

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 378 527**

51 Int. Cl.:
C12N 15/12 (2006.01) **A61P 35/00** (2006.01)
C07K 14/435 (2006.01)
C07K 14/47 (2006.01)
C12N 15/10 (2006.01)
C12N 15/66 (2006.01)
C12N 15/11 (2006.01)
C12Q 1/68 (2006.01)
C07K 16/18 (2006.01)
G01N 33/53 (2006.01)
A61K 51/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **00945990 .0**
- 96 Fecha de presentación: **22.06.2000**
- 97 Número de publicación de la solicitud: **1192253**
- 97 Fecha de publicación de la solicitud: **03.04.2002**

54 Título: **Polipéptido ICBP90 y sus fragmentos, polinucleótidos que codifican dichos polipéptidos y aplicaciones para el diagnóstico y el tratamiento del cáncer**

30 Prioridad:
22.06.1999 FR 9907935

45 Fecha de publicación de la mención BOPI:
13.04.2012

45 Fecha de la publicación del folleto de la patente:
13.04.2012

73 Titular/es:
**INSTITUT NATIONAL DE LA SANTE ET DE LA
RECHERCHE MEDICALE (INSERM)
101, RUE DE TOLBIAC
75654 PARIS CÉDEX 13, FR**

72 Inventor/es:
**BRONNER, Christian;
HOPFNER, Raphael;
MOUSLI, Marc;
JELTSCH, Jean-Marc;
LUTZ, Yves y
OUDET, Pierre**

74 Agente/Representante:
Carpintero López, Mario

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).

ES 2 378 527 T3

DESCRIPCIÓN

Polipéptido ICBP90 y sus fragmentos, polinucleótidos que codifican dichos polipéptidos y aplicaciones para el diagnóstico y el tratamiento del cáncer

5 La presente invención se refiere a un nuevo polipéptido ICBP90 y a sus fragmentos, a la clonación del ADNc y los polinucleótidos que codifican dichos polipéptidos, a vectores de clonación y/o de expresión que incluyen dichos polinucleótidos, a células transformadas por dichos vectores y a anticuerpos específicos dirigidos contra dichos polipéptidos. La invención se refiere igualmente a procedimientos y a kits de diagnóstico de cánceres, a un procedimiento y un kit de cribado de ligandos del polipéptido de la invención y a compuestos utilizables a modo de medicamento para la prevención y/o el tratamiento de cánceres.

10 Las ADN topoisomerasas son proteínas nucleares altamente conservadas en el transcurso de la evolución cuyo papel principal es controlar la conformación y topología del ADN en el núcleo, que se alteran constantemente por los diferentes procesos biológicos que implican el ADN, tales como por ejemplo transcripción y replicación. Las topoisomerasas ejercen su acción cortando el ADN y religando estas lesiones después de realizar el cambio conformacional adecuado.

15 En los mamíferos, y en particular el hombre, existen actualmente al menos cinco genes diferentes que codifican una topoisomerasa y al menos dos pseudogenes adicionales (para revisión, véase Nitiss 1998). Así, la topoisomerasa I, codificada por el gen TOP1, retira los supergiros presentes en el ADN cortando solo una hebra. Las dos topoisomerasas de tipo II que existen en el hombre, denominadas TopII α y TopII β , alteran la topología del ADN al introducir escisiones bicatenarias transitorias (para revisión, véase Wang 1996). Por último, existen dos topoisomerasas de tipo III codificadas por dos genes localizados en 17p 11.2-12 y 22q 11-12 y que actúan únicamente sobre los supergiros negativos del ADN.

20 En las células tumorales, las topoisomerasas de tipo II desempeñan un papel muy importante; en estas células en crecimiento y división rápida, existe una gran necesidad de mantener las moléculas de ADN en una conformación correcta, ya que son necesarias tasas de transcripción y replicación elevadas. Así, las tasas de topoisomerasas II son generalmente más elevadas en las células tumorales humanas que en los tejidos normales del mismo origen. Sin embargo, la tasa de expresión elevada de la topoisomerasa II α en las células tumorales puede variar entre dos tumores de naturaleza diferente que afectan al mismo tejido. Por ejemplo, el núcleo de células de carcinoma pulmonar microcítico presenta una tasa más elevada de topoisomerasa II α que el núcleo de células de carcinomas pulmonares no microcíticos (Guinee *et al.*, 1996). De la misma manera, la tasa de topoisomerasa II α en células A549 es tres veces más elevada en que células PC3, procediendo estas dos líneas celulares de adenocarcinoma de epitelio pulmonar (Yamasaki *et al.*, 1996).

25 Estas comprobaciones hacen pensar que la topoisomerasa II α puede considerarse como un marcador de proliferación celular para ciertos tipos de cáncer. El proceso canceroso se caracteriza por una proliferación celular anormal debida en parte a la pérdida de la inhibición por contacto, así pues, la topoisomerasa II α parece una diana privilegiada de fármacos quimioterapéuticos para el tratamiento del cáncer (Pommier *et al.* 1994), y los tratamientos anticancerosos actuales recurren en gran medida a inhibidores de topoisomerasas.

30 La mayoría de estos inhibidores ejercen sus efectos citotóxicos estabilizando el complejo de escisión de ADN. Los fármacos como las antraciclinas [doxorubicina (adriamicina) o epipodofilotoxinas (tal como etopósido (VP-16) o tenipósido (VM26))], acridinas (tal como mAMSA) y antracenedionas (por ejemplo, mitoxantrona) son ejemplos de fármacos inhibidores de topoisomerasas II que estabilizan el complejo de escisión. Más recientemente, se ha desarrollado una nueva clase de inhibidores de topoisomerasas II; estos inhibidores actúan al nivel de la actividad catalítica y no estabilizan ya el complejo de escisión. El fármaco fostriecina es un ejemplo de ellos (Boritzki *et al.*, 1988). A día de hoy, se utilizan estos diferentes fármacos en tratamientos anticancerosos curativos y paliativos.

35 No obstante, uno de los mayores problemas encontrados en los tratamientos anticancerosos actuales que utilizan inhibidores de topoisomerasa es la emergencia de resistencia a fármacos (Kubo *et al.*, 1995). Estas resistencias son o bien por la sobreexpresión de las bombas que permite el flujo de salida de fármacos al exterior de las células antes de que alcancen su diana (por ejemplo, P-glucoproteína, proteína asociada a la multiresistencia de fármacos (MRP)), o bien por el cambio de la tasa de expresión de topoisomerasa II α (Deffie *et al.*, 1989; Fry *et al.*, 1991), o bien por los dos (para revisión, véase Isaacs *et al.*, 1998).

40 Así pues, uno de los aspectos de la presente invención es comprender los mecanismos de regulación de la expresión del gen de topoisomerasa II α con el fin de desarrollar una alternativa al fenómeno de resistencia a los fármacos observado para ciertos cánceres, y esto con el enfoque de mejorar el tratamiento preventivo y curativos de los cánceres.

45 Existen dos tipos de topoisomerasas de tipo II que difieren en su perfil de expresión: la topoisomerasa II α (Top II α) (170 kDa), localizada esencialmente en el nucleoplasma al nivel de los centrómeros de los cromosomas mitóticos, interviene en procesos biológicos fundamentales como son replicación, condensación de cromosomas y transcripción. La topoisomerasa II β (Top II β) (180 kDa) parece estar más bien implicada en la transcripción de ARN ribosómicos dada la localización nucleolar de esta enzima. Las dos topoisomerasas de tipo II humanas están

localizadas en dos cromosomas diferentes (17q21-22 para la topoisomerasa II α y 3p24 para la topoisomerasa II β) (Tsai-Plugfelder *et al.*, 1988; Drake *et al.*, 1989; Chung *et al.*, 1989; Jenkins *et al.*, 1992; Austin *et al.*, 1993).

Al contrario que la topoisomerasa II β , cuya expresión se caracteriza por una constancia relativa, la topoisomerasa II α presenta una variación de la expresión en función del estado de proliferación de las células y de su posición en el ciclo celular. La expresión del ARN mensajero (ARNm) es más elevada en las células en proliferación que en las células detenidas en confluencia. La expresión de la topoisomerasa II α aumenta en el transcurso de la fase S del ciclo celular, alcanzando un máximo al final de la fase G2/M (Goswami *et al.*, 1996), y siendo el nivel de ARN mensajero diez veces más elevado al final de la fase S que durante la fase G1. Igualmente, parece existir un acoplamiento entre la síntesis y la degradación de topoisomerasa II α y la condensación/descondensación cromosómica (Heck *et al.*, 1988).

Los conocimientos actuales referentes a la regulación del gen de topoisomerasa II α resultan en conjunto bastante escasos. Recientemente, se ha descrito una región promotora de aproximadamente 650 pares de bases por Hochhauser *et al.* (1992), que presenta todas las características de un gen doméstico, siendo dos ejemplos la ausencia de la secuencia TATA y la riqueza moderada de sitios GC (presencia especialmente de una secuencia Sp1 que puede reemplazar a la secuencia TATA). La presencia de 5 secuencias CCAAT invertidas o ICB (inverted CCAAT box en inglés) es otra particularidad de este tipo de promotor.

Se han descrito factores de transcripción que interaccionan con el promotor del gen de topoisomerasa II α humana; se pueden citar c-myb (Brandt *et al.*, 1997), p53 (Sandri *et al.*, 1996), ATF (Lim *et al.*, 1998), Sp1 y Sp3 (Kubo *et al.*, 1995). Cualesquiera que sean, aparte de NF-Y (igualmente denominado CBF, ACF y CP1, referencia en Isaacs *et al.*, 1996), los factores de transcripción que actúan sobre las secuencias ICB del promotor del gen de topoisomerasa II α humana no se han identificado y caracterizado todos todavía; sin embargo, Herzog y Zwelling (1997) han determinado dos proteínas de un peso molecular aparente de 90 kDa y 140 kDa que ligan ICB1 respectivamente con ICB4 e ICB5. Isaacs y sus colaboradores (1996) han propuesto que el NFY, así como otra proteína no identificada, reconocen una secuencia ICB de la región promotora del gen de topoisomerasa II α ; han mostrado igualmente que las mutaciones de ICB2 anulaban completamente la disminución de la actividad promotora observada normalmente en las células detenidas en confluencia (Isaac *et al.*, 1996). Han identificado a NFY como un componente de un complejo inducido por la proliferación y que se liga *in vitro* con la secuencia ICB2 del promotor del gen de topoisomerasa II α humana, aunque NF-Y siga siendo detectable en las células detenidas en confluencia (Isaac *et al.*, 1996). Han propuesto que la ICB2 actúa como regulador negativo del promotor del gen de topoisomerasa II α de células detenidas en confluencia y que esta represión puede suprimirse en las células proliferativas. La secuencia ICB2 del promotor del gen de topoisomerasa II α desempeña pues un papel primordial en la detención del proceso proliferativo normal cuando las células llegan a confluencia.

Los factores de transcripción que se ligan con la secuencia ICB, así como la secuencia ICB misma, constituyen pues dianas moleculares para controlar la tasa de expresión de la topoisomerasa II α . Al intervenir sobre dichos factores, es posible considerar controlar la expresión del gen de topoisomerasa II α y como consecuencia la proliferación celular.

La presente invención tiene como objeto la determinación de nuevos factores de transcripción que se ligan con la secuencia ICB implicada en el control de la proliferación celular.

Se ha utilizado una técnica reciente, denominada sistema "híbrido simple", que permite aislar clones de ADNc que codifican proteínas de unión con ADN específico de ciertas secuencias. Este sistema presenta una doble ventaja, puesto que es capaz no solo de revelar las interacciones ADN-proteína *in vivo* en levadura, sino también de dar acceso directamente a los ADN complementarios (ADNc) que codifican las proteínas candidatas que tienen actividad de factor de transcripción. El sistema se basa principalmente en la construcción de una cepa de levadura de ensayo según el principio planteado por Wang y Reed (1993). Esta cepa de levadura permite el cribado de bancos de ADNc, determinando la interacción ADN-proteína *in vivo* mediante el sesgo de activación de un gen informador integrado en el genoma de la levadura de ensayo.

La presente invención tiene pues como objeto un polipéptido aislado denominado ICBP90 (inverted CCAAT box binding protein en inglés) de secuencia aminoacídica SEQ ID N°2. Esta secuencia comprende:

- a) un dominio de "ubiquitina" que comprende la secuencia aminoacídica 1 a 75 de la secuencia de SEQ ID N°2;
- b) un dominio de "dedo de cinc" de tipo C4HC3 que comprende la secuencia aminoacídica 310 a 366 de la secuencia SEQ ID N°2 y un dominio de "dedo de cinc" de tipo C3HC4 que comprende la secuencia aminoacídica 724 a 763 de la secuencia SEQ ID N°2;
- c) un supuesto dominio de "cremallera de leucina" que comprende la secuencia aminoacídica 58 a 80 de la secuencia SEQ ID N°2;
- d) dos dominios de localización nuclear potenciales que comprenden las secuencias aminoacídicas 581 a 600 y 648 a 670 de la secuencia SEQ ID N°2;

e) un sitio de fosforilación por una tirosina cinasa que comprende la secuencia aminoacídica 452 a 458 de la secuencia SEQ ID N°2;

f) sitios de fosforilación por una proteína cinasa dependiente de cAMP/cGMP que comprende las secuencias aminoacídicas 246 a 249, 295 a 298 y 648 a 651 de la secuencia SEQ ID N°2;

5 g) sitios de fosforilación por una caseína cinasa II que comprenden las secuencias aminoacídicas 23 a 26, 57 a 60, 91 a 94, 109 a 112, 165 a 168, 265 a 268, 354 a 357 y 669 a 672 de la secuencia SEQ ID N°2;

h) sitios de fosforilación por una proteína cinasa C que comprenden las secuencias aminoacídicas 82 a 84, 104 a 106, 160 a 162, 173 a 175, 251 a 253, 301 a 303, 380 a 382, 393 a 395, 504 a 506, 529 a 531, 625 a 627 y 639 a 641 de la secuencia SEQ ID N°2.

10 Ha de observarse que Fujimori *et al.* (*Mammalian Genome*, 9: 1032) describen por otro lado una proteína nuclear de murino denominada Np95 cuya secuencia nucleotídica presenta una identidad de secuencia de un 77,6% con la secuencia nucleotídica de SEQ ID NO1 de la invención, y cuya secuencia proteica presenta una identidad de secuencia de un 73,6% con la SEQ ID N°2 de la invención. Sin embargo, el perfil de expresión de Np95 difiere en gran medida del de la proteína ICBP90 de la invención: al contrario que ICBP90, la Np95 se expresa en gran medida en bazo, pulmón y testículos. Las funciones biológicas de estas dos proteínas parecen pues claramente distintas a pesar de la homología de secuencia.

15

La presente invención hace referencia igualmente a un polipéptido aislado caracterizado porque comprende un polipéptido elegido entre:

20 a) un polipéptido de secuencias SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8; o el fragmento C-terminal a partir del residuo D263 de la SEQ ID N°2.

Debe entenderse que la invención se refiere a polipéptidos obtenidos mediante purificación a partir de fuentes naturales, o bien obtenidas por recombinación genética, o también mediante síntesis química y que pueden comprender entonces aminoácidos no naturales.

En la presente descripción, se utilizará el término polipéptido para designar igualmente a una proteína o péptido.

25 Se pretende designar como polipéptido variante al conjunto de polipéptidos mutados que pueden existir naturalmente, particularmente en el ser humano, y que corresponden especialmente a truncamientos, sustituciones, deleciones y/o adiciones de residuos aminoacídicos. Los polipéptidos homólogos según la invención conservan al menos un dominio elegido entre el dominio de unión a ADN y/o el dominio de interacción con otra proteína.

30 Se pretende designar como polipéptido homólogo a los polipéptidos que presentan, con relación al polipéptido natural ICBP90, ciertas modificaciones como, particularmente, una deleción, adición o sustitución de al menos un aminoácido, un truncamiento, alargamiento y/o fusión quimérica. Entre los polipéptidos homólogos, se prefieren aquellos cuya secuencia aminoacídica presenta al menos un 80% de homología, preferiblemente un 90%, más preferiblemente un 95% y aún más preferiblemente un 97% de homología con las secuencias aminoacídicas de polipéptidos según la invención. En el caso de una sustitución, se reemplazan uno o varios aminoácidos consecutivos o no consecutivos por aminoácidos "equivalentes". La expresión aminoácido "equivalente" pretende designar aquí cualquier aminoácido susceptible de ser sustituido por un aminoácido de estructura básica sin modificar sin embargo las características o propiedades funcionales esenciales, como sus actividades biológicas, de los polipéptidos correspondientes, tales como la inducción *in vivo* de anticuerpos capaces de reconocer el polipéptido cuya secuencia aminoacídica está comprendida en la secuencia aminoacídica SEQ ID N°2, o uno de sus fragmentos definidos anteriormente, y especialmente las secuencias aminoacídicas SEQ ID N°4, SEQ ID N°6 y SEQ ID N°8. Estos aminoácidos equivalentes pueden determinarse o bien basándose en su homología estructural con los aminoácidos a los que sustituye, o bien con los resultados de ensayos de actividad biológica cruzada a los que son susceptibles de dar lugar los diferentes polipéptidos. A modo de ejemplo, se mencionarán las posibilidades de sustituciones susceptibles de efectuarse sin que resulte una modificación profunda de las actividades biológicas de los polipéptidos modificados correspondientes, los reemplazos, por ejemplo, de leucina por valina o isoleucina, de ácido aspártico por ácido glutámico, de glutamina por asparagina, de arginina por lisina, etc., siendo naturalmente factibles las sustituciones inversas en las mismas condiciones.

45 Se pretende designar como fragmento biológicamente activo particularmente un fragmento de secuencia aminoacídica de un polipéptido según la invención que presente al menos una de las características o propiedades funcionales de los polipéptidos según la invención, especialmente que: (i) pueda ser reconocido por un anticuerpo específico de un polipéptido según la invención; (ii) presente al menos uno de los dominios o regiones tales como se definen a continuación; (iii) pueda ligarse con ADN y especialmente con las secuencias CCAATT y/o CCAAT inversa; (iv) pueda modular la tasa de expresión del gen de topoisomerasa IIa; (v) pueda modular la proliferación celular.

55 Se pretende designar como fragmento de polipéptido un polipéptido que comprende como mínimo 5 aminoácidos, preferiblemente 7 aminoácidos, más preferiblemente 10 y aún más preferiblemente 15 aminoácidos. Los fragmentos de polipéptido según la invención obtenidos mediante escisión de dicho polipéptido por una enzima proteolítica, por

un reactivo químico, o también disponiendo dicho polipéptido en un entorno muy ácido, forman parte igualmente de la invención.

El polipéptido según la invención puede asociarse igualmente con otros polipéptidos mediante interacciones proteína-proteína. Se pretende designar como interacciones proteína-proteína las asociaciones que ponen directamente en contacto al menos dos proteínas. Así, el polipéptido de la invención puede dimerizarse formando homodímeros o heterodímeros, o asociarse en forma de homomultímeros o heteromultímeros. El polipéptido según la invención puede interactuar igualmente con otro polipéptido para ejercer su función; así, el polipéptido según la invención puede poseer, además de su dominio de unión a ADN, un dominio de acción sobre la transcripción que ejerza su efecto mediante interacciones proteína-proteína con otros componentes proteicos de la maquinaria transcripcional. Se pretende designar como componente proteico de la maquinaria transcripcional todos los factores de transcripción necesarios para la realización y la regulación de la reacción de transcripción.

El polipéptido según la invención se caracteriza porque puede ligarse con una secuencia de ADN y porque comprende al menos un dominio de fijación al ADN seleccionado del grupo compuesto por un dominio de "dedo de cinc" (zinc-finger) y un dominio de "cremallera de leucina, siendo la secuencia de ADN con la que se liga dicho polipéptido una secuencia CCAAT, preferiblemente una secuencia CCAAT inversa (inverted CCAAT box : ICB).

Se pretende designar como enlace con una secuencia de ADN una interacción específica entre el polipéptido de la invención y una secuencia de ADN mediante una serie de enlaces débiles realizados entre los aminoácidos de la proteína y las bases. El polipéptido según la invención posee al menos un dominio de unión a ADN que contiene al menos uno de los motivos proteicos conocidos susceptibles de interactuar con ADN, es decir, la estructura en dedo de guante a la que está asociado un átomo de cinc ("zinc finger"), la estructura de hélice-vuelta-hélice, la estructura de hélice-bucle-hélice y la cremallera de leucina ("leucine zipper").

Se pretende designar como motivo en dedo de guante ("zinc finger") una secuencia de una veintena de aminoácidos que tienen una forma espacial en dedo de guante. Existen dos tipos: los que contienen cuatro cisteínas (C4) y los que contienen dos cisteínas y dos histidinas (C2H2). Estos aminoácidos definen la naturaleza del dedo de guante y se sitúan en su base, y un ión Zn^{++} se sitúa en el centro del cuadrado formado por estos cuatro aminoácidos. El polipéptido según la invención posee potencialmente dos motivos de tipo C4.

Se pretende designar como motivo de tipo "leucine zipper" motivos que pertenecen a factores de transcripción diméricos que son homodímeros o heterodímeros. El monómero está constituido por una secuencia de tipo básico que interactúa específicamente con el ADN y un dominio hidrófobo de hélice α que interactúa con el dominio homólogo de la otra cadena. En este dominio, se encuentra una leucina cada 7 aminoácidos, es decir, en cada vuelta de hélice. Todas estas leucinas están alineadas y la interacción se realiza a su nivel entre los dos monómeros. El péptido según la invención posee potencialmente un motivo de tipo "cremallera de leucina".

La invención se refiere igualmente a un polinucleótido aislado caracterizado porque codifica un polipéptido de secuencia SEQ ID N° 1 tal como se define anteriormente. Preferiblemente, el polinucleótido según la invención posee la secuencia SEQ ID N°1.

