



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 329 467**

51 Int. Cl.:

**C12N 15/12** (2006.01)

**A61K 38/18** (2006.01)

**A61K 39/395** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04022325 .7**

96 Fecha de presentación : **03.02.1999**

97 Número de publicación de la solicitud: **1510579**

97 Fecha de publicación de la solicitud: **02.03.2005**

54 Título: **Utilización de heterogubulina como factor de crecimiento de células epiteliales.**

30 Prioridad: **04.02.1998 US 20598**

45 Fecha de publicación de la mención BOPI:  
**26.11.2009**

45 Fecha de la publicación del folleto de la patente:  
**26.11.2009**

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

72 Inventor/es: **Sliwkowski, Mark y**  
**Kern, Jeffrey A.**

74 Agente: **Ponti Sales, Adelaida**

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

## DESCRIPCIÓN

Utilización de heterogubulina como factor de crecimiento de células epiteliales.

## 5 Antecedentes de la invención

## Campo de la invención

Esta invención hace referencia a la utilización de ligandos de HER2, HER3 y/o HER4, en particular a polipéptidos de heregulina como factores de crecimiento de células epiteliales.

## Descripción de los antecedentes y de materias relacionadas

La familia HER (ErbB) pertenece a la superfamilia de subclase I de receptores tirosina quinasa y consiste en tres receptores distintos, HER2, HER3 y HER4. Uno de los ligandos de esta familia ErbB es la proteína heregulina (HRG), una proteína que contiene un multidominio con al menos 15 isoformas distintas.

La transducción de señales que regulan el crecimiento y la diferenciación celular se regula en parte por la fosforilación de varias proteínas celulares, siendo las proteínas tirosina quinasas las enzimas que catalizan este proceso. Se cree que los receptores protein tirosina quinasa dirigen directamente el crecimiento celular mediante la fosforilación de tirosinas estimulada por ligandos de sustratos intracelulares. En los receptores protein tirosina quinasa del factor de crecimiento de la subfamilia de clase I se incluye el receptor del factor de crecimiento epidérmico de 170 KDa (EGFR), que está codificado por el gen *erbB1*. El *erbB1* está implicado causalmente en procesos malignos en humanos. Por ello, se ha observado un incremento de la expresión de este gen en los carcinomas más agresivos de mama, de vejiga, de pulmón y de estómago.

El segundo miembro de la subfamilia de clase I, *p<sup>158neu</sup>* se identificó originalmente como el producto del gen transformante de neuroblastomas de ratas tratadas químicamente. El gen *neu* (también denominado *erbB2* y HER2) codifica un receptor protein quinasa de 185 KDa. La amplificación y la sobreexpresión del gen HER2 humano se correlaciona con un pronóstico peor tanto en el cáncer de mama como en el de ovario (Slamon y col., Science 235:177-182 (1987); y Slamon y col., Science 244:707-712 (1989). La sobreexpresión de HER2 se ha correlacionado con otros carcinomas incluyendo los carcinomas de estómago, endometrio, glándula salivar, pulmón, riñón, colon y vejiga. Según ello, Slamon y col., en la patente americana 4.968.603 se describen y reivindican los ensayos diagnósticos para la determinación de la amplificación o la expresión del gen HER2 en células tumorales. Slamon y colaboradores descubrieron que la presencia de múltiples copias génicas del oncogene HER2 en las células tumorales era una indicación de que la enfermedad probablemente se propaga más allá del foco tumoral primario, con lo que, en dicho caso, la enfermedad puede requerir un tratamiento más agresivo que el indicado por otros factores diagnósticos. Slamon y colaboradores concluyeron que el análisis de la amplificación del gen HER2, junto con la determinación del estado del nódulo linfático, proporciona una mejor utilidad pronóstica.

También se ha descrito otro gen relacionado, denominado *erbB3* o HER3. Véase la patente americana 5.183.884; Kraus y col., Proc. Natl. Acad. Sci. 86:9193-9197 (1989); la patente europea EP 444.961A1; y Kraus y col., Proc. Natl. Acad. Sci. USA 90:2900-2904 (1993). Kraus y col., (1989) descubrieron que los niveles acentuadamente elevados del mRNA de *erbB3*, presentes en ciertas líneas celulares de tumores mamarios humanos, eran indicadores de que el *erbB3*, al igual que el *erbB1* y el *erbB2* podían desempeñar una función en los procesos malignos en el hombre. También, Kraus y col., (1993) mostraron que la activación dependiente de EGF del dominio catalítico de *erbB3* de un receptor quimérico EGFR/*erbB3* daba como resultado una respuesta proliferativa en las células NIH-3T3 transfectadas. Actualmente se piensa que es el resultado del *erbB1* o *erbB2* endógenos sobre NIH3T3. Además, estas investigaciones demostraron que algunas líneas celulares de tumores mamarios humanos presentan un aumento de la fosforilación de tirosina de *erbB3* en el estado de equilibrio, lo que indica que este receptor puede desempeñar una función en los procesos malignos en el hombre. La función del *erbB3* en el cáncer también ha sido explorada por otros investigadores, hallándose sobreexpresado en el cáncer de mama (Lemoine y col., Br. J Cancer 66:1116-1121 (1992)), cáncer gastrointestinal (Poller y col., J. Pathol. 168:275-280 (1992), Rajkumar y col., J. Pathol. 170:271-278 (1993), y Sanidas y col., Int. J Cancer 54:935-940 (1993)) y en cánceres de páncreas (Lemoine y coll., J. Pathol. 168:269-273 (1992) y Fries y col., Clinical Cancer Research 1:1413-1420 (1995)).

La subfamilia de clase I de los receptores proteína tirosina quinasa del factor de crecimiento ha sido ampliada posteriormente para incluir al receptor HER4/*erbB4*. Véase la patente europea 599.274; Plowman y col., Proc. Natl. Acad. Sci. USA 90:1746-1750 (1993); y Plowman y col., Nature 366:473-475 (1993). Plowman y col., hallaron que el aumento de la expresión de HER4 se correlacionaba estrechamente con ciertos carcinomas de origen epitelial, incluyendo los adenocarcinomas de mama. Los procedimientos diagnósticos para la detección de las condiciones neoplásicas en el hombre (especialmente los cánceres de mama), que evalúan la expresión de HER4 se describen en la patente europea 599.274.

La búsqueda del activador del oncogen HER2 ha conducido al descubrimiento de una familia de polipéptidos de heregulina. Estas proteínas parecen ser el resultado del ajuste alternativo de un solo gen que se ha localizado mediante mapeo en el brazo corto del cromosoma 8 humano por Orr-Urtreger y col., Proc Natl Acad Sci. USA 90:1867-1871 (1993). Véase también Lee y Wood, Genomics, 16:790-791 (1993).

Holmes y col., aislaron y clonaron una familia de activadores polipeptídicos para el receptor HER2 que denominaron heregulina- $\alpha$  (HRG- $\alpha$ ), heregulina- $\beta$ 1 (HRG- $\beta$ 1), heregulina- $\beta$ 2 (HRG- $\beta$ 2), de tipo-herregulina- $\beta$ 2 (tipo-HRG- $\beta$ 2) y heregulina- $\beta$ 3 (HRG- $\beta$ 3). Véase Holmes y col., Science 256:1205-1210 (1992); patente internacional WO 92/20798; y patente americana U.S. 5.367.060. El polipéptido de 45 Kda, HRG- $\alpha$ , se purificó a partir del medio condicionado de la línea celular de cáncer de mama MDA-MB-231. Estos investigadores demostraron la capacidad de los polipéptidos de heregulina purificados para activar la fosforilación de tirosinas del receptor HER2 en células de tumor mamario MCF7. Además, se demostró la actividad mitogénica de los polipéptidos de heregulina sobre las células SK-BR-3 (que expresan niveles elevados del receptor HER2). Al igual que otros factores de crecimiento que pertenecen a la familia EGF, los polipéptidos HRG solubles parecen derivar de un precursor unido a membrana (denominado pro-HRG) que se procesa proteolíticamente para liberar la forma soluble de 45 KDa. Estos pro-HRGs carecen de un péptido señal N-terminal.

Mientras que las heregulinas son idénticas en los primeros 213 residuos aminoacídicos, se clasifican en dos tipos principales,  $\alpha$  y  $\beta$ , basándose en dos dominios similares al EGF variantes que difieren en sus porciones C-terminales. Sin embargo, estos dominios similares al EGF son idénticos en la región que contiene los seis residuos de cisteína. Según una comparación de secuencia aminoacídica, Holmes y col., hallaron que entre la primera y la sexta cisteína en el dominio de tipo-EGF, las HRGs eran similares en un 45% al factor tipo-EGF de unión a la heparina (HB-EGF), idénticos en un 35% a la amfiredulina (AR), idénticos en un 32% al TGF- $\alpha$  y en un 27% al EGF.

El factor de diferenciación *neu* de 44KDa (NDF), el equivalente de rata al humano HRG, fue el primero que se describió por Peles y col., Cell, 69:205-216 (1992); y Wen y col., Cell, 69:559-572 (1992). Al igual que los polipéptidos HRG, el NDF tiene un dominio homólogo con la inmunoglobulina (1G) seguido por un dominio de tipo-EGF y carece de un péptido señal N-terminal. Wen y col., J Mol. Cell Biol., 14(3):1909-1919 (1994) llevaron a cabo un “clonaje exhaustivo” para ampliar la familia de los NDFs. Este trabajo demostró seis pro-NDFs fibroblásticos distintos. Adoptando la nomenclatura de Holmes y col., los NDFs se clasificaron como polipéptidos  $\alpha$  o  $\beta$  según la secuencia de los dominios de tipo-EGF. Las isoformas 1 a 4 se caracterizan por la longitud variable del fragmento de membrana (entre el dominio de tipo-EGF y el dominio transmembrana). También, se describen las isoformas a, b y c que tienen dominios citoplasmáticos de longitud variable. Estos investigadores concluyen que las distintas isoformas de NDF se generan mediante ajuste alternativo y realizan funciones específicas de tejido distintas. Véase también la patente europea 505.148; la patente internacional WO 93/22424; y la patente internacional 94/28133 referentes al NDF.

Falls y col., Cell, 72:801-815 (1993) describen otro miembro de la familia de heregulinas que denominaron polipéptidos con actividad inductora del receptor de acetil-colina (ARIA). El polipéptido ARIA del pollo estimula la síntesis de los receptores de acetilcolina del músculo. Véase la patente internacional WO 94/08007. La ARIA es una heregulina de tipo  $\beta$  que carece de la región espaciadora entera rica en sitios de glicosilación entre el dominio de tipo-Ig y el dominio de tipo-EGF de la HRG $\alpha$  y la HRG $\beta$ 1- $\beta$ 3.

Marchionni y col., Nature, 362:313-318 (1993) identificaron varias proteínas bovinas que denominaron factores de crecimiento gliales (GGFs). Estos GGFs comparten el dominio de tipo-Ig y el dominio de tipo-EGF con las otras proteínas de heregulina descritas anteriormente, pero también un dominio *kringle*. Los GGFs, generalmente, no tienen la región espaciadora glicosilada de forma completa entre el dominio de tipo-Ig y el dominio de tipo-EGF. Únicamente uno de los GGFs, GGFII, posee un péptido señal N-terminal. Véase también la patente internacional WO 94/00140; la patente internacional WO 94/04560; la patente internacional WO 94/26298; y la WO 95/32724 que hacen referencia a los GGFs y sus utilidades.

Ho y col., J Biol Chem 270(4):14523-14532 (1995) describen otro miembro de la familia de heregulinas denominado factor derivado de las neuronas motoras y sensoriales (SMDF). Esta proteína tiene un dominio de tipo-EGF característico de todos los polipéptidos de heregulina, pero un dominio N-terminal distinto. La diferencia estructural principal entre el SMDF y los otros polipéptidos de heregulina es la ausencia en el SMDF del dominio de tipo-Ig y de la región espaciadora “glico” característica de los otros polipéptidos de heregulina. Otra característica del SMDF es la presencia de dos regiones de aminoácidos hidrofóbicos próximas al extremo N-terminal.

Aunque los polipéptidos de heregulina se identificaron en primer lugar según su capacidad para activar el receptor HER2 (véase Holmes y col., *supra*), se descubrió que ciertas células de ovario que expresaban *neu* y fibroblastos transfectados con *neu* no se unían ni interactuaban con el NDF y tampoco respondían al NDF para desencadenar la fosforilación de tirosinas (Peles y col., EMBO J. 12:961-971 (1993)). Ello indicaba que otro componente celular era necesario para otorgar una respuesta completa a heregulina. Carraway y col., demostraron a continuación que I<sup>125</sup>-rHRG $\beta$ 1<sub>177-244</sub> se unía a fibroblastos NIH-3T3 transfectados de forma estable con el erbB3 bobino, pero no así con las células parentales no-transfectadas. Según ello, concluyeron que el ErbB3 es un receptor de HRG que facilita la fosforilación de residuos tirosina intrínsecos así como también la fosforilación del receptor ErbB2 en las células que expresan ambos receptores. Carraway y col., J Biol Chem 269(19):1403-1406 (1994). Silwoski y col., J. Biol. Chem. 269(20):14661-14665 (1994) hallaron que las células transfectadas sólo con HER3 mostraban baja afinidad por la heregulina, mientras que las células transfectadas con el HER2 y el HER3 mostraban mayor afinidad.

Esta observación se correlaciona con el “receptor *cross-talking*” descrito anteriormente por Kokai y col., Cell 58:287-292 (1989); Stern y col., EMBO J 7:995-1001 (1988); y King y col., 4:13 (1989). Estos investigadores hallaron que la unión de EGF con el EGFR resultaba en la activación del dominio quinasa de EGFR y en la fosforilación

cruzada de p185<sup>HER2</sup>. Se piensa que ello es el resultado de la heterodimerización del receptor inducida por ligando y la fosforilación cruzada concomitante de los receptores en el heterodímero (Wada y col., Cell 61:1339-1347 (1990)).

Plowman y sus colegas estudiaron de forma similar la activación de p185<sup>HER4</sup>/p185<sup>HER2</sup>. Expresaron el p185<sup>HER2</sup> solo, el p185<sup>HER4</sup> solo, o los dos receptores juntos en linfocitos T humanos y demostraron que la heregulina es capaz de estimular la fosforilación de tirosinas del p185<sup>HER4</sup>, pero que únicamente era capaz de estimular la fosforilación de p185<sup>HER2</sup> en las células que expresaban ambos receptores, Plowman y col., Nature 336:473-475 (1993).

La función biológica de la heregulina ha sido investigada por diversos grupos. Por ejemplo, Falls y col., (citados anteriormente) hallaron que ARIA desempeña una función en la diferenciación de miotubos, afectando a la síntesis y la concentración de receptores de neurotransmisores en las células del músculo postsináptico de las neuronas motoras. Corfas y Fischbach demostraron que ARIA también aumenta el número de canales de sodio en el músculo del pollo. J. Neuroscience, 13 (5):2118-2125 (1993). También se demostró que el GGFII es mitogénico para los mioblastos humanos quiescentes subconfluentes y que la diferenciación de mioblastos humanos clonales en presencia continua de GGFII resultaba en un mayor número de miotubos al cabo de 6 días de diferenciación (Sklar y col., J Cell Biochem., Abst. W462, 18D, 540 (1994). Véase también, la patente internacional WO 94/26298, publicada el 24 de noviembre de 1994.

Holmes y col., *supra*, hallaron que la HRG ejercía un efecto mitogénico sobre las líneas celulares mamarias (como SK-BR-3 y MCF-7). También se describió actividad mitogénica de los GGFS sobre las células de Schwann. Véase, p.ej., Brockes y col., J. Biol. Chem. 255(18):8374-8377 (1980); Lemke y Brockes, J. Neurosci. 4:75-83 (1984, Lemke y Brockes, J. Neurosci. 4:75-83 (1986); Brockes, J. Methods in Enzym., 147:217-225 (1987) y Marchionni y col., *supra*. Las células de Schwann constituyen células gliales importantes que proporcionan las membranas de mielina alrededor de los axones de las neuronas, formando así las fibras nerviosas. En consecuencia, es evidente que las células de Schwann desempeñan un papel importante en el desarrollo, función y regeneración de los nervios periféricos, cuyas implicaciones desde un punto de vista terapéutico han sido indicadas por Levi y col., J. Neuroscience 14(3):1309-1319 (1994). Levi y col. Discuten el potencial para la construcción de una prótesis celular que comprenda células de Schwann humanas que podría transplantarse en áreas lesionadas de la médula espinal. Los procedimientos de cultivo de las células de Schwann *ex vivo* se han descrito en, por ejemplo, la patente internacional WO 94/00140 y Li y col., J. Neuroscience 16(6):2012-2019 (1996).

Pinkas-Kramarski y col., hallaron que el NFD parece expresarse en las neuronas y células glías en el cerebro embrionario y de rata adulta y en cultivos primarios de células de cerebro de rata. Por ello, sugieren que pueda actuar como un factor de supervivencia y maduración astrocitos (Pinkas-Kramarski y col., PNAS, USA 91:9387-9391 (1994)). Meyer y Brichmeier, PNAS, USA 91: 1064-1068 (1994) analizaron la expresión de la heregulina durante la embriogénesis murina y en el animal perinatal utilizando la hibridación *in situ* y los experimentos de protección a la RNasa. Véase también Meyer y col., Development 124 (18):3575-3586 (1997). Estos autores concluyen que, basándose en la expresión de esta molécula, la heregulina desempeña un papel *in vivo* como un factor mesenquimal y neuronal. De modo similar, Danilenko y col., Abstract 3101, FASEB 8(4-5):A535 (1994); Danilenko y col., Journal of Clinical Investigation 95(2):842-851 (1995), hallaron que la interacción del NDF y el receptor HER2 era importante para dirigir la migración y la diferenciación epidérmica durante la reparación de la herida.

Ram y col., Journal of Cellular Physiology 163:589-596 (1995) evaluaron la actividad mitogénica de NDF en la línea de células epiteliales de mama humana inmortalizada MCF-10A. Danilenko y col., J. Clin Invest 95:842-851 (1995) investigaron si el NDF influenciaría la migración epidérmica en un modelo *in vivo* de reparación de la herida de grosor-profundidad parcial producida por excisión. Estos autores no hallaron diferencias estadísticas significativas en los queratinocitos proliferantes basales y superbasales en las heridas controles respecto a las heridas tratadas con rhNDF- $\alpha_2$ . Marikovsky y col., Oncogene 10:1403-1411 (1995), estudiaron las respuestas proliferativas de una línea celular continua de queratinocitos BALBA/MK aneuploide y evaluaron los efectos de las isoformas  $\alpha$  y  $\beta$  del NDF sobre los queratinocitos epidérmicos.

La relación entre la estructura y la función de proteínas nuevas puede investigarse utilizando cualquier variedad de técnicas de análisis mutacional disponible. En los ejemplos de dichas técnicas se incluye la mutagénesis por rastreo de alaninas y la expresión en fagémidos. El rastreo de alaninas puede utilizarse para identificar los residuos activos (es decir, residuos que tengan un efecto significativo sobre la función de la proteína) en una proteína o dominio proteico. Por ejemplo, Cunningham y Wells utilizaron el rastreo de alaninas para identificar los residuos en la hormona de crecimiento humana que eran importantes para la unión con su receptor. Cunningham y Wells, Science 244:1081-1085 (1989). En el rastreo de alaninas, un gen que codifica la proteína o un dominio proteico a escanear se inserta en un vector de expresión y se lleva a cabo la mutagénesis para generar una serie de vectores que codifican las proteínas o dominios en los que sus residuos se convierten secuencialmente en alanina. Las proteínas el dominio codificados o se expresan a partir de estos vectores y se analizan las actividades de las variantes de sustitución con alanina para identificar aquellos que presentan una actividad alterada. Una alteración de la actividad es indicativa de que el residuo en la posición sustituida por alanina era un residuo activo.

La expresión de fagémidos se desarrolló para permitir el rastreo de un gran número de polipéptidos variantes para una actividad de unión particular. Smith y Parmely demostraron que pueden "expresarse" de forma eficiente péptidos foráneos en la superficie de fagos filamentosos insertando fragmentos génicos cortos en el gen III del fago fd. Smith, Science 228:1315-1317 (1985); Parmley y Smith, Gene 73:305-318 (1985). Del gen de la proteína de la cubierta III

existen aproximadamente 5 copias en uno de los extremos de la partícula fágica. El fago modificado se denominó “fago de fusión” porque expresa péptidos foráneos fusionados con el gen de la proteína de la cubierta III. Puesto que cada partícula de fago fusionado expresa aproximadamente cinco copias de la proteína de fusión, este modo de expresión en fagos se denominó “expresión polivalente”.

Scott y col., y Cwirla y col., mostraron que las librerías de fagos de fusión podían rastrearse mediante selección por afinidad secuencial como el “*panning*”. Scott y col., Science 249:386-390 (1990); Cwirla y col., PNAS USA 87:6378-6382 (1990). Sin embargo, los primeros esfuerzos para seleccionar los fagos de fusión de alta afinidad no tuvieron éxito, presuntamente debido a la polivalencia de las partículas fágicas. Este problema se solventó con el desarrollo de un sistema de expresión en fagos “monovalentes”, en el que la proteína de fusión se expresa a un nivel bajo a partir de un fagémido y un fago auxiliar proporciona de la proteína la cubierta de tipo salvaje en exceso. Bass y col., Proteins 8:309-314 (1990); Lowman y col., Biochem. 30:10832-10838 (1991). La expresión en fagos monovalentes se puede utilizar para generar y rastrear un gran número de polipéptidos variantes y así aislar los que se unan con una elevada afinidad a una diana de interés.

En los Estados Unidos de América nacen cada año aproximadamente 50.000 niños con un peso al nacer inferior a 1,5 Kg. Y alrededor de dos tercios de estos niños con muy bajo peso al nacer presentan evidencias de inmadurez pulmonar, que se manifiesta por una dificultada respiratoria poco tiempo después del nacimiento. La mayoría de estos niños requieren ventilación mecánica. El síndrome de la dificultad respiratoria, causado por la producción insuficiente de tensoactivo pulmonar, así como la inmadurez estructural del pulmón, son responsables de las dificultades respiratorias observadas en estos neonatos nacidos prematuramente. Para una transferencia de oxígeno eficiente desde la interfaz aire-líquido del pulmón a la circulación sistémica, es necesario que los alvéolos estén bien desarrollados. Las proteínas tensoactivas son críticas para reducir la tensión superficial alveolar a un volumen pulmonar bajo y para prevenir el colapso alveolar.

Por todo ello, continúa existiendo la necesidad de hallar un procedimiento para el tratamiento del síndrome de la dificultad respiratoria y para otras enfermedades asociadas con el desarrollo inmaduro del pulmón y con la producción de tensoactivos pulmonares.

## Resumen de la invención

En general, un objeto de la presente invención es proporcionar medios de inducción del crecimiento y desarrollo de células epiteliales con el propósito de promover la reparación y la curación de la lesión o herida tisular.

Según ello, uno objeto de la presente invención es proporcionar un medio de tratamiento del síndrome de la dificultad respiratoria en pacientes, principalmente en pacientes humanos, que necesiten de dicho tratamiento. Otro objetivo es proporcionar un medio de inducción del crecimiento y desarrollo de las células epiteliales del pulmón. Otro objetivo es proporcionar un medio para aumentar la producción de la proteína A tensoactiva pulmonar en el pulmón de personas con una transferencia de oxígeno defectuosa en los alvéolos. La presente invención es útil en el tratamiento de niños/neonatos con dificultad respiratoria, así como también en jóvenes y adultos con una función pulmonar pobre debido a una lesión o herida pulmonar.

En un aspecto de esta invención, se ha descubierto ahora que estos objetivos y el objetivo más amplio de tratamiento de las condiciones asociadas con la lesión o herida de células epiteliales se alcanza mediante administración a un paciente que lo requiera de una cantidad efectiva de un ligando heregulina, preferentemente un polipéptido o un fragmento de ésta. Estos polipéptidos de heregulina (HRG), incluyen la HRG- $\alpha$ , la HRG- $\beta$ 1, la HRG- $\beta$ 2, la HRG- $\beta$ 3 y otros polipéptidos HRG que reaccionen de forma cruzada con los anticuerpos dirigidos contra estos miembros de la familia y/o que sean esencialmente homólogos, tal como se define más abajo, incluyendo las variantes de HRG como sus fragmentos N-terminal y C-terminal. Una HRG preferida es el ligando descrito en la Fig. 1A-1D, denominado HRG- $\alpha$ . Otras HRGs preferidas son los ligandos descritos en la Fig. 2A-2E, denominados HRG- $\beta$ 1; los descritos en la Fig. 3A-3E denominados HRG- $\beta$ 2 y los descritos en la Fig. 4A-4C, denominados HRG- $\beta$ 3. Del mismo modo, la invención proporciona medios como se define en las reivindicaciones.

En general, la invención proporciona medios de regeneración y/o reparación de la lesión de células epiteliales estimulando el crecimiento y la proliferación de dichas células epiteliales, en particular las células epiteliales ductales y ciliadas. Los tipos de lesiones y sus causas son muy variadas, por ejemplo, como se define en las reivindicaciones, la lesión causada por una incisión o resección quirúrgica, la inhalación o aspiración química de productos químicos o humo, la ulceración química o bioquímica, la lesión celular causada por la infección vírica o bacteriana, etc. Las células epiteliales que pueden estar afectadas por los medios de la invención incluyen cualquier célula epitelial que exprese HER2, HER3 y/o HER4; las células en cuestión se localizan, por ejemplo, en el pulmón, la mucosa gástrica, el endometrio, los oviductos, las glándulas mamarias, el páncreas, las glándulas salivares, etc. El medio de la invención estimula el crecimiento y la proliferación de células epiteliales, la reparación y el restablecimiento de las barreras celulares de los órganos, permitiendo que los tejidos afectados desarrollen sus funciones fisiológicas normales más rápidamente. Por ejemplo, la lesión en las células epiteliales del pulmón causada por inhalación de humo genera un enfisema. El tratamiento de las células pulmonares con el medicamento preparado de acuerdo con la invención regenera la capa-barrera de células epiteliales pulmonares, mejora la oxigenación y acelera el desarrollo de una barrera contra la infección. De manera similar, la lesión celular causada por la aspiración de ácido gástrico puede tratarse con el medicamento preparado de acuerdo con la invención para facilitar la regeneración de las células epiteliales.

**Descripción resumida de las figuras**

Figura 1A-1D muestra la secuencia aminoacídica deducida (SEC ID N°:1) de la secuencia del cDNA (SEC ID N°:2) contenida en un clon obtenido de acuerdo con la patente americana U.S. 5.367.060. La metionina iniciadora (Met) de la HRG- $\alpha$  se halla en la posición 45.

Figura 2A-2E muestra la secuencia aminoacídica deducida (SEC ID N°:3) y la secuencia del cDNA (SEC ID N°:4) de una secuencia codificante potencial de un clon obtenido de acuerdo con la patente americana U.S. 5.367.060 para HRG- $\beta$ 1. La Met de iniciación es la Met31.

Figura 3A-3E muestra la secuencia aminoacídica deducida (SEC ID N°:5) y la secuencia del cDNA (SEC ID N°:6) de una secuencia nucleotídica de un clon obtenido de acuerdo con la patente americana U.S. 5.367.060 para HRG- $\beta$ 2.

Figura 4A-4C muestra la secuencia aminoacídica deducida (SEC ID N°:7) y la secuencia del cDNA (SEC ID N°:8) de un clon obtenido de acuerdo con la patente americana U.S. 5.367.060 para HRG- $\beta$ 3.

Figura 5A-4D muestra la secuencia aminoacídica deducida (SEC ID N°:9) y la secuencia del cDNA (SEC ID N°:10) de un clon obtenido de acuerdo con la patente americana U.S. 5.367.060 para la proteína de tipo-HRG- $\beta$ 2.

Figura 6A-6C muestra una comparación de homologías aminoacídicas de varias heregulinas conocidas,  $\alpha$ ,  $\beta$ 1,  $\beta$ 2, tipo- $\beta$ 2 y  $\beta$ 3 en orden descendente y muestra las inserciones, deleciones y sustituciones aminoacídicas que caracterizan estas formas de HRG (SEC ID N°: 1, 3, 5, 9 y 7).

Figura 7A-7C muestra la secuencia aminoacídica deducida (SEC ID N°:11) y la secuencia del cDNA (SEC ID N°:12) de HRG- $\gamma$  obtenida tal como se describe en la patente americana U.S. 08/891.845. La región hidrofóbica está subrayada. El dominio de tipo-EGF está sombreado, los residuos cisteína en el dominio de tipo-EGF están dentro de un círculo. Los sitios de N-glicosilación están marcados sobre la secuencia nucleotídica con un ( ).

Figura 8 muestra la secuencia del cDNA (SEC ID N°:13) y la secuencia aminoacídica (SEC ID N°:14) de SMDF, obtenidas tal como se describe en la patente americana 08/339.517. Un dominio de tipo-EGF y los dominios apolares y no cargados (es decir, "apolar I" que consisten en los residuos aproximadamente 48-62 y el "apolar II" que consiste en los residuos aproximadamente 76-100) están subrayados. Las cisteínas en el dominio de tipo-EGF y el grupo de cisteínas del único dominio N-terminal ("NTD-grupo cys" están encuadradas. El codón de parada está indicado por la letra "O".

**Descripción detallada de las realizaciones preferidas**

Los ligandos HRG, en particular los polipéptidos y anticuerpos agonistas de lo mismo, tienen afinidad por y estimulan la autofosforilación de los receptores HER2, HER3 y/o HER4 o combinaciones de éstos. Se incluyen dentro de la definición de ligandos HRG, además de HRG- $\alpha$ , HRG- $\beta$ 1, HRG- $\beta$ 2, HRG- $\beta$ 3 y de tipo-HRG- $\beta$ 2, otros polipéptidos de unión al receptor HER2, HER3 y/o HER4 con una identidad de secuencia aminoacídica importante con HRG- $\alpha$  o HRG- $\beta$ 1. Dichos polipéptidos adicionales se hallan dentro de la definición de HRG como una familia de ligandos polipeptídicos que se unen a los receptores HER2, HER3 y/o HER4.

Los polipéptidos de heregulina se unen con diversa afinidad a los receptores HER2, HER3 y/o HER4. En general, con los receptores HER3 y HER4 se unen con elevada afinidad. También se sabe que la heterodimerización de HER2 con HER3 y HER2 con HER4 tiene lugar con la consiguiente fosforilación cruzada del receptor, tal como se describió más arriba. En la presente invención, el crecimiento y/o la proliferación de las células epiteliales se induce cuando una proteína heregulina interactúa y se une con una molécula de receptor individual o un dímero, induciéndose de este modo la fosforilación del receptor. La unión y la activación de HER2, HER3, HER4 o combinaciones de lo mismo, incluye la activación de cualquier forma del receptor necesaria para la activación y función biológica de éste, incluyendo las formas monoméricas y diméricas del receptor. Las formas diméricas del receptor pueden ser referidas más adelante como HER2/HER3, HER2/HER4 y HER3/HER4.

**I. Definiciones**

En general, tanto en la descripción como en los ejemplos y reivindicaciones se mantendrá la definición indicada para los términos y frases siguientes.

Ligando "heregulina" (HRG) se define aquí como cualquier ligando aislado, preferentemente una secuencia polipeptídica que posea una propiedad biológica de un polipéptido HRG natural. Los ligandos abarcados dentro del ámbito de la presente invención incluyen las proteínas heregulina, indicadas anteriormente, NDF, ARIA y el factor de crecimiento GGF, así como los polipéptidos de SMDF y HRG discutidos con detalle en la presente invención. Estos polipéptidos de heregulina, NDF, ARIA y GGF son bien conocidos en la materia. La HRG incluye los polipéptidos que se muestran en las Figs. 1A-1D, 2A-2E, 3A-3E, 4A-4C, 5A-5D, 6A-6C, 7A-7C y 8 y sus análogos de mamíferos. Se incluyen las variantes HRG como la  $\gamma$ -HRG descrita en la patente internacional WO 98/02541, publicada el 22 de enero de 1998; las variantes descritas en la patente internacional WO 98/35036, publicada el 13 de agosto de 1998; y las variantes SMDF descritas en la patente americana US 5.770.567 publicada el 23 de junio de 1998 (patente interna-

cional WO 96/15244). Estas variantes pueden prepararse por lo procedimientos descritos más adelante, opcionalmente junto con las técnicas conocidas en la materia del rastro por alaninas y la expresión en fagos. Cunningham y Wells, Science 244:1081-85 (1989); Bass y col., Proteins 8:309-14 (1990); Lowman y col., Biochem. 30:10832-38 (1991).

