biespecífico", que es un anticuerpo que tiene dos dominios de unión al antígeno diferentes, y cada dominio se dirige hacia un epítipo diferente.

35 Se puede utilizar la técnica de expresión en fagos para seleccionar los genes que codifican anticuerpos con actividades de unión hacia los polipéptidos de la invención a partir de repertorios de genes V amplificados mediante PCR de linfocitos de seres humanos que se han cribado para que posean los anticuerpos relevantes, o de bibliotecas sin exposición previa (McCafferty, J. *et al.*, (1990), *Nature* 348, 552-554; Marks, J. *et al.*, (1992) *Biotechnology* 10, 779-783). La afinidad de estos anticuerpos se puede mejorar también mediante reordenamiento aleatorio de cadenas (Clackson, T. *et al.*, (1991) *Nature* 352, 624-628).

40 Los anticuerpos generados mediante las técnicas anteriores, ya sean policlonales o monoclonales, tienen una utilidad adicional, ya que se pueden emplear como reactivos en inmunoensayos, radioinmunoensayos (RIA) o ensayos inmunoabsorbentes ligados a enzimas (ELISA). En estas aplicaciones, los anticuerpos se pueden marcar con un reactivo detectable analíticamente tal como un radioisótopo, una molécula fluorescente o una enzima.

45 Las moléculas de ácidos nucleicos de la invención se pueden usar en los métodos y aplicaciones descritas en la presente memoria.

50 Las moléculas de ácidos nucleicos de la presente invención pueden estar en forma de ARN, tal como mARN, o en forma de ADN, que incluye, por ejemplo, cADN, ADN sintético o ADN genómico. Tales moléculas de ácidos nucleicos se pueden obtener mediante clonación, mediante técnicas de síntesis química o mediante una combinación de las mismas. Las moléculas de ácidos nucleicos se pueden preparar, por ejemplo, mediante síntesis química mediante el uso de técnicas tales como síntesis química de fosforamidita en fase sólida, a partir de bibliotecas genómicas o de cADN, o mediante separación a partir de un organismo. Se pueden generar moléculas de ARN en general mediante la transcripción *in vitro* o *in vivo* de secuencias de ADN.

55 Las moléculas de ácidos nucleicos pueden ser bicatenarias o monocatenarias. El ADN monocatenario puede ser la cadena codificante, también conocida como cadena directa, o puede ser la cadena no codificante, también conocida como cadena inversa.

60 La expresión "molécula de ácido nucleico" también incluye los análogos de ADN y ARN, tales como aquellos que contienen esqueletos modificados, y ácidos peptidonucleicos (APN). El término "APN", tal como se usa en la presente memoria, se refiere a una molécula inversa o un agente anti-gen que comprende un oligonucleótido de al menos cinco nucleótidos de longitud unido a un esqueleto peptídico de residuos de aminoácidos, que preferiblemente finaliza en una lisina. La lisina terminal confiere solubilidad a la composición. Los APNs pueden estar pegila-

dos para prolongar su duración en una célula, en la que se unen de manera preferente al ADN y ARN monocatenario complementario y detienen la elongación del transcrito (Nielsen, P. E. *et al.* (1993) *Anticancer Drug Des.* 8:53-63).

5 Una molécula de ácido nucleico que codifica el polipéptido de SEQ ID N°:14 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico mostrada en SEQ ID N°:13. Una molécula de ácido nucleico que codifica el polipéptido de SEQ ID N°:34 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico mostrada en SEQ ID N°:33. Una molécula de ácido nucleico que codifica el polipéptido de SEQ ID N°:36 puede ser idéntica a la secuencia codificante de la molécula de ácido nucleico mostrada en SEQ ID N°:35.

10 Estas moléculas también pueden tener una secuencia diferente que, como resultado de la degeneración del código genético, codifica un polipéptido de SEQ ID N°:14, SEQ ID N°:34 o SEQ ID N°:36. Tales moléculas de ácidos nucleicos pueden incluir, pero sin limitación, la secuencia codificante del polipéptido maduro propiamente dicho; la secuencia codificante del polipéptido maduro y secuencias codificantes adicionales, tales como las que codifican una secuencia líder o secretora, tal como una secuencia de pro-, pre- o prepro-polipéptido; la secuencia codificante del polipéptido maduro, con o sin las secuencias codificantes adicionales anteriormente mencionadas, junto con otras secuencias no codificantes adicionales, que incluyen las secuencias de 5' y 3' no codificantes, tales como las secuencias transcritas no traducidas, que desempeñan un papel en la transcripción (que incluyen las señales de terminación), la unión al ribosoma y la estabilidad del mRNA. Las moléculas de ácidos nucleicos también pueden incluir secuencias adicionales que codifican aminoácidos adicionales, tales como aquellos que proporcionan funciones adicionales.

25 Las moléculas de ácidos nucleicos de la invención también se pueden modificar, mediante el uso de métodos conocidos en general en la técnica, por una diversidad de razones, que incluyen la modificación de la clonación, el procesamiento, y/o la expresión del producto génico (el polipéptido). El reordenamiento aleatorio del ADN mediante fragmentación aleatoria y reensamblaje por PCR de fragmentos génicos y oligonucleótidos sintéticos está incluido en las técnicas que se pueden usar para modificar las secuencias de nucleótidos. La mutagénesis dirigida se puede usar para insertar nuevos sitios de restricción, alterar los patrones de glicosilación, cambiar la preferencia de los codones, producir variantes de corte y empalme, introducir mutaciones, etc.

30 Las moléculas de ácidos nucleicos que codifican un polipéptido del primer aspecto de la invención se pueden ligar a una secuencia heteróloga, de forma que la molécula de ácido nucleico combinada codifique una proteína de fusión. Tales moléculas de ácidos nucleicos combinadas están incluidas en la invención. Por ejemplo, para cribar bibliotecas de péptidos en busca de inhibidores de la actividad del polipéptido, puede ser útil expresar, mediante el uso de tal molécula de ácido nucleico combinada, una proteína de fusión que puede ser reconocida por un anticuerpo disponible comercialmente. Una proteína de fusión también se puede modificar para que contenga un sitio de escisión localizado entre la secuencia del polipéptido de la invención y la secuencia de una proteína heteróloga, de manera que el polipéptido se puede escindir y purificar a partir de la proteína heteróloga.

40 Tal como se discute adicionalmente más adelante con respecto a los ensayos que se pueden utilizar según la invención, una molécula de ácido nucleico como se describió anteriormente se puede usar en forma de una sonda de hibridación para ARN, cADN o ADN genómico, para aislar cADNs de longitud completa y clones genómicos que codifican los polipéptidos INSP005 y para aislar cADN y clones genómicos de genes homólogos u ortólogos que tienen una elevada similitud de secuencias hacia el gen que codifica estos polipéptidos.

45 A este respecto, se pueden utilizar los siguientes métodos, entre otros conocidos en la técnica, y se discuten más adelante con fines ilustrativos. Los métodos para la secuenciación y el análisis del ADN se conocen bien y están disponibles en general en la técnica, y se pueden usar, de hecho, para poner en práctica muchas de las realizaciones de la invención discutidas en la presente memoria. Tales métodos pueden emplear enzimas tales como el fragmento Klenow de la ADN polimerasa I, Sequenasa (US Biochemical Corp, Cleveland, OH), Taq polimerasa (Perkin Elmer), polimerasa T7 termoestable (Amersham, Chicago, IL), o combinaciones de polimerasas y exonucleasas con corrección de errores tales como las halladas en el sistema de amplificación ELONGASE comercializado por Gibco/BRL (Gaithersburg, MD). Preferiblemente, el procedimiento de secuenciación se puede automatizar mediante el uso de aparatos tales como el Hamilton Micro Lab 2200 (Hamilton, Reno, NV), el Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) y el ABI Catalyst y los secuenciadores de ADN 373 y 377 (Perkin Elmer).

55 Un método para aislar una molécula de ácido nucleico que codifica un polipéptido con una función equivalente a la de los polipéptidos INSP005 es analizar una biblioteca genómica o de cADN con una sonda natural o diseñada artificialmente mediante el uso de procedimientos estándar reconocidos en la técnica (véase, por ejemplo, "Current Protocols in Molecular Biology", Ausubel *et al.* (eds). Greene Publishing Association and John Wiley Interscience, Nueva York, 1989, 1992). Las sondas que comprenden al menos 15, preferiblemente al menos 30, y más preferiblemente al menos 50, bases contiguas que corresponden, o que son complementarias, a las secuencias de ácidos nucleicos del gen codificante adecuado (SEQ ID N°:13, SEQ ID N°:33 o SEQ ID N°:35), son sondas especialmente útiles. Tales sondas se pueden marcar con un reactivo detectable analíticamente para facilitar su identificación. Los reactivos útiles incluyen, pero sin limitación, radioisótopos, tintes fluorescentes y enzimas que son capaces de catalizar la formación de un producto detectable. Mediante el uso de estas sondas, el técnico de experiencia habitual será

capaz de aislar copias complementarias de polinucleótidos de ADN genómico, cADN o ARN que codifican proteínas de interés de seres humanos, mamíferos u otras fuentes animales, y de cribar tales fuentes en busca de secuencias relacionadas, por ejemplo, en busca de miembros adicionales de la familia, tipo y/o subtipo.

5 En muchos casos, las secuencias de cADN aisladas serán incompletas, ya que la región que codifica el polipéptido estará recortada, normalmente en el extremo 5'. Existen varios métodos para obtener cADNs de tamaño completo, o para prolongar los cADNs cortos. Tales secuencias se pueden prolongar mediante la utilización de una secuencia de nucleótidos parcial, y empleando diversos métodos conocidos en la técnica para detectar secuencias en posición 5' tales como promotores y elementos reguladores. Por ejemplo, un método que se puede emplear se basa en el método de amplificación rápida de los extremos del cADN (RACE; véase, por ejemplo, Frohman *et al.*, PNAS USA 85, 8998-9002, 1988). Las modificaciones recientes de esta técnica, ejemplificadas por la técnica Marathon™ (Clontech Laboratories Inc.), por ejemplo, han simplificado significativamente la búsqueda de cADNs más largos. Una técnica ligeramente diferente, denominada PCR de "sitio de restricción", usa cebadores universales para recuperar una secuencia de ácido nucleico desconocida adyacente a un locus conocido (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). También se puede usar la PCR inversa para amplificar o prolongar secuencias mediante el uso de cebadores divergentes basados en una región conocida (Triglia, T. *et al.* (1988) Nucleic Acids Res. 16:8186). Otro método que se puede usar es la PCR de captura, que implica la amplificación mediante PCR de fragmentos de ADN adyacentes a una secuencia conocida en ADN cromosómico artificial de levadura y humano (Lagerstrom, M. *et al.* (1991) PCR Methods Applic., 1, 111-119). Otro método que se puede usar para recuperar secuencias conocidas es el de Parker, J. D. *et al.* (1991); Nucleic Acids Res. 19:3055-3060). Además, se puede usar PCR, cebadores anidados, y bibliotecas PromoterFinder™ para recorrer el ADN genómico (Clontech, Palo Alto, CA). Este procedimiento evita la necesidad de cribar bibliotecas, y es útil para hallar uniones intrón/exón.

25 Cuando se criban cADNs de tamaño completo, es preferible usar bibliotecas que se han seleccionado por tamaño para que incluyan cADNs mayores. Además, son preferibles las bibliotecas cebadas aleatoriamente, ya que contendrán más secuencias que contienen las regiones 5' de los genes. El uso de una biblioteca cebada aleatoriamente puede ser especialmente preferible para situaciones en las que una biblioteca de oligo d(T) no produce un cADN de tamaño completo. Las bibliotecas genómicas pueden ser útiles para la prolongación de una secuencia hasta las regiones reguladoras no transcritas de 5'.

30 En una realización de la invención, las moléculas de ácidos nucleicos de la presente invención se pueden usar para la localización cromosómica. En esta técnica, una molécula de ácido nucleico se dirige específicamente hacia, y puede hibridar con, una localización particular en un cromosoma humano individual. La cartografía de secuencias relevantes en cromosomas según la presente invención es una etapa importante en la correlación confirmatoria de esas secuencias con la enfermedad asociada al gen. Una vez que una secuencia se ha cartografiado en una localización cromosómica precisa, la posición física de la secuencia en el cromosoma se puede correlacionar con los datos de cartografía genética. Tales datos se hallan, por ejemplo, en V. McKusick, Mendelian Inheritance in Man (disponible en línea a través de la Biblioteca Médica Welch de la Universidad Johns Hopkins). Las relaciones entre los genes y las enfermedades que se han cartografiado en la misma región cromosómica se identifican después por medio del análisis de ligamiento (herencia conjunta de genes físicamente adyacentes). Esto proporciona información valiosa a los investigadores que buscan genes de enfermedades mediante el uso de clonación posicional u otras técnicas de descubrimiento de genes. Una vez que la enfermedad o el síndrome se ha localizado de manera aproximada mediante ligamiento genético en una región genómica particular, cualquier secuencia que se cartografie en ese área puede representar genes asociados o regulatorios para la investigación posterior. La molécula de ácido nucleico se puede usar también para detectar diferencias en la localización cromosómica debida a translocación, inversión, etc. entre individuos normales, portadores, o afectados.

50 Las moléculas de ácidos nucleicos de la presente invención también son valiosas para la localización de tejidos. Tales técnicas permiten la determinación de patrones de expresión del polipéptido en tejidos mediante la detección de los mARNs que los codifican. Estas técnicas incluyen las técnicas de hibridación *in situ* y las técnicas de amplificación de nucleótidos, tales como PCR. Los resultados de estos estudios proporcionan una indicación de las funciones normales del polipéptido en el organismo. Además, los estudios comparativos del patrón de expresión normal de los mARNs con los de los mARNs codificados por un gen mutante proporcionan percepciones valiosas del papel de los polipéptidos mutantes en la enfermedad. Tal expresión inadecuada puede ser de naturaleza temporal, espacial o cuantitativa.

60 También se pueden utilizar aproximaciones de silenciamiento génico para inhibir la expresión endógena de un gen que codifica un polipéptido de la invención. El ARN de interferencia (ARNi) (Elbashir, S M *et al.*, Nature 2001, 411, 494-498) es un método de silenciamiento génico post-transcripcional específico de secuencia que se puede emplear. Se sintetizan oligonucleótidos de dsARN *in vitro* y se introducen en una célula. La unión específica de secuencia de estos oligonucleótidos de dsARN desencadena la degradación del mARN seleccionado como objetivo, reduciendo o eliminando la expresión de la proteína seleccionada como objetivo.

La eficacia de las aproximaciones de silenciamiento génico indicadas anteriormente se puede determinar por medio de la medida de la expresión del polipéptido (por ejemplo, mediante transferencia de Western), y a nivel del ARN mediante el uso de metodologías basadas en TaqMan.

5 Los vectores de la presente invención comprenden las moléculas de ácidos nucleicos de la invención, y pueden ser vectores de clonación o de expresión. Las células hospedadoras de la invención, que se pueden transformar, transfectar o transducir con los vectores de la invención, pueden ser procarióticas o eucarióticas.

10 Los polipéptidos de la invención se pueden preparar de forma recombinante mediante la expresión de sus moléculas de ácidos nucleicos codificantes en vectores contenidos dentro de una célula hospedadora. Los expertos en la técnica conocen bien tales métodos de expresión, y muchos se describen con detalle en Sambrook *et al.* (anteriormente mencionado) y Fernandez y Hoefler (1998, eds. "Gene expression systems. Using nature for the art of expression". Academic Press, San Diego, Londres, Boston, Nueva York, Sidney, Tokio, Toronto).

15 En general, se puede usar cualquier sistema o vector que sea adecuado para mantener, propagar o expresar moléculas de ácidos nucleicos para producir un polipéptido en el hospedador necesario. La secuencia de nucleótidos adecuada se puede insertar en un sistema de expresión mediante cualquiera de una diversidad de técnicas rutinarias y muy conocidas, tales como, por ejemplo, las descritas en Sambrook *et al.*, (anteriormente mencionado). En general, el gen codificante se puede colocar bajo el control de un elemento de control tal como un promotor, sitio de unión al ribosoma (para la expresión bacteriana), y, opcionalmente, un operador, de manera que la secuencia de ADN que codifica el polipéptido deseado se transcribe hasta ARN en la célula hospedadora transformada.

25 Los ejemplos de sistemas de expresión adecuados incluyen, por ejemplo, sistemas cromosómicos, episómicos y derivados de virus, que incluyen, por ejemplo, vectores derivados de: plásmidos bacterianos, bacteriófagos, transposones, episomas de levaduras, elementos de inserción, elementos cromosómicos de levaduras, virus tales como baculovirus, papovavirus tales como SV40, virus vaccinia, adenovirus, virus de viruela aviar, virus de pseudorrabia y retrovirus, o combinaciones de los mismos, tales como las descritas de elementos genéticos de plásmidos y bacteriófagos, que incluyen los cósmidos y fagémidos. También se pueden emplear cromosomas artificiales humanos (HACs) para administrar fragmentos más largos de ADN que se pueden contener y expresar en un plásmido.

30 Los sistemas de expresión especialmente adecuados incluyen microorganismos tales como bacterias transformadas con vectores de expresión de ADN de bacteriófagos recombinantes, plásmidos o cósmidos; levaduras transformadas con vectores de expresión de levaduras; sistemas de células de insectos infectadas con vectores de expresión de virus (por ejemplo, baculovirus); sistemas de células vegetales transformadas con vectores de expresión de virus (por ejemplo, virus del mosaico de la coliflor, CaMV; virus del mosaico del tabaco, TMV) o con vectores de expresión bacterianos (por ejemplo, plásmidos Ti o pBR322); o sistemas de células animales. También se pueden emplear sistemas de traducción sin células para producir los polipéptidos de la invención.

40 La introducción de moléculas de ácidos nucleicos que codifican un polipéptido de la presente invención en células hospedadoras se puede llevar a cabo mediante métodos descritos en muchos manuales de laboratorio habituales, tales como Davis *et al.*, Basic Methods in Molecular Biology (1986) y Sambrook *et al.*, (anteriormente mencionado). Los métodos especialmente adecuados incluyen la transcripción mediante fosfato cálcico, transfección mediada por DEAE-dextrano, transvección, microinyección, transfección mediada por lípidos catiónicos, electroporación, transducción, transferencia por raspado, introducción balística o infección (véase Sambrook *et al.*, 1989 [anteriormente mencionado]; Ausubel *et al.*, 1991 [anteriormente mencionado]; Spector, Goldman y Leinwald, 1998). En las células eucarióticas, los sistemas de expresión pueden ser transitorios (por ejemplo, episómicos) o permanentes (integración cromosómica) según las necesidades del sistema.

50 La molécula de ácido nucleico codificante puede incluir o no una secuencia que codifica una secuencia de control, tal como un péptido señal o secuencia líder, según se desee, por ejemplo, para la secreción del polipéptido traducido al interior del retículo endoplásmico, al espacio periplásmico o al medio extracelular. Estas señales pueden ser endógenas para el polipéptido, o pueden ser señales heterólogas. Las secuencias líder pueden ser eliminadas por el hospedador bacteriano en el procesamiento postraduccional.

55 Además de las secuencias de control, puede ser deseable añadir secuencias reguladoras que posibiliten la regulación de la expresión del polipéptido respecto del crecimiento de la célula hospedadora. Los ejemplos de secuencias reguladoras son aquellas que provocan que la expresión de un gen aumente o disminuya en respuesta a un estímulo químico o físico, que incluye la presencia de un compuesto regulador o diversas condiciones de temperatura o condiciones metabólicas. Las secuencias reguladoras son las regiones no traducidas del vector, tales como potenciadores, promotores y regiones sin traducir de 5' y 3'. Estas interactúan con las proteínas celulares del hospedador para llevar a cabo la transcripción y la traducción. Tales secuencias reguladoras pueden variar de intensidad y especificidad. Dependiendo del sistema del vector y del hospedador utilizado, se puede usar cualquier número de elementos de transcripción y traducción adecuados, que incluyen promotores constitutivos e inducibles. Por ejemplo, cuando se realiza la clonación en sistemas bacterianos, se pueden usar promotores inducibles tales como el promotor de lacZ híbrido del fagémido Bluescript (Stratagene, La Jolla, CA) o el plásmido pSport1™ (Gibco BRL), y

similares. Se puede usar el promotor de polihedrina de baculovirus en células de insecto. Los promotores o potenciadores derivados de los genomas de células vegetales (por ejemplo, choque térmico, RUBISCO y genes de proteínas de almacenamiento) o de virus vegetales (por ejemplo, promotores o secuencias líder virales) se pueden clonar en el vector. En los sistemas de células mamíferas, son preferibles los promotores de genes mamíferos o de virus de mamíferos. Si es necesario generar una línea celular que contenga múltiples copias de la secuencia, se pueden usar vectores basados en SV40 o EBV con un marcador seleccionable adecuado.

Se construye un vector de expresión de forma que la secuencia codificante de ácido nucleico particular esté localizada en el vector con las secuencias reguladoras adecuadas, y la posición y la orientación de la secuencia codificante con respecto a las secuencias reguladoras es tal que la secuencia codificante se transcribe bajo el "control" de las secuencias reguladoras, es decir, la ARN polimerasa que se une a la molécula de ADN en las secuencias de control transcribe la secuencia codificante. En algunos casos, puede ser necesario modificar la secuencia de manera que se pueda unir a las secuencias de control con la orientación adecuada; es decir, para mantener el marco de lectura.

Las secuencias de control y otras secuencias reguladoras se pueden ligar a la secuencia codificante de ácido nucleico antes de la inserción en un vector. De manera alternativa, la secuencia codificante se puede clonar directamente en un vector de expresión que ya contenga las secuencias de control y un sitio de restricción adecuado.

Para la producción a largo plazo y de alto rendimiento de un polipéptido recombinante, se prefiere la expresión estable. Por ejemplo, se pueden transformar líneas celulares que expresan de manera estable el polipéptido de interés mediante el uso de vectores de expresión que pueden contener orígenes de replicación virales y/o elementos de expresión endógenos y un gen marcador seleccionable en el mismo vector o en un vector distinto. Tras la introducción del vector, se puede dejar que las células crezcan durante 1-2 días en un medio enriquecido antes de cambiarlas al medio selectivo. El propósito del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el cultivo y la recuperación de las células que expresan de manera eficaz las secuencias introducidas. Los clones resistentes de células transformadas de manera estable se pueden hacer proliferar mediante el uso de técnicas de cultivo de tejidos adecuadas para ese tipo de células.

Las líneas celulares mamíferas disponibles como hospedadores para la expresión se conocen en la técnica, e incluyen muchas líneas celulares inmortalizadas disponibles de la American Type Culture Collection (ATCC) que incluyen, pero sin limitación, células de ovario de hámster chino (CHO), HeLa, riñón de hámster recién nacido (BHK), riñón de mono (COS), C127, 3T3, BHK, HEK 293, melanoma de Bowes y carcinoma hepatocelular humano (por ejemplo Hep G2), y otras líneas celulares.

En el sistema de baculovirus, los materiales para los sistemas de expresión baculovirus/células de insecto están disponibles comercialmente en forma de equipos, entre otros, de Invitrogen, San Diego CA (el equipo "MaxBac"). Los expertos en la técnica conocen en general estos métodos, y se describen completamente en Summers y Smith, Texas Agricultural Experiment Station Bulletin N° 1555 (1987). Las células hospedadoras especialmente adecuadas para el uso en este sistema incluyen las células de insecto tales como las células de S2 *Drosophila* y Sf9 de *Spodoptera*.

Existen muchos sistemas de expresión genética mediante cultivos de células vegetales y de plantas completas en la técnica. Los ejemplos de sistemas de expresión genética celular vegetal adecuados incluyen los descritos en las pat. de EE.UU. n° 5.693.506; 5.659.122; y 5.608.143. Se han descrito ejemplos adicionales de expresión genética en cultivos de células vegetales en Zenk, *Phytochemistry* 30, 3861-3863 (1991).

En particular, se pueden usar todas las plantas de las cuales se puedan aislar protoplastos y cultivarlos para proporcionar plantas regeneradas completas, de manera que se recuperan las plantas completas que contienen el gen transferido. Prácticamente todas las plantas se pueden regenerar a partir de las células o tejidos cultivados, que incluyen, pero sin limitación, todas las especies principales de caña de azúcar, remolacha azucarera, algodón, árboles frutales y otros árboles, legumbres y hortalizas.

Los ejemplos especialmente preferidos de células hospedadoras bacterianas incluyen células de *streptococci*, *staphylococci*, *E. coli*, *Streptomyces* y *Bacillus subtilis*.

Los ejemplos de células hospedadoras especialmente adecuadas para la expresión fúngica incluyen las células de levaduras (por ejemplo, *S. cerevisiae*) y *Aspergillus*.

Se conoce en la técnica un gran número de sistemas de selección que se pueden usar para recuperar las líneas celulares transformadas. Los ejemplos incluyen los genes de la timidina quinasa del virus herpes simple (Wigler, M. *et al.* (1977) *Cell* 11:223-32) y adenina fosforribosiltransferasa (Lowy, I. *et al.* (1980) *Cell* 22:817-23) que se pueden emplear en células tk- o aprt±, respectivamente.