La invención se refiere igualmente al polinucleótido aislado, caracterizado porque comprende un polinucleótido elegido entre:

a) un polinucleótido de secuencias SEQ ID N°1, SEQ ID N°3, SEQ ID N°5 o SEQ ID N°7 o que codifica el fragmento C-terminal a partir del residuo D236 de la SEQ ID N°2.

En la presente descripción, se pretenderá designar como polinucleótido, oligonucleótido, secuencia polinucleotídica, secuencia nucleotídica o ácido nucleico un fragmento de ADN, tanto un ADN bicatenario o un ADN monocatenario como productos de la transcripción de dichos ADN, y/o un fragmento de ARN, comprendiendo o no dichos fragmentos naturales aislados o de síntesis nucleótidos no naturales, y designando un encadenamiento preciso de nucleótidos, modificados o no, que permiten definir un fragmento o una región de un ácido nucleico.

Se pretende designar como polinucleótido de secuencia complementaria cualquier ADN cuyos nucleótidos sean complementarios de los de SEQ ID N° 1, SEQ ID N°3, SEQ ID N°5, SEQ ID N°7 o de una parte de las SEQ ID N° 1, SEQ ID N°3, SEQ ID N°5, SEQ ID N°7 y cuya orientación esté invertida.

Se entiende por porcentaje de homología en el sentido de la presente invención un porcentaje de identidad entre las bases de dos polinucleótidos, siendo este porcentaje puramente estadístico y estando repartidas las diferencias entre los dos polinucleótidos al azar sobre toda su longitud. Según la invención, los polinucleótidos de secuencia nucleica homóloga presentan una tasa de homología de al menos un 80%, preferiblemente un 90%, más preferiblemente un 95% y aún más preferiblemente un 97%.

Una hibridación en condiciones de alto rigor significa que las condiciones de temperatura y de fuerza iónica se eligen de tal manera que permitan el mantenimiento de la hibridación entre dos fragmentos de ADN complementarios. A modo ilustrativo, las condiciones de rigor alto de la etapa de hibridación, con el fin de definir los fragmentos polinucleotídicos descritos anteriormente, son ventajosamente las siguientes:

La hibridación ADN-ADN o ADN-ARN se realiza en dos etapas: (1) prehibridación a 42°C durante 3 horas en tampón fosfato (20 mM, pH 7,5) que contiene 5 x SSC (1 x SSC corresponde a una solución de NaCl 0,15 M + citrato de sodio 0,015 M), 50% de formamida, 7% de dodecilsulfato de sodio (SDS), 10 x de solución de Denhard, 5% de sulfato de dextrano y 1% de ADN de esperma de salmón; (2) hibridación propiamente dicha durante 20 horas a una temperatura dependiente del tamaño de la sonda (concretamente: 42°C para una sonda de tamaño > 100 nucleótidos) seguido de 2 lavados de 20 minutos con 2 x SSC + 2% de SDS y 1 lavado de 20 minutos a 20°C con 0,1 x SSC + 0,1% de SDS. El último lavado se practica con 0,1 x SSC + 0,1% de SDS durante 30 minutos a 60°C para una sonda de tamaño > 100 nucleótidos. Las condiciones de hibridación de alto rigor descritas anteriormente para un polinucleótido de tamaño definido se adaptarán por el especialista en la materia para los oligonucleótidos de tamaño más grande o más pequeño, según las enseñanzas de Sambrook *et al.*, 1989.

Ventajosamente, un fragmento nucleotídico que responda a la definición precedente tendrá al menos 15 nucleótidos, preferiblemente al menos 21 nucleótidos, aún más preferiblemente al menos 30 nucleótidos consecutivos de la secuencia de la que procede.

Se pretende designar como EST ("expressed sequence tag") las secuencias expresadas caracterizadas en un banco de ADN complementario (ADNc) y utilizadas como baliza cartográfica del ADN genómico.

Dicho polinucleótido se caracteriza porque está marcado directa o indirectamente con un compuesto radiactivo o un compuesto no radiactivo. Puede utilizarse como cebador para la amplificación o polimerización de secuencias nucleicas, o como sonda para la detección de secuencias nucleicas. Los fragmentos de polinucleótidos que pueden utilizarse como sonda o como cebador en procedimientos de detección, identificación, valoración o amplificación de secuencias nucleicas presentarán un tamaño mínimo de 9 bases, preferiblemente de 18 bases, más preferiblemente de 36 bases. Por último, es posible utilizar un polinucleótido según la invención como secuencia de ácido nucleico codificante o anticodificante para controlar la expresión del correspondiente producto proteico.

Las secuencias de polinucleótidos según la invención no marcadas pueden utilizarse directamente como sonda, cebador u oligonucleótido; sin embargo, las secuencias utilizadas están generalmente marcadas para obtener secuencias utilizables por numerosas aplicaciones. El marcaje de cebadores, sondas u oligonucleótidos según la invención se realiza mediante elementos radiactivos o mediante moléculas no radiactivas; entre los isótopos radiactivos utilizados, se pueden citar ³²P, ³³P, ³⁵S, ³H o ¹²⁵I. Las entidades no radiactivas se seleccionan entre ligandos tales como biotina, avidina, estreptavidina, dioxigenina, haptenos, colorantes, agentes luminiscentes tales como agentes radioluminiscentes, quimioluminiscentes, fluorescentes y fosforescentes.

Los polinucleótidos según la invención pueden utilizarse así como cebador y/o sonda en procedimientos que aplican especialmente la técnica PCR (reacción en cadena de la polimerasa) (Erich, 1989; Innis *et al.*, 1990 y Rolfs *et al.*, 1991). Esta técnica necesita la elección de pares de cebadores oligonucleotídicos que flanqueen el fragmento que debe amplificarse. Por ejemplo, se puede referir a la técnica descrita en la patente estadounidense U.S. N° 4.683.202. Los fragmentos amplificados pueden identificarse, por ejemplo, después de una electroforesis en gel de agarosa o poliacrilamida, o después de una técnica cromatográfica como la filtración en gel o la cromatografía por intercambio de iones. La especificidad de la amplificación puede controlarse mediante hibridación molecular utilizando como sonda las secuencias nucleotídicas de polinucleótidos de la invención, plásmidos que contienen estas secuencias o sus productos de amplificación. Los fragmentos nucleotídicos amplificados pueden utilizarse como reactivos en reacciones de hibridación con el fin de determinar la presencia, en una muestra biológica, de un ácido nucleico diana de secuencia complementaria a la de dichos fragmentos nucleotídicos amplificados.

La invención hace igualmente referencia a fragmentos nucleotídicos susceptibles de obtenerse mediante amplificación con la ayuda de cebadores según la invención.

Pueden emplearse ventajosamente otras técnicas de amplificación de ácido nucleico diana como alternativa a la PCR (de tipo PCR) con la ayuda de un par de cebadores de secuencias nucleotídicas según la invención. Se pretenderá designar como de tipo PCR todos los procedimientos que aplican reproducciones directas o indirectas de secuencias de ácidos nucleicos, o bien en los que los sistemas de marcado se han amplificado; estas técnicas son ciertamente conocidas, en general, se trata de la amplificación de ADN por una polimerasa; cuando la muestra de origen es un ARN, conviene efectuar previamente una transcripción inversa. Existen actualmente muy numerosos procedimientos que permiten esta amplificación como, por ejemplo, la técnica SDA (Strand Displacement Amplification) o técnica de amplificación por desplazamiento de hebra (Walker *et al.*, 1992), la técnica TAS (Transcription-based Amplification System, o sistema de amplificación basado en la transcripción) descrita por Kwoh *et al.* en 1989, la técnica 3SR (Self-Sustained Sequence Replication, o replicación de secuencia automantenida) descrita por Guatelli *et al.* en 1990, la técnica NASBA (Nucleic Acid Sequence Based Amplification, o amplificación basada en la secuencia de ácido nucleico) descrita por Kievitis *et al.* en 1991, la técnica TMA (Transcription Mediated Amplification, o amplificación mediante por la transcripción), la técnica LCR (Ligase Chain Reaction, o reacción en cadena de la ligasa) descrita por Landegren *et al.* en 1988 y perfeccionada por Barany *et al.* en 1991, que emplea una ligasa termoestable, la técnica de RCR (Repair Chain Reaction, o reacción en cadena de reparación) descrita por Segev en 1992, la técnica CPR (Cycling Probe Reaction, o reacción de sonda cíclica) descrita por Duck *et al.* en 1990, la técnica de amplificación por replicasa Q-beta descrita por Miele *et al.* en 1983 y

perfeccionada especialmente por Chu *et al.* en 1986 y Lizardi *et al.* en 1988, y después por Burg *et al.* así como por Stone *et al.* en 1996.

En el caso en que el polinucleótido diana a detectar sea un ARN, por ejemplo un ARNm, se utilizará ventajosamente, antes de la aplicación de una reacción de amplificación con la ayuda de cebadores según la invención o en la aplicación de un procedimiento de detección con la ayuda de sondas de la invención, una enzima de tipo transcriptasa inversa con el fin de obtener un ADNc a partir del ARN contenido en la muestra biológica. El ADNc servirá entonces como diana para los cebadores o las sondas aplicadas en el procedimiento de amplificación o detección según la invención.

Las sondas nucleotídicas hibridan específicamente con una molécula de ADN o ARN de polinucleótido, más particularmente con la secuencia SEQ ID N° 1 que codifica el polipéptido ICBP90, en condiciones de hibridación de alto rigor tales como las dadas en forma de ejemplo anteriormente.

La técnica de hibridación puede realizarse de diversas maneras (Matthews *et al.*, 1988). El procedimiento más general consiste en inmovilizar el ácido nucleico extraído de células de diferentes tejidos o de células en cultivo sobre un soporte (tal como nitrocelulosa, nailon o poliestireno) e incubar, en condiciones bien definidas, el ácido nucleico diana inmovilizado con la sonda. Después de la hibridación, se elimina en exceso de sonda y se detectan las moléculas híbridas formadas mediante el procedimiento apropiado (medida de la radiactividad, fluorescencia o actividad enzimática ligada a la sonda).

Según otro modo de aplicación de las sondas nucleicas, estas últimas pueden utilizarse como sonda de captura. En este caso, dicha "sonda de captura" se inmoviliza sobre un soporte y sirve para capturar mediante hibridación específica el ácido nucleico diana obtenido a partir de la muestra biológica a ensayar, y a continuación se detecta el ácido nucleico diana gracias a una segunda sonda, denominada "sonda de detección" marcada con un elemento fácilmente detectable.

La invención se refiere a un vector recombinante de clonación de un polinucleótido según la invención y/o de expresión de un polipéptido según la invención, caracterizado porque contiene un polinucleótido según la invención tal como se describe anteriormente. El vector según la invención se caracteriza porque comprende elementos que permiten la expresión y eventualmente la secreción de dichas secuencias en una célula hospedadora. Estos vectores son útiles para transformar células hospedadoras con el fin de clonar o expresar las secuencias nucleotídicas de la invención. Son vectores particulares, por ejemplo, vectores de origen vírico o plasmídico. Entre estos vectores, se prefieren aquellos de la serie pGEX (Pharmacia) para expresión en bacterias o pSG5 (Stratagene, La Jolla, CA, EE.UU.) para expresión en sistema eucariótico.

Según un modo particular de realización, el vector según la invención comprende elementos de control de la expresión de polipéptidos; estos elementos de control se eligen preferiblemente entre (i) la secuencia promotora del gen ICBP90 según la invención que corresponde a la secuencia SEQ ID N°12 ; (ii) un polinucleótido cuya secuencia es complementaria de la secuencia SEQ ID N° 12 ; (iii) un polinucleótido cuya secuencia comprende al menos un 80% de identidad con un polinucleótido definido en (i) o (ii); (iv) un polinucleótido que hibrida en condiciones de alto rigor con la secuencia polinucleotídica definida en (i), (ii) o (iii). Las herramientas informáticas a disposición del especialista en la materia le permiten identificar fácilmente las secuencias reguladoras promotoras necesarias y suficientes para el control de la expresión génica, especialmente las secuencias TATA, CCAAT, GC, así como las secuencias reguladoras estimulantes ("potenciadoras") o inhibitoras ("silenciadoras") que controlan la expresión en cis de los genes según la invención.

La invención comprende además células hospedadoras, especialmente células eucarióticas y procarióticas, caracterizadas porque comprenden los vectores según la invención. Preferiblemente, las células hospedadoras se transforman en condiciones que permiten la expresión de un polipéptido recombinante según la invención. El hospedador celular puede elegirse entre las células bacterianas (Olins y Lee, 1993), o igualmente células animales, particularmente los cultivos de células de mamíferos (Edwards y Aruffo, 1993), o igualmente células de insecto en las que se puedan utilizar los procedimientos que aplican baculovirus por ejemplo (Luckow, 1993). Estas células pueden obtenerse mediante la introducción en las células hospedadoras de una secuencia nucleotídica insertada en un vector tal como se define anteriormente, y el cultivo entonces de dichas células en condiciones que permitan la replicación y/o expresión de la secuencia nucleotídica transfectada.

La invención se refiere igualmente a un procedimiento de preparación de un polipéptido caracterizado porque se aplica un vector según la invención. Más particularmente, la invención hace referencia a un procedimiento de preparación de un polipéptido recombinante caracterizado porque se cultivan células transformadas según la invención en condiciones que permitan la expresión de dicho polipéptido recombinante y que se recupere dicho polipéptido recombinante.

El polipéptido según la invención es susceptible de obtenerse según un procedimiento de la invención y según técnicas de producción de polipéptidos recombinantes conocidas por el especialista en la materia. La presente invención se refiere pues a un polipéptido recombinante susceptible de obtenerse mediante el procedimiento presentado anteriormente. En este caso, la secuencia de ácido nucleico utilizada se dispone bajo el control de

señales que permiten su expresión en un hospedador celular. Un sistema eficaz de producción de un polipéptido recombinante necesita disponer de un vector, por ejemplo de origen plasmídico o vírico, y de una célula hospedadora compatible. El vector debe comprender un promotor, señales de iniciación y de terminación de la traducción, así como regiones apropiadas de regulación de la transcripción. Se debe poder mantener de forma estable en la célula y puede poseer eventualmente señales particulares que especifiquen la secreción del polipéptido traducido. Estas diferentes señales de control se eligen en función del hospedador celular utilizado. A este respecto, las secuencias de ácido nucleico según la invención pueden insertarse en vectores de replicación autónoma dentro del hospedador elegido o de vectores integrativos del hospedador elegido. Dichos vectores se prepararán según procedimientos utilizados normalmente por el especialista en la materia, y los clones resultantes pueden introducirse en un hospedador apropiado mediante procedimientos estándares tales como, por ejemplo, transfección por precipitación con fosfato de calcio, lipofección, electroporación o choque térmico.

Los polipéptidos recombinantes obtenidos como se indica anteriormente pueden presentarse tanto en forma glucosilada como no glucosilada y pueden presentar o no la estructura terciaria natural.

Los polipéptidos obtenidos mediante síntesis química y que pueden comprender aminoácidos no naturales correspondientes a dichos polipéptidos recombinantes están igualmente comprendidos en la invención. Los péptidos según la invención pueden prepararse igualmente mediante técnicas clásicas en el campo de la síntesis de péptidos. Esta síntesis puede realizarse en solución homogénea o en fase sólida.

Los procedimientos de purificación de polipéptido recombinante utilizados son conocidos por el especialista en la materia. El polipéptido recombinante puede purificarse a partir de lisados y extractos celulares, del sobrenadante del medio de cultivo, mediante procedimientos utilizados individualmente o en combinación, tales como fraccionamiento, procedimientos de cromatografía, técnicas de inmunoafinidad con la ayuda de anticuerpos monoclonales o policlonales específicos, etc.

Una variante preferida consiste en producir un polipéptido recombinante fusionado con una proteína "portadora" (proteína quimérica). La ventaja de este sistema es que permite una estabilización y una disminución de la proteólisis del producto recombinante, un aumento de la solubilidad en el transcurso de la renaturalización *in vitro* y/o una simplificación de la purificación cuando la pareja de fusión posee afinidad por un ligando específico.

La invención se refiere igualmente a un anticuerpo monoclonal o policlonal y a sus fragmentos, caracterizados porque se ligan específicamente con un polipéptido elegido entre las SEQ ID NO: 2, 4 o 8. Los anticuerpos quiméricos, anticuerpos humanizados y anticuerpos monocatenarios forman parte igualmente de la invención. Los fragmentos de anticuerpo según la invención son preferiblemente fragmentos Fab o F(ab')₂.

Los polipéptidos según la invención permiten preparar anticuerpos monoclonales o policlonales. Los anticuerpos monoclonales podrán prepararse ventajosamente a partir de hibridomas según la técnica descrita por Kohler y Milstein en 1975. Los inventores han empleado esta técnica para obtener un hibridoma productor de un nuevo anticuerpo monoclonal altamente específico de un epítipo de la proteína ICBP90.

Los anticuerpos policlonales podrán prepararse, por ejemplo, mediante inmunización de un animal, particularmente un ratón, con un polipéptido según la invención asociado a un coadyuvante de la respuesta inmunitaria, y después purificación de los anticuerpos específicos contenidos en el suero de los animales inmunizados en una columna de afinidad sobre la que se ha fijado previamente el polipéptido que ha servido de antígeno. Los anticuerpos policlonales según la invención pueden prepararse también mediante purificación en columna de afinidad sobre la que se ha inmovilizado anteriormente un polipéptido según la invención.

Se pueden citar igualmente anticuerpos monoclonales específicos de la proteína ICBP90 humana y capaces de inhibir la interacción entre ICBP90 y la secuencia de ADN con la que se liga específicamente la proteína ICBP90. Según otro modo de realización, el anticuerpo monoclonal específico de la proteína ICBP90 humana es capaz de inhibir la interacción entre ICBP90 y las proteínas con las que interacciona ICBP90, siendo preferiblemente dichas proteínas la ICBP90 misma o proteínas del complejo de transcripción. Se pretende designar como proteínas del complejo de transcripción todas las proteínas que intervienen en la reacción de transcripción, sea iniciación, elongación o terminación de la transcripción.

Los anticuerpos de la invención podrán marcarse igualmente de la misma manera que se describe anteriormente para las sondas nucleicas de la invención, y preferiblemente con un marcador de tipo enzimático, fluorescente o radiactivo.

Por otro lado, además de su utilización para la purificación de polipéptidos, los anticuerpos de la invención, particularmente los anticuerpos monoclonales, pueden utilizarse igualmente para la detección de estos polipéptidos en una muestra biológica.

Constituyen así un medio de análisis de la expresión de polipéptido según la invención, por ejemplo mediante inmunofluorescencia, marcaje con oro o inmunocombinados enzimáticos.

Más generalmente, los anticuerpos de la invención pueden aplicarse ventajosamente en cualquier situación en que deba observarse la expresión de un polipéptido según la invención, y más particularmente en inmunocitoquímica, inmunohistoquímica o en experimentos de "transferencia Western".

5 Así, la invención se refiere a un procedimiento de detección y/o valoración de un polipéptido según la invención en una muestra biológica, caracterizado porque comprende las etapas siguientes de poner en contacto la muestra biológica con los anticuerpos según la invención y determinar entonces el complejo antígeno-anticuerpo formado. Este procedimiento puede utilizarse en inmunocitoquímica para la localización celular del polipéptido según la invención y en inmunohistoquímica para evaluar la proliferación celular.

10 Entra igualmente dentro del marco de la invención un estuche para la detección y/o valoración de un polipéptido según la invención en una muestra biológica, caracterizado porque comprende los elementos siguientes: (i) un anticuerpo monoclonal o policlonal tal como se describe anteriormente; (ii) llegado el caso, reactivos para la constitución del medio propicio para la reacción inmunológica; (ii) reactivos que permitan la detección de los complejos antígeno-anticuerpo producidos por la reacción inmunológica. Este kit es especialmente útil para la realización de experimentos de transferencia Western; estos permiten estudiar la regulación de la expresión del polipéptido según la invención a partir de tejidos o células. Este kit es igualmente útil en experimentos de inmunoprecipitación para determinar especialmente las proteínas que interaccionan con el polipéptido según la invención.

15 Puede aplicarse cualquier procedimiento clásico para realizar dicha detección y/o valoración. A modo de ejemplo, un procedimiento preferido pone en juego procesos inmunoenzimáticos según la técnica ELISA, mediante inmunofluorescencia o radioinmunología (RIA) o equivalente.

Se puede citar igualmente un procedimiento de detección y/o valoración de ácido nucleico en una muestra biológica, caracterizado porque comprende las etapas siguientes: (i) aislamiento de ADN a partir de la muestra biológica a analizar, u obtención de un ADNc a partir del ARN de la muestra biológica; (ii) amplificación específica del ADN que codifica el polipéptido según la invención con la ayuda de cebadores; (iii) análisis de los productos de amplificación.

25 Un estuche útil para la detección y/o valoración de un ácido nucleico en una muestra biológica mencionada anteriormente comprende los elementos siguientes: (i) un par de cebadores nucleicos según la invención, (ii) los reactivos necesarios para efectuar una reacción de amplificación de ADN, y eventualmente (iii) un componente que permita verificar la secuencia del fragmento amplificado, más particularmente una sonda según la invención.

30 Otro procedimiento de detección y/o valoración de ácido nucleico en una muestra biológica comprende las etapas siguientes: (i) la puesta en contacto de una sonda según la invención con una muestra biológica; (ii) la detección y/o valoración del híbrido formado entre dicha sonda y el ADN de la muestra biológica.