5 El término “una célula epitelial normal” significa una célula epitelial que no ha sido transformada, es decir, no es cancerosa ni está inmortalizada. Además, las células epiteliales normales son preferentemente no aneuploides. Existe aneuploidía cuando el núcleo de una célula no contiene un múltiplo exacto del número haploide de cromosomas, es decir uno o más cromosomas están presentes en mayor o menor número que el resto. Las propiedades típicas de las células transformadas que se hallan fuera del ámbito de la presente invención incluyen la capacidad para formar  
10 tumores cuando se implantan en ratones inmuno-deprimidos (ratones atímicos desnudos), la capacidad para crecer en suspensión o en medio semi-sólido como el agar, la pérdida de inhibición por contacto que permite que las células se acumulen formando colonias o focos, la pérdida de la dependencia de los factores de crecimiento o del suero, la muerte celular si se inhibe el crecimiento celular y desorganización de los filamentos de actina. Se incluyen de forma específica dentro de la presente invención las células epiteliales normales que no formarán tumores en los  
15 ratones, las células que crecen adheridas al plástico o vidrio (son dependientes del anclaje), las que presentan inhibición por contacto, necesitan suero que contenga hormonas y factores de crecimiento, las que permanecen viables si el crecimiento se para por falta de suero y las que contienen filamentos de actina bien organizados. Aunque las células epiteliales normales no son preferentemente células cultivadas, también son adecuadas para la presente invención las células epiteliales no-inmortalizadas y no-transformadas aisladas de tejidos de mamífero. Estas células aisladas  
20 pueden cultivarse durante diversas generaciones (hasta aproximadamente 10 o 50 generaciones) en presencia de una heregulina con el fin de inducir el crecimiento y/o proliferación de una muestra de células aisladas, es decir para crecer la muestra. La muestra crecida se reintroduce a continuación en el mamífero con el objetivo de repoblar el tejido de células epiteliales (repitelización). Ello es particularmente útil para reparar el tejido lesionado o dañado.

25 Una célula “epitelial” es una célula localizada en una capa celular avascular que recubre la superficie libre (cutánea, mucosa o serosa) de un órgano o forra un tubo o cavidad del cuerpo de un animal. Las células epiteliales del pulmón incluyen las células epiteliales bronquiales, las células de tipo II y las células Clara. El término “célula epitelial”, tal como se utiliza aquí, es consistente con la definición reconocida en la materia de células epiteliales del epitelio. Véase, por ejemplo, la definición en el Diccionario Enciclopédico Médico Taber, Edición 12, (1973) F.A. publicación  
30 de Davis Company.

El término “característica biológica” para los propósitos de la presente invención hace referencia a una función biológica o antigénica *in vivo* o actividad que se realice directa o indirectamente por una secuencia HRG (bien en su conformación nativa o desnaturalizada), o mediante una subsecuencia de la misma. Las funciones biológicas incluyen la unión al receptor, cualquier actividad enzimática o actividad moduladora de la actividad enzimática, cualquier  
35 actividad de unión al transportador, cualquier actividad hormonal, cualquier actividad promotora o inhibidora de la adhesión celular a una matriz extracelular o moléculas de superficie celular, o cualquier función estructural. Sin embargo, las funciones biológicas no incluyen las funciones antigénicas, es decir, poseer un epitopo o sitio antigénico capaz de reaccionar de forma cruzada con los anticuerpos generados contra un polipéptido HRG producido de forma  
40 natural.

El término HRG “activa biológicamente” se define aquí como un polipéptido que comparte una función biológica de una secuencia HRG que puede (aunque no lo necesite) poseer además una función antigénica. Un efecto o función principal de la HRG es actuar como un polipéptido ligando con actividad biológica cualitativa de unión a HER2, HER3  
45 y/o HER4, dando como resultado la activación del receptor de tirosina quinasa (un “ligando activador”). Un test para los ligandos de activación es el ensayo de autofosforilación de la tirosina de HRG, que se describe más adelante. Se incluyen dentro del ámbito de la presente invención, las HRG con secuencias aminoácídicas maduras traducidas de la HRG humana completa; los derivados desglicosilados o no glicosilados de la HRG, las variantes de secuencia aminoácídica de la secuencia de HRG y los derivados de HRG, capaces de presentar una característica biológica en común con la HRG. Aunque la HRG nativa es un polipéptido unido a membrana, también se incluyen dentro de  
50 esta definición las formas solubles, como las formas carentes de un dominio transmembrana funcional. En particular, se incluyen los fragmentos polipeptídicos de la secuencia de HRG que tienen un extremo N-terminal en cualquiera de los residuos S216 a A227 y su extremo C-terminal en cualquier residuo desde el K268 al R286 y las secuencias homólogas que se muestran en la Fig. 6A-C, de aquí en adelante referido en todas las HRGs como dominio del factor de crecimiento (GFD).  
55

Una HRG “activa antigénicamente” se define como un polipéptido que posee una función antigénica de una HRG y que puede (aunque no lo necesite) tener además una función biológica.

60 En realizaciones preferidas, la HRG activa antigénicamente es un polipéptido que se une con una afinidad de al menos  $10^{-9}$  l/mol a un anticuerpo producido contra una secuencia HRG natural. Generalmente el polipéptido se une con una afinidad de al menos  $10^{-8}$  l/mol. Más preferentemente, la HRG activa antigénicamente es un polipéptido que se une a un anticuerpo producido contra una de las HRGs en su conformación nativa. La HRG en su conformación nativa corresponde generalmente a la HRG hallada en la naturaleza, que no ha sido desnaturalizada por agentes caotrópicos, calor ni otros tratamientos que modifican esencialmente la estructura tridimensional de la HRG, según se determina,  
65 por ejemplo, por migración en geles de separación de tamaño no reductores ni desnaturalizantes. El anticuerpo utilizado en esta determinación puede ser un anticuerpo policlonal de conejo generado mediante formulación de la HRG nativa de una especie distinta al conejo en adyuvante completo de Freund, inyectando subcutáneamente la formulación

y estimulando la respuesta inmune mediante inyección intraperitoneal de la formulación hasta alcanzar la meseta en la curva de titulación de anticuerpos anti-HRG.

Generalmente, la HRG activa biológica o antigénicamente tendrá una secuencia aminoacídica con al menos una identidad de secuencia del 75% con una secuencia HRG dada, más preferentemente al menos un 80%, más preferentemente al menos un 90% y más preferentemente al menos un 95%. La identidad u homología respecto a una secuencia de HRG se define aquí como el porcentaje de residuos aminoacídicos en la secuencia candidata que son idénticos con los residuos HRG de la Fig. 6A-6C, tras alinear las secuencias e introducir los huecos necesarios, para lograr el porcentaje máximo de identidad y sin considerar ninguna sustitución conservadora como parte de la identidad de secuencia. En cuanto a la identidad de secuencia, no se deberá considerar que ésta se vea afectada por las extensiones en los extremos N- o C-terminal o las extensiones internas, ni tampoco por las delecciones o inserciones en la secuencia de HRG.

En consecuencia, los polipéptidos HRG activos antigénica y biológicamente que son el objetivo de esta invención incluyen cada una de las SECUENCIAS completas de HRG; los fragmentos con una secuencia consecutiva de al menos 5, 10, 15, 20, 25, 30 o 40 residuos de aminoácidos procedentes de la secuencia HRG; las variantes de secuencia aminoacídica de la secuencia HRG en donde un residuo aminoacídico se ha insertado en el extremo N- o C-terminal de, o dentro, de la secuencia HRG o su fragmento, tal como se definió anteriormente; variantes de secuencia aminoacídica de la secuencia HRG o su fragmento tal como se definió anteriormente en donde se ha realizado una sustitución aminoacídica. Los polipéptidos HRG incluyen aquellos polipéptidos que contienen mutaciones predeterminadas, por ejemplo, mediante mutagénesis dirigida o por PCR y otras especies animales de polipéptidos HRG como el HRG de conejo, rata, cerdo, primate no-humano, equino, murino y ovino y los alelos u otras variantes naturales de las secuencias anteriores y las secuencias humanas; derivados de HRG o sus fragmentos tal como se definió anteriormente, en donde la HRG o sus fragmentos se han modificado covalentemente mediante sustitución, enzimáticamente, químicamente o por otros medios adecuados con una porción distinta al aminoácido natural (por ejemplo una porción como una enzima o un radioisótopo); variantes de glicosilación de HRG (inserción de un sitio de glicosilación o delección de cualquier sitio de glicosilación mediante delección, inserción, o sustitución del aminoácido adecuado); y las formas solubles de HRG, como HRG-GFD o las que carecen de un dominio transmembrana funcional.

El término “aislado” significa un ligando, como la HRG, que ha sido identificado y separado y/o recuperado de un componente de su entorno natural. Los componentes contaminantes de su entorno natural son materiales que podrían interferir con las utilizaciones diagnósticas o terapéuticas de la HRG e incluyen proteínas, hormonas y otras sustancias. En realizaciones preferidas, la HRG se purificará (1) a un grado superior al 95% en peso de proteína, según se determina por el procedimiento de Lowry u otros procedimientos de determinación de proteínas validados y más preferentemente superior al 99% en peso, (2) a un grado suficiente para obtener al menos 15 residuos de la secuencia aminoacídica N-terminal o interna utilizando la mejor secuencia aminoacídica comercial disponible en la actualidad, o (3) a homogeneidad mediante SDS-PAGE y utilizando azul de Coomassie o, preferentemente tinción de plata. La HRG aislada incluye la HRG *in situ* dentro de las células recombinantes, ya que al menos un componente del entorno natural de HRG no estará presente. La HRG aislada incluye también la HRG de una especie en un cultivo de células recombinantes de otra especie, con lo que la HRG en dichas circunstancias será una fuente de polipéptidos. Sin embargo, por lo general la HRG aislada se preparará al menos mediante una etapa de purificación.

De acuerdo con esta invención, el ácido nucleico de HRG es un RNA o DNA que contiene más de 10 bases que codifican una HRG activa biológica o antigénicamente y es complementario a la secuencia aminoacídica codificante de dicha HRG, o hibrida con la secuencia de ácido nucleico que codifica dicha HRG, permaneciendo unido de forma estable a ésta en condiciones astringentes.

Preferentemente, el ácido nucleico de HRG que codifica un polipéptido que comparte al menos el 75% de identidad de secuencia, más preferentemente al menos el 80%, todavía más preferentemente al menos el 85%, todavía mejor un 90%, y mucho mejor el 95%, con una secuencia HRG. Preferentemente, el ácido nucleico de HRG que hibrida contiene al menos 20, más preferentemente 40 y todavía más preferentemente 90 bases.

El ácido nucleico de HRG incluye un ácido nucleico que se puede identificar y separar del ácido nucleico con el que generalmente el ácido nucleico de HRG se halla asociado en la fuente original. El ácido nucleico de HRG así aislado se halla en una forma distinta a la natural. Este ácido nucleico aislado codificante de HRG incluye el ácido nucleico de HRG de células que normalmente expresan HRG, en donde el ácido nucleico se halla en una localización cromosómica distinta a la de las células naturales o, dicho de otro modo, está flanqueado por una secuencia de DNA distinta a la hallada en la naturaleza. El ácido nucleico que codifica HRG puede utilizarse en ensayos específicos de hibridación, en particular aquellas porciones de secuencia codificante de HRG que no hibridan con otras secuencias de DNA conocidas. “Condiciones astringentes” son aquellas condiciones que (1) emplean baja fuerza iónica y elevada temperatura para el lavado, por ejemplo, NaCl 0,015 M/citrato sódico 0,0015 M/NaDodSO<sub>4</sub> al 0,1% a 50°C; (2) utilizan durante la hibridación un agente desnaturalizante como la formamida, por ejemplo, formamida al 50% (vol/vol) con albúmina sérica bovina al 0,1%, Ficoll al 0,1%, polivinilpirrolidona al 0,1%, tampón fosfato sódico 50 mM a pH 6,5 con NaCl 750 mM, citrato sódico 75 mM a 42°C; o (3) utilizan formamida al 50%, 5XSSC (NaCl 0,75 M, citrato sódico 0,075 M), fosfato sódico 50 mM (pH 6,8), pirofosfato sódico al 0,1%, solución de 5x Denhardt, DNA de esperma de salmón sonificado (50g/ml), SDS al 0,1% y sulfato de dextrano al 10% a 42°C, con lavados a 42°C en 0,2X SSC y SDS al 0,1%.

El término “secuencias control” hace referencia a las secuencias de DNA necesarias para la expresión de una secuencia codificante unida operativamente a un organismo huésped determinado. Las secuencias control que son adecuadas para procariotas, por ejemplo, incluyen un promotor, opcionalmente una secuencia operadora, un sitio de unión a ribosomas y posiblemente, otras secuencias todavía poco conocidas. Las células eucariotas son conocidas por utilizar promotores, señales de poliadenilación e intensificadores de la transcripción.

El ácido nucleico está “unido operativamente” cuando está colocado en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el DNA para una presecuencia o líder secretor está unido operativamente al DNA codificante de un polipéptido si éste se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o secuencia intensificadora está unida operativamente a una secuencia codificante si afecta la transcripción de la secuencia; o un sitio de unión a ribosomas está unido operativamente a una secuencia codificante si está colocado en una posición que facilite la traducción. En general, “unido operativamente” significa que las secuencias de DNA que están unidas son contiguas y, en el caso de un líder secretor, contiguas y en la misma pauta de lectura. Sin embargo, los intensificadores no tienen porque ser contiguos. La unión se logra mediante ligación con dianas de restricción adecuadas. Si dichas dianas no existieran, entonces se utilizan adaptadores o engarces oligonucleotídicos sintéticos de acuerdo con la práctica convencional.

Un elemento “exógeno” se define aquí como una secuencia de ácido nucleico que es foránea a la célula, u homóloga a la célula pero en una posición en la célula huésped que no es la habitual.

Tal como se utilizan aquí, las expresiones “célula”, “línea celular” y “cultivo celular” se utilizan de modo intercambiable, incluyéndose la progenie en todas las denominaciones. En consecuencia, las palabras “transformantes” y “células transformadas” incluyen la célula primaria en cuestión y los cultivos derivados de la misma, independientemente del número de transferencias. Se entenderá que toda la progenie no tiene porqué ser idéntica en cuanto a contenido de DNA debido a mutaciones deliberadas o inadvertidas que puedan haber sucedido. Se incluye también, la progenie mutante que tiene la misma función o actividad biológica que la analizada en la célula original transformada. Cuando se utilicen denominaciones distintas, la terminología será evidente según el contexto.

Los “plásmidos” se denominan por una “p” minúscula y/o seguida de letras mayúsculas y/o números. Los plásmidos de inicio en la presente invención están disponibles comercialmente, son de disponibilidad pública sin restricciones, o pueden construirse a partir de plásmidos disponibles de acuerdo con procedimientos publicados. Además, otros plásmidos equivalentes son conocidos en la materia y serán evidentes a un experto en la materia.

La “digestión por enzimas de restricción” del DNA hace referencia al corte catalítico del DNA con una enzima que actúa únicamente en ciertos sitios en el DNA. dichas enzimas son denominadas endonucleasas de restricción, y los sitios específicos para cada enzima, dianas de restricción. Las diversas enzimas de restricción utilizadas en la presente invención están disponibles comercialmente y se utilizan las condiciones de reacción, cofactores y otros requisitos establecidos por los fabricantes de las enzimas. Las enzimas de restricción se denominan generalmente por abreviaciones compuestas de una letra mayúscula, seguida por otras letras que representan el microorganismo de donde se obtuvo originalmente la enzima de restricción determinada y a continuación el número de la enzima en cuestión. En general, se utiliza aproximadamente 1 mg de plásmido o fragmento de DNA con aproximadamente 1-2 unidades de enzima en 20 ml de solución tamponada. Los tampones adecuados y las cantidades de sustrato precisas para cada enzima en particular se especifican por el fabricante. La incubación suele ser de 1 hora a 37°C, pero puede variar de acuerdo con las instrucciones del fabricante. Después de la incubación, se elimina la proteína o el polipéptido mediante extracción con fenol y cloroformo y se recupera el ácido nucleico digerido de la fracción acuosa mediante precipitación con etanol. La digestión con una enzima de restricción puede proseguirse con la hidrólisis de fosfatasa alcalina bacteriana de los fosfatos 5' terminales para evitar que los dos extremos cortados de restricción de un fragmento de DNA se recircularicen o formen un bucle cerrado que impediría la inserción de otros fragmento de DNA en la diana de restricción. Salvo que se indique lo contrario, la digestión de los plásmidos no se prosigue por la desfosforilación 5' terminal. Los procedimientos y reactivos para la desfosforilación son los convencionales descritos en las secciones 1.56-1.61 de Sambrook y col., (Molecular Cloning: A Laboratory Manual New York: Cold Spring Harbor Laboratory Press, 1989).

La “recuperación” o el “aislamiento” de un fragmento de DNA dado de una digestión de restricción significa la separación del fragmento digerido en un gel de poliacrilamida o agarosa mediante electroforesis, la identificación del fragmento de interés por comparación de su movilidad respecto a los fragmentos de DNA de peso molecular conocido, la eliminación de la sección del gel que contiene el fragmento deseado y la separación del DNA del gel. Este procedimiento de aislamiento es bien conocida en la materia, véase, por ejemplo, Lawn y col., Nucleic Acids Res. 9:6103-6114 (1981) y Goeddel y col., Nucleic Acids Res. 8:4057, 1980).

El análisis de “transferencia de Northern” es un procedimiento utilizado para identificar secuencias de RNA que hibridan con una sonda conocida, un oligonucleótido, un fragmento de DNA, un cDNA o un fragmento de lo mismo, o un fragmento de RNA. La sonda se marca con un radioisótopo como el P<sup>32</sup>, o mediante biotinilación, o con una enzima. El RNA a analizar se separa electroforéticamente en un gel de agarosa o poliacrilamida, se transfiere a una membrana de nitrocelulosa, nylon u otro tipo y se hibrida con la sonda, utilizando técnicas estándares bien conocidas en la materia como las descritas en las secciones 7.39-7.52 de Sambrook y col., *supra*.

La “ligación” hace referencia al proceso de formación de enlaces fosfodiéster entre los dos fragmentos de ácido nucleico. Para ligar los fragmentos de DNA juntos, sus extremos deben ser compatibles. En algunos casos, los extremos serán directamente compatibles después de la digestión con endonucleasas. Sin embargo, a veces puede ser necesario convertir en primer lugar los extremos cohesivos en romos mediante digestión con endonucleasas de restricción para hacerlos compatibles para la ligación. Para convertir los extremos en romos, el DNA se trata en un tampón adecuado durante 15 min a 15°C con aproximadamente 10 unidades del fragmento Klenow de DNA polimerasa 1 o la T4 DNA polimerasa en presencia de los cuatro desoxiribonucleótidos trifosfato. El DNA se purifica por extracción con fenol-cloroformo y se precipita con etanol. Los fragmentos de DNA que se han de ligar se ponen en una solución a aproximadamente cantidades equimolares. La solución también contendrá ATP, tampón de ligasa y una ligasa como la T4 DNA ligasa a aproximadamente 10 unidades por cada 0,5 mg de DNA. Si el DNA se ha de ligar en un vector, éste se lineariza mediante digestión con la(s) endonucleasa(s) de restricción adecuada(s). El fragmento linearizado se trata con fosfatasa alcalina bacteriana, o con fosfatasa intestinal bovina para prevenir la auto-ligación durante la etapa de ligación.

La “preparación” de DNA a partir de células quiere decir el aislamiento del DNA plasmídico de un cultivo de células huésped. Los procedimientos comúnmente utilizados para la preparación de DNA son las preparaciones plasmídicas a pequeña y gran escala descritas en las secciones 1.25-1.33 de Sambrook y col., *supra*.

Los “oligonucleótidos” son polidesoxinucleótidos de corta longitud, de cadena sencilla o doble que se sintetizan químicamente mediante procedimientos conocidos (como el fosfotriéster, el fosfito o la química de la fosforamidita, utilizando técnicas de fase sólida como las descritas en la patente europea EP 266.032, publicada el 4 de mayo de 1988, o mediante intermediarios desoxinucleósidos H-fosfonato, tal como se describe por Froehler y col., Nucl. Acids Res. 14: 5399-5407, 1986. Y a continuación se purifican en geles de poliacrilamida.

La técnica de la “reacción en cadena de polimerasa” o “PCR”, tal como se utiliza aquí, hace referencia a un procedimiento en el que en cuestión de minutos un fragmento de ácido nucleico, RNA y/o DNA se amplifica tal como se describe en la patente americana U.S. 4.683.195, publicada el 28 de julio de 1987. En general, para que puedan diseñarse los cebadores oligonucleotídicos, es necesario que la información de secuencia de los extremos de la región de interés esté disponible; estos cebadores serán idénticos o similares a la secuencia de las cadenas opuestas del molde a amplificar. Los nucleótidos del extremo 5' terminal de los dos cebadores pueden coincidir con los extremos del material amplificado. La PCR puede utilizarse para amplificar las secuencias de RNA específicas, las secuencias de DNA específica del DNA genómico total y del cDNA transcrito a partir del RNA celular total, de secuencias de bacteriófago o plasmídicas, etc. Véase por ejemplo, Mullis y col., Cold Spring Harbor Symp. Quant. Biol. 51:263 (1987); Elrich, ed., PCR Technology, (Stockton Press, NY, 1989). Tal como se utiliza aquí, la PCR se considera uno de los procedimientos, pero no el único, de la reacción de una polimerasa de ácido nucleico para amplificar una muestra de ácido nucleico en un ensayo, que comprende la utilización de un ácido nucleico conocido (DNA o RNA) como cebador y utiliza una polimerasa de ácido nucleico para amplificar o generar un fragmento específico de ácido nucleico, o para amplificar o generar un fragmento de ácido nucleico complementario de un ácido nucleico determinado.

El ensayo de autofosforilación de tirosinas de HRG para detectar la presencia o la bioactividad de ligandos HRG puede utilizarse para controlar la purificación de un ligando para los receptores de HER2 y HER3. Este ensayo se basa en asumir que un ligando específico para el receptor estimulará la autofosforilación del receptor, en analogía con el receptor de EGF y la estimulación de su propia autofosforilación. Véase Sadich y col., Anal. Biochem. 235:207-214 (1996). Las células MDA-MDB 453 o las células MCF7 que contienen niveles elevados de p185<sup>HER2</sup> pero negligibles de receptores de EGF humanos, se obtuvieron del American Type Culture Collection, Rockville, Md. (ATCC N° HTB-131) y se mantuvieron en el cultivo de tejidos con suero fetal bovino al 10% en DMEM/Ham F12 (1:1). Para el ensayo, las células se tripsinizaron y plaquearon a 150.000 células/pocillo en placas de 24 pocillos (Costar). Después de la incubación con medio que contenía suero durante toda la noche, las células se colocaron en medio sin suero durante 2-18 horas antes del ensayo. Se añadieron 100 µl de las muestras a cada pocillo. Las células se incubaron durante 5-30 minutos (típicamente 30 min) a 37°C y se eliminó el medio. A continuación, se trataron las células de cada pocillo con 100 µl de un tampón desnaturalizante en gel con SDS (SEPROSOL, Enpotech, Inc.) y las placas se calentaron a 100°C durante 5 minutos para disolver las células y desnaturalizar las proteínas, se migraron alícuotas de cada pocillo en una electroforesis en geles de gradiente al 5-50% con SDS (NOVEX, Encinitas, CA) de acuerdo con las direcciones del fabricante. Después de que el colorante alcanzara la parte inferior del gel, se finalizó la electroforesis y se colocó una membrana de PVDF (PRPBLot, ABI) sobre el gel, transfiriéndose las proteínas del gel a la membrana en una cámara de transferencia (BioRad) a 200 mAmps durante 30-60 min. Después de la transferencia, las membranas se incubaron con solución salina tamponada con TRIS y un tampón detergente TWEEN-20 al 0,1% y BSA al 5% durante 2-18 h para bloquear la unión inespecífica, y luego se trataron con un anticuerpo anti-fosfotirosina (Upstate Biological, N.Y.). A continuación, las membranas de transferencia se trataron con el anticuerpo anti-ratón de cabra conjugado con fosfatasa alcalina. Los geles se revelaron utilizando el sistema PROTOBLot de Promega. Después de secar las membranas, se cuantificó la densidad de las bandas correspondientes con el p185<sup>HER2</sup> en cada muestra con un SCANJET Plus Scanner de Hewlett Packard acoplado a un ordenador Macintosh. El número de receptores por célula en las células MDA-MB-453 fue tal que en estas condiciones experimentales, la proteína del receptor p185<sup>HER2</sup> es la principal proteína marcada.

La “microsecuenciación de proteínas” se realizó basándose en los procedimientos siguientes. Las proteínas de la etapa final de HPLC se secuenciaron directamente por degradación de Edman automatizada con un secuenciador de fase de gas de Applied Biosystems, modelo 470A, equipado con un analizador de aminoácidos 120A PTH, o se secuenciaron después de digerirse con diversas enzimas o agentes químicos. Los aminoácidos PTH se integraron uti-

lizando el sistema de datos CHROMPERFECT (Justice Innovations. Palo Alto). La interpretación de la secuencia se realizó en un ordenador VAX 11/785 Digital Equipment Corporation tal como se describió en Henzel y col., J. Chromatography 404:41-52 (1987). En algunos casos, se migraron alícuotas de las fracciones de HPLC en electroforesis en geles de poliacrilamida-SDS al 5-20%, se electrotransfirieron a una membrana de PVDF (PROBLOTT, ABI, Foster City, CA) y se tiñeron con Azul de Coomassie Brillante (Matsudaira, P. J. Biol. Chem. 262:10035-10038, 1987). La proteína específica se recortó de la membrana de transferencia para la secuenciación N-terminal. Para determinar las secuencias de proteínas internas, las fracciones de HPLV se secaron al vacío (SPEEDVAC), se resuspendieron en un tampón adecuado y digirieron con bromuro de cianógeno, la enzima específica de lisina Lys-C (Wako Chemicals, Richmond, VA) o Asp-N (Boehringer Mannheim, Indianapolis, Ind). Después de la digestión, los péptidos resultantes se secuenciaron como una mezcla, o se resolvieron por HPLC en una columna C4 con un gradiente de propanol en TFA al 0,1% antes de la secuenciación, tal como se describió anteriormente.

Los “anticuerpos” (Abs) y las “inmunoglobulinas” (Igs) son glicoproteínas que tienen las mismas características estructurales. Mientras que los anticuerpos presentan especificidad de unión contra un antígeno específico, las inmunoglobulinas incluyen tanto los anticuerpos como otras moléculas de tipo anticuerpo que carecen de especificidad antigénica. Los polipéptidos de la última clase se producen, por ejemplo, a niveles bajos por el sistema linfático y a niveles mayores por mielomas.

La digestión con papaína de los anticuerpos produce dos fragmentos de unión al antígeno idénticos, denominados fragmentos “Fab”, cada uno con un sitio de unión al antígeno y un fragmento “Fc” residual, cuyo nombre refleja su capacidad para cristalizar fácilmente. El tratamiento con pepsina produce un fragmento F(ab')<sub>2</sub> que tiene dos sitios de combinación antigénica y es todavía capaz de unirse con el antígeno.

“Fv” es el fragmento de anticuerpo mínimo que contiene un sitio de reconocimiento y de unión al antígeno completo. Esta región consiste en un dímero de una cadena pesada y un dominio variable de cadena ligera en asociación estrecha, no-covalente. En esta configuración interactúan las tres regiones hipervariables de cada dominio variable para definir un sitio de unión al antígeno en la superficie del dímero V<sub>H</sub>-V<sub>L</sub>. En conjunto, las seis regiones hipervariables confieren al anticuerpo especificidad de unión por el antígeno. Sin embargo, incluso un solo dominio variable (o la mitad de un Fv que comprenda las tres regiones hipervariables específicas para un antígeno) tiene la capacidad para reconocer y unirse al antígeno, aunque con una afinidad inferior que el sitio de unión completo.

El fragmento Fab también contiene el dominio constante de cadena ligera y el primer dominio constante (CH1) de cadena pesada. Los fragmentos Fab' difieren de los fragmentos Fab por la adición de unos pocos residuos en el extremo carboxi-terminal del dominio CH1 de cadena pesada incluyendo una o más cisteínas de la región bisagra del anticuerpo. Fab'-SH es la denominación para Fab', en la cual los residuos cisteína de los dominios constantes tienen un grupo tiol libre. Los fragmentos de anticuerpo F(ab')<sub>2</sub> se producen originalmente como parejas de fragmentos Fab' con las cisteínas bisagra entre ellos. También se conocen otros acoplantes químicos de los fragmentos de anticuerpo.

Las “cadenas ligeras” de anticuerpos (inmunoglobulinas) de cualquier especie de vertebrados pueden clasificarse en uno o dos tipos claramente distintos, denominados kappa (K) y lambda (γ), según las secuencias aminoácidas de sus dominios constantes.

Dependiendo de la secuencia aminoácida del dominio constante de sus cadenas pesadas, las inmunoglobulinas pueden clasificarse en clases distintas. Existen cinco clases principales de inmunoglobulinas: IgA, IgD, IgE, IgG e IgM y algunas de éstas pueden además dividirse en subclases (isotipos), p.ej., IgG1, IgG3, IgG3, IgG4, IgA e IgA2. Los dominios de cadena constante que corresponden con las clases distintas de inmunoglobulinas se denominan α, δ, ε, γ y μ, respectivamente. Las estructuras de la subunidad y las configuraciones en tres dimensiones de clases distintas de inmunoglobulinas se conocen con profundidad.

El término “anticuerpo” se utiliza aquí en el sentido amplio y abarca específicamente los anticuerpos monoclonales (incluyendo los anticuerpos monoclonales de longitud completa), los anticuerpos policlonales, los anticuerpos multiespecíficos (p.ej., los anticuerpos biespecíficos) y los fragmentos de anticuerpos siempre que presenten la actividad biológica deseada.

Los “fragmentos de anticuerpo” comprenden una porción de un anticuerpo de longitud completa, generalmente el dominio de unión al antígeno o dominio variable de lo mismo. Ejemplos de fragmentos de anticuerpos incluyen los fragmentos Fab, Fab', F(ab')<sub>2</sub> y Fv; los anticuerpos bivalentes o “diacuerpos”; los anticuerpos lineales; las moléculas de anticuerpo de cadena sencilla; y los anticuerpos multiespecíficos formados de fragmentos de anticuerpos.

El término “anticuerpo monoclonal”, tal como se utiliza aquí, hace referencia a un anticuerpo obtenido de una población homogénea de anticuerpos, es decir, los anticuerpos individuales que comprenden la población son idénticos excepto por mutaciones posibles que ocurran de forma natural y puedan estar en cantidades minoritarias. Los anticuerpos monoclonales son altamente específicos y están dirigidos contra un sitio antigénico único. Además, al contrario que las preparaciones de anticuerpos convencionales (policlonales) que incluyen anticuerpos distintos dirigidos contra determinantes distintos (epitopos), cada anticuerpo monoclonal está dirigido contra un único determinante en el antígeno. El modificador “monoclonal” indica el carácter del anticuerpo como el obtenido de una población homogénea de anticuerpos, y no debe considerarse que la producción de anticuerpos necesite ningún procedimiento particular. Por ejemplo, los anticuerpos monoclonales para utilizar de acuerdo con la presente invención pueden generarse por el

procedimiento del hibridoma descrito por primera vez por Kohler y col., Nature 256:495 (1975), o por procedimientos del DNA recombinante (véase, p.ej., la patente americana U.S. 4.816.567). Los “anticuerpos monoclonales” también pueden aislarse de librerías de fagos utilizando la técnica descrita por Clackson y col., Nature 352:624-628 (1991) y Marks y col., J. Mol. Biol. 222:581-597 (1991), por ejemplo.

Los anticuerpos monoclonales de la presente invención incluyen específicamente anticuerpos (inmunoglobulinas) “quiméricos” en los que una porción de la cadena ligera y/o pesada es idéntica u homóloga a las secuencias correspondientes en los anticuerpos derivados de una especie particular o perteneciente a una clase o subclase de anticuerpo determinada, mientras que la(s) cadena(s) restante(s) es idéntica u homóloga a las secuencias correspondientes en los anticuerpos derivados de otras especies o pertenecientes a otra clase o subclase de anticuerpos, así como también los fragmentos de dichos anticuerpos, siempre que presenten la actividad biológica deseada (patente americana U.S. 4.816.567; y Morrison y col., Proc. Natl. Acad. Sci. USA 81:6851-6855 (1984)).