Además, se puede usar la resistencia a antimetabolitos, antibióticos o herbicidas como base para la selección; por ejemplo, la dihidrofolato reductasa (DHFR) que confiere resistencia a metotrexato (Wigler, M. *et al.* (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, que confiere resistencia a los aminoglicósidos neomicina y G-418 (Colbere-Garapin, F. *et al.* (1981) J. Mol. Biol. 150:1-14) y als o pat, que confieren resistencia a clorsulfuron y fosfotricina acetiltransferasa, respectivamente. Se han descrito genes seleccionables adicionales, cuyos ejemplos serán evidentes para los expertos en la técnica.

Aunque la presencia o ausencia de expresión del gen marcador indica que el gen de interés también está presente, puede ser necesario confirmar su presencia y expresión. Por ejemplo, si la secuencia relevante se inserta en una secuencia de un gen marcador, las células transformadas que contienen las secuencias adecuadas se pueden identificar por la ausencia de la función del gen marcador. De manera alternativa, se puede colocar un gen marcador en tándem con una secuencia que codifica un polipéptido de la invención bajo el control de un único promotor. La expresión del gen marcador en respuesta a la inducción o selección también indica normalmente la expresión del gen en tándem.

De manera alternativa, las células hospedadoras que contienen una secuencia de ácido nucleico que codifica un polipéptido de la invención y que expresan dicho polipéptido se pueden identificar mediante una diversidad de procedimientos conocidos para los expertos en la técnica. Estos procedimientos incluyen, pero sin limitación, hibridaciones ADN-ADN o ADN-ARN y bioensayos de proteínas, por ejemplo, separación de células activada por fluorescencia (FACS) o técnicas de inmunoensayos (tales como el ensayo inmunoabsorbente ligado a enzimas [ELISA] y el radioinmunoensayo [RIA]), que incluyen las técnicas basadas en membranas, disoluciones, o chips para la detección y/o cuantificación del ácido nucleico o la proteína (véase Hampton, R. *et al.* (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) y Maddox, D. E. *et al.* (1983) J. Exp. Med, 158, 1211-1216).

Los expertos en la técnica conocen una amplia diversidad de marcadores y técnicas de conjugación, y se pueden usar en diversos ensayos de ácidos nucleicos y aminoácidos. Los medios para producir sondas de hibridación o de PCR marcadas para detectar las secuencias relacionadas con las moléculas de ácidos nucleicos que codifican los polipéptidos de la presente invención incluyen el marcaje con oligonucleótidos, el traslado de mella, el marcaje en los extremos o la amplificación mediante PCR usando un polinucleótido marcado. De manera alternativa, las secuencias que codifican el polipéptido de la invención se pueden clonar en un vector para la producción de una sonda de mRNA. Tales vectores se conocen en la técnica, están disponibles comercialmente, y se pueden usar para sintetizar sondas de ARN *in vitro* mediante la adición de una ARN polimerasa adecuada tal como T7, T3 o SP6 y nucleótidos marcados. Estos procedimientos se pueden llevar a cabo mediante el uso de una diversidad de equipos disponibles comercialmente (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); y U.S. Biochemical Corp., Cleveland, OH)).

Las moléculas indicadoras o marcadores adecuados, que se pueden usar para facilitar la detección, incluyen los radionúclidos, las enzimas y los agentes fluorescentes, quimioluminiscentes o cromogénicos, así como sustratos, cofactores, inhibidores, partículas magnéticas, y similares.

Las moléculas de ácidos nucleicos según la presente invención se pueden usar también para crear animales transgénicos, en particular roedores. Tales animales transgénicos forman un aspecto adicional de la presente invención. Esto se puede realizar localmente mediante la modificación de células somáticas, o mediante la terapia de la línea germinal para incorporar modificaciones heredables. Tales animales transgénicos pueden ser especialmente útiles en la generación de modelos de animales para moléculas farmacológicas eficaces como moduladores de los polipéptidos de la presente invención.

El polipéptido se puede recuperar y purificar a partir de cultivos de células recombinantes mediante métodos muy conocidos, que incluyen la precipitación con sulfato amónico o etanol, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía en fosfocelulosa, cromatografía de interacción hidrofóbica, cromatografía de afinidad, cromatografía de hidroxapatito y cromatografía de lectina. La cromatografía líquida de alto rendimiento es especialmente útil para la purificación. Se pueden emplear técnicas muy conocidas para replegar las proteínas para regenerar una conformación activa cuando el polipéptido se desnaturaliza durante el aislamiento y/o la purificación.

También se pueden usar construcciones de vectores especializados para facilitar la purificación de las proteínas, según se desee, uniendo secuencias que codifican los polipéptidos de la invención a una secuencia de nucleótidos que codifica un dominio polipeptídico que facilitará la purificación de las proteínas solubles. Los ejemplos de tales dominios que facilitan la purificación incluyen los péptidos quelantes de metales tales como los módulos de histidina-triptófano que permiten la purificación con metales inmovilizados, los dominios de proteína A que permiten la purificación con inmunoglobulina inmovilizada, y el dominio utilizado en el sistema de extensión/purificación por afinidad FLAGS (Immunex Corp., Seattle, WA). La inclusión de secuencias ligadoras escindibles tales como las específicas del Factor XA o de la enteroquinasa (Invitrogen, San Diego, CA) entre el dominio de purificación y el polipéptido de la invención se pueden usar para facilitar la purificación. Tal vector de expresión posibilita la expresión de una proteína de fusión que contiene el polipéptido de la invención fusionado a varios residuos de histidina que preceden a un sitio

de escisión de tiorredoxina o enteroquinasa. Los residuos de histidina facilitan la purificación mediante IMAC (cromatografía de afinidad con iones metálicos inmovilizados descrita en Porath, J. *et al.* (1992), *Prot. Exp. Purif.* 3: 263-281) mientras el sitio de escisión de tiorredoxina o enteroquinasa proporciona un medio para la purificación del polipéptido a partir de la proteína de fusión. Se proporciona una discusión de vectores que contienen proteínas de fusión en Kroll, D. J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

Si el polipéptido se va a expresar para el uso en ensayos de cribado, generalmente se prefiere que se produzca en la superficie de la célula hospedadora en la que se expresa. En este caso, las células hospedadoras se pueden recoger antes del uso en el ensayo de cribado, por ejemplo mediante el uso de técnicas tales como separación de células activada por fluorescencia (FACS) o técnicas de inmunoafinidad. Si el polipéptido se secreta en el medio, se puede recuperar el medio para recuperar y purificar el polipéptido expresado. Si el polipéptido se produce de manera intracelular, primero se deben lisar las células antes de recuperar el polipéptido.

El polipéptido de la invención se puede usar para cribar bibliotecas de compuestos en cualquiera de una diversidad de técnicas de cribado de fármacos. Tales compuestos pueden activar (actuar como agonistas) o inhibir (actuar como antagonistas) el nivel de expresión del gen o la actividad del polipéptido de la invención, y forman un aspecto adicional de la presente invención. Los compuestos preferidos son eficaces para alterar la expresión de un gen natural que codifica un polipéptido del primer aspecto de la invención o para regular la actividad de un polipéptido del primer aspecto de la invención.

Los compuestos agonistas o antagonistas se pueden aislar, por ejemplo, a partir de células, preparaciones exentas de células, bibliotecas químicas o mezclas de productos naturales. Estos agonistas o antagonistas pueden ser sustratos naturales o modificados, ligandos, enzimas, receptores o moléculas miméticas estructurales o funcionales. Para una revisión adecuada de tales técnicas de cribado, véase Coligan *et al.*, *Current Protocols in Immunology* 1(2): Capítulo 5 (1991).

Los compuestos que es más probable que sean buenos antagonistas son las moléculas que se unen al polipéptido de la invención sin inducir el efecto biológico del polipéptido tras unirse a él. Los antagonistas potenciales incluyen moléculas orgánicas pequeñas, péptidos, polipéptidos y anticuerpos que se unen al polipéptido de la invención y por tanto inhiben o eliminan su actividad. De esta manera, se puede inhibir la unión del polipéptido a moléculas de unión celulares normales, de manera que se evita la actividad biológica normal del polipéptido.

El polipéptido de la invención que se emplea en tal técnica de cribado puede estar libre en disolución, fijado a un soporte sólido, asociado a una superficie celular o localizado de manera intracelular. En general, tales procedimientos de cribado pueden implicar el uso de células o membranas celulares adecuadas que expresan el polipéptido que están en contacto con un compuesto de ensayo para observar la unión, o la estimulación o la inhibición de una respuesta funcional. La respuesta funcional de las células puestas en contacto con el compuesto de ensayo se compara después con las células de control que no se pusieron en contacto con el compuesto de ensayo. Tal ensayo puede determinar si el compuesto de ensayo da como resultado una señal generada por la activación del polipéptido, mediante el uso de un sistema de detección adecuado. Los inhibidores de la activación se ensayan en general en presencia de un agonista conocido, y se observa el efecto sobre la activación por el agonista en presencia del compuesto de ensayo.

Los polipéptidos INSP005 de la presente invención pueden modular una diversidad de procesos fisiológicos y patológicos, que incluyen procesos reproductivos tales como la maduración o la fertilización del huevo. Así, la actividad biológica de los polipéptidos INSP005 se puede examinar en sistemas que permiten el estudio de tales actividades moduladoras, mediante el uso de una diversidad de ensayos adecuados. Por ejemplo, los ensayos posibles incluyen la medida de la fertilización del ovocito y/o las tasas de gestación tras la inducción de la ovulación, la medida de las tasas de implantación del embrión, o en el caso de la infertilidad masculina, la medida de la motilidad espermática (Luo C. W. *et al.*, *J. Biol. Chem.* 276 (10), 6913-6921 (2001)).

Se describe un método para identificar un compuesto agonista o antagonista de un polipéptido de la presente invención que comprende:

- (a) poner en contacto una célula que expresa en su superficie el polipéptido según el primer aspecto de la invención, y el polipéptido está asociado a un segundo componente capaz de proporcionar una señal detectable en respuesta a la unión de un compuesto al polipéptido, con un compuesto a cribar en condiciones que permiten la unión al polipéptido; y
- (b) determinar si el compuesto se une y activa o inhibe el polipéptido midiendo el nivel de una señal generada a partir de la interacción del compuesto con el polipéptido.

Un método adicional para identificar un agonista o antagonista de un polipéptido de la invención puede comprender:

(a) poner en contacto una célula que expresa en su superficie el polipéptido, y el polipéptido está asociado a un segundo componente capaz de proporcionar una señal detectable en respuesta a la unión de un compuesto al polipéptido, con un compuesto a cribar en condiciones que permiten la unión al polipéptido; y

5 (b) determinar si el compuesto se une y activa o inhibe el polipéptido comparando el nivel de una señal generada a partir de la interacción del compuesto con el polipéptido con el nivel de una señal en ausencia del compuesto.

10 Los métodos generales que se describieron anteriormente pueden comprender además llevar a cabo la identificación del agonista o antagonista en presencia de un ligando marcado o sin marcar para el polipéptido.

Se describe un método para identificar un agonista o antagonista de un polipéptido de la presente invención que comprende:

15 determinar la inhibición de la unión de un ligando a células que tienen un polipéptido de la invención en su superficie, o a membranas celulares que contienen tal polipéptido, en presencia de un compuesto candidato en condiciones que permiten la unión al polipéptido, y determinar la cantidad de ligando unido al polipéptido. Un compuesto capaz de provocar la reducción de la unión de un ligando se considera que es un agonista o antagonista. Preferiblemente, el ligando está marcado.

20 Más en particular, un método de cribado en busca de un compuesto antagonista o agonista de un polipéptido puede comprender las etapas de:

25 (a) incubar un ligando marcado con una célula completa que expresa un polipéptido según la invención en la superficie celular, o una membrana celular que contiene un polipéptido de la invención,

(b) medir la cantidad de ligando marcado unido a la célula completa o a la membrana celular;

30 (c) añadir un compuesto candidato a una mezcla de ligando marcado y la célula completa o membrana celular de la etapa (a) y permitir que la mezcla alcance el equilibrio;

(d) medir la cantidad de ligando marcado unido a la célula completa o a la membrana celular tras la etapa (c); y

35 (e) comparar la diferencia del ligando marcado unido en la etapa (b) y (d), de manera que el compuesto que provoca la reducción de la unión en la etapa (d) se considera que es un agonista o antagonista.

Se puede descubrir que los polipéptidos modulan una diversidad de procesos fisiológicos y patológicos de una manera dependiente de la dosis en los ensayos anteriormente descritos.

40 En algunas de las realizaciones descritas anteriormente, se pueden usar ensayos de unión simple, en los que se detecta la adherencia de un compuesto de ensayo a una superficie que alberga el polipéptido por medio de un marcador asociado directamente o indirectamente al compuesto de ensayo o en un ensayo que implica la competición con un competidor marcado. En otra realización, se pueden usar ensayos de cribado de fármacos competitivos en los que los anticuerpos neutralizantes que son capaces de unirse al polipéptido compiten específicamente con un compuesto de ensayo por la unión. De esta manera, se pueden usar los anticuerpos para detectar la presencia de cualquier compuesto de ensayo que posea una afinidad de unión específica hacia el polipéptido.

50 También se pueden diseñar ensayos para detectar el efecto de los compuestos de ensayo añadidos sobre la producción de mRNA que codifica el polipéptido en las células. Por ejemplo, se puede construir un ELISA que mide los niveles secretados o asociados a células del polipéptido mediante el uso de anticuerpos monoclonales o policlonales mediante métodos habituales conocidos en la técnica, y esto se puede usar para buscar compuestos que pueden inhibir o aumentar la producción del polipéptido a partir de células o tejidos manipulados de manera adecuada. Después se puede medir la formación de complejos de unión entre el polipéptido y el compuesto que se está ensayando.

55 Los métodos de ensayo que también se incluyen en los términos de la presente invención son aquellos que implican el uso de genes y polipéptidos de la invención en ensayos de sobreexpresión o de supresión. Tales ensayos implican la manipulación de los niveles de estos genes/polipéptidos en las células y la determinación del impacto de este acto de manipulación sobre la fisiología de las células manipuladas. Por ejemplo, tales experimentos revelan detalles de las rutas de señalización y las rutas metabólicas en las que están implicados los genes/polipéptidos particulares, generan información con respecto a las identidades de los polipéptidos con los que interaccionan los polipéptidos estudiados, y proporcionan pistas en cuanto a los métodos por los que se regulan los genes y proteínas relacionadas.

60

- Otra técnica para el cribado de fármacos que se puede usar proporciona el cribado de alto rendimiento de compuestos que tienen una afinidad de unión adecuada hacia el polipéptido de interés (véase la solicitud de patente internacional WO84/03564). En este método, se sintetiza un gran número de diferentes compuestos de ensayo pequeños en un sustrato sólido, que después se pueden hacer reaccionar con el polipéptido de la invención y se lavan. Una manera de inmovilizar el polipéptido es usar anticuerpos no neutralizantes. El polipéptido unido se puede detectar después mediante el uso de métodos que se conocen bien en la técnica. El polipéptido purificado se puede revestir también directamente sobre placas para el uso en las técnicas de cribado de fármacos anteriormente mencionadas.
- El polipéptido de la invención se puede usar para identificar receptores asociados a membrana o receptores solubles, por medio de métodos de unión a receptores estándar que se conocen en la técnica, tales como ensayos de unión y de entrecruzamiento de ligandos en los que el polipéptido se marca con un isótopo radiactivo, se modifica químicamente, o se fusiona a una secuencia peptídica que facilita su detección o purificación, y se incuba con una fuente del supuesto receptor (por ejemplo, una composición de células, membranas celulares, sobrenadantes celulares, extractos tisulares, o líquidos corporales). La eficacia de la unión se puede medir mediante el uso de técnicas bioquímicas tales como resonancia de plasmones superficiales (suministrada por Biacore AB, Uppsala, Suecia) y espectroscopia. Se pueden usar ensayos de unión para la purificación y la clonación del receptor, pero también pueden identificar agonistas y antagonistas del polipéptido, que compiten con la unión del polipéptido a su receptor. Los métodos habituales para llevar a cabo ensayos de cribado se conocen muy bien en la técnica.
- La invención incluye los agonistas, antagonistas, ligandos, sustratos y enzimas, y otros compuestos que modulan la actividad o la antigenicidad del polipéptido de la invención descubiertos mediante los métodos que se describieron anteriormente.
- La invención también proporciona composiciones farmacéuticas que comprenden un polipéptido, ácido nucleico, ligando o compuesto de la invención en combinación con un vehículo farmacéutico adecuado. Estas composiciones pueden ser adecuadas como reactivos terapéuticos o diagnósticos, o como otras composiciones inmunógenas, como se resume con detalle más adelante.
- Según la terminología usada en la presente memoria, una composición que contiene un polipéptido, ácido nucleico, ligando o compuesto [X] está "sustancialmente exenta de" impurezas [en la presente memoria, Y] cuando al menos un 85% en peso del total de X+Y en la composición es X. Preferiblemente, X comprende al menos alrededor de un 90% en peso del total de X+Y en la composición, más preferiblemente al menos alrededor de un 95%, 98% o incluso un 99% en peso.
- Las composiciones farmacéuticas deberían comprender preferiblemente una cantidad terapéuticamente eficaz del polipéptido, molécula de ácido nucleico, ligando, o compuesto de la invención. La expresión "cantidad terapéuticamente eficaz", tal como se usa en la presente memoria, se refiere a una cantidad de agente terapéutico necesaria para tratar, mejorar, o prevenir una enfermedad o afección seleccionada como objetivo, o para exhibir un efecto terapéutico o preventivo detectable. Para cualquier compuesto, se puede estimar inicialmente la dosis terapéuticamente eficaz en ensayos de cultivos celulares, por ejemplo, de células neoplásicas, o en modelos de animales, normalmente ratones, conejos, perros, o cerdos. El modelo de animal también se puede usar para determinar el intervalo de concentraciones adecuado y la vía de administración. Tal información se puede usar después para determinar las dosis útiles y las vías de administración en seres humanos.
- La cantidad eficaz exacta para un sujeto humano dependerá de la gravedad del estado patológico, la salud general del sujeto, la edad, el peso, y el sexo del sujeto, la dieta, el momento y la frecuencia de la administración, la(s) combinación(es) de fármacos, la sensibilidad a la reacción, y la tolerancia/respuesta a la terapia. Esta cantidad se puede determinar mediante experimentación rutinaria, y está dentro del juicio del clínico. En general, una dosis eficaz será de 0,01 mg/kg a 50 mg/kg, preferiblemente 0,05 mg/kg a 10 mg/kg. Las composiciones se pueden administrar individualmente a un paciente, o se pueden administrar en combinación con otros agentes, fármacos u hormonas.
- Una composición farmacéutica también puede contener un vehículo farmacéuticamente aceptable, para la administración de un agente terapéutico. Tales vehículos incluyen anticuerpos y otros polipéptidos, genes y otros agentes terapéuticos tales como liposomas, con tal de que el vehículo no induzca por sí mismo la producción de anticuerpos dañinos hacia el individuo que recibe la composición, y que se pueda administrar sin una toxicidad indeseada. Los vehículos adecuados pueden ser macromoléculas grandes, metabolizadas lentamente, tales como proteínas, polisacáridos, poli(ácidos lácticos), poli(ácidos glicólicos), aminoácidos poliméricos, copolímeros de aminoácidos y partículas virales inactivas.
- Se pueden usar sales farmacéuticamente aceptables, por ejemplo, sales de ácidos minerales tales como hidroclo-ros, hidrobromuros, fosfatos, sulfatos, y similares; y sales de ácidos orgánicos tales como acetatos, propionatos, malonatos, benzoatos, y similares. Una discusión minuciosa de los vehículos farmacéuticamente aceptables está disponible en Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Los vehículos farmacéuticamente aceptables de las composiciones terapéuticas pueden contener adicionalmente líquidos tales como agua, solución salina, glicerol y etanol. Además, en tales composiciones puede haber presentes sustancias auxiliares, tales como agentes humectantes o emulsionantes, sustancias tamponadoras del pH, y similares. Tales vehículos permiten que las composiciones farmacéuticas se formulen como comprimidos, píldoras, grageas, cápsulas, líquidos, geles, jarabes, suspensiones espesas, suspensiones, y similares, para la ingestión por parte del paciente.

Una vez formuladas, las composiciones de la invención se pueden administrar directamente al sujeto. Los sujetos a tratar pueden ser animales; en particular, se pueden tratar sujetos humanos.

Las composiciones farmacéuticas utilizadas en esta invención se pueden administrar mediante cualquier vía que incluye, pero sin limitación, las aplicaciones oral, intravenosa, intramuscular, intraarterial, intramedular, intratecal, intraventricular, transdérmica o transcutánea (por ejemplo, véase el documento WO98/20734), subcutánea, intraperitoneal, intranasal, enteral, tópica, sublingual, intravaginal o rectal. También se pueden usar pistolas de genes o hiposprays para administrar las composiciones farmacéuticas de la invención. En general, las composiciones terapéuticas se pueden preparar en forma de inyectables, en forma de disoluciones líquidas o suspensiones; también se pueden preparar formas sólidas adecuadas para la disolución, o suspensión, en vehículos líquidos antes de la inyección.

La administración directa de las composiciones se llevará a cabo en general mediante inyección, de manera subcutánea, intraperitoneal, intravenosa o intramuscular, o se administrará en el espacio intersticial de un tejido. Las composiciones también se pueden administrar en una lesión. El tratamiento puede ser mediante un calendario de dosis única o un calendario de dosis múltiples.

Si la actividad del polipéptido de la invención es excesiva en un estado patológico particular, hay disponibles varias aproximaciones. Una aproximación comprende administrar a un sujeto un compuesto inhibidor (antagonista) como se describió anteriormente, junto con un vehículo farmacéuticamente aceptable en una cantidad eficaz para inhibir la función del polipéptido, bloqueando la unión de los ligandos, sustratos, enzimas, receptores, o inhibiendo una segunda señal, y por lo tanto aliviando la afección anormal. Preferiblemente, tales antagonistas son anticuerpos. Lo más preferiblemente, tales anticuerpos son quiméricos y/o humanizados para minimizar su inmunogenicidad, como se describió previamente.

En otra aproximación, se pueden administrar formas solubles del polipéptido que retienen la afinidad de unión hacia el ligando, sustrato, enzima o receptor en cuestión. En general, el polipéptido se puede administrar en forma de fragmentos que retienen las porciones relevantes.

En una aproximación alternativa, se puede inhibir la expresión del gen que codifica el polipéptido mediante el uso de técnicas de bloqueo de la expresión, tales como el uso de moléculas de ácidos nucleicos inversas (como se describió anteriormente), generadas internamente o administradas por separado. Se pueden obtener modificaciones de la expresión génica diseñando secuencias complementarias o moléculas inversas (ADN, ARN, o APN) respecto de regiones del control, de 5' o reguladoras (secuencia señal, promotores, potenciadores e intrones) del gen que codifica el polipéptido. De manera similar, se puede llevar a cabo la inhibición mediante el uso de la metodología de emparejamiento de bases en "triple hélice". El emparejamiento en triple hélice es útil, ya que provoca la inhibición de la capacidad de la doble hélice de abrirse lo suficiente para la unión de las polimerasas, factores de transcripción, o moléculas reguladoras. Se han descrito en la bibliografía avances terapéuticos recientes que usan ADN triple (Gee, J. E. *et al.* (1994) En: Huber, B. E. y B. I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing Co., Mt. Kisco, NY). También se puede diseñar la secuencia complementaria o la molécula inversa para que bloquee la traducción del mRNA impidiendo que el transcrito se una a los ribosomas. Tales oligonucleótidos se pueden administrar o se pueden generar *in situ* a partir de la expresión *in vivo*.

Además, se puede impedir la expresión del polipéptido de la invención mediante el uso de ribozimas específicas hacia su secuencia de mRNA codificante. Las ribozimas son ARNs catalíticamente activos que pueden ser naturales o sintéticos (véase, por ejemplo, Usman, N, *et al.*, *Curr. Opin. Struct. Biol* (1996) 6(4), 527-33). Las ribozimas sintéticas se pueden diseñar para escindir específicamente mARNs en posiciones seleccionadas, por lo que se impide la traducción de los mARNs hasta el polipéptido funcional. Las ribozimas se pueden sintetizar con un esqueleto de fosfato de ribosa natural y bases naturales, tal como se hallan normalmente en las moléculas de ARN. De manera alternativa, las ribozimas se pueden sintetizar con esqueletos no naturales, por ejemplo, 2'-O-metil ARN, para proporcionar protección a la degradación por ribonucleasas, y pueden contener bases modificadas.

Las moléculas de ADN se pueden modificar para incrementar la estabilidad intracelular y la semivida. Las modificaciones posibles incluyen, pero sin limitación, la adición de secuencias flanqueantes en los extremos 5' y/o 3' de la molécula o el uso de fosforotioato o 2' O-metilo en vez de enlaces fosfodiéster en el esqueleto de la molécula. Este concepto es inherente en la producción de APNs, y se puede ampliar en todas estas moléculas mediante la inclusión de bases no tradicionales tales como inosina, queosina y butosina, así como formas modificadas de manera similar

mediante acetilo, metilo, tio de adenina, citidina, guanina, timina y uridina que no son reconocidas fácilmente por las endonucleasas endógenas.