35 Un estuche útil para esta detección y/o valoración de ácido nucleico en una muestra biológica comprende los elementos siguientes: (i) una sonda según la invención, (ii) los reactivos necesarios para la aplicación de una reacción de hibridación, llegado el caso, (iii) un par de cebadores según la invención, así como los reactivos necesarios para una reacción de amplificación del ADN.

La invención se refiere particularmente a los procedimientos según la invención y descritos anteriormente para la detección y el diagnóstico de proliferación celular, y más particularmente de proliferación celular de origen canceroso.

40 La invención se refiere igualmente a un procedimiento de cribado de ligandos susceptibles de afectar a la actividad de transcripción de un gen cuyo promotor comprende las secuencias CCAAT y/o CCAAT inversa susceptibles de ligarse con un polipéptido según la invención, estando caracterizado dicho procedimiento porque comprende las etapas siguientes de puesta en contacto de dicho polipéptido y uno o varios ligandos potenciales en presencia de los reactivos necesarios para la aplicación de una reacción de transcripción o de detección y/o la medida de la actividad transcripcional. Es igualmente uno de los objetos de la invención proporcionar un kit o un estuche para el cribado de ligandos susceptibles de afectar a la actividad transcripcional de un gen cuyo promotor comprende las secuencias CCAAT y/o CCAAT inversa susceptibles de ligarse con un polipéptido según la invención, caracterizado porque comprende los elementos siguiente: (i) un polipéptido según la invención; (ii) un ligando; (iii) los reactivos necesarios para la aplicación de una reacción de transcripción.

50 El polipéptido ICBP90 según la invención presenta una función de receptor nuclear. Por receptor nuclear, se pretende designar un polipéptido que posea las propiedades esenciales de los receptores nucleares de hormonas. Esta superfamilia de genes contiene entre otros los receptores nucleares de ácido retinoico (RAR, RXR,...), los receptores nucleares de hormonas esteroideas (glucocorticoides, mineralocorticoides, progesterona, andrógeno, estrógeno) y los receptores nucleares de hormonas tiroideas (hormona T3). Es pues igualmente uno de los objetos de la presente invención proporcionar un procedimiento de cribado de ligandos susceptibles de afectar a la función de "receptor nuclear" del polipéptido según la invención. Dicho procedimiento comprende las etapas de:

- 55 a) puesta en contacto del polipéptido de la invención con uno o varios ligandos potenciales en presencia de los reactivos necesarios;

b) detección y/o medida de la actividad transcripcional de un gen cuyo promotor comprende secuencias nucleotídicas con las que son susceptibles de ligarse el polipéptido de la invención. Preferiblemente, dichas secuencias nucleotídicas son secuencias CCAAT y/o CCAAT inversa (ICB).

5 Las técnicas de detección y/o medida de la actividad transcripcional son conocidas por el especialista en la materia. Conviene citar especialmente las tecnologías de transferencia Northern y PCR-TI, que pueden aplicarse con los polinucleótidos de la invención utilizados respectivamente como sonda o como cebador.

10 Se pretende definir como ligando todos los compuestos susceptibles de interaccionar con el polipéptido según la invención para formar un complejo susceptible de afectar a la actividad transcripcional, es decir, de aumentar, disminuir, modular o anular la transcripción de un gen bajo el control de un promotor que contiene una secuencia de ADN con la que se liga el polipéptido de la invención.

15 Dicho ligando es pues susceptible de tener actividad agonista o antagonista. Entre los ligandos según la invención, conviene citar las moléculas biológicas que interaccionan con el polipéptido según la invención, así como todos los compuestos químicos de síntesis. Entre los ligandos, conviene citar igualmente el anticuerpo según la invención, así como un oligonucleótido que presente una identidad de secuencia con la secuencia nucleotídica CCAAT y/o CCAAT inversa; dicho ligando es susceptible de constituir un inhibidor del polipéptido según la invención.

La invención hace referencia igualmente al ligando susceptible de obtenerse mediante los procedimientos de cribado precedentes.

20 Se pretende definir igualmente como ligando cualquier compuesto susceptible de ligarse con la secuencia de ADN de unión del polipéptido según la invención. Dicho ligando constituye un inhibidor competitivo del polipéptido según la invención para el enlace con la secuencia de ADN.

De preferencia, la muestra biológica según la invención en la que se realiza la detección y valoración está constituida por un fluido corporal, por ejemplo, un suero humano o animal, sangre, saliva, moco pulmonar o biopsias. Entra igualmente en la definición de muestra biológica de la invención el líquido biológico resultante de un lavado broncoalveolar obtenido igualmente en los análisis de diagnóstico de cánceres de las vías respiratorias profundas.

25 Según otro aspecto, la invención se refiere a un compuesto caracterizado porque se elige entre un anticuerpo, polipéptido, ligando, polinucleótido, oligonucleótido o vector según la invención a modo de medicamento, y especialmente como principios activos de medicamento; estos compuestos estarán preferiblemente en forma soluble, asociados a un vehículo farmacéuticamente aceptable. Se pretende designar como vehículo farmacéuticamente aceptable cualquier tipo de vehículo empleado habitualmente en la preparación de
30 composiciones inyectables, es decir, un diluyente o agente de suspensión tal como una solución salina isotónica o tamponada. Preferiblemente estos compuestos se administrarán por vía sistémica, particularmente por vía intravenosa, intramuscular, intradérmica u oral. Sus modos de administración, posologías y formas galénicas óptimas pueden determinarse según los criterios tenidos generalmente en cuenta en el establecimiento de un tratamiento adaptado a un paciente como, por ejemplo, la edad o el peso corporal del paciente, la gravedad de su estado general, la tolerancia al tratamiento y los efectos secundarios comprobados, etc.
35

Según otro aspecto, la invención se refiere a un compuesto caracterizado porque se elige entre un polipéptido, polinucleótido, polinucleótido anticodificante, anticuerpo, vector, célula o ligando según la invención a modo de medicamento, y especialmente como principios activos de medicamento; estos compuestos estarán preferiblemente
40 en forma soluble, asociados a un vehículo farmacéuticamente aceptable. Se pretende designar como vehículo farmacéuticamente aceptable cualquier tipo de vehículo empleado habitualmente en la preparación de composiciones inyectables, es decir, un diluyente o agente de suspensión tal como una solución salina isotónica o tamponada. Preferiblemente, estos compuestos se administrarán por vía sistémica, en particular por vía intravenosa, intramuscular, intradérmica u oral. Sus modos de administración, posologías y formas galénicas óptimas pueden determinarse según los criterios tomados generalmente en cuenta en el establecimiento de un tratamiento adaptado
45 a un paciente como, por ejemplo, la edad o el peso corporal del paciente, la gravedad de su estado general, la tolerancia al tratamiento y los efectos secundarios comprobados, etc. Cuando el agente es un polipéptido, antagonista, ligando o polinucleótido, por ejemplo una composición anticodificante o un vector, puede introducirse en los tejidos o células hospedadores mediante un cierto número de formas, incluyendo infección vírica, microinyección o fusión de vesículas. Se puede utilizar igualmente la inyección a chorro para una administración intramuscular como se describe por Furth *et al.* (1992). Se puede depositar el polinucleótido sobre micropartículas de oro, y suministrarlo
50 por vía intradérmica con ayuda de un dispositivo de bombardeo de partículas o una "pistola de genes" como se describe en la bibliografía (véase, por ejemplo, Tang *et al.* (1992) en que los microproyectiles de oro se revisten con el polinucleótido de la invención, preferiblemente el polinucleótido anticodificante de la invención, y después se bombardean en las células de piel.

55 El compuesto según la invención se utiliza para la preparación de un medicamento destinado a modular, aumentar o disminuir la proliferación celular.

La invención hace referencia igualmente a una composición farmacéutica para el tratamiento preventivo y curativo del cáncer, caracterizada porque contiene una cantidad terapéuticamente eficaz de un compuesto según la

invención y un vehículo farmacéuticamente aceptable. Según un modo de realización preferido, la composición farmacéutica se caracteriza porque contiene un anticuerpo según la invención como agente orientador, conjugado con al menos un agente seleccionado entre el grupo de agentes antiproliferativos, antineoplásicos o citotóxicos. Estos agentes son radioisótopos o entidades no isotópicas. La conjugación del anticuerpo de la presente invención con un agente antiproliferativo, antineoplásico o citotóxico puede utilizarse para detener el desarrollo de cánceres y para inducir la regresión y/o eliminación de la masa tumoral. Preferiblemente, se introduce el anticuerpo o fragmento de anticuerpo así conjugado en el paciente aquejado de cáncer y se suministra a los sitios tumorales por vía oral o parenteral en un líquido transportador farmacéuticamente aceptable tal como una solución de sal fisiológica. Como alternativa, puede perfundirse una solución o suspensión de anticuerpo o fragmento de anticuerpo conjugado con un agente directamente en el tejido epitelial maligno, utilizándose este procedimiento preferiblemente en el caso en que el cáncer no haya metastatizado.

Los radioisótopos preferidos conjugados con anticuerpos monoclonales empleados para la terapia son radioisótopos emisores de rayos gamma, y preferiblemente yodo¹³¹, itrio⁹⁰, oro¹⁹⁹, paladio¹⁰⁰, cobre⁶⁷, bismuto²¹⁷ y antimonio²¹¹. Los radioisótopos emisores de rayos beta y alfa pueden utilizarse igualmente para la terapia. Las entidades no isotópicas conjugadas con los anticuerpos monoclonales empleadas para la terapia son múltiples y variadas; se pueden citar: (i) antimetabolitos tales como los agentes antifolato, metotrexato; (ii) análogos de purinas y pirimidinas (mercaptapurina, fluorouracilo, 5-azacitidina); (iii) antibióticos; (iv) lectinas (ricina, abrina) y (v) toxinas bacterianas (toxina diftérica).

El anticuerpo según la invención puede utilizarse igualmente como agente orientador para orientar células citotóxicas tales como linfocitos T humanos, monocitos o linfocitos NK al sitio del tumor metastásico o no. Las células citotóxicas pueden unirse al anticuerpo mediante el receptor Fc situado en la superficie de estas células o mediante un anticuerpo intermedio que presenta una doble especificidad por ejemplo; dichos anticuerpos biespecíficos para orientación a células cancerosas pueden producirse fusionando una célula inmunitaria productora del anticuerpo de la presente invención o el hibridoma de la presente invención con una célula productora de un anticuerpo dirigido contra la célula citotóxica diana. Los anticuerpos biespecíficos pueden producirse igualmente mediante acoplamiento químico de dos anticuerpos que tienen la especificidad deseada. El anticuerpo según la invención permite igualmente orientar vehículos de suministro, agentes antiproliferativos, antineoplásicos o citotóxicos al sitio del tumor metastásico o no. Se pretende designar como vehículos de suministro los liposomas y las partículas víricas. En ciertos casos, se podrán prever elementos orientadores que aseguren una expresión específica de ciertos tejidos o células de forma que se puedan limitar las zonas de expresión de los polipéptidos según la invención.

La invención se refiere igualmente a un producto que comprende al menos un compuesto según la invención y al menos un agente anticanceroso como producto de combinación para utilización simultánea, separada o extendida en el tiempo en terapia anticancerosa.

Por último, la invención se refiere a una composición para la detección, localización y formación de imágenes de cánceres, que comprende un anticuerpo según la invención tal que el anticuerpo esté marcado directa o indirectamente con un marcador generador de señal seleccionado entre los isótopos radiactivos y las entidades no isotópicas tales como se definen anteriormente. La invención tiene igualmente como objeto un procedimiento de detección, localización y formación de imágenes del cáncer que comprende las etapas de (i) inyección parenteral en un ser humano de una composición según la invención; (ii) acumulación después de un tiempo suficiente al nivel de las células cancerosas del anticuerpo marcado, y penetración entonces del anticuerpo marcado en el interior de dichas células, sin que dicho anticuerpo se ligue sustancialmente con células normales, (iii) detección de la señal mediante un detector de señal y (iv) conversión de la señal detectada en una imagen de las células cancerosas.

Aparecerán otras características y ventajas de la invención como consecuencia de la descripción con los ejemplos representados a continuación. En estos ejemplos, se referirá a las figuras siguientes.

Figura 1: Expresión de la proteína ICBP90 en células HeLa (células tumorales) y en fibroblastos pulmonares humanos en cultivo primario (células no tumorales).

Se ha realizado la detección de la proteína endógena ICBP90 en extractos totales de proteínas de células HeLa en confluencia (carril 1) o en proliferación (carril 2) y en extractos totales de proteínas de fibroblastos pulmonares humanos en cultivo primario en confluencia (carril 3) o en proliferación (carril 4). Después de migrar por gel de poliacrilamida al 8% en presencia de SDS, se transfieren las proteínas a membrana de nitrocelulosa mediante electrotransferencia. Se realiza el revelado con la ayuda del anticuerpo 1RC1C-10 diluido 1/4000 (concentración inicial 2 mg/ml) y un anticuerpo secundario acoplado con fosfatasa alcalina y dirigido contra las cadenas pesadas de anticuerpo de ratón. En los carriles correspondientes a los extractos de células HeLa, se observa una banda mayoritaria de 97 kDa; para las células HeLa en proliferación, aparecen bandas suplementarias de tamaños inferiores a 97 kDa (carril 2). En los fibroblastos pulmonares humanos en confluencia, no se expresa la proteína endógena (carril 3) y aparece cuando las células empiezan a proliferar (carril 4). Estas observaciones sugieren que la proteína endógena ICBP90 es un marcador de la proliferación celular para células normales (fibroblastos), mientras que para las células tumorales sería un marcador para cualquier estado celular.

Figura 2: Inmunoprecipitación de la proteína endógena

Se realiza la inmunoprecipitación con extractos proteicos totales de células MOLT-4. Se fija el anticuerpo 1RC1C-10 sobre bolas de proteína G-Sepharose y se pone en contacto entonces con los extractos proteicos durante 2 horas a temperatura ambiente. Después del lavado, se precipitan los complejos de bolas/1RC1C-10/proteína mediante centrifugación, se analizan mediante migración en gel de poliacrilamida al 8% en presencia de SDS y se transfieren entonces a membrana de nitrocelulosa y se revelan como se indica en la figura 1. Se observa una banda única de 97 kDa, así como una banda de 45 kDa que corresponde a la cadena pesada de 1RC1C-10.

Figura 3: Localización nuclear de la proteína endógena

Se han utilizado células HeLa para examinar la expresión endógena de la proteína ICBP90 *in situ* con la ayuda del anticuerpo 1RC1C-10 y de un anticuerpo secundario anti-ratón acoplado con el fluorocromo CY3. El marcaje está localizado exclusivamente en el núcleo. El nucleolo y citoplasma no están marcados.

Figura 4: Expresión de ICBP59 endógena en células en proliferación

Se ha observado la proteína endógena en cortes de parafina de apéndice humano. Después del desparafinado y pretratamiento por calentamiento en tampón ácido (desvelado de los sitios antigénicos), se incuban los cortes durante 16 horas con el anticuerpo 1RC1C-10 diluido 1/10000 (concentración inicial de 2 mg/ml). Se realiza el revelado mediante la puesta en contacto con un anticuerpo secundario biotinilado y entonces incubación con el complejo estreptavidina-peroxidasa. Se realiza igualmente una contratinción de los núcleos con hematoxilina de Harris. El marcaje por 1RC1C-10 está localizado esencialmente en las zonas de proliferación celular. Las células marcadas se encuentran en las criptas glandulares (CB) así como en las zonas germinativas (ger).

Figura 5: Expresión de ICBP-59 en diversos tejidos humanos

Se ha evaluado el nivel de expresión de ARNm correspondiente a ICBP59 en una transferencia puntual de ARN que comprende 50 tejidos humanos diferentes. Se hibridó la transferencia durante 16 horas a 68°C con una sonda de ADNc radiactivo (³²P) de 679 pb en una solución de hibridación ExpressHyb (Clontech). Después de los lavados, se realiza un revelado por autorradiografía (exposición durante una semana a 80°C). Los tejidos que presentan el nivel más alto de expresión son timo adulto y fetal así como médula ósea adulta e hígado fetal.

Figura 6: Secuencia nucleotídica de ICBP90

El ADNc que codifica ICBP90 comprende 2379 pb. Las porciones de secuencia indicadas en negrita son aquellas que no aparecen en las bases de datos de EST humanas (human dbest). Las otras partes de la secuencia existen en diversas EST:

de 1 a 325 : EST n° AI083773.

de 367 a 865 : EST n° AA811055.

de 940 a 1857 : EST n° AA488755, EST n° AA129794 y EST n° AA354253.

Figura 7: Secuencia proteica de ICBP90

Se deduce la secuencia aminoacídica de ICBP90 mediante la traducción de la secuencia nucleotídica de la figura 6. ICBP90 comprende 793 residuos y presenta un peso molecular teórico de 89,758 kDa. El p*K*_i es de 7,7. Los aminoácidos indicados en negrita corresponden a ICBP-59.

Figura 8: Detección de ICBP90 en los sueros de pacientes que tienen marcadores séricos elevados de tumores sólidos

Se diluye un volumen de 2 µl de suero de cada paciente en 1 ml de tampón PBS (solución salina tamponada con fosfato 1X) que contiene 0,1% de Tween 20, seguido de diluciones crecientes realizadas en el mismo tampón como se indica en la figura. Se filtra una muestra de 0,5 ml de cada dilución por membrana de nitrocelulosa con la ayuda de un aparato "Slot Blot BioRad". Se bloquea la membrana en presencia de tampón PBS (que contiene 0,1% de Tween 20 y 5% de leche) durante 1 hora a temperatura ambiente. Se revela la proteína ICBP90 con la ayuda del anticuerpo 1RC-1C10 (1 ng/ml) y el anticuerpo secundario (anti-ratón acoplado con peroxidasa diluido 1/5000). Se revelan las bandas por quimioluminiscencia mediante exposición durante 10 segundos a una película X-MAT (Kodak).

Figura 9: Organización estructural del gen ICBP90

A. Exones representados por recuadros: los recuadros grises representan los exones codificantes; los recuadros blancos representan los exones no codificantes. El tamaño de los exones se indica en pb en cada recuadro y los nombres de los exones se indican por encima de los recuadros. Los intrones se mencionan de manera esquemática por líneas finas y su tamaño aproximado se indica en pb. Se indican un supuesto sitio de iniciación de la

transcripción y una señal de consenso de poliadenilación. El ATG es el codón de inicio de la traducción y TGA el codón de terminación de la traducción.

- 5 **B.** Secuencia de la región 5' flanqueante del gen ICBP90 (Seq ID N° 12) (número de acceso a Genbank N° AF220.226, depositado el 30 de diciembre de 1999). Los exones están en mayúsculas y los intrones en minúsculas. El codón de inicio ATG está en mayúsculas en negrita, y las secuencias ricas en GC (GC) y las secuencias CCAAT (CB) se representan en minúsculas en negrita.

Figura 10: Análisis del promotor de ICBP90

Se fusionaron las secuencias del promotor de ICBP90 con la secuencia del gen "informador" CAT en el vector pBLCAT2, que se transfirió a continuación en células COS-1.

- 10 Se representa una representación esquemática de estas construcciones a la izquierda, refiriéndose el número a los nucleótidos en dirección 5' del codón de iniciación. Las actividades CAT relativas de los extractos celulares correspondientes a la actividad del promotor TK mínimo se expresan en porcentaje (a partir de tres experimentos de transfección independientes) y se indican a la derecha.

Figura 11: Análisis por transferencia Northern y transferencia Western de la expresión de ICBP90.

- 15 **A.** Se efectuó la hibridación Northern en una membrana de transferencia Northern cuyos depósitos de ARN provienen de líneas celulares cancerosas de diferentes órganos. Se marca con digoxigenina una sonda específica de ICBP90, sintetizada mediante PCR, y se utiliza para la detección del ARNm de ICBP90. Se mencionan los tamaños de los ARNm a la derecha de la línea 7.

- 20 Las líneas 1 a 7 contienen ARN procedentes respectivamente de la línea leucémica promielocítica HL-60, de Hela 53, de células K562 de leucemia mielogenosa crónica, de células de leucemia linfoblástica MOLT-4, de células Raji de linfoma de Burkitt, de células SW480 de adenocarcinoma colorrectal y de células A549 de carcinoma pulmonar.

El histograma muestra las tasas de expresión de los ARNm correspondientes a las bandas de 5,1 kb y 4,3 kb expresados en porcentaje de tasa de expresión del ARNm de 5,1 kb de células HL-60 (línea 1, figura 11A).

B. Análisis de transferencia Western de la expresión de ICBP90 en células MOLT-4 y Hela.

- 25 Se prepararon lisados de células totales de células Hela y MOLT-4 en proliferación. Se analizó la expresión de ICBP90 en transferencia Western utilizando el anticuerpo 1RC1C-10.

EJEMPLO 1: DETERMINACIÓN DE UNA NUEVA PROTEÍNA DE UNIÓN CON LA SECUENCIA ICB

1.1. Construcción informadora para cribado del banco

- 30 El sistema de híbrido simple es una técnica potente que permite detectar *in vivo* en levadura la interacción de proteínas con secuencias de ADN específicas cribando los bancos de ADNc. Esto permite evaluar directamente el ADNc correspondiente de la proteína a ligar. Varios estudios han permitido identificar la nueva proteína con este procedimiento. Estos procedimientos describen muy bien los protocolos utilizados (Inouye *et al.*, 1994; Wang y Reed, 1993).