Las formas “humanizadas” de anticuerpos no-humanos (p.ej., murinos) son anticuerpos quiméricos que contienen una secuencia derivada de una inmunoglobulina no-humana. En la mayoría de los casos, los anticuerpos humanizados son inmunoglobulinas (anticuerpo receptor) en los que se han reemplazado residuos de la región hipervariable del receptor por residuos de la región hipervariable de una especie no-humana (anticuerpo donante) como por ejemplo el ratón, la rata, el conejo o primates no-humanos con la especificidad y capacidad deseadas. En algunos casos, los residuos de la región de entramado (FR) de la inmunoglobulina humana se sustituyen con los residuos no-humanos correspondientes. Además, los anticuerpos humanizados pueden comprender residuos que no se hallen en el anticuerpo receptor o en el anticuerpo donante. Estas modificaciones se realizan para perfeccionar la generación del anticuerpo. En general, el anticuerpo humanizado comprenderá todos o al menos uno, y generalmente dos, dominios variables, en los que todas o esencialmente todas las regiones hipervariables corresponden con las de una inmunoglobulina no-humana y todas o esencialmente todas las FRs son las de una secuencia de inmunoglobulina humana. El anticuerpo humanizado comprenderá opcionalmente una porción por lo menos de una región constante de inmunoglobulina (Fc), generalmente la de una inmunoglobulina humana. Para detalles adicionales, véase Jones y col., Nature 321:522-525 (1986); Reichmann y col., Nature 332:323-329 (1988); y Presta, Curr. Op. Struct. Biol. 2:593-596 (1992).

Los fragmentos de anticuerpo “Fv cadena sencilla” o “sFv” comprenden los dominios  $V_H$  y  $V_L$  del anticuerpo, en donde estos dominios presentan una sola cadena polipeptídica. En general, el polipéptido Fv comprende además un engarce polipeptídico entre los dominios  $V_H$  y  $V_L$  que permite que los sFv formen la estructura deseada para la unión con el antígeno. Para una revisión de sFv véase Pluckthun en The Pharmacology of Monoclonal Antibodies, vol., 113, Rosenberg and Moore, eds. Springer-Verlag, New York, pp. 269-315 (1994).

El término “diacuerpos” hace referencia a fragmentos de anticuerpos pequeños con los dos sitios de unión al antígeno, que comprenden un dominio variable de cadena pesada ( $V_H$ ) conectada con un dominio variable de cadena ligera ( $V_L$ ) en la misma cadena polipeptídica ( $V_H$ - $V_L$ ). Utilizando un engarce que es demasiado corto para permitir el emparejamiento entre los dos dominios en la misma cadena, se obliga a los dominios a emparejarse con los dominios complementarios de otra cadena y crear sitios de unión al antígeno. Los diacuerpos se describen más completamente en, por ejemplo, la patente europea EP 404.097; la patente internacional WO 93/11161; y Hollinger y col., Proc. Natl. Acad. Sci. USA 90:6444-64448 (1993).

La expresión “anticuerpos lineales” cuando se utiliza en esta solicitud hace referencia a los anticuerpos descritos en Zapata y col., Protein Eng. 8 (10):1507-1062 (1995). En resumen, estos anticuerpos comprenden una pareja de segmentos Fd en tándem ( $V_H$ - $C_H$ I- $V_H$ - $C_H$ I) que forman parte de una pareja de regiones de unión al antígeno. Los anticuerpos lineales pueden ser biespecíficos o monoespecíficos.

## II. Utilización y preparación de secuencias HRG

### H. Preparación de secuencias HRG, incluyendo las variantes

El sistema a utilizar en la preparación de una secuencia HRG dependerá de la secuencia HRG particular seleccionada. Si la secuencia HRG es suficientemente pequeña puede prepararse mediante procedimientos de síntesis *in vitro* de polipéptidos. Sin embargo, la HRG se preparará más frecuentemente en cultivo de células recombinantes utilizando los sistemas de huésped-vector descritos más adelante. Una HRG adecuada incluye cualquier HRG activa biológica y antigénicamente.

En general, se utilizarán las células huésped de mamífero, aunque dichos huéspedes pueden carecer de los sistemas post-traduccionales para procesar las preprosecuencias PRO de manera normal. Si las células huésped contienen dichos sistemas, entonces será posible recuperar fragmentos del subdominio natural como HRG-GFD de los cultivos. Si no, entonces se puede lograr el procesamiento adecuado transformando los huéspedes con la enzima(s) requerida(s) o suministrándola(s) en un procedimiento *in vitro*. Sin embargo, no es necesario transformar las células con los genes prepro o estructurales completos para una HRG seleccionada cuando se sólo desee producir fragmentos de secuencias HRG como, p.ej., HRG-GFD. Por ejemplo, se liga un codón de inicio al extremo 5' del DNA que codifica una HRG-GFD, este DNA se utilizó para transformar células huésped y el producto se expresa directamente como la forma Met N-terminal (si se desea, la Met foránea puede eliminarse *in vitro* o con desmetonilasas N-terminales endógenas). Alternativamente, la HRG-GFD se expresa como una fusión con una secuencia señal reconocida por la célula huésped,

la cual procesará y secretará la HRG-GFD madura tal como se describe más adelante. De igual manera se producen variantes de secuencia aminoacídica de las secuencias HRG-GFD.

Las secuencias HRG localizadas entre el primer residuo maduro N-terminal y el primer residuo N-terminal de la secuencia HRG-GFD, denominado HRG-NTD puede funcionar al menos en parte como una secuencia señal no convencional, o como un precursor/transportador circulante normal para HRG-GFD con una sola actividad biológica. La HRG-NTD se produce de la misma manera que la molécula de longitud completa, pero a partir de la expresión de un DNA que tiene un codón de parada en el extremo C-terminal de HRG-NTD. Además, las variantes de HRG se expresan a partir del DNA que codifica la proteína en la que tanto el dominio GFD como el NTD están en su propia orientación pero contienen una inserción, delección o sustitución aminoacídica en el sitio de corte GFD-NTD (localizado dentro de la secuencia VKC) que inhibe o evita la rotura proteolítica del sitio de unión NTD-GFD *in vivo*, y en donde un codón de parada se halla en el extremo 3' de la secuencia codificante GFD. En un ejemplo de este grupo de variantes (denominadas HRG-NTDXGFD), (1) el residuo lisina hallado en la secuencia de unión de NTD-GFD, VKC, se deleciona o (preferentemente) se sustituye por otro residuo distinto al arginil, como el histidil, alanil o seril y (2) se introduce un codón de parada en la secuencia RCT o RCQ en lugar de cisteinil, o treonil (para HRG- $\alpha$ ) o glutaminil (para HRG- $\beta$ ).

Un ligando HRG- $\alpha$  preferido con afinidad de unión por p158<sup>HER</sup> comprende los aminoácidos 226-265 de la figura 1A-D. El ligando HRG- $\alpha$  puede comprender además de 1-20 aminoácidos adicionales antes del aminoácido 226 y de 1-20 aminoácidos después del aminoácido 265. Un ligando de HRG- $\beta$  preferido con afinidad de unión por p185<sup>HER2</sup> comprende los aminoácidos 226-265 de la figura 2A-E. Este ligando HRG- $\beta$  puede comprender hasta 1-20 aminoácidos adicionales antes del 226 y de 1-20 aminoácidos después del 265.

Tal como indicó más arriba, otras secuencias HRG que se preparan de acuerdo con la presente invención corresponden a las de GFD. Éstas se sintetizan *in vitro* o se producen en el cultivo de células recombinantes. Se producen de forma más barata en levaduras o en *E. coli* mediante secreción bajo el control de una señal heteróloga de HRG, tal como se describe más adelante, aunque la preparación en células animales también se contempla utilizando una señal de proteína de mamífero como la de tPA, UK o una proteína viral de secreción. El GFD puede ser la secuencia de una HRG nativa o puede ser una variante de lo mismo tal como se describe más adelante. Las secuencias GFD incluyen aquellas en las que uno o más residuos de un miembro de la familia EGF se sustituyen en la secuencia de GFD.

Una HRG adicional es aquella que contiene el GFD y la secuencia entre el C-terminal del GFD y el N-terminal del dominio transmembrana (siendo este último denominado el dominio de corte C-terminal o CTC). En esta variante (HRG-GFD-CTC) el codón de inicio del DNA está presente en el extremo 5' de la secuencia señal heteróloga de HRG o adyacente al extremo 5' de la región codificante de GFD y un codón de parada se halla en lugar de los tres primeros residuos del dominio extracelular (ECD) o de los tres primeros residuos de la región transmembrana. Además, en algunas variantes HRG-GFD-CTC, los codones se modifican en el sitio de proteólisis GFD-CTC mediante sustitución, inserción o delección. El sitio de proteólisis de GFD-CTC es el dominio que contiene el residuo C-terminal de GFD y aproximadamente 5 residuos en posición N-terminal y 5 residuos en C-terminal respecto a este residuo. Se sabe que la Met-227 terminal y la Val-229 terminal de HRG- $\alpha$ -GFD son biológicamente activas. El C-terminal para HRG- $\alpha$ -GFD puede ser Met-227, Lys-228, Gln-230, Asn-231 o Gln-232 y para HRG- $\beta$ -GFD puede ser Met-226, Ala-227, Ser-228, Phe-229, Trp-230, o Lys231/Ser231. El extremo C-terminal nativo se determina fácilmente mediante secuenciación C-terminal, aunque no es necesario que la HRG-GFD tenga el extremo nativo tan largo, ya que la secuencia GFD posee la actividad deseada. En algunas realizaciones de variantes HRG-GFD-CTC, los cambios aminoacídicos en el CTC se rastrean por su capacidad para resistir la proteólisis *in vitro* e inhibir la proteasa responsable de la generación de HRG-GFD.

Las variantes HRG-ECD se generan proporcionando un codón de parada en la misma localización que las variantes HRG-GFD-CTC. HRG-ECD puede comprender una cualquiera o más de una de las variantes descritas anteriormente relativas a sus subfragmentos, por ejemplo, variantes GFD-CTC que contengan modificaciones del sitio de proteólisis CTC.

Si se desea preparar polipéptidos de HRG más largos y los extremos 5' o 3' de la HRG dada no se han descrito aquí, puede ser necesario preparar ácidos nucleicos en los que los dominios ausentes se sustituyan por regiones homólogas de los ácidos nucleicos HRG más completos. Alternativamente, los dominios ausentes pueden obtenerse hibridando librerías con los DNAs descritos en las Figuras, o con fragmentos de lo mismo.

#### A. Aislamiento del DNA que codifica la Heregulina

El DNA que codifica la HRG puede obtenerse de cualquier librería de cDNA a partir de un tejido que exprese el mRNA de HRG y lo haga a un nivel detectable. El gen HRG- $\alpha$  puede obtenerse de este modo de una librería genómica. Se pueden utilizar procedimientos similares para el aislamiento de otra HRG, como por ejemplo los genes codificantes de HRG- $\beta$ 1, HRG- $\beta$ 2, o HRG- $\beta$ 3.

Las librerías se rastrean con sondas diseñadas para identificar el gen de interés o la proteína codificada por éste. Para las librerías de expresión de cDNA, las sondas adecuadas incluyen anticuerpos monoclonales o policlonales que reconocen y se unen específicamente a HRG- $\alpha$ ; oligonucleótidos de aproximadamente 20-80 bases de longitud que

codifican porciones conocidas o posibles del cDNA de HRG- $\alpha$  de la misma o distinta especie; y/o cDNAs complementarios u homólogos o fragmentos de lo mismo que codifican el mismo gen o uno similar. Las sondas adecuadas para el rastreo de librerías de DNA genómicas incluyen, pero no se limitan a, oligonucleótidos; cDNAs o fragmentos de lo mismo que codifican el mismo o un gen similar; y/o DNAs genómicos homólogos o fragmentos de lo mismo.

5 El rastreo de la librería de cDNA o genómica con la sonda seleccionada se puede realizar utilizando procedimientos estándares, tal como se describen en los capítulos 10-12 de Sambrook y col., *supra*.

Un método alternativo para aislar el gen que codifica HRG- $\alpha$  es utilizar la reacción en cadena de la polimerasa (PCR), tal como se describe en la sección 14 de Sambrook y col., *supra*. Este procedimiento requiere la utilización de sondas oligonucleotídicas que hibridarán con HRG- $\alpha$ . Las estrategias para la selección de oligonucleótidos se describen más adelante.

10

Otro procedimiento alternativo para obtener el gen de interés es sintetizarlo químicamente utilizando uno de los procedimientos descritos en Engels y col., (Agnew. Chem. Int. Ed Engl. 28:716-734, 1989). Estos procedimientos incluyen el triéster, el fosfito, la fosforamidita y los procedimientos del H-fosfonato, la PCR y otros procedimientos de auto-cebado y la síntesis de oligonucleótidos en soportes sólidos. Estos procedimientos pueden utilizarse si se conoce toda la secuencia de ácido nucleico del gen, o si está disponible la secuencia de ácido nucleico complementaria a la cadena codificante, o alternativamente, si la secuencia aminoacídica es conocida, se puede deducir la secuencia de ácido nucleico potencial utilizando residuos codificantes y preferidos para cada residuo aminoacídico.

15

Un procedimiento preferido para la práctica de la presente invención es utilizar secuencias oligonucleotídicas seleccionadas cuidadosamente para rastrear librerías de cDNA de diversos tejidos, preferentemente de líneas de células humanas de mama, colon, glándulas salivares, placenta, feto, cerebro y carcinomas. Otras fuentes biológicas de DNA que codifica un ligando de tipo-herregulina incluye otros mamíferos y pájaros. Entre los mamíferos preferidos están los miembros de las especies siguientes: bovina, ovina, equina, murina y roedores.

20

Las secuencias oligonucleotídicas seleccionadas como sondas deberían ser de longitud suficiente y suficientemente no ambiguas para minimizar falsos positivos. La secuencia nucleotídica actual puede, por ejemplo, estar basada en secuencias nucleotídicas conservadas o altamente homólogas o en regiones de HRG- $\alpha$ . Los oligonucleótidos pueden estar degenerados en más de una posición. La utilización de oligonucleótidos degenerados puede ser de particular importancia si se rastrea una librería en una especie en la que no se conoce su utilización preferencial de codones. El oligonucleótido debe estar marcado de modo que pueda detectarse tras hibridación con el DNA en la librería que se rastrea. El procedimiento preferido de marcaje es utilizar el ATP marcado con P<sup>32</sup> por la polinucleótido quinasa, tal como se conoce en la materia, para marcar isotópicamente el oligonucleótido. Sin embargo, se pueden utilizar otros procedimientos para marcar el oligonucleótido, incluyendo, pero sin limitarse a, la biotilación o el marcaje enzimático.

30

De interés particular es un ácido nucleico HRG- $\alpha$  que codifica un polipéptido de tamaño completo. En algunas realizaciones preferidas, la secuencia de ácido nucleico incluye la secuencia señal de la HRG- $\alpha$  nativa. El ácido nucleico con toda la secuencia codificante de la proteína se obtiene mediante rastreo de librerías de cDNA o genómicas, y, si es necesario, utilizando procedimientos de extensión del cebador tal como se describe en la sección 7.79 de Sambrook y col., *supra*, para detectar precursores y procesar los intermediarios del mRNA que no han sido retro-transcritos a cDNA.

35

El DNA codificante de la HRG de las Figuras 1A-1D puede utilizarse para aislar el DNA que codifica el ligando análogo de otra especie animal mediante procedimientos de hibridación ya discutidos anteriormente. Los animales preferidos son los mamíferos, en particular los bovinos, ovinos, equinos, felinos, caninos y roedores y más especialmente las ratas, los ratones y los conejos.

40

#### 50 B. Variantes de secuencia aminoacídica de la heregulina

Las variantes de secuencia aminoacídica de HRG se preparan introduciendo los cambios nucleotídicos adecuados en el DNA de HRG, o mediante la síntesis *in vitro* del polipéptido HRG deseado. Dichas variantes incluyen, por ejemplo, deleciones de, o inserciones o sustituciones de residuos en la secuencia aminoacídica mostrada para las secuencias HRG humanas. Cualquier combinación de deleción, inserción y sustitución puede realizarse para lograr la construcción final, siempre que ésta posea las características deseadas. Se excluyen del ámbito de la invención las variantes de HRG o secuencias polipeptídicas que no sean nuevas ni las más indicadas según los antecedentes en la materia. Los cambios aminoacídicos también pueden alterar procesos post-traduccionales de HRG- $\alpha$ , como cambiar el número o la posición de los sitios de glicosilación, alterar las características de anclaje a la membrana, alterar la localización intra-celular de HRG mediante inserción, deleción, o afectando la secuencia líder de la HRG nativa, o modificando su susceptibilidad al corte proteolítico.

55

La secuencia HRG puede procesarse proteolíticamente para crear diversos fragmentos HRG. Las secuencias HRG-GFD de HRG- $\alpha$  contienen secuencias aminoacídicas entre las cisteína 226 y la cisteína 265 de HRG. El extremo amino-terminal del fragmento de HRG- $\alpha$  puede resultar del corte de cualquier enlace peptídico entre la alanina 1 y la cisteína 226, preferentemente adyacente a una arginina, lisina, valina o metionina y más preferentemente entre la metionina 45 y la serina 46. El extremo carboxi-terminal del fragmento de HRG- $\alpha$  puede resultar del corte de cualquier enlace peptídico entre la cisteína 265, preferentemente adyacente a una arginina, lisina, valina o metionina

60

y más preferentemente entre la lisina 272 y la valina 273, entre la lisina 278 y la alanina 279, o entre la lisina 285 y la arginina 286. Los ligandos de HRG- $\alpha$  resultantes de dicho procesamiento proteolítico son los ligandos preferidos.

Las HRG- $\beta$ -GFD son análogas a las descritas más arriba para HRG- $\alpha$ -GFD. Cada HRG- $\beta$ -GFD contiene el segmento polipeptídico desde la cisteína 212 a la 251 de la figura 2A-E. El extremo amino-terminal del fragmento HRG- $\beta$ 1 puede resultar del corte de cualquier enlace peptídico entre la alanina 1 y la cisteína 212, preferentemente adyacente a una arginina, lisina, valina o metionina y más preferentemente entre la metionina 31 y la serina 32. El extremo carboxi-terminal del fragmento HRG- $\beta$ 1 puede resultar del corte de cualquier enlace peptídico entre la cisteína 251 de la Figura 2A-2E, preferentemente adyacente a una arginina, lisina, valina, o metionina y más preferentemente entre la valina 255 y la metionina 256, entre la lisina 261 y la histidina 262, entre la lisina 276 y la alanina 277, o entre la lisina 301 y la treonina 302. De modo similar, los ligandos HRG- $\beta$ 1 resultantes de dicho procesamiento proteolítico se hallan entre los ligandos preferidos. De igual modo, el procesamiento para producir ligandos de fragmentos preferidos de HRG- $\beta$ 2 según la Fig. 3A-3E y la HRG- $\beta$ 3 según la Fig. 4A-4C puede lograrse cortando las secuencias HRG de las Figs. 3A-3E y 4A-4C, preferentemente adyacentes a una arginina, lisina, valina o metionina.

En el diseño de variantes de secuencia aminoacídica de HRG, la localización del sitio de mutación y la naturaleza de la mutación dependerá de las características a modificar. Los sitios de mutación pueden modificarse individualmente o en series, p.ej., (1) sustituyendo primero con aminoácidos conservadores y luego con otros más radicales dependiendo de los resultados logrados, (2) delecionando el residuo diana, o (3) insertando residuos de otros ligandos del receptor adyacentes al sitio localizado.

Un procedimiento de utilidad para la identificación de ciertos residuos o regiones del polipéptido HRG que sean localizaciones preferidas para la mutagénesis es el denominado "mutagénesis por rastreo de alaninas", tal como se describe por Cunningham y Wells (Science, 244:1081-1085, 1989). Aquí, se identifica un residuo o grupo de residuos diana (p.ej., residuos cargados como la arg, asp, his, lys y glu) y se reemplazan por un aminoácido neutro o cargado negativamente (preferentemente la alanina o la polialanina) para afectar la interacción de los aminoácidos con el entorno acuoso dentro o fuera de la célula. Aquellos dominios que demuestran una sensibilidad funcional a las sustituciones se remodelan a continuación introduciendo otras variantes en los sitios de sustitución. En consecuencia, mientras que el sitio para la introducción de una variación de secuencia aminoacídica está predeterminado, la naturaleza de la mutación *per se* no tiene porqué estarlo. Por ejemplo, para optimizar la realización de una mutación en un sitio dado, se puede realizar el rastreo de alaninas o la mutagénesis aleatoria en el codón o región diana y a continuación, rastrear la combinación óptima que tenga la actividad deseada de las variantes HRG.

En la construcción de variantes de secuencia aminoacídica existen dos variables principales: la localización del sitio de mutación y la naturaleza de la mutación. Estas variantes de la secuencia HRG pueden representar alelos naturales (que no requerirán manipulación del DNA de HRG) o formas mutantes predeterminadas generadas mediante mutagénesis del DNA, para obtener un alelo o una variante que no se halle en la naturaleza. En general, la localización y la naturaleza de la mutación elegida dependerán de la característica de la HRG a modificar. Obviamente, dichas variantes, que por ejemplo convierten la HRG en un ligando de receptor conocido, no se incluyen dentro del ámbito de la invención, tampoco las otras variantes de HRG o secuencias polipeptídicas que no sean nuevas ni las más indicadas según las referencias en la materia.

Las delecciones de secuencia aminoacídica abarcan generalmente de 1 a 30 residuos, más preferentemente de 1 a 10 residuos y típicamente de 1 a 5 residuos contiguos. Las delecciones se pueden introducir en las regiones de baja homología con otros precursores de la familia EGF para modificar la actividad HRG. Las delecciones de HRG en áreas de homología sustancial con las secuencias de la familia EGF modificarán probablemente la actividad biológica de HRG de forma más significativa. El número de delecciones consecutivas se seleccionará para preservar la estructura terciaria de HRG en el dominio afectado, p.ej., la unión de cisteínas, la hoja plegada beta o la hélice alfa.

Las inserciones de secuencia aminoacídica incluyen las fusiones amino- y/o carboxi-terminal cuya longitud varía desde 1 residuo a polipéptidos que contienen un centenar o más de residuos, así como la inserción intrasecuencia de uno o múltiples residuos aminoacídicos. Las inserciones intrasecuencia (es decir, inserciones dentro de la secuencia de HRG) pueden ser de aproximadamente 1 a 10 residuos, más preferentemente de 1 a 5 y todavía mejor de 1 a 3 residuos aminoacídicos. Ejemplos de inserciones terminales incluyen la HRG con un residuo N-terminal (un artefacto de la expresión directa de HRG en un cultivo de bacterias recombinantes) y la fusión de una secuencia señal N-terminal de HRG heteróloga para facilitar la secreción de la HRG madura en las células huésped recombinantes. Dichas secuencias señal se obtendrán generalmente de, y en consecuencia son homólogas a, las especies de huéspedes escogidas. Las secuencias adecuadas incluyen STII, tPA o Ipp para *E. coli*, el factor alfa para levaduras y las señales virales como gD del virus herpes para las células de mamífero.

Otras variantes de inserción de HRG incluyen la fusión del extremo N- o C-terminal de HRG de un polipéptido inmunogénico, p.ej., los polipéptidos de bacterias como la beta-lactamasa o una enzima codificada por el locus *trp* de *E. coli*, o la proteína de levadura, la albúmina sérica bovina y los polipéptidos quimiotácticos. También se contemplan las fusiones C-terminales de HRG-ECD con proteínas que tienen una vida media larga como las regiones constantes de inmunoglobulina (u otras regiones de inmunoglobulina), la albúmina, o la ferritina, tal como se describe en la patente internacional WO 89/02922, publicada el 6 de abril de 1989.

## ES 2 329 467 T3

Otro grupo de variantes son las variantes de sustitución. En estas variantes de HRG, se ha sustituido por lo menos un residuo aminoacídico por otro distinto en su lugar. Los sitios de mayor interés para la mutagénesis sustitucional incluyen los sitios identificados como sitios activos de HRG, y los sitios en donde los aminoácidos hallados en los ligandos HRG de diversas especies son diferentes en términos de volumen de cadena lateral, carga y/o hidrofobicidad.

5 Un sub-dominio similar de HRG-GFD que tenga actividad biológica como un factor de crecimiento es el segmento C-terminal, en particular dentro de la secuencia desde aproximadamente la glicina 218 a la valina 226 (HRG- $\alpha$ ), y la glicina 218 a la lisina 228/serina 228 (HRG- $\beta$ ) según la analogía con la subsecuencia de EGF con actividad EGF.

Otros sitios de interés son aquellos residuos idénticos en ligandos de tipo HRG obtenidos de especies diversas.

10 Estas posiciones pueden ser importantes para la actividad biológica de la HRG. Estos sitios, especialmente aquellos que se hallan dentro de una secuencia que comprende al menos tres sitios conservados de forma idéntica, se sustituyen de forma relativamente conservada. Dichas sustituciones conservadas se muestran en la Tabla 1 bajo el encabezamiento de "sustituciones preferidas". Si dichas sustituciones resultan en un cambio de la actividad biológica, se introducen y rastrean más cambios sustanciales, denominados ejemplos de sustituciones en la Tabla 1, o como se describe más  
15 adelante respecto a las clases de aminoácidos.

TABLA 1

| 20 | Residuo original | Ejemplos de sustituciones           | Sustituciones preferidas |
|----|------------------|-------------------------------------|--------------------------|
| 25 |                  |                                     |                          |
|    | Ala (A)          | Val; leu; ile                       | Val                      |
|    | Arg (R)          | Lys, gln, asn                       | Lys                      |
| 30 | Asn (N)          | Gln, hi, lys, arg                   | Gln                      |
|    | Asp (D)          | Glu                                 | Glu                      |
|    | Cys (C)          | Ser                                 | Ser                      |
| 35 | Gln (Q)          | Asn                                 | Asn                      |
|    | Glu (E)          | Asp                                 | Asp                      |
|    | Gly (G)          | Pro                                 | Pro                      |
| 40 | His (H)          | Asn, gln, lys, arg                  | Arg                      |
|    | Ile (I)          | Leu, val, met, ala, phe, norleucina | Leu                      |
| 45 | Leu (L)          | Norleucina, ile, val, met, ala, phe | Ile                      |
|    | Lys (K)          | Arg, gln, asn,                      | Arg                      |
|    | Met (M)          | Leu, phe, ile                       | Leu                      |
| 50 | Phe (F)          | Leu, val, ile, ala                  | Leu                      |
|    | Pro (P)          | Gly                                 | Gly                      |
|    | Ser (S)          | Thr                                 | Thr                      |
| 55 | Thr (T)          | Ser                                 | Ser                      |
|    | Trp (W)          | Tyr                                 | Tyr                      |
|    | Tyr (Y)          | Trp, phe, thr, ser                  | Phe                      |
| 60 | Val (V)          | Il, leu, met, phe, ala, norleucina  | leu                      |

Las modificaciones en la función o la identidad inmunológica de HRG se realizan seleccionando sustituciones que difieran significativamente en su efecto o manteniendo (a) la estructura del esqueleto del polipéptido en el área de la sustitución, por ejemplo, como una hoja o conformación helicoidal, (b) la carga o la hidrofobicidad de la molécula en el sitio diana, o (c) el volumen de la cadena lateral. Los residuos naturales se dividen en los siguientes grupos basándose en las propiedades de su cadena lateral común:

## ES 2 329 467 T3

1) hidrofóbicos: norleucina, met, ala, val, leu, ile;

2) hidrofílicos neutros: cys, ser, thr;

3) ácidos: asp, glu;

4) básicos: asn, gln, his, lys, arg;

5) residuos que influyen la orientación de la cadena lateral: gly, pro y

6) aromáticos: trp, tyr, phe.

Las sustituciones no conservadoras implican el cambio de un miembro de una de estas clases por otro. Dichos residuos sustituidos pueden introducirse en regiones HRG que sean homólogas con otros ligandos de receptor, o más preferentemente, en regiones no-homólogas de la molécula.

En una realización de la invención, se inactivan uno o más sitios de corte de proteasas presentes en la molécula. Estos sitios se identifican por la inspección de la secuencia aminoacídica codificada. Si se identifican los sitios de corte de proteasas, se vuelven inactivos al corte proteolítico por sustitución del residuo diana por otro residuo, preferentemente un residuo básico como la glutamina o un residuo hidrofílico como la serina; delecionando el residuo; o insertando un residuo prolil inmediatamente después del residuo.

En otra realización, se sustituyeron por otro residuo (preferentemente de acuerdo con la Tabla 1), o se delecionó cualquier residuo metionil distinto al residuo metionil de inicio de la secuencia señal, o cualquiera de los tres residuos N- o C-terminal para cada residuo metionil. Alternativamente, se insertaron 1-3 residuos adyacentes a dichos sitios.

También puede sustituirse cualquier residuo cisteína no implicado en mantener la conformación adecuada de HRG, en general por serina, para mejorar la estabilidad oxidativa de la molécula y prevenir uniones aberrantes.

Sitios adecuados para las sustituciones, deleciones o inserciones, o para su utilización como fragmentos, incluyen los siguientes numerados desde el extremo N-terminal de HRG- $\alpha$  de la Figura 1-A-1D:

1) adición de sitios de glicosaminoglicano potenciales en los dipéptidos serina-glicina en las posiciones: 42-43, 64-65, 151-152;

2) glicosilación potencial unida a asparragina en las posiciones: 164, 170, 208 y 437 (NDS), 164-166, (NIT) 170-172, (NTS) 208-210 y 609-611 (NTS);

3) O-glicosilación potencial en un grupo de serinas y treoninas en 209-218;

4) cisteínas en 226, 234, 254, 256 y 265;

5) dominio transmembrana en 287-309;

6) bucle I limitado por las cisteínas 226 y 240;

7) bucle 2 limitado por las cisteínas 234 y 254;

8) bucle 3 limitado por las cisteínas 256 y 265; y

9) sitios de procesamiento potencial de proteasas en 2-3, 8-9, 23-24, 33-34, 36-37, 45-46, 48-49, 62-63, 66-67, 86-87, 110-111, 123-124, 134-135, 142-143, 272-273, 278-279 y 285-286;

Las regiones análogas en la HRG- $\beta$ 1 pueden determinarse por referencia a su secuencia. Los aminoácidos de HRG- $\beta$ 1 análogos pueden mutarse o modificarse tal como se describió para la HRG- $\alpha$ . Las regiones análogas en HRG- $\beta$ 2 también pueden determinarse por referencia a su secuencia. Asimismo, las regiones análogas en HRG- $\beta$ 3 también pueden determinarse por referencia a su secuencia. Los aminoácidos análogos de HRG- $\beta$ 2 pueden mutarse o modificarse tal como se describió anteriormente para HRG- $\alpha$  o HRG- $\beta$ 1. Además, los aminoácidos análogos de HRG- $\beta$ 3 pueden mutarse o modificarse tal como se discutió para HRG- $\alpha$ , HRG- $\beta$ 1, o HRG- $\beta$ 2.

Otra variante HRG es la HRG- $\gamma$  (o gamma-heregulina). La  $\gamma$ -heregulina es cualquier secuencia polipeptídica que posea al menos una propiedad biológica de secuencia nativa  $\gamma$ -HRG con la SEC ID N°:11. La propiedad biológica de esta variante es la misma que la citada anteriormente. La variante abarca no sólo el polipéptido aislado de una fuente de HRG- $\gamma$  nativa, como las células MDA-MB-175 o de otra fuente, como especies de otros animales, sino también el polipéptido preparado por procedimientos recombinantes o sintéticos. También incluye las formas de variantes funcionales, las variantes alélicas, las isoformas naturales y los análogos de lo mismo. A veces, la HRG- $\gamma$  es una "HRG- $\gamma$

nativa" que hace referencia a un polipéptido HRG- $\gamma$  endógeno que se ha aislado de un mamífero. La HRG- $\gamma$  también puede ser una "HRG- $\gamma$  de secuencia nativa" en la medida en que tiene la misma secuencia aminoacídica que la HRG- $\gamma$  nativa (p.ej., la HRG- $\gamma$  humana que se muestra en la Fig. 7A-7C). Las variantes de secuencia aminoacídica de la secuencia nativa se preparan introduciendo cambios nucleotídicos en el DNA de secuencia nativa, o mediante síntesis *in vitro* del polipéptido deseado. Dichas variantes incluyen, por ejemplo, deleciones, o inserciones o sustituciones de residuos dentro de la secuencia aminoacídica que se muestra para la proteína humana en la Fig. 7A-7C, tal como se describió anteriormente para otra HRG. Para una obtener una construcción final, será realizar cualquier combinación de deleción, inserción y sustitución para lograr una, siempre que dicha construcción final posea las características deseadas. Los cambios aminoacídicos también pueden alterar los procesos post-traduccionales de la secuencia nativa, como el cambio del número o posición de los sitios de O-glicosilación.