5 Para tratar afecciones anormales relacionadas con la infra-expresión del polipéptido de la invención y su actividad, también hay disponibles varias aproximaciones. Una aproximación comprende administrar a un sujeto una cantidad terapéuticamente eficaz de un compuesto que activa el polipéptido, es decir, un agonista como se describió anteriormente, para aliviar la afección anormal. De manera alternativa, se puede administrar una cantidad terapéutica del polipéptido en combinación con un vehículo farmacéutico adecuado para restaurar el equilibrio fisiológico relevante del polipéptido.

10 Se puede emplear la terapia génica para llevar a cabo la producción endógena del polipéptido por las células relevantes del sujeto. La terapia génica se usa para tratar permanentemente la producción inadecuada del polipéptido sustituyendo un gen defectuoso por un gen terapéutico corregido.

15 La terapia génica de la presente invención se puede dar *in vivo* o *ex vivo*. La terapia génica *ex vivo* requiere el aislamiento y la purificación de células del paciente, la introducción de un gen terapéutico y la introducción de las células alteradas genéticamente de nuevo al paciente. En contraste, la terapia génica *in vivo* no requiere el aislamiento y la purificación de las células de un paciente.

20 El gen terapéutico generalmente se "empaqueta" para la administración a un paciente. Los vehículos de administración génica pueden ser no virales, tales como liposomas, o virus deficientes de replicación, tales como adenovirus, como describió Berkner, K. L., en Curr. Top. Microbiol. Immunol., 158, 39-66 (1992) o vectores de virus adenoasociados (VAA), como describió Muzyczka, N., en Curr. Top. Microbiol. Immunol., 158, 97-129 (1992) y la pat. de EE.UU. nº 5.252.479. Por ejemplo, una molécula de ácido nucleico que codifica un polipéptido de la invención se puede modificar para la expresión en un vector retroviral deficiente de replicación. Esta construcción de expresión se puede aislar después e introducirla en una célula de empaquetamiento transducida con un vector plasmídico retroviral que contiene ARN que codifica el polipéptido, de manera que la célula de empaquetamiento produce ahora partículas virales infecciosas que contienen el gen de interés. Estas células productoras se pueden administrar a un sujeto para la modificación de las células *in vivo* y la expresión del polipéptido *in vivo* (véase el Capítulo 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (y las referencias citadas en ese documento) en Human Molecular Genetics (1996), T Strachan y A P Read, BIOS Scientific Publishers Ltd).

35 Otra aproximación es la administración de "ADN desnudo", en la que el gen terapéutico se inyecta directamente en el torrente sanguíneo o en el tejido muscular.

En las situaciones en las que los polipéptidos o las moléculas de ácidos nucleicos de la invención son los agentes causantes de la enfermedad, la invención posibilita que se puedan usar en vacunas para generar anticuerpos hacia el agente causante de la enfermedad. Cuando el polipéptido o la molécula de ácido nucleico anteriormente mencionados es uno que está estimulado, el desarrollo de la vacuna puede implicar la generación de anticuerpos o células T hacia tales agentes (como se describió en el documento WO00/29428).

40 La administración genética de anticuerpos que se unen a los polipéptidos según la invención también se puede llevar a cabo, por ejemplo, como se describió en la solicitud de patente internacional WO98/55607.

45 Esta invención también se refiere al uso de moléculas de ácidos nucleicos según la presente invención como reactivos de diagnóstico. La detección de una forma mutada del gen caracterizado por las moléculas de ácidos nucleicos de la invención que está asociada a una disfunción proporcionará una herramienta de diagnóstico que se puede añadir a, o definir, el diagnóstico de una enfermedad, o la susceptibilidad hacia una enfermedad, que resulta de la infra-expresión, sobre-expresión o expresión espacial o temporal alterada del gen. Los individuos que portan mutaciones en el gen se pueden detectar a nivel del ADN mediante una diversidad de técnicas.

50 Las moléculas de ácidos nucleicos para el diagnóstico se pueden obtener a partir de las células de un sujeto, tal como a partir de la sangre, orina, saliva, biopsia tisular o material de autopsia. El ADN genómico se puede usar directamente para la detección o se puede amplificar enzimáticamente mediante el uso de PCR, reacción en cadena de la ligasa (LCR), amplificación mediante desplazamiento de cadenas (SDA), u otras técnicas de amplificación (véase Saiki *et al.*, Nature, 324, 163-166 (1986); Bej, *et al.*, Crit. Rev. Biochem. Molec. Biol., 26, 301-334 (1991); Birkenmeyer *et al.*, J. Virol. Meth., 35, 117-126 (1991); Van Brunt, J., Bio/Technology, 8, 291-294 (1990)) antes del análisis.

60 En una realización, este aspecto de la invención proporciona un método para diagnosticar una enfermedad en un paciente, que comprende determinar el nivel de expresión de un gen natural que codifica un polipéptido según la invención y comparar dicho nivel de expresión respecto de un nivel de control, en el que un nivel que es diferente respecto de dicho nivel de control es indicativo de la enfermedad. El método puede comprender las etapas de:

- a) poner en contacto una muestra de tejido del paciente con una sonda de ácido nucleico en condiciones rigurosas que permiten la formación de un complejo híbrido entre una molécula de ácido nucleico de la invención y la sonda;
- b) poner en contacto una muestra de control con dicha sonda en las mismas condiciones usadas en la etapa a);
- c) y detectar la presencia de complejos híbridos en dichas muestras;

en el que la detección de niveles del complejo híbrido en la muestra del paciente que difiere de los niveles del complejo híbrido en la muestra de control es indicativa de la enfermedad.

Un aspecto adicional de la invención comprende un método de diagnóstico que comprende las etapas de:

- a) obtener una muestra de tejido de un paciente que se va a ensayar en busca de la enfermedad;
- b) aislar una molécula de ácido nucleico según la invención de dicha muestra de tejido; y
- c) diagnosticar la enfermedad en el paciente detectando la presencia de una mutación en la molécula de ácido nucleico que está asociada a la enfermedad.

Para ayudar en la detección de moléculas de ácidos nucleicos en los métodos anteriormente descritos, se puede incluir una etapa de amplificación, por ejemplo mediante el uso de PCR.

Las deleciones y las inserciones se pueden detectar mediante un cambio en el tamaño del producto amplificado en comparación con el genotipo normal. Las mutaciones puntuales se pueden identificar hibridando el ADN amplificado a ARN marcado de la invención o, de manera alternativa, secuencias de ADN inversas marcadas de la invención. Las secuencias perfectamente coincidentes se pueden distinguir de las moléculas dobles emparejadas erróneamente mediante digestión con RNasas o determinando diferencias en las temperaturas de fusión. La presencia o la ausencia de la mutación en el paciente se puede detectar poniendo en contacto el ADN con una sonda de ácido nucleico que hibrida con el ADN en condiciones rigurosas para formar una molécula bicatenaria híbrida, y la molécula bicatenaria híbrida tiene una porción sin hibridar de la cadena de la sonda de ácido nucleico en cualquier porción que corresponda a una mutación asociada a la enfermedad; y detectando la presencia o ausencia de una porción sin hibridar de la cadena de la sonda como indicación de la presencia o ausencia de una mutación asociada a la enfermedad en la porción correspondiente de la cadena de ADN.

Tales diagnósticos son especialmente útiles para los ensayos prenatales e incluso neonatales.

Las mutaciones puntuales y otras diferencias en las secuencias entre el gen de referencia y los genes "mutantes" se pueden identificar mediante otras técnicas muy conocidas, tales como secuenciación directa del ADN o polimorfismo conformacional monocatenario (véase, Orita *et al.*, Genomics, 5, 874-879 (1989)). Por ejemplo, se puede usar un cebador de secuenciación con un producto de PCR bicatenario o con una molécula molde monocatenaria generada mediante una PCR modificada. La determinación de la secuencia se lleva a cabo mediante procedimientos convencionales con nucleótidos radiomarcados o mediante procedimientos de secuenciación automáticos con marcadores fluorescentes. También se pueden usar segmentos de ADN clonados como sondas para detectar segmentos de ADN específicos. La sensibilidad de este método aumenta en gran medida cuando se combina con la PCR. Además, se pueden detectar mutaciones puntuales y otras variaciones de la secuencia, tales como polimorfismos, como se describió anteriormente, por ejemplo, por medio del uso de oligonucleótidos específicos de alelos para la amplificación mediante PCR de secuencias que difieren en oligonucleótidos individuales.

Las diferencias en las secuencias de ADN se pueden detectar también mediante las alteraciones de la movilidad electroforética de los fragmentos de ADN en geles, con o sin agentes desnaturizantes, o mediante secuenciación directa del ADN (por ejemplo, Myers *et al.*, Science (1985) 230:1242). También se pueden revelar los cambios en la secuencia en lugares específicos mediante ensayos de protección de nucleasas, tales como la protección de RNasa y S1 o el método de escisión química (véase, Cotton *et al.*, Proc. Natl. Acad. Sci. USA (1985) 85:4397-4401).

Además de la electroforesis en gel convencional y la secuenciación del ADN, también se pueden detectar mutaciones tales como microdeleciones, aneuploidías, translocaciones, inversiones, mediante análisis in situ (véase, por ejemplo, Keller *et al.*, DNA Probes, 2ª Ed., Stockton Press, Nueva York, N.Y., EE.UU. (1993)), es decir, se pueden analizar las mutaciones en las secuencias de ADN o ARN en células sin la necesidad de su aislamiento y/o inmovilización en una membrana. La hibridación in situ con fluorescencia (FISH) es actualmente el método aplicado más habitualmente, y han aparecido numerosas revisiones de FISH (véase, por ejemplo, Trachuck *et al.*, Science, 250, 559-562 (1990), y Trask *et al.*, Trends, Genet., 7, 149-154 (1991)).

Se puede construir una matriz de sondas oligonucleotídicas que comprende una molécula de ácido nucleico según la invención para llevar a cabo un cribado eficaz de variantes genéticas, mutaciones y polimorfismos. Se conocen muy

bien los métodos de la técnica de matrices, y tienen una aplicabilidad general y se pueden usar para abordar una diversidad de cuestiones en genética molecular que incluyen la expresión génica, el ligamiento genético, y la variabilidad genética (véase, por ejemplo: M. Chee *et al.*, Science (1996), vol. 274, págs. 610-613).

5 En una realización, se prepara la matriz y se usa según los métodos descritos en la solicitud PCT WO95/11995 (Chee *et al.*); Lockhart, D. J. *et al.* (1996) Nat. Biotech. 14: 1675-1680); y Schena, M. *et al.* (1996) Proc. Natl. Acad. Sci. 93: 10614-10619). Los pares de oligonucleótidos pueden oscilar de dos a más de un millón. Los oligómeros se sintetizan en áreas designadas sobre un sustrato mediante el uso de un procedimiento químico dirigido por la luz. El sustrato puede ser papel, nailon u otro tipo de membrana, filtro, chip, portaobjetos de vidrio o cualquier otro soporte sólido adecuado. En otro aspecto, se puede sintetizar un oligonucleótido sobre la superficie del sustrato mediante el uso de un procedimiento de acoplamiento químico y un aparato de aplicación de chorro de tinta, como se describe en la solicitud PCT WO95/25116 (Baldeschweiler *et al.*). En otro aspecto, se puede usar una matriz "cuadriculada" análoga a una transferencia en punto (o ranura) para ordenar y unir fragmentos de cADN u oligonucleótidos en la superficie de un sustrato mediante el uso de un sistema de vacío, o procedimientos de unión térmica, mediante UV, mecánica o química. Se puede producir una matriz, tal como se describió anteriormente, a mano o mediante el uso de dispositivos disponibles (aparato de transferencia en ranura o de transferencia en punto), materiales (cualquier soporte sólido adecuado), y aparatos (que incluyen los instrumentos robotizados), y pueden contener 8, 24, 96, 384, 1536 ó 6144 oligonucleótidos, o cualquier otro número entre dos y más de un millón que se preste al uso eficaz de la instrumentación disponible comercialmente.

20 Además de los métodos discutidos anteriormente, se pueden diagnosticar enfermedades mediante métodos que comprenden determinar, a partir de una muestra derivada de un sujeto, un nivel anormalmente disminuido o incrementado de polipéptido o mRNA. La expresión disminuida o incrementada se puede medir a nivel del ARN mediante el uso de cualquiera de los métodos conocidos en la técnica para la cuantificación de polinucleótidos, tales como, por ejemplo, la amplificación de ácidos nucleicos, por ejemplo PCR, RT-PCR, protección de RNasas, transferencia de Northern y otros métodos de hibridación.

25 Las técnicas de ensayo que se pueden usar para determinar los niveles de un polipéptido de la presente invención en una muestra derivada de un hospedador son muy conocidas para los expertos en la técnica, y se discutieron con cierto detalle anteriormente (e incluyen los radioinmunoensayos, ensayos de unión competitiva, análisis de transferencia de Western y ensayos de ELISA). Este aspecto de la invención proporciona un método de diagnóstico que comprende las etapas de: (a) poner en contacto un ligando como se describió anteriormente con una muestra biológica en condiciones adecuadas para la formación de un complejo ligando-polipéptido; y (b) detectar dicho complejo.

30 Los protocolos tales como ELISA, RIA, y FACS para medir los niveles de polipéptido pueden proporcionar adicionalmente una base para diagnosticar niveles alterados o anormales de expresión del polipéptido. Los valores normales o estándar para la expresión del polipéptido se establecen combinando líquidos corporales o extractos celulares tomados de sujetos mamíferos normales, preferiblemente seres humanos, con anticuerpos hacia el polipéptido en condiciones adecuadas para la formación de los complejos. La cantidad de formación de complejos estándar se puede cuantificar mediante diversos métodos, tales como los medios fotométricos.

35 Los anticuerpos que se unen específicamente a un polipéptido de la invención se pueden usar para el diagnóstico de afecciones o enfermedades caracterizadas por la expresión del polipéptido, o en ensayos para monitorizar pacientes que están siendo tratados con los polipéptidos, las moléculas de ácidos nucleicos, los ligandos y otros compuestos de la invención. Los anticuerpos útiles para fines de diagnóstico se pueden preparar de la misma manera que los descritos anteriormente para la terapia. Los ensayos de diagnóstico del polipéptido incluyen métodos que utilizan el anticuerpo y un marcador para detectar el polipéptido en líquidos corporales humanos o en extractos de células o tejidos. Los anticuerpos se pueden usar con o sin modificaciones, y se pueden marcar uniéndolos, de manera covalente o no covalente, a una molécula indicadora. Se puede usar una amplia diversidad de moléculas indicadoras conocidas en la técnica, varias de las cuales se describieron anteriormente.

40 Las cantidades de polipéptido expresado en una muestra de un sujeto, un control y las muestras patológicas de tejidos obtenidos mediante biopsia se comparan con los valores patrón. La desviación entre los valores patrón y los valores de un sujeto establece los parámetros para diagnosticar la enfermedad. Se pueden usar ensayos de diagnóstico para distinguir entre la ausencia, la presencia, y el exceso de expresión del polipéptido y para monitorizar la regulación de los niveles del polipéptido durante una intervención terapéutica. Tales ensayos se pueden usar también para estudiar la eficacia de un régimen de tratamiento terapéutico particular en estudios en animales, en ensayos clínicos o en la monitorización del tratamiento de un paciente individual.

45 Un equipo de diagnóstico de la presente invención puede comprender:

- (a) una molécula de ácido nucleico de la presente invención;
- (b) un polipéptido de la presente invención; o

50

(c) un ligando de la presente invención.

En un aspecto de la invención, un equipo de diagnóstico puede comprender un primer recipiente que contiene una sonda de ácido nucleico que hibrida en condiciones rigurosas con una molécula de ácido nucleico según la invención; un segundo recipiente que contiene cebadores útiles para amplificar la molécula de ácido nucleico; e instrucciones para el uso de la sonda y de los cebadores para facilitar el diagnóstico de la enfermedad. El equipo puede comprender además un tercer recipiente que alberga un agente para digerir el ARN sin hibridar.

En un aspecto alternativo de la invención, un equipo de diagnóstico puede comprender una matriz de moléculas de ácidos nucleicos, al menos una de las cuales puede ser una molécula de ácido nucleico según la invención.

Diversos aspectos y realizaciones de la presente invención se describirán a continuación con más detalle a modo de ejemplo, con referencia particular a los polipéptidos INSP005.

Breve descripción de las Figuras

Figura 1: Resumen de los resultados de búsquedas en bases de datos mediante el uso de la secuencia polipeptídica predicha de INSP005 como secuencia problema (alineamientos de secuencias mostrados).

Figura 2: Tabla de bibliotecas de cADN humano usadas en la investigación mediante clonación de INSP005.

Figura 3: Secuencia de nucleótidos del polipéptido predicho INSP005 y secuencia de aminoácidos predicha.

Figura 4: Tabla de cebadores de clonación de INSP005.

Figura 5: Secuencia de nucleótidos de 3' y secuencia de aminoácidos de INSP005 identificada mediante PCR RACE.

Figura 6: Tabla de cebadores usados durante la secuenciación de INSP005.

Figura 7: INSP005a supuesto de tamaño completo clonado a partir de cADN de útero humano.

Figura 8: Blastp de INSP005a frente a la base de datos NCBI-nr (diez mayores coincidencias y alineaciones más relacionadas mostradas).

Figura 9: Mapa del plásmido de clonación de INSP005a PCR4-TOPO-IPAAAIPAAA7883-1.

Figura 10: INSP005b supuesto de tamaño completo clonado a partir de una mezcla de cADNs derivados de fibroblastos primarios de pulmón humano, queratinocitos y líquido sinovial de osteoartritis.

Figura 11: Blastp de INSP005b frente a la base de datos NCBI-nr (diez mayores coincidencias y alineaciones más relacionadas mostradas).

Figura 12: Mapa del plásmido de clonación de INSP005b PCR-TOPO-IPAAA78836-2.

Figura 13: Alineación múltiple de la secuencia polipeptídica predicha de INSP005, la secuencia polipeptídica clonada de INSP005a, la secuencia polipeptídica clonada de INSP005b y ciertas secuencias de interés de la técnica anterior.

Figura 14: Datos de predicción del péptido señal mediante SignalP para el polipéptido INSP005b.

Figura 15A: Efecto de la administración de hIL-6 o plásmido de INSP005b sobre los niveles séricos de ASAT.

Figura 15B: Efecto de la administración de hIL-6 o plásmido de INSP005b sobre los niveles séricos de ALAT.

Figura 16A: Efecto de la administración de hIL-6 o plásmido de INSP005b sobre los niveles séricos de mL-6.

Figura 16B: Efecto de la administración de hIL-6 o plásmido de INSP005b sobre los niveles séricos de TNF-alfa.

Ejemplos

Ejemplo 1: Polipéptido Predicho INSP005

Se usó una secuencia del polipéptido INSP005 (SEQ ID N°:37) predicha mediante técnicas bioinformáticas patentadas como secuencia problema para búsquedas en las siguientes bases de datos:

NCBI-nr	NCBI-nt	NCBI-pat-aa
NCBI-pat-nt	NCBI-month-aa	NCBI-month-nt
NCBI-est		

Los resultados de estas búsquedas se resumen en la Figura 1, que muestra dos alineaciones de secuencias relevantes. Los encabezamientos de la Figura 1 indican qué algoritmos de búsqueda/alineamiento se usaron y en qué base de datos se buscó. Estos resultados demuestran que la coincidencia más relacionada con la secuencia polipeptídica predicha de INSP005 es la enzima de eclosión EHE4 de *Anguilla japonica* (anguila japonesa). Estas búsquedas también identificaron otras tres secuencias de interés de la técnica anterior, que se discuten con más detalle más adelante.

Los miembros de la familia similar a coriolisina/astacina de metaloproteasas se han implicado en el endurecimiento del corion de óvulos de peces ovíparos tras la fertilización (para un ejemplo véase Shibata *et al.* (2000) J. Biol. Chem vol. 275, N° 12 pág. 8349). Este cambio post-fertilización impide la polispermia y corresponde a la formación de membranas de fertilización en erizo de mar, anfibios y la reacción de la zona en mamíferos. También se han implicado en la hidrólisis del corion endurecido en el momento de la eclosión y en la hidrólisis del corion de los óvulos sin fertilizar.

Como se describió anteriormente, la identificación de metaloproteasas nuevas es de gran importancia en la comprensión creciente de las rutas subyacentes que conducen a ciertos estados patológicos en los que están implicadas estas proteínas, y en el desarrollo de terapias génicas o farmacológicas más eficaces para tratar estos trastornos. De manera similar, la identificación de miembros adicionales de la familia similar a astacina/coriolisina de metaloproteasas es de gran importancia en la comprensión creciente de las rutas subyacentes que conducen a ciertos estados patológicos en los que están implicadas estas proteínas, y en el desarrollo de terapias génicas o farmacológicas más eficaces para tratar estos trastornos.

Ejemplo 2: Resumen de la clonación de INSP005

1.1 Bibliotecas de cADN

Se adquirieron bibliotecas de cADN humano (en vectores de bacteriófago lambda (λ)) de Stratagene o Clontech, o se prepararon en el Serono Pharmaceutical Research Institute en vectores λ ZAP o λ GT10 según el protocolo del fabricante (Stratagene). Se preparó ADN de bacteriófago λ a partir de cultivos a pequeña escala de una cepa hospedadora de *E. coli* infectada mediante el uso del sistema de purificación de ADN Wizard Lambda Preps según las instrucciones del fabricante (Promega Corporation, Madison WI). La lista de bibliotecas y de cepas hospedadoras usadas se muestra en la Figura 2. Se usaron ocho mezclas (A-H) de cinco bibliotecas diferentes (100 ng/ μ l de ADN de fago) en las reacciones de PCR posteriores.

1.2 Generación de moldes de cADN transcritos inversamente

Se aisló el ARN total a partir de células humanas primarias, líneas celulares humanas y tejidos humanos mediante el uso del reactivo Trizol™ (Invitrogen) según las instrucciones del fabricante, o se adquirió de Clontech, Invitrogen o Ambion. La calidad y la concentración del ARN se analizaron mediante el uso de un bioanalizador Agilent 2100.

Para la síntesis de cADN, la mezcla de reacción contuvo: 1 μ l de cebador de oligo (dT)₁₅ (500 μ g/ml, Promega, n° de cat. C 1101), 2 μ g de ARN total, 1 μ l de dNTPs 10 mM en un volumen de 12 μ l. La mezcla se calentó a 65 °C durante 5 min y después se enfrió sobre hielo. Después se añadieron los siguientes reactivos: 4 μ l de tampón de la primera cadena 5X, 2 μ l de DTT (0,1 M), 1 μ l de inhibidor de ribonucleasas recombinante RNaseOut (40 unidades/ μ l, Promega, n° de cat. N 2511) y se incubó a 42 °C durante 2 min antes de la adición de 1 μ l (200 unidades) de Superscript II (Invitrogen, n° de cat. 18064-014). La mezcla se incubó a 42 °C durante 50 min y después se calentó a 70 °C durante 15 min. Para eliminar el molde de ARN, se añadió 1 μ l (2 unidades) de RNasa H de *E. coli* (Invitrogen, n° de cat. 18021-014) y la mezcla de reacción se incubó adicionalmente a 37 °C durante 20 min. La mezcla de reacción final se diluyó hasta 200 μ l con agua estéril y se almacenó a -80 °C. Se generaron mezclas de cADN mezclando volúmenes iguales de 5 moldes de cADN diferentes.

1.3 PCR de cADNs virtuales de ADN de biblioteca de fagos

Se obtuvo un cADN parcial que codificaba INSP005 en forma de un producto de amplificación de PCR de 248 pb mediante el uso de cebadores de clonación específicos de genes (CP1 y CP2, Figura 3 y Figura 4). La PCR se llevó a cabo en un volumen final de 50 μ l que contenía tampón AmpliTaq™ 1X, 200 μ M de dNTPs, 50 pmoles de cada uno de los cebadores de clonación, 2,5 unidades de AmpliTaq™ (Perkin Elmer) y 100 ng de cada mezcla de ADN de bibliotecas de fagos mediante el uso de un aparato MJ Research DNA Engine, programado como sigue: 94 °C, 1

min; 40 ciclos de 94 °C, 1 min, x °C, e y min y 72 °C, (en los que x es la T_m más baja - 5 °C e y = 1 min por kb de producto); seguido de 1 ciclo a 72 °C durante 7 min y un ciclo de espera a 4 °C.

5 Los productos de amplificación se visualizaron en geles de agarosa del 0,8 % en tampón TAE 1 X (Invitrogen), y los productos de PCR que migraban a la masa molecular predicha se purificaron a partir del gel mediante el uso del sistema de purificación de ADN Wizard PCR Preps (Promega). Los productos de PCR eluidos en 50 µl de agua estéril se subclonaron directamente o se almacenaron a -20 °C.