- 35 Brevemente, se sintetizaron los oligonucleótidos siguientes 5'-AATTC**GATTGGTTCTGATTGGTTCTGATTGGTTCTT**-3' y 5'-CTAGAAGA**CCAATCAGAACCAATCAGAACCAATCG**-3'. Se hibridan a continuación estos nucleótidos. Según las instrucciones del fabricante (Clontech, Palo Alto, CA), la construcción informadora diana posee tres copias en serie de la secuencia ICB2 (ICB2X3). Como se menciona anteriormente, se subraya una copia de ICB2 y las secuencias CCAAT se representan en negrita. Para determinar la especificidad de unión de las proteínas con la secuencia ICB, se sintetizaron e hibridaron los oligonucleótidos siguientes que contienen tres copias en serie de la secuencia GC1 (GC1X3) igualmente presentes en el promotor:

5'- AATTC**GGGGCGGGGCCGGGGCGGGGCCGGGGCGGGGCT**-3'

5'- CTAGAG**CCCCGCCCCGGCCCCGCCCCGGCCCCGCCCCGG**-3'

- 45 Se clonan los fragmentos de ADN diana resultantes en el poliligador de un plásmido integrativo pHISi-1 (Clontech) mediante ligamiento de los extremos cohesivos al nivel del sitio EcoRI y XbaI, en dirección 5' del promotor mínimo del gen *his3*. Se utiliza la cepa de levadura YM4271 (Clontech) para la transformación y se seleccionan las colonias de levadura que hayan integrado el plásmido en su genoma en medio sintético Dropout que no contiene histidina. Se aislaron dos clones: uno para ICB2 y otro para la secuencia GC 1.

1.2. Cribado del banco

Se utiliza un banco de ADNc de la línea celular Jurkat clonada en el sitio EcoRI del poliligador en dirección 3' de GAL4-AD del vector pGAD10 (Clontech) para el cribado según las instrucciones del fabricante. Se seleccionan los clones positivos y se cultivan entonces en medio selectivo desprovisto de histidina y leucina. Se recupera el ADN plasmídico de estos clones y se introduce mediante electroporación en bacterias *Escherichia coli* XL1-blue. Se realizó la secuenciación de los insertos sobre una matriz de ADN plasmídico purificada a partir de un cultivo de 1,5 ml utilizando un kit de minipreparación (Bio-Rad, Hercules, CA, EE.UU.). Se cribó un banco de ADNc de timo humano clonado en λ gt10 (Clontech) mediante hibridación en placa, para recuperar un ADNc que codifica la parte N-terminal de la proteína.

1.3. Descubrimiento de ICBP-59

Se secuenciaron los ADNc de cuatro clones que habían cumplido los criterios de selección del sistema híbrido simple y se analizaron los perfiles con la ayuda de bases de datos informáticos (Genbank, EMBL, PDB, Swissprot) con el fin de determinar la naturaleza de las proteínas codificadas. Dos clones corresponden a proteínas ribosómicas (hRS12 y hRS4), uno a una serina-treonina cinasa (STPLK-1) y el cuarto a una proteína humana de un peso molecular teórico de 59 kDa (calculado a partir de la secuencia traducida) y no catalogada.

Se clonaron los ADNc que codifican hRS4, hRS12 e ICBP-59, y obtenidos mediante digestión con EcoRI de clones positivos obtenidos en el vector pGAD10, en el sitio EcoRI del vector de expresión pGEX-4T-1 (Pharmacia). Se transformaron a continuación los ADN recombinantes en una cepa de *Escherichia coli* adaptada (BL21). Se utilizaron 500 ml de cultivo del clon seleccionado cuando se alcanzó una densidad óptica de 0,5. Se indujo la sobreexpresión de las proteínas de interés por IPTG (1 mM) durante 2 h a 37°C. El vector pGEX-4T-1 conduce a la obtención de grandes cantidades de proteínas en forma fusionada con glutation-S-transferasa (GST). Se purificaron a continuación las proteínas de fusión con glutation-S-transferasa (GST) utilizando bolas de Sepharose acoplada con glutation (Pharmacia), seguido de un corte durante la noche con trombina (0,05 U/ml) a 4°C (Pharmacia).

Para ensayar la capacidad de la proteína de peso molecular 59 kDa de ligarse específicamente con las secuencias ICB1 y/o ICB2, se marcaron tres copias en serie de ICB2 (ICB2X3, secuencias descritas anteriormente) al nivel terminal con fósforo 32 P utilizando la polinucleótido cinasa T4 (New England Biolabs) y [γ - 32 P]ATP (160 mCi/mmol, ICN Irvine, CA, EE.UU.). Para examinar la especificidad de la unión, se sintetizaron oligonucleótidos que contenían solamente una copia de la secuencia CCAAT:

ICB1: 5'-AGTCAGGG**ATTGG**CTGGTCTG-3';

5'- CAGACCAG**CCAAT**CCCTGACT-3'

ICB2: 5'-AAGCTACG**ATTGG**TTCTTCTG-3';

5'-CAGAAGA**CCAAT**CGTAGCTT-3'.

Se incubó la proteína ICBP-59 purificada (1 μ g) con 1 ng de oligonucleótido marcado en su extremo terminal con fósforo 32 P en glicerol al 12%, HEPES-NaOH 12 mM (pH 7,9), KCl 60 mM, Tris-HCl 4 mM (pH 7,9), 100 ng de BSA, DTT 0,6 mM y 100 ng de poli(dI/dC) en 20 μ l (Inouye *et al.*, 1994). Después de 30 minutos de incubación a temperatura ambiente, se carga la mezcla de reacción sobre geles de poli(acrilamida) al 6%. En experimentos competitivos, se añade la cantidad indicada de oligonucleótidos no marcados a la mezcla de reacción 10 min antes de la adición de proteínas. Para examinar las propiedades de unión de ICBP90 a la secuencia ICB2, se utiliza el mismo protocolo, con la diferencia de que el oligonucleótido marcado contiene solamente una copia de la secuencia CCAAT tal como se describe a continuación:

ICB2: 5'-ATAAAGGCAAGCTACG**ATTGG**TTCTTCTGGACGGAGAC-3'

5'-GTCTCCGTCCAGAAGA**CCAAT**CGTAGCTTGCCTTTTAT-3'.

Se estudia la especificidad de unión utilizando un nucleótido no marcado que contiene una secuencia GC del promotor del gen de la topoisomerasa IIa humana:

5'-GAATTCCGAGGGT**AAAGGGGCGGGG**TTGAGGCAGATGCCA-3'

5'-TGGCATCTGCCTCA**ACCCCGCCCC**TTTACCCTCGAATTC-3'.

Estos experimentos de retardo de la migración sobre gel de acrilamida han permitido determinar que la nueva proteína humana de 59 kDa es capaz de ligarse con una secuencia de ADN de tipo ICB, y esto de manera específica. Se ha denominado a esta proteína ICBP-59 (por Inverted CCAT Box Binding Protein en inglés, proteína de unión a la secuencia CCAAT invertida de 59 kDa).

EJEMPLO 2: CARACTERIZACIÓN DE LA PROTEÍNA ICBP90**2.1. Síntesis de anticuerpo**

Se sintetizan anticuerpos monoclonales de ratón en el laboratorio mediante inyección de la proteína ICBP-59 mediante procedimientos tradicionales (Brou *et al.*, 1993); la proteína se purificó anteriormente mediante un sistema de fusión GST. Se seleccionaron dos anticuerpos monoclonales 1RC1C-10 y 1RC1H-12 por su rendimiento para detectar la proteína endógena correspondiente a la proteína ICBP-59 a la vez en experimentos de transferencia Western y en experimentos de inmunocitoquímica. Antes de la utilización, se purifican los anticuerpos sobre columna de DEAE-celulosa (DE52, Whatmann) a partir de líquidos de ascitis.

2.2. Determinación de la proteína endógena por transferencia Western

Con el fin de detectar la proteína endógena correspondiente a ICBP-59, se ha utilizado en un primer momento 1RC1C-10 en la transferencia Western (0,4 µg/ml de anticuerpo monoclonal 1RC1C-10) sobre extractos nucleares de células HeLa en situación de proliferación y de confluencia (Figura 1). Se cultivan las células COS-1 y HeLa tal como se describe anteriormente (Brou *et al.*, 1993; Gaub *et al.*, 1998; Rochette-Egly *et al.*, 1997). Se cultivan las células MOLT-4 al aire al 100% en RPMI suplementado con 10% de suero fetal de ternero. Se preparan y cultivan fibroblastos pulmonares humanos en cultivo primario DMEM/F12 tal como se describe anteriormente (Kassel *et al.*, 1998). Se adquirieron los extractos nucleares de células Jurkat en Sigma, mientras que los de MOLT-4 y HL60 se prepararon tal como se describe precedentemente (Lavie *et al.*, 1999). Las células HeLa en fase de crecimiento y los fibroblastos pulmonares humanos se obtienen mediante la depleción de suero del cultivo durante 30 h, seguido de reintroducción durante 16 horas en suero fetal de ternero (v/v) al 10%. Se detiene la proliferación cuando la confluencia alcanza de 60 a 70%. Las células detenidas en confluencia (confluencia de 100%) de manera concomitante omitiendo la etapa de depleción de suero. Para estos dos tipos celulares, se preparan lisados celulares brutos recogiendo las células en PBS (solución salina tamponada con fosfato) seguido de una etapa de sonicación. Para los experimentos de inmunotransferencia, se cargan lisados de células totales y extractos nucleares sobre geles de poliacrilamida con SDS al 8% para realizar una electroforesis unidimensional. Se transfieren las proteínas sobre membranas de nitrocelulosa bloqueadas con un reactivo de bloqueo al 10% (Roche Molecular Biochemicals, Mannheim, Alemania) y se incuban con el anticuerpo monoclonal purificado (1RC1C-10) a la concentración de 0,5 µg/ml. Se utiliza un anticuerpo de cabra anti-ratón acoplado con fosfatasa alcalina (fragmentos Fab, Roche Molecular Biochemicals) a una dilución 1/2.500. Se detectan las señales utilizando cloruro de 4-nitroazul de tetrazolio/fosfato de 5-bromo-4-cloro-3-indolilo como sustrato.

Estos experimentos muestran que la proteína endógena presenta un peso molecular aparente de aproximadamente 97 kDa. Además, se observa que las formas de la proteína varían en función de la naturaleza tumoral o no tumoral de las células, así como del estado de confluencia o proliferación de las células. Efectivamente, en los carriles correspondientes a los extractos de células HeLa, se observa una banda mayoritaria de 97 kDa; para las células HeLa en proliferación, aparecen bandas suplementarias de tamaños inferiores a 97 kDa (carril 2). En los fibroblastos pulmonares humanos en confluencia, no se expresa la proteína endógena (carril 3) y aparece cuando las células empiezan a proliferar (carril 4). Estas observaciones sugieren que la proteína endógena ICBP90 es un marcador de la proliferación celular para células normales (fibroblastos), mientras que para las células tumorales sería un marcador en cualquier estado celular.

La utilización del anticuerpo monoclonal en experimentos de inmunoprecipitación con extractos de proteínas nucleares, seguido de una transferencia Western, conduce de la misma manera a la determinación de una proteína de 97 kDa (Figura 2).

Los resultados obtenidos en transferencia Western, tanto para los extractos de proteínas nucleares como para las inmunoprecipitaciones, muestran que la proteína de 59 kDa aislada con la ayuda del sistema híbrido simple solo constituye un fragmento de la proteína endógena humana correspondiente, en este caso el fragmento C-terminal a partir del residuo D263. Ha hecho falta entonces emprender un nuevo cribado del banco de ADNc.

2.3. Análisis de transferencia puntual de ARN de múltiples tejidos humanos

Con el fin de elegir un banco que dé las mayores posibilidades de aislar la proteína completa, se ha querido identificar un tejido humano que expresara el ARN mensajero (ARNm) correspondiente en gran cantidad. Con la ayuda de una sonda de ADNc que recubre una parte de la secuencia de ICBP59 y marcada con ³²P, se ha ensayado la expresión del ARNm de interés en 50 tejidos humanos diferentes en una transferencia puntual de ARN. Brevemente, se sintetizó una sonda de 678 pares de bases correspondiente a la secuencia aminoacídica 269 a 500 de ICBP90 mediante PCR utilizando la polimerasa Taq (Sigma, St Louis, MO, EE.UU.). Se purifica la sonda marcada mediante cebado aleatorio utilizando dCTP-³²P sobre columnas de Sephadex G50 (Pharmacia, Uppsala, Suecia).

Se hibrida una transferencia puntual de ARN de múltiples órganos que contiene ARN poli(A)+ de 50 tejidos humanos diferentes durante 20 horas en condiciones de alto rigor en un medio ExpressHyb (Clontech) a 68°C con una sonda marcada con ³²P. Se realizan lavados de alto rigor con 0,1 x SSC, 0,1% SDS de a 68°C (De Vries *et al.* 1996).

Los resultados obtenidos (figura 5) muestran que los tejidos que expresan más el ARNm de la proteína ICBP-59 son timo adulto y fetal, así como médula ósea adulta e hígado fetal. Para aislar la proteína entera, se hace pues la selección de un banco de ADNc de timo adulto.

2.4. Cribado del banco y clonación de ICBP90

- 5 El cribado del banco ha permitido obtener varios clones de aproximadamente 4000 pares de bases (pb) que comprenden un marco de lectura abierto de 2379 pb (Fig. 6). Esta secuencia codifica una proteína de 793 aminoácidos (Fig.7) cuyo peso molecular teórico (calculado a partir de la secuencia traducida) es de 89,758 kDa. Se ha denominado a esta proteína ICBP90 (por Inverted CCAT Box Binding Protein en inglés, proteína de unión a la secuencia CCAAT invertida de 90 kDa) por analogía con la denominación utilizada para la proteína inicial de 59 kDa.
- 10 Se sintetizó el ADNc de ICBP90 (2379 pb) mediante PCR utilizando la ADN polimerasa Deep Vent (New England Biolabs, Beverly, MA, EE.UU.) y los oligonucleótidos utilizados en el transcurso de esta reacción de PCR estaban cercanos al sitio de EcoRI. Se subclonó a continuación el producto de reacción en un vector pGEX-4T-1 (Pharmacia) para la expresión de la proteína de fusión GST en BL21. Se induce la sobreexpresión por IPTG (1 mM) durante 4 h a 25°C. Se purifica a continuación la proteína ICBP90.

15 2.5. Inmuncitoquímica e inmunohistoquímica.

Se aplicó igualmente la observación directa de la proteína ICBP90 sobre células y tejidos.

- Se transfectaron células COS-1 como se describe anteriormente (Brou *et al.*, 1993; Gaub *et al.*, 1998) con el vector pSG5 (Stratagene, La Jolla, CA) en el que se subclonó el ADNc de ICBP90 (2379 pb) en el sitio de restricción EcoRI. Se sintetiza el ADNc mediante reacción de polimerización en cadena (PCR) utilizando la polimerasa Deep Vent (New England Biolabs) y los oligonucleótidos que flanquean el sitio de restricción EcoRI. Se verifica la construcción plasmídica mediante secuenciación. Se realiza el inmunomarcaje de las células HeLa y las células COS-1 transfectadas tal como se describe anteriormente (Brou *et al.*, 1993) respectivamente con los anticuerpos monoclonales 1RC1C-10 y 1RC1H-12. Se realiza un marcate indirecto con inmunoperoxidasa de ICBP90 y de topoisomerasa II α como se describe anteriormente (Rio *et al.*, 1987, Devys *et al.*, 1993). Se embeben apéndices humanos en parafina y se fijan en formalina tamponada al 10% (Sigma). Se incuban los cortes en serie (3 μ m) durante una noche a temperatura ambiente con el anticuerpo 1RC1C-10 y con el anticuerpo anti-topoisomerasa II α (NeoMarkers, Union City, CA, EE.UU.). Se visualizan los anticuerpos ligados específicamente mediante un complejo que utiliza estreptavidina-biotina (procedimiento LAB/LSAB, Dako LSAB2 System Kit; DAKO, Carpinteria, CA, EE.UU.).
- 20
- 25
- 30 En la inmuncitoquímica, el anticuerpo 1RC1C-10 marca el núcleo de células HeLa, mientras que el nucleolo y el conjunto del citoplasma no están marcados (Figura 3). En inmunohistoquímica, los cortes de apéndice humano en parafina muestran un marcate localizado esencialmente en zonas de proliferación celular (Figura 4). Efectivamente, las células marcadas están alojadas en las criptas glandulares (CG), así como en las zonas germinativas (ger). Se obtiene un marcate idéntico cuando se utiliza un anticuerpo anti-topoisomerasa II α , que es una enzima expresada únicamente en células en proliferación (resultados no ilustrados).
- 35

2.6. Investigaciones con BLAST y predicciones de dominios

Se realizaron los estudios con BLAST en línea a partir de informaciones del National Center for Biotechnology Information del National Institute of Health (Bethesda, MD, EE.UU.). Se utilizan SCANPROSITE y PROFILESCAN para el análisis proteico (Infobiogen, Villejuif, Francia).

- 40 ICBP90 comprende un dominio "de tipo ubiquitina" en sus primeros 80 aminoácidos, dos sitios de localización nuclear potenciales en la parte C-terminal y dos dominios en dedo de cinc ("zinc-finger"), en que uno estaría implicado en la unión al ADN y el otro en las interacciones proteína-proteína. Están igualmente presentes varios sitios potenciales de fosforilación para proteína cinasa C, caseína cinasa II, así como una tirosina cinasa.

- 45 La producción y purificación de ICBP90 con la ayuda del sistema de fusión GST (mismo procedimiento que el utilizado para ICBP-59) ha permitido finalmente ensayar la capacidad de la proteína completa de ligarse con secuencias de ADN de tipo ICB. Su comportamiento es idéntico en todos los puntos al observado para ICBP-59.

- 50 En definitiva, se ha aislado una nueva proteína humana que se ha denominado ICBP90 por las razones mencionadas anteriormente. Su peso molecular teórico es 89,758 kDa y su peso molecular aparente sobre gel de acrilamida es 97 kDa. Esta proteína no está solo localizada exclusivamente en el núcleo de células humanas, sino que presenta igualmente la capacidad de ligarse con secuencias de ADN de manera específica, en este caso con secuencias de tipo CCAAT. Por estas razones, se cree que ICBP90 tiene la posibilidad de modular la expresión de genes cuyo promotor está provisto de secuencias CCAAT, eventualmente en posición inversa (ICB). El gen de la topoisomerasa II α , que interesa más particularmente y que comprende 5 secuencias ICB en su promotor, parece una de las dianas privilegiadas de ICBP90.

Estos experimentos han permitido determinar las características notables del anticuerpo 1RC1C-10, que solo marca únicamente las células en proliferación en el caso de células no cancerosas y marca las células cancerosas en proliferación o quiescencia; es utilizable en 4 técnicas diferentes (transferencia Western, inmunocitoquímica, inmunohistología e inmunoprecipitación); posee una muy buena afinidad y permite utilizar una dilución 1/150.000 en inmunohistoquímica (concretamente, 13 ng/ml); por último, su utilización no genera casi ruido de fondo.

Las aplicaciones futuras de 1RC1C-10 se sitúan en primer lugar en los campos del diagnóstico y la investigación fundamental. Para el diagnóstico en anatomía patológica, por ejemplo, sería perfectamente posible dar cuenta del estado proliferativo de un tejido canceroso dado. Respecto a la investigación fundamental, están en marcha investigaciones de laboratorio con el fin de determinar la contribución exacta de ICBP90 en los mecanismos de proliferación de células normales y de células cancerosas. Así pues, para el estudio de la expresión de ICBP90 en función del ciclo celular, de su localización nuclear precisa y de su interacción con otras proteínas celulares, la utilización del anticuerpo será inevitable.

Por el momento, no se ha estudiado la expresión de ICBP90 en función del ciclo celular. No obstante, en el caso en que las líneas de células cancerosas sean confluentes o cuando estén en proliferación, no pueden detectarse diferencias significativas en la expresión de ICBP90 (Fig. 1) al menos en lo que se refiere a la forma de 97 kDa. En contraposición, en células confluentes no cancerosas (células de músculo liso bronquiales humanas) la expresión de ICBP90 es difícilmente detectable (resultados no ilustrados). Esto se confirma en los cortes histológicos en que ninguna célula quiescente se marca por el anticuerpo. Es posible en consecuencia que la ICBP90 sea expresada en cualquier fase del ciclo celular en células cancerosas, mientras que su expresión variaría en función de cada fase en células no cancerosas. Esto vuelve la utilización del anticuerpo extremadamente interesante, en el sentido de que se tendría a disposición un marcador de la proliferación celular de tejidos cancerosos que no dependería de la fase del ciclo celular, al contrario que otros marcadores de la proliferación celular tales como Ki-67, topoisomerasa II α , ciclina E y ciclina B1. Efectivamente, el final de la fase S se caracteriza por una expresión muy baja de Ki-67, la ciclina E marca las células al final de la fase G1 hasta la mitad de la fase S y la ciclina B2 marca las células en fase G2/M (para revisión, Darzynkiewicz *et al.*, 1994). Por otro lado, se ha mostrado que el PCNA (Proliferating Cell Nuclear Antigen en inglés, antígeno nuclear de célula proliferante) sobreestima el número de células en proliferación en ciertos tipos de tejidos (Roskell y Biddolph, 1999).

La ICBP90 desempeña un papel importante en la proliferación celular al regular la expresión de genes tales como el de topoisomerasa II α . Las diferentes estrategias que tienden a bloquear la acción de esta proteína deben permitir modificar la proliferación celular. Así, la utilización del anticuerpo 1RC1C-10, así como la utilización de péptidos que imitan la interacción ADN/ICBP90 sin generar por ello el efecto fisiológico posterior, constituyen una posibilidad interesante. El diseño de estos péptidos se inspiraría directamente en la secuencia proteica de ICBP90 que se ha descrito. Una forma truncada correspondiente a ICBP59 podría ser, por ejemplo, uno de los primeros candidatos.

Puede considerarse el bloqueo puro y duro de la expresión de ICBP90 con el objetivo de eliminar completamente su influencia sobre los genes y, por extensión sobre la proliferación; puede hacerse mediante un enfoque clásico obteniendo inhibidores de la proteína, o bien utilizando un enfoque más moderno correspondiente a la técnica de interferencia por ARN bicatenario (ARN interferente o ARNi) tal como se ha descrito recientemente por Kennerdell y Carthew (1998).