Otra variante es la referida al polipéptido como el factor derivado de neuronas sensoriales y motoras (SMDF), cuyas secuencias de ácido nucleico y aminoacídicas (SEC ID N°: 13 y 14), que se muestran en la Fig. 8, pueden prepararse tal como se describe en la patente internacional WO 96/15244. Los polipéptidos SMDF de la invención presentan propiedades de unión a los receptores de HER2/HER3 y estimulación del crecimiento y diferenciación de las células epiteliales de forma similar a los polipéptidos HRG descritos anteriormente. Las variantes de secuencia aminoacídica de SMDF de secuencia nativa se preparan introduciendo los cambios de nucleótidos adecuados en la secuencia nativa del DNA de SMDF, o mediante síntesis *in vitro* del polipéptido SMDF, tal como se indicó anteriormente para otra HRG. Dichas variantes incluyen, por ejemplo, deleciones, inserciones o sustituciones de residuos dentro de la secuencia aminoacídica de SMDF que se muestra en la Fig. 8. Para una obtener una construcción final, será realizar cualquier combinación de deleción, inserción y sustitución para lograr una, siempre que dicha construcción final posea las características deseadas. Los cambios aminoacídicos también pueden alterar los procesos post-traduccionales de SMDF de secuencia nativa, como un cambio en el número o posición de los sitios de O-glicosilación.

Variantes adicionales incluyen polipéptidos en los que la variante tiene una sustitución aminoacídica en un residuo seleccionado y correspondiente a un residuo de la heregulina- $\beta$ 1 humana nativa de 645 aminoácidos seleccionada a partir de:

S177, H178, L179. V180, K181, E194. E186,  
K187, T188, V191, N192, G193, G194, E195,  
M198, V 199, K200, D201, N204, P205, S206,  
R207, Y208, L209, K211. P213, N214, E215,  
T217, G218, D219, Q222, N223, Y224, S228, y F229.

En una variación de esta realización, la sustitución aminoacídica no es una sustitución del residuo seleccionado por un residuo de un factor de crecimiento epitelial (EGF) correspondiente con el residuo seleccionado.

Otras variantes de heregulina- $\beta$ 1 incluyen una sustitución aminoacídica seleccionado a partir de los residuos siguientes:

S177W; H178S, E, R, or A; V180Q, I o E;  
K181P o A; A183G; E184V, W, K, R G, o N;  
K185E, S, Q, o G; E186R; K187E or A; T188Q;  
E195Q; F197Y; M198R o K; K200R; D201T o I;  
P205T o Y; S206K. H, G, P, o R; R207Y;  
Y208R o L; L209M or G; K211R; P213S, T, N, o K;  
N214L, K, S, o E; F216M; N223H o W; y M226L.

En una variación de esta realización, la variante de heregulina incluye sustituciones de aminoácidos que presentan al menos un incremento de 50 veces en la afinidad por el receptor HER3, que también se acompaña por un incremento en la afinidad por el receptor HER4. La variante específica incluye los residuos siguientes:

## ES 2 329 467 T3

A183G, E184W, K185D, E186R, K187E, T188G, M226I;  
 A183D, E184K, K185S, E186R, K187E, T188G, M226I;  
 5 F197Y, M198K, K200R, D201I, M226I;  
 P205Y, S206G, R207Y, Y208L, L209M;  
 P205Y, S206R, R207Y, Y208R, L209M, M226I;  
 10 P205T, S206H, R207Y, Y208R, L209M;  
 P205T, S206K, R207Y, Y208R, L209G;  
 N223W. M226I;  
 15 N223H. M226I;  
 S177W, H178E, K181P, A183G, E184W, K185D, E186R. K187E,  
 T188G, M226I;  
 20 P203Y, S206G, R207Y, Y208L, L209M, M226I;  
 A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T:  
 A183G, K185E, E186R, K187E, T188G, P205Y, S206G, R207Y,  
 25 Y208L, L209M;  
 A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T,  
 P205Y, S206G, R207Y, Y208L, L209M;  
 30 A183G, K185E, E186R, K187E, T188G, M226I;  
 F197Y, M198R, D201T. P205Y, S206G, R207Y, Y208L, L209M:  
 35 F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, L209M.  
 M226I;  
 F197Y, M198R, D201T, M226I;  
 40 A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T,  
 M226I;  
 A183G, K185E, E186R, K187E, T188G, P205Y, S206G, R207Y,  
 45 Y208L, L209M, M226I;  
 A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T,  
 P205Y, S206G, R207Y, Y208L, L209M,  
 50 M226I;  
 F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, L209M,  
 N223H, M226I; and  
 55 A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T,  
 P205Y, S206G, R207Y, Y208L, L209M,  
 60 N223H. M226I.

Además de incluir una o más de las sustituciones aminoacídicas aquí descritas, la variante de heregulina puede tener o no otras modificaciones, como la sustitución aminoacídica, una inserción de al menos un aminoácido, una delección de al menos un aminoácido, o una modificación química. Por ejemplo, la invención proporciona una variante de heregulina que es un fragmento de ésta. En una variación de esta realización, el fragmento incluye residuos correspondientes a una porción de la heregulina- $\beta$ 1 humana que se extiende desde aproximadamente el residuo 175 al residuo

230 (es decir, el dominio de tipo-EGF). Por ejemplo, el fragmento puede extenderse desde el residuo 177 al 244 y puede prepararse mediante técnicas recombinantes (rHRG $\beta$ 1-177-244).

Las variantes de DNA codificante de la secuencia aminoacídica de la HRG se preparan mediante procedimientos conocidos en la materia. Estos procedimientos incluyen, pero no se limitan a, el aislamiento a partir de una fuente natural (en el caso de variantes de secuencia aminoacídica que ocurren de forma natural), o la preparación mediante mutagénesis mediada por oligonucleótidos (o mutagénesis dirigida), mutagénesis por PCR y la mutagénesis de *casete* de una versión variante o no de la HRG preparada con anterioridad. Estas técnicas pueden utilizar el ácido nucleico de la HRG (DNA o RNA) o el ácido nucleico complementario al ácido nucleico de HRG.

La mutagénesis mediada por oligonucleótidos es un procedimiento preferido para preparar las variantes de sustitución, delección e inserción del DNA de HRG. Esta técnica es bien conocida en la materia tal como se describe por Adelman y col., DNA, 2:183 (1983). En resumen, se modifica el DNA de HRG hibridando un oligonucleótido que codifica la mutación deseada con un molde de DNA, en donde el molde es la forma de cadena sencilla de un plásmido o bacteriófago que contiene la secuencia de DNA nativa o inalterada de la HRG. Después de la hibridación, se utiliza una polimerasa de DNA para sintetizar una segunda cadena complementaria completa del molde, que en consecuencia incorporará el cebador oligonucleotídico y codificará la alteración seleccionada en el DNA de HRG.

En general, se utilizan oligonucleótidos de menos de 25 nucleótidos de longitud. Un oligonucleótido óptimo tendrá de 12 a 15 nucleótidos que sean completamente complementarios con el molde por cada lado del nucleótido codificante de la mutación. Ello asegura que el oligonucleótido hibridará adecuadamente con la molécula molde de DNA de cadena sencilla. Los oligonucleótidos se sintetizan fácilmente utilizando técnicas conocidas en la materia como las descritas por Crea y col., (Proc. Natl. Acad. Sci. USA, 75:576, 1978).

El molde de DNA de cadena sencilla puede también generarse desnaturalizando el DNA del plásmido (u otro) de cadena doble mediante técnicas estándares.

Para la alteración de la secuencia de DNA nativo (para generar variantes de secuencia aminoacídica, por ejemplo), se hibrida el oligonucleótido a un molde de cadena sencilla en condiciones de hibridación adecuadas. Una enzima polimerizante de DNA, generalmente el fragmento Klenow de la DNA polimerasa I, se añade para sintetizar la cadena complementaria del molde utilizando el oligonucleótido como cebador en la síntesis. De este modo se forma una molécula heterodúplex en la que una cadena de DNA codifica la forma mutada de HRG y la otra cadena (el molde original) codifica la forma nativa de la secuencia inalterada de HRG. La molécula heterodúplex se transforma a continuación en una célula huésped adecuada, generalmente una procarionota como *E. coli* JM101. Una vez las células están crecidas, se siembran en placas de agarosa y se rastrean utilizando el oligonucleótido marcado radioactivamente con fosfato-P<sup>32</sup> para identificar las colonias bacterianas que contienen el DNA mutado. La región mutada se extrae a continuación y se coloca en un vector adecuado para la producción de proteína, en general un vector de expresión del tipo utilizado para la transformación de un huésped adecuado.

El procedimiento descrito más arriba puede modificarse para crear una molécula de homodúplex, en la que ambas cadenas del plásmido contengan la mutación. Las modificaciones son las siguientes: el oligonucleótido de cadena sencilla se anilla al molde de cadena sencilla, tal como se describió anteriormente. Una mezcla de tres desoxiribonucleótidos, desoxiriboadenosina (dATP), desoxiriboguanosina (dGTP) y desoxiribotimidina (dTTP), se combina con un tio-desoxiribocitosina denominado dCTP-(aS) (que puede obtenerse de Amersham Corporation). Esta mezcla se añade al complejo molde-oligonucleótido. Tras la adición de la polimerasa de DNA, se genera una cadena de cDNA idéntica al molde con excepción de las bases mutadas. Además, esta nueva cadena de DNA contendrá dCTP-(aS) en lugar de dCTP, que sirve para protegerlo de la digestión con endonucleasas de restricción. Después de que la cadena molde del heterodúplex de cadena doble se melle con una enzima de restricción adecuada, se puede digerir la cadena molde con la nucleasa ExoIII u otra nucleasa adecuada pasada la región que contenga los sitios a mutar. A continuación, se para la reacción para dejar una molécula que sea parcialmente de cadena sencilla. Se forma así un homodúplex de DNA de cadena doble utilizando la polimerasa de DNA en presencia de los cuatro desoxiribonucleótidos trifosfatos, ATP y DNA ligasa. Esta molécula de homodúplex puede transformarse en una célula huésped adecuada como *E. coli* JM101, tal como se describió más arriba.

El DNA que codifica los mutantes de HRG con más de un aminoácido sustituido puede generarse de diversas maneras. Si los aminoácidos se localizan próximos unos de otros en la cadena polipeptídica, pueden mutarse simultáneamente utilizando un oligonucleótido que codifique para todos las sustituciones aminoacídicas deseadas. Si, al contrario, los aminoácidos se localizan a cierta distancia unos de otros (separados por más de 10 aminoácidos aproximadamente), es más difícil generar un oligonucleótido que codifique todos los cambios deseados. En su lugar, puede utilizarse uno de los dos procedimientos alternativos.

En el primer procedimiento, se genera un oligonucleótido independiente para cada sustitución de aminoácido. Los oligonucleótidos se anillan a continuación con el DNA molde de cadena sencilla simultáneamente y así la segunda cadena de DNA que se sintetiza a partir del molde codificará todas las sustituciones de aminoácidos deseadas.

El procedimiento alternativo implica dos o más tandas de mutagénesis para producir el mutante deseado. La primera tanda es tal como se describió para los mutantes sencillos: se utiliza el DNA de tipo salvaje para el molde y

un oligonucleótidos que codifica la(s) primera(s) sustitución(es) de aminoácidos deseados se hibrida con el molde, generándose en consecuencia la molécula de DNA heterodúplex. La segunda tanda de mutagénesis utiliza el DNA producido en la primera tanda de mutagénesis como el molde. En consecuencia, este molde contiene una o más mutaciones. El oligonucleótido codificante de la(s) sustitución(es) aminoacídica adicional se hibrida con el molde, con lo que la cadena de DNA resultante codifica ahora para las mutaciones de ambas tandas de mutagénesis. Este DNA resultante puede utilizarse como un molde en una tercera tanda de mutagénesis, y así sucesivamente.

La mutagénesis por PCR también es adecuada para generar variantes aminoacídicas de HRG. Mientras que la descripción siguiente hace referencia al DNA, se entiende que la técnica también halle aplicación con RNA. La técnica de PCR hace referencia en general al siguiente procedimiento (véase, Erlich, *supra*, el capítulo de R. Higuchi, p. 61-70). Cuando se utiliza como material de inicio para una PCR, cantidades pequeñas de DNA molde, es posible utilizar cebadores que difieran ligeramente en la secuencia de la región correspondiente en el DNA molde, para generar cantidades relativamente grandes de un fragmento de DNA específico que difiera de la secuencia molde únicamente en aquellas posiciones donde los cebadores difieren del molde. Para introducir una mutación en un DNA plasmídico, se diseña uno de los cebadores para que solape con la posición de la mutación y contenga dicha mutación, la secuencia del otro cebador debe ser idéntica a un fragmento de la secuencia de la cadena opuesta del plásmido, aunque esta secuencia puede estar localizada en cualquier sitio del DNA plasmídico. Sin embargo, es preferible que la secuencia del segundo cebador se localice en los siguientes 200 nucleótidos, para que al final pueda secuenciarse fácilmente toda la región amplificada del DNA sintetizado con los dos cebadores. La amplificación por PCR utilizando una pareja de cebadores similares a la descrita más arriba resulta en una población de fragmentos de DNA que difiere en la posición de la mutación especificada por el cebador y posiblemente en otras posiciones, ya que en la reacción de copiado del molde se producen errores.

Si la proporción de molde respecto a producto es extremadamente baja, la gran mayoría de fragmentos de DNA producidos habrán incorporado la(s) mutación(es) deseada(s). Este producto se utiliza para sustituir la región correspondiente en el plásmido que sirvió como molde de PCR utilizando tecnología estándar del DNA. Las mutaciones en posiciones separadas pueden introducirse simultáneamente utilizando un segundo cebador, o realizando una segunda PCR con cebadores mutantes distintos y ligando los dos fragmentos de PCR resultantes con el fragmento del vector en una ligación de tres (o más etapas).

En un ejemplo específico de mutagénesis por PCR, se lineariza el DNA plasmídico molde (1 mg) mediante digestión con una endonucleasa de restricción que tenga una diana única en el DNA plasmídico fuera de la región a amplificar. Se añaden 100 ng del material anterior a la mezcla de PCR que contiene el tampón de PCR con los cuatro desoxinucleótidos trifosfatos incluidos en el equipo GENEAMP (de Perkin-Elmer Cetus, Norwalk, CT y Emeryville, CA), y 25 pmoles de cada uno de los cebadores oligonucleotídicos hasta un volumen final de 50 ml. La mezcla de reacción se cubre con 35 ml de aceite mineral. La reacción se desnaturaliza durante 5 minutos a 100°C, se coloca rápidamente en hielo y se añade 1 ml de DNA polimerasa de *Thermus aquaticus* (Taq) (5 unidades/ml, de Perkin-Elmer Cetus, Norwalk, CT y Emeryville, CA) por debajo de la capa de aceite mineral. La mezcla de reacción se inserta en un Termociclador de DNA (de Perkin-Elmer Cetus) y se procede con el programa siguiente:

2 min a 55°C

30 s a 72°C, luego 19 ciclos de lo siguiente:

30 s a 94°C

30 s a 55°C y

30 s a 72°C.

Al término del programa, el vial de la reacción se extrae del termociclador y la fase acuosa se transfiere a un nuevo vial, se realiza una extracción con fenol/cloroformo (50:50:vol), una precipitación con etanol y se extrae el DNA mediante procedimientos estándares. Este material se somete a continuación a tratamientos adecuados para la inserción en un vector.

Otro procedimiento para la preparación de variantes, la mutagénesis en casete, se basa en la técnica descrita por Wells y col., (Gene, 34:315, 1985). El material de partida es el plásmido (u otro vector) que comprende el DNA de la HRG a mutar. En primer lugar, se identifica el codón/codones en el DNA de HRG a mutar. Debe existir una sola diana de restricción en cada lado de la mutación identificada. Si no existieran las dianas de restricción, se pueden generar utilizando el procedimiento de la mutagénesis mediada por oligonucleótidos descrita anteriormente para introducirlos en localizaciones adecuadas en el DNA de HRG. Después de introducir las dianas de restricción en el plásmido, éste se digiere por dichas dianas para su linearización. A continuación, utilizando los procedimientos estándares, se sintetiza un oligonucleótido de doble cadena que codifica la secuencia de DNA entre ambas dianas de restricción pero con la mutación deseada. Las dos cadenas se sintetizan separadamente y luego se hibridan juntas mediante técnicas estándares. A este oligonucleótido de doble cadena se le denomina casete. Este casete se diseña con extremos 3' y 5' compatibles con los extremos del plásmido linearizado, de modo que pueda ser ligado directamente en el plásmido. Este plásmido contiene ahora la secuencia mutada del DNA de HRG.

### C. Inserción del DNA en un vector de clonaje

El cDNA o el DNA genómico que codifica una HRG nativa o variante se inserta en un vector replicable para su posterior clonaje (amplificación del DNA) o para su expresión. Existen en la actualidad muchos vectores disponibles, cuya selección dependerá de: 1) si se ha de utilizar para la amplificación del DNA o la expresión; 2) el tamaño del DNA a insertar en el vector, y 3) la célula huésped a transformar con el vector. Cada vector contiene varios componentes dependiendo de su función (amplificación o expresión del DNA) y de la célula huésped para la cual es compatible. Los componentes del vector incluyen generalmente, sin estar limitados por ello, uno o más de los siguientes: una secuencia señal, un origen de replicación, uno o más genes marcadores, un elemento intensificador de la transcripción, un promotor y una secuencia de terminación de la transcripción.

#### (I) Componente de secuencia señal

En general, la secuencia señal puede ser un componente del vector, o puede ser una parte del DNA de HRG que se inserta en el vector. El DNA de la HRG nativa contiene probablemente una secuencia señal en el extremo amino-terminal (extremo 5' del DNA que codifica HRG) del polipéptido que se corta post-traduccionalmente para formar el ligando polipeptídico de HRG madura que se une al receptor HER2/HER3, aunque no se ha demostrado la presencia de una estructura señal convencional. La HRG nativa es secretada por las células pero permanece alojada en la membrana porque contiene un dominio transmembrana y una región citoplasmática en la región carboxi-terminal del polipéptido. En consecuencia, en una versión soluble de la HRG secretada, se delecta el dominio carboxi-terminal de la molécula, incluyendo el dominio transmembrana. Esta variante truncada del polipéptido HRG podrá secretarse por las células siempre que el DNA de la variante truncada codifique una secuencia señal que sea reconocida por el huésped.

La HRG de la presente invención puede expresarse no sólo directamente, sino también como una fusión con un polipéptido heterólogo, preferentemente una secuencia señal u otro polipéptido que tenga un sitio de corte específico en el extremo N- o C-terminal de la proteína madura o polipéptido. En general, la secuencia señal puede ser un componente del vector, o puede ser una parte del DNA de HRG que se inserta en el vector. Se incluyen dentro del ámbito de la presente invención las HRG con la secuencia señal nativa delectada y sustituida por una secuencia señal heteróloga. La secuencia señal heteróloga seleccionada debería ser una señal que sea reconocida y procesada, es decir se corte, por una peptidasa señal de la célula huésped. En las células huésped procariotas, donde no se reconoce ni procesa la secuencia señal de HRG nativa, dicha secuencia señal se sustituye por una secuencia señal procariota seleccionada, por ejemplo, del grupo compuesto por la fosfatasa alcalina, la penicilinas, Ipp, o los líderes de la enterotoxina II termo-estable. Para la secreción en levaduras, la secuencia señal de HRG nativa puede sustituirse por la invertasa de levadura, el factor alfa, o los líderes de la fosfatasa ácida. En la expresión en células de mamífero, la secuencia señal de HRG nativa es satisfactoria, aunque también son adecuadas otras secuencias señal de mamíferos.

#### (II) Origen del componente de replicación

Tanto los vectores de expresión como los de clonaje contienen una secuencia de ácido nucleico que permite que el vector sea replicable en una o más células seleccionadas. En general, en los vectores de clonaje esta secuencia permite la replicación del vector independientemente del DNA cromosómico del huésped e incluye los orígenes de replicación o las secuencias de replicación autónomas. Dichas secuencias son bien conocidas para muchas bacterias, levaduras y virus. El origen de replicación del plásmido pBR322 es adecuado para la mayoría de bacterias Gram-negativas, el origen del plásmido 2m es adecuado para levaduras y varios orígenes virales (SV40, poliovirus, adenovirus, VSV o BPV) son de utilidad para los vectores de clonaje en células de mamífero. Por lo general, no es necesario el origen de replicación para los vectores de expresión en mamíferos (el origen SV40 se utiliza comúnmente sólo porque contiene el promotor temprano).

Muchos de los vectores de expresión son vectores "lanzadera", es decir, son capaces de replicarse en al menos una clase de organismo, pero pueden transfectarse en otro organismo para su expresión. Por ejemplo, un vector se clona en *E. coli* y luego el mismo vector se transfecta en levaduras o células de mamífero para la expresión aún cuando no sea capaz de replicarse independientemente del cromosoma de la célula huésped.

El DNA puede amplificarse mediante inserción en el genoma huésped. Ello se logra utilizando especies de *Bacillus* como huésped, por ejemplo, incluyendo en el vector una secuencia de DNA complementaria a una secuencia hallada en el DNA genómico de *Bacillus*. La transfección de *Bacillus* con este vector da como resultado la recombinación homóloga con el genoma y la inserción del DNA de HRG. Sin embargo, la recuperación del DNA genómico que codifica HRG es más compleja que la de un vector replicado exógenamente, porque se requiere la digestión con una enzima de restricción para excisar el DNA de HRG. El DNA puede amplificarse por PCR y transfectarse directamente en las células huésped sin ningún componente de replicación.

#### (III) Componente de selección génica

Los vectores de clonaje y expresión deben contener un gen de selección, también denominado un marcador seleccionable. Este gen codifica una proteína necesaria para la supervivencia o el crecimiento de células transformadas en un medio de cultivo selectivo, por lo que las células huésped no transformadas con el vector que contiene el gen de selección no sobreviven en el medio de cultivo. Genes de selección típicos son aquellos que codifican proteínas que

(a) confieren resistencia a los antibióticos o a otras toxinas, p.ej., la ampicilina, la neomicina, el metotrexato, o la tetraciclina, (b) complementan deficiencias auxotróficas, o (c) suministran nutrientes críticos no disponibles del medio, p.ej., el gen que codifica la racemasa D-alanina de *Bacilli*.

- 5 Un ejemplo de selección es la utilización de un fármaco para parar el crecimiento de una célula huésped. Aquellas células que se transformen satisfactoriamente con un gen heterólogo expresarán una proteína que les conferirá resistencia al fármaco y por ello sobrevivirán al régimen de selección. Ejemplos de dicha selección dominante incluyen la utilización de neomicina (Southern y col., J. Molec. Appl. Genet. 1:327-1982), el ácido micofenólico (Mulligan y col., Science 209:1422, 1980) o la higromicina (Sugden y col., Mol. Cell. Biol. 5:410-413, 1985). Los tres ejemplos dados  
10 anteriormente utilizan genes bacterianos bajo el control eucariota para proporcionar la resistencia al fármaco G418 o neomicina (genética), xprt (ácido micofenólico), o higromicina, respectivamente.

- Otro ejemplo de marcadores seleccionables adecuados para las células de mamífero son aquellos que permiten la identificación de células competentes para incorporar el ácido nucleico de HRG, como la dihidrofolato reductasa  
15 (DHFR) o la timidina quinasa. Los transformantes de células de mamífero se someten a una presión de selección donde únicamente los transformantes están adaptados para sobrevivir gracias a que contienen el gen marcador. La presión de selección se impone cultivando los transformantes en condiciones determinadas, en las que se cambia sucesivamente la concentración del agente de selección en el medio y por ello se induce la amplificación del gen de selección y del DNA que codifica la HRG. La amplificación es el proceso por el cual, genes con mayor demanda de producción de  
20 una proteína crítica para el crecimiento, se repiten en tándem en los cromosomas de generaciones sucesivas de células recombinantes. En consecuencia, se sintetizan cantidades crecientes de HRG a partir del DNA amplificado.

- Por ejemplo, las células transformadas con el gen de selección DHFR se identifican en primer lugar cultivando todos los transformantes en un medio de cultivo que contiene metotrexato (Mtx), un antagonista competitivo del  
25 DHFR. Cuando se utiliza el DHFR de tipo salvaje, una célula huésped adecuada es la línea celular de ovario de hámster Chino (CHO) deficiente en actividad DHFR, preparada y crecida en cultivo tal como se describe por Urlaub y Chasin, Proc. Natl. Acad. Sci. USA, 77:4216, 1980. A continuación, las células transformadas se exponen a niveles crecientes de metotrexato, lo que conduce a la síntesis de copias múltiples del gen DHFR, y, consecuentemente, copias múltiples del otro componente de DNA que contiene los vectores de expresión, como el DNA que codifica la  
30 HRG. Esta técnica de amplificación puede utilizarse con cualquier huésped adecuado, p.ej., ATCC N° CCL61 CHO-K1, independientemente de la presencia de DHFR endógeno, si se utiliza, por ejemplo, un gen mutante DHFR que sea altamente resistente al MTx (patente europea 117.060). Alternativamente, las células huésped (en particular los huéspedes de tipo salvaje que contienen DHFR endógeno) transformadas o co-transformadas con las secuencias de DNA que codifican HRG, la proteína DHFR de tipo salvaje, y otro marcador seleccionable como la aminoglicosido  
35 3'-fosfotransferasa (APH) puede seleccionarse por crecimiento celular en medio que contenga un agente de selección para el marcador seleccionable como un antibiótico aminoglicósido, p.ej., kanamicina, neomicina, o G418 (véase la patente americana U.S. 4.965.199).

- Un gen de selección adecuado para utilizar con levaduras es el gen *trp1* presente en el plásmido de levaduras  
40 YRp7 (Stinchcomb y col., Nature, 282: 39, 1979; Kingsman y col., Gene, 7:141, 1979; o Tschemper y col., Gene, 10:157 (1980)). El gen *trp1* proporciona un marcador seleccionable para una cepa mutante de levaduras carentes de la capacidad para crecer en triptófano, por ejemplo, ATCC N° 44076 o PEP4-1 (Jones, Genetics, 85:12, 1977). La presencia de la lesión *trp1* en el genoma de la célula huésped de levaduras proporciona un entorno eficaz para detectar la transformación por el crecimiento en ausencia de triptófano. De modo similar, las cepas de levadura Leu2-deficientes  
45 (ATCC 20.622 o 38.626) se complementan con plásmidos conocidos que contienen el gen *Leu2*.

#### (IV) Componente promotor

- Los vectores de expresión y clonaje contienen generalmente un promotor que es reconocido por los organismos  
50 huésped y está unido operativamente al ácido nucleico de HRG. Los promotores son secuencias no traducidas localizadas cadena arriba (5') del codón de inicio de un gen estructural (generalmente en unas 100 a 1000 pb) que controlan la transcripción y la traducción de una secuencia de ácido nucleico determinada, como la HRG a la que está unida operativamente. Dichos promotores se agrupan generalmente en dos clases, los inducibles y los constitutivos. Los promotores inducibles son promotores que inician niveles elevados de transcripción a partir del DNA que está bajo  
55 su control en respuesta a algún cambio en las condiciones de cultivo, p.ej., la presencia o ausencia de un nutriente o un cambio en la temperatura. En la actualidad, se conoce un gran número de promotores que son reconocidos por diversas células huésped potenciales bien conocidas. Estos promotores están unidos operativamente al DNA que codifica la HRG tras haber eliminado el promotor de la fuente de DNA mediante digestión con enzimas de restricción e insertado la secuencia promotora aislada en el vector. Para la amplificación y/o expresión directa del DNA de HRG,  
60 se pueden utilizar tanto la secuencia promotora de la HRG nativa como cualquiera de los muchos promotores heterólogos existentes. Sin embargo, son preferibles los promotores heterólogos, ya que permiten una transcripción mayor y rendimientos superiores de la HRG expresada en comparación con el promotor de la HRG nativa.

- Los promotores adecuados para utilizar con huéspedes procariotas incluyen los sistemas de promotores de la  $\beta$ -lactamasa y la lactosa (Chang y col., Nautre, 275:615, 1978; Goeddel y col., Nature 281:544, 1979), la fosfatasa alcalina, un sistema de promotor del triptófano (*trp*) (Goeddel, Nucleic Acids Res., 8:4057, 1980 y la patente europea EP 36.776), tPA (patente americana U.S. 5.641.655) y los promotores híbridos como el promotor *tac* (deBoer y col.,

Proc. Natl. Acad. Sci. USA 80:21-25, 1983). Sin embargo, también son adecuados otros promotores bacterianos conocidos. Sus secuencias nucleotídicas se han publicado, con lo que se facilita al experto en la materia su ligación con el DNA que codifica la HRG (Siebenlist y col., Cell 20:269, 1980) utilizando engarces o adaptadores para suministrar cualquier diana de restricción requerida. Los promotores para utilizar en sistemas bacterianos contienen generalmente una secuencia Shine-Dalgarno (S.D.) unida operativamente al DNA que codifica la HRG.

Las secuencias promotoras para utilizar con huéspedes de levaduras incluyen los promotores de la 3-fosfoglicerato quinasa (Hitzeman y col., J. Biol. Chem., 255:2073, 1980) o de otras enzimas glicolíticas (Hess y col., J. Adv. Enzyme Reg 7:149, 1968; y Holland, Biochemistry 17:4900, 1978), como la enolasa, la gliceraldehído-3-fosfato deshidrogenasa, la hexoquinasa, la piruvato descarboxilasa, la fosfofructoquinasa, la glucosa-6-fosfato isomerasa, la 3-fosfoglicerato mutasa, la piruvato quinasa, la triosafofosfato isomerasa, la fosfoglucosa isomerasa y la glucoquinasa.

Otros promotores de levaduras, como los promotores inducibles que tienen la ventaja adicional de inducir una transcripción controlada por las condiciones de crecimiento, son las regiones promotoras de la alcohol deshidrogenasa 2, la isocitocromo C, la fosfatasa ácida, las enzimas degradadoras asociadas con el metabolismo del nitrógeno, la metalotioneína, la gliceraldehído-3-fosfato deshidrogenasa y las enzimas responsables de la utilización de la maltosa y la galactosa. Los vectores y promotores adecuados para utilizar en la expresión de levaduras se describen por Hitzeman y col., patente europea EP 73.657A. Las secuencias intensificadoras de la transcripción de levaduras también se utilizan de forma ventajosa con los promotores de levaduras.

Se conocen las secuencias promotoras de eucariotas. En general, todos los genes eucariotas tienen una región rica en AT localizada a aproximadamente 25-30 bases cadena arriba del sitio de inicio de la transcripción. Otra secuencia a 70 a 80 bases cadena arriba del codón de inicio de la transcripción de muchos genes es una región CXCAAT, en donde X puede ser cualquier nucleótido. En el extremo 3' de muchos genes eucariotas se halla una secuencia AATAAA que puede ser la señal para la adición de una cola poliA en el extremo 3' de la secuencia codificante. Todas estas secuencias se insertan de forma adecuada en los vectores de expresión de mamíferos.

La transcripción del gen HRG a partir de vectores en células huésped de mamífero se controla mediante promotores obtenidos de los genomas de virus como el virus del poliovirus, el virus de la viruela aviar (patente inglesa UK 2.211.504, publicada el 5 de julio de 1989), los adenovirus (como el adenovirus 2), el virus del papiloma, el virus del sarcoma aviar, el citomegalovirus, un retrovirus, el virus de la hepatitis B y más preferentemente el Simian Virus 40 (SV40), de promotores heterólogos de mamífero, p.ej., el promotor de la actina o un promotor de inmunoglobulina, de los promotores de la proteína de choque térmico y del promotor asociado normalmente con la secuencia HRG, siempre que los promotores sean compatibles con los sistemas de células huéspedes.

Los promotores temprano y tardío del virus SV40 se obtienen de forma conveniente como un fragmento de restricción que también contiene el origen de replicación del virus SV40 (Fiers y col., Nature 273:113 (1978); Mulligan y Berg, Science, 209:1422-1427 (1980); Pavlakakis y col., Proc. Natl. Acad. Sci. USA, 78:7398-7402 (1981)). El promotor inmediatamente temprano del citomegalovirus humano se obtiene de forma adecuada como un fragmento de restricción HindIII (Greenaway y col., Gene, 18:355-360 (1982)). Un sistema para la expresión del DNA en huéspedes mamíferos utilizando el virus del papiloma bovino como un vector se describe en la patente americana U.S. 4.419.446. Una modificación de este sistema se describe en la patente americana U.S. 4.601.978. Véase también Gray y col., Nature, 295:503-508 (1982) sobre la expresión del cDNA que codifica el interferón inmune en células de mono; Reyes y col., Nature, 297:598-601 (1982) sobre la expresión del cDNA del interferón- $\beta$  humano en células de ratón bajo el control de un promotor de timidina quinasa del virus herpes simplex; Canaani y Berg, Proc. Natl. Acad. Sci. USA, 79:5166-5170 (1982) sobre la expresión del gen del interferón- $\beta$ 1 humano en células de ratón y conejo en cultivo; y Gorman y col., Proc. Natl. Acad. Sci. USA, 79:6777-6781 (1982) sobre la expresión de secuencias CAT bacterianas en células de riñón de mono CV-1, fibroblastos embrionales de pollo, células de ovario de hámster chino, células HeLa y células de ratón NIH-3T3 utilizando como promotor la secuencia de repetición terminal larga del virus del sarcoma de Rous.