1.4 Cebadores de clonación específicos de gen para PCR

10 Se diseñaron pares de cebadores de PCR que tenían una longitud de entre 18 y 25 bases para amplificar la secuencia de tamaño completo y parcial del cADN virtual mediante el uso del Primer Designer Software (Scientific & Educational Software, PO Box 72045, Durham, NC 27722-2045, EE.UU.). Los cebadores de PCR se optimizaron para que tuvieran una T_m cercana a 55 ± 10 °C y un contenido de GC del 40-60%. Se seleccionaron cebadores que tenían una selectividad elevada hacia la secuencia objetivo INSP005 (poco o ningún cebado inespecífico).

1.5 Subclonación de los productos de PCR

20 Los productos de PCR se subclonaron en el vector de clonación modificado mediante topoisomerasa I (pCRII TOPO) mediante el uso del equipo de clonación TA adquirido de Invitrogen Corporation con el uso de las condiciones especificadas por el fabricante. Brevemente, 4 µl de producto de PCR purificado a partir de la amplificación de la mezcla N de bibliotecas humanas se incubaron durante 15 min a temperatura ambiente con 1 µl de vector TOPO y 1 µl de solución salina. La mezcla de reacción se transformó después en la cepa de *E. coli* TOP10 (Invitrogen) como sigue: una alícuota de 50 µl de células One Shot TOP10 se descongeló en hielo y se añadieron 2 µl de reacción de TOPO. 25 La mezcla se incubó durante 15 min en hielo y después se sometió a choque térmico mediante incubación a 42 °C durante exactamente 30 s. Las muestras se devolvieron al hielo y se añadieron 250 µl de medios SOC calientes (temperatura ambiente). Las muestras se incubaron con agitación (220 rpm) durante 1 h a 37 °C. La mezcla de transformación se colocó después en placas de caldo L (LB) que contenían ampicilina (100 µg/ml) y se incubaron durante la noche a 37 °C. Las colonias resistentes a ampicilina que contenían insertos de cADN se identificaron 30 mediante PCR de las colonias.

1.6 PCR de colonias

35 Las colonias se inocularon en 50 µl de agua estéril mediante el uso de un palillo estéril. Una alícuota de 10 µl del inóculo se sometió después a PCR en un volumen de reacción total de 20 µl como se describió anteriormente, excepto que los cebadores usados fueron SP6 y T7. Las condiciones cíclicas fueron las siguientes: 94 °C, 2 min; 30 ciclos de 94 °C, 30 seg, 47 °C, 30 seg y 72 °C durante 1 min; 1 ciclo, 72 °C, 7 min. Las muestras se mantuvieron después a 4 °C (ciclo de espera) antes del análisis posterior.

40 Los productos de la reacción de PCR se analizaron en geles de agarosa del 1% en tampón TAE 1 X. Las colonias que proporcionaron el tamaño del producto de PCR esperado (cADN de 248 pb + 185 pb debido al sitio de clonación múltiple o MCS) se cultivaron durante la noche a 37 °C en 5 ml de caldo L (LB) que contenía ampicilina (100 µg/ml), con agitación a 220 rpm a 37 °C

1.7 Preparación y secuenciación de ADN plasmídico

50 Se preparó ADN plasmídico Miniprep a partir de cultivos de 5 ml mediante el uso de un sistema robótico Qiaprep Turbo 9600 (Qiagen) o un equipo Wizard Plus SV Minipreps (Promega, n° de cat. 1460) según las instrucciones del fabricante. El ADN plasmídico se eluyó en 100 µl de agua estéril. Se midió la concentración de ADN mediante el uso de un fotómetro Eppendorf BO. Se sometió el ADN plasmídico (200-500 ng) a secuenciación del ADN con cebador T7 y cebador SP6 mediante el uso del sistema BigDyeTerminator (Applied Biosystems, n° de cat. 4390246) según las instrucciones del fabricante. Las reacciones de secuenciación se purificaron mediante el uso de columnas Dye-Ex (Qiagen) o placas de limpieza Montage SEQ 96 (Millipore, n° de cat. LSKS09624) y después se analizaron en un secuenciador Applied Biosystems 3700.

1.8 Identificación de la secuencia de tamaño completo de INSP005 mediante el uso de PCR RACE.

60 La secuencia predicha del ORF de INSP005 se muestra en la Figura 3. Los intentos de aislar la secuencia codificante de tamaño completo mediante PCR fracasaron con las bibliotecas analizadas, mediante el uso de pares de cebadores para amplificar la predicción de tamaño completo o una versión más corta que usaba un 2^a sitio de inicio predicho en M96 en el marco de lectura abierto. Las secuencias relacionadas más cercanas a INSP005 son la metalopeptidasa similar a astacina en *Anguilla japonica* y coriolisina H en *Oryzias latipes*. INSP005 parece ser un ortólogo humano de la coriolisina H. Las coriolisinas están implicadas en el endurecimiento del corion de óvulos de peces ovíparos tras la fertilización, lo que sugiere que el útero puede ser una fuente adecuada de mRNA de INSP005.

La elección de este tejido estuvo apoyada adicionalmente por el hallazgo de una única EST, B1061462 derivado de un tumor de útero humano.

5 Para identificar la secuencia codificante completa, se llevó a cabo una PCR RACE con cADN preparado a partir de ARN de útero (adquirido de Clontech) mediante el uso del equipo GeneRacer (Invitrogen, cat. nº L1502-01) según las instrucciones del fabricante. Para la amplificación de los extremos 3', la primera PCR se llevó a cabo en un volumen de reacción de 50 µl que contenía 1 µl de cADN RACE Ready, 5 µl de tampón High Fidelity 10X, 1 µl de dNTPs (10 mM), 2 µl de MgSO₄ 50 mM, 3 µl de cebador de 3' GeneRacer (10 µM), 1 µl de cebador específico del gen (78836-GR1-3') (10 µM) y 2,5 unidades (0,5 µl) de Platinum Taq ADN polimerasa Hi Fi (Invitrogen). Las condiciones cíclicas fueron las siguientes: 94 °C, 2 min; 5 ciclos de 94 °C, 30 s y 72 °C, 2 min; 5 ciclos de 94 °C, 30 s y 70 °C, 5 min; 25 ciclos de 94 °C, 30 s, 65 °C, 30 s y 68 °C, 5 min; una extensión final a 68 °C durante 10 min y un ciclo de espera de 4 °C. Después se usó un µl de la reacción de amplificación como molde para una PCR anidada que se llevó a cabo en un volumen de reacción final de 50 µl con los mismos reactivos que anteriormente, excepto por los cebadores. Los cebadores para la PCR anidada fueron 1 µl de cebador anidado de 3' GeneRacer (10 µM) y 1 µl de cebador específico del gen anidado (78836-GR1nest-3') (10 µM). Las condiciones cíclicas fueron 94 °C, 2 min; 25 ciclos de 94 °C, 30 s, 65 °C, 30 s y 68 °C, 5 min; una extensión final a 68 °C durante 10 min y un ciclo de espera de 4 °C. Los productos de la PCR se purificaron en un gel, se subclonaron en el vector pCR4-TOPO y se secuenciaron como se describió anteriormente. Todos los cebadores usados se enumeran en la Figura 4. La secuencia de nucleótidos y la secuencia de aminoácidos del producto de 3' RACE se muestran en la Figura 5. La secuencia de aminoácidos codificada por el producto de 3' RACE tiene un extremo C-terminal prolongado, que diverge de la predicción después de la posición del nucleótido 85, lo que sugiere una forma sometida a un corte y empalme alternativo.

1.9 Clonación de la secuencia codificante de tamaño completo de INSP005 mediante PCR

25 La secuencia codificante de tamaño completo supuesta de INSP005 se clonó a partir de cADN de útero humano (preparado como se describió en la sección 1.2) mediante PCR en 50 µl de mezcla de reacción de PCR que contenía 2 µl de cADN de útero, 5 µl de tampón High Fidelity 10X, 1 µl de dNTPs (10 mM), 2 µl de MgSO₄ 50 mM, 1 µl de cebador 78836-FL-F específico del gen (10 µM), 1 µl de cebador inverso 78836-FL-R específico del gen (10 µM) y 2,5 unidades (0,5 µl) de Platinum Taq ADN polimerasa Hi Fi (Invitrogen). Las condiciones cíclicas fueron 94 °C, 2 min; 40 ciclos de 94 °C, 30 s, 55 °C, 30 s y 68 °C, 1 min 30 s; una extensión final a 68 °C durante 10 min y un ciclo de espera de 4 °C. Los productos de amplificación se visualizaron en geles de agarosa del 0,8 % en tampón TAE 1 X (Invitrogen) y los productos de PCR que migraban a la masa molecular predicha (1048 pb) se purificaron a partir del gel mediante el uso del sistema de purificación de ADN Wizard PCR Preps (Promega). Los productos de PCR se eluyeron en 50 µl de agua estéril y se subclonaron en el vector pCR4 TOPO como se describió en la sección 35 1.4. Varias colonias resistentes a ampicilina se sometieron a una PCR de colonias como se describió en la sección 1.5, excepto por el tiempo de extensión en la reacción de amplificación, que fue de 2 min. Las colonias que contenían el inserto de tamaño correcto (1048 pb + 99 pb debido al MCS) se cultivaron durante la noche a 37 °C en 5 ml de caldo L (LB) que contenía ampicilina (100 µg/ml), con agitación a 220 rpm a 37 °C. Se preparó ADN plasmídico Miniprep a partir de cultivos de 5 ml mediante el uso de un sistema robótico Qiaprep Turbo 9600 (Qiagen) o un equipo Wizard Plus SV Minipreps (Promega, nº de cat. 1460) según las instrucciones del fabricante, y se secuenciaron 200-500 ng de ADN mini-prep como se describió en la sección 1.7 con cebadores T3 y T7 (Figura 6). La secuencia clonada se proporciona en la Figura 7. La alineación de aminoácidos de la secuencia clonada (INSP005a) con la secuencia predicha se muestra en la Figura 13. El mapa del plásmido resultante, pCR4-TOPO-IPAAA78836-1 (SEQ ID Nº:38; ID. de plásmido Nº 13164) se muestra en la Figura 9.

2.0 Identificación de bibliotecas/moldes de cADN que contienen INSP005

Los productos de PCR obtenidos con CP1 y CP2 y que migraban en el tamaño correcto (248 pb) se identificaron en la mezcla N de bibliotecas (bibliotecas 18, 19, 20 y 21). Se aisló un cADN que codificaba un supuesto INSP005 de tamaño completo (INSP005a) a partir de cADN de útero mediante el uso de cebadores 78836-FL-F y 78836-FL-R. El cebador 78836-FL-F se localiza en el exón 3 de la secuencia predicha. No se obtuvieron productos de PCR mediante el uso del cebador inverso (78836-FL-R) con cebadores localizados en el exón 1 de la predicción.

55 Una segunda versión de tamaño completo supuesta de INSP005 (INSP005b) que contenía un extremo 5' alternativo se clonó en una mezcla de cADNs derivados de fibroblastos primarios de pulmón humano, queratinocitos y líquido sinovial de osteoartritis mediante el uso de los cebadores 78836-FL2-F y 78836-FL-R, pero no se detectó en el útero. El producto de PCR resultante (1313 pb - Figura 10) se subclonó en el vector pCR4 TOPO mediante el uso del equipo de clonación TOPO-TA y se secuenció como se describió en las secciones 1.5-1.7. El mapa del plásmido resultante, pCR4-TOPO-IPAAA78836-2 (SEQ ID Nº:39; ID. de plásmido Nº 13296) se muestra en la Figura 12.

2.1 Resumen de los resultados de la clonación

Los intentos de clonar el polipéptido predicho INSP005 de tamaño completo identificaron dos variantes del polipéptido predicho INSP005, en la presente memoria denominadas INSP005a e INSP005b (Figura 13; SEQ ID Nº:14 y SEQ ID Nº:34, respectivamente). Como se describió anteriormente, los polipéptidos INSP005a e INSP005b (y el

polipéptido maduro INSP005b) se denominan en la presente memoria polipéptidos INSP005, a diferencia del polipéptido predicho INSP005.

5 Las secuencias de nucleótidos y de aminoácidos para los exones predichos en los polipéptidos INSP005a e INSP005b se proporcionan en SEQ ID N°s 1-12 y SEQ ID N°s 15-32, respectivamente. Como se describió anteriormente, las secuencias de nucleótidos de tamaño completo supuestas de los polipéptidos INSP005a e INSP005b se proporcionan en SEQ ID N°s 13 y 33, respectivamente. Las secuencias de aminoácidos de los polipéptidos INSP005a e INSP005b se proporcionan en SEQ ID N°s 14 y 34, respectivamente.

10 Las relaciones entre los polipéptidos INSP005a e INSP005b y el polipéptido predicho INSP005 y tres secuencias de la técnica anterior de interés se muestran en la Figura 13, que proporciona una alineación de las secuencias. Estas relaciones se describirán con detalle a continuación.

15 INSP005a es una versión de tamaño completo supuesta del polipéptido predicho INSP005 a partir de una biblioteca de cADN de útero. Esta secuencia difiere de la predicción de INSP005 original que tiene un extremo 5' truncado, que comienza en la metionina 3 del polipéptido predicho INSP005 original (véase la Figura 13). INSP005a también tiene un extremo 3' prolongado que incorpora un exón extra respecto del polipéptido predicho INSP005. INSP005a tiene seis exones predichos en total. Estas diferencias no se predijeron debido a la baja homología de esos elementos de secuencia respecto de otras metaloproteinasas. Además, hay un aminoácido alternativo usado en la posición 22 de
20 INSP005a en comparación con el polipéptido predicho INSP005. No se predice que INSP005a contenga un péptido señal. INSP005a no tiene una metionina de inicio en posición 5' alternativa en el marco de lectura antes de un codón STOP en posición 5'.

25 La secuencia polipeptídica mostrada en SEQ ID N°:14 (INSP005a) se usó como secuencia problema en BLAST respecto de la base de datos de secuencias no redundantes del NCBI. Las diez mayores coincidencias son todas enzimas relacionadas con la eclosión de huevos de *Anguilla japonica* o proteasas coriolíticas, y se alinean respecto de la secuencia problema con valores E muy significativos (de e^{-115} a $2e^{-41}$) (Figura 8). La Figura 8 también muestra la alineación de la secuencia problema del polipéptido INSP005a respecto de la secuencia de la mayor coincidencia anotada bioquímicamente, la enzima de eclosión HE13 de *Anguilla japonica*. Estos resultados proporcionan pruebas
30 sólidas de que el polipéptido INSP005a es una metaloproteasa, más específicamente es una metaloproteasa similar a coriolisina/astacina.

35 INSP005b es una versión de tamaño completo supuesta del polipéptido predicho INSP005 clonado a partir de una mezcla de cADNs derivados de fibroblastos primarios de pulmón humano, queratinocitos y líquido sinovial de osteoartritis. INSP005b incluye la secuencia del polipéptido predicho INSP005 original, aunque se usan dos aminoácidos alternativos en las posiciones 117 y 222. También contiene tres exones nuevos en posición 5' y un exón en posición 3', lo que hace que INSP005b sea un polipéptido de nueve exones. El exón final se comparte con INSP005a. INSP005b no se detectó en útero. Estas diferencias no se predijeron debido a la baja homología de esos elementos de secuencia respecto de otras metaloproteinasas. Tal como se describió anteriormente, se predice que
40 INSP005b contiene un péptido señal con un sitio de escisión entre los aminoácidos 23 y 24 (SEQ ID N°s 35 y 36; Figura 14).

45 La secuencia polipeptídica mostrada en SEQ ID N°:34 (INSP005b) se usó como secuencia problema en BLAST respecto de la base de datos de secuencias no redundantes del NCBI. Las diez mayores coincidencias son todas enzimas relacionadas con la eclosión de huevos de *Anguilla japonica* o proteasas coriolíticas, y se alinean respecto de la secuencia problema con valores E muy significativos (de e^{-152} a $4e^{-46}$) (Figura 11). La Figura 11 también muestra la alineación de la secuencia problema del polipéptido clonado respecto de la secuencia de la mayor coincidencia anotada bioquímicamente, la enzima de eclosión HE13 de *Anguilla japonica*. Estos resultados proporcionan pruebas sólidas de que el polipéptido INSP005b es una metaloproteasa, más específicamente es una metaloproteasa similar a coriolisina/astacina.
50

Los primeros 7 exones de INSP005b coinciden con una secuencia de nucleótidos descrita en el documento WO200216566-A2, a la que se dio el número de acceso AX443328 (véase la Figura 1 y la Figura 13), aunque el exón final de 3' no se describe en el documento WO200216566-A2 (Applera Corp). Las moléculas nucleotídicas y polipeptídicas de la presente invención excluyen específicamente las descritas en el documento WO200216566-A2.
55

Una secuencia de interés adicional de la técnica anterior es una EST sometida a corte y empalme (BI061462.1; véase la Figura 1) de tumor de útero que cubre el exón 1 de INSP005a y los exones 2, 3 y 4 de INSP005b. La dirección de EST no se proporciona en el informe, y es difícil llegar a una conclusión sobre la presencia de una metionina de inicio a partir de la traducción. Sin embargo, las moléculas nucleotídicas y polipeptídicas de la presente invención excluyen específicamente las secuencias descritas en la EST BI061462.1.
60

Otra secuencia de interés de la técnica anterior, con el número de acceso AX526191 (Lexicon) (Figuras 1 y 13), se describe como cADN en la entrada de la base de datos relevante (descrita en el documento WO02/066624), y no se hace referencia a un posible papel reproductivo. Incluye INSP005a y los exones 2-8 de INSP005b. Sin embargo, se
65

usa un aminoácido alternativo en la posición 127 de INSP005a en comparación con el aminoácido correspondiente de INSP005b y la secuencia de AX526191 (Lexicon). La metionina de inicio de AX526191 está cubierta por la EST de tumor de útero descrita anteriormente. Se predice un péptido señal para AX526191 con una probabilidad de 0,875. Las moléculas nucleotídicas y polipeptídicas de la presente invención no incluyen las secuencias descritas en el documento WO02/066624, que incluyen la que tiene el número de acceso AX526191.

La Figura 13 también resalta los residuos del sitio activo, que son idénticos en cada uno de los polipéptidos mostrados. Esto proporciona pruebas más convincentes de que los polipéptidos INSP005a e INSP005b son metaloproteasas.

Los polipéptidos INSP005a e INSP005b, por lo tanto, representan metaloproteasas nuevas, y hay pruebas sólidas de que son miembros de la familia similar a coriolisina/astacina de metaloproteasas. Los polipéptidos INSP005a e INSP005b, por lo tanto, pueden desempeñar papeles importantes en procesos fisiológicos y patológicos en seres humanos, en particular en procesos reproductivos.

Ejemplo 3: Expresión y purificación de INSP005b clonado y marcado con His

3.1 Expresión

Se mantuvieron células 293 de riñón embrionario humano que expresaban el antígeno nuclear del virus de Epstein-Barr (HEK293-EBNA, Invitrogen) en suspensión en medio exento de suero Ex-cell VPRO (reserva de siembra, medio de mantenimiento, JRH). De dieciséis a 20 horas antes de la transfección (Día 1), las células se sembraron en matraces 2x T225 (50 ml por matraz en DMEM/F12 (1:1) que contenían medio de siembra FBS del 2% (JRH) a una densidad de 2×10^5 células/ml). El siguiente día (transfección: día 0) la transfección tuvo lugar mediante el uso del reactivo JetPEI™ (2 µl/µg de ADN plasmídico, transfección PolyPlus). Para cada matraz, se co-transfectaron 113 µg de cADN (plásmido nº 13403) con 2,3 µg de GFP (gen indicador fluorescente). La mezcla de transfección se añadió después a los matraces 2xT225 y se incubó a 37 °C (5% de CO₂) durante 6 días. Para incrementar la cantidad de material, este procedimiento se repitió con dos matraces extra para generar 200 ml en total. Se llevó a cabo la confirmación de la transfección positiva mediante examen fluorescente cualitativo en el día 6 (Axiovert 10 Zeiss).

En el día 6 (día de la recogida), se mezclaron los sobrenadantes (200 ml) de los cuatro matraces y se centrifugaron (4 °C, 400 g) y se colocaron en un recipiente que tenía un único identificador.

Se guardó una alícuota (500 µl) para el QC de la proteína marcada con 6His (QC de bioprocesamiento interno). La hoja de entrega correspondiente se puede hallar en el libro de notas de T. Battle 11140, pág. 28.

Con fines de producción extra, el lote 2 se produjo en una transfección en tubo de centrifuga de 500 ml, como sigue:

Se mantuvieron células 293 de riñón embrionario humano que expresaban el antígeno nuclear del virus de Epstein-Barr (HEK293-EBNA, Invitrogen) en suspensión en medio exento de suero Ex-cell VPRO (reserva de siembra, medio de mantenimiento, JRH). En el día de la transfección, las células se contaron, se centrifugaron (velocidad baja) y el sedimento se re-suspendió en el volumen deseado de medio FEME (véase más adelante) complementado con un 1% de FCS para proporcionar una concentración celular de 1×10^6 células viables/ml. El cADN nº 13403 se diluyó hasta 2 mg/litro (co-transfectado con un 2% de eGFP) en FEME (200 ml/litro). Después se añadió el agente de transfección Polietilenimina (4 mg/litro) a la disolución de cADN, se agitó en vórtex y se incubó a temperatura ambiente durante 10 minutos (lo que generó la mezcla de transfección).

Esta mezcla de transfección se añadió después al tubo de centrifuga y se incubó durante 90 minutos en un incubador de CO₂ (5% de CO₂ y 37 °C). Después de 90 minutos se añadió medio FEME nuevo (1% FCS) para doblar el volumen inicial del tubo de centrifuga. El tubo de centrifuga se incubó después durante 6 días. En el día 6 (día de la recogida), se centrifugó (4 °C, 400 g) el sobrenadante del tubo de centrifuga (500 ml) y se colocó en un recipiente que tenía un único identificador con el número de plásmido y el número de fermentación.

Se guardó una alícuota (500 µl) para el QC de la proteína marcada con 6His (QC de bioprocesamiento interno).

3.2 Procedimiento de purificación

La muestra de medio de cultivo de 200 ml que contenía la proteína recombinante con un marcador de 6His C-terminal se diluyó hasta un volumen final de 400 ml con tampón A frío (NaH₂PO₄ 50 mM; NaCl 600 mM; 8,7% (p/v) de glicerol, pH 7,5). La muestra se filtró a través de un filtro estéril de 0,22 µm (Millipore, unidad de filtración de 500 ml) y se almacenó a 4 °C en una botella de medios cuadrada estéril de 500 ml (Nalgene).

La muestra de medio de cultivo de 500 ml se diluyó hasta un volumen final de 1000 ml con tampón A frío. La muestra se filtró a través de un filtro estéril de 0,22 µm (Millipore, unidad de filtración de 500 ml) y se almacenó a 4 °C en una botella de medios cuadrada estéril de 1000 ml (Nalgene).

Las purificaciones se llevaron a cabo a 4 °C en la estación de trabajo VISION (Applied Biosystems) conectada a un cargador de muestras automático (Labomatic). El procedimiento de purificación estaba compuesto de dos etapas secuenciales, cromatografía de afinidad de iones metálicos en una columna Poros 20 MC (Applied Biosystems) cargada con iones de Ni (4,6 x 50 mm, 0,83 ml), seguido de un intercambio de tampones en una columna de filtración en gel con medio Sephadex G-25 (Amersham Pharmacia) (1,0 x 15 cm).

Para la primera etapa de cromatografía la columna de afinidad de iones metálicos se regeneró con 30 volúmenes de columna de disolución de EDTA (EDTA 100 mM; NaCl 1 M; pH 8,0), se recargó con iones de Ni por medio del lavado con 15 volúmenes de columna de una disolución de NiSO₄ 100 mM, se lavó con 10 volúmenes de columna de tampón A, seguido de 7 volúmenes de columna de tampón B (NaH₂PO₄ 50 mM; NaCl 600 mM; 8,7% (p/v) de glicerol, imidazol 400 mM, pH 7,5), y finalmente se equilibró con 15 volúmenes de columna de tampón A que contenía imidazol 15 mM. La muestra se transfirió mediante el cargador de muestras Labomatic a un bucle de muestra de 200 ml y posteriormente se cargó en la columna de afinidad de iones metálicos de Ni a un caudal de 10 ml/min. El procedimiento de carga se repitió 2 y 5 veces, respectivamente, para transferir todo el volumen de muestra (400 ó 1000 ml) a la columna de Ni. La columna se lavó posteriormente con 12 volúmenes de columna de tampón A, seguido de 28 volúmenes de columna de tampón A que contenía imidazol 20 mM. Durante el lavado con imidazol 20 mM, las proteínas contaminantes unidas débilmente se eluyeron de la columna. La proteína marcada con His recombinante se eluyó finalmente con 10 volúmenes de columna de tampón B a un caudal de 2 ml/min, y la proteína eluida se recogió en una fracción de 1,6 ml.