EJEMPLO 3: AISLAMIENTO Y CARACTERIZACIÓN DEL GEN DE ICBP90

3.1. Materiales y procedimientos

3.1.1. Construcción y cribado de un banco genómico placentario humano

Después de digestión parcial con la enzima MboI, se fraccionó el ADN genómico placentario en función del tamaño en un gradiente de 10 a 40% de sacarosa. Se ligaron fragmentos de ADN de 15 kb con un vector λ GEM12 previamente digerido con BamHI (Promega, Madison WI, EE.UU.). Después de empaquetar, se valoraron las partículas de los fagos λ en células TAP 90. El banco genómico contiene 3×10^6 unidades formadoras de placas (ufp). Se expusieron para el análisis 10^6 clones. Se marcó una sonda de 620 pb correspondiente al extremo 5' terminal del ADNc de ICBP90 utilizada para el cribado con $\alpha^{32}\text{P}$ -dCTP mediante un procedimiento de cebado aleatorio (random priming) (Sambrook *et al.*, 1989). Se utiliza la sonda marcada según un protocolo clásico de hibridación en placa para cribar el banco genómico (Sambrook *et al.*, 1989). Se realizó la hibridación a 68°C en 5X SSC (NaCl 15 mM, citrato de sodio 1,5 mM, pH 7,0), 5 X de solución de Denhardt, ADN de esperma de salmón 100 $\mu\text{g}/\text{ml}$ y 0,1% de SDS, seguido de 30 minutos de lavados con 2X SSC, 0,1% de SDS a temperatura ambiente.

Se realizaron dos etapas de cribado para purificar un clon positivo. Se digirió a continuación el clon positivo con la enzima NotI y se subclonaron dos fragmentos de 6 y 10 kb en el vector pBluescript-SK+ (Stratagene, La Jolla CA, EE.UU.) según un protocolo estándar (Sambrook *et al.*, 1989).

3.1.2. Cribado del banco de ADNc de timo humano

Se cribó un banco de λ GT10 del extremo 5' de ADNc de timo humano (Clontech, Palo Alto, CA, EE.UU.) mediante hibridación en placa utilizando la sonda de ADNc de 679 pb sintetizada tal como en el párrafo relativo al análisis por

transferencia Northern. Se detectaron las señales utilizando cloruro de 4-nitroazul de tetrazolio y fosfato de 5-bromo-4-cloro-3-indolilo como sustrato.

3.1.3. Reacción de polimerización en cadena (PCR) de ADN genómico placentario

5 Se preparó ADN genómico placentario según un procedimiento convencional (Sambrook *et al.*, 1989). Para la región 5' del gen ICBP90, los inventores utilizaron el kit PCR Advantage®-GC Genomic de Clontech, que está adaptado a regiones ricas en GC del ADN genómico. Para cubrir las regiones 3' flanqueantes, se utilizaron la polimerasa Taq (Sigma, St. Louis, MO, EE.UU.) y su tampón correspondiente. Se realizaron las reacciones según las instrucciones del fabricante, utilizando 250 ng de ADN genómico placentario como matriz en un volumen final de 50 µl. Con el fin de obtener la amplificación de los intrones de longitudes 19 kb y 8,7 kb, se utilizó el sistema PCR Expand™ 20kb^{plus} (Roche Diagnostics, Mannheim, Alemania).

Se realizó la reacción en 100 µl utilizando 125 ng de ADN genómico placentario por reacción.

3.1.4. Construcciones plasmídicas y ensayos CAT

15 Se obtuvieron una serie de fragmentos diferentes por PCR de la región 5' flanqueante del gen de ICBP90 utilizando cebadores de 20 nucleótidos con el fin de obtener las construcciones descritas en la Figura 10. Estas contienen un sitio de restricción BamHI y el ADN genómico placentario se utilizó como cebador. Se digirieron los productos de PCR y se subclonaron en dirección 5' del gen informador cloranfenicol acetiltransferasa (CAT) de un vector que contenía el promotor mínimo de timidina cinasa (pBICAT2). Se verificaron las construcciones plasmídicas por secuenciación. Se cultivaron células COS-1 en medio Dulbecco modificado por Eagle (DMEM) suplementado con 5% de suero fetal de ternero. Después de la exposición, se transfirieron las células con las diferentes construcciones plasmídicas (5 µg) utilizando la técnica de coprecipitación con fosfato de calcio (Banerji *et al.*, 1981). Se realizaron a 20 continuación los análisis de expresión de CAT como se describe en otro lugar (Goetz *et al.* (1996)).

3.1.5. Localización cromosómica del gen ICBP90

25 Se prepararon cromosomas metafásicos a partir de leucocitos humanos de sangre periférica según protocolos estándares (Haddad *et al.* (1988)). Brevemente, se marcó una sonda de 10 kb, correspondiente a un fragmento 5' terminal del clon de 16 kb aislado a partir del cribado del banco de ADN genómico placentario, con biotina-16-dUTP (Roche Diagnostics) mediante "traslación de mella". Se precipita a continuación la sonda con un exceso (50X) de ADN humano Cot-1 (Life Technologies, Rockville MD), resuspendido en 50% de formamida, 1X SSC, prehibridado durante 2 horas a 37°C y entonces hibridado durante una noche a 37°C. Se realiza la detección utilizando avidina-FITC (Vector Laboratories, Burlingame CA). Se contratificaron los cromosomas con 4',6-diamino-2-fenilindol (Sigma).

3.1.6. Análisis de transferencia Northern y de transferencia Western

35 Se prehibridó una membrana de transferencia Northern que contenía 2 µg de ARN poliA+ por línea, procedente de 7 líneas celulares humanas cancerosas diferentes (Clontech) en Express Hyb (Clontech) y se hibridó entonces con la sonda específica de ICBP90 en Express Hyb a 68°C durante 2 horas. Se preparó la sonda bicatenaria marcada con digoxigenina mediante amplificación por PCR de un fragmento de 676 pb a partir de ADNc de ICBP90 (nucleótidos 806 a 1485; número de acceso a Genbank AF 129 507) según las instrucciones del fabricante (Roche Diagnostics).

40 Después de purificación mediante una columna de cromatografía Micro Bio-Spin 30 (Bio-Rad, Hercules, CA), se calentó a 95°C durante 15 minutos la sonda específica de ICBP90 (5 ng/ml) y se enfrió entonces en hielo antes de la adición de la solución de hibridación. Se realizaron los lavados después de hibridación dos veces con 2X SSC, 0,1% de SDS (30 minutos por lavado a temperatura ambiente) y entonces dos veces con 0,1X SSC, 0,1% de SDS (30 minutos por lavado a 68°C). Se trató la membrana con solución A (ácido málico 0,1 M, NaCl 0,15 M a pH 7,5) y se bloqueó entonces mediante incubación con 1% de agente bloqueante (Roche Diagnostics) en tampón A durante 30 minutos a temperatura ambiente.

45 Se añadió un anticuerpo conjugado con fosfatasa alcalina dirigido contra digoxigenina (fragmento Fab, Roche Diagnostics) (150 mU/ml) y se incubó entonces durante 30 minutos a temperatura ambiente. Se lavó a continuación la membrana dos veces con la solución A y se equilibró entonces con Tris-HCl 0,1 M, NaCl 0,1 M, pH 9,5. Para la detección por quimioluminiscencia, los inventores han utilizado el agente 3-(4-metoxiespiro-{1,2-dietano-3,2'-(5'-cloro)tricyclo[3.3-1.1^{3,7}]decan}-4-il)fenilfosfato de disodio (Roche Diagnostics) según las instrucciones del fabricante. Se cuantificaron las bandas de ARNm utilizando el software NIH Image 1.62 y se expresaron en porcentaje de la banda de ARNm más abundante (es decir, la banda de 5,1 kb de las células HL-60).

50 Se realizó el análisis de transferencia Western como se describe en otro lugar (Hopfner *et al.* (2000)). Se detectaron las señales utilizando cloruro de 4-nitroazul de tetrazolio/fosfato de 5-bromo-4-cloro-3-indolilo como sustrato.

3.1.7. Herramientas de investigación del alineamiento local de bases, predicciones de sitios de iniciación de la transcripción y de la señal poliA

Se realizaron las investigaciones de alineamiento local de bases por el Centre National d'Information en Biotechnologie del National Institute of Health (Bethesda, MD, EE.UU.). Se realizaron el cribado del banco de factores de transcripción con el programa informático Mat Inspector, las predicciones de sitios de iniciación de la transcripción (TSS) con Neural Network, así como la predicción de la señal poliA, por el Baylor College of Medicine (Reese *et al.* (1996)).

3.2. Resultados

3.2.1. Aislamiento y caracterización del gen de ICBP90

Se cribó un banco de ADN complementario de placenta humana clonado en el fago lambda GEM 12 con la ayuda de una sonda de ADN. El cribado condujo a la purificación de un solo clon positivo que tiene un inserto de 16 kb. El análisis de secuencia permitió determinar que contenía una secuencia intrónica larga de 10 kb y que contenía 3 exones (denominados B, C y D en la Figura 9A). Todos los demás cribados, incluyendo especialmente los realizados por PCR en bancos de BAC (cromosoma artificial bacteriano) o YAC (cromosoma artificial de levadura) no permitieron aislar otros clones positivos. Por consiguiente, se ha decidido determinar el resto de la organización del gen por PCR directamente en el ADN genómico de placenta humana. La mayor dificultad fue obtener el lado 5' del intrón de 19 kb. Así, se eligieron los cebadores en el exón A (cebador codificante) y en el lado 5' del clon de 16 kb (cebador anticodificante). El exón E y el intrón de 8,7 kb se amplificaron utilizando un cebador codificante en el exón D y el cebador anticodificante en el exón F. Finalmente, se determinó la secuencia completa del exón F hasta la señal de poliadenilación utilizando un cebador codificante elegido al inicio del exón F y el cebador anticodificante en el lado 3' de una EST (nº de referencia en GenBank AW297533) homóloga de la secuencia del gen de ICBP90. La secuencia completa del gen de ICBP90 muestra que está compuesta por 6 exones codificantes cuyo tamaño varía de 100 pares de bases a 3453 pares de bases. La mayoría de las conexiones exón/intrón responden a secuencias consenso para los sitios aceptores y dadores de corte y empalme. Se encontró una secuencia consenso de poliadenilación (AATAAA) en la región 3', es decir, 1152 nucleótidos después del codón de terminación en la Figura 9A.

3.2.2 La región 5' del gen de ICBP90

El cribado de un banco de ADN complementario de timo humano clonado en el fago lambda gt 10 condujo a la obtención de dos poblaciones de ADNc que se distinguen entre sí en su región 5', exactamente 10 pares de bases en dirección 5' del codón de iniciación, es decir, en la región 5' no traducida. Estas dos poblaciones de ADNc predicen la existencia de dos exones alternativos en 5' denominados exón I y exón II (Figura 9A). Se ha observado que los exones I y II están ligados a un sitio de corte y empalme alternativo interno del exón A. Además, se ha encontrado en una base de datos una EST (nº de referencia en GenBank AI084125) correspondiente a los nucleótidos 1290 a 1356 (Figura 9B). Se determinaron las posiciones de estos dos exones y de la EST en el interior del locus mediante PCR. Para ello, se han utilizado los cebadores correspondientes a los 18 primeros nucleótidos de cada exón y un cebador anticodificante elegido en el primer exón traducido (exón A). Esta estrategia ha permitido reconstruir la región 5' tal como se representa en las Figuras 9A y 9B, con el exón I correspondiente a los nucleótidos 1 a 134 y el exón II correspondiente a los nucleótidos 676 a 725. La secuencia EST (AI084125) es contigua al sitio de corte y empalme interno del exón A. Aún no se ha determinado con precisión el inicio de los exones I, II y A, ya que sus secuencias se han deducido a partir de cribados de bancos de ADNc (Figura 9A).

Se encontraron cuatro secuencias GC (GC1 a GC4) en la región 5' (Figura 9B). Estas secuencias representan sitios potenciales de enlace para el factor de transcripción Sp1, pero solamente una secuencia (GC3) corresponde a una secuencia consenso, es decir, GGGGCGGGG. Además, se han encontrado dos secuencias CCAAT (CB1 y CB2). Los análisis predictivos de secuencia sugieren que existen dos regiones promotoras en la región 5', es decir, antes del codón de iniciación (ATG). Se han predicho dos sitios potenciales de iniciación de la transcripción en las posiciones 571 y 827. El primero sigue a la secuencia consenso de enlace a Sp1 y el segundo sigue a la secuencia GC1 (respectivamente entre los exones I y II y los exones II y A). Con el fin de ver si estas dos regiones son funcionales como región promotora, se prepararon varias construcciones plasmídicas que contenían un gen informador (gen de cloranfenicol acetiltransferasa, CAT) en dirección 3' de las diferentes regiones promotoras potenciales. Se transfirieron células COS con estas construcciones plasmídicas. La Figura 10 muestra los resultados obtenidos y que corresponden con el porcentaje de aumento de la actividad basal. Se obtuvo la actividad máxima con la construcción plasmídica que contiene 1114 pares de bases en dirección 5' del sitio de iniciación de la traducción, con un aumento de un 236,7% de la actividad promotora basal (promotor mínimo del gen de timidina cinasa). La construcción plasmídica que contiene 642 pares de bases en dirección 5' del ATG condujo a un aumento de un 115,6%, mientras que la construcción plasmídica que contiene únicamente la secuencia entre el exón I y el exón II mostraba una actividad relativamente débil con un aumento de únicamente un 22,8% (Figura 10). Estos resultados sugieren la existencia de una región promotora entre los exones II y A.

3.2.3. Localización cromosómica del gen de ICBP90

Se realizó la localización cromosómica del gen de ICBP90 mediante hibridación *in situ* de fluorescencia (FISH). El gen de ICBP90 está localizado en el cromosoma 19p13.3 en una región telomérica. Una investigación realizada en Genbank mostró que una región de 6 Mb en la banda cromosómica 19p13.3 de un banco de cósmidos específico del cromosoma 19 (híbrido hombre/hámster 5HL2-B) contiene 147 nucleótidos que codifican los aminoácidos 746 a 793 de ICBP90. Estas secuencias se localizó entre los marcadores STS (sitio de secuencia marcada) D19S883 y D19S325.

3.2.4 Expresión de ICBP90 en diferentes líneas celulares

La ICBP90 participa en la regulación de la expresión del gen TopII α (Hopfner *et al.* (2000)). Como TopII α se expresa de manera diferencial en diferentes tumores y líneas celulares, la ICBP90 misma es susceptible de tener una regulación compleja en términos de actividad de expresión génica.

En una primera etapa hacia la comprensión de los mecanismos que regulan la expresión del gen de ICBP90, se analizó el ARNm de ICBP90 en diferentes líneas celulares. Se estudió el ARNm de ICBP90 en la línea celular HL60 derivada de leucemia promielocítica (línea 1), células Hela S3 (línea 2), células de leucemia linfoblástica MOLT-4, células Raji de linfoma de Burkitt (línea 5), adenocarcinoma colorrectal SW 480 (línea 6) y células A549 de carcinoma de pulmón (línea 7) (Figura 11A).

Se observan dos bandas de ARNm de 4,3 y 5,1 kb. Las cantidades relativas de ARNm en las bandas varían según el tipo celular. El histograma de la Figura 11A muestra las tasas de ARNm en las bandas de cada una de las líneas celulares, expresadas en porcentaje de la cantidad máxima observada en bandas de ARNm de 5,1 kb en las células HL-60 (línea 1, Figura 11A). En las células MOLT-4, solo se observa la banda de ARNm de 4,3 kb, mientras que en las células de leucemias promielocíticas, la banda de 5,1 kb es predominante. En las células Raji de linfoma de Burkitt, solo se detecta la banda de 5,1 kb. Aproximadamente, se observan cantidades iguales de los dos tipos de ARNm en las demás líneas celulares, es decir, las células Hela, K562, A549, SW 480. Para las células HL-60, no obstante, el ARNm de 5,1 kb se expresa más que el ARNm de 4,3 kb. Se han emprendido otros análisis de células HeLa para confirmar que los 2 transcritos proceden de la transcripción del gen de ICBP90. La utilización de una sonda de ADNc de 626 pb marcada con digoxigenina localizada inmediatamente en dirección 5' de la señal poliA (es decir, el exón F) y utilizada como sonda para experimentos de transferencia Northern, produjo los mismos resultados, es decir, la aparición de dos bandas de ARNm de 4,3 kb y 5,1 kb. Este resultado confirma que las dos formas de ARNm se generan a partir de un solo gen.

Los inventores han estudiado igualmente la expresión de la proteína ICBP90 con el fin de determinar si estas dos isoformas de ARNm son susceptibles de codificar dos proteínas diferentes.

La Figura 11B muestra el perfil de expresión de ICBP90 en los extractos proteicos de células MOLT-4 y Hela. Aunque se observa una sola banda de 97 kDa en las células MOLT-4, en las células HeLa, al lado de la banda de 97 kDa que está duplicada, se observan varias otras bandas de un peso molecular inferior. Estos resultados sugieren que, en las células MOLT-4, un ARNm codifica una forma única de ICBP90. Al contrario, en las células HeLa, los dos ARNm son susceptibles de conducir a la producción de diferentes isoformas de ICBP90.

3.3 Comentarios

El gen de ICBP90 se extiende sobre aproximadamente 35,8 kb. Se han identificado por los inventores 6 exones traducidos y 2 exones no traducidos y, de hecho, 7 intrones. Los dos dominios en dedo de cinc de ICBP90 están codificados por el mismo exón (exón F), al contrario que el gen del receptor de estrógenos humanos, en el que cada uno de los supuestos dedos de cinc del dominio de unión a ADN del receptor está codificado separadamente (Ponglikitmongkol *et al.* (1988)). El dominio "de tipo ubiquitina" de ICBP90 está codificado por los exones A y B, mientras que el dominio de "cremallera de leucina" está codificado por el exón B. De manera interesante, solo el exón F es susceptible de codificar una proteína funcional, puesto que codifica dos señales de localización nuclear, dominios en dedo de cinc y varios supuestos sitios de fosforilación. Se encontraron dos grandes intrones de 8,7 kb y 19 kb.

El gen de ICPB90 se localizó en la región cromosómica 19p13.3. Se localizaron varios otros genes en esta región, por ejemplo el factor nuclear I/C (igualmente un factor de transcripción que se liga con CCAAT, (Qian *et al.* (1995)). De manera interesante, se describió una translocación atípica t (7; 19) en la leucemia mielomonocítica aguda que implica un sitio frágil en el locus 19p13.3 (Sherer *et al.* (1991)). Igualmente, se ha sugerido que los genes implicados en el desarrollo de carcinomas pancreáticos están localizados en 19p13.3 y 19q13.1-13.2 (Hoglund *et al.* (1998)). Las redistribuciones de las bandas de 14q32.3 y 19p13.3 de una deleción preferida del brazo corto del cromosoma 1 constituyen alteraciones cromosómicas no aleatorias en el mieloma múltiple y la leucemia de células plasmáticas (Taniwaki *et al.* (1996)). Se localizaron otros genes en esta región; incluyen un gen implicado en el adenocarcinoma del síndrome de Peutz-Jeghers (Gruba *et al.* (1998)). Igualmente, se sugirió que el supuesto gen supresor tumoral para adenoma maligno está localizado en D19S216 al nivel de la banda cromosómica 19p13.3, que desempeña un papel importante en la tumorigénesis del adenoma maligno ((Lee *et al.* (1998)).

El análisis de la secuencia de la región 5' del gen de ICBP90 reveló la existencia de varios exones no traducidos con una región promotor entre los exones II y A, y probablemente un segundo promotor más débil localizado entre los exones I y II. La región promotora entre los exones II y A es un promotor sin secuencia TATA, sugiriendo que el gen de ICBP90 puede ser un gen doméstico, al menos cuando está implicado este promotor. En este sentido, se parece mucho a las regiones promotoras de los genes ATF α (Goetz *et al.*, 1996), CRE-BP1 / ATF 2 (Nagase *et al.*, 1990) y TopII α (Hochhauser *et al.*, 1992), que no contienen secuencias TATA canónicas sino varios sitios de unión de SP-1.

Las secuencias GC y/o CCAAT son susceptibles de estar implicadas en la regulación de la expresión del gen de ICBP90 mediante los factores de transcripción SP-1 y las proteínas de enlace a CCAAT. Además, dado que la proteína ICBP90 es una proteína de unión con CCAAT, la ICBP90 es igualmente susceptible de regular su propia expresión.

Se cribó un banco de datos de factores de transcripción con la ayuda del programa informático Mat Inspector del Baylor College of Medicine, y se identificaron numerosos sitios de unión de factores de transcripción en la secuencia precedente al codón ATG (Figura 9B). Entre estos sitios de unión a factores de transcripción, es interesante observar sitios de unión del factor de transcripción AP-2 regulado en el transcurso del desarrollo y que controla la expresión de genes tales como DR-nm 23 (Martínez *et al.* (1997)), los sitios de unión de la proteína mieloides en "dedo de cinc" MZF 1, que está implicada en la regulación de la hematopoyesis (Hromas *et al.* (1996)).

El análisis de transferencia Northern demostró que existen dos poblaciones de ARNm de 4,3 kb y 5,1 kb. De manera interesante, cada población presenta una especificidad celular. Por ejemplo, las células linfoblásticas MOLT-4 expresan solamente el ARNm de 4,3 kb, mientras que en las células Raji de linfoma de Burkitt (linfocitos B maduros), se observa solo el transcrito de 5,1 kb. Las células HL-60 expresan más ARNm de 5,1 kb que ARNm de 4,3 kb. Las células HL-60 y las células Raji de linfoma de Burkitt están más diferenciadas que las células MOLT-4, sugiriendo que la tasa de expresión del transcrito de 5,1 kb con relación a la de 4,3 kb puede estar correlacionada directamente con el estado de diferenciación de las células.