#### (v) *Componente del elemento intensificador de la transcripción*

La transcripción de un DNA que codifica la HRG de la presente invención por eucariotas superiores se aumenta a menudo insertando una secuencia intensificadora de la transcripción en el vector. Estas secuencias son elementos de actuación en cis del DNA, generalmente a aproximadamente 10-300 pb, que actúan sobre un promotor para incrementar su transcripción. Las secuencias intensificadoras de la transcripción tienen una posición y orientación relativamente independiente y se han hallado en 5' (Laimins y col., Proc. Natl. Acad. Sci. USA, 78:993, 1981) y 3' (Lusky y col., Mol. Cell Biol., 3:1108, 1983) de la unidad de transcripción, dentro de un intrón, así como dentro de la propia secuencia codificante (Osborne y col., Mol. Cell. Bio., 4:1293, 1984). En la actualidad, se conocen muchas secuencias intensificadoras de la transcripción de genes de mamíferos (globina, elastasa, albúmina,  $\alpha$ -fetoproteína e insulina). De forma característica, sin embargo, se utilizará una secuencia intensificadora de la transcripción de un virus de células eucariotas. Los ejemplos incluyen la secuencia intensificadora de la transcripción de SV40 en el lado tardío del origen de replicación (pb 100-270), la secuencia intensificadora de la transcripción del promotor temprano del citomegalovirus, la secuencia intensificadora de la transcripción del poliovirus en el lado tardío del origen de replicación y las secuencias intensificadoras de la transcripción de adenovirus (véase también Yaniv, Nature, 297:17-18 (1982) sobre elementos intensificadores de la activación de promotores eucariotas). La secuencia intensificadora de la transcripción

se puede ayustar en el vector en una posición 5' o 3' al DNA de HRG, pero se localiza preferentemente en 5' del promotor.

(vi) *Componente de finalización de la transcripción*

Los vectores de expresión utilizados en las células huésped eucariotas (levaduras, hongos, insectos, plantas, animales, humanos, o células nucleadas de otros organismos multicelulares) contendrán también secuencias necesarias para la finalización de la transcripción y para la estabilización del mRNA. Dichas secuencias están disponibles comúnmente de las regiones no traducidas 5', y ocasionalmente 3', de los DNAs o cDNAs eucariotas o virales. Estas regiones contienen segmentos de nucleótidos transcritos como fragmentos poliadenilados en la porción no traducida del mRNA que codifica la HRG. Las regiones 3' no traducidas también incluyen los sitios de finalización de la transcripción.

Para la construcción de vectores adecuados que contengan uno o más de los anteriores componentes enunciados, las secuencias codificantes y control deseadas, se utilizan técnicas de ligación estándar. Los plásmidos aislados, o los fragmentos de DNA se cortan, se recortan sus extremos y se religan en la forma deseada para generar los plásmidos requeridos.

En el análisis para confirmar que las secuencias de los plásmidos construidos son correctas, se utilizan mezclas de ligación para transformar la cepa de *E. coli* K12 294 (ATCC 31.446) y los transformantes con éxito se seleccionan por su resistencia a ampicilina o tetraciclina, según sea el caso. A continuación, se amplifican los plásmidos de los transformantes, se analizan mediante digestión con endonucleasas de restricción y/o se secuencian por el procedimiento de Messing y col., *Nucleic Acids Res.* 9:309 (1981) o por el procedimiento de Maxam y Col., *Methods in Enzymology* 65:499 (1980).

En la práctica de la presente invención, son de especial interés los vectores de expresión que proporcionan la expresión transitoria del DNA que codifica la HRG en las células de mamífero. En general, la expresión transitoria implica la utilización de un vector de expresión que sea capaz de replicarse de forma eficaz, de modo que la célula huésped acumule muchas copias del vector de expresión y, a su vez, sintetice niveles elevados de un polipéptido deseado codificado por el vector de expresión. Los sistemas de expresión transitorios, comprenden un vector de expresión y una célula huésped adecuada, que permiten la identificación positiva conveniente de los polipéptidos codificados por el DNA clonado, así como el rastreo rápido de dichos polipéptidos por sus propiedades biológicas o fisiológicas deseadas. Por ello, los sistemas de expresión transitoria son de especial interés en la presente invención con el propósito de identificar análogos y variantes de HRG que tengan una actividad similar a HRG. Dicho sistema de expresión transitoria se describe en la patente europea americana U.S. 5.024.939.

Otros procedimientos, vectores y células huésped adecuadas para la síntesis de HRG en el cultivo de células de vertebrados se describen en Gething y col., *Nature* 293:620-625; Mantel y col., *Nature*, 281:40-46, 1979; Levinson y col., patentes europeas EP 117.060 y EP 117.058. Un plásmido de expresión de utilidad para la expresión en el cultivo de células de mamífero de HRG es el pRK5 (patente europea EP 307.247).

D. *Selección y transformación de células huésped*

Las células huésped adecuadas para el clonaje o la expresión de los vectores de la presente invención son las procariotas, levaduras o eucariotas superiores, tal como se describió anteriormente. Las procariotas adecuadas incluyen las eubacterias, como los organismos Gram-negativos o los Gram-positivos, por ejemplo, *E. coli*, *Bacilli* como *B. subtilis*, *Pseudomonas* especies como *P. aeruginosa*, *Salmonella typhimurium* o *Serratia marcescans*. Un huésped de clonaje preferido de *E. coli* es *E. coli* 294 (ATCC 31.446), aunque también son adecuadas otras cepas como *E. coli* B, *E. coli* 1776 (ATCC 31.537) y *E. coli* W3110 (ATCC 27.325). Estos ejemplos son ilustrativos en lugar de limitantes. Preferentemente, la célula huésped debería secretar cantidades mínimas de enzimas proteolíticas. Alternativamente, son adecuados los procedimientos *in vitro* de clonaje, p.ej., la PCR u otras reacciones de la polimerasa de ácidos nucleicos.

Además de los procariotas, los microbios eucariotas como los hongos filamentosos o las levaduras son huéspedes adecuados para los vectores que codifican la HRG. *Saccharomyces cerevisiae*, o la levadura del pan, es el más utilizado de entre los huéspedes eucariotas inferiores. Sin embargo, otros géneros, especies y cepas también están generalmente disponibles y son de utilidad en la presente invención, como *Schizosaccharomyces pombe* (Beach u Nurse, 249:140 (1981); patente europea EP 139.383, publicada el 2 de mayo de 1985), huéspedes de *Kluyveromyces* (patente americana U.S.S.N. 4.943.529) como por ejemplo *K. lactis* (Louvencourt y col., *J. Bacteriol.* 737 (1983); *K. fragilis*, *K. bulgaricus*, *K. thermotolerans* y *K. marxianus*, *Yarrowia* (patente europea EP 402.226); *Pichia pastoris* (patente europea EP 183.070), Streekrishna y col., *J. Basic Microbiol.* 28:265-278 (1988); *Candida*, *Trichoderma reesia* (patente europea EP 244.234); *Neurospora crassa* (Case y col., *Proc. Natl. Acad. Sci. USA*, 76:5259-5263 (1979) y hongos filamentosos como *Neurospora*, *Penicillium*, *Tolypoctadium* (patente internacional WO 91/00357, publicada el 10 de enero de 1991) y huéspedes de *Aspergillus* como *A. nidulans* (Ballance y col., *Biochem. Biophys. Res. Commun.*, 112:284-289 (1983); Tilburn y col., *Gene* 26:205-221 (1983); Yelton y col., *Proc. Natl. Acad. Sci. USA*, 81:1470-1474 (1984) y *A. niger* (Kelly y Hynes, *EMBO J.*, 4:475-479 (1985)).

Las células huésped adecuadas para la expresión del polipéptido HRG glicosilado provienen de organismos multicelulares. Dichas células huésped son capaces de realizar un procesamiento complejo y de presentar actividades de glicosilación. En principio, cualquier eucariota superior es adecuada para trabajar, bien sea en cultivo de vertebrados o invertebrados. Ejemplos de células de invertebrados incluyen las células de plantas e insectos. Se han identificado numerosas cepas y variantes de baculovirus y las correspondientes células huésped de insectos permisivas como *Spodoptera frugiperda* (oruga), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (mosca de la fruta) y *Bombyx mori* (ver por ejemplo, Luckow y col., Biol Technology, 6:47-55 (1988); Miller y col., in Genetic Engineering, Setlow, JK y col., eds., Vol 8 (Plenum Publishing, 1986), pág. 277-279; y Maeda y col., Nature, 315:592-594 (1985)). Varias de estas cepas víricas son de disponibilidad pública, p.ej., la variante L-1 de *Autographa californica* NPV y la cepa Bm-5 de *Bombix mori* NPV, pudiéndose utilizar dichos virus tal como se describe en la presente invención, en particular para la transfección de células de *Spodoptera frugiperda*. También pueden utilizarse como huéspedes, los cultivos de plantas del algodón, el maíz, la patata, la soja, la petunia, el tomate y el tabaco. En general, las células de plantas se transfectan mediante incubación con ciertas cepas de la bacteria *Agrobacterium tumefaciens*, previamente manipulada para contener el DNA de HRG. Durante la incubación del cultivo de células vegetales con *A. tumefaciens*, el DNA que codifica HRG se transfiere a la célula vegetal, de modo que ésta se transfecta y en condiciones adecuadas expresará el DNA de HRG. Además, también están disponibles secuencias reguladoras y secuencias señal compatibles con las células vegetales, como el promotor y las secuencias de señalización de la sintasa nopalina (Depicker y col., J. Mol. Appl. Gen., 1:561 (1982)). Además, los segmentos del DNA aislados de la región cadena arriba del gen T-DNA 780 son capaces de activar, o aumentar los niveles de transcripción de los genes expresables en plantas en el tejido vegetal que contiene el DNA recombinante (véase la patente europea EP 321.196, publicada el 21 de junio de 1989).

Sin embargo, se ha prestado un interés mucho mayor a las células de vertebrados, con lo que la propagación de células de vertebrados en cultivo (cultivo de tejidos) se ha convertido en un procedimiento de rutina en los últimos años (Tissue Culture, Academic Press, Kruse y Patterson, ed. (1973)). Ejemplos de líneas de células huésped de mamíferos son la línea CV1 de riñón de mono transformada con SV40 (COS-7, ATCC CRL 1651); la línea de riñón embrionario humano (células 293 o 293 subclonadas para el crecimiento en un cultivo en suspensión, Graham y col., J. Gen Virol., 36:59, 1977); células de riñón de bebé hámster (BHK, ATCC CCL 10); células de ovario de hámster chino/DHFR (CHO, Urlaub y Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); células de sertoli de ratón (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); células de riñón de mono (CV1 ATCC CCL 70); células de riñón de mono verde africano (VERO 76, ATCC CRL-1587); células de carcinoma cervical (HELA, ATCC CCL2); células de riñón canino (MDCK; ATCC CCL 34); células de hígado de rata búfalo (BRL 3A, ATCC CRL 1442); células de pulmón humanas (W138, ATCC CCL75); células hepáticas humanas (HepG2, HB 8065); células de tumor mamario de ratón (MMT 060562, ATCC CCL51); células TRI (Mather y col., Annals N.Y. Acad. Sci. 383:44-68 (1982)); células MRC 5; células FS4; y una línea celular de hepatoma humano (Hep G2). Las células huésped preferidas son las embrionarias humanas 293 y las células de ovario de hámster chino.

Las células huésped se transfectan y transforman preferentemente con los vectores de expresión o clonaje de la presente invención descritos anteriormente y se cultivan en medios con nutrientes convencionales modificados de forma adecuada para la inducción de promotores, selección de transformantes o la amplificación de los genes codificantes de las secuencias deseadas.

La transfección hace referencia a la incorporación de un vector de expresión por una célula huésped independientemente de que se exprese o no cualquier secuencia codificante. Se conocen numerosos procedimientos de transfección en la materia, por ejemplo, el procedimiento del  $\text{CaPO}_4$  y la electroporación. En general, se reconoce que la transfección ha sido satisfactoria cuando se tiene cualquier indicación de que el vector es operativo dentro de la célula huésped.

La transformación significa la introducción del DNA en un organismo para que dicho DNA se replice, bien como elemento extracromosómico o mediante integración cromosómica. Dependiendo de la célula huésped utilizada, la transformación se realiza utilizando técnicas estándares adecuadas a dichas células. El tratamiento del calcio utiliza el cloruro cálcico, tal como se describe en la sección 1.82 de Sambrook y col., *supra*, se utiliza generalmente para procariotas u otras células que contienen barreras esenciales de pared celular. La infección con *Agrobacterium tumefaciens* se utiliza para la transformación de ciertas células vegetales, tal como se describió por Shaw y col., Gene, 23:315 (1983) y la patente internacional WO 89/0589, publicada el 29 de junio de 1989. Para las células de mamífero sin dichas paredes celulares, es preferible el procedimiento del fosfato cálcico descrito en las secciones 16.30-16.37 de Sambrook y col., *supra*. Los aspectos generales de las transformaciones de sistemas de células huésped de mamíferos se han descrito en Axel en la patente americana U.S. 4.399.216, publicada el 16 de agosto de 1983. Las transformaciones en levaduras se llevan a cabo por lo general de acuerdo con el procedimiento de Van Sollingen y col., J. Bact., 130:946 (1977) y Hsiao y col., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). Sin embargo, también se pueden utilizar otros procedimientos para la introducción de DNA en las células como la inyección nuclear, la electroporación o la fusión de protoplastos.

#### E. Cultivo de células huésped

Las células procariotas utilizadas para producir el polipéptido HRG de la presente invención se cultivan en un medio adecuado tal como se describe en general por Sambrook y col., *supra*.

Las células huésped de mamíferos utilizadas para producir la HRG de esta invención pueden cultivarse en muy diversos medios. Son adecuados para el cultivo de células huésped los medios disponibles comercialmente como el Ham F10 (Sigma), el Minimal Essential Medium (MEM, Sigma), el JRPMI-1640 (Sigma) y el Medio Eagle Modificado por Dulbecco (DMEM, Sigma). Además, cualquier medio de los descritos por Ham y Wallace, Meth. Enz. 58:44 (1979), Barnes y Sato, Anal. Biochem. 102:255 (1980), las patentes americanas U.S. 4.767.704; 4.657.866; 4.927.762; o 4.560.655; las patentes internacionales WO 90/03430; WO 87/00195 y la patente americana U.S. Re. 30.085 o la U.S. 5.122.469 pueden utilizarse como medio de cultivo para las células huésped. Cualquiera de estos medios puede suplementarse si fuera necesario con hormonas y/o otros factores de crecimiento (como la insulina, la transferrina o el factor de crecimiento epitelial), sales (como el cloruro sódico, el calcio, el magnesio y el fosfato), tampones (como el HEPES), nucleósidos (como la adenosina y la timidina), antibióticos (como el fármaco Gentamicina<sup>TM</sup>), oligoelementos (definidos como compuestos orgánicos generalmente presentes a concentraciones finales del orden micromolar) y glucosa o cualquier otra fuente de energía equivalente. También se puede incluir cualquier otro suplemento a concentraciones adecuadas, que será conocido por los expertos en la materia. Las condiciones de cultivo, como la temperatura, pH y otras, son las previamente utilizadas por las células huésped seleccionadas para la expresión y serán evidentes a un experto en la materia.

Las células huésped referidas en esta descripción abarcan las células en un cultivo *in vitro*, así como las células que se hallan dentro de un animal huésped.

Además, se prevé que la HRG de esta invención pueda producirse mediante recombinación homóloga, o mediante procedimientos de producción recombinante utilizando elementos control introducidos en las células que contienen el DNA que codifica la HRG y actualmente en utilización. Por ejemplo, se inserta un elemento promotor/secuencia intensificadora de la transcripción potente, un supresor o un elemento modulador de la transcripción exógena en el genoma de la célula huésped deseada a proximidad y orientación suficientes para modificar la transcripción del DNA que codifica la HRG deseada. El elemento control no codifica la HRG de la presente invención, pero el DNA está presente en el genoma de la célula huésped. A continuación, se rastrean las células que generen la HRG de la presente invención, o presenten niveles aumentados o disminuidos de expresión, según se desee.

#### F. Detección de la amplificación/expresión génica

Se puede cuantificar directamente la amplificación y/o expresión génicas en una muestra, por ejemplo, mediante transferencia de Southern convencional, transferencia de Northern para cuantificar la transcripción del mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)), transferencia de mancha (análisis de DNA) o hibridación *in situ*, utilizando una sonda marcada adecuadamente basada en las secuencias proporcionadas aquí. Se pueden utilizar diversos marcajes, más comúnmente los radioisótopos, en particular el P<sup>32</sup>. Sin embargo, también se pueden utilizar otras técnicas, como la utilización de nucleótidos modificados con biotina para su introducción en un polinucleótido. La biotina se utiliza a continuación como el sitio de unión de la avidina, o de los anticuerpos que pueden marcarse con una gran variedad de marcajes, como los radionúclidos, los fluorescentes, las enzimas o similares. Alternativamente, se pueden utilizar anticuerpos que reconozcan dúplex específicos, incluyendo dúplex de DNA, dúplex de RNA y dúplex híbridos de DNA-RNA o dúplex de DNA-proteína. Los anticuerpos a su vez se pueden marcar y el ensayo se puede llevar a cabo si el dúplex está unido a una superficie. De este modo, se puede detectar la presencia del anticuerpo unido al dúplex después de la formación del dúplex en la superficie.

Alternativamente, la expresión génica se puede cuantificar mediante procedimientos inmunológicos, como la tinción inmunohistoquímica de secciones de tejido y del ensayo del cultivo de células o de fluidos corporales, con el fin de cuantificar directamente la expresión del producto génico. Para las técnicas de tinción inmunohistoquímicas, la muestra celular se prepara normalmente mediante deshidratación y fijación, seguido por la reacción con anticuerpos específicos para el producto del gen acoplado si los marcajes son visualmente detectables como marcajes enzimáticos, marcajes fluorescentes, marcajes luminiscentes y similares. Una técnica de tinción particularmente sensible y adecuada para utilizar en la presente invención se describe por Hsu y col., Am. J. Clin. Path., 75:734-738 (1980).

Los anticuerpos útiles para la tinción inmunohistoquímica y/o el ensayo de muestras de fluidos pueden ser monoclonales o policlonales y pueden prepararse en cualquier mamífero. De forma adecuada, los anticuerpos pueden prepararse contra un polipéptido HRG o contra un péptido sintético basado en las secuencias del DNA proporcionadas aquí, tal como se describe más adelante.

#### G. Purificación de los polipéptidos de heregulina

La HRG se recupera de una fracción de membrana celular. Alternativamente, un fragmento de HRG soluble, cortado proteolíticamente o expresado de forma truncada se recupera del medio de cultivo como un polipéptido soluble. La HRG se recupera de los lisados de células huéspedes si se expresa directamente sin una secuencia señal.

Si la HRG se expresa en una célula recombinante distinta a la de origen humano, la HRG estará completamente libre de proteínas o polipéptidos de origen humano. Sin embargo, es deseable purificar la HRG de proteínas o polipéptidos celulares recombinantes para obtener preparaciones esencialmente homogéneas como la HRG. En una primera etapa, el medio de cultivo o lisado se centrifuga para eliminar los restos celulares y a continuación, se separan las fracciones de membrana y proteínas solubles. La HRG se purifica de la fracción de proteínas solubles (que requiere la

presencia de una proteasa) y de la fracción de membranas del lisado del cultivo, dependiendo de que la HRG esté unida o no a la membrana. Los procedimientos siguientes son ejemplos adecuados de procedimientos de purificación: fraccionamiento en columnas de inmunoafinidad o de intercambio iónico; precipitación con etanol; HPLC de fase inversa; cromatografía en sílice, heparina-Sepharose o una resina de intercambio catiónico como la DEAE; cromatografía; SDS-PAGE; precipitación con sulfato amónico; y filtración en gel utilizando, por ejemplo, Sephadex G-75.

Las variantes de HRG en donde los residuos se delecionan, insertan o sustituyen se recuperan de la misma manera que la HRG nativa, teniendo en cuenta cualquier cambio esencial en las propiedades ocasionadas por la variación. Por ejemplo, la preparación de una fusión HRG con otra proteína o polipéptido, p.ej., un antígeno viral o bacteriano, facilita la purificación, ya que puede utilizarse una columna de inmunoafinidad que contenga el anticuerpo contra el antígeno para adsorber la fusión. Se pueden utilizar columnas de inmunoafinidad, como una columna anti-HRG policlonal de conejo para adsorber la variante HRG, uniéndola por lo menos a uno de los epitopos inmunes restantes. Un inhibidor de proteasa como el fenilmetilsulfonilfluoruro (PMSF) también puede ser útil para inhibir la degradación proteolítica durante la purificación y pueden incluirse antibióticos para evitar el crecimiento de contaminantes fortuitos. Un experto en la materia apreciará si los procedimientos de purificación adecuados para la HRG nativa necesitan modificarse para ajustar los cambios en el carácter de las variantes HRG, o después de la expresión en un cultivo de células recombinantes.

#### H. Modificaciones covalentes de HRG

Las modificaciones covalentes de los polipéptidos de HRG se incluyen dentro del ámbito de la presente invención. Opcionalmente, se modifican covalentemente tanto la HRG nativa como sus variantes de secuencia aminoacídica. Un tipo de modificación covalente, incluida dentro del ámbito de esta invención, es un fragmento polipeptídico de HRG. Los fragmentos de HRG, como HRG-GDF, que tienen hasta 40 residuos aminoacídicos se preparan de forma conveniente mediante síntesis química, o mediante corte enzimático o químico del polipéptido HRG de longitud completa, o del polipéptido variante de HRG. En la molécula, se introducen otros tipos de modificaciones covalentes de HRG, o de sus fragmentos, haciendo reaccionar residuos aminoacídicos diana de HRG, o fragmentos de lo mismo con un agente de derivación orgánico capaz de reaccionar con las cadenas laterales seleccionadas de los residuos N- o C-terminal.

Los residuos cisteinil reaccionan más frecuentemente con los  $\alpha$ -haloacetatos (y las correspondientes aminas), como el ácido cloroacético o la cloroacetamida para producir derivados de carboximetil o carboximidometil. Los residuos cisteinil también se derivan mediante reacción con bromotrifluoroacetona, ácido  $\alpha$ -bromo- $\beta$ (5-imidozolil)propiónico, cloroacetil fosfato, N-alkilmaleimidias, 3-nitro-2-piridil disulfuro, metil 2-piridil disulfuro, p-cloromercuribenzoato, 2-cloromercuri-4-nitrofenol, o cloro-7-nitrobenzo-2-oxa-1,3-diazol.

Los residuos histidil se derivan mediante reacción con dietilpirocarbonato a pH 5,5-7,0, porque este agente es relativamente específico para la cadena lateral histidil. También es útil el para-bromofenacil bromuro; la reacción se realiza preferentemente en cacodilato sódico 0,1 M a pH 6,0.

Los residuos lisinil y amino terminales se hacen reaccionar con ácido succínico u otros anhídridos de ácido carboxílico. La derivación con estos agentes tiene el efecto de invertir la carga de los residuos lisinil. Otros agentes adecuados para la derivación de residuos con grupo  $\alpha$ -amino incluyen los imidoésteres como el metil picolimidato; el piridoxal fosfato; el piridoxal; el cloroborohidruro; el ácido trinitobencenosulfónico; la O-metilisourea; la 2,4-pentanediona; y la reacción de transaminasas catalizada con glioxilato.

Los residuos arginil se modifican mediante reacción con uno o varios reactivos convencionales, entre ellos se hallan el fenilglioxal, la 2,3-butanediona, la 1,2-ciclohexanedina y la nitrhidrina. La derivación de residuos arginina requiere que la reacción se realice en condiciones alcalinas debido al elevado pKa del grupo funcional de guanidina. Además, estos reactivos pueden reaccionar con los grupos de lisina así como con el grupo épsilon-amina de la arginina.

La modificación específica de residuos tirosil puede realizarse, introduciendo marcajes espectrales en los residuos tirosil mediante reacción con compuestos de diazonio aromático de tetranitrometano. Se utilizan más frecuentemente el N-acetilimidazol y el tetrametano para formar especies O-acetil tirosil y derivados 3-nitro, respectivamente. Los residuos tirosil se yodan con  $I^{125}$  o  $I^{131}$  para preparar proteínas marcadas para utilizar en el radioinmunoensayo, siendo adecuado el procedimiento de cloramina T descrito anteriormente.

Los grupos laterales carboxil (aspartil o glutamil) se modifican selectivamente mediante reacción con carbodii-midas ( $R'-N=C=N-R'$ ), en donde R y R' son diferentes grupos alquil, como el 1-ciclohexil-3-(2-morfolinil-4-etil) carbodiimida o el 1-etil-3-(4-azonia-4,4-dimetilfenil) carbodiimida. Además, los residuos aspartil y glutamil se convierten en asparaguinil y glutamil reaccionándolos con iones amonio.

La derivación con agentes bifuncionales es útil para la unión de HRG a una matriz soporte insoluble en agua o a una superficie a utilizar en un procedimiento de purificación de anticuerpos anti-HRG y vice versa. Los agentes de unión utilizados comúnmente incluyen, p.ej., la 1,1-bis(diazoacetil)-2-feniletano, el glutaraldehído, los ésteres de N-hidroxisuccinimida, por ejemplo, los ésteres con el ácido 4-azidosalicílico, los imidoésteres homobifuncionales, incluyendo ésteres de disuccinimidil como el 3,3'-ditiobis(succinimidilpropionato) y las maleimidias bifuncionales

como la bis-N-maleimida-1,8-octano. Los agentes de derivación como el metil-3-[(p-azidofenil)ditio]propioimidato producen intermediarios fotoactivables que son capaces de formar uniones en presencia de la luz. Alternativamente, las matrices reactivas insolubles en agua como los carbohidratos activados con bromuro de cianógeno y los sustratos reactivos descritos en las patentes americanas U.S. 3.969.287; 3.691.016; 4.195.128; 4.247.642; 4.229.537 y 4.330.440 se utilizan para la inmovilización de proteínas.

Los residuos glutaminil y asparraguinil se desaminan frecuentemente en los residuos correspondientes glutamil y aspartil. Alternativamente, estos residuos se desaminan en condiciones ligeramente ácidas. Cualquier forma de estos residuos se halla dentro del ámbito de la presente invención.

Otras modificaciones incluyen la hidroxilación de prolina y lisina, la fosforilación de grupos hidroxilos de residuos seril o treonil, la metilación de grupos  $\alpha$ -amino de las cadenas laterales de lisina, arginina e histidina (T.E. Creighton, *Proteins: Structure and Molecular Properties*, WH. Freeman & Co., San Francisco, pag. 79-86 (1983), la acetilación del grupo amino N-terminal y la amidación del grupo carboxilo C-terminal.

La HRG se fusiona opcionalmente con un polipéptido heterólogo a HRG. El polipéptido heterólogo es opcionalmente una secuencia de anclaje como la hallada en el factor de aceleración de la desintegración (DAF); una proteína de la cubierta de un fago, por ejemplo las proteínas de los genes III u VIII de M13. Estos polipéptidos heterólogos están unidos covalentemente a HRG a través de cadenas laterales o por los residuos terminales.

La HRG también se modifica covalentemente alterando su patrón nativo de glicosilación. En estas realizaciones, uno o más sustituyentes de carbohidratos se modifican añadiendo, eliminando o variando los componentes monosacáridos en un sitio determinado, o mediante la modificación de residuos en la HRG, de modo que se añaden o delecionan dichos sitios de glicosilación.

La glicosilación de polipéptidos, típicamente, está ligada a N- u O-. Ligada a N-, o N-glicosilación, hace referencia a la unión de la porción de carbohidratos a la cadena lateral de un residuo de asparragina. Las secuencias tri-peptídicas asparragina-X-serina y asparragina-X-treonina, en donde X es cualquier aminoácido a excepción de la prolina, son las secuencias de reconocimiento para la unión enzimática de la porción de carbohidrato a la cadena lateral de la asparragina. Así, la presencia de cualquiera de estas secuencias tripeptídicas en un polipéptido crea un sitio potencial de glicosilación. La O-glicosilación hace referencia a la unión de uno de los azúcares N-acetilgalactosamina, galactosa, o xilosa, a un ácido hidroxiamino, más comúnmente la serina o la treonina, aunque también puede utilizarse la 5-hidroxiprolina o la 5-hidroxilisina.

Los sitios de glicosilación se añaden a la HRG mediante alteración de su secuencia aminoacídica, a fin de contener una o más de las secuencias tripeptídicas anteriormente descritas (para sitios de N-glicosilación). La alteración también se puede realizar mediante adición de, o sustitución por, uno o más residuos serina o treonina a la HRG (para sitios de O-glicosilación). Para mayor facilidad, la HRG se modifica preferentemente mediante cambios a nivel del DNA, en particular mutando el DNA que codifica para la HRG en las bases preseleccionadas y generando así los codones que se traducirán en los aminoácidos deseados.

El acoplamiento químico o enzimático de los glicósidos a la HRG incrementa el número de sustituyentes de carbohidratos. Estos procedimientos son ventajosos ya que no requieren la producción del polipéptido en una célula huésped capaz de realizar N- u O-glicosilación. Dependiendo del modo de acoplamiento utilizado, el azúcar (o azúcares) puede unirse a (a) una arginina e histidina, (b) grupos carboxilo libres, (c) grupos sulfidrilos libres como los de la cisteína, (d) grupos hidroxilo libres como los de la serina, treonina, o hidroxiprolina, (e) residuos aromáticos como los de la fenilalanina, la tirosina, o el triptófano, o (f) el grupo amida de la glutamina. Estos procedimientos se describen en la patente internacional WO 87/05330, publicada el 11 de setiembre y en Aplin y Wriston (CRC Crit. Rev. Biochem., pag. 259-306 (1981)).

Las porciones de carbohidrato presentes en una HRG también se eliminan química o enzimáticamente. La deglicosilación química requiere la exposición del polipéptido al compuesto de ácido trifluorometansulfónico, o a un compuesto equivalente. Este tratamiento da como resultado el corte de todos o la mayoría de azúcares, excepto del azúcar de unión (N-acetilglucosamina o N-acetilgalactosamina), dejando el polipéptido intacto. La deglicosilación química se describe por Hakimuddin y col. (Arch. Biochem. Biophys., 259:52 (1987)) y por Edge y col., (Anal. Biochem., 118:131 (1981)). Las porciones de carbohidrato se eliminan de la HRG mediante diversas endo- y exoglicosidasas, tal como se describe por Thotakura y col., (Meth. Enzymol., 138:350 (1987)).

La glicosilación también se suprime por tunicamicina, tal como se describe por Duskin y col., (J. Biol. Chem., 257:3105 (1982)). La tunicamicina bloquea la formación de las uniones proteína-N-glicósido.

La HRG también se modifica mediante unión de la HRG a diversos polímeros no proteicos, p.ej., el polietilén glicol, el propilén glicol o los polioxialquilenos, de la forma descrita en las patentes americanas U.S. 4.640.835; 4.496.689; 4.301.144; 4.670.417; 4.791.192 o 4.179.337.

Un modo preferido para aumentar la vida media circulante *in vivo* de la HRG no unida a membrana es conjugarla a un polímero que le confiera una vida media prolongada, como el polietilén glicol (PEG). (Maxfield y col., Polymer

16, 505-509 (1975); Bailey, F.E. y col., en *Nonionic Surfactants* (Schick, MJ, ed.) pag. 794-821 (1967); Abuchowski A. y col., *J. Biol. Chem.*, 252:3582-3586 (1977); Abuchowski A. y col., *Cancer Biochem Biophys.* 7:175-186 (1984); Katre, N.V. y col., *Proc. Natl. Acad. Sci.* 84, 1487-1491 (1987); Goodoon R y col., *Bio Technology*, 8:343-346, (1990). También se ha descrito que la conjugación con PEG reduce la inmunogenicidad y la toxicidad (Abuchowski A. y col., *J. Biol. Chem.*, 252:3582-3586 (1977)).

La HRG también se inmoviliza en microcápsulas preparadas, por ejemplo, mediante técnicas de coacervación o polimerización interfacial (por ejemplo, microcápsulas de hidroximetilcelulosa o de gelatina y microcápsulas de poli-[metilmetacrilato], respectivamente), en sistemas de liberación del fármaco coloidal (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nanopartículas y nanocápsulas), o en macroemulsiones. Dichas técnicas se describen en Remington Pharmaceutical Sciences, 16ª ed., Oslo, A., Ed. (1980).