Para la segunda etapa de cromatografía, la columna de filtración en gel de Sephadex G-25 se regeneró con 2 ml de tampón D (NaCl 1,137 M; KCl 2,7 mM; KH₂PO₄ 1,5 mM; Na₂HPO₄ 8 mM; pH 7,2), y se equilibró posteriormente con 4 volúmenes de columna de tampón C (NaCl 137 mM; KCl 2,7 mM; KH₂PO₄ 1,5 mM; Na₂HPO₄ 8 mM; 20% (p/v) de glicerol; pH 7,4). La fracción del pico eluida de la columna de Ni se cargó automáticamente por medio del cargador de muestras integrado en VISION en la columna de Sephadex G-25, y la proteína se eluyó con tampón C a un caudal de 2 ml/min. La muestra de proteína de la columna de Sephadex G-25 se recuperó en una fracción de 2,2 ml. La fracción se filtró a través de un filtro de centrifugación estéril de 0,22 µm (Millipore), se congeló y se almacenó a -80 °C. Se analizó una alícuota de la muestra mediante SDS-PAGE (gel NuPAGE del 4-12%; Novex) mediante tinción de Coomassie y transferencia de Western con anticuerpos anti-His.

Tinción de Coomassie. El gel NuPAGE se tiñó con una disolución de tinción R250 de azul de Coomassie al 0,1% (30% de metanol, 10% de ácido acético) a temperatura ambiente durante 1 h, y posteriormente se destiñó con un 20% de metanol, 7,5% de ácido acético hasta que el fondo fue claro y las bandas de proteína fueron claramente visibles.

Transferencia de Western. Tras la electroforesis, las proteínas se electrotransfirieron desde el gel a una membrana de nitrocelulosa a 290 mA durante 1 hora a 4 °C. La membrana se bloqueó con un 5% de leche en polvo en tampón E (NaCl 137 mM; KCl 2,7 mM; KH₂PO₄ 1,5 mM; Na₂HPO₄ 8 mM; 0,1% de Tween 20, pH 7,4) durante 1 h a temperatura ambiente, y posteriormente se incubó con una mezcla de 2 anticuerpos anti-His policlonales de conejo (G-18 y H-15, 0,2 µg/ml cada uno; Santa Cruz) en un 2,5% de leche en polvo en tampón E durante la noche a 4 °C. Después de 1 hora adicional de incubación a temperatura ambiente, la membrana se lavó con tampón E (3 x 10 min), y después se incubó con un anticuerpo secundario anti-conejo conjugado a HRP (DAKO, HRP 0399) diluido 1/3000 en tampón E que contenía un 2,5% de leche en polvo durante 2 horas a temperatura ambiente. Después de lavar con tampón E (3 x 10 minutos), la membrana se reveló con el equipo ECL (Amersham Pharmacia) durante 1 min. La membrana se expuso posteriormente a Hyperfilm (Amersham Pharmacia), la película se reveló y se analizó visualmente la imagen de la transferencia de Western.

Ensayo de proteínas. Se determinó la concentración de proteínas mediante el uso del equipo de ensayo de proteínas con BCA (Pierce) con albúmina de suero bovino como patrón. Se recuperaron 78 y 90 µg de proteína purificada a partir de las muestras de medios de cultivo de 200 ml y 500 ml, respectivamente.

Ejemplo 4: INSP005b en modelo de ratón de hepatitis fulminante

4.1 Introducción

Para caracterizar INSP005b *in vivo*, se usó la técnica de electrotransferencia muscular para expresar la proteína INSP005b en la circulación de animales de TN y tratados con ConA. No se observaron cambios significativos en los niveles de transaminasas en suero o en los niveles de citocinas TNF-alfa, IFN-gamma, IL-6, IL-4 o MCP-1 tras la electrotransferencia de INSP005b en animales de TN. Los animales sometidos a electroporación se expusieron después a ConA para determinar el efecto de INSP005b sobre los niveles de citocinas en suero y los niveles de transaminasas.

4.2 Antecedentes - Hepatitis inducida por Concanavalina A (ConA)

La enfermedad hepática tóxica representa un problema sanitario mundial en seres humanos, para la cual todavía se tienen que descubrir tratamientos farmacológicos. Por ejemplo, la hepatitis crónica activa que conduce a cirrosis hepática es un estado patológico en el que las células del parénquima hepático son destruidas progresivamente por las células T activadas. La toxicidad hepática inducida por ConA es uno de los tres modelos experimentales de lesión hepática apoptótica y necrótica dependiente de células T descritos en ratones. Los ratones sensibilizados mediante Gal N (D-Galactosamina) expuestos a AB monoclonales anti-CD3 activantes o con superantígeno SEB desarrollan lesiones hepáticas apoptóticas y lesiones necróticas secundarias graves (Kusters S, Gastroenterology. Agosto de 1996; 111(2):462-71). La inyección de la lectina vegetal mitógena de células T ConA a ratones sin sensibilizar da como resultado también la apoptosis hepática que precede a la necrosis. ConA induce la liberación de TNF-alfa e IFN-gamma sistémica y otras citocinas diversas. Tanto TNF-alfa como IFN-gamma son mediadores críticos de la lesión hepática. La liberación de transaminasas 8 horas después de la exposición indica una destrucción hepática grave. Se ha demostrado que están implicados varios tipos de células en la lesión hepática, las células T CD4, macrófagos y células citotóxicas naturales (Kaneko J Exp Med 2000, 191, 105-114). Los anticuerpos anti-CD4 bloquean la activación de las células T y por lo tanto la lesión hepática (Tiegs *et al.* 1992, J Clin Invest 90, 196-203). El pre-tratamiento de ratones con anticuerpos monoclonales hacia CD8 no produjo protección, mientras la deleción de los macrófagos previno la inducción de la hepatitis.

El presente estudio se llevó a cabo para investigar el papel de INSP005b, una proteína similar a coriolisina, en la hepatitis inducida por ConA. Se ha demostrado que varias citosinas son críticas para inducir o para conferir protección hacia la lesión hepática inducida por ConA. TNF-alfa, por ejemplo, es una de las primeras citocinas producidas tras la inyección de ConA, y los anticuerpos anti-TNF-alfa confieren protección contra la enfermedad (Seino *et al.* 2001, Annals of surgery 234, 681). IFN-gamma parece ser también un mediador crítico en la lesión hepática, ya que el antisuero anti-IFN-gamma protege significativamente a los ratones, tal como se mide mediante los niveles disminuidos de transaminasas en la sangre de los animales tratados con ConA (véase, Kusters *et al.*, anteriormente mencionado). En la lesión hepática, se observó una producción incrementada de IFN-gamma en los pacientes con hepatitis autoinmunitaria o viral. Además, los ratones transgénicos que expresan IFN-gamma en el hígado desarrollan una lesión hepática que se parece a la hepatitis activa crónica (Toyonaga *et al.* 1994, PNAS 91, 614-618). IFN-gamma también puede ser citotóxico para los hepatocitos, ya que el IFN-gamma *in vitro* induce la muerte celular en hepatocitos de ratón que fue acelerada por TNF (Morita *et al.* 1995, Hepatology 21, 1585-1593).

Se ha descrito que otras moléculas son protectoras en el modelo de ConA. Una única administración de rIL-6 inhibió completamente la liberación de transaminasas (Mizuhara *et al.* 1994, J. Exp. Med. 179, 1529-1537).

4.3 Electrotransferencia de cADN en fibras musculares para conseguir la expresión sistémica de una proteína de interés

Entre las técnicas no virales para la transferencia génica *in vivo*, la inyección directa de ADN plasmídico en el músculo y la electroporación subsiguiente es simple, barata y segura. La naturaleza post-mitótica y la longevidad de las miofibras permite la expresión estable de los genes transfectados, aunque el ADN transfectado no experimenta normalmente la integración cromosómica (Somari *et al.* 2000, Molecular Therapy 2, 178). Varios informes han demostrado que se puede conseguir la secreción de las proteínas producidas por el músculo al torrente sanguíneo tras la electroporación de los cADNs correspondientes (Rizzuto *et al.* PNAS, 1996, 6417; Aihara H *et al.*, 1998, Nature Biotech 16, 867). Además, se ha demostrado la eficacia *in vivo* de Epo e IL-18BP expresados en músculo en modelos de enfermedades (Rizzuto, 2000, Human Gene Therapy 41, 1891; Mallat, 2001, Circulation research 89, 41).

4.4 Materiales y Métodos

4.4.1 Animales

En todos los estudios se usaron ratones C57/BL6 macho (8 semanas de edad). En general, se usaron 10 animales por grupo experimental. Los ratones se mantuvieron en condiciones estándar en un ciclo de luz-oscuridad de 12 horas, se les proporcionaron alimentos y agua irradiados a voluntad.

4.4.2 Electrotransferencia Muscular

4.4.2.1 Elección del vector

Los genes de IL6 e INSP005b (IPAAA78836-2) marcados con His o StrepII se clonaron en el vector pDEST12.2 compatible con Gateway que contenía el promotor de CMV.

4.4.2.2 Protocolo de Electroporación

Los ratones se anestesiaron con gas (isoflurano, Baxter, Ref: ZDG9623). Las extremidades posteriores se afeitaron, y se aplicó un gel ecográfico. Se inyectó hialuronidasa en el músculo tibialis posterior (20 U en 50 µl de NaCl estéril del 0,9%, Sigma, Ref. H3631). Después de 10 min, se inyectaron 100 µg de plásmido (50 µg por pata en 25 µl de NaCl estéril del 0,9%) en el mismo músculo. El ADN se preparó en el tampón de PBS-L-Glutamato (6 mg/ml; L-Glutamato, Sigma P4761) antes de la inyección intramuscular. Para la electrotransferencia, se aplicó el campo eléctrico para cada pata con el ElectroSquarePorator BTX, ref ECM830 a 75 voltios durante 20 ms para cada pulso, 10 pulsos con un intervalo de 1 segundo de una manera unipolar con 2 electrodos cilíndricos (de un tamaño de 0,5 mm de diámetro).

4.4.3 El Modelo de ConA

4.4.3.1 Inyección i.v. de ConA y toma de muestras de sangre

Se adquirieron ratones hembra C57/B16 de 8 semanas de edad de IFFA CREDO. Se inyectó ConA (Sigma, ref. C7275) a 18 mg/kg i.v. y se tomaron muestras de sangre a las 1,30 y 8 horas tras la inyección. En el momento del sacrificio, se tomó sangre del corazón.

4.4.3.2 Detección de citocinas y transaminasas en las muestras de sangre

Se midieron los niveles de las citocinas IL2, IL5, IL4, TNF-alfa e IFN-gamma mediante el uso de del ensayo de CBA de TH1/TH2. Se detectó TNF-alfa, IL-6, MCP1, IFN-alfa, IL-10 e IL-12 mediante el uso del ensayo de CBA de inflamación. Los parámetros sanguíneos de las transaminasas se determinaron mediante el uso del instrumento COBAS (Hitachi).

4.4.3.3 Electrotransferencia de INSP0005b e IL-6

En el día 0 se llevó a cabo la electrotransferencia de pDEST12.2.INSP005b, pDEST12.2-hIL-6, así como de un control con un vector vacío (véase anteriormente el protocolo de electrotransferencia). En el día 5 tras la electrotransferencia, se inyectó i.v. ConA (18 mg/kg) y se tomaron muestras de sangre en 2 momentos (1,30, 8 horas). Se llevaron a cabo las medidas de citocinas y de ASAT y ALAT como se describió anteriormente.

4.5 Resultados

Se ha descubierto que INSP005b protege de la lesión hepática en un modelo de ratón que imita la hepatitis fulminante tras la administración sistémica de la proteína mediante el uso de electrotransferencia. Las Figuras 15A y 15B demuestran que los animales a los que se electrotransfirió INSP005b muestran una disminución de los niveles de transaminasas en comparación con los animales con un control de vector vacío 8 horas después de la exposición a ConA. Además, los niveles de las citocinas TNF-alfa e IL-6 se reducen significativamente en estos animales (Figuras 16A y 16B). El efecto es similar al obtenido con el vector de control positivo pDEST12.2hIL-6-SII.

4.6 Conclusión

Estos resultados demuestran que la administración de cADN de INSP005b en un modelo *in vivo* de hepatitis fulminante disminuye los niveles de TNF-alfa y m-IL-6 en el suero, y tuvo un efecto significativo sobre la reducción de los niveles de ASAT y ALAT medidos en el suero.

La disminución de los niveles de ASAT y ALAT se podría deber a los niveles disminuidos de TNF-alfa e IL-6. TNF-alfa es una citocina importante implicada en la lesión hepática tras inyección de ConA. En este modelo de ratón de hepatitis, se produce TNF-alfa principalmente por los macrófagos hepáticos, las denominadas células de Kupfer. Los anticuerpos anti-TNF-alfa confieren protección contra la enfermedad (Seino *et al.* 2001, *Annals of surgery* 234, 681).

Información de las Secuencias:

SEQ ID N°: 1 (exón 1 de la secuencia nucleotídica de INSP005A)

1 ATGGGTGGTA GTGGTGTCTGTT GGAGGTCCCC TTCCTGCTCT CCAGCAAGTA
51 CG

5

SEQ ID N°: 2 (exón 1 de la secuencia proteica de INSP005A)

1 MGGSGVVEVP FLLSSKYD

SEQ ID N°: 3 (exón 2 de la secuencia nucleotídica de INSP005A)

1 ATGAGCCCAG CCGCCAGGTC ATCCTGGAGG CTCTTGCCGA GTTTGAACGT
51 TCCACGTGCA TCAGGTTTGT CACCTATCAG GACCAGAGAG ACTTCATTTT
101 CATCATCCCC ATGTATGG

10

SEQ ID N°: 4 (exón 2 de la secuencia proteica de INSP005A)

1 EPSRQVILEA LAEFERSTCI RFVTYQDQRD FISIIPMYG

15

SEQ ID N°: 5 (exón 3 de la secuencia nucleotídica de INSP005A)

1 GTGCTTCTCG AGTGTGGGGC GCAGTGGAGG GATGCAGGTG GTCTCCCTGG
51 CGCCACGTG TCTCCAGAAG GGCCGGGGCA TTGTCTTCA TGAGCTCATG
101 CATGTGCTGG GCTTCTGGCA CGAGCACACG CGGGCCGACC GGGACCGCTA
151 TATCCGTGTC AACTGGAACG AGATCCTGCC AG

SEQ ID N°: 6 (exón 3 de la secuencia proteica de INSP005A)

1 CFSSVGRSGG MQVVSLAPTC LQKGRGIVLH ELMHVLGFWH EHTRADRDRY
51 IRVNWNEILP G

20

SEQ ID N°: 7 (exón 4 de la secuencia nucleotídica de INSP005A)

1 GCTTTGAAAT CAACTTCATC AAGTCTCAGA GCAGCAACAT GCTGACGCCC
51 TATGACTACT CCTCTGTGAT GCACTATGGG AG

SEQ ID N°: 8 (exón 4 de la secuencia proteica de INSP005A)

1 FEINFIKSQS SNMLTPYDYS SVMHYGR

25

SEQ ID N°: 9 (exón 5 de la secuencia nucleotídica de INSP005A)

1 GCTCGCCTTC AGCCGGCGTG GGCTGCCAC CATCACACCA CTTTGGGCCC
51 CCAGTGCCA CATCGGCCAG CGATGGAACC TGAGTGCCCTC GGACATCACC
101 CGGGTCTCA AACTCTACGG CTGCAGCCA AGTGGCCCA GGCCCCGTGG
151 GAGAG

30

SEQ ID N°: 10 (exón 5 de la secuencia proteica de INSP005A)

1 LAFSRRGLPT ITPLWAPSVH IGQRWNLSAS DITRVLKLYG CSPSGPRPRG
51 RG

SEQ ID N°: 11 (exón 6 de la secuencia nucleotídica de INSP005A)

1 GGTCCCATGC CCACAGCACT GGTAGGAGCC CCGCCCCGGC CTCCCTATCT
 51 CTGCAGCGGC TTTTGGAGGC ACTGTCGGCG GAATCCAGGA GCCCCGACCC
 101 CAGTGGTTCC AGTGCGGGAG GCCAGCCCGT TCCTGCAGGG CCTGGGGAGA
 151 GCCACATGG GTGGGAGTCC CCTGCCCTGA AAAAGCTCAG TGCAGAGGCC
 201 TCGGCAAGGC AGCCTCAGAC CCTAGCTTCC TCCCCAAGAT CAAGGCCTGG
 251 AGCAGGTGCC CCCGGTGTG CTCAGGAGCA GTCCTGGCTG GCCGGAGTGT
 301 CCACCAAGCC CACAGTCCCA TCTTCAGAAG CAGGAATCCA GCCAGTCCCT
 351 GTCCAGGGAA GCCCAGCTCT GCCAGGGGGC TGTGTACCTA GAAATCATTT
 401 CAAGGGGATG TCCGAAGAT

SEQ ID N°: 12 (exón 6 de la secuencia proteica de INSP005A)

1 SHAHSTGRSP APASLSLQRL LEALSAESRS PDPSSGSSAGG QPVPAGPGES
 51 PHGWESPALK KLSAEASARQ PQTCLASSPRS RPGAGAPGVA QEQSWLAGVS
 101 TKPTVPSSEA GIQPVVQGS PALPGGCVPR NHFKGMSD

5

SEQ ID N°: 13 (secuencia nucleotídica completa de INSP005A)

1 ATGGGTGGTA GTGGTGTCTG GGAGGTCCCC TTCCTGCTCT CCAGCAAGTA
 51 CGATGAGCCC AGCCGCCAGG TCATCCTGGA GGCTCTTGCG GAGTTTGAAC
 101 GTTCCACGTG CATCAGGTTT GTCACCTATC AGGACCAGAG AGACTTCATT
 151 TCCATCATCC CCATGTATGG GTGCTTCTCG AGTGTGGGGC GCAGTGGAGG
 201 GATGCAGGTG GTCTCCCTGG CGCCACGTG TCTCCAGAAG GGCCGGGGCA
 251 TTGTCCTTCA TGAGCTCATG CATGTGCTGG GCTTCTGGCA CGAGCACACG
 301 CGGGCCGACC GGGACCGCTA TATCCGTGTC AACTGGAACG AGATCCTGCC
 351 AGGCTTTGAA ATCAACTTCA TCAAGTCTCA GAGCAGCAAC ATGCTGACGC
 401 CCTATGACTA CTCCTCTGTG ATGCACTATG GGAGGCTCGC CTTAGCCGG
 451 CGTGGGCTGC CCACCATCAC ACCACTTTGG GCCCCAGTG TCCACATCGG
 501 CCAGCGATGG AACCTGAGTG CCTCGGACAT CACCCGGGTC CTCAAACTCT
 551 ACGGCTGCAG CCCAAGTGGC CCCAGGCCCC GTGGGAGAGG GTCCCATGCC
 601 CACAGCACTG GTAGGAGCCC CGCCCCGGCC TCCCTATCTC TGCAGCGGCT
 651 TTTGGAGGCA CTGTCGGCGG AATCCAGGAG CCCCAGCCCC AGTGGTTCCA
 701 GTGCGGGAGG CCAGCCCGTT CCTGCAGGGC CTGGGGAGAG CCCACATGGG
 751 TGGGAGTCCC CTGCCCTGAA AAAGCTCAGT GCAGAGGCCT CGGCAAGGCA
 801 GCCTCAGACC CTAGCTTCCT CCCCAGATC AAGGCCTGGA GCAGGTGCCC
 851 CCGGTGTTGC TCAGGAGCAG TCCTGGCTGG CCGGAGTGTC CACCAAGCCC
 901 ACAGTCCCAT CTTCAGAAGC AGGAATCCAG CCAGTCCCTG TCCAGGGAAG
 951 CCCAGCTCTG CCAGGGGGCT GTGTACCTAG AAATCATTTT AAGGGGATGT
 1001 CCGAAGAT

10 **SEQ ID N°: 14 (secuencia proteica completa de INSP005A)**

1 MGGSGVVEVP FLLSSKYDEP SRQVILEALA EFERSTCIRF VTYQDQRDFI
 51 SIIPMYGCFS SVGRSGGMQV VSLAPTCLQK GRGIVLHELM HVLGFWHEHT
 101 RADRDYIRV NWNELPGFE INFIKSQSSN MLTPYDYSSV MHYGRLAFSR
 151 RGLPTITPLW APSVHIGQRW NLSASDITRV LKLYGCSPSG PRPRGRGSHA
 201 HSTGRSPAPA SLSLQRLLEA LSAESRSPDP SGSSAGGQPV PAGPGESPHG
 251 WESPALKKLS AEASARQPQT LASSPRSRPG AGAPGVAQEQ SWLAGVSTKP
 301 TVPSSEAGIQ PVPVQGSPAL PGGCVPRNHF KGMSD

SEQ ID N°: 15 (exón 1 de la secuencia nucleotídica de INSP005B)

1 ATGGAGGGTG TAGGGGGTCT CTGGCCTTGG GTGCTGGGTC TGCTCTCCTT
 51 GCCAG

5

SEQ ID N°: 16 (exón 1 de la secuencia proteica de INSP005B)

1 MEGVGGLWPW VLGLLSLPG

SEQ ID N°: 17 (exón 2 de la secuencia nucleotídica de INSP005B)

1 GTGTGATCCT AGGAGCGCCC CTGGCCTCCA GCTGCGCAGG AGCCTGTGGT
 51 ACCAGCTTCC CAGATGGCCT CACCCCTGAG GGAACCCAGG CCTCCGGGGA
 101 CAAGGACATT CCTGCAATTA ACCAAG

10

SEQ ID N°: 18 (exón 2 de la secuencia proteica de INSP005B)

1 VILGAPLASS CAGACGTSFP DGLTPEGTQA SGDKDIPAIN QG

15

SEQ ID N°: 19 (exón 3 de la secuencia nucleotídica de INSP005B)

1 GGCTCATCCT GGAAGAAACC CCAGAGAGCA GCTTCTCAT CGAGGGGGAC
 51 ATCATCCGGC CG

SEQ ID N°: 20 (exón 3 de la secuencia proteica de INSP005B)

1 LILEETPESS FLIEGDIIRP

20

SEQ ID N°: 21 (exón 4 de la secuencia nucleotídica de INSP005B)

1 AGTCCCTTCC GACTGCTGTC AGCAACCAGC AACAAATGGC CCATGGGTGG
 51 TAGTGGTGTC GTGGAGGTCC CCTTCTGCT CTCCAGCAAG TACG

25

SEQ ID N°: 22 (exón 4 de la secuencia proteica de INSP005B)

1 SPFRLLSATS NKWPMGGSGV VEVFPLLSSK YD

SEQ ID N°: 23 (exón 5 de la secuencia nucleotídica de INSP005B)

1 ATGAGCCAG CCGCCAGGTC ATCCTGGAGG CTCTTGCGGA GTTTGAACGT
 51 TCCACGTGCA TCAGTTTGT CACCTATCAG GACCAGAGAG ACTTCATTTT
 101 CATCATCCCC ATGTATGG

30

SEQ ID N°: 24 (exón 5 de la secuencia proteica de INSP005B)

1 EPSRQVILEA LAEFERSTCI RFVTYQDQRD FISIIPMYG

SEQ ID N°: 25 (exón 6 de la secuencia nucleotídica de INSP005B)

1 GTGCTTCTCG AGTGTGGGGC GCAGTGGAGG GATGCAGGTG GTCTCCCTGG
 51 CGCCCACGTG TCTCCAGAAG GGCCGGGGCA TTGTCTTCA TGAGCTCATG
 101 CATGTGCTGG GCTTCTGGCA CGAGCACACG CGGGCCGACC GGGACCGCTA
 151 TATCCGTGTC AACTGGAACG AGATCCTGCC AG

SEQ ID N°: 26 (exón 6 de la secuencia proteica de INSP005B)

1 CFSSVGRSGG MQVVS LAPTC LQKGRGIVLH ELMHVLGFWH EHTRADRDRY
 51 IRVNWNEILP G

5

SEQ ID N°: 27 (exón 7 de la secuencia nucleotídica de INSP005B)

1 GCTTTGAAAT CAACTTCATC AAGTCTCGGA GCAGCAACAT GCTGACGCCC
 51 TATGACTACT CCTCTGTGAT GCACTATGGG AG

SEQ ID N°: 28 (exón 7 de la secuencia proteica de INSP005B)

1 FEINFIKRSR SNMLTPYDYS SVMHYGR

10

SEQ ID N°: 29 (exón 8 de la secuencia nucleotídica de INSP005B)

1 GCTCGCCTTC AGCCGGCGTG GGCTGCCCAC CATCACACCA CTTTGGGCCC
 51 CCAGTGTCCA CATCGGCCAG CGATGGAACC TGAGTGCCTC GGACATCACC
 101 CGGGTCTCA AACTCTACGG CTGCAGCCCA AGTGGCCCCA GGCCCCGTGG
 151 GAGAG

15

SEQ ID N°: 30 (exón 8 de la secuencia proteica de INSP005B)

1 LAFSRRGLPT ITPLWAPSVH IGQRWNLSAS DITRVLKLYG CSPSGPRPRG
 51 RG

SEQ ID N°: 31 (exón 9 de la secuencia nucleotídica de INSP005B)