De manera interesante, se identificó una etiqueta de secuencia expresada (EST, Expressed Sequence Tag en inglés) correspondiente a la secuencia 5' del exón A a partir de oligodendroglioma anaplásico (número de acceso a Genbank AI 084 125), mientras que se aisló una EST correspondiente a la inclusión del exón II a partir de una mezcla de tumores de células germinales (nº de acceso a Genbank AI 968 662). Los resultados de los inventores sugieren pues que la regulación de los transcritos de ICBP90 es comparable con lo que pasa para el receptor de estrógenos. De hecho, se informó de 6 transcritos diferentes que codifican una proteína común, pero diferentes en la región 5' no traducida por el hecho de un corte y empalme alternativo de los exones en dirección 5' (Flouriot *et al.*, 1998 y Grandien, 1996).

El análisis de transferencia Western muestra una banda mayoritaria de 97 kDa en las células MOLT-4, mientras que se observan varias bandas en las células HeLa (Figura 11B). Estos datos están de acuerdo con la existencia de varios ARNm de ICBP90 y/o isoformas de la proteína ICBP90 cuya tasa de expresión puede controlarse de manera específica de célula.

Se han descrito dos isoformas proteicas para el receptor de estrógenos (Griffin *et al.*, 1999) que difieren entre sí por los 41 aminoácidos N-terminales. La doble banda a 97 kDa observada a partir de células HeLa (Figura 11B) es pues susceptible de representar dos isoformas diferentes por su extremo N-terminal. Por este hecho, el exón A que codifica 47 aminoácidos se corta y empalma fuera del marco de lectura, y por consiguiente, la región proteica codificante comienza con el exón B. No obstante, es igualmente posible que existan otros exones susceptibles de transcribirse en otros tejidos.

Igualmente, el intrón de 8,7 kb (es decir, entre el exón D y E) es susceptible de contener una región promotora que pueda conducir a isoformas de ICBP90 de pesos moleculares inferiores a los observados en células HeLa en proliferación (Figura 11B). De manera interesante, la especificidad tisular de los diferentes ARNm de receptor de estrógenos se determina mediante diferentes promotores cuya actividad parecía estar alterada en líneas celulares de cáncer de mama (Flouriot *et al.*, 1998).

El conjunto de estos resultados sugiere que el gen de ICBP90 y la proteína ICBP90 presentan características comunes con los miembros de la familia del receptor de ácido retinoico, esteroides u hormonas tiroideas en lo que respecta a las estructuras génica y proteica.

Efectivamente, los inventores han demostrado experimentalmente, utilizando la técnica de híbrido doble, la existencia de interacciones entre la proteína ICBP90 y TIP60 (Tat Interactive Protein en inglés, proteína interactiva con Tat, 60 kDa). La proteína TIP60 se ha descrito muy recientemente por ser un coactivador del receptor nuclear, especialmente el receptor de andrógenos (Brady ME *et al.*, 1999).

Por esto, la ICBP90 es susceptible de desempeñar el papel de un receptor nuclear con el que se liga un ligando endógeno. Está pues igualmente dentro del alcance de la presente invención utilizar el polipéptido ICBP90 de la invención para aislar, cribar e identificar el ligando endógeno. Está igualmente dentro del alcance de la invención utilizar el polipéptido ICBP90 de la invención para aislar, cribar e identificar moléculas naturales o de síntesis, biológicas o químicas, agonistas o antagonistas de este ligando natural.

REFERENCIAS

- Austin *et al.* (1993), Biochim. Biophys. Acta, 1172, 283-291.
- Banerji, J. *et al.* (1981), Cell, 27: 299-308.
- Barany, F. (1991), Proc. Natl. Acad. Sci. USA, 88, 189-193.
- 5 Boritzki, T. J. *et al.* (1988), Biochem. Pharmacol., 37, 4063-4068.
- Brady, M. E. *et al.* (1999), J. Biol. Chem., 274: 17599-17604.
- Brandt, T.L. *et al.* (1997), J. Biol. Chem., 272, 6278-6284.
- Brou, C. *et al.* (1993), EMBO J., 12, 489-499.
- Buckholz, R. G. (1993), Curr. Op. Biotechnology, 4, 538-542.
- 10 Burg, J.L. *et al.* (1996), Mol. and Cell. Probes, 10, 257-271.
- Chu, B.C.F. *et al.* (1986), Nucleic Acids Res., 14, 5591-5603.
- Chung, T.D.Y. *et al.* (1989), Proc. Natl. Acad. Sci. USA, 86, 9431-9435.
- Darzynkiewicz *et al.* (1994), Methods in Cells Biology, 41, 421-435.
- Deffie, A.M. *et al.* (1989), Cancer Res., 49, 58-62.
- 15 De Vries, L. *et al.* (1995), Proc. Natl. Acad. Sci. USA, 92, 11916-11920.
- Devys *et al.* (1993), Nature Genet., 4, 335-340.
- Drake, F. H. *et al.*, Biochemistry, 28, 8154-8160.
- Duck, P. *et al.* (1990), Biotechniques, 9, 142-147.
- Edwards, C.P. y Aruffo, A. (1993), Curr. Op. Biotechnology, 4, 558-563.
- 20 Erlich, H.A. (1989), New York: Stockton Press.
- Flouriot *et al.* (1998), Mol. Endocrinol., 12: 1239-254.
- Fry, A.M. *et al.* (1991), Cancer Res., 51, 6592-6595.
- Furth *et al.* (1992), Anal. Biochem., 205: 365.
- Gaub, M.P. *et al.* (1998), J. Histochem Cytochem., 46, 1103-1111.
- 25 Goetz, J. *et al.* (1996), J. Biol. Chem., 271: 29589-29598.
- Goswami, P.C. *et al.* (1996), Mol. Cell. Biol., 16, 1500-1508.
- Grandien (1996), Mol. Cell. Endocrinol., 116: 207-212).
- Griffin *et al.* (1999), Mol. Endocrinol. 13: 1571-1587.
- Gruba *et al.* (1998), Cancer Res., 58: 5267-5270.
- 30 Guatelli, J.C. *et al.* (1990), Proc. Natl. Acad. Sci. USA, 87, 1874-1878.
- Guinee, D.G. *et al.* (1996), Cancer, 78, 729-735.
- Haddad *et al.* (1988), Human Genet., 103: 619-625.
- Heck, M.M. *et al.* (1988), Proc. Natl. Acad. Sci. USA, 85, 1086-1090.
- Herzog, C.E. y Zwelling, L. A. (1997), Biochem. Biophys. Res. Commun., 232, 608-612.
- 35 Hochhauser, D. *et al.* (1992), J. Biol. Chem., 267, 18961-18965.
- Hoglund *et al.* (1998), Genes Chromosomes Cancer, 21: 8-16.
- Hopfner *et al.* (2000), Cancer Res., 60: 121-128.

- Hromas *et al.* (1996), Curr. Top. Microbiol. Chem., 211: 159-164.
- Innis, M.A. *et al.* (1990), Academic Press.
- Inouye, C. *et al.* (1994), DNA Cell Biol., 13, 731-742.
- Isaacs, R.J. *et al.*, Biochim. Biophys. Acta, 1400, 121-137.
- 5 Isaacs, R.J. *et al.* (1996), J. Biol. Chem., 271, 16741-16747.
- Jenkins, J.R. *et al.* (1992), Nucleic Acids Res., 20, 5587-5592.
- Kassel, O. *et al.* (1998), Mol. Pharmacol., 54, 1073-1079.
- Kennerdell, J.R. y Carthew, R.W. (1998), Cell, 95, 1017-1026.
- Kievitis, T. *et al.* (1991), J. Virol. Methods, 35, 273-286.
- 10 Kohler, G. *et al.* (1975), Nature, 256 (5517), 495-497.
- Kubo, T. *et al.* (1995), Cancer Res., 55, 3860-3864.
- Kwoh, D.Y. *et al.* (1989), Proc. Natl. Acad. Sci. USA, 86, 1173-1177.
- Landegren, U. *et al.* (1988), Science, 241, 1077-1080.
- Lavie, J. *et al.* (1999), J. Biol. Chem., 274, 2308-2314.
- 15 Lee *et al.* (1998), Cancer Res., 58 : 1140-1143.
- Lim, K. *et al.* (1998), Biochem. Mol. Biol. Int., 46, 35-42.
- Lizardi, P.M. *et al.* (1988), Bio/technology, 6, 1197-1202.
- Luckow, V.A. (1993), Curr. Op. Biotechnology, 4, 564-572.
- Martinez *et al.* (1997), Cancer Res., 57: 1180-1187.
- 20 Matthews, J.A. *et al.* (1988), Anal. Biochem., 169: 1-25.
- Miele, E.A. *et al.* (1983), J. Mol. Biol., 171: 281-295.
- Nagase *et al.* (1990), J. Biol. Chem., 265: 17300-17306.
- Nitiss, J.L. (1998), Biochim. Biophys. Acta, 1400, 63-81.
- Olins, P.O. y Lee, S.C. (1993), Curr. Op. Biotechnology, 4, 520-525.
- 25 Pommier, Y. *et al.* (1994), Cancer Invest., 12, 530-542.
- Ponglikitmongkol *et al.* (1988), EMBO J. 7 : 3385-3388.
- Rio, M.C. *et al.* (1987), Proc. Natl. Acad. Sci. USA, 84, 9243-9247.
- Qian *et al.* (1995), Genomics, 28: 66-73.
- 30 Reese *et al.* (1996), "Large Scale Sequencing Specific Neural Networks for Promoter and Splice Recognition. Biocomputing : Proceedings of the 1996 Pacific Symposium. Edited by Lawrence Hunter and Terri E. World Scientific", Singapur, 1996, 27 de enero de 1996.
- Rochette-Egly, C. *et al.* (1997), Cell, 90, 97-107.
- Roskell, D.E. y Biddolph, S.C. (1999), Eur. J. Med. Res. 26, 105-106.
- Rolfs, A. *et al.* (1991), Berlin: Springer-Verlag.
- 35 Sambrook, J. *et al.* (1989), "Molecular Cloning: A Laboratory Manual", Sec. Ed. Cold Spring Harbor Lab., Cold Spring Harbor, Nueva York.
- Sandri, M.I. *et al.* (1996), Nucleic Acids Res., 24, 4464-4470.
- Segev, D. (1992), Kessler C. Springer Verlag, Berlin, Nueva York, 197-205.

Sherer *et al.* (1991), Cancer Genet. Cytogenet., 57: 169-173.

Stone, B.B. *et al.* (1996). Mol. and Cell. Probes, 10: 359-370.

Tang *et al.* (1992), Nature, 356: 152.

Taniwaki *et al.* (1996), Leuk. Lymphoma, 21: 25-30.

5 Tsai-plugfelder, M. *et al.* (1988), Proc. Natl. Acad. Sci. USA, 85, 7177-7181.

Walker, G.T. *et al.* (1992), Nucleic Acids Res., 20: 1691-1696.

Wang, J.C. (1996), Ann. Rev. Biochem., 65, 635-692.

Wang, M.M. y Reed, R.R. (1993), Nature (Londres), 364, 121-126.

Yamazaki *et al.* (1996), Acta Oncol., 35, 417-423.

10

LISTA DE SECUENCIAS

- <110> ADEREGEM
- <120> Polipéptido ICBP90 y sus fragmentos, polinucleótidos que codifican dichos polipéptidos y aplicaciones para el diagnóstico y tratamiento del cáncer
- 5 <130> d18243
- <140>
- <141>
- <160> 12
- <170> PatentIn Ver. 2.1
- 10 <210> 1
- <211> 2382
- <212> ADN
- <213> *Homo sapiens*
- <220>
- 15 <221> CDS
- <222> (1)..(2382)
- <400> 1

ES 2 378 527 T3

atg tgg atc cag gtt cgg acc atg gat ggg agg cag acc cac acg gtg	48
Met Trp Ile Gln Val Arg Thr Met Asp Gly Arg Gln Thr His Thr Val	
1 5 10 15	
gac tcg ctg tcc agg ctg acc aag gtg gag gag ctg agg cgg aag atc	96
Asp Ser Leu Ser Arg Leu Thr Lys Val Glu Glu Leu Arg Arg Lys Ile	
20 25 30	
cag gag ctg ttc cac gtg gag cca ggc ctg cag agg ctg ttc tac agg	144
Gln Glu Leu Phe His Val Glu Pro Gly Leu Gln Arg Leu Phe Tyr Arg	
35 40 45	
ggc aaa cag atg gag gac ggc cat acc ctc ttc gac tac gag gtc cgc	192
Gly Lys Gln Met Glu Asp Gly His Thr Leu Phe Asp Tyr Glu Val Arg	
50 55 60	
ctg aat gac acc atc cag ctc ctg gtc cgc cag agc ctc gtg ctc ccc	240
Leu Asn Asp Thr Ile Gln Leu Leu Val Arg Gln Ser Leu Val Leu Pro	
65 70 75 80	
cac agc acc aag gag cgg gac tcc gag ctc tcc gac acc gac tcc ggc	288
His Ser Thr Lys Glu Arg Asp Ser Glu Leu Ser Asp Thr Asp Ser Gly	
85 90 95	
tgc tgc ctg ggc cag agt gag tca gac aag tcc tcc acc cac ggt gag	336
Cys Cys Leu Gly Gln Ser Glu Ser Asp Lys Ser Ser Thr His Gly Glu	
100 105 110	
gcg gcc gcc gag act gac agc agg cca gcc gat gag gac atg tgg gat	384
Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp Glu Asp Met Trp Asp	
115 120 125	
gag acg gaa ttg ggg ctg tac aag gtc aat gag tac gtc gat gct cgg	432
Glu Thr Glu Leu Gly Leu Tyr Lys Val Asn Glu Tyr Val Asp Ala Arg	

ES 2 378 527 T3

130		135		140		
gac acg aac atg ggg gcg tgg ttt gag gcg cag gtg gtc agg gtg acg						480
Asp Thr Asn Met Gly Ala Trp Phe Glu Ala Gln Val Val Arg Val Thr						
145		150		155		160
cgg aag gcc ccc tcc cgg gac gag ccc tgc agc tcc acg tcc agg ccg						528
Arg Lys Ala Pro Ser Arg Asp Glu Pro Cys Ser Ser Thr Ser Arg Pro						
		165		170		175
gcg ctg gag gag gac gtc att tac cac gtg aaa tac gac gac tac ccg						576
Ala Leu Glu Glu Asp Val Ile Tyr His Val Lys Tyr Asp Asp Tyr Pro						
		180		185		190
gag aac ggc gtg gtc cag atg aac tcc agg gac gtc cga gcg cgc gcc						624
Glu Asn Gly Val Val Gln Met Asn Ser Arg Asp Val Arg Ala Arg Ala						
		195		200		205
cgc acc atc atc aag tgg cag gac ctg gag gtg ggc cag gtg gtc atg						672
Arg Thr Ile Ile Lys Trp Gln Asp Leu Glu Val Gly Gln Val Val Met						
		210		215		220
ctc aac tac aac ccc gac aac ccc aag gag cgg ggc ttc tgg tac gac						720
Leu Asn Tyr Asn Pro Asp Asn Pro Lys Glu Arg Gly Phe Trp Tyr Asp						
		225		230		235
gcg gag atc tcc agg aag cgc gag acc agg acg gcg cgg gaa ctc tac						768
Ala Glu Ile Ser Arg Lys Arg Glu Thr Arg Thr Ala Arg Glu Leu Tyr						
		245		250		255
gcc aac gtg gtg ctg ggg gat gat tct ctg aac gac tgt cgg atc atc						816
Ala Asn Val Val Leu Gly Asp Asp Ser Leu Asn Asp Cys Arg Ile Ile						
		260		265		270
ttc gtg gac gaa gtc ttc aag att gag cgg ccg ggt gaa ggg agc ccc						864
Phe Val Asp Glu Val Phe Lys Ile Glu Arg Pro Gly Glu Gly Ser Pro						
		275		280		285
atg gtt gac aac ccc atg aga cgg aag agc ggg ccg tcc tgc aag cac						912
Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His						
		290		295		300
tgc aag gac gac gtg aac aga ctc tgc agg gtc tgc gcc tgc cac ctg						960
Cys Lys Asp Asp Val Asn Arg Leu Cys Arg Val Cys Ala Cys His Leu						
		305		310		315
tgc ggg ggc cgg cag gac ccc gac aag cag ctc atg tgc gat gag tgc						1008
Cys Gly Gly Arg Gln Asp Pro Asp Lys Gln Leu Met Cys Asp Glu Cys						
		325		330		335
gac atg gcc ttc cac atc tac tgc ctg gac ccg ccc ctc agc agt gtt						1056
Asp Met Ala Phe His Ile Tyr Cys Leu Asp Pro Pro Leu Ser Ser Val						
		340		345		350
ccc agc gag gac gag tgg tac tgc cct gag tgc cgg aat gat gcc agc						1104
Pro Ser Glu Asp Glu Trp Tyr Cys Pro Glu Cys Arg Asn Asp Ala Ser						
		355		360		365
gag gtg gta ctg gcg gga gag cgg ctg aga gag agc aag aag aat gcg						1152
Glu Val Val Leu Ala Gly Glu Arg Leu Arg Glu Ser Lys Lys Asn Ala						
		370		375		380

ES 2 378 527 T3

aag atg gcc tcg gcc aca tcg tcc tca cag cgg gac tgg ggc aag ggc	1200
Lys Met Ala Ser Ala Thr Ser Ser Ser Gln Arg Asp Trp Gly Lys Gly	
385 390 395 400	
atg gcc tgt gtg ggc cgc acc aag gaa tgt acc atc gtc ccg tcc aac	1248
Met Ala Cys Val Gly Arg Thr Lys Glu Cys Thr Ile Val Pro Ser Asn	
405 410 415	
cac tac gga ccc atc ccg ggg atc ccc gtg ggc acc atg tgg cgg ttc	1296
His Tyr Gly Pro Ile Pro Gly Ile Pro Val Gly Thr Met Trp Arg Phe	
420 425 430	
cga gtc cag gtc agc gag tcg ggt gtc cat cgg ccc cac gtg gct ggc	1344
Arg Val Gln Val Ser Glu Ser Gly Val His Arg Pro His Val Ala Gly	
435 440 445	
atc cat ggc cgg agc aac gac gga tcg tac tcc cta gtc ctg gcg ggg	1392
Ile His Gly Arg Ser Asn Asp Gly Ser Tyr Ser Leu Val Leu Ala Gly	
450 455 460	
ggc tat gag gat gat gtg gac cat ggg aat ttt ttc aca tac acg ggt	1440
Gly Tyr Glu Asp Asp Val Asp His Gly Asn Phe Phe Thr Tyr Thr Gly	
465 470 475 480	
agt ggt ggt cga gat ctt tcc ggc aac aag agg acc gcg gaa cag tct	1488
Ser Gly Gly Arg Asp Leu Ser Gly Asn Lys Arg Thr Ala Glu Gln Ser	
485 490 495	
tgt gat cag aaa ctc acc aac acc aac agg gcg ctg gct ctc aac tgc	1536
Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu Ala Leu Asn Cys	
500 505 510	
ttt gct ccc atc aat gac caa gaa ggg gcc gag gcc aag gac tgg cgg	1584
Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala Lys Asp Trp Arg	
515 520 525	
tcg ggg aag ccg gtc agg gtg gtg cgc aat gtc aag ggt ggc aag aat	1632
Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys Gly Gly Lys Asn	
530 535 540	
agc aag tac gcc ccc gct gag ggc aac cgc tac gat ggc atc tac aag	1680
Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp Gly Ile Tyr Lys	
545 550 555 560	
gtt gtg aaa tac tgg ccc gag aag ggg aag tcc ggg ttt ctc gtg tgg	1728
Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly Phe Leu Val Trp	
565 570 575	
cgc tac ctt ctg cgg agg gac gat gat gag cct ggc cct tgg acg aag	1776
Arg Tyr Leu Leu Arg Arg Asp Asp Glu Pro Gly Pro Trp Thr Lys	
580 585 590	
gag ggg aag gac cgg atc aag aag ctg ggg ctg acc atg cag tat cca	1824
Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr Met Gln Tyr Pro	
595 600 605	
gaa ggc tac ctg gaa gcc ctg gcc aac cga gag cga gag aag gag aac	1872
Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg Glu Lys Glu Asn	
610 615 620	

ES 2 378 527 T3

agc aag agg gag gag gag gag cag cag gag ggg ggc ttc gcg tcc ccc 1920
 Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly Gly Phe Ala Ser Pro
 625 630 635 640
 agg acg ggc aag ggc aag tgg aag cgg aag tcg gca gga ggt ggc ccg 1968
 Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser Ala Gly Gly Gly Pro
 645 650 655
 agc agg gcc ggg tcc ccg cgc cgg aca tcc aag aaa acc aag gtg gag 2016
 Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys Lys Thr Lys Val Glu
 660 665 670
 ccc tac agt ctc acg gcc cag cag agc agc ctc atc aga gag gac aag 2064
 Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu Ile Arg Glu Asp Lys
 675 680 685
 agc aac gcc aag ctg tgg aat gag gtc ctg gcg tca ctc aag gac cgg 2112
 Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala Ser Leu Lys Asp Arg
 690 695 700
 ccg gcg agc ggc agc ccg ttc cag ttg ttc ctg agt aaa gtg gag gag 2160
 Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu Ser Lys Val Glu Glu
 705 710 715 720
 acg ttc cag tgt atc tgc tgt cag gag ctg gtg ttc cgg ccc atc acg 2208
 Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val Phe Arg Pro Ile Thr
 725 730 735
 acc gtg tgc cag cac aac gtg tgc aag gac tgc ctg gac aga tcc ttt 2256
 Thr Val Cys Gln His Asn Val Cys Lys Asp Cys Leu Asp Arg Ser Phe
 740 745 750
 cgg gca cag gtg ttc agc tgc cct gcc tgc cgc tac gac ctg ggc cgc 2304
 Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg Tyr Asp Leu Gly Arg
 755 760 765
 agc tat gcc atg cag gtg aac cag cct ctg cag acc gtc ctc aac cag 2352
 Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln Thr Val Leu Asn Gln
 770 775 780
 ctc ttc ccc ggc tac ggc aat ggc cgg tga 2382
 Leu Phe Pro Gly Tyr Gly Asn Gly Arg
 785 790