Los expertos en la materia serán capaces de rastrear variantes con el fin de seleccionar la variante óptima para el propósito deseado. Por ejemplo, un cambio en el carácter inmunológico de HRG, como un cambio en la afinidad por un antígeno dado o por el receptor HER2, se cuantifica mediante un inmunoensayo de tipo competitivo utilizando un estándar o control como p.ej., una HRG nativa (en particular la HRG-GFD nativa). También se ensayan, mediante procedimientos bien conocidos en la materia, otras modificaciones potenciales de las propiedades de la proteína o del polipéptido, como la estabilidad redox o térmica, la hidrofobicidad, la susceptibilidad a la degradación proteolítica, la estabilidad en el cultivo de células recombinantes, o la tendencia a agregarse con vehículos o en multímeros.

## 2. Composiciones terapéuticas, administración y utilización de Heregulinas

La HRG se utiliza en la presente invención para inducir el crecimiento epitelial, por ejemplo el crecimiento, la proliferación y diferenciación de células epiteliales de pulmón, y para aumentar la producción de proteína A tensoactiva por las células pulmonares. Estos efectos permiten el tratamiento con el medicamento preparado de acuerdo con la presente invención de enfermedades asociadas con la lesión tisular, por ejemplo, la enfermedad de obstrucción pulmonar crónica (EPOC), incluyendo los subtipos de lo mismo como la bronquitis crónica, el enfisema, el asma, etc., las enfermedades pulmonares neonatales incluyendo el síndrome de la dificultad respiratoria neonatal, el síndrome de aspiración del meconio, la enfermedad pulmonar crónica del neonato, la hernia diafragmática congénita, etc., las lesiones agudas del pulmón incluyendo la inhalación de humo o de sustancias químicas, la pneumonitis debida a aspiración, radiación, etc., el ahogo, la fibrosis quística y otras enfermedades traumáticas de las células epiteliales, incluyendo las lesiones asociadas con heridas quirúrgicas y las resecciones, las lesiones, las úlceras y rasgado de tejidos.

Una indicación preferida para el tratamiento con el medio de la invención es el tratamiento de EPOC. La EPOC es un espectro de enfermedades respiratorias inflamatorias crónicas caracterizadas por tos, esputo, disnea, limitación del flujo de aire e intercambio de gas deficiente. La EPOC es común en individuos de edad avanzada y presenta un patrón de disminución gradual de la función pulmonar. De forma característica, un paciente presentará una tos crónica con esputo claro que empeora a una tos con un esputo espeso acompañado por un intercambio de aire escaso. Estas condiciones conducen frecuentemente a enfermedades del corazón y a la muerte. Muchas personas con EPOC tendrán bronquitis crónica junto con enfisema. La presente invención es especialmente importante porque para, ralentiza y/o revierte el proceso de destrucción pulmonar en pacientes EPOC. En esta acción, el medicamento preparado de acuerdo con la invención es muy diferente de los tratamientos típicos para EPOC, en los que se tratan los síntomas pero no la destrucción subyacente del tejido de células pulmonares y su función.

El medio de la invención puede, sin embargo, combinarse con o administrarse junto con otras terapias para el tratamiento de enfermedades pulmonares como la EPOC. Por ejemplo, el medicamento preparado de acuerdo con la presente invención puede utilizarse junto con la administración de un broncodilatador anticolinérgico como el bromuro de ipratropio (ATROVENT, disponible por Boehringer Ingelheim) o el tiotropio, un agonista del receptor beta adrenérgico como el albuterol (PROVENT disponible por Scherig) o el salmeterol, esteroides como la prednisona, el ácido retinoico, los inhibidores de la fosfodiesterasa, los antagonistas del endotelio, los inhibidores de metaloproteinasas, los inhibidores de la elastasa, los inhibidores de radicales libres, los inhibidores de proteinasa serina, los inhibidores de la elastasa de neutrófilos, las composiciones de tensoactivos pulmonares como el beractant (SURVANTA, disponible por Ross LaBS.), PDGF, FGF, EGF, la hormona del crecimiento u otros factores de crecimiento proteicos, etc., o combinaciones de lo mismo. La cantidad relativa de la HRG y de los compuestos adicionales puede determinarse fácilmente por el clínico considerando los síntomas individuales del paciente. Se prevé que estas composiciones y las cantidades relativas de los componentes varíen en función de las necesidades de un paciente, por lo que se controlarán y ajustarán mediante análisis físicoquímicos y médicos convencionales de la función pulmonar.

Las formulaciones terapéuticas de la HRG se preparan para el almacenamiento mezclando la proteína HRG con el grado de pureza deseado y con vehículos, excipientes o estabilizantes opcionales fisiológicamente aceptables (Remington Pharmaceutical Sciences, *supra*), en la forma de una pasta liofilizada o soluciones acuosas. Los vehículos, excipientes o estabilizantes aceptables no son tóxicos para el receptor en las dosis y concentraciones utilizadas e incluyen tampones como el fosfato, el citrato y otros ácidos orgánicos; antioxidantes como el ácido ascórbico; polipéptidos de peso molecular bajo (menos de 10 residuos); proteínas como la albúmina sérica, la gelatina, o las inmunoglobulinas; los polímeros hidrofílicos como la polivinilpirrolidona; los aminoácidos como la glicina, la glutamina, la asparragina, la arginina o la lisina; los monosacáridos, disacáridos y otros carbohidratos incluyendo la glucosa, la manosa o las dextrinas; los agentes quelantes como el EDTA; los alcoholes de azúcares como el manitol o el sorbitol; las parejas

de iones formadores de sales; y/o los tensoactivos no-iónicos como el TWEEN, PLURONICS o el polietilén glicol (PEG).

La HRG a utilizar en la administración *in vivo* debe ser estéril. Ello se logra fácilmente mediante filtración a través de membranas de filtración estériles, antes o después de la liofilización y reconstitución. La HRG se guardará generalmente en la forma liofilizada o en solución.

Las composiciones terapéuticas de HRG se coloca generalmente en recipientes con un puerto de acceso estéril, por ejemplo, una bolsa de solución intravenosa o un vial con un tapón agujereable por una aguja de inyección hipodérmica.

La vía de administración de la HRG está de acuerdo con los procedimientos conocidos, p. ej., la inyección o la infusión por vía intravenosa, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial, administración de aerosol en polvo o líquido para la nariz o el pulmón o vías intralesionales, o mediante sistemas de liberación prolongada tal como se indica más adelante. El ligando HRG puede administrarse de forma continua mediante infusión o mediante inyección de bolo. Un anticuerpo agonista se administra preferentemente de la misma manera o mediante administración a la circulación sanguínea o linfa.

La HRG, la variante o fragmentos de HRG pueden secarse mediante pulverización, o liofilización por pulverización utilizando técnicas conocidas (Yeo y col., *Biotech. Y Bioeng.*, 41:341-346 (1993); Gombotz y col., PCT/US90/02421).

Ejemplos adecuados de preparaciones de liberación prolongada incluyen las matrices semipermeables de polímeros hidrofóbicos sólidos que contienen la proteína, y se presentan en formatos diversos, p.ej., películas, o microcápsulas. Ejemplos de matrices de liberación prolongada incluyen los poliésteres, los hidrogeles (p.ej., poli(2-hidroximetilmetacrilato) tal como se describió por Langer y col., *J. Biomed. Mater. Res.*, 15:167-277 (1981) y Langer, *Chem. Tech.*, 12:98-105 (1982) o el poli-vinilalcohol, los poliláctidos (patente americana U.S. 3.773.919; patente europea EP 58.481), los copolímeros del ácido L-glutámico y el gamma-etil-L-glutamato (Sidman y col., *Biopolymers*, 22:547-556 (1983)), el acetato de etilén-vinilo no degradable (Langer y col., *supra*), los copolímeros de ácido glicólico-ácido láctico degradables como el LUPRPN DEPOT (microesferas inyectables compuestas de copolímero ácido láctico-ácido glicólico y el acetato leuprolido), y el ácido poli-D-(-)-3-hidroxibutírico (patente europea EP 133.988). Mientras que los polímeros como el acetato de etilén-vinil y el ácido láctico-ácido glicólico permiten la liberación de moléculas durante más de 100 días, ciertos hidrogeles liberan proteínas durante períodos más cortos. Cuando las proteínas encapsuladas permanecen en el cuerpo durante un período de tiempo largo, se pueden desnaturalizar o agregar como resultado de la exposición a la humedad a 37°C, lo que produce una pérdida de la actividad biológica y cambios posibles en la inmunogenicidad. Se pueden idear estrategias racionales para la estabilización de proteínas dependiendo del mecanismo implicado. Por ejemplo, si el mecanismo de agregación es la formación de un enlace S-S intermolecular a través del intercambio de tio-disulfuro, se puede lograr la estabilización modificando los residuos sulfidrilos, liofilizando a partir de las soluciones ácidas, controlando el contenido de humedad, utilizando aditivos adecuados y desarrollando composiciones de matrices poliméricas específicas.

Las composiciones de HRG de liberación prolongada también incluyen la HRG en liposomas. Los liposomas que contienen la HRG se preparan mediante procedimientos conocidos *per se*: patente DE 3.218.121; Epstein y col., *Proc. Natl. Acad. Sci. USA*, 82:3688-3691 (1985); Hwang y col., *Proc. Natl. Acad. Sci. USA*, 77:4030-4034 (1980); patente europea EP 52.332; EP 36.676; EP 88.046; EP 143.494; EP 142.641; solicitud de patente japonesa 83-118008; patente americana U.S. 4.485.045 y 4.544.545; y patente europea EP 102.324. Generalmente, los liposomas son del tipo unilamelar pequeño (aproximadamente 200-800 Angstroms) en los que el contenido lipídico es superior a aproximadamente 30 moles. El % de colesterol, la proporción seleccionada se ajusta para una terapia óptima de HRG. Los liposomas con un tiempo de circulación aumentado se describen en la patente americana U.S. 5.013.556.

Una cantidad efectiva de HRG a utilizar terapéuticamente dependerá, por ejemplo, de los objetivos terapéuticos, la vía de administración y la condición del paciente. Según ello, será necesario para el clínico titular la dosificación y modificar la vía de administración con el fin de obtener un efecto terapéutico óptimo. Una dosificación diaria típica puede hallarse en el intervalo de aproximadamente 1 µg/Kg y hasta 100 mg/Kg o más, dependiendo de los factores mencionados anteriormente. En general, el clínico administrará la HRG hasta alcanzar una dosificación que logre el efecto deseado. El progreso de esta terapia se es fácilmente controlable mediante ensayos convencionales, por ejemplo, la producción de la proteína A tensoactiva.

En otra realización, se pueden obtener células epiteliales o aislarse de un tejido de mamífero para obtener una muestra de células epiteliales normales utilizando técnicas bien conocidas en la materia (biopsia, etc.). Esta muestra se trata a continuación con una proteína heregulina con el fin de inducir el crecimiento y/o la proliferación de células epiteliales en la muestra, con lo que aumenta la población de células epiteliales primarias. En general, la heregulina se añadirá al cultivo de células epiteliales *in vitro* a una concentración de aproximadamente 0,1 a 100 nM, preferentemente de 1-50 nM. Si se desea, las células epiteliales primarias se pueden cultivar *in vitro* durante varias generaciones con el fin de aumentar la población de células epiteliales. A continuación las células se cultivan en condiciones adecuadas para el cultivo de células de mamífero, tal como se describió anteriormente. Después de la expansión, la muestra crecida se reintroduce en el mamífero con el propósito de regenerar el epitelio en el tejido del mamífero. Por ejemplo, pueden obtenerse células epiteliales del pulmón de un paciente con un enfisema o una enfermedad pulmonar obstructiva, y a continuación, las células se pueden crecer y reintroducir en el pulmón con el fin de regenerar más rápidamente

el epitelio del tejido pulmonar dañado y en consecuencia restablecer la función pulmonar. Las células así crecidas pueden reintroducirse en el pulmón mediante aspiración o intubación utilizando procedimientos bien conocidos en la materia.

- 5 Los procedimientos y procesos aquí descritos en relación con la HRG- $\alpha$  o la HRG pueden aplicarse en general a otras HRG, como la HRG- $\beta$ 1, HRG- $\beta$ 2 y HRG- $\beta$ 3 y a variantes de la mismo, así como a los anticuerpos. Los ejemplos siguientes se ofrecen a título ilustrativo y no limitante.

## Ejemplos

### 10 Ejemplo 1

#### *Preparación de heregulinas*

- 15 (a) Se aislaron las heregulinas HRG- $\alpha$ , HRG- $\beta$ 1, HRG- $\beta$ 2, de tipo-HRG- $\beta$ 2 y HRG- $\beta$ 3, se clonaron, se expresaron y se aislaron del medio de cultivo tal como se describe en la patente americana U.S. 5.367.060.

(b) Los polipéptidos SMDF se prepararon tal como se describe en la patente internacional WO 96/15244.

- 20 (c) El polipéptido  $\gamma$ -HRG se preparó y caracterizó tal como se describe más adelante.

Reactivos: El dominio de tipo-EGF de la HRG $\beta$ 1<sub>(177-244)</sub> se expresó en *E. coli*, se purificó y radioyodinó tal como se describió anteriormente (Sliwkowski y col., J. Biol. Chem. 14661-14665 (1994)). Los anticuerpos monoclonales anti-HER2, 2C4 y 4D5 corresponden a los descritos anteriormente por Fendly y col., Cancer Research 50:1550-1558 (1990)).

30 Inmunoadhesinas-HER3 y -HER4: se diseñó un único sitio M1 en un plásmido de expresión que expresaba la cadena pesada de la IgG humana en la región codificante del dominio bisagra de la inmunoglobulina. Los sitios M1 también se diseñaron en un grupo de plásmidos de expresión de HER en la región codificante de las uniones ECD/TM de estos receptores. Toda la mutagénesis se realizó utilizando el procedimiento de Kunkel (Kunkel, T., Proc. Natl. Acad. Sci. USA. 82:488 (1985)). Los sitios M1 se utilizaron para generar las construcciones de fusión HER-IgG adecuadas. Dichas uniones de fusión de las diversas quimeras HER-IgG fueron: para HER, E<sup>646</sup><sub>HER2</sub>-(TR)-DKTH<sup>224</sup><sub>VH</sub>; para HER3, L<sup>636</sup><sub>HER3</sub>-(TR)-DKTH<sup>224</sup><sub>VH</sub>; para HER4, G<sup>640</sup><sub>HER4</sub>-(TR)-DKTH<sup>224</sup><sub>VH</sub>. La secuencia conservada TR se derivó del sitio M1. La expresión final de las construcciones se realizó en un plásmido de tipo pRK, en donde la expresión eucariota está dirigida por un promotor de CMV (Gorman y col., DNA Prot. Eng. Tehc. 2:3-10 (1990)).

40 Para obtener la proteína para los experimentos *in vitro*, se transfectaron las células HEK-293 con los plásmidos de expresión adecuados (Gorman y col., Huang y col., Nucleic Acids Res. 18:937-947 (1990)). El medio con suero se reemplazó con medio sin suero 15 horas después de la transfección y las células transfectadas se incubaron durante 5-7 días. El medio condicionado resultante se cosechó y se pasó por una columna de Proteína A (Pharmacia HiTrap<sup>TM</sup> de 1 ml). Las fusiones de IgG purificadas se eluyeron con ácido cítrico 0,1 M (pH 4,2) en tubos con Tris 1 M pH 9,0. Las proteínas eluidas se dializaron a continuación en PBS y se concentraron utilizando filtros Centri-prep-30 (Amicon). Se añadió glicerol a una concentración final del 25% y el material se almacenó a -20°C. Las concentraciones del material se determinaron mediante un ELISA-Fc.

45 Cultivo de células: Las líneas de cáncer de mama humano MDA-MB-175, MDA-MB-231, SK-BR-3 y MCF7 se obtuvieron del American Type Culture Collection y se mantuvieron en una mezcla de 50:50 de medio F12 HAM y Dulbecco Modified Eagle Medium (DMEM), suplementado con FBS al 10% inactivado al calor, glutamina 2 mM y penicilina-estreptomycin al 10%.

50 Generación y caracterización de la librería de cDNA: Se purificó el RNA total de las células MDA-MB-175 utilizando el procedimiento del isotiocianato de guanidinio-cloruro de cesio (Sambrook y col., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, (1988)). El RNA poli-A+ se aisló utilizando Dynabeads de oligo(dT) (DYNAL), tal como recomienda el fabricante. Se realizó la síntesis de la primera y segunda cadena con un equipo de síntesis de cDNA de Gibco BRL. Se generaron recombinantes de cDNA  $\lambda$ gt10 utilizando un sistema de clonaje de cDNA de Amersham. El empaquetamiento se realizó *in vitro* con extracto de empaquetamiento GigapackII (Stratagene). El fragmento PstI-XhoI del cDNA de HRG $\beta$ 3 (nt 144-618) se marcó mediante *random-priming* y se rastrearon 1x10<sup>6</sup> calvas. Se verificaron los clones positivos y se purificaron mediante un segundo y tercer cribaje. El DNA del fago se aisló como un fragmento BamHI y se subclonó en la diana correspondiente de pBluescript SK. El clon 5 se secuenció completamente utilizando el equipo de secuenciación de DNA Sequenase versión 2.0 (United States Biochemicals, Inc.). Se secuenciaron ambas cadenas.

60 Sistema de expresión en bacterias: Un fragmento del cDNA del clon 5 (nt 1690-2722) se subclonó en el vector de fusión pET-32 TRX (Novagen). Este fragmento BglII-BglII se insertó en la diana BamHI del plásmido pET32a. La expresión de la proteína trxy-HRG (aminoácidos 455-768) en *E. coli* se indujo tal como recomienda el fabricante.

Purificación de  $\gamma$ -HRG recombinante: Se recolectaron las células *E. coli* que expresaban *trx* $\gamma$ -HRG y se resuspendieron en 9 ml/g en Tris 50 mM, pH 8,0. Se añadió lisozima a una concentración final de 0,2 mg/ml y la solución se agitó en hielo durante 1 h. A continuación se añadió Dnasa I (10  $\mu$ g/ml) y MgCl<sub>2</sub> (4 mM). La solución se sonificó durante 30 min, recogiendo el sedimento. La fracción del sedimento se disolvió en 250ml/g en Gdn HCl 6 M, Tris HCl 0,1 M, pH 8,8. Las proteínas solubilizadas se sulfitolisaron añadiendo 1/10 de volumen de Na<sub>2</sub>SO<sub>3</sub> 1 M. La reacción se continuó durante 1,5 h a temperatura ambiente, a continuación se purificó la proteína mediante cromatografía de filtración en gel con una columna de grado prep High Load Superdex™ 75 (Pharmacia). El replegamiento se inició mediante adición de cisteína 1 M, luego se añadió metionina 10 mM como antioxidante y se incubó durante toda la noche a temperatura ambiente. La concentración de proteínas se determinó mediante el análisis cuantitativo de aminoácidos.

Hibridación de Northern y Southern: el RNA total se aisló por el procedimiento de Chomczynski y col., Anal. Biochem. 162:156-159 (1987). Se aisló el poli(A)+ utilizando columnas de oligo(dT) de celulosa (Qiagen), tal como recomienda el fabricante. El RNA se desnaturalizó y se fraccionó por tamaño en un gel de agarosa al 1%/formaldehído al 0,8% y se transfirió a una membrana de nylon (Hybond, Amersham). El RNA se fijó por UV (UV Strataliker, Stratagene). La prehibridación se llevó a cabo a 42°C en formamida al 50%/SDS al 1%/NaCl 1 M, sulfato de dextrano al 10% y DNA de esperma de salmón a 100  $\mu$ g/ml durante al menos 2 horas. Para la hibridación se utilizaron sondas de cDNA, bien fragmentos de restricción con la secuencia complementaria al dominio de tipo-EGF de HRG $\beta$ 3 o un fragmento KpnI-AvaII del cDNA que codifica únicamente la secuencia de  $\gamma$ -HRG (nt 1238-1866), marcadas por *random-priming* (Prime-it II, Stratagene). La hibridación se llevó a cabo en la misma solución a 42°C con las sondas marcadas con P<sup>32</sup> durante 16 horas. Las transferencias se lavaron varias veces con 2xSSC/SDS al 1% a temperatura ambiente, se lavaron dos veces con la misma solución a 65°C durante 20 minutos y por último se lavaron con 0,2xSSC/SDS al 0,1% a temperatura ambiente durante 15 min. Las transferencias se secaron al aire y se expusieron con una película autoradiográfica Du Pont Reflection™ y con pantallas intensificadoras a -80°C durante 7-40 horas. Las transferencias de Northern (Clontech) de múltiples tejidos de 2  $\mu$ g de poli(A)+ de bazo, timo, próstata, testículos, ovario, intestino delgado, colon, leucocitos de sangre periférica, corazón, cerebro, pulmón, músculo esquelético, riñón y páncreas, se hibridaron con una sonda marcada radioactivamente del cDNA de  $\gamma$ -HRG (nt 841-1447) según las instrucciones del fabricante.

Se aisló el DNA genómico de MDA-MB-175 y MDA-MB-231 tal como se describe en Sambrook y col., *supra*. El DNA se digirió con distintas enzimas de restricción, antes de tratar el gel con HCl 0,25 N y transferirlo a una membrana de nylon (Hybond, Amersham). El fragmento BglII-NdeI del cDNA de  $\gamma$ -HRG (nt 1690-2351) también se marcó isotópicamente mediante *random-priming* y se utilizó como una sonda de hibridación. La prehibridación se llevó a cabo en 6xSSC/5xDenhardt's (SDS al 0,75%, sulfato de dextrano al 1% y DNA de esperma de salmón a 100  $\mu$ g/ml a 68°C durante 4 h. La hibridación se llevó a cabo con una sonda marcada radioactivamente durante toda la noche. Se utilizaron las mismas condiciones de lavado que para las transferencias de Northern, excepto que se añadió una etapa de lavado con 0,2xSSC/0,1% SDS a 68°C y la detección se prosiguió tal como se describió más arriba.

Ensayo de unión de HRG-I<sup>125</sup>: se realizaron ensayos de unión en pocillos rompibles Nunc. Las placas se recubrieron con 100  $\mu$ l de anticuerpo de cabra anti-humano (Boehringer Mannheim) a 5  $\mu$ g/ml a 4°C durante toda la noche en tampón carbonato 50 mM (pH 9,6). Las placas se lavaron dos veces con tampón de lavado (PBS/0,5% Tween-20) y se bloquearon con 100  $\mu$ l de 1% BSA/PBS durante 30 min. El tampón se retiró y cada pocillo se incubó con 15 ng de proteína de fusión IgG en 1% BSA/PBS con agitación vigorosa durante 1,5 horas. Las placas se lavaron tres veces con tampón y se llevó a cabo la unión competitiva añadiendo cantidades diversas de  $\gamma$ -HRG y HRG $\beta$ 1-I<sup>125</sup> con agitación vigorosa. Después de la incubación durante 1,5-2 horas, los pocillos se lavaron tres veces con tampón de lavado, se drenaron y se contaron los pocillos individualmente utilizando un contador- $\gamma$  100 Series Iso Data.

Ensayo de fosforilación de tirosinas: se plaquearon células MCF7 a 1x10<sup>5</sup> células/pocillo en placas de 24 pocillos en F12/DMEM con FBS al 10%. Al cabo de 48 horas, se lavaron las células con F12/DMEM sin suero y se dejaron sin suero durante 6 horas. Diversas concentraciones de  $\gamma$ -HRG truncada expresada en bacterias (es decir, *trx* $\gamma$ -HRG 0 pM, 22 pM, 66 pM, 200 pM y 600 pM) o medio condicionado no purificado de células MDA-MB-175 se prepararon en tampón de unión (BSA al 1% en F12/DMEM) y se añadieron a cada pocillo. Al cabo de 8 min de incubación a temperatura ambiente, se aspiró cuidadosamente el medio y se pararon las reacciones añadiendo 100  $\mu$ l de tampón de muestra (SDS al 5%, 2-mercaptoetanol al 0,25%, Tris-HCl 25 mM pH 6,8). Alícuotas de 20  $\mu$ l de cada muestra se migraron en un gel de gradiente del 4-12% (Novex) y luego, se transfirieron las electroforesis a una membrana de nitrocelulosa. Las inmunotransferencias se revelaron y la banda reactiva predominante de un PM aproximado de 180 KDa se cuantificó por densitometría.

Producción y caracterización del medio condicionado de las células MDA-MB 175. Las células se sembraron en frascos T175 y se crecieron hasta alcanzar el 70-80% de confluencia (~ 2,5x10<sup>7</sup> células/frasco). Seguidamente, se lavaron las células con PBS y se crecieron sin suero en medio F12/DMEM durante 3-4 días. El medio se recogió a continuación, se filtró y se concentró utilizando una célula de ultrafiltración con membranas YM10 Diaflo (Amicon). La  $\gamma$ -HRG se visualizó en el medio condicionado de las células MDA-MB-175 mediante análisis de transferencia de Western en condiciones no reductoras. La  $\gamma$ -HRG se purificó parcialmente por HPLC usando una columna de fase inversa C4. La HRG $\beta$ 1 de tamaño completo (carril 1) y la  $\gamma$ -HRG semi pura (carril 2) expresadas en células CHO se migraron en una electroforesis, se realizó una transferencia de Western y la membrana se hibridó con heterodí-

meros HER2/HER4 IgG. Se observó una banda de ~ 64 KDa en el carril que contenía el sobrenadante parcialmente purificado, mientras que la HRG- $\beta$ 1 de tamaño completo y expresada en CHO migró como una proteína de 45 KDa.

Ensayo de proliferación celular con cristal violeta. Las líneas tumorales se sembraron en placas de 96 pocillos a las siguientes densidades:  $2 \times 10^4$  células/pocillo para MDA-MB-175 y  $1 \times 10^4$  células/pocillo para SK-BR-3 en un medio contenía FBS al 1%. Las células se dejaron adherir durante 2 h. Se añadieron anticuerpos monoclonales, inmunoadhesinas (10  $\mu$ g/ml) o únicamente medio y las células se incubaron de nuevo durante 2 horas a 37°C. Se añadió rHRG $\beta$ 1<sub>177-244</sub> a una concentración final de 1 nM, o 100 nM para neutralizar la inmunoadhesión y las células se incubaron durante 4 días. Las monocapas se lavaron con PBS y se tiñeron/fijaron con cristal violeta al 0,5%. Las placas se secaron al aire, el colorante se diluyó con citrato sódico 0,1 M (pH 4,2) en etanol (50:50) y se cuantificó la absorbancia a 540 nm.

Aislamiento y análisis de secuenciación de  $\gamma$ -HRG. Para caracterizar el transcrito de heregulina en las células MDA-MB-175, se construyó una librería de cDNA en  $\lambda$ gt10 con el mRNA derivado de esta línea celular. La librería se rastreó con una sonda de cDNA correspondiente con el dominio de tipo-EGF y parte de la secuencia N-terminal de HRG $\beta$ 3 y se identificaron diversos clones. Uno de los clones que parecía contener la secuencia del cDNA de tamaño completo, se aisló y secuenció. La Fig. 7A-7D muestra la secuencia nucleotídica y la secuencia de aminoácidos predicha de  $\gamma$ -HRG. La única pauta de lectura abierta predicha de 2303 pb comienza con un codón ATG en el nt 334. Este codón de inicio se halla en un contexto de secuencia nucleotídica conocido por ser un sitio potencial de iniciación de la transcripción (Kozak, Nucleic Acid Research 15:8125-8148 (1987)). Se hallaron diversos codones de terminación cadena arriba del codón de inicio El codón de parada TAG en el nt 2637 se prosigue por la secuencia 3' no codificante, idéntica a las otras secuencias de isoformas de HRG e incluye una señal de poliadenilación seguida por una región rica en A. La pauta de lectura abierta codifica una proteína de 768 residuos de aminoácidos con un peso molecular calculado de 84,2 KDa.

(d) La selección de las variantes de HRG- $\beta$ 1 que contenían residuos correspondientes con el dominio de tipo-EGF (HRG- $\beta$ 1 177-228) se realizó mediante expresión en fagos monovalentes. Para estas variantes, los números de los residuos se muestran entre paréntesis indicando la posición del residuo en el dominio mínimo de tipo-EGF (es decir, HRG- $\beta$ 1 EGF 1-52).

Las variantes de HRG- $\beta$ 1 EGF se prepararon y seleccionaron para la unión con HER-3-Ig utilizando la expresión en fagos monovalentes, de acuerdo con el procedimiento de Bass y col., Proteins 8:309-314 (1990). Tal como se describe con mayor detalle más adelante, se preparó un vector fagémido, en el que HRG- $\beta$ 1 EGF se fusionó con un fragmento C-terminal de la proteína pIII de la cubierta de M13. Se realizó la mutagénesis de Kunkel para introducir codones de parada en este vector en los sitios seleccionados para la mutagénesis aleatoria. Esta etapa asegura que el vector de inicio sea incapaz de expresar el polipéptido de tipo salvaje. Tramos de 4 a 6 residuos por librería se mutaron aleatoriamente de forma lineal, excepto para las 6 cisteínas, Phe189 (HRG $\beta$ 1 EGF Phe 13) y los dos residuos más C-terminales. La Phe 189 no se alteró, porque este residuo está conservado como un residuo aromático en el EGF y el TGF- $\alpha$  y forma una interacción con la Tyr208 (HRG- $\beta$ 1 EGF Tyr32) Jacobsen y col., Biochemistry 35:3402-17 (1996). En consecuencia, HRG- $\beta$ 1 EGF se preparó en 8 librerías, denominadas A-E, G, H e I.

La librería E que comprendía HRG- $\beta$ 1 202-209 (HRG- $\beta$ 1 EGF 26-33), contenía una delección de tres residuos. La región delecionada corresponde con un giro desordenado entre la segunda y la tercera hoja  $\beta$  de la HRG- $\beta$ 1 EGF y los aminoácidos equivalentes están ausentes en EGF y TGF- $\alpha$ . Una variante control de HRG- $\beta$ 1 EGF, en la que se delecionó HRG- $\beta$ 1 202-204 (HRG- $\beta$ 1 EGF 26-28) de HRG8 (HRG63), se unió con HER-3-Ig con una afinidad similar a la de tipo salvaje.

Se creó una librería adicional (F) para mutar aleatoriamente un fragmento de superficie compuesto de cadenas laterales de la primera y segunda hoja  $\beta$ , incluyendo HRG- $\beta$ 1 178, 180, 198 y 200 (HRG- $\beta$ 1 EGF2, 4, 22 y 24).

Los sitios seleccionados en los vectores de inicio se mutaron aleatoriamente según el procedimiento de mutagénesis de Kunkel para producir librerías de HRG- $\beta$ 1 EGF. Se produjeron fagos que expresaban HRG- $\beta$ 1 EGFs a partir de librerías en condiciones tales, estadísticamente, que cada fago expresaba no más de una copia de HRG- $\beta$ 1 EGF mutada. Véase Bass y col., *supra*. Estos fagos se seleccionaron por su unión a (cribados contra) HER3-Ig inmovilizada en una placa de ELISA. Los fagos unidos se eluyeron y se utilizaron para reinfectar las células huésped, que se utilizaron para producir nuevos fagos para otra tanda de selección. Este proceso se repitió seis a siete veces por cada librería. Por último, se secuenciaron 12 clones de los fago seleccionados de cada librería.