1 GGTC CATGC CCACAGCACT GGTAGGAGCC CCGCTCCGGC CTCCCTATCT
 51 CTGCAGCGGC TTTTGGAGGC ACTGTCGGCG GAATCCAGGA GCCCCGACCC
 101 CAGTGGTTCC AGTGC GGAG GCCAGCCCGT TCCTGCAGGG CCTGGGGAGA
 151 GCCACATGG GTGGGAGTCC CCTGCCCTGA AAAAGCTCAG TGCAGAGGCC
 201 TCGGCAAGGC AGCCTCAGAC CCTAGCTTCC TCCCAAGAT CAAGGCCTGG
 251 AGCAGGTGCC CCCGGTGTG CTCAGGAGCA GTCCTGGCTG GCCGGAGTGT
 301 CCACCAAGCC CACAGTCCCA TCTTCAGAAG CAGGAATCCA GCCAGTCCCT
 351 GTCCAGGGAA GCCCAGCTCT GCCAGGGGGC TGTGTACCTA GAAATCATTT
 401 CAAGGGGATG TCCGAAGAT

20

SEQ ID N°: 32 (exón 9 de la secuencia proteica de INSP005B)

1 SHAHSTGRSP APASLSLQRL LEALSAESRS PDPSGSSAGG QPVPAGPGES
 51 PHGWESPALK KLSAEASARQ PQT LASSPRS RPGAGAPGVA QEQSWLAGVS
 101 TKPTVPSSEA GIQVPVQGS PALPGGCVPR NHFKGMSED

SEQ ID N°: 33 (secuencia nucleotídica completa de INSP005B)

1 ATGGAGGGTG TAGGGGGTCT CTGGCCTTGG GTGCTGGGTC TGCTCTCCTT
 51 GCCAGGTGTG ATCCTAGGAG CGCCCTGGC CTCCAGCTGC GCAGGAGCCT
 101 GTGGTACCAG CTTCCCAGAT GGCCTCACCC CTGAGGGAAC CCAGGCCTCC
 151 GGGGACAAGG ACATTCCCTGC AATTAACCAA GGGCTCATCC TGGAAGAAAC
 201 CCCAGAGAGC AGCTTCCTCA TCGAGGGGGA CATCATCCGG CCGAGTCCCT
 251 TCCGACTGCT GTCAGCAACC AGCAACAAAT GGCCCATGGG TGGTAGTGGT
 301 GTCGTGGAGG TCCCCTTCCT GCTCTCCAGC AAGTACGATG AGCCAGCCG
 351 CCAGGTCATC CTGGAGGCTC TTGCGGAGTT TGAACGTTCC ACGTGCATCA
 401 GTTTGTGTCAC CTATCAGGAC CAGAGAGACT TCATTTCCAT CATCCCCATG
 451 TATGGGTGCT TCTCGAGTGT GGGGCGCAGT GGAGGGATGC AGGTGGTCTC
 501 CCTGGCGCCC ACGTGTCTCC AGAAGGGCCG GGCATTGTC CTTTCATGAGC
 551 TCATGCATGT GCTGGGCTTC TGGCACGAGC ACACGCGGGC CGACCGGGAC
 601 CGCTATATCC GTGTCAACTG GAACGAGATC CTGCCAGGCT TTGAAATCAA
 651 CTTTCATCAAG TCTCGGAGCA GCAACATGCT GACGCCCTAT GACTACTCCT
 701 CTGTGATGCA CTATGGGAGG CTCGCCCTCA GCCGGCGTGG GCTGCCACC
 751 ATCACACCAC TTTGGGCCCC CAGTGTCCAC ATCGGCCAGC GATGGAACCT
 801 GAGTGCCCTCG GACATCACCC GGGTCCTCAA ACTCTACGGC TGCAGCCCAA
 851 GTGGCCCCAG GCCCCGTGGG AGAGGGTCCC ATGCCACAG CACTGGTAGG
 901 AGCCCCGCTC CGGCCTCCCT ATCTCTGCAG CGGCTTTTGG AGGCACGTGC
 951 GGCGBAATCC AGGAGCCCCG ACCCCAGTGG TTCCAGTGCG GGAGGCCAGC
 1001 CCGTTCCTGC AGGGCCTGGG GAGAGCCCAC ATGGGTGGGA GTCCCCTGCC
 1051 CTGAAAAGC TCAGTGCAGA GGCCTCGGCA AGGCAGCCTC AGACCCTAGC
 1101 TTCCTCCCCA AGATCAAGGC CTGGAGCAGG TGCCCCGGT GTTGCTCAGG
 1151 AGCAGTCTCG GCTGGCCGGA GTGTCCACCA AGCCACAGT CCCATCTTCA
 1201 GAAGCAGGAA TCCAGCCAGT CCCTGTCCAG GGAAGCCAG CTCTGCCAGG
 1251 GGGCTGTGTA CCTAGAAATC ATTTCAAGGG GATGTCCGAA GAT

SEQ ID Nº: 34 (secuencia proteica completa de INSP005B)

1 MEGVGGLWPW VLGLLSLPGV ILGAPLASSC AGACGTSFPD GLTPEGTQAS
 51 GDKDIPAINQ GLILEETPES SFLIEGDIIR PSPFRLLSAT SNKWPMMGGSG
 101 VVEVPFLLSS KYDEPSRQVI LEALAEFERS TCIRFVTYQD QRDFISIIPM
 151 YGCFSSVGRS GGMQVVS LAP TCLQKGRGIV LHELMHVLGF WHEHTRDRD
 201 RYIRVNWNEI LPGFEINFIK SRSSNMLTPY DYSSVMHYGR LAFSRRGLPT
 251 ITPLWAPSVH IGQRWNLSAS DITRVLKLYG CSPSGPRPRG RGS SHAHSTGR
 301 SPAPASLSLQ RLEALS AES RSPDPSGSSA GGQVPVAGPG ESPHWESPA
 351 LKKLSAEASA RQPQTLASSP RSRPGAGAPG VAQEQSWLAG VSTKPTVPSS
 401 EAGIQPVPVQ GSPALPGGCV PRNHFKGMSE D

SEQ ID N°: 35 (secuencia nucleotídica madura de INSP005b)

1 GCGCCCCTGG CCTCCAGCTG CGCAGGAGCC TGTGGTACCA GCTTCCCAGA
 51 TGGCCTCACC CCTGAGGGAA CCCAGGCCTC CGGGGACAAG GACATTCCCTG
 101 CAATTAACCA AGGGCTCATC CTGGAAGAAA CCCCAGAGAG CAGCTTCCCTC
 151 ATCGAGGGGG ACATCATCCG GCCGAGTCCC TTCCGACTGC TGTCAGCAAC
 201 CAGCAACAAA TGGCCCATGG GTGGTAGTGG TGTCGTGGAG GTCCCCTTCC
 251 TGCTCTCCAG CAAGTACGAT GAGCCCAGCC GCCAGGTCAT CCTGGAGGCT
 301 CTTGCGGAGT TTGAACGTTT CACGTGCATC AGGTTTGTC A CCTATCAGGA
 351 CCAGAGAGAC TTCATTTCCA TCATCCCCAT GTATGGGTGC TTCTCGAGTG
 401 TGGGGCGCAG TGGAGGGATG CAGGTGGTCT CCCTGGCGCC CACGTGTCTC
 451 CAGAAGGGCC GGGGCATTGT CCTTCATGAG CTCATGCATG TGCTGGGCTT
 501 CTGGCACGAG CACACGCGGG CCGACCGGGA CCGCTATATC CGTGTCAACT
 551 GGAACGAGAT CCTGCCAGGC TTTGAAATCA ACTTCATCAA GTCTCGGAGC
 601 AGCAACATGC TGACGCCCTA TGACTACTCC TCTGTGATGC ACTATGGGAG
 651 GCTCGCCTTC AGCCGGCGTG GGCTGCCCAC CATCACACCA CTTTGGGCCC
 701 CCAGTGTCCA CATCGGCCAG CGATGGAACC TGAGTGCCTC GGACATCACC
 751 CGGGTCCTCA AACTCTACGG CTGCAGCCCA AGTGGCCCCA GGCCCCGTGG
 801 GAGAGGGTCC CATGCCACA GCACTGGTAG GAGCCCCGCT CCGGCCTCCC
 851 TATCTCTGCA GCGGCTTTTG GAGGCACTGT CGGCGGAATC CAGGAGCCCC
 901 GACCCAGTG GTTCCAGTGC GGGAGGCCAG CCCGTTCCCTG CAGGGCCTGG
 951 GGAGAGCCCA CATGGGTGGG AGTCCCCTGC CCTGAAAAAG CTCAGTGCAG
 1001 AGGCCTCGGC AAGGCAGCCT CAGACCCTAG CTTCCCTCCC AAGATCAAGG
 1051 CCTGGAGCAG GTGCCCCCGG TGTTGCTCAG GAGCAGTCCT GGCTGGCCGG
 1101 AGTGTCCACC AAGCCCACAG TCCCATCTTC AGAAGCAGGA ATCCAGCCAG
 1151 TCCCTGTCCA GGAAGCCCA GCTCTGCCAG GGGGCTGTGT ACCTAGAAAT
 1201 CATTTCAAGG GGATGTCCGA AGAT

5

SEQ ID N°: 36 (secuencia polipeptídica madura de INSP005b)

1 APLASSCAGA CGTSFPDGLT PEGTQASGDK DIPAINQGLI LEETPESSFL
 51 IEGDIIRPSP FRLLSATSNK WPMGGSGVVE VPFLSSKYD EPSRQVILEA
 101 LAEFERSTCI RFVTYQDQRD FISIIPMYGC FSSVGRSGGM QVVSLAPTCL
 151 QKGRGIVLHE LMHVLGFWHE HTRADRDRYI RVNWNELPG FEINFIKSRS
 201 SNMLTPYDYS SVMHYGRLAF SRRGLPTITP LWAPSVHIGQ RWNLSASDIT
 251 RVLKLYGCSP SGPRPRGRGS HAHSTGRSPA PASLSLQRL EALSAESRSP
 301 DPSGSSAGGQ PVPAGPGESP HGWESPALKK LSAEASARQP QTLASSPRSR
 351 PGAGAPGVAQ EQSWLAGVST KPTVPSSEAG IQVPVQGSF ALPGGCVPRN
 401 HFKGMSED

SEQ ID N°: 37 (secuencia polipeptídica predicha de INSP005)

ES 2 374 392 T3

1 MLRLWDFNPG GALSDLALGL RMEEGGYSC AGACGTSFPD GLTPEGTQAS GDKDIPAINQ
 61 GLILEETPES SFLIEGDIIR PSPFRLLSAT SNKWPMGGSG VVEVPFLLSS KYDEPSHQVI
 121 LEALAEFERS TCIRFVTYQD QRDFISIIPM YGCFSSVGRS GGMQVVSLAP TCLQKGRGIV
 181 LHELMHVLGF WHEHTRADRD RYIRVNWNEI LPGFEINFIK SQSSNMLTPY DYSSVMHYGR
 241 LAFSRRGLPT ITPLWAPSVH IGQRWNLSAS DITRVLKLYG CSPSGPRPRG RGEWHGRKVT

SEQ ID N°: 38 (secuencia nucleotídica del plásmido pCR4 TOPO IPAAA78836-1)

1 AGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC
 61 ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC
 121 TCACTCATTG GGCACCCAG GCTTTACTACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA
 181 TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCTCA
 241 GAATTAACCC TCACTAAAGG GACTAGTCCT GCAGGTTTAA ACGAATTCGC CCTTAGCCAC
 301 AGGCTTAATC TTCGGACATC CCCTTGAAAT GATTTCTAGG TACACAGCCC CCTGGCAGAG
 361 CTGGGCTTCC CTGGACAGGG ACTGGCTGGA TTCCTGCTTC TGAAGATGGG ACTGTGGGCT
 421 TGGTGGACAC TCCGGCCAGC CAGGACTGCT CCTGAGCAAC ACCGGGGCA CCTGCTCCAG
 481 GCCTTGATCT TGGGGAGGAA GCTAGGGTCT GAGGCTGCCT TGCCGAGGCC TCTGCACTGA
 541 GCTTTTTTCAG GGCAGGGGAC TCCCACCCAT GTGGGCTCTC CCCAGGCCCT GCAGGAACGG
 601 GCTGGCCTCC CGCACTGGAA CCACTGGGGT CGGGGCTCCT GGATTCGCC GACAGTGCCT
 661 CCAAAAGCCG CTGCAGAGAT AGGGAGGCCG GGGCGGGCT CCTACCAGTG CTGTGGGCAT

721 GGGACCTCT CCCACGGGGC CTGGGGCCAC TTGGGCTGCA GCCGTAGAGT TTGAGGACCC
 781 GGGTGATGTC CGAGGCAC TC AGGTTCATC GCTGGCCGAT GTGGACACTG GGGGCCAAA
 841 GTGGTGTGAT GGTGGGCAGC CCACGCCGGC TGAAGGCGAG CCTCCCATAG TGCATCACAG
 901 AGGAGTAGTC ATAGGGCGTC AGCATGTTGC TGCTCTGAGA CTTGATGAAG TTGATTTCAA
 961 AGCCTGGCAG GATCTCGTTC CAGTTGACAC GGATATAGCG GTCCCCGGTCG GCCCGCGTGT
 1021 GCTCGTGCCA GAAGCCCAGC ACATGCATGA GCTCATGAAG GACAATGCC CCGCCCTTCT
 1081 GGAGACACGT GGGCGCCAGG GAGACCACCT GCATCCCTCC ACTGCGCCCC ACACTCGAGA
 1141 AGCACCCATA CATGGGGATG ATGGAAATGA AGTCTCTCTG GTCCTGATAG GTGACAAACC
 1201 TGATGCACGT GGAACGTTCA AACTCCGCAA GAGCCTCCAG GATGACCTGG CCGCTGGGCT
 1261 CATCGTACTT GCTGGAGAGC AGGAAGGGGA CCTCCACGAC ACCACTACCA CCCATGGGCC
 1321 ATTTGTTGCT GGTGCTGAC AGAAGGGCGA ATTCGCGGCC GCTAAATTCA ATTCGCCCTA
 1381 TAGTGAGTCG TATTACAATT CACTGGCCGT CGTTTTACAA CGTCGTGACT GGGAAAACCC
 1441 TGGCGTTACC CAACTTAATC GCCTTGCAGC ACATCCCCCT TTCGCCAGCT GGCCTAATAG
 1501 CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTATACG TACGGCAGTT
 1561 TAAGGTTTAC ACCTATAAAA GAGAGAGCCG TTATCGTCTG TTTGTGGATG TACAGAGTGA
 1621 TATTATTGAC ACGCCGGGGC GACGGATGGT GATCCCCCTG GCCAGTGCAC GTCGTCTGTC
 1681 AGATAAAGTC TCCCGTGAAC TTTACCCGGT GGTGCATATC GGGGATGAAA GCTGGCGCAT
 1741 GATGACCACC GATATGGCCA GTGTGCCGGT CTCCGTTATC GGGGAAGAAG TGGCTGATCT
 1801 CAGCCACCGC GAAAATGACA TCAAAAACGC CATTAACTG ATGTTCTGGG GAATATAAAT
 1861 GTCAGGCATG AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTACGT AGAAAGCCAG
 1921 TCCGCAGAAA CGGTGCTGAC CCCGGATGAA TGTCAGCTAC TGGGCTATCT GGACAAGGGA
 1981 AAACGCAAGC GCAAAGAGAA AGCAGGTAGC TTGCAGTGGG CTTACATGGC GATAGCTAGA
 2041 CTGGGCGGTT TTATGGACAG CAAGCGAACC GGAATTGCCA GCTGGGGCGC CCTCTGGTAA
 2101 GGTTGGGAAG CCTGCAAAG TAAACTGGAT GGCTTTCTCG CCGCCAAGGA TCTGATGGCG
 2161 CAGGGGATCA AGCTCTGATC AAGAGACAGG ATGAGGATCG TTTGCGATGA TTGAACAAGA
 2221 TGGATTGCAC GCAGGTTCTC CGGCCGCTTG GGTGGAGAGG CTATTCGGCT ATGACTGGGC
 2281 ACAACAGACA ATCGGCTGCT CTGATGCCGC CGTGTTCGG CTGTCAGCGC AGGGGCGCCC
 2341 GGTTCTTTTT GTCAAGACCG ACCTGTCCGG TGCCCTGAAT GAACTGCAAG ACGAGGCAGC
 2401 GCGGCTATCG TGGCTGGCCA CGACGGGCGT TCCTTGCGCA GCTGTGCTCG ACGTTGTAC
 2461 TGAAGCGGGA AGGGACTGGC TGCTATTGGG CGAAGTGCCG GGGCAGGATC TCCTGTCATC
 2521 TCACCTTGCT CCTGCCGAGA AAGTATCCAT CATGGCTGAT GCAATGCGGC GGCTGCATAC
 2581 GCTTGATCCG GCTACCTGCC CATTCGACCA CCAAGCGAAA CATCGCATCG AGCGAGCAGC
 2641 TACTCGGATG GAAGCCGGTC TTGTCGATCA GGATGATCTG GACGAAGAGC ATCAGGGGCT
 2701 CGCGCCAGCC GAACTGTTCTG CCAGGCTCAA GGCGAGCATG CCCGACGGCG AGGATCTCGT
 2761 CGTGACCCAT GGCATGCCCT GCTTGCCGAA TATCATGGTG GAAAATGGCC GCTTTTCTGG
 2821 ATTCATCGAC TGTGGCCGGC TGGGTGTGGC GGACCCTAT CAGGACATAG CGTTGGCTAC
 2881 CCGTGATATT GCTGAAGAGC TTGGCGGCGA ATGGGCTGAC CGCTTCCTCG TGCTTTACGG
 2941 TATCGCCGCT CCCGATTCGC AGCGCATCGC CTTCTATCGC CTTCTTGACG AGTTCCTCTG
 3001 AATTATTAAC GCTTACAATT TCCTGATGCG GTATTTCTC CTTACGCATC TGTGCGGTAT
 3061 TTCACACCGC ATACAGGTGG CACTTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA
 3121 TTTTTCTAAA TACATCAAA TATGTATCCG CTCATGAGAC AATAACCCTG ATAAATGCTT

3181 CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT TCCGTGTCGC CCTTATTCCC
 3241 TTTTTTGCGG CATTTTGCCT TCCTGTTTTT GCTCACCCAG AAACGCTGGT GAAAGTAAAA
 3301 GATGCTGAAG ATCAGTTGGG TGCACGAGTG GGTACATCG AACTGGATCT CAACAGCGGT
 3361 AAGATCCTTG AGAGTTTTCG CCCCAGAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT
 3421 CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCCGGGC AAGAGCAACT CGGTCGCCGC
 3481 ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG TCACAGAAAA GCATCTTACG
 3541 GATGGCATGA CAGTAAGAGA ATTATGCAGT GCTGCCATAA CCATGAGTGA TAACACTGCG
 3601 GCCAACTTAC TTCTGACAAC GATCGGAGGA CCGAAGGAGC TAACCGCTTT TTTGCACAAC
 3661 ATGGGGGATC ATGTAACTCG CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA
 3721 AACGACGAGC GTGACACCAC GATGCCGTGA GCAATGGCAA CAACGTTGCG CAACTATTA
 3781 ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA TAGACTGGAT GGAGGCGGAT
 3841 AAAGTTGCAG GACCACTTCT GCGCTCGGCC CTTCGGCTG GCTGGTTTTAT TGCTGATAAA
 3901 TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT ATCATTGCAG CACTGGGGCC AGATGGTAAAG
 3961 CCCTCCCGTA TCGTAGTTAT CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT
 4021 AGACAGATCG CTGAGATAGG TGCCCTACTG ATTAAGCATT GGTAAGTGT AGACCAAGTT
 4081 TACTCATATA TACTTTAGAT TGATTTAAAA CTTCATTTTT AATTTAAAAG GATCTAGGTG
 4141 AAGATCCTTT TTGATAATCT CATGACCAAA ATCCCTTAAC GTGAGTTTTT GTTCCACTGA
 4201 GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA
 4261 ATCTGCTGCT TGCAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGT TT GCCGGATCAA
 4321 GAGCTACCAA CTCTTTTTCC GAAGGTAACT GGCTTCAGCA GAGCGCAGAT ACCAAACTACT
 4381 GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA ACTCTGTAGC ACCGCCTACA
 4441 TACCTCGCTC TGCTAATCCT GTTACCAGTG GCTGCTGCCA GTGGCGATAA GTCGTGTCTT
 4501 ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC AGCGGTCCGG CTGAACGGGG
 4561 GGTTCGTGCA CACAGCCAG CTGGAGCGA ACGACCTACA CCGAAGTGA ATACCTACAG
 4621 CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA
 4681 AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC CAGGGGAAA CGCTGGTAT
 4741 CTTTATAGTC CTGTCCGGTT TCGCCACCTC TGAAGTGTAGC GTCGATTTTT GTGATGCTCG
 4801 TCAGGGGGGC GGAGCCTATG GAAAACGCC AGCAACGGCG CCTTTTTTACG GTTCTGGGC
 4861 TTTTGCTGGC CTTTGTCTCA CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC
 4921 CGTATTACCG CCTTTGAGTG AGCTGATACC GCTCGCCGCA GCCGAACGAC CGAGCGCAGC
 4981 GAGTCAGTGA GCGAGGAAGC GGAAG

SEQ ID N°: 39 (secuencia nucleotídica del plásmido XpCR4TOPO IPAAA78836-2)

1 AGCGCCCAAT ACGCAAACCG CCTCTCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC
 61 ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC
 121 TCACTCATTA GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA
 181 TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCTCA
 241 GAATTAACCC TCACTAAAGG GACTAGTCCT GCAGGTTTAA ACGAATTCGC CCTTAGCCAC
 301 AGGCTTAATC TTCGGACATC CCCTTGAAAT GATTTCTAGG TACACAGCCC CCTGGCAGAG
 361 CTGGGCTTCC CTGGACAGGG ACTGGCTGGA TTCTTGCTTC TGAAGATGGG ACTGTGGGCT

421 TGGTGGACAC TCCGGCCAGC CAGGACTGCT CCTGAGCAAC ACCGGGGGCA CCTGCTCCAG
 481 GCCTTGATCT TGGGGAGGAA GCTAGGGTCT GAGGCTGCCT TGCCGAGGCC TCTGCACTGA
 541 GCTTTTTTCAG GGCAGGGGAC TCCCACCCAT GTGGGCTCTC CCCAGGCCCT GCAGGAACGG
 601 GCTGGCCTCC CGCACTGGAA CCACTGGGGT CGGGGCTCCT GGATTCCGCC GACAGTGCCT
 661 CCAAAAAGCCG CTGCAGAGAT AGGGAGGCCG GAGCGGGGCT CCTACCAGTG CTGTGGGCAT
 721 GGGACCCTCT CCCACGGGGC CTGGGGCCAC TTGGGCTGCA GCCGTAGAGT TTGAGGACCC
 781 GGGTGATGTC CGAGGCACTC AGGTTCCATC GCTGGCCGAT GTGGACACTG GGGGCCAAA
 841 GTGGTGTGAT GGTGGGCAGC CCACGCCGGC TGAAGGCGAG CCTCCCATAG TGCATCACAG
 901 AGGAGTAGTC ATAGGGCGTC AGCATGTTGC TGCTCCGAGA CTTGATGAAG TTGATTTCAA
 961 AGCCTGGCAG GATCTCGTTC CAGTTGACAC GGATATAGCG GTCCCAGTCC GCCCAGTGT
 1021 GCTCGTGCCA GAAGCCCAGC ACATGCATGA GTCATGAAG GACAATGCC CGGCCCTTCT
 1081 GGAGACACGT GGGGCCCAGG GAGACCACCT GCATCCCTCC ACTGCGCCCC ACACTCGAGA
 1141 AGCACCCATA CATGGGGATG ATGGAAATGA AGTCTCTCTG GTCCGTGATAG GTGACAAAAC
 1201 TGATGCACGT GGAACGTTCA AACTCCGCAA GAGCCTCCAG GATGACCTGG CGGCTGGGCT
 1261 CATCGTACTT GCTGGAGAGC AGGAAGGGGA CCTCCACGAC ACCACTACCA CCCATGGGCC
 1321 ATTTGTTGCT GGTGCTGAC AGCAGTCGGA AGGGACTCGG CCGGATGATG TCCCCCTCGA
 1381 TGAGGAAGCT GCTCTCTGGG GTTTCTTCCA GGATGAGCCC TTGGTTAATT GCAGGAATGT
 1441 CCTTGTCCCC GGAGGCCTGG GTTCCCTCAG GGGTGAGGCC ATCTGGGAAG CTGGTACCAC
 1501 AGGCTCCTGC GCAGCTGGAG GCCAGGGGCG CTCCTAGGAT CACACCTGGC AAGGAGAGCA
 1561 GACCCAGCAC CCAAGGCCAG AGACCCCTA CACCCTCCAT GGTAGAAAGG GCGAATTCGC
 1621 GGCCGCTAAA TTCAATTCGC CCTATAGTGA GTCGTATTAC AATTCACTGG CCGTCTTTT
 1681 ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC
 1741 CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT
 1801 GCGCAGCCTA TACGTACGGC AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCCGTTATCG
 1861 TCTGTTTGTG GATGTACAGA GTGATATTAT TGACACGCCG GGGCGACGGA TGGTGATCCC
 1921 CCTGGCCAGT GCACGTCTGC TGTCAGATAA AGTCTCCCGT GAACTTTACC CGGTGGTGCA
 1981 TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT
 2041 TATCGGGGAA GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA ACGCCATTA
 2101 CCTGATGTTT TGGGGAATAT AAATGTCAGG CATGAGATTA TCAAAAAGGA TCTTCACCTA
 2161 GATCCTTTTC ACGTAGAAAG CCAGTCCGCA GAAACGGTGC TGACCCCGGA TGAATGTCAG
 2221 CTACTGGGCT ATCTGGACAA GGGAAAACGC AAGCGCAAAG AGAAAGCAGG TAGCTTGCAG
 2281 TGGGCTTACA TGGCGATAGC TAGACTGGGC GGTTTTATGG ACAGCAAGCG AACCGGAATT
 2341 GCCAGCTGGG GCGCCCTCTG GTAAGGTTGG GAAGCCCTGC AAAGTAACT GGATGGCTTT
 2401 CTCGCCGCCA AGGATCTGAT GGCAGAGGGG ATCAAGCTCT GATCAAGAGA CAGGATGAGG
 2461 ATCGTTTCGC ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG CTTGGGTGGA
 2521 GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC TGCTCTGATG CCGCCGTGTT
 2581 CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT TTTTGTCAAG ACCGACCTGT CCGGTGCCCT
 2641 GAATGAAC TG CAAGACGAGG CAGCGCGGCT ATCGTGGCTG GCCACGACGG GCGTTCCTTG
 2701 CGCAGCTGTG CTCGACGTTG TCACTGAAGC GGAAGGGGAC TGGCTGCTAT TGGCGAAGT
 2761 GCCGGGGCAG GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT CCATCATGGC
 2821 TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC TGCCCATTCG ACCACCAAGC