- <210> 2
- <211> 793
- <212> PRT
- 5 <213> *Homo sapiens*
- <400> 2

ES 2 378 527 T3

Met Trp Ile Gln Val Arg Thr Met Asp Gly Arg Gln Thr His Thr Val
1 5 10 15
Asp Ser Leu Ser Arg Leu Thr Lys Val Glu Glu Leu Arg Arg Lys Ile
20 25 30
Gln Glu Leu Phe His Val Glu Pro Gly Leu Gln Arg Leu Phe Tyr Arg
35 40 45
Gly Lys Gln Met Glu Asp Gly His Thr Leu Phe Asp Tyr Glu Val Arg

ES 2 378 527 T3

50																
Leu 65	Asn	Asp	Thr	Ile	Gln 70	Leu	Leu	Val	Arg	Gln 75	Ser	Leu	Val	Leu	Pro	80
His	Ser	Thr	Lys	Glu 85	Arg	Asp	Ser	Glu	Leu 90	Ser	Asp	Thr	Asp	Ser	Gly	95
Cys	Cys	Leu	Gly 100	Gln	Ser	Glu	Ser	Asp 105	Lys	Ser	Ser	Thr	His	Gly	Glu	110
Ala	Ala	Ala 115	Glu	Thr	Asp	Ser	Arg	Pro 120	Ala	Asp	Glu	Asp	Met	Trp	Asp	125
Glu	Thr	Glu	Leu	Gly	Leu	Tyr 135	Lys	Val	Asn	Glu	Tyr	Val	Asp	Ala	Arg	140
Asp	Thr	Asn	Met	Gly	Ala 150	Trp	Phe	Glu	Ala	Gln	Val	Val	Arg	Val	Thr	160
Arg	Lys	Ala	Pro	Ser 165	Arg	Asp	Glu	Pro	Cys	Ser	Ser	Thr	Ser	Arg	Pro	175
Ala	Leu	Glu	Glu 180	Asp	Val	Ile	Tyr	His	Val 185	Lys	Tyr	Asp	Asp	Tyr	Pro	190
Glu	Asn	Gly	Val 195	Val	Gln	Met	Asn	Ser	Arg	Asp	Val	Arg	Ala	Arg	Ala	205
Arg	Thr	Ile	Ile	Lys	Trp	Gln 215	Asp	Leu	Glu	Val	Gly	Gln	Val	Val	Met	220
Leu	Asn	Tyr	Asn	Pro	Asp 230	Asn	Pro	Lys	Glu	Arg	Gly	Phe	Trp	Tyr	Asp	240
Ala	Glu	Ile	Ser	Arg 245	Lys	Arg	Glu	Thr	Arg	Thr	Ala	Arg	Glu	Leu	Tyr	255
Ala	Asn	Val	Val 260	Leu	Gly	Asp	Asp	Ser	Leu	Asn	Asp	Cys	Arg	Ile	Ile	270
Phe	Val	Asp	Glu	Val	Phe	Lys	Ile	Glu	Arg	Pro	Gly	Glu	Gly	Ser	Pro	285
Met	Val	Asp	Asn	Pro	Met	Arg	Arg	Lys	Ser	Gly	Pro	Ser	Cys	Lys	His	300
Cys	Lys	Asp	Asp	Val	Asn 310	Arg	Leu	Cys	Arg	Val	Cys	Ala	Cys	His	Leu	320
Cys	Gly	Gly	Arg	Gln 325	Asp	Pro	Asp	Lys	Gln	Leu	Met	Cys	Asp	Glu	Cys	335
Asp	Met	Ala	Phe	His	Ile	Tyr	Cys	Leu	Asp	Pro	Pro	Leu	Ser	Ser	Val	345
Pro	Ser	Glu	Asp	Glu	Trp	Tyr	Cys	Pro	Glu	Cys	Arg	Asn	Asp	Ala	Ser	355
Glu	Val	Val	Leu	Ala	Gly	Glu	Arg	Leu	Arg	Glu	Ser	Lys	Lys	Asn	Ala	370
						375										380

ES 2 378 527 T3

Lys Met Ala Ser Ala Thr Ser Ser Ser Gln Arg Asp Trp Gly Lys Gly
 385 390 395 400
 Met Ala Cys Val Gly Arg Thr Lys Glu Cys Thr Ile Val Pro Ser Asn
 405 410 415
 His Tyr Gly Pro Ile Pro Gly Ile Pro Val Gly Thr Met Trp Arg Phe
 420 425 430
 Arg Val Gln Val Ser Glu Ser Gly Val His Arg Pro His Val Ala Gly
 435 440 445
 Ile His Gly Arg Ser Asn Asp Gly Ser Tyr Ser Leu Val Leu Ala Gly
 450 455 460
 Gly Tyr Glu Asp Asp Val Asp His Gly Asn Phe Phe Thr Tyr Thr Gly
 465 470 475 480
 Ser Gly Gly Arg Asp Leu Ser Gly Asn Lys Arg Thr Ala Glu Gln Ser
 485 490 495
 Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu Ala Leu Asn Cys
 500 505 510
 Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala Lys Asp Trp Arg
 515 520 525
 Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys Gly Gly Lys Asn
 530 535 540
 Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp Gly Ile Tyr Lys
 545 550 555 560
 Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly Phe Leu Val Trp
 565 570 575
 Arg Tyr Leu Leu Arg Arg Asp Asp Asp Glu Pro Gly Pro Trp Thr Lys
 580 585 590
 Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr Met Gln Tyr Pro
 595 600 605
 Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg Glu Lys Glu Asn
 610 615 620
 Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly Gly Phe Ala Ser Pro
 625 630 635 640
 Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser Ala Gly Gly Gly Pro
 645 650 655
 Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys Lys Thr Lys Val Glu
 660 665 670
 Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu Ile Arg Glu Asp Lys
 675 680 685
 Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala Ser Leu Lys Asp Arg
 690 695 700

ES 2 378 527 T3

Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu Ser Lys Val Glu Glu
 705 710 715 720
 Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val Phe Arg Pro Ile Thr
 725 730 735
 Thr Val Cys Gln His Asn Val Cys Lys Asp Cys Leu Asp Arg Ser Phe
 740 745 750
 Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg Tyr Asp Leu Gly Arg
 755 760 765
 Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln Thr Val Leu Asn Gln
 770 775 780
 Leu Phe Pro Gly Tyr Gly Asn Gly Arg
 785 790

- <210> 3
- <211> 45
- <212> ADN
- 5 <213> *Homo sapiens*
- <220>
- <221> CDS
- <222> (1)..(45)
- <400> 3

acc cac ggt gag gcg gcc gcc gag act gac agc agg cca gcc gat
 Thr His Gly Glu Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp
 1 5 10 15

45

- <210> 4
- <211> 15
- <212> PRT
- <213> *Homo sapiens*
- 15 <400> 4

Thr His Gly Glu Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp
 1 5 10 15

- <210> 5
- <211> 78
- <212> ADN
- 20 <213> *Homo sapiens*
- <220>
- <221> CDS
- <222> (1)..(78)
- <400> 5

ES 2 378 527 T3

atg gtt gac aac ccc atg aga cgg aag agc ggg ccg tcc tgc aag cac 48
 Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
 1 5 10 15

tgc aag gac gac gtg aac aga ctc tgc agc 78

Cys Lys Asp Asp Val Asn Arg Leu Cys Ser
 20 25

<210> 6

<211> 26

<212> PRT

5 <213> *Homo sapiens*

<400> 6

Met Val Asp Asn Pro Met Arg Arg Lys Ser Gly Pro Ser Cys Lys His
 1 5 10 15

Cys Lys Asp Asp Val Asn Arg Leu Cys Ser
 20 25

<210> 7

<211> 525

10 <212> ADN

<213> *Homo sapiens*

<220>

<221> CDS

<222> (1)..(525)

15 <400> 7

ES 2 378 527 T3

cga gag aag gag aac agc aag agg gag gag gag gag cag cag gag ggg 48
 Arg Glu Lys Glu Asn Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly
 1 5 10 15

ggc ttc gcg tcc ccc agg acg ggc aag ggc aag tgg aag cgg aag tcg 96
 Gly Phe Ala Ser Pro Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser
 20 25 30

gca gga ggt ggc ccg agc agg gcc ggg tcc ccg cgc cgg aca tcc aag 144
 Ala Gly Gly Gly Pro Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys
 35 40 45

aaa acc aag gtg gag ccc tac agt ctc acg gcc cag cag agc agc ctc 192
 Lys Thr Lys Val Glu Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu
 50 55 60

atc aga gag gac aag agc aac gcc aag ctg tgg aat gag gtc ctg gcg 240
 Ile Arg Glu Asp Lys Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala
 65 70 75 80

tca ctc aag gac cgg ccg gcg agc ggc agc ccg ttc cag ttg ttc ctg 288
 Ser Leu Lys Asp Arg Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu
 85 90 95

agt aaa gtg gag gag acg ttc cag tgt atc tgc tgt cag gag ctg gtg 336
 Ser Lys Val Glu Glu Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val
 100 105 110

ttc cgg ccc atc acg acc gtg tgc cag cac aac gtg tgc aag gac tgc 384
 Phe Arg Pro Ile Thr Thr Val Cys Gln His Asn Val Cys Lys Asp Cys
 115 120 125

ctg gac aga tcc ttt cgg gca cag gtg ttc agc tgc cct gcc tgc cgc 432

Leu Asp Arg Ser Phe Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg
 130 135 140

tac gac ctg ggc cgc agc tat gcc atg cag gtg aac cag cct ctg cag 480
 Tyr Asp Leu Gly Arg Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln
 145 150 155 160

acc gtc ctc aac cag ctc ttc ccc gcc tac ggc aat ggc cgg tga 525
 Thr Val Leu Asn Gln Leu Phe Pro Gly Tyr Gly Asn Gly Arg
 165 170 175

<210> 8

<211> 174

<212> PRT

5 <213> *Homo sapiens*

<400> 8

ES 2 378 527 T3

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

<210> 9
 <211> 324
 <212> ADN

5 <213> *Homo sapiens*

<400> 9

atgtggatcc aggttcggac catggatggg aggcagaccc acacgggtgga ctgctgtcc 60
 aggctgacca aggtggagga gctgaggcgg aagatccagg agctgttcca cgtggagcca 120

ggcctgcaga ggctgttcta caggggcaaa cagatggagg acggccatac cctcttcgac 180
 tacgaggtcc gcctgaatga caccatccag ctctgtgtec gccagagcct cgtgctcccc 240
 cacagcacca aggagcggga ctccgagctc tccgacaccg actccggctg ctgcctgggc 300
 cagagtgagt cagacaagtc ctcc 324

<210> 10
 <211> 495
 <212> ADN

10 <213> *Homo sapiens*

<400> 10

ES 2 378 527 T3

```

gaggacatgt gggatgagac ggaattgggg ctgtacaagg tcaatgagta cgctcgatgct 60
cgggacacga acatgggggc gtggtttgag gcgcaggtgg tcagggtgac gcggaaggcc 120
ccctcccggg acgagccctg cagctccacg tccaggccgg cgctggagga ggacgtcatt 180
taccacgtga aatacgacga ctaccggag aacggcgtgg tccagatgaa ctccagggac 240
gtccgagcgc gcgcccgcac catcatcaag tggcaggacc tggaggtggg ccaggtggtc 300
atgctcaact acaaccccga caacccaag gagcggggct tctggtacga cgcggagatc 360
tccaggaagc gcgagaccag gacggcgcgg gaactctacg ccaacgtggt gctgggggat 420
gattctctga acgactgtcg gatcatcttc gtggacgaag tcttcaagat tgagcggccg 480
ggtgaaggga gcccc                                     495

```

<210> 11

<211> 915

<212> ADN

5 <213> *Homo sapiens*

<400> 11

```

gtctgcgcct gccacctgtg cgggggcccgg caggaccccg acaagcagct catgtgcgat 60
gagtgcgaca tggcettcca catctactgc ctggaccgcg ccctcagcag tgttcccagc 120
gaggacgagt ggtactgccc tgagtgccgg aatgatgcc a gcgaggtggt actggcggga 180
gagcggctga gagagagcaa gaagaatgcg aagatggcct cggccacatc gtcctcacag 240
cgggactggg gcaagggcat ggctgtgtg ggccgcacca agaatgtac catcgtcccg 300
tccaaccact acggacccat cccggggatc cccgtgggca ccatgtggcg gttccgagtc 360
caggtcagcg agtcgggtgt ccatcggccc cacgtggctg gcatccatgg ccggagcaac 420
gacggatcgt actccctagt cctggcgggg ggctatgagg atgatgtgga ccatgggaat 480
tttttcacat acacgggtag tggtggtcga gatctttccg gcaacaagag gaccgcggaa 540
cagtcttgtg atcagaaact caccaacacc aacagggcgc tggtctcaa ctgctttgct 600
cccatcaatg accaagaagg ggccgaggcc aaggactggc ggtcggggaa gccggtcagg 660
gtggtgcgca atgtcaaggg tggcaagaat agcaagtacg cccccgctga gggcaaccgc 720
tacgatggca tctacaaggt tgtgaaatac tggcccgaga aggggaagtc cgggtttctc 780
gtgtggcgct accttctgcg gagggacgat gatgagcctg gcccttggac gaaggagggg 840
aaggaccgga tcaagaagct ggggctgacc atgcagtatc cagaaggcta cctggaagcc 900
ctggccaacc gagag                                     915

```

<210> 12

<211> 1366

10 <212> ADN

<213> *Homo sapiens*

<400> 12

ES 2 378 527 T3

```

ggcagcgttt gccgagcggg cgtccgggt cgcacgcaag tccgcgcggg gtccgggcca 60
cgcacgcggg ttcategcca tccccagccg ggccagggcg gcaggcagac aagctgttcg 120
cggcgaccgg agaggtgagc gggcgggccc ggtcggggtg ccagcccggg ccggggcgac 180
ggggctcggg aactttgcaa aactttcccg cgcggccagc ccggggcgac gcatgtcccg 240
cactctgtcc cgggatccag ggccctccct tccacctaac cctcgggaat cgttcccccg 300
cacacatccg gctggagccg ggaccagcgc tgcgtccccg gagcccggcg gggggtcgag 360
cgcgcggggg gggggagggc ctggcgagcc gccggggagg atgtcaggct ccgcgcctgc 420
gcgcggggcg ccccgcgatt caattgtcgc gcccagagccc gatttcgcgc gccctgagtt 480

ccccgggagc atctgggcca atggggagcg agcggggcgg ggcggccggg tgctgcggag 540
ccaataagag gcggtcaag tgaagggggg cgggacttga cgagcggggg ccccctctgt 600
agtcccggcg gcgggggtgg gcgtgggctc gctggcgca cccgcgcggg ccagtgggag 660
tgcgggaggg acgccgaggg tccagggtt ggagggggcg gagctgccgg gggttggagg 720
tcgaggtgag tcgcggggcg cgcgcgctcg cgggtggccg ggacggggcg cggttacat 780
ggccaccgcg gggcgggccc ggtcgcgcac gcgcgcgggg ggggccggca aggagggggg 840
gcgtgggcac cgaggggtcc cggggtccgc ggatctcggg tggggtttt cccatttcag 900
tggcaattgg ttaagttccc ccgggacct ctgaagttcc ggcccgcgct ggactttctg 960
ggattecctc ttccgtaaat aggaatccga ggaatgaat aatcaatgaa tgaatgaata 1020
aacgaaccaa ctcgggccac ttggcccggg cctcctttct cctctggtcg tggggaagga 1080
gggatgggtt ggaccttctg cttttctttc aatccctct ttcattctc ctccctctc 1140
aatcttcaac acttggtag tcgttaatgc cttaagtgt taattgttg tgtctggtcc 1200
tggccagggt ctggtgtac aggaggactg gaagggcatc ctgggagttt cctggtgtcc 1260
acaggccgga caaaagcaac cccgactcct tagagcatgg catggctcag aggtgctggt 1320
aaaactgatg ggggttatg ctgtccctcc cctcagcgcc gacacc 1366

```

REIVINDICACIONES

1. Polipéptido aislado denominado ICBP90 (inverted CCAAT box, binding protein 90) de secuencia aminoacídica SEQ ID N°2.
- 5 2. Polipéptido aislado caracterizado porque comprende un polipéptido elegido entre los polipéptidos de secuencias SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8 o el fragmento C-terminal a partir del residuo D263 de la SEQ ID N°2.
3. Polipéptido según la reivindicación 1 o 2, caracterizado porque la secuencia de ADN con la que se liga dicho polipéptido es una secuencia CCAAT, preferiblemente una secuencia CCAAT inversa (inverted ccaat box:icb ICB).
- 10 4. Polinucleótido aislado caracterizado porque codifica un polipéptido según la reivindicación 1.
5. Polinucleótido según la reivindicación 4 de secuencia SEQ ID N°1.
6. Polinucleótido aislado caracterizado porque comprende un polinucleótido elegido entre un polinucleótido de secuencias SEQ ID N°1, SEQ ID N°3, SEQ ID N°5, SEQ ID N°7 o que codifica el fragmento C-terminal a partir del residuo D263 de la SEQ ID N°2.
- 15 7. Vector recombinante de clonación de un polinucleótido según una de las reivindicaciones 4 a 6 y/o de expresión de un polipéptido según una de las reivindicaciones 1 a 3, caracterizado porque contiene un polinucleótido según una cualquiera de las reivindicaciones 4 a 6.
8. Vector según la reivindicación 7 caracterizado porque incluye los elementos que permiten la expresión y eventualmente la secreción de dicho polipéptido en una célula hospedadora.
- 20 9. Vector según una cualquiera de las reivindicaciones 7 a 8, caracterizado porque los elementos que permiten la expresión de dicho polipéptido se eligen entre:
 - a) el polinucleótido aislado de secuencia SED ID N°12;
 - b) un polinucleótido cuya secuencia es complementaria de la secuencia del polinucleótido definido en a);
 - 25 c) un polinucleótido cuya secuencia comprende al menos un 80% de identidad con un polinucleótido definido en a) o b);
 - d) un polinucleótido que hibrida en condiciones de alto rigor con una secuencia del polinucleótido definido en a), b) o c).
10. Célula hospedadora caracterizada porque comprende un vector según una de las reivindicaciones 7, 8 y 9.
- 30 11. Procedimiento de preparación de un polipéptido recombinante caracterizado porque se cultiva una célula hospedadora según la reivindicación 10 en condiciones que permiten la expresión y eventualmente la secreción de dicho polipéptido recombinante y porque se recupera dicho polipéptido recombinante.
12. Polipéptido recombinante susceptible de obtenerse mediante un procedimiento según la reivindicación 11.
13. Anticuerpo monoclonal o policlonal y sus fragmentos, caracterizado porque se liga específicamente con un polipéptido elegido entre: SEQ ID NO: 2, 4, 6 o 8.
- 35 14. Procedimiento de detección y/o valoración de un polipéptido según una de las reivindicaciones 1 a 3 en una muestra biológica, caracterizado porque comprende las etapas siguientes:
 - a) puesta en contacto de la muestra biológica con un anticuerpo según la reivindicación 13;
 - b) determinación del complejo antígeno-anticuerpo formado.
- 40 15. Estuche para la aplicación de un procedimiento según la reivindicación 14 en una muestra biológica mediante reacción inmunológica, caracterizado porque comprende los elementos siguientes:
 - a) un anticuerpo monoclonal o policlonal según la reivindicación 13;
 - b) llegado el caso, los reactivos para la constitución del medio propicio a la reacción inmunológica;
 - c) los reactivos que permiten la detección del complejo antígeno-anticuerpo producido por la reacción inmunológica.
- 45 16. Procedimiento según la reivindicación 14 para el diagnóstico de proliferación celular.