# ES 2 329 467 T3

TABLA 2

| Variantes de la librería A |                    |     |     |     |     |
|----------------------------|--------------------|-----|-----|-----|-----|
|                            | Posición en HRG-β1 |     |     |     |     |
| Variantes N°               | 177                | 178 | 179 | 180 | 181 |
| Tipo salvaje               | S                  | H   | L   | V   | K   |
| 1                          | W                  | R   | .   | .   | P   |
| 2                          | W                  | S   | .   | Q   | P   |
| 3,5,10                     | W                  | E   | .   | .   | P   |
| 4                          | W                  | S   | .   | .   | .   |
| 6                          | W                  | S   | .   | I   | P   |
| 7                          | W                  | R   | .   | .   | A   |
| 8                          | W                  | A   | .   | .   | P   |
| 9                          | W                  | S   | .   | Q   | .   |
| 11                         | W                  | E   | .   | .   | A   |
| 12                         | W                  | S   | .   | E   | P   |

TABLA 3

| Variantes de la librería B |                    |     |     |     |     |     |
|----------------------------|--------------------|-----|-----|-----|-----|-----|
|                            | Posición en HRG-β1 |     |     |     |     |     |
| Variantes N°               | 183                | 184 | 185 | 186 | 187 | 188 |
| Tipo salvaje               | A                  | E   | K   | E   | K   | T   |
| 1*                         | G                  | V   | G   | R   | D   | G   |
| 2*                         | G                  | G   | E   | R   | E   | G   |
| 3                          | G                  | .   | E   | R   | E   | G   |
| 4*,5*                      | G                  | W   | D   | R   | E   | G   |
| 6*                         | G                  | V   | Q   | R   | E   | G   |
| 7                          | G                  | .   | E   | R   | A   | G   |
| 8                          | G                  | K   | E   | R   | E   | G   |
| 9*                         | T                  | N   | S   | R   | E   | G   |
| 10*                        | D                  | K   | S   | R   | E   | G   |
| 11*                        | G                  | .   | D   | R   | .   | Q   |
| 12                         | G                  | R   | E   | R   | E   | G   |

\* Variante que también contiene la mutación Met226Ile

# ES 2 329 467 T3

TABLA 4

| Variantes de la librería C |                            |     |     |     |     |
|----------------------------|----------------------------|-----|-----|-----|-----|
|                            | Posición en HRG- $\beta$ 1 |     |     |     |     |
| Variantes N°               | 191                        | 192 | 193 | 194 | 195 |
| Tipo salvaje               | V                          | N   | G   | G   | E   |
| 1,1,4,5,7-12               | .                          | .   | .   | .   | .   |
| 3                          |                            |     |     |     | V   |
| 6                          |                            |     |     |     | Q   |

TABLA 5

| Variantes de la librería D |                            |     |     |     |     |
|----------------------------|----------------------------|-----|-----|-----|-----|
|                            | Posición en HRG- $\beta$ 1 |     |     |     |     |
| Variantes N°               | 197                        | 198 | 199 | 200 | 201 |
| Tipo salvaje               | F                          | M   | V   | K   | D   |
| 1*,2*,8*,12*               | Y                          | K   | .   | R   | I   |
| 3                          | .                          | R   | .   | .   | T   |
| 4,5,7,9                    | Y                          | R   | .   | .   | T   |
| 6                          | Y                          | .   | I   | .   | Y   |
| 10                         | Y                          | .   | .   | .   | T   |
| 11                         | M                          | R   | .   | R   | T   |

\* Variante que también contiene la mutación Met226Ile

TABLA 6

| Variantes de la librería E |                            |     |     |     |     |
|----------------------------|----------------------------|-----|-----|-----|-----|
|                            | Posición en HRG- $\beta$ 1 |     |     |     |     |
| Variantes N°               | 205                        | 206 | 207 | 208 | 209 |
| Tipo salvaje               | P                          | S   | R   | Y   | L   |
| 1                          | T                          | P   | Y   | L   | M   |
| 2,4                        | Y                          | G   | Y   | L   | M   |
| 3*                         | Y                          | R   | Y   | R   | M   |
| 5,12                       | T                          | H   | Y   | R   | G   |
| 6                          | T                          | H   | Y   | R   | M   |
| 7*                         | Y                          | K   | Y   | R   | M   |
| 8,9                        | T                          | K   | Y   | R   | G   |
| 10                         | Y                          | K   | Y   | R   | .   |

\* Variante que también contiene la mutación Met226Ile

TABLA 7

| Variantes de la librería G |                            |     |     |     |     |     |
|----------------------------|----------------------------|-----|-----|-----|-----|-----|
|                            | Posición en HRG- $\beta$ 1 |     |     |     |     |     |
| Variantes N°               | 211                        | 212 | 213 | 214 | 215 | 216 |
| Tipo salvaje               | K                          | C   | P   | N   | N   | F   |
| 1, 5, 6, 10, 12            | R                          | .   | S   | L   | L   | .   |
| 2                          | R                          | .   | S   | E   | E   | .   |
| 3                          | .                          | .   | .   | K   | K   | M   |
| 4                          | R                          | .   | T   | V   | V   | Y   |
| 7, 8                       | R                          | .   | T   | V   | V   | Y   |
| 9                          | .                          | .   | N   | S   | .   | .   |
| 11                         | R                          | .   | K   | K   | .   | .   |

## Ejemplo 2

*Expresión de HER2/HER3 en pulmón de rata embrionario*

Se microdisecionaron los pulmones de rata de embriones (E) de 16, 18, 20 días y de ratas post-natales (P) de 7 y 14 días y de adultos. El tejido aislado del pulmón se homogeneizó en un tampón con inhibidor de proteasas estándar y cantidades iguales de proteína se migraron en un gel de SDS-PAGE (4-20%), se transfirieron a una membrana de nitrocelulosa y se identificaron con anticuerpos específicos (HER2, HER3, HER4-Santa Cruz Biological, San José, CA, y HRG, 3G11, Genentech, Inc.) mediante técnicas de inmunofluorescencia.

El análisis de las transferencias indica que HER2 se expresa a niveles elevados durante el desarrollo en el útero, aparece el día E18 y disminuye después del nacimiento hasta alcanzar los niveles bajos del adulto. La expresión de HER3 alcanza un pico máximo en E18 y luego disminuye después del nacimiento hasta los niveles bajos del adulto. HER4 no se identifica en ningún momento durante el desarrollo del pulmón. Una preforma de HRG puede estar presente (proteína de 75 KDa) el día E16, con un máximo el E18 y luego disminuye a los niveles bajos del adulto. Estos resultados sugieren que el sistema HER/HRG se modula durante el desarrollo del pulmón y el receptor activo puede ser un heterodímero HER2/HER3. El período de tiempo durante el cual la expresión de los receptores es máxima (E18) representa el estadio pseudoglandular, la proliferación de células epiteliales pulmonares es la mayor de cualquier otro período. Durante el estadio canalicular, la diferenciación se inicia con la aparición de pneumocitos de tipo I (producción de tensoactivos) y tipo II. Ello indica un papel para la interacción HRG/HER en cualquiera de estos procesos o en ambos.

## Ejemplo 3

*Expresión de HER2/HER3 en el pulmón humano fetal*

El pulmón humano fetal se obtuvo de embriones del segundo trimestre (17-22 semanas). El tejido del pulmón se cultivó en medio Weymouth sin suero en una interfaz de fluido aéreo a 37°C y en atmósfera de humedad con CO<sub>2</sub> al 5%. El tejido se cosechó el día 0 (el día de obtención del tejido), y después de 1-4 días (D) de cultivo (D1-D14), se homogeneizó en un tampón con inhibidor de proteasas estándar y cantidades equimolares de proteína se sometieron a una electroforesis de SDS-PAGE (4-20%), se transfirieron a una membrana de nitrocelulosa y se identificaron con anticuerpos específicos.

HER2 se expresó el día D0 e incrementó el nivel durante el cultivo *in vitro*. HER3 está presente a niveles bajos a indetectables en D0 e incrementa su expresión durante el cultivo *in vitro*. HER4 no se observó en ningún momento durante el desarrollo del pulmón. La HRG no se identificó en D0, sin embargo, se identificó una proteína de 75 KDa en D1-D2 y su expresión continuó aumentado durante el tiempo de cultivo. Este modelo de explantes humanos recapitula parte del programa de desarrollo normal del pulmón durante los 5 días de cultivo. El desarrollo tiene lugar rápidamente, y tanto la proliferación de las células epiteliales como su diferenciación se producen conjuntamente durante la formación del espacio aéreo. Estos resultados indican que el sistema HER2/HRG también se modula durante el desarrollo del pulmón humano *in vitro*, y el receptor activo puede ser un heterodímero HER2/HER3. El tercer

trimestre representa el estadio pseudoglandular tardío (días 42-112) y el estadio canalicular (días 112-196) en el desarrollo del pulmón humano. Al igual que en la rata, también tiene lugar la proliferación celular y la formación de vías aéreas prospectivas durante el estadio pseudoglandular. Durante el estadio canalicular tiene lugar la diferenciación del epitelio.

#### Ejemplo 4

##### *Expresión de HER2 y HER3 en el pulmón humano*

HER2 y HER3 se expresan exclusivamente en el epitelio pulmonar durante el desarrollo del pulmón. Se cultivó el pulmón humano del segundo trimestre *in vitro*, tal como se describió más arriba. El tejido se cosechó diariamente, se congeló y se cortaron secciones de 5 micras. Las secciones se montaron en cubreobjetos de vidrio y se realizó la inmunohistoquímica utilizando procedimientos ABC estándares. HER2, HER3 y HER4 se identificaron utilizando anticuerpos específicos tal como se describió anteriormente.

Como control, se tiñó el tejido de pulmón en el día D0 y D5 con un anticuerpo irrelevante sin que se observara ninguna señal. En D0, el pulmón estaba relativamente poco formado. La mayor parte de los tejidos estaban compuestos de células mesenquimatosas y se desarrollaron espacios aéreos temprano. En D5, los espacios aéreos eran claramente identificables, con un tejido mesenquimatoso más delgado y con proliferación del epitelio pulmonar necesario para recubrir los espacios aéreos en crecimiento.

Utilizando la tinción de HER2 en D0 y D5, se demostró que HER2 estaba presente de forma no uniforme en el tejido de pulmón D0. No se observó expresión en el tejido mesenquimatoso. En D5, HER2 permanece uniforme y restringido al epitelio pulmonar.

Utilizando la tinción control de HER3 en D0 y D5, se identificó HER3 en el tejido pulmonar D0. La señal de tinción fue relativamente inferior a la de HER2 y tampoco homogénea, lo que sugiere que existen áreas específicas epiteliales que expresan el receptor HER3. En D5, la expresión fue más homogénea en el epitelio pulmonar, pero tampoco uniforme. La expresión continuó siendo específica de epitelio.

#### Ejemplo 5

##### *Preparación de rHRGβ1<sub>177-244</sub>*

El fragmento del dominio de tipo-EGF HRGβ1<sub>177-244</sub> se amplificó del vector pHL89 (descrito por Holmes y col., Science 256:1205-1210 (1992)) mediante PCR con cebadores que contenían extremos con dianas protuberantes NsiI/XbaI. El fragmento se insertó en vector de expresión fagémido pam-g3 mediante digestión con enzimas de restricción y ligación en las mismas dianas para generar la construcción pHRG2-g3 (177-244). El vector pam-g3 es un derivado de pHGHam-g3, diseñado para la expresión en fagos de la hormona de crecimiento humana (hGH), descrito por Lowman y col., Biochemistry 30:10832-10838 (1991), pam-g3 se produjo mediante eliminación del gen hGH presente en pHGHam-g3 y sustitución por un fragmento de relleno, que proporciona espacio para el corte por las dianas de restricción utilizadas en el clonaje. El fragmento HRG-β1 se unió al residuo 247 de pIII.

El dominio de tipo-EGF HRG-β1 expresado de la anterior construcción, descrita más arriba, se denominó sin la “p” y “g3” que aparecen en el nombre de la construcción. En consecuencia, el dominio de tipo-EGF HRG-β1 expresado de la construcción pHRG2-g3 se denominó “HRG2”.

El dominio se expresó de forma monovalente en los fagos como una proteína de fusión pIII, tal como se describe por Bass y col., Proteins 78:309-314 (1990).

De forma similar, las variantes HRG-β1<sub>147-227</sub>, HRG-β1<sub>147-244</sub> y HRG-β1<sub>177-227</sub> se prepararon y expresaron tal como se describió anteriormente.

#### Ejemplo 6

##### *rHRGβ1<sub>177-244</sub> provoca un desarrollo acelerado del pulmón *in vitro**

La rHRGβ1<sub>177-244</sub> provoca un desarrollo acelerado del pulmón *in vitro*. Para determinar si los receptores HER2/HER3 expresados eran funcionales y comprobar el papel de la estimulación de HRG durante el desarrollo del pulmón *in vitro*, se añadió al cultivo *in vitro* rHRGβ1<sub>177-244</sub> a 10 nM. El cultivo se cosechó en el día D5, se congeló y se cortaron secciones de 5 micras para su análisis.

La morfología, en comparación con las muestras control no tratadas fue extremadamente distinta. Hubo una proliferación marcada del epitelio. Los espacios aéreos que normalmente están revestidos de una única capa de células

tenían ahora un espesor de 2-3 células. Los cambios fueron dependientes de la dosis con una mayor respuesta de células epiteliales a concentraciones mayores. HER2 y HER3 fueron todavía identificables y únicamente en el epitelio.

## 5 Ejemplo 7

### *Diferenciación del pulmón humano*

La diferenciación de las células epiteliales del pulmón humano tuvo lugar después del tratamiento con HRG. La diferenciación se cuantificó por la producción de proteína A tensoactiva (SPA). Todas las secciones se tiñeron para SPA. El explante de pulmón humano fue positivo para la tinción para SPA, localizándose en las células epiteliales de los ductos prealveolares. El explante humano expuesto a 10 nM de HRG mostró un efecto sobre la producción de SPA. Como un control de diferenciación, se expuso un explante de pulmón a dibutiril cAMP 1 mM. También se realizó un control negativo.

### **Referencias citadas en la descripción**

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

### **Documentos de patente citados en la descripción**

- US 4968603 A, Slamon [0004]
- US 5183884 A [0005]
- EP 444961 A1 [0005]
- EP 599274 A [0006] [0006]
- WO 9220798 A [0008]
- US 5367060 A [0008] [0030] [0030] [0030] [0030] [0205]
- EP 505148 A [0010]
- WO 9322424 A [0010]
- WO 9428133 A [0010]
- WO 9408007 A [0011]
- WO 9218627 A [0012]
- WO 9400140 A [0012] [0018]
- WO 9404560 A [0012]
- WO 9426298 A [0012] [0017]
- WO 9532724 A [0012]
- US 08891845 B [0030]
- US 08339517 B [0030]
- WO 9802541 A [0034]
- WO 9835036 A [0034]
- US 5770567 A [0034]
- WO 9615244 A [0034] [0111] [0205]
- EP 266032 A [0057]

## ES 2 329 467 T3

- US 4683195 A [0058]
- US 4816567 A [0069] [0070]
- 5 • EP 404097 A [0073]
- WO 9311161 A [0073]
- WO 8902922 A [0099]
- 10 • EP 117060 A [0139] [0153]
- US 4965199 A [0139]
- 15 • EP 36776 A [0142]
- US 5641655 A [0142]
- EP 73657 A, Hitzeman [0144]
- 20 • GB 2211504 A [0146]
- US 4419446 A [0147]
- 25 • US 4601978 A [0147]
- US 5024939 A [0152]
- EP 117058 A [0153]
- 30 • EP 307247 A [0153]
- EP 139383 A [0155]
- 35 • US SN4943529 A [0155]
- EP 402226 A [0155]
- EP 183070 A [0155]
- 40 • EP 244234 A [0155]
- WO 9100357 A [0155]
- 45 • EP 321196 A [0156]
- WO 8905859 A [0160]
- US 4399216 A [0160]
- 50 • US 4767704 A [0162]
- US 4657866 A [0162]
- 55 • US 4927762 A [0162]
- US 4560655 A [0162]
- WO 9003430 A [0162]
- 60 • WO 8700195 A [0162]
- US RE30985 E [0162]
- 65 • US 5122469 A [0162]
- US 3969287 A [0178]

## ES 2 329 467 T3

- US 3691016 A [0178]
- US 4195128 A [0178]
- 5 • US 4247642 A [0178]
- US 4229537 A [0178]
- US 4330440 A [0178]
- 10 • WO 8705330 A [0185]
- US 4640835 A [0188]
- 15 • US 4496689 A [0188]
- US 4301144 A [0188]
- US 4670417 A [0188]
- 20 • US 4791192 A [0188]
- US 4179337 A [0188]
- 25 • US 9002421 W [0199]
- US 3773919 A [0200]
- EP 58481 A [0200]
- 30 • EP 133988 A [0200]
- DE 3218121 [0201]
- 35 • EP 52322 A [0201]
- EP 36676 A [0201]
- EP 88046 A [0201]
- 40 • EP 143949 A [0201]
- EP 142641 A [0201]
- 45 • JP 58118008 A [0201]
- US 4485045 A [0201]
- US 4544545 A [0201]
- 50 • EP 102324 A [0201]
- US 5013556 A [0201]
- 55 • US 09020598 B [0245].

### Documentos que no son patentes citados en la descripción

- **Slamon et al.** *Science*, 1987, vol. 235, 177-182 [0004]
- 60 • **Slamon et al.** *Science*, 1989, vol. 244, 707-712 [0004]
- **Kraus et al.** *Proc. Natl. Acad. Sci. USA*, 1989, vol. 86, 9193-9197 [0005]
- 65 • **Kraus et al.** *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 2900-2904 [0005]
- **Lemoine et al.** *Br. J. Cancer*, 1992, vol. 66, 1116-1121 [0005]

- Poller *et al. J. Pathol.*, 1992, vol. 168, 275-280 [0005]
- Rajkumer *et al. J. Pathol.*, 1993, vol. 170, 271-278 [0005]
- 5 • Sanidas *et al. Int. J. Cancer*, 1993, vol. 54, 935-940 [0005]
- Lemoine. *J. Pathol.*, 1992, vol. 168, 269-273 [0005]
- Friess *et al. Clinical Cancer Research*, vol. 1, 1413-1420 [0005]
- 10 • Plowman *et al. Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 1746-1750 [0006]
- Plowman *et al. Nature*, 1993, vol. 366, 473-475 [0006]
- 15 • Op-Urtreger *et al. Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 1867-1871 [0007]
- Lee; Wood. *Genomics*, 1993, vol. 16, 790-791 [0007]
- Holmes *et al. Science*, 1992, vol. 256, 1205-1210 [0008] [0238]
- 20 • Peles *et al. Cell*, 1992, vol. 69, 205-216 [0010]
- Wen *et al. Cell*, 1992, vol. 69, 559-572 [0010]
- 25 • Wen *et al. Mol. Cell. Biol.*, 1994, vol. 14 (3), 1909-1919 [0010]
- Falls *et al. Cell*, 1993, vol. 72, 801-815 [0011]
- Marchionni *et al. Nature*, 1993, vol. 362, 312-318 [0012]
- 30 • Ho *et al. J. Biol. Chem.*, 1995, vol. 270 (4), 14523-14532 [0013]
- Peles *et al. EMBO J.*, 1993, vol. 12, 961-971 [0014]
- 35 • Carraway *et al. J. Biol. Chem.*, 1994, vol. 269 (19), 14303-14306 [0014]
- Sliwkowski *et al. J. Biol. Chem.*, 1994, vol. 269 (20), 14661-14665 [0014]
- Kokai *et al. Cell*, 1989, vol. 58, 287-292 [0015]
- 40 • Stem *et al. EMBO J.*, 1988, vol. 7, 995-1001 [0015]
- King *et al. EMBO J.*, 1989, vol. 4, 13-18 [0015]
- 45 • Wada *et al. Cell*, 1990, vol. 61, 1339-1347 [0015]
- Plowman *et al. Nature*, 1993, vol. 336, 473-475 [0016]
- Corfas; Fischbach. *J. Neuroscience*, 1993, vol. 13 (5), 2118-2125 [0017]
- 50 • Sklar *et al. J. Cell Biochem.*, 1994, vol. W462 (18D), 540 [0017]
- Brockes *et al. J. Biol. Chem.*, 1980, vol. 255 (18), 8374-8377 [0018]
- 55 • Lernke; Brockes. *J. Neurosci.*, 1984, vol. 4, 75-83 [0018]
- Lemke; Brockes. *J. Neurosci.*, 1984, vol. 4, 75-83 [0018]
- Brockes *et al. Ann. Neurol.*, 1986, vol. 20 (3), 317-322 [0018]
- 60 • Brockes, *J. Methods in Enzym.*, 1987, vol. 147, 217-225 [0018]
- Levi *et al. J. Neuroscience*, 1994, vol. 14 (3), 1309-1319 [0018]
- 65 • Li *et al. J. Neuroscience*, 1996, vol. 16 (6), 2012-2019 [0018]
- Pinkas-Kramarski *et al. PNAS, USA*, 1994, vol. 91, 9387-9391 [0019]

- **Meyer**; Birchmeier. *PNAS*, USA, 1994, vol. 91, 1064-1068 [0019]
- **Meyer et al.** *Development*, 1997, vol. 124 (18), 3575-3586 [0019]
- 5 • **Danilenko et al.** Abstract 3101, *FASEB*, 1994, vol. 8 (4-5), A535 [0019]
- **Danilenko et al.** *Journal of Clinical Investigation*, 1995, vol. 95 (2), 842-851 [0019]
- **Ram et al.** *Journal of Cellular Physiology*, 1995, vol. 163, 589-596 [0020]
- 10 • **Danilenko et al.** *J. Clin. Invest.*, 1995, vol. 93, 842-851 [0020]
- **Marikovsky et al.** *Oncogene*, 1995, vol. 10, 1403-1411 [0020]
- 15 • **Cunningham; Wells.** *Science*, 1989, vol. 244, 1081-1085 [0021] [0095]
- **Smith.** *Science*, 1985, vol. 228, 1315-1317 [0022]
- **Parmley; Smith.** *Gene*, 1985, vol. 73, 305-318 [0022]
- 20 • **Scott et al.** *Science*, 1990, vol. 249, 386-390 [0023]
- **Cwirla et al.** *PNAS USA*, 1990, vol. 87, 6378-6382 [0023]
- 25 • **Bass et al.** *Proteins*, 1990, vol. 8, 309-314 [0023] [0221][0240]
- **Lowman et al.** *Biochem.*, 1991, vol. 30, 10832-10838 [0023]
- **Cunningham; Wells.** *Science*, 1989, vol. 244, 1081-85 [0034]
- 30 • **Bass et al.** *Proteins*, 1990, vol. 8, 309-14 [0034]
- **Lowman et al.** *Biochem.*, 1991, vol. 30, 10832-38 [0034]
- 35 • **Sambrook et al.** *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press*, 1989, 1.56-1.61 [0052]
- **Lawn et al.** *Nucleic Acids Res*, 1981, vol. 9, 6103-6114 [0053]
- 40 • **Goeddel et al.** *Nucleic Acids Res*, 1980, vol. 8, 4057 [0053]
- **Froehler et al.** *Nucl. Acids Res.*, 1986, vol. 14, 5399-5407 [0057]
- **Mullis et al.** *Cold Spring Harbor Symp. Quant. Biol.*, 1987, vol. 51, 263 [0058]
- 45 • *PCR Technology. Stockton Press*, 1989 [0058]
- **Sadich et al.** *Anal. Biochem.*, 1996, vol. 235, 207-214 [0059]
- 50 • **Henzel et al.** *J. Chromatography*, 1987, vol. 404, 41-52 [0060]
- **Matsudaira, P.** *J. Biol. Chem*, 1987, vol. 262, 10035-10038 [0060]
- **Kohler et al.** *Nature*, 1975, vol. 256, 495 [0069]
- 55 • **Clackson et al.** *Nature*, 1991, vol. 352, 624-628 [0069]
- **Marks et al.** *J. Mol. Biol*, 1991, vol. 222, 581-597 [0069]
- 60 • **Morrison et al.** *Proc. Natl. Acad. Sci. USA*, 1984, vol. 81, 6851-6855 [0070]
- **Jones et al.** *Nature*, 1986, vol. 321, 522-525 [0071]
- **Reichmann et al.** *Nature*, 1988, vol. 332, 323-329 [0071]
- 65 • **Presta.** *Curr. Op. Struct. Biol.*, 1992, vol. 2, 593-596 [0071]

## ES 2 329 467 T3

- **Pluckthun.** The Pharmacology of Monoclonal Antibodies. *Springer-Verlag*, 1994, vol. 113, 269-315 [0072]
- **Hollinger et al.** *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 6444-6448 [0073]
- 5 • **Zapata et al.** *Protein Eng.*, 1995, vol. 8 (10), 1057-1062 [0074]
- **Engels et al.** *Agnew. Chem. Int. Ed. Engl.*, 1989, vol. 28, 716-734 [0086]
- **Adelman et al.** *DNA*, 1983, vol. 2, 183 [0117]
- 10 • **Crea et al.** *Prac. Natl. Acad. Sci. USA*, 1978, vol. 75, 5765 [0118]
- **Wells et al.** *Gene*, 1985, vol. 34, 315 [0129]
- 15 • **Southern et al.** *J. Molec. Appl. Genet.*, 1982, vol. 1, 327 [0137]
- **Mulligan et al.** *Science*, 1980, vol. 209, 1422 [0137]
- **Sugden et al.** *Mol. Cell. Biol.*, 1985, vol. 5, 410-413 [0137]
- 20 • **Urlaub; Chasin.** *Proc. Natl. Acad. Sci. USA*, 1980, vol. 77, 4216 [0139]
- **Stinchcomb et al.** *Nature*, 1979, vol. 282, 39 [0140]
- 25 • **Kingsman et al.** *Gene*, 1979, vol. 7, 141 [0140]
- **Tschemper et al.** *Gene*, 1980, vol. 10, 157 [0140]
- **Jones.** *Genetics*, 1977, vol. 85, 12 [0140]
- 30 • **Chang et al.** *Nature*, 1978, vol. 275, 615 [0142]
- **Goeddel et al.** *Nature*, 1979, vol. 281, 544 [0142]
- 35 • **Goeddel.** *Nucleic Acids Res.*, 1980, vol. 8, 4057 [0142]
- **deBoer et al.** *Proc. Natl. Acad. Sci. USA*, 1983, vol. 80, 21-25 [0142]
- **Siebenlist et al.** *Cell*, 1980, vol. 20, 269 [0142]
- 40 • **Hitzeman et al.** *J. Biol. Chem*, 1980, vol. 255, 2073 [0143]
- **Hess et al.** *J. Adv. Enzyme Reg*, 1968, vol. 7, 149 [0143]
- 45 • **Holland.** *Biochemistry*, 1978, vol. 17, 4900 [0143]
- **Fiers et al.** *Nature*, 1978, vol. 273, 113 [0147]
- **Mulligan; Berg.** *Science*, 1980, vol. 209, 1422-1427 [0147]
- 50 • **Pavlakakis et al.** *Proc. Natl. Acad. Sci. USA*, 1981, vol. 78, 7398-7402 [0147]
- **Greenaway et al.** *Gene*, 1982, vol. 18, 355-360 [0147]
- 55 • **Gray et al.** *Nature*, 1982, vol. 295, 503-508 [0147]
- **Reyes et al.** *Nature*, vol. 297, 598-601 [0147]
- **Canaani; Berg.** *Proc. Natl. Acad. Sci. USA*, 1982, vol. 79, 5166-5170 [0147]
- 60 • **Gorman et al.** *Proc. Natl. Acad. Sci. USA*, 1982, vol. 79, 6777-6781 [0147]
- **Laimins et al.** *Proc. Natl. Acad. Sci. USA*, 1981, vol. 78, 993 [0148]
- 65 • **Lusky et al.** *Mol. Cell Bio*, 1983, vol. 3, 1108 [0148]
- **Banerji et al.** *Cell*, 1993, vol. 33, 729 [0148]

## ES 2 329 467 T3

- **Osborne et al.** *Mol. Cell Bio.*, 1984, vol. 4, 1293 [0148]
- **Yaniv.** *Nature*, 1982, vol. 297, 17-18 [0148]
- 5 • **Messing et al.** *Nucleic Acids Res*, 1981, vol. 9, 309 [0151]
- **Maxam et al.** *Methods in Enzymology*, 1980, vol. 65, 499 [0151]
- **Gething et al.** *Nature*, 1981, vol. 293, 620-625 [0153]
- 10 • **Mantei et al.** *Nature*, 1979, vol. 281, 40-46 [0153]
- **Beach; Nurse.** *Nature*, 1981, vol. 290, 140 [0155]
- 15 • **Louvencourt et al.** *J. Bacteriol.*, 1983, 737 [0155]
- **Sreekrishna et al.** *J. Basic Microbiol.*, 1988, vol. 28, 265-278 [0155]
- **Case et al.** *Proc. Natl. Acad. Sci. USA*, 1979, vol. 76, 5259-5263 [0155]
- 20 • **Ballance et al.** *Biochem. Biophys. Res. Commun.*, 1983, vol. 112, 284-289 [0155]
- **Tilburn et al.** *Gene*, 1983, vol. 26, 205-221 [0155]
- 25 • **Yelton et al.** *Proc. Natl. Acad. Sci USA*, 1984, vol. 81, 1470-1474 [0155]
- **Kelly; Hynes.** *EMBO J.*, 1985, vol. 4, 475-479 [0155]
- **Luckow et al.** *Bio/Technology*, 1988, vol. 6, 47-55 [0156]
- 30 • **Miller et al.** *Genetic Engineering. Plenum Publishing*, 1986, vol. 8, 277-279 [0156]
- **Maeda et al.** *Nature*, 1985, vol. 315, 592-594 [0156]
- 35 • **Depicker et al.** *J. Mol. Appl. Gen.*, 1992, vol. 1, 561 [0156]
- Tissue Culture. *Academic Press*, 1973 [0157]
- **Graham et al.** *J. Gen Virol.*, 1977, vol. 36, 59 [0157]
- 40 • **Urlaub; Chasin.** *Proc. Natl. Acad Sci. USA*, 1980, vol. 77, 4216 [0157]
- **Mather.** *Biol. Reprod.*, 1980, vol. 23, 243-251 [0157]
- 45 • **Mather et al.** *Annals N.Y. Acad Sci.*, 1982, vol. 383, 44-68 [0157]
- **Shaw et al.** *Gene*, 1983, vol. 23, 313 [0160]
- Van **Solingen et al.** *J. Bact.*, 1977, vol. 130, 946 [0160]
- 50 • **Hsiao et al.** *Proc. Natl. Acad Sci. (USA)*, 1979, vol. 76, 3829 [0160]
- **Ham; Wallace.** *Meth. Enz.*, 1979, vol. 58, 44 [0162]
- 55 • **Barnes; Sato.** *Anal. Biochem.*, 1980, vol. 102, 255 [0162]
- **Thomas.** *Proc. Natl. Acad Sci USA*, 1980, vol. 77, 5201-5205 [0165]
- **Hsu et al.** *Am. J. Clin. Path.*, 1980, vol. 75, 734-738 [0166]
- 60 • T. E. Creighton. *Proteins: Structure and Molecular Properties.* W. H. *Freeman & Co*, 1983, 79-86 [0180]
- **Aplin; Wriston.** *CRC Crit. Rev. Biochem*, 1981, 259-306 [0185]
- 65 • **Hakimuddin et al.** *Arch. Biochem. Biophys.*, 1987, vol. 259, 52 [0186]
- **Edge.** *Anal. Biochem*, 1981, vol. 118, 131 [0186]

## ES 2 329 467 T3

- **Thotakura et al.** *Mesh. Enzymol.*, 1987, vol. 138, 350 [0186]
- **Duskin et al.** *J. Biol. Chem.*, 1982, vol. 237, 3105 [0187]
- 5 • **Maxfield et al.** *Polymer*, 1975, vol. 16, 505-509 [0189]
- **Bailey, F. E. et al.** *Nonionic Surfactants*. 1967, 794-821 [0189]
- **Abuchowski A. et al.** *J. Biol. Chem.*, 1977, vol. 252, 3382-3386 [0189]
- 10 • **Abuchowski, A. et al.** *Cancer Blochem. Biophys.*, 1984, vol. 7, 175-186 [0189]
- **Katre, N.V. et al.** *Proc. Natl. Acad Sci*, 1987, vol. 84, 1487-1491 [0189]
- 15 • **Goodson, R.** *Bio Technology*, 1990, vol. 8, 343-346 [0189]
- **Abuchowski, A. et al.** *J. Biol. Chem*, 1977, vol. 252, 3578-3581 [0189]
- **Remington's Pharmaceutical Sciences**. 1980 [0190]
- 20 • **Yeo et al.** *Biotech. and Bioeng.*, 1993, vol. 41, 341-346 [0199]
- **Langer et al.** *J. Biomed Mater. Res.*, 1981, vol. 15, 167-277 [0200]
- 25 • **Langer.** *Chem. Tech.*, 1982, vol. 12, 98-105 [0200]
- **Sidman et al.** *Biopolymers*, 1983, vol. 22, 547-556 [0200]
- **Epstein.** *Proc. Natl. Acad. Sci. USA.*, 1995, vol. 82, 3688-3692 [0201]
- 30 • **Hwang et al.** *Proc. Natl. Acad Sci. USA*, 1980, vol. 77, 4030-4034 [0201]
- **Sliwkowski et al.** *J. Biol. Chem*, 1994, vol. 269, 14661-14665 [0206]
- 35 • **Fendly et al.** *Cancer Research*, 1990, vol. 50, 1550-1558 [0206]
- **Kunkel, T.** *Proc. Natl. Acad Sci. U.S.A.*, 1985, vol. 82, 488 [0207]
- **Gorman et al.** *DNA Prot. Eng. Tech.*, 1990, vol. 2, 3-10 [0207]
- 40 • **Huang et al.** *Nucleic Acids Res*, 1990, vol. 18, 937-947 [0208]
- **Sambrook et al.** *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press*, 1989 [0210]
- 45 • **Chomczynski et al.** *Anal Biochem*, 1987, vol. 162, 156-159 [0213]
- **Kozak.** *Nucleic Acid Research*, 1987, vol. 15, 8125-8148 [0219]
- **Jacobsen et al.** *Biochemistry*, 1996, vol. 35, 3402-17
- 50 • **Lowman et al.** *Biochemistry*, 1991, vol. 30, 10832-10838 [0238].