ES 2 374 392 T3

2881 GAAACATCGC ATCGAGCGAG CACGTACTCG GATGGAAGCC GGTCTTGTCG ATCAGGATGA
 2941 TCTGGACGAA GAGCATCAGG GGCTCGCGCC AGCCGAACTG TTCGCCAGGC TCAAGGCGAG
 3001 CATGCCCGAC GGCGAGGATC TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT
 3061 GGTGGAAAAT GGCCGCTTTT CTGGATTTCAT CACTGTGGC CGGCTGGGTG TGGCGACCG
 3121 CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA GAGCTTGGCG GCGAATGGGC
 3181 TGACCGCTTC CTCGTGCTTT ACGGTATCGC CGCTCCCGAT TCGCAGCGCA TCGCCTTCTA
 3241 TCGCCTTCTT GACGAGTTCT TCTGAATTAT TAACGCTTAC AATTTCTGA TCGGTATTT
 3301 TCTCCTTACG CATCTGTGCG GTATTTTACA CCGCATAACAG GTGGCACTTT TCGGGGAAAT
 3361 GTGCGCGGAA CCCCATTATT TTTATTTTTC TAAATACATT CAAATATGTA TCCGCTCATG
 3421 AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA
 3481 CATTTCCGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCGTG TTTTGCTCAC
 3541 CCAGAAACGC TGGTAAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTAC
 3601 ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT
 3661 CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC
 3721 GGGCAAGAGC AACTCGGTGC CCGCATAAC TATCTCAGA ATGACTTGGT TGAGTACTCA
 3781 CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGTGCC
 3841 ATAACCATGA GTGATAACAC TGCGGCCAAC TTAATTCTGA CAACGATCGG AGGACCGAAG
 3901 GAGCTAACCG CTTTTTGC CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA
 3961 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG
 4021 GCAACAACGT TGCGCAAACT ATTAAC TGCC GAACTACTTA CTCTAGCTTC CCGCAACAA
 4081 TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTCCG
 4141 GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT
 4201 GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT
 4261 CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCCTC ACTGATTAAG
 4321 CATTTGGTAACT TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATTT AAACTTCAT
 4381 TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT
 4441 TAACGTGAGT TTTCTGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT
 4501 TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAA CAAAAAACC ACCGCTACCA
 4561 GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC
 4621 AGCAGAGCGC AGATACCAA TACTGTCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC
 4681 AAGAACTCTG TAGCACC GCC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT
 4741 GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG
 4801 GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGA GCGAACGACC
 4861 TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT TCCCGAAGGG
 4921 AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG
 4981 CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCTGTGCG GGTTCGCCA CCTCTGACTT
 5041 GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAA CGCCAGCAAC
 5101 GCGGCCTTTT TACGGTTCCT GGGCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCTGCG
 5161 TTATCCCTG ATCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC
 5221 CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAG

REIVINDICACIONES

1. Un polipéptido que tiene actividad de metaloproteasa que comprende o consiste en la secuencia de aminoácidos indicada en SEQ ID N°: 14, SEQ ID N°:34 o SEQ ID N°:36.
2. Una proteína de fusión que comprende un polipéptido según la reivindicación 1.
3. La proteína de fusión de la reivindicación 2, en la que dicha proteína comprende un marcador de histidina.
4. Una molécula de ácido nucleico purificada que codifica un polipéptido según cualquiera de las reivindicaciones precedentes.
5. Una molécula de ácido nucleico purificada según la reivindicación 4, que comprende o consiste en la secuencia de ácido nucleico indicada en SEQ ID N°: 13, SEQ ID N°:33 o SEQ ID N°:35.
6. Un vector que comprende una molécula de ácido nucleico como se indicó en la reivindicación 4 o la reivindicación 5.
7. Una célula hospedadora que comprende un vector según la reivindicación 6.
8. Un anticuerpo que se une específicamente a un polipéptido según cualquiera de las reivindicaciones 1 a 3.
9. El anticuerpo de la reivindicación 8 que inhibe la actividad de metaloproteasa de un polipéptido según cualquiera de las reivindicaciones 1 a 3.
10. Un polipéptido según cualquiera de las reivindicaciones 1 a 3, una molécula de ácido nucleico según la reivindicación 4 o la reivindicación 5, un vector según la reivindicación 6, una célula hospedadora según la reivindicación 7, o un anticuerpo según la reivindicación 8 o la reivindicación 9, para el uso de la terapia o el diagnóstico de una enfermedad, y en el que dicha enfermedad es una enfermedad autoinmunitaria, enfermedad hepática viral o aguda, que incluye insuficiencia hepática alcohólica, o una enfermedad inflamatoria.
11. Un método para diagnosticar una enfermedad en un paciente, que comprende determinar *in vitro* el nivel de expresión de un gen natural que codifica un polipéptido según cualquiera de las reivindicaciones 1 a 3, o determinar la actividad de un polipéptido según cualquiera de las reivindicaciones 1 a 3, en un tejido de dicho paciente y comparar dicho nivel de expresión o actividad con un nivel de control, en el que un nivel que es diferente de dicho nivel de control es indicativo de la enfermedad, y en el que dicha enfermedad es una enfermedad autoinmunitaria, enfermedad hepática viral o aguda, que incluye insuficiencia hepática alcohólica, o una enfermedad inflamatoria.
12. Un método según la reivindicación 11, que comprende las etapas de: (a) poner en contacto un anticuerpo según la reivindicación 8 o la reivindicación 9 con una muestra biológica en condiciones adecuadas para la formación de un complejo anticuerpo-polipéptido; y (b) detectar dicho complejo.
13. Un método según la reivindicación 11, que comprende las etapas de:
 - a. poner en contacto una muestra de tejido del paciente con una sonda de ácido nucleico en condiciones rigurosas que permiten la formación de un complejo híbrido entre una molécula de ácido nucleico según la reivindicación 4 o la reivindicación 5 y la sonda;
 - b. poner en contacto una muestra de control con dicha sonda en las mismas condiciones usadas en la etapa a); y
 - c. detectar la presencia de los complejos híbridos en dichas muestras; en el que la detección de niveles del complejo híbrido en la muestra del paciente que difieren de los niveles del complejo híbrido en la muestra de control es indicativa de la enfermedad.
14. Un método según la reivindicación 11, que comprende:
 - a. poner en contacto una muestra de ácido nucleico de tejido del paciente con un cebador de ácido nucleico en condiciones rigurosas que permiten la formación de un complejo híbrido entre una molécula de ácido nucleico según la reivindicación 4 o la reivindicación 5 y el cebador;
 - b. poner en contacto una muestra de control con dicho cebador en las mismas condiciones usadas en la etapa a); y
 - c. amplificar el ácido nucleico de la muestra; y
 - d. detectar el nivel de ácido nucleico amplificado a partir de las muestras del paciente y del control; en el que la detección de niveles del ácido nucleico amplificado en la muestra del paciente que difieren significativamente de los niveles del ácido nucleico amplificado en la muestra de control es indicativa de la enfermedad.
15. El uso de un polipéptido según cualquiera de las reivindicaciones 1 a 3 como una metaloproteasa.

16. Una composición farmacéutica que comprende un polipéptido según cualquiera de las reivindicaciones 1 a 3, una molécula de ácido nucleico según la reivindicación 4 o la reivindicación 5, un vector según la reivindicación 6, una célula hospedadora según la reivindicación 7, o un anticuerpo según la reivindicación 8 o la reivindicación 9.
- 5 17. Un polipéptido según cualquiera de las reivindicaciones 1 a 3, una molécula de ácido nucleico según la reivindicación 4 o la reivindicación 5, un vector según la reivindicación 6, una célula hospedadora según la reivindicación 7, un anticuerpo según la reivindicación 8 o la reivindicación 9, o una composición farmacéutica según la reivindicación 16, para el uso en la fabricación de un medicamento para el tratamiento de una enfermedad autoinmunitaria, enfermedad hepática viral o aguda, que incluye insuficiencia hepática alcohólica, o una enfermedad inflamatoria.
- 10 18. Un método para monitorizar el tratamiento terapéutico de una enfermedad en un paciente *ex vivo*, que comprende monitorizar a lo largo de un periodo de tiempo el nivel de expresión o actividad de un polipéptido según cualquiera de las reivindicaciones 1 a 3, o el nivel de expresión de una molécula de ácido nucleico según la reivindicación 4 o la reivindicación 5 en un tejido de dicho paciente, en el que la alteración de dicho nivel de expresión o actividad a lo largo del periodo de tiempo hacia un nivel de control es indicativa de la regresión de dicha enfermedad, y en el que dicha enfermedad es una enfermedad autoinmunitaria, enfermedad hepática viral o aguda, que incluye insuficiencia hepática alcohólica, o una enfermedad inflamatoria.
- 15 19. Un método para la identificación de un compuesto que es eficaz en el tratamiento y/o el diagnóstico de una enfermedad, que comprende poner en contacto un polipéptido según cualquiera de las reivindicaciones 1 a 3, o una molécula de ácido nucleico según la reivindicación 4 o la reivindicación 5 con uno o más compuestos que se sospecha que poseen afinidad de unión por dicho polipéptido o molécula de ácido nucleico, y seleccionar un compuesto que se une específicamente a dicha molécula de ácido nucleico o polipéptido, en el que dicha enfermedad es una enfermedad autoinmunitaria, enfermedad hepática viral o aguda, que incluye insuficiencia hepática alcohólica, o una enfermedad inflamatoria.
- 20 25

FIGURA 1

Blastp frente a NCBI-nr

```
>dbj|BAB68513.1| enzima de eclosión EHE4 [Anguilla japonica]
      Longitud = 271

Puntuación = 197 bits (502), Esperado = 1e-49
Identidades = 103/233 (44%), Positivos = 141/233 (60%), Huecos = 5/233 (2%)

Problema: 52  DKDIPAINQGLILEETPESSFLIEGDIIRPSPFRLLSATSNNK--WPMGGSGVVEVPFLLS 109
              D D I  ++      S L+EGD+I  +  +  +N+ W  G+VEVP+ +S
Objetivo: 41  DPDDVDITTSILQSNNGSSEILMEGDLIVSNTRNAMKCWNNQCLWRKSSDGLVEVPYTVS 100

Problema: 110 SKYDEPSHQVILEALAEFERSTCIRFVTVYQDQRFISIIIPMYGCFSSVGRSGGMQVVSLSLA 169
              +++   + I A+ F  TCIRFV  QRDFISI  GC+S +GR+GG QVVSLSLA
Objetivo: 101 NEFSYYHKKRIENAMKTFNTETCIRFVPRSSQRDFISIESRDCYSYLGRTGGKQVVSLSLA 160

Problema: 170 PT-CLQKGRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINFIKSQSSNMLT 228
              C+  GI+ HEL H LGF+HEHTR+DRD Y+++NW  + P  NF  ++N+ T_
Objetivo: 161 RYGCYV--HGI IQHELNHALGFYHEHTRSDDRDEYVKINWENVAPHTIYNFQTQDTNNLNT 218

Problema: 229 PYDYSSVMHYGRLAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKLYGC 281
              PYDY+S+MHYGR AFS  G+ TITP+  P+  IGQR ++S  DI R+ KLY C
Objetivo: 219 PYDYTSIMHYGRTAFSTNGMDTITPVPNPNQSIGQRRSMSGDILRIKKLYSC 271
```

Tblastn frente a NCBI-est

```
Tejido = Tumor de útero

>gb|BI061462.1|BI061462 IL3-UT0117-070301-494-H12 UT0117 cADN de Homo sapiens.
      Longitud = 652

Puntuación = 175 bits (443), Esperado = 2e-42
Identidades = 85/86 (98%), Positivos = 85/86 (98%)
Marco = -2

Problema: 29  SCAGACGTSFPDGLTPEGTQASGDKDIPAINQGLILEETPESSFLIEGDIIRPSPFRLLS 88
              SCAGACGTSFPDGLTPEGTQASGDKDIPAINQGLILEETPESSFLIEGDIIRPSPFRLLS
Objetivo: 546 SCAGACGTSFPDGLTPEGTQASGDKDIPAINQGLILEETPESSFLIEGDIIRPSPFRLLS 367

Problema: 89  ATSNKWPMGGSGVVEVPFLLSSKYDE 114
              ATSNKWPMGGSGVVEVPFLLSSKY  E
Objetivo: 366 ATSNKWPMGGSGVVEVPFLLSSKYGE 289
```

2/18

FIGURA 2

Biblioteca	Fuente tisular/celular	Vector	Cepa hospedadora	Proveedor	Nº de cat.
1	cerebro fetal humano	Zap II	XL1-Blue MRF'	Stratagene	936206
2	ovario humano	GT10	LE392	Clontech	HL1098a
3	hipófisis humana	GT10	LE392	Clontech	HL1097a
4	placenta humana	GT11	LE392	Clontech	HL1075b
5	testículo humano	GT11	LE392	Clontech	HL1010b
6	sustancia negra humana	GT10	LE392	interno	
7	cerebro fetal humano	GT10	LE392	interno	
8	corteza cerebral humana	GT10	LE392	interno	
9	colon humano	GT10	LE392	Clontech	HL1034a
10	cerebro fetal humano	GT10	LE392	Clontech	HL1065a
11	pulmón fetal humano	GT10	LE392	Clontech	HL1072a
12	riñón fetal humano	GT10	LE392	Clontech	HL1071a
13	hígado fetal humano	GT10	LE392	Clontech	HL1064a
14	médula ósea humana	GT10	LE392	Clontech	HL1058a
15	monocitos de sangre periférica humana	GT10	LE392	Clontech	HL1050a
16	placenta humana	GT10	LE392	interno	
17	SHSYSY humano	GT10	LE392	interno	
18	línea celular U373 humana	GT10	LE392	interno	
19	línea celular CFPoc-I humana	Uni Zap	XL1-Blue MRF'	Stratagene	936206
20	retina humana	GT10	LE392	Clontech	HL1132a
21	vejiga urinaria humana	GT10	LE392	interno	
22	plaquetas humanas	Uni Zap	XL1-Blue MRF'	interno	
23	neuroblastoma humano Kan + TS	GT10	LE392	interno	
24	músculo liso bronquial humano	GT10	LE392	interno	
25	músculo liso bronquial humano	GT10	LE392	interno	
26	timo humano	GT10	LE392	Clontech	HL1127a
27	tramo de 5' de bazo humano	GT11	LE392	Clontech	HL1134b
28	monocitos de sangre periférica humana	GT10	LE392	Clontech	HL1050a
29	testículo humano	GT10	LE392	Clontech	HL1065a
30	cerebro fetal humano	GT10	LE392	Clontech	HL1065a
31	sustancia negra humana	GT10	LE392	Clontech	HL1093a
32	placenta humana nº 11	GT11	LE392	Clontech	HL1075b
33	cerebro fetal humano	GT10	LE392	Clontech	por encargo
34	placenta humana nº 59	GT10	LE392	Clontech	HL5014a
35	hipófisis humana	GT10	LE392	Clontech	HL1097a
36	páncreas humano nº 63	Uni Zap XR	XL1-Blue MRF'	Stratagene	937208
37	placenta humana nº 19	GT11	LE392	Clontech	HL1008
38	tramo de 5' de hígado humano	GT11	LE392	Clontech	HL1115b
39	útero humano	Zap-CMV XR	XL1-Blue MRF'	Stratagene	980207
40	biblioteca de cADN de insertos grandes de riñón humano	TriplEx2	XL1-Blue	Clontech	HL5507u

FIGURA 3

```

1  AGGTCCTTGT GGACAATAGC TATTCTTCTT GGCTCTGTGC CTTCCCTTCA CTGGGTGCAG
61  GTGACTGTGG GGGTGTCCCC AAATGCTGCC CAGCGCTGAC ATGCTCCGCC TCTGGGATTT
                                     m l r l w d

121 CAATCCAGGT GGGGCCCTGA GTGACCTGGC TCTGGGGCTC AGGGGTATGG AGGAGGGGGG
    f n p g   g a l s d l a l g l r g m e e g

181 ATATAGCTGC GCAGGAGCCT GTGGTACCAG CTTCCAGAT GGCCTCACCC CTGAGGGAAC
    g y s c   a g a c g t s f p d g l t p e g

241 CCAGGCCTCC GGGGACAAGG ACATTCTTGC AATTAACCAA GGGCTCATCC TGAAGAAAC
    t q a s   g d k d i p a i n q g l i l e e

301 CCCAGAGAGC AGCTTCCTCA TCGAGGGGGA CATCATCCGG CCGAGTCCCT TCCGACTGCT
    t p e s   s f l i e g d i i r p s p f r l

361 GTCAGCAACC AGCAACAAAT GGCCCATGGG TGGTAGTGGT GTCGTGGAGG TCCCTTCTCT
    l s a t   s n k w p m g g s g v v e v p f

421 GCTCTCCAGC AAGTACGATG AGCCCAGCCA TCAGGTCATC CTGGAGGCTC TTGCGGAGTT
    l l s s   k y d e p s h q v i l e a l a e

481 TGAACGTTCC ACGTGCATCA GGTTTGTAC CATTACAGGAC CAGAGAGACT TCATTTCCAT
    f e r s   t c i r f v t y q d q r d f i s

541 CATCCCCATG TATGGGTGCT TCTCGAGTGT GGGGCGCAGT GGAGGGATGC AGGTGGTCTC
    i i p m   y g c f s s v g r s g g m q v v

601 CCTGGCGCCC ACGTGTCTCC AGAAGGGCCG GGGCATTGTC CTTTCATGAGC TCATGCATGT
    s l a p   t c l q k g r g i v l h e l m h
                                     CP1
661 GCTGGGCTTC TGGCAGGAGC ACACGCGGGC CGACCGGGAC CGCTATATCC GTGTCAACTG
    v l g f   w h e h t r a d r d r y i r v n
                                     ----->
721 GAACGAGATC CTGCCAGGCT TTGAAATCAA CTTTCATCAAG TCTCAGAGCA GCAACATGCT
    w n e i   l p g f e i n f i k s q s s n m

781 GACGCCCTAT GACTACTCCT CTGTGATGCA CTATGGGAGG CTCGCCTTCA GCCGGCGTGG
    l t p y   d y s s v m h y g r l a f s r r
                                     78836-GR1-3'
841 GCTGCCACCC ATCACACCAC TTTGGGCCCC CAGTGTCCAC ATCGGCCAGC GATGGAACCT
    g l p t   i t p l w a p s v h i g q r w n
                                     ----->
901 GAGTGCCTCG GACATCACCC GGGTCTCAA ACTCTACGGC TGCAGCCCAA GTGGCCCCAG
    l s a s   d i t r v l k l y g c s p s g p
    -----> 78836-GR1nest-3'  CP2
961 GCCCCGTGGG AGAGGTGAGT GGCATGGCAG GAAGGTGACT TGAACCTGGA GAAGGCGCCT
    r p r g   r g e w h g r k v t -

1021 GTGCTCTAAT GGTGTCAGGG AGGGTGACAA GGAGGGAGAT GAGGTTGCAG GGGGAGCAGG
1081 GTGAGATCAC GGGGGCTTGC CAC
    
```

Posición y sentido de los cebadores de PCR  

FIGURA 4

Cebador	Nombre	Secuencia (5'-3')
CP1	4C5	ACC GCT ATA TCC GTG TCA A
CP2	4C6	GCT GCA GCC GTA GAG TTT
GeneRacer 3'		GCT GTC AAC GAT ACG CTA CGT AAC G
78836-GR1-3'		AGT GTC CAC ATC GGC CAG CGA TGG AA
		CGC TAC GTA ACG GCA TGA CAG TG
78836-GR1nest-3'		ATG GAA CCT GAG TGC CTC GGA CAT C
78836-FL-F	4C7	CTG TCA GCA ACC AGC AAC AA
78836-FL-R	9B2	AGC CAC AGG CTT AAT CTT CG
78836-FL2-F	9E6	TCT ACC ATG GAG GGT GTA GG

FIGURA 5

1 ATGGAACCTG AGTGCCTCGG ACATCACCCG GTCTCTCAA CTCTACGGCT GCAGCCCAAG
 w n l s a s d i t r v l k l y g c s p
 61 TGGCCCCAGG CCCCGTGGGA GAGGGTCCCA TGCCCACAGC ACTGGTAGGA GCCCCGCTCC
 s g p r p r g r g s h a h s t g r s p a
 121 GGCCTCCCTA TCTCTGCAGC GGCTTTTGGG GGCACGTGCG GCGGAATCCA GGAGCCCCGA
 p a s l s l q r l l e a l s a e s r s p
 181 CCCAGTGGT TCCAGTGCGG GAGGCCAGCC CGTTCC TGCA GGGCTGGGG AGAGCCCACA
 d p s g s s a g g q p v p a g p g e s p
 241 TGGGTGGGAG TCCCCTGCC TGAAAAAGCT CAGTGCAGAG GCCTCGGCAA GGCAGCCTCA
 h g w e s p a l k k l s a e a s a r q p
 301 GACCCTAGCT TCCTCCCCAA GATCAAGGCC TGGAGCAGGT GCCCCCGGTG TTGCTCAGGA
 q t l a s s p r s r p g a g a p g v a q
 361 GCAGTCCTGG CTGGCCGGAG TGTCCACCAA GCCACAGTC CCATCTTCAG AAGCAGGAAT
 e q s w l a g v s t k p t v p s s e a g
 421 CCAGCCAGTC CCTGTCCAGG GAAGCCCAGC TCTGCCAGGG GGCTGTGTAC CTAGAAATCA
 i q p v p v q g s p a l p g g c v p r n
 481 TTTCAAGGGG ATGTCCGAAG ATTAAGCCTG TGGCTTCTGT CCCCAAGTAG GGAGGGCATC
 h f k g m s e d
 541 CTCTGCCCAG TGGAGCTGGG TCGTCTACCT CTTGGCTCCT TTGGGCCACA CCACTGTCTT
 601 CCAGCCCCAA CCTACCACC CATCTCAGAG GGCCAGGACT CTTCCTGT CTCTCTCAC
 661 TGTGTTCCTC TAAGGGCTCC TAGGGCCAGG GGTTCCTCTA GCTCTGCCAC AGGGGAAGGC
 721 AGGCCTGGCT GTGCCTGCTC TTGACTTTTG CCCAGCCCTG GTGGATGCTG GGAATGGGAG
 781 GTGACATTCT CCAGGGACAG GTCCTGGAAG GGGTGGGGAA GAGGTAGGTT CCAGCCCCGC
 841 AGAACCCTGG AATCCCTCCT GTGCCTGAGG CCCTGCCCC CAGCATGGAC TAATGGTGTC
 901 CCTACCTCTC CCTCAGGGCA GCCCTGTGGC TGGGACCTG GGAACAGCCT CCCATCCCAC
 961 CCAACATGCC CAAGTGTGGG GGAATGTTCT ACAGCAGTGT AGCCTCCAGC CTTCTCTCC
 1021 AGGAGGCTTT GAGAGCCCAA CTTACTCCCC TGCAGAGCAG GAAGGTGGTA GGTCAAGTGT
 1081 GGCCACCATT GGGGAGACGA GAAAGAAGTG GGGCCCCACC AGATTGCACA ATGGGAACCT
 1141 CAGCTGGCCC CTGAACAGAG GACTCAGTTG TCTCCACCCT ACACCGCTAT TCCCTGGAGC
 1201 TCAGCCAGGC GCAGCCTTGG AAGGAGAAAG GGC TGGGGTT ACCTGGCTTG TCCTCTCCA
 1261 GGAAAGCCCC CTPCCTCCTC TGCCCCAGCT CCCAGCCTGG CCTCCTCCAG GCAGGCCCTA
 1321 CTCCTCTGCC CCAGCTCCGG CTTTCCCCTAT GAGGTTTGTC CCAGGCATGA AGAAAGCATC
 1381 CAGGGTGCCA ATGAGTGGGC CTAGGCCAGA GGCCCCCAG TCCCCAAGGG TACTGTTTTG
 1441 GTGGCCTTTC AGAGGGTCAA GGAAGCCCTG CTTGGGGTAG AAGGGCAGG AGCCCCACAT
 1501 GTTGGGGGAG GAAATAAAGT GGAGTGTGCT GTGCTGAAAA AAAAAAAAAA AAAA