17. Procedimiento de cribado de ligando susceptible de afectar a la actividad transcripcional de un gen cuyo promotor comprende las secuencias CCAAT y/o CCAAT inversa (ICB) susceptibles de ligarse con un polipéptido según las reivindicaciones 1 a 8 y que incluye las etapas siguientes:
- 5 a) puesta en contacto de dicho polipéptido y de uno o varios ligandos potenciales en presencia de los reactivos necesarios para la aplicación de una reacción de transcripción;
- b) detección y/o medida de la actividad transcripcional.
18. Procedimiento de cribado de ligando susceptible de afectar a la función de "receptor nuclear" del polipéptido según las reivindicaciones 1 a 3 y que comprende las etapas siguientes:
- 10 a) puesta en contacto de dicho polipéptido y de uno o varios ligandos potenciales en presencia de los reactivos necesarios;
- b) detección y/o medida de la actividad transcripcional de un gen cuyo promotor comprende las secuencias CCAAT y/o CCAAT inversa (ICB) susceptibles de ligarse con dicho polipéptido.
19. Estuche para llevar a cabo un procedimiento según las reivindicaciones 17 y 18 en una muestra biológica, caracterizado porque comprende los elementos siguientes:
- 15 a) un polipéptido según las reivindicaciones 1 a 3;
- b) un ligando;
- c) los reactivos necesarios para llevar a cabo una reacción de transcripción.
20. Compuesto caracterizado porque se elige entre:
- a) un polipéptido según una de las reivindicaciones 1 a 3;
- 20 b) un polinucleótido según una de las reivindicaciones 4 a 6;
- c) un vector según una de las reivindicaciones 7 a 9;
- d) una célula según la reivindicación 10;
- e) un anticuerpo según la reivindicación 13;
- a modo de medicamento.
- 25 21. Compuesto según la reivindicación 20 a modo de medicamento destinado a la prevención y/o el tratamiento del cáncer.
22. Utilización de un compuesto según las reivindicaciones 20 y 21 para la preparación de un medicamento destinado a modular, aumentar o disminuir la proliferación celular.
- 30 23. Composición farmacéutica para el tratamiento preventivo y curativo del cáncer, caracterizada porque contiene una cantidad terapéuticamente eficaz de un compuesto según una de las reivindicaciones 20 y 21 y un vehículo farmacéuticamente aceptable.
24. Composición farmacéutica caracterizada porque comprende un anticuerpo según la reivindicación 13 como agente orientador conjugado con al menos un agente seleccionado entre el grupo de agentes antiproliferativos, antineoplásicos o citotóxicos.
- 35 25. Producto que comprende al menos un compuesto según las reivindicaciones 20 y 21 y al menos otro agente anticanceroso como producto de combinación para una utilización simultánea, separada o extendida en el tiempo en terapia anticancerosa.
- 40 26. Composición para la detección, localización y formación de imágenes de cánceres, que comprende un anticuerpo según la reivindicación 13, estando el anticuerpo marcado directa o indirectamente con un marcador generador de señal seleccionado entre los isótopos radiactivos y las entidades no isotópicas.

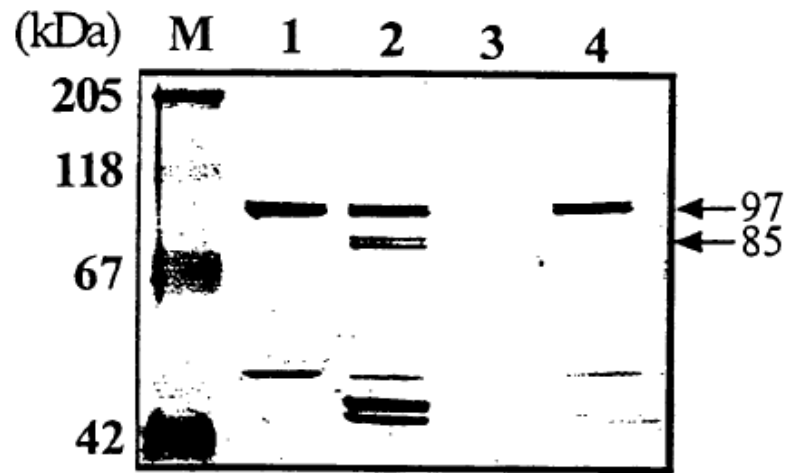


FIG.1

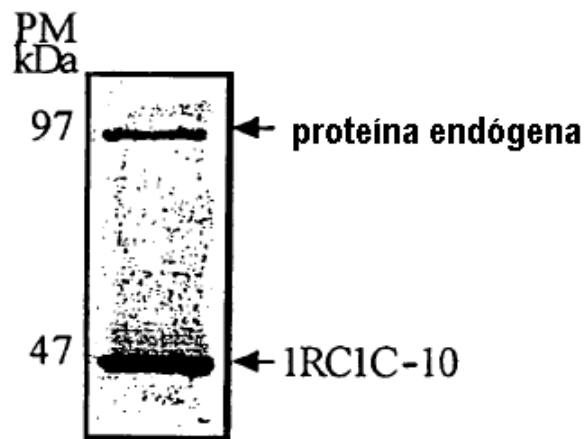


FIG. 2



FIG. 3



FIG. 4

FIG_5

	1	2	3	4	5	6	7	8
A								
B								
C								
D								
E								
F								
G								
H								
A	cerebro completo	amígdala	núcleo caudado	cerebelo	corteza cerebral	lóbulo frontal	hipocampo	bulbo raquídeo
B	lóbulo occipital	putamen	sustancia negra	lóbulo temporal	tálamo	núcleo subtalámico	médula espinal	
C	corazón	aorta	músculo esquelético	colon	vejiga	útero	próstata	estómago
D	testículo	ovario	páncreas	glándula pituitaria	glándula suprarrenal	glándula tiroidea	glándula salival	glándula mamaria
E	riñón	hígado	intestino delgado	bazo	tímo	leucocito periférico	ganglio linfático	médula ósea
F	apéndice	pulmón	arteria traqueal	placenta				
G	cerebro fetal	corazón fetal	riñón fetal	hígado fetal	bazo fetal	tímo fetal	pulmón fetal	
H	ARN total de levadura 100 ng	ARNt de levadura 100 ng	ARNr de E. coli 100 ng	ADN de E. coli 100 ng	Poli r(A) 100 ng	ADN de Cot humano 100 ng	ADN humano 100 ng	ADN humano 500 ng

1 ATGTGGATCC AGGTTGGAC CATGGATGG AGGCAGACC 40 | 50 | 60 | 70 | 80
 81 GCTGAGGGG AAGATCCAGG AGCTGTTCCA CGTGGAGCCA ACACGGTGA CTCGCTGTCC AGGCTGACCA AGTGGAGGA 80
 161 ACGGCCATAC CCTCTTCGAC TACGAGGTCC CCTGAATGA GGCCTGCAGA GTCCTTCTA CAGGGGCAA CATGTGAGG 160
 241 CACAGCACCA AGGAGCGGGA CTCCGAGTTC TCCGACAACG CACTATCCAG CTCCTGGTCC GCCAGAGCCT CGTGTCCCC 240
 321 CTCACCCAC GGTGAGCGG CCGCCGAGAC TGACAGCAAG ACTCCGGCTG CTGCCCTGGC CAGAGTGAGT CAGACRAGTC 320
 401 TGTACAAGGT CAATGAGTAC GTCGATGCTC GGGACACGAA CCAGCCGATG AGGACATGTG GATGAGAGC GAATGGGGC 400
 481 CGGAAGGCCC CCTCCGGGA CGAGCCCTGC AGCTCCACGT CATGGGGCG TGGTTTGAGG CGCAGGTGGT CAGGTGACG 480
 561 ATACGACGAC TACCCGAGA ACGGCGTGGT CCAGATGAAC CCAGGCCGC GCTGGAGGAG GACGTCATTT ACCACGTGAA 560
 641 GGCAGGACTT GGAGGTGGC CAGGTGGTCA TGCTAACTA TCCAGGAGC TCCGAGCGG CGCCCCAC ATCATCAAGT 640
 721 GCGGAGATCT CCAGGACCG CGAGACGAG ACGGCCGGG AACTCTAOC CAACGTGGT CTGGGGGATG ATTCCTGAA 800
 801 CGACTGTCGG ATCATCTTCG TGGACGAGT CTTCAAGATT GAGCGCCGG GTGAAGGGAG CCCATGGTT GACNACCCA 880
 881 TGAGACGGAA GAGCGGCGG TCCTGCAAGC ACTGCAAGGA CGACGTGAC AGACTGTGA GGTCTGCG CTGCCACTG 960
 961 TCGGGGGCC GGCAGNCCC CGACAAGCAG CTCATGTGG ATGAGTGGA CATGGCTTC CACATCTACT GCCTGGACCC 1040
 1041 GCCCCTCAGC AGTGTCCCA GCGAGGACGA GTGGTACTGC CTTGAGTGC GGAATGATC CAGCGAGGTG GTACTGGCGG 1120
 1121 GAGACCGCTT GAGAGAGAG AAGAAGATG CGAAGATGGC CTCGGCCACA TCGTCTCAC AGCGGACTG GGGCAAGGC 1200
 1201 ATGGCTGTG TGGCCGAC CAAGGANTGT ACCATCGTCC CGTCCNACCA CTACGGACOC ATCCCCGGGA TCCCCGTGG 1280
 1281 CACCATGTGG CGGTTCGAG TCCAGGTCAG CGAGTCGGGT GTCCATCGC CCCACGTGC TGGCATCCAT GGCCGGAGCA 1360
 1361 ACGAOGGATC GTACTCCCTA GTCCTGCGG GGGGCTATGA GGATGATG GACCATGGGA ATTTTTTCAC ATACAGGGT 1440
 1441 AGTGTGGTC GAGATCTTC CGGCAACCAAG AGGACCGGG AACAGTCTG TGATCAGAA CTCACCAACA CCAACAGGC 1520
 1521 GCTGGCTCTC AACTGCTTTG CTCCCATCAA TGACCAGAA GGGCCGAGG CCAAGACTG GCGGTGCGG AAGCCGTCA 1600
 1601 GGTGTGTGG CAATGTCAAG GGTGGCAAGA ATAGCAAGTA CGCCCCCGT GAGGGCAAC GCTACGATGG CATCTACAAG 1680
 1681 GTTGTGAAAT ACTGGCCGA GAAGGGGAG TCCGGGTTTC TCGTGTGGG CTACCTTCTG CGGAGGGAGC ATGATGAGCC 1760
 1761 TGGCCCTTGG ACGAAGGAG GGAAGGACC GATCAAGAA CTGGGGCTGA CCATGCAATA TCCAGAAGGC TACCCTGGAAG 1840
 1841 CCTGGCCAA CCGAGAGCGA GAGAAGGAGA ACAGCAAGAG GGAGGAGGAG GAGCAGCAG AGGGGGGCTT CGCGTCCCC 1920
 1921 AGGACGGGA AGGGCAAGT GAAGCGGAAG TCGGJAGGAG GTGGCCCGAG CAGGCCGGG TCCCCCGCC GGACATCCAA 2000
 2001 GAAACCAAG GTGGACCCCT ACAGTCTCAC GGCCACGAG AGCAGCCTCA TCAGAGAGGA CAAGAGCAAC GCCAAGCTGT 2080
 2081 GGAATGAGGT CCTGGGTCA CTCAAGGACC GGCAGGCGAG CGCAGCCCG TTCCAGTTGT TCCGTGAGTAA AGTGGAGGAG 2160
 2161 ACGTTCCAGT GTATCTGCTG TCAGGAGCTG GTGTTCCGGC CCATCACGAC CGTGTGCCAG CACACCTGT GCAAGGACTG 2240
 2241 CCTGGACAGA TCCTTTGGG CACAGGTGT CAGCTGCCCT GCCTGCCGT ACGACCTGGG CGCAGCTAT GCCATGACAGG 2320
 2321 TGAACCAAGC TCTGCAGACC GTCTCAACC AGCTCTTCCC CGGCTACGG ANTGGCGGT GA 2382

FIG-6

		10		20		30		40		50		60		70		80																																																																
1	MW	IQR	TMDG	RQ	HTV	DLSL	RL	TKV	EELRR	KIQ	ELF	HVEP	GL	QLF	YR	GK	QM	EDG	H	T	L	F	D	YE	VR	L	N	D	T	I	Q	LL	V	R	Q	S	L	V	L	P	80																																							
81	HST	KER	DSEL	SD	T	S	G	C	CLG	Q	S	E	D	K	S	S	T	H	G	E	A	A	E	T	D	S	R	PA	ED	M	W	D	E	T	EL	G	L	Y	K	V	N	E	Y	VD	AR	D	T	N	M	G	A	WF	E	A	Q	V	V	R	V	T	160																			
161	RK	AP	S	R	D	E	P	C	S	S	T	S	R	P	A	L	E	D	V	I	Y	H	V	K	Y	D	Y	P	E	N	G	V	V	Q	M	N	S	R	D	V	R	A	R	A	R	T	I	I	K	W	Q	L	E	V	G	Q	V	V	M	L	N	Y	N	P	D	NP	K	E	R	G	F	W	Y	D	240					
241	AE	I	S	R	K	R	E	T	R	T	A	R	E	L	Y	A	N	V	V	L	G	D	D	S	L	N	D	C	R	I	I	F	V	D	E	V	F	K	I	ER	P	G	E	G	S	P	M	V	D	N	P	M	R	R	K	S	G	P	S	C	H	C	K	D	D	D	V	N	RL	C	R	V	C	A	C	H	L	320		
321	CG	GR	Q	P	D	K	Q	L	M	C	D	E	C	D	M	A	F	H	I	Y	C	L	D	P	P	L	S	S	V	P	S	E	D	E	W	Y	C	PE	C	R	N	D	A	S	E	V	V	L	A	G	E	R	L	R	E	S	K	K	N	A	K	M	A	S	A	T	S	S	S	Q	R	D	W	G	K	G	400			
401	MA	C	V	G	R	T	K	E	C	T	I	V	P	S	N	H	Y	G	P	I	P	G	I	P	V	G	T	M	W	R	F	R	V	Q	V	S	E	S	G	V	H	R	P	H	V	A	G	I	H	G	R	S	N	D	G	S	Y	S	L	V	L	A	G	G	Y	E	D	D	V	D	H	G	N	F	F	T	Y	T	G	480
481	SG	GR	D	L	S	G	N	K	R	T	A	E	Q	S	C	D	Q	K	L	T	N	T	N	R	A	L	A	L	N	C	F	A	P	I	N	D	Q	E	G	A	E	A	K	D	W	R	S	G	K	P	V	R	V	V	R	N	V	K	G	G	K	N	S	K	Y	A	P	A	E	G	N	R	Y	D	G	I	Y	K	560	
561	V	K	Y	W	P	E	K	G	K	S	G	F	L	V	W	R	Y	L	L	R	R	D	D	E	P	G	P	W	T	K	E	G	K	D	R	I	K	K	L	G	L	T	M	Q	Y	P	E	G	Y	L	E	A	L	A	N	R	R	E	E	K	E	N	S	K	R	E	E	E	E	Q	Q	E	G	G	F	A	S	P	640	
641	RT	G	K	G	K	W	R	R	K	S	A	G	G	G	P	S	R	A	G	S	P	R	R	T	S	K	T	K	V	E	P	Y	S	L	T	A	Q	Q	S	S	L	I	R	E	D	K	S	N	A	K	L	W	N	E	V	L	A	S	L	K	D	R	P	A	S	A	G	S	P	F	Q	L	F	L	S	K	V	E	E	720
721	TF	Q	C	I	C	C	Q	E	L	V	F	R	P	I	T	T	V	C	Q	H	N	V	C	K	D	C	L	D	R	S	F	R	A	Q	V	F	S	C	P	A	C	R	Y	D	L	G	R	S	Y	A	M	Q	V	N	Q	P	L	Q	T	V	L	N	Q	L	F	P	P	G	Y	G	N	G	R	*	794					

FIG-7

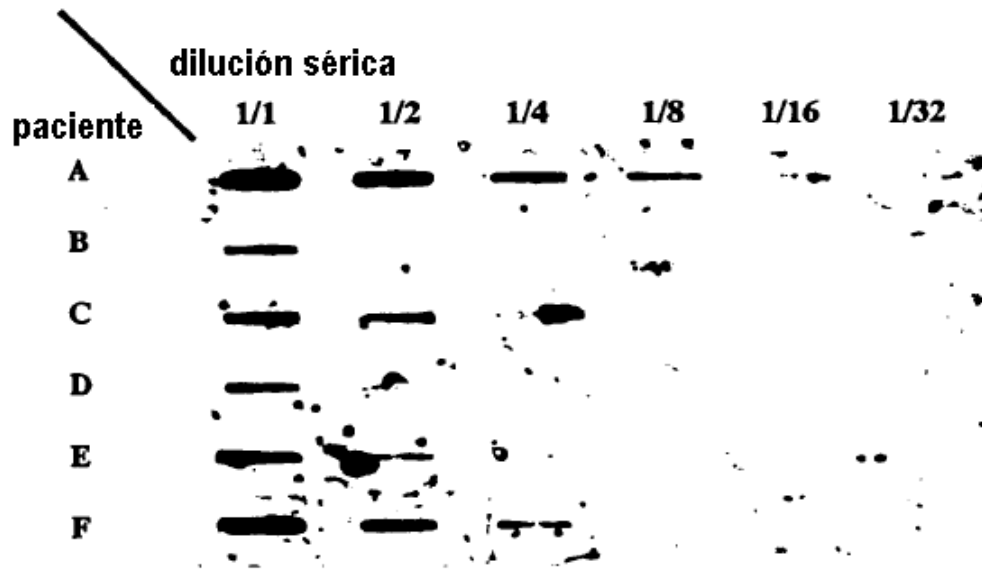
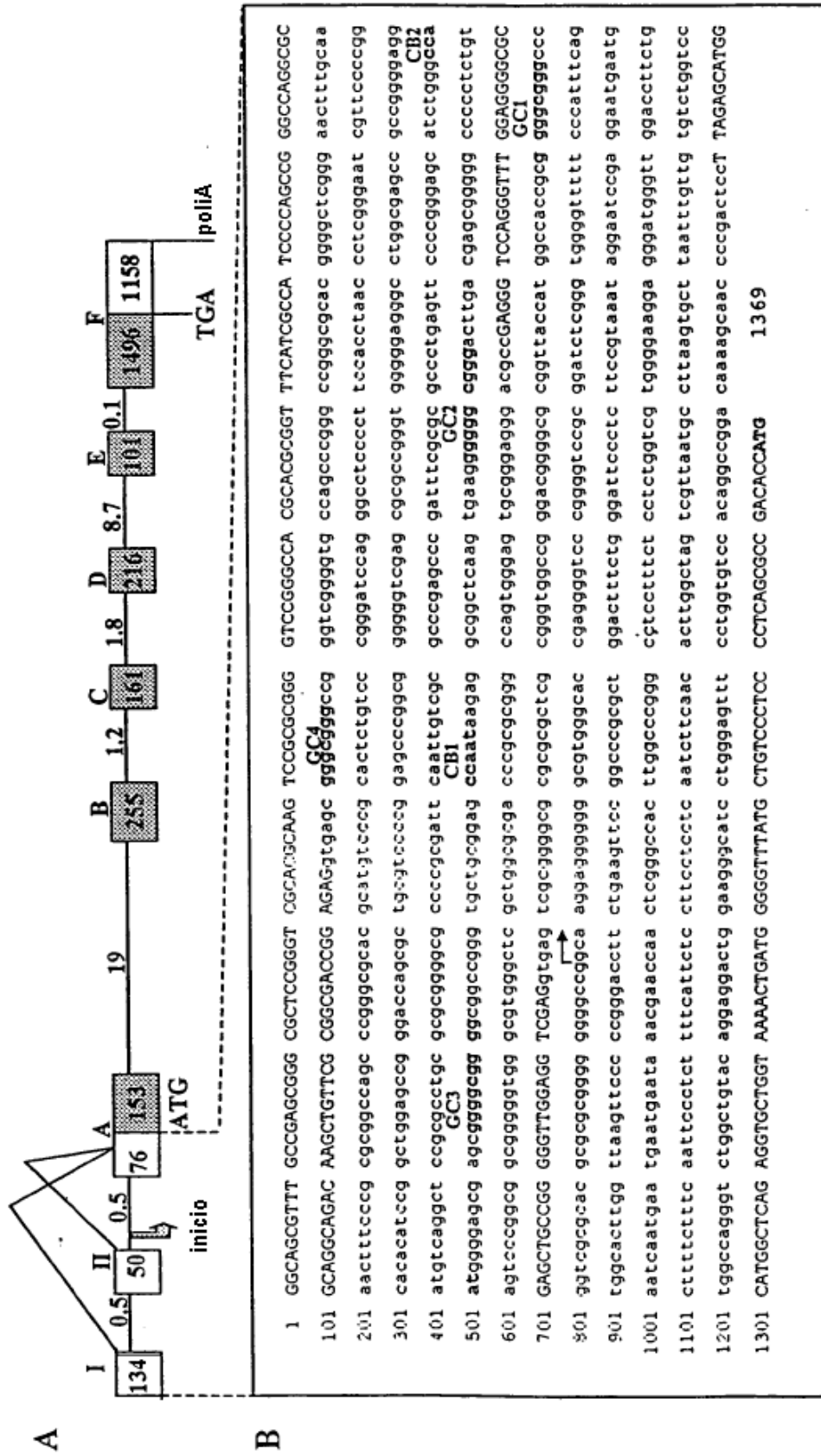


FIG-8

FIG-9



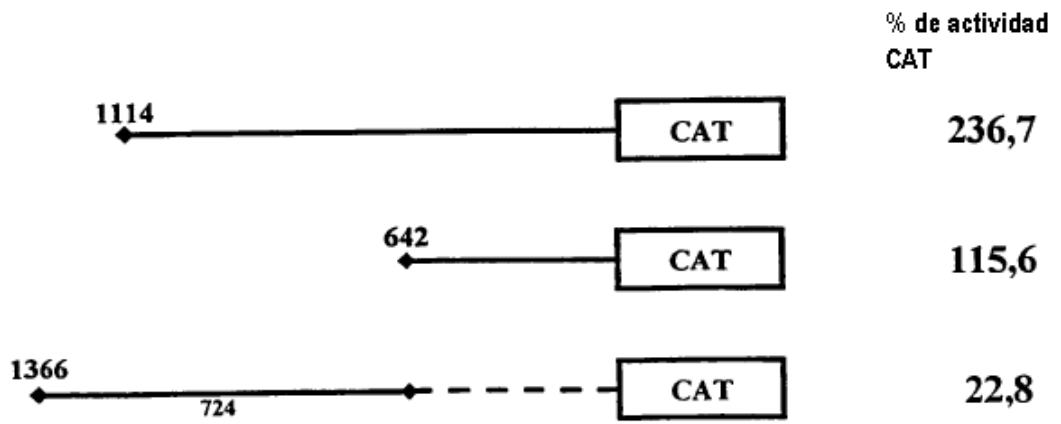


FIG 10