TAA Codón de parada

AATAAA Sitio de poliadenilación consenso

FIGURA 6

Cebador	Secuencia (5'-3')
T3	ATT AAC CCT CAC TAA AGG GA
T7	TAA TAC GAC TCA CTA TAG GG
SP6	ATT TAG GTG ACA CTA TAG

8/18

FIGURA 8

Búsqueda= INSP005a
(336 letras)

Base de datos: Todas las traducciones no superfluas de GenBank CDS+PDB+SwissProt+PIR+PRF
1.247.039 secuencias; 397.579.747 letras totales

Buscando.....hecho

Secuencias que producen alineaciones significativas:	Puntuación (bits)	Valor E
ref XP 141346.1 similar a la enzima de eclosión EHE7 [Anguilla japon...	416	e-115
dbj BAB68518.1 enzima de eclosión EHE13 [Anguilla japónica]	187	2e-46
dbj BAB68515.1 enzima de eclosión EHE7 [Anguilla japónica]	186	4e-46
dbj BAB68516.1 enzima de eclosión EHE10 [Anguilla japónica]	186	4e-46
dbj BAB68513.1 enzima de eclosión EHE4 [Anguilla japónica]	186	5e-46
dbj BAB68517.1 enzima de eclosión EHE12 [Anguilla japónica]	183	3e-45
dbj BAB68514.1 enzima de eclosión EHE6 [Anguilla japónica]	183	3e-45
dbj BAB68519.1 enzima de eclosión EHE14 [Anguilla japónica]	182	7e-45
pir C48826 proteinasa de eclosión sumamente coriolítica (EC 3.4.24.-) H...	171	1e-41
dbj BAA12146.1 coriolisina H [Oryzias latipes]	171	2e-41

Alineación máxima respecto de la metaloproteínasa conocida:

>dbj|BAB68518.1| enzima de eclosión EHE13 [Anguilla japónica]
Longitud = 271

Puntuación = 187 bits (475), Esperado = 2e-46
Identidades = 93/183 (50%), Positivos = 124/183 (66%), Huecos = 3/183 (1%)

```

Problema: 5   GVVEVPFLLSSKYDEPSRQVILEALAEFERSTCIRFVTYQDQDFISIIIPMYGCFSSVGR 64
              G+VEVP+ +SS++ ++ I A+ F TCIRFV QRDFISI GC+S +GR
Objetivo: 91  GLVEVPYTVSSEFSYYHKKRIENAMETFNTECIRFVPRSSQRDFISIESRDGCYSYLGR 150

Problema: 65  SGGMQVVSLAPT-CLQKGRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINF 123
              +GG QVVSLA C+ GI+ HEL H LGF+HEHTR+DRD Y+++NW + P NF
Objetivo: 151 TGGKQVVSLARYGCVY--HGIIQHELNHALGFYHEHTRSDRDEYVKINWENVAPHTIYNF 208

Problema: 124 IKSQSSNMLTPYDYSSVMHYGR LAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKL 183
              + ++N+ TPYDY+S+MHYGR AFS G+ TITP+ P+ IGQR ++S DI R+ KL
Objetivo: 209 QEQDTNNLNTPYDYTSIMHYGR TAFSTNGMDTITVPNPNQSIGQRSMKGDILRINKL 268

Problema: 184 YGC 186
              Y C
Objetivo: 269 YSC 271
    
```

9/18

FIGURA 9

Molécula: PCR4 TOPO-IPAAA78836-1, ADN Circular de 5005 bps
 Nombre de Archivo: 13164.cm5, fecha 24 de oct. de 2002

Descripción: Ligadura de producto de PCR invertido de 78836_F2/R8 en el vector* lineal PCR4-TOPO

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	205	221	M13	sitio de cebado rev
MARCADOR	243		T3	
REGIÓN	262	294		Poliligador'
REGIÓN	294	294		sitio de clonación TOPO'
GEN	1315	308	C IPAAA78836-1	
REGIÓN	1342	295	C	producto de PCR insertado
REGIÓN	1343	1360		'Poliligador
REGIÓN	1343	1343		'sitio de clonación TOPO
MARCADOR	1395		C T7	
REGIÓN	1403	1418	M13	
GEN	2207	3001	KanR	
GEN	3205	4065	AmpR	
REGIÓN	4210	4883	pUC ori	

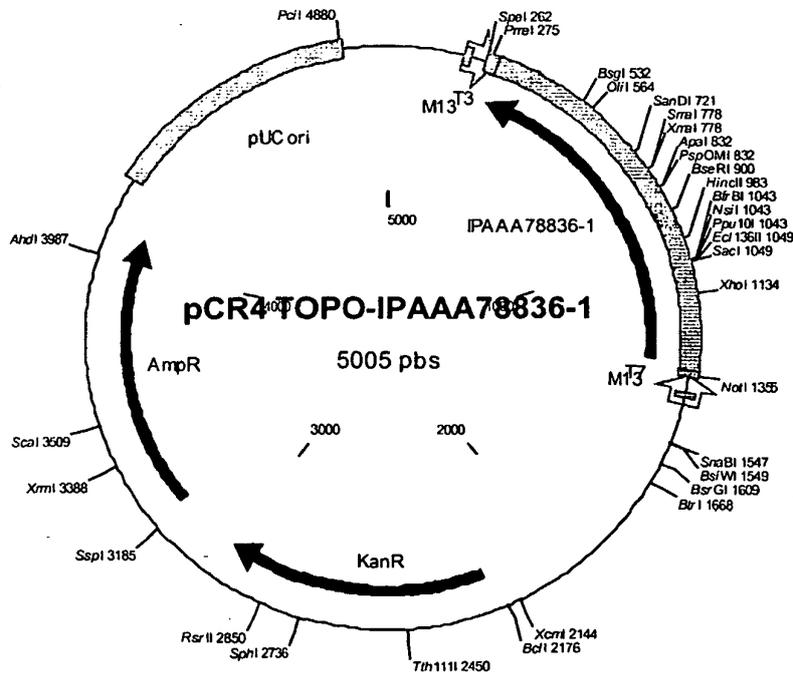


FIGURA 10

78836-FL2-F

1 TTCTACCATG GAGGGTGTAG GGGGTCTCTG GCCTTGGGTG CTGGGTCTGC TCTCCTTGCC
 m e g v g g l w p w v l g l l s l

61 AGGTGTGATC CTAGGAGCGC CCCTGGCCTC CAGCTGCGCA GGAGCCTGTG GTACCAGCTT
 p g v i l g a p l a s s c a g a c g t s

121 CCCAGATGGC CTCACCCCTG AGGGAACCCA GGCCTCCGGG GACAAGGACA TTCCTGCAAT
 f p d g l t p e g t q a s g d k d i p a

181 TAACCAAGGG CTCATCCTGG AAGAAACCCC AGAGAGCAGC TTCCTCATCG AGGGGGACAT
 i n q g l i l e e t p e s s f l i e g d

241 CATCCGGCCG AGTCCCTTCC GACTGCTGTC AGCAACCAGC AACAAATGGC CCATGGGTGG
 i i r p s p f r l l s a t s n k w p m g

301 TAGTGGTGTC GTGGAGGTCC CCTTCCTGCT CTCCAGCAAG TACGATGAGC CCAGCCGCCA
 g s g v v e v p f l l s s k y d e p s r

361 GGTCA TCCTG GAGGCTCTTG CCGAGTTTGA ACGTTCACG TGCATCAGGT TTGTACACCTA
 q v i l e a l a e f e r s t c i r f v t

421 TCAGGACCAG AGAGACTTCA TPTCCATCAT CCCCATGTAT GGGTGCTTCT CGAGTGTGGG
 y q d q r d f i s i i p m y g c f s s v

481 GCGCAGTGA GGGATGCAGG TGGTCTCCCT GGCGCCACG TGTCTCCAGA AGGGCCGGGG
 g r s g g m q v v s l a p t c l q k g r

541 CATTGTCCTT CATGAGCTCA TGCATGTGCT GGGCTTCTGG CACGAGCACA CGCGGGCCGA
 g i v l h e l m h v l g f w h e h t r a

601 CCGGGACCGC TATATCCGTG TCAACTGGAA CGAGATCCTG CCAGGCTTTG AAATCAACTT
 d r d r y i r v n w n e i l p g f e i n

661 CATCAAGTCT CGGAGCAGCA ACATGCTGAC GCCCTATGAC TACTCCTCTG TGATGCACTA
 f i k s r s s n m l t p y d y s s v m h

721 TGGGAGGCTC GCCTTCAGCC GCGGTGGGCT GCCCACCATC ACACCACTTT GGGCCCCCAG
 y g r l a f s r r g l p t i t p l w a p

781 TGTCCACATC GGCCAGCGAT GGAACCTGAG TGCCTCGGAC ATCACCCGGG TCCTCAAAC T
 s v h i g q r w n l s a s d i t r v l k

841 CTACGGCTGC AGCCCAAGTG GCCCCAGGCC CCGTGGGAGA GGGTCCCATG CCCACAGCAC
 l y g c s p s g p r p r g r g s h a h s

901 TGGTAGGAGC CCCGCTCCGG CCTCCCTATC TCTGCAGCGG CTTTTGGAGG CACTGTCCGGC
 t g r s p a p a s l s l q r l l e a l s

961 GGAATCCAGG AGCCCCGACC CCAGTGGTTC CAGTGCGGGA GGCCAGCCCC TTCCTGCAGG
 a e s r s p d p s g s s a g g q p v p a

1021 GCCTGGGGAG AGCCCACATG GGTGGGAGTC CCCTGCCCTG AAAAACTCA GTGCAGAGGC
 g p g e s p h g w e s p a l k k l s a e

11/18

1081 CTCGGCAAGG CAGCCTCAGA CCCTAGCTTC CTCCCCAAGA TCAAGGCCTG GAGCAGGTGC
a s a r q p q t l a s s p r s r p g a g

1141 CCCC GG TGT GCTCAGGAGC AGTCCTGGCT GGCCGGAGTG TCCACCAAGC CCACAGTCCC
a p g v a q e q s w l a g v s t k p t v

1201 ATCTTCAGAA GCAGGAATCC AGCCAGTCCC TGTCAGGGA AGCCAGCTC TGCCAGGGGG
p s s e a g i q p v p v q g s p a l p g

1261 CTGTGTACCT AGAAATCATT TCAAGGGGAT GTCCGAAGAT TAAGCCTGTG GCT
g c v p r n h f k g m s e d

78836-FL-R

12/18

FIGURA 11

Búsqueda= INSP005b
(431 letras)

Base de datos: Todas las traducciones no superfluas de GenBank CDS+PDB+SwissProt+PIR+PRF
1.247.039 secuencias; 397.579.747 letras totales

Buscando..... hecho

Secuencias que producen alineaciones significativas:	Puntuación (bits)	Valor E
ref XP 141346.1 similar a la enzima de eclosión EHE7 [Anguilla japon...	540	e-152
dbj BAB68513.1 enzima de eclosión EHE4 [Anguilla japónica]	198	1e-49
dbj BAB68518.1 enzima de eclosión EHE13 [Anguilla japónica]	198	1e-49
dbj BAB68516.1 enzima de eclosión EHE10 [Anguilla japónica]	197	3e-49
dbj BAB68515.1 enzima de eclosión EHE7 [Anguilla japónica]	196	4e-49
dbj BAB68514.1 enzima de eclosión EHE6 [Anguilla japónica]	196	7e-49
dbj BAB68517.11 enzima de eclosión EHE12 [Anguilla japónica]	194	3e-48
dbj BAB68519.1 enzima de eclosión EHE14 [Anguilla japónica]	191	1e-47
pir C48826 proteinasa de eclosión sumamente coriolítica (EC 3.4.24.-) H...	187	3e-46
dbj BAA12146.1 coriolisina H [Oryzias latipes]	186	4e-46

Alineación máxima respecto de la metaloproteínasa conocida:

>dbj|BAB68518.1| enzima de eclosión EHE13 [Anguilla japónica]
Longitud = 271

Puntuación = 198 bits (503), Esperado = 1e-49
Identities = 103/233 (44%), Positivos = 144/233 (61%), Huecos = 5/233 (2%)

```

Problema: 52  DKDIPAINQGLILEETPESSFLIEGDIIRPSPFRLLSATSNK--WPMGGSGVVEVPFLLS 109
              D D  I  ++      S  L+EGD++  +   ++  +N+  W    G+VEVP+ +S
Objetivo: 41  DPDDLITARILQSNNGSSEILMEGDMVVSNTRNAINCWNNQCLWRKSSDGLVEVPYTVS 100

Problema: 110 SKYDEPSRQVILEALAEFERSTCIRFVTYQDQRDFISIIIPMYGCFSSVGRSGGMQVVSLSA 169
              S++   ++ I  A+  F   TCIRFV   QRDFISI   GC+S +GR+GG QVVSLSA
Objetivo: 101 SEFSYHKKRIENAMETFNTETCIRFVPRSSQRDFISIESRDGCYSYLGRTGGKQVVSLSA 160

Problema: 170 PT-CLQKGRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINFIKSRSSNMLT 228
              C+   GI+  HEL  H  LGF+HEHTR+DRD Y+++NW  + P   NF +  ++N+  T
Objetivo: 161 RYGCVY--HGIIQHELNHALGFYHEHTRSDRDEYVKINWENVAPHTIYNFQEODTNNLNT 218

Problema: 229 PYDYSSVMHYGR LAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKLYGC 281
              PYDY+S+MHYGR  AFS  G+  TITP+  P+  IGQR ++S  DI  R+  KLY  C
Objetivo: 219 PYDYTSIMHYGR TAFSTNGMDTITPVPNPNQSIGQRRSMSGDILRINKLYSC.
    
```

13/18

FIGURA 12

Molécula: pCR4 TOPO-IPAAA78836-2, ADN Circular de 5269 pbs
 Nombre de Archivo: 13296.cm5, fecha 24 de Oct de 2002

Descripción: Ligadura de IPAAA78836v2 invertido en el vector* lineal pCR4-TOPO

Características de la molécula:

Tipo	Inicio	Fin	Nombre	Descripción
REGIÓN	205	221	M13	sitio de cebado rev
MARCADOR	243		T3	
REGIÓN	262	294		'Poliligador'
REGIÓN	294	294		sitio de clonación TOPO'
GEN	1600	307	C IPAAA78836-2	
REGIÓN	1607	1624		'Poliligador
REGIÓN	1607	1607		'sitio de clonación TOPO
MARCADOR	1659		C T7	
REGIÓN	1667	1682	M13	
GEN	2471	3265	KanR	
GEN	3469	4329	AmpR	
REGIÓN	4474	5147	pUC ori	

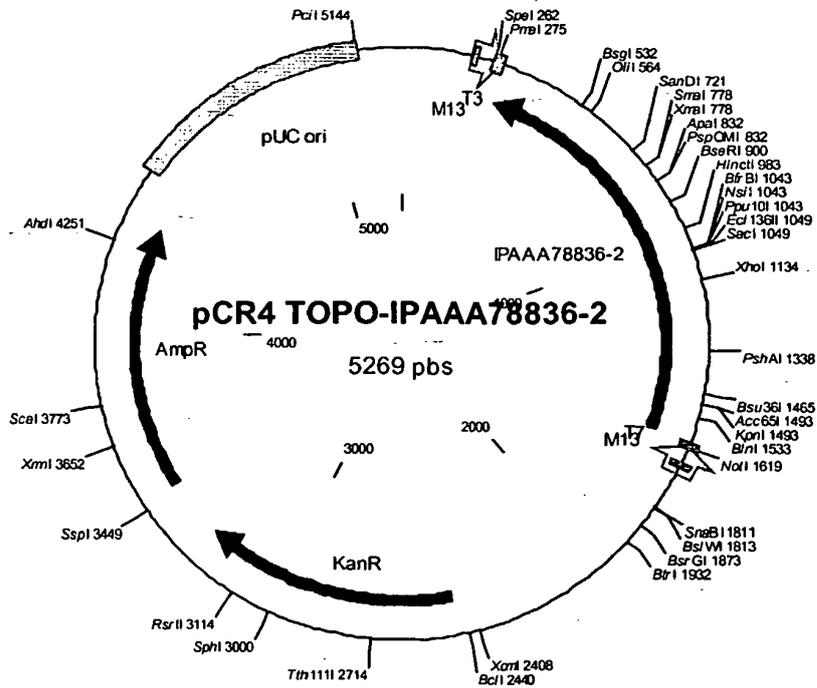
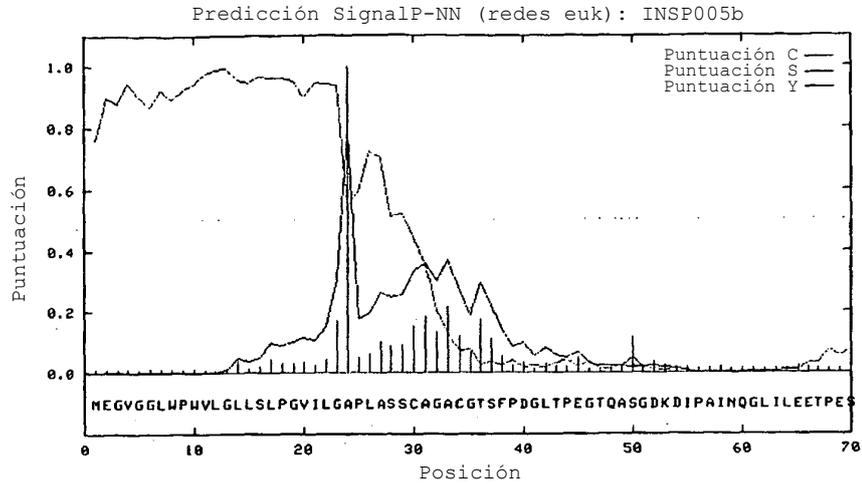


FIGURA 14

>INSP005b

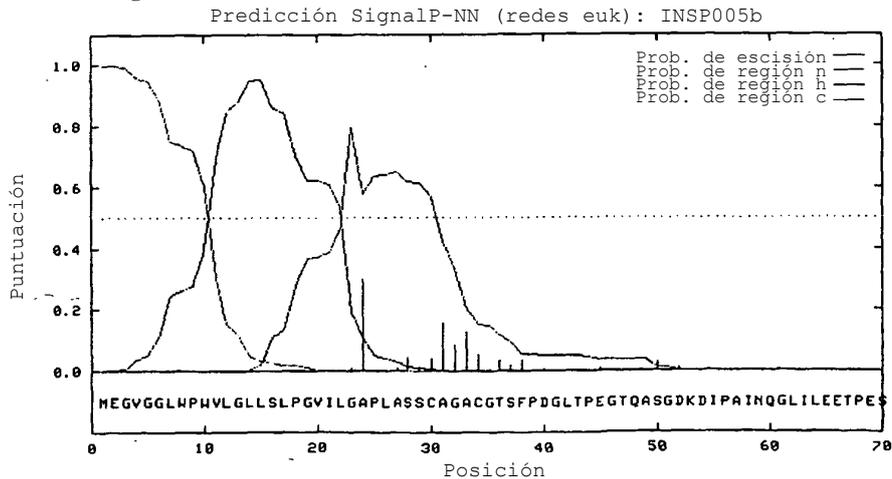
Resultado SignalP-NN:



datos

```
>INSP005b                                longitud = 70
# Medida Posición Valor Umbral péptido señal?
max. C    24      1,000  0,33  SI
max. Y    24      0,789  0,32  SI
max. S    13      0,991  0,82  SI
mean S    1-23    0,929  0,47  SI
# Sitio de escisión más probable entre pos. 23 y 24: ILG-AP
```

Resultado SignalP-HMM:



#datos

```
>INSP005b
Predicción: Péptido señal
Probabilidad de péptido señal: 0,996
Probabilidad de anclaje señal: 0,003
Probabilidad de sitio de escisión máx.: 0,302 entre la pos. 23 y 24
```

FIGURA 15A

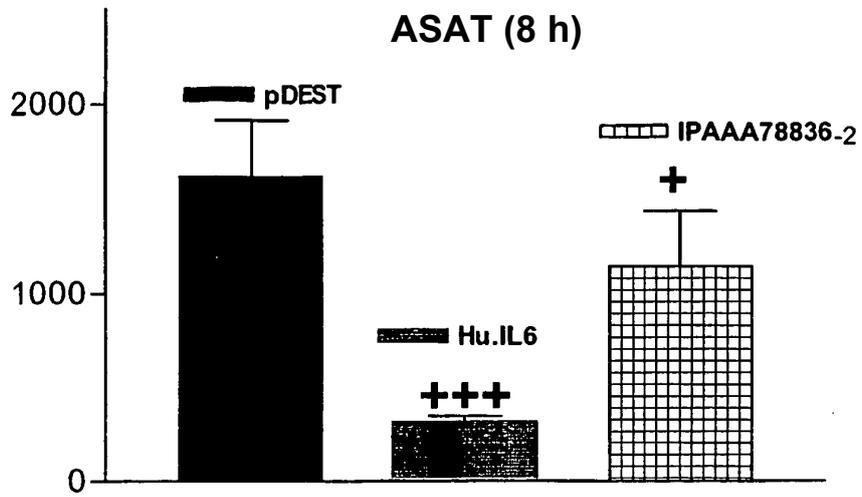


FIGURA 15B

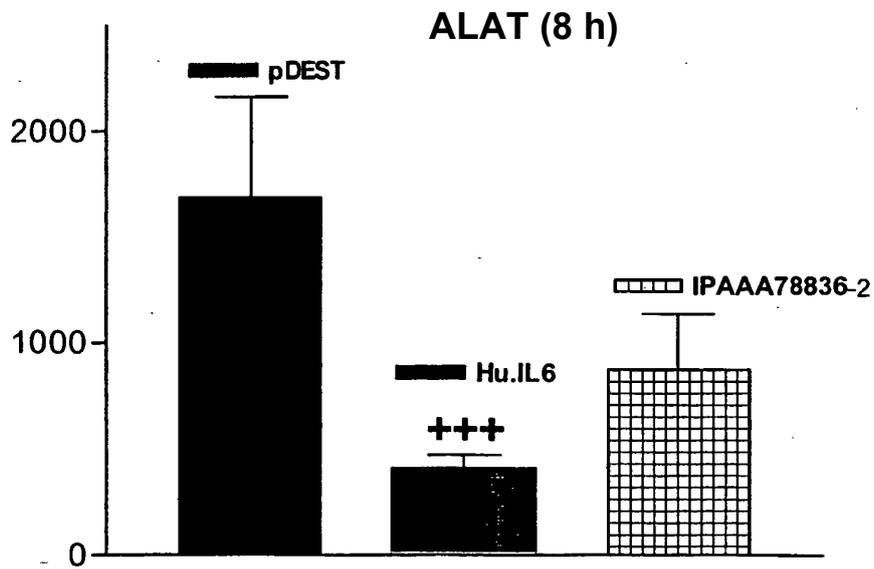


FIGURA 16A

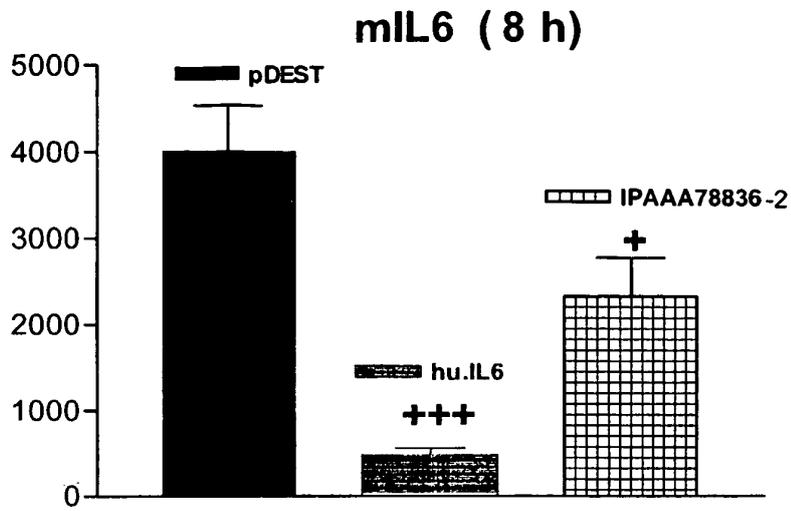


FIGURA 16B