# REIVINDICACIONES

1. Utilización de un ácido nucleico que codifica un ligando activador de HER2, HER3 y/o HER4, el cual es un polipéptido de heregulina (HRG) o un fragmento del mismo capaz de unirse al receptor HER2, HER3 y/o HER4 en la fabricación de un medicamento para el tratamiento de dificultades respiratorias, la enfermedad pulmonar obstructiva crónica (EPOC), enfisema, enfermedad pulmonar del neonato, lesiones agudas del pulmón, pneumonitis debida a aspiración o radicación, ahogo, fibrosis quística y otras enfermedades traumáticas de las células epiteliales, incluyendo las lesiones asociadas con heridas quirúrgicas y las resecciones, las lesiones, las úlceras y rasgado de tejidos.

2. Utilización según la reivindicación 1, en donde el ligando activador es una HRG humano o un fragmento de la misma.

3. Utilización según la reivindicación 2, en donde la HRG humana o fragmento de la misma es una HRG humano recombinante o fragmento de la misma.

4. Utilización según la reivindicación 1, en donde el ligando activador se selecciona de entre el grupo consistente en la HRG- $\alpha$ , - $\beta$ 1, - $\beta$ 2, tipo- $\beta$ 2 y  $\beta$ 3 y fragmentos de éstas.

5. Utilización según la reivindicación 1, en donde el ligando activador es la  $\gamma$ -HRG o un fragmento de ésta.

6. Utilización según la reivindicación 1 en donde la HRG es rHRG- $\beta$ 1-177-244.

7. Utilización según la reivindicación 1 en donde el ligando activador es una variante de HRG.

8. Utilización según la reivindicación 7 en donde la secuencia de variante de HRG codifica al menos una sustitución aminoacídica en un residuo correspondiente con un residuo de la heregulina- $\beta$ 1 humana nativa de 645 aminoácidos seleccionada a partir de:

S177, H178, L179, V180, K181, E184, E186, K187, T188,  
V191,  
N192, G193, G194, E195, M198, V199, K200, D201, N204,  
P205,  
S206, R207, Y208, L209, K211, P213, N214, E215, T217,  
G218,  
D219, Q222, N223, Y224, S228 o F229.

9. Utilización según la reivindicación 8 en donde la secuencia de la variante HRG codifica al menos una sustitución aminoacídica en un residuo correspondiente con un residuo de la heregulina- $\beta$ 1 humana nativa de 645 aminoácidos seleccionada a partir de P213, N214 y E215.

10. Utilización según la reivindicación 7 en donde la secuencia de variante de la HRG codifica una variante de heregulina- $\beta$ 1, que incluye una o más sustituciones aminoacídicas seleccionadas a partir de:

- a) S177W;
- b) H178 se substituye con S, E, R, o A;
- c) V180 se substituye con Q, I, o E;
- d) K181 se substituye con P o A;
- e) A183G;
- f) E184 se substituye con V, W, K, R, G, o N;
- g) K185 se substituye con E, S, Q, o G;
- h) E186R;
- i) K187 se substituye con E o A;

## ES 2 329 467 T3

j) T188 se substituye con Q o G;

k) E195Q;

l) M198 se substituye con R o K;

m) V199I

n) D201 se substituye con T o I;

o) P205 se substituye con T o Y;

p) S206 se substituye con K, H, G, P o R;

q) R207Y;

r) Y208 se substituye con R o L;

s) L209 se substituye con M o G;

t) K211R;

u) P213 se substituye con S, T, N, o K;

v) N214 se substituye con L, K, S, o E;

w) F216 se substituye con M o Y;

x) N223 se substituye con H o W;

y) M226I

z) F197Y; y

zz K200R

11. Utilización según la reivindicación 7 en donde la secuencia de variante de la HRG codifica un conjunto de substituciones aminoacídicas, y en donde el conjunto de substituciones aminoacídicas se selecciona de un grupo consistente en:

a) A183G, E184W, K185D, E186R, K187E, T188G, y M226I;

b) A183D, E184K, K185S, E186R, K187E, T188G, y M226I;

c) F197Y, M198K, K200R, D201I, y M226I;

d) P205Y, S206G, R207Y, Y208L, y L209M;

e) P205Y, S206R, R207Y, Y208R, L209M, y M226I;

f) P205T, S206H, R207Y, Y208R, y L209M;

g) P205T, S206K, R207Y, Y208R, y L209G;

h) N223W y M226I;

i) N223H y M226I;

j) S177W, H178E, K181P, A183G, E184W, K185D, E186R, K187E, T188G, y M226I;

k) P205Y, S206G, R207Y, Y208L, L209M, y M226I;

l) A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T;

m) A183G, K185E, E186R, K187E, T188G, P205Y, S206G, R207Y, Y208L, y L209M;

## ES 2 329 467 T3

n) A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, y L209M;

o) A183G, K185E, E186R, K187E, T188G, y M226I;

p) F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, y L209M;

q) F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, L209M, y M226I;

r) F197Y, M198R, D201T, y M226I;

s) A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T, y M226I;

t) A183G, K185E, E186R, K187E, T188G, P205Y, S206G, R207Y, Y208L, L209M, y M226I;

u) A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, L209M, y M226I;

v) F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, L209M, N223H, y M226I; y

w) A183G, K185E, E186R, K187E, T188G, F197Y, M198R, D201T, P205Y, S206G, R207Y, Y208L, L209M, N223H, y M226I.

12. Utilización de una célula huésped que comprende un ácido nucleico como se define en cualquiera de las reivindicaciones 1 a 11 en la fabricación de un medicamento para el tratamiento de dificultades respiratorias, la enfermedad pulmonar obstructiva crónica, enfisema, enfermedad pulmonar neonatal, lesiones agudas del pulmón, pneumonitis debida a aspiración o radicación, ahogo, fibrosis quística y otras enfermedades traumáticas de las células epiteliales, incluyendo las lesiones asociadas con heridas quirúrgicas y las resecciones, las lesiones, las úlceras y rasgado de tejidos.

13. Utilización según cualquiera de las reivindicaciones 1 a 12, en donde la dificultad respiratoria es una dificultad respiratoria neonatal o infantil.

14. Utilización según cualquiera de las reivindicaciones 1 a 12, en donde la enfermedad es un subtipo de EPOC seleccionada de entre bronquitis crónica, enfisema y asma.

15. Utilización según cualquiera de las reivindicaciones 1 a 12 en donde la enfermedad pulmonar de neonatos se selecciona de entre el síndrome de dificultad respiratoria neonatal, el síndrome de aspiración del meconio, la enfermedad pulmonar crónica del neonato, la hernia diafrágica congénita.

16. Utilización según cualquiera de las reivindicaciones 1 a 12 en donde la lesión pulmonar aguda es humo o inhalación química.

17. Utilización según cualquiera de las reivindicaciones 1 a 16, en donde la HRG induce la proliferación de células epiteliales.

18. Utilización según la reivindicación 17, en donde las células epiteliales con células epiteliales pulmonares.

19. Ácido nucleico o célula huésped como se define en cualquiera de las reivindicaciones 1-12, 17 y 18, para utilizarse en el tratamiento de una condición seleccionada de entre dificultad respiratoria, enfermedad pulmonar obstructiva crónica (EPOC), enfisema, enfermedad pulmonar neonatal, lesiones agudas del pulmón, pneumonitis debida a aspiración o radicación, ahogo, fibrosis quística y otras enfermedades traumáticas de las células epiteliales, incluyendo las lesiones asociadas con heridas quirúrgicas y las resecciones, las lesiones, las úlceras y rasgado de tejidos.

20. Ácido nucleico o célula huésped según la reivindicación 19, en donde la condición es tal y como se define en cualquiera de las reivindicaciones 13-16.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| GG  | GCG | CGA | GCG | CCT | CAG | CGC | GGC | CGC | TCG | CTC | TCC | CCC | 38  |
| Ala | Arg | Ala | Pro | Gln | Arg | Gly | Arg | Ser | Leu | Ser | Pro |     |     |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     |
| TCG | AGG | GAC | AAA | CTT | TTC | CCA | AAC | CCG | ATC | CGA | GCC | CTT | 77  |
| Ser | Arg | Asp | Lys | Leu | Phe | Pro | Asn | Pro | Ile | Arg | Ala | Leu |     |
|     | 15  |     |     |     |     |     | 20  |     |     |     |     | 25  |     |
| GGA | CCA | AAC | TCG | CCT | GCG | CCG | AGA | GCC | GTC | CGC | GTA | GAG | 116 |
| Gly | Pro | Asn | Ser | Pro | Ala | Pro | Arg | Ala | Val | Arg | Val | Glu |     |
|     |     |     | 30  |     |     |     |     |     | 35  |     |     |     |     |
| CGC | TCC | GTC | TCC | GGC | GAG | ATG | TCC | GAG | CGC | AAA | GAA | GGC | 155 |
| Arg | Ser | Val | Ser | Gly | Glu | Met | Ser | Glu | Arg | Lys | Glu | Gly |     |
|     | 40  |     |     |     |     | 45  |     |     |     |     | 50  |     |     |
| AGA | GCG | AAA | GGG | AAG | GGC | AAG | AAG | AAG | GAG | CGA | GGC | TCC | 194 |
| Arg | Gly | Lys | Gly | Lys | Gly | Lys | Lys | Lys | Glu | Arg | Gly | Ser |     |
|     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |     |
| GGC | AAG | AAG | CCG | GAG | TCC | GCG | GCG | GGC | AGC | CAG | AGC | CCA | 233 |
| Gly | Lys | Lys | Pro | Glu | Ser | Ala | Ala | Gly | Ser | Gln | Ser | Pro |     |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |
| GCC | TTG | CCT | CCC | CGA | TTG | AAA | GAG | ATG | AAA | AGC | CAG | GAA | 272 |
| Ala | Leu | Pro | Pro | Arg | Leu | Lys | Glu | Met | Lys | Ser | Gln | Glu |     |
|     |     | 80  |     |     |     |     | 85  |     |     |     |     | 90  |     |
| TCG | GCT | GCA | GGT | TCC | AAA | CTA | GTC | CTT | CGG | TGT | GAA | ACC | 311 |
| Ser | Ala | Ala | Gly | Ser | Lys | Leu | Val | Leu | Arg | Cys | Glu | Thr |     |
|     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     |
| AGT | TCT | GAA | TAC | TCC | TCT | CTC | AGA | TTC | AAG | TGG | TTC | AAG | 350 |
| Ser | Ser | Glu | Tyr | Ser | Ser | Leu | Arg | Phe | Lys | Trp | Phe | Lys |     |
|     | 105 |     |     |     |     |     | 110 |     |     |     | 115 |     |     |
| AAT | GGG | AAT | GAA | TTG | AAT | CGA | AAA | AAC | AAA | CCA | CAA | AAT | 389 |
| Asn | Gly | Asn | Glu | Leu | Asn | Arg | Lys | Asn | Lys | Pro | Gln | Asn |     |
|     |     |     | 120 |     |     |     |     | 125 |     |     |     |     |     |
| ATC | AAG | ATA | CAA | AAA | AAG | CCA | GGG | AAG | TCA | GAA | CTT | CGC | 428 |
| Ile | Lys | Ile | Gln | Lys | Lys | Pro | Gly | Lys | Ser | Glu | Leu | Arg |     |
| 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |
| ATT | AAC | AAA | GCA | TCA | CTG | GCT | GAT | TCT | GGA | GAG | TAT | ATG | 467 |
| Ile | Asn | Lys | Ala | Ser | Leu | Ala | Asp | Ser | Gly | Glu | Tyr | Met |     |
|     |     | 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |
| TGC | AAA | GTG | ATC | AGC | AAA | TTA | GGA | AAT | GAC | AGT | GCC | TCT | 506 |
| Cys | Lys | Val | Ile | Ser | Lys | Leu | Gly | Asn | Asp | Ser | Ala | Ser |     |
|     |     |     |     | 160 |     |     |     |     | 165 |     |     |     |     |

**FIG. 1A**

```

GCC AAT ATC ACC ATC GTG GAA TCA AAC GAG ATC ATC ACT 545
Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile Thr
170 175 180

GGT ATG CCA GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA 584
Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
185 190

GAG TCT CCC ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA 623
Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala
195 200 205

AAT ACT TCT TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA 662
Asn Thr Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr
210 215 220

AGC CAT CTT GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC 701
Ser His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe
225 230

TGT GTG AAT GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT 740
Cys Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu
235 240 245

TCA AAC CCC TCG AGA TAC TTG TGC AAG TGC CAA CCT GGA 779
Ser Asn Pro Ser Arg Tyr Leu Cys Lys Cys Gln Pro Gly
250 255

TTC ACT GGA GCA AGA TGT ACT GAG AAT GTG CCC ATG AAA 818
Phe Thr Gly Ala Arg Cys Thr Glu Asn Val Pro Met Lys
260 265 270

GTC CAA AAC CAA GAA AAG GCG GAG GAG CTG TAC CAG AAG 857
Val Gln Asn Gln Glu Lys Ala Glu Glu Leu Tyr Gln Lys
275 280 285

AGA GTG CTG ACC ATA ACC GGC ATC TGC ATC GCC CTC CTT 896
Arg Val Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu
290 295

GTG GTC GGC ATC ATG TGT GTG GTG GCC TAC TGC AAA ACC 935
Val Val Gly Ile Met Cys Val Val Ala Tyr Cys Lys Thr
300 305 310

AAG AAA CAG CGG AAA AAG CTG CAT GAC CGT CTT CGG CAG 974
Lys Lys Gln Arg Lys Lys Leu His Asp Arg Leu Arg Gln
315 320

AGC CTT CGG TCT GAA CGA AAC AAT ATG ATG AAC ATT GCC 1013
Ser Leu Arg Ser Glu Arg Asn Asn Met Met Asn Ile Ala
325 330 335

```

**FIG.\_1B**

|                                                          |                                                         |
|----------------------------------------------------------|---------------------------------------------------------|
| AAT GGG CCT CAC CAT CCT AAC CCA CCC CCC GAG AAT GTC 1052 | Asn Gly Pro His His Pro Asn Pro Pro Pro Glu Asn Val 350 |
| CAG CTG GTG AAT CAA TAC GTA TCT AAA AAC GTC ATC TCC 1091 | Gln Leu Val Asn Gln Tyr Val Ser Lys Asn Val Ile Ser 360 |
| AGT GAG CAT ATT GTT GAG AGA GAA GCA GAG ACA TCC TTT 1130 | Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser Phe 375 |
| TCC ACC AGT CAC TAT ACT TCC ACA GCC CAT CAC TCC ACT 1169 | Ser Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr 385 |
| ACT GTC ACC CAG ACT CCT AGC CAC AGC TGG AGC AAC GGA 1208 | Thr Val Thr Gln Thr Pro Ser His Ser Trp Ser Asn Gly 400 |
| CAC ACT GAA AGC ATC CTT TCC GAA AGC CAC TCT GTA ATC 1247 | His Thr Glu Ser Ile Leu Ser Glu Ser His Ser Val Ile 415 |
| GTG ATG TCA TCC GTA GAA AAC AGT AGG CAC AGC AGC CCA 1286 | Val Met Ser Ser Val Glu Asn Ser Arg His Ser Ser Pro 425 |
| ACT GGG GGC CCA AGA GGA CGT CTT AAT GGC ACA GGA GGC 1325 | Thr Gly Gly Pro Arg Gly Arg Leu Asn Gly Thr Gly Gly 440 |
| CCT CGT GAA TGT AAC AGC TTC CTC AGG CAT GCC AGA GAA 1364 | Pro Arg Glu Cys Asn Ser Phe Leu Arg His Ala Arg Glu 450 |
| ACC CCT GAT TCC TAC CGA GAC TCT CCT CAT AGT GAA AGG 1403 | Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser Glu Arg 465 |
| TAT GTG TCA GCC ATG ACC ACC CCG GCT CGT ATG TCA CCT 1442 | Tyr Val Ser Ala Met Thr Thr Pro Ala Arg Met Ser Pro 480 |
| GTA GAT TTC CAC ACG CCA AGC TCC CCC AAA TCG CCC CCT 1481 | Val Asp Phe His Thr Pro Ser Ser Pro Lys Ser Pro Pro 490 |
| TCG GAA ATG TCT CCA CCC GTG TCC AGC ATG ACG GTG TCC 1520 | Ser Glu Met Ser Pro Pro Val Ser Ser Met Thr Val Ser 505 |

**FIG. 1C**

|                                                          |         |
|----------------------------------------------------------|---------|
| ATG CCT TCC ATG GCG GTC AGC CCC TTC ATG GAA GAA GAG 1559 |         |
| Met Pro Ser Met Ala Val Ser Pro Phe Met Glu Glu Glu      |         |
| 510                                                      | 515     |
| AGA CCT CTA CTT CTC GTG ACA CCA CCA AGG CTG CGG GAG 1598 |         |
| Arg Pro Leu Leu Leu Val Thr Pro Pro Arg Leu Arg Glu      |         |
| 520                                                      | 525 530 |
| AAG AAG TTT GAC CAT CAC CCT CAG CAG TTC AGC TCC TTC 1637 |         |
| Lys Lys Phe Asp His His Pro Gln Gln Phe Ser Ser Phe      |         |
| 535                                                      | 540 545 |
| CAC CAC AAC CCC GCG CAT GAC AGT AAC AGC CTC CCT GCT 1676 |         |
| His His Asn Pro Ala His Asp Ser Asn Ser Leu Pro Ala      |         |
| 550                                                      | 555     |
| AGC CCC TTG AGG ATA GTG GAG GAT GAG GAG TAT GAA ACG 1715 |         |
| Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr Glu Thr      |         |
| 560                                                      | 565 570 |
| ACC CAA GAG TAC GAG CCA GCC CAA GAG CCT GTT AAG AAA 1754 |         |
| Thr Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys      |         |
| 575                                                      | 580     |
| CTC GCC AAT AGC CGG CGG GCC AAA AGA ACC AAG CCC AAT 1793 |         |
| Leu Ala Asn Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn      |         |
| 585                                                      | 590 595 |
| GGC CAC ATT GCT AAC AGA TTG GAA GTG GAC AGC AAC ACA 1832 |         |
| Gly His Ile Ala Asn Arg Leu Glu Val Asp Ser Asn Thr      |         |
| 600                                                      | 605 610 |
| AGC TCC CAG AGC AGT AAC TCA GAG AGT GAA ACA GAA GAT 1871 |         |
| Ser Ser Gln Ser Ser Asn Ser Glu Ser Glu Thr Glu Asp      |         |
| 615                                                      | 620     |
| GAA AGA GTA GGT GAA GAT ACG CCT TTC CTG GGC ATA CAG 1910 |         |
| Glu Arg Val Gly Glu Asp Thr Pro Phe Leu Gly Ile Gln      |         |
| 625                                                      | 630 635 |
| AAC CCC CTG GCA GCC AGT CTT GAG GCA ACA CCT GCC TTC 1949 |         |
| Asn Pro Leu Ala Ala Ser Leu Glu Ala Thr Pro Ala Phe      |         |
| 640                                                      | 645     |
| CGC CTG GCT GAC AGC AGG ACT AAC CCA GCA GGC CGC TTC 1988 |         |
| Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala Gly Arg Phe      |         |
| 650                                                      | 655 660 |
| TCG ACA CAG GAA GAA ATC CAG G 2010                       |         |
| Ser Thr Gln Glu Glu Ile Gln                              |         |
| 665                                                      | 669     |

**FIG. 1D**

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| GG  | GAC | AAA | CTT | TTC | CCA | AAC | CCG | ATC | CGA | GCC | CTT | GGA | 38  |
|     | Asp | Lys | Leu | Phe | Pro | Asn | Pro | Ile | Arg | Ala | Leu | Gly |     |
|     | 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |
| CCA | AAC | TCG | CCT | GCG | CCG | AGA | GCC | GTC | CGC | GTA | GAG | CGC | 77  |
| Pro | Asn | Ser | Pro | Ala | Pro | Arg | Ala | Val | Arg | Val | Glu | Arg |     |
|     |     | 15  |     |     |     | 20  |     |     |     |     |     | 25  |     |
| TCC | GTC | TCC | GGC | GAG | ATG | TCC | GAG | CGC | AAA | GAA | GGC | AGA | 116 |
| Ser | Val | Ser | Gly | Glu | Met | Ser | Glu | Arg | Lys | Glu | Gly | Arg |     |
|     |     |     | 30  |     |     |     |     |     | 35  |     |     |     |     |
| GGC | AAA | GGG | AAG | GGC | AAG | AAG | AAG | GAG | CGA | GGC | TCC | GGC | 155 |
| Gly | Lys | Gly | Lys | Gly | Lys | Lys | Lys | Glu | Arg | Gly | Ser | Gly |     |
|     | 40  |     |     |     | 45  |     |     |     |     |     | 50  |     |     |
| AAG | AAG | CCG | GAG | TCC | GCG | GCG | GCG | AGC | CAG | AGC | CCA | GCC | 194 |
| Lys | Lys | Pro | Glu | Ser | Ala | Ala | Gly | Ser | Gln | Ser | Pro | Ala |     |
|     |     | 55  |     |     |     |     |     | 60  |     |     |     |     |     |
| TTG | CCT | CCC | CAA | TTG | AAA | GAG | ATG | AAA | AGC | CAG | GAA | TCG | 233 |
| Leu | Pro | Pro | Gln | Leu | Lys | Glu | Met | Lys | Ser | Gln | Glu | Ser |     |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |
| GCT | GCA | GGT | TCC | AAA | CTA | GTC | CTT | CGG | TGT | GAA | ACC | AGT | 272 |
| Ala | Ala | Gly | Ser | Lys | Leu | Val | Leu | Arg | Cys | Glu | Thr | Ser |     |
|     |     | 80  |     |     |     |     | 85  |     |     |     |     | 90  |     |
| TCT | GAA | TAC | TCC | TCT | CTC | AGA | TTC | AAG | TGG | TTC | AAG | AAT | 311 |
| Ser | Glu | Tyr | Ser | Ser | Leu | Arg | Phe | Lys | Trp | Phe | Lys | Asn |     |
|     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     |
| GGG | AAT | GAA | TTG | AAT | CGA | AAA | AAC | AAA | CCA | CAA | AAT | ATC | 350 |
| Gly | Asn | Glu | Leu | Asn | Arg | Lys | Asn | Lys | Pro | Gln | Asn | Ile |     |
|     | 105 |     |     |     |     | 110 |     |     |     |     | 115 |     |     |
| AAG | ATA | CAA | AAA | AAG | CCA | GGG | AAG | TCA | GAA | CTT | CGC | ATT | 389 |
| Lys | Ile | Gln | Lys | Lys | Pro | Gly | Lys | Ser | Glu | Leu | Arg | Ile |     |
|     |     |     | 120 |     |     |     |     | 125 |     |     |     |     |     |
| AAC | AAA | GCA | TCA | CTG | GCT | GAT | TCT | GGA | GAG | TAT | ATG | TGC | 428 |
| Asn | Lys | Ala | Ser | Leu | Ala | Asp | Ser | Gly | Glu | Tyr | Met | Cys |     |
| 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |
| AAA | GTG | ATC | AGC | AAA | TTA | GGA | AAT | GAC | AGT | GCC | TCT | GCC | 467 |
| Lys | Val | Ile | Ser | Lys | Leu | Gly | Asn | Asp | Ser | Ala | Ser | Ala |     |
|     |     | 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |
| AAT | ATC | ACC | ATC | GTG | GAA | TCA | AAC | GAG | ATC | ATC | ACT | GGT | 506 |
| Asn | Ile | Thr | Ile | Val | Glu | Ser | Asn | Glu | Ile | Ile | Thr | Gly |     |
|     |     |     |     | 160 |     |     |     |     | 165 |     |     |     |     |

FIG. 2A

```

ATG CCA GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA GAG 545
Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu
170 175 180

TCT CCC ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA AAT 584
Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn
185 190

ACT TCT TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA AGC 623
Thr Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser
195 200 205

CAT CTT GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC TGT 662
His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys
210 215 220

GTG AAT GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT TCA 701
Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser
225 230

AAC CCC TCG AGA TAC TTG TGC AAG TGC CCA AAT GAG TTT 740
Asn Pro Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe
235 240 245

ACT GGT GAT CGC TGC CAA AAC TAC GTA ATG GCC AGC TTC 779
Thr Gly Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe
250 255

TAC AAG CAT CTT GGG ATT GAA TTT ATG GAG GCG GAG GAG 818
Tyr Lys His Leu Gly Ile Glu Phe Met Glu Ala Glu Glu
260 265 270

CTG TAC CAG AAG AGA GTG CTG ACC ATA ACC GGC ATC TGC 857
Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys
275 280 285

ATC GCC CTC CTT GTG GTC GGC ATC ATG TGT GTG GTG GCC 896
Ile Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala
290 295

TAC TGC AAA ACC AAG AAA CAG CGG AAA AAG CTG CAT GAC 935
Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp
300 305 310

CGT CTT CGG CAG AGC CTT CGG TCT GAA CGA AAC AAT ATG 974
Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn Met
315 320

ATG AAC ATT GCC AAT GGG CCT CAC CAT CCT AAC CCA CCC 1013
Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro
325 330 335

```

**FIG.\_2B**

|     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| CCC | GAG | AAT | GTC | CAG | CTG | GTG | AAT | CAA | TAC | GTA | TCT | AAA | 1052 |
| Pro | Glu | Asn | Val | Gln | Leu | Val | Asn | Gln | Tyr | Val | Ser | Lys |      |
|     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |      |
| AAC | GTC | ATC | TCC | AGT | GAG | CAT | ATT | GTT | GAG | AGA | GAA | GCA | 1091 |
| Asn | Val | Ile | Ser | Ser | Glu | His | Ile | Val | Glu | Arg | Glu | Ala |      |
|     |     |     | 355 |     |     |     |     | 360 |     |     |     |     |      |
| GAG | ACA | TCC | TTT | TCC | ACC | AGT | CAC | TAT | ACT | TCC | ACA | GCC | 1130 |
| Glu | Thr | Ser | Phe | Ser | Thr | Ser | His | Tyr | Thr | Ser | Thr | Ala |      |
|     | 365 |     |     |     |     | 370 |     |     |     |     | 375 |     |      |
| CAT | CAC | TCC | ACT | ACT | GTC | ACC | CAG | ACT | CCT | AGC | CAC | AGC | 1169 |
| His | His | Ser | Thr | Thr | Val | Thr | Gln | Thr | Pro | Ser | His | Ser |      |
|     |     |     | 380 |     |     |     |     | 385 |     |     |     |     |      |
| TGG | AGC | AAC | GGA | CAC | ACT | GAA | AGC | ATC | CTT | TCC | GAA | AGC | 1208 |
| Trp | Ser | Asn | Gly | His | Thr | Glu | Ser | Ile | Leu | Ser | Glu | Ser |      |
| 390 |     |     |     |     | 395 |     |     |     | 400 |     |     |     |      |
| CAC | TCT | GTA | ATC | GTG | ATG | TCA | TCC | GTA | GAA | AAC | AGT | AGG | 1247 |
| His | Ser | Val | Ile | Val | Met | Ser | Ser | Val | Glu | Asn | Ser | Arg |      |
|     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |      |
| CAC | AGC | AGC | CCA | ACT | GGG | GGC | CCA | AGA | GGA | CGT | CTT | AAT | 1286 |
| His | Ser | Ser | Pro | Thr | Gly | Gly | Pro | Arg | Gly | Arg | Leu | Asn |      |
|     |     |     |     | 420 |     |     |     |     | 425 |     |     |     |      |
| GGC | ACA | GGA | GGC | CCT | CGT | GAA | TGT | AAC | AGC | TTC | CTC | AGG | 1325 |
| Gly | Thr | Gly | Gly | Pro | Arg | Glu | Cys | Asn | Ser | Phe | Leu | Arg |      |
|     | 430 |     |     |     |     | 435 |     |     |     |     | 440 |     |      |
| CAT | GCC | AGA | GAA | ACC | CCT | GAT | TCC | TAC | CGA | GAC | TCT | CCT | 1364 |
| His | Ala | Arg | Glu | Thr | Pro | Asp | Ser | Tyr | Arg | Asp | Ser | Pro |      |
|     |     |     | 445 |     |     |     |     | 450 |     |     |     |     |      |
| CAT | AGT | GAA | AGG | TAT | GTG | TCA | GCC | ATG | ACC | ACC | CCG | GCT | 1403 |
| His | Ser | Glu | Arg | Tyr | Val | Ser | Ala | Met | Thr | Thr | Pro | Ala |      |
| 455 |     |     |     |     | 460 |     |     |     |     | 465 |     |     |      |
| CGT | ATG | TCA | CCT | GTA | GAT | TTC | CAC | ACG | CCA | AGC | TCC | CCC | 1442 |
| Arg | Met | Ser | Pro | Val | Asp | Phe | His | Thr | Pro | Ser | Ser | Pro |      |
|     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |      |
| AAA | TCG | CCC | CCT | TCG | GAA | ATG | TCT | CCA | CCC | GTG | TCC | AGC | 1481 |
| Lys | Ser | Pro | Pro | Ser | Glu | Met | Ser | Pro | Pro | Val | Ser | Ser |      |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |      |
| ATG | ACG | GTG | TCC | ATG | CCT | TCC | ATG | GCG | GTC | AGC | CCC | TTC | 1520 |
| Met | Thr | Val | Ser | Met | Pro | Ser | Met | Ala | Val | Ser | Pro | Phe |      |
|     | 495 |     |     |     |     | 500 |     |     |     |     | 505 |     |      |

FIG.\_2C

|     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| ATG | GAA | GAA | GAG | AGA | CCT | CTA | CTT | CTC | GTG | ACA | CCA | CCA | 1559 |
| Met | Glu | Glu | Glu | Arg | Pro | Leu | Leu | Leu | Val | Thr | Pro | Pro |      |
|     |     |     | 510 |     |     |     |     | 515 |     |     |     |     |      |
| AGG | CTG | CGG | GAG | AAG | AAG | TTT | GAC | CAT | CAC | CCT | CAG | CAG | 1598 |
| Arg | Leu | Arg | Glu | Lys | Lys | Phe | Asp | His | His | Pro | Gln | Gln |      |
| 520 |     |     |     |     | 525 |     |     |     |     | 530 |     |     |      |
| TTC | AGC | TCC | TTC | CAC | CAC | AAC | CCC | GCG | CAT | GAC | AGT | AAC | 1637 |
| Phe | Ser | Ser | Phe | His | His | Asn | Pro | Ala | His | Asp | Ser | Asn |      |
|     |     | 535 |     |     |     |     | 540 |     |     |     |     | 545 |      |
| AGC | CTC | CCT | GCT | AGC | CCC | TTG | AGG | ATA | GTG | GAG | GAT | GAG | 1676 |
| Ser | Leu | Pro | Ala | Ser | Pro | Leu | Arg | Ile | Val | Glu | Asp | Glu |      |
|     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |      |
| GAG | TAT | GAA | ACG | ACC | CAA | GAG | TAC | GAG | CCA | GCC | CAA | GAG | 1715 |
| Glu | Tyr | Glu | Thr | Thr | Gln | Glu | Tyr | Glu | Pro | Ala | Gln | Glu |      |
|     | 560 |     |     |     |     | 565 |     |     |     |     | 570 |     |      |
| CCT | GTT | AAG | AAA | CTC | GCC | AAT | AGC | CGG | CGG | GCC | AAA | AGA | 1754 |
| Pro | Val | Lys | Lys | Leu | Ala | Asn | Ser | Arg | Arg | Ala | Lys | Arg |      |
|     |     |     | 575 |     |     |     |     | 580 |     |     |     |     |      |
| ACC | AAG | CCC | AAT | GGC | CAC | ATT | GCT | AAC | AGA | TTG | GAA | GTG | 1793 |
| Thr | Lys | Pro | Asn | Gly | His | Ile | Ala | Asn | Arg | Leu | Glu | Val |      |
| 585 |     |     |     |     | 590 |     |     |     |     | 595 |     |     |      |
| GAC | AGC | AAC | ACA | AGC | TCC | CAG | AGC | AGT | AAC | TCA | GAG | AGT | 1832 |
| Asp | Ser | Asn | Thr | Ser | Ser | Gln | Ser | Ser | Asn | Ser | Glu | Ser |      |
|     |     | 600 |     |     |     |     | 605 |     |     |     |     | 610 |      |

**FIG..2D**

|                                                         |         |
|---------------------------------------------------------|---------|
| GAA ACA GAA GAT GAA AGA GTA GGT GAA GAT ACG CCT TTC     | 1871    |
| Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr Pro Phe     |         |
| 615                                                     | 620     |
| CTG GGC ATA CAG AAC CCC CTG GCA GCC AGT CTT GAG GCA     | 1910    |
| Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala     |         |
| 625                                                     | 630 635 |
| ACA CCT GCC TTC CGC CTG GCT GAC AGC AGG ACT AAC CCA     | 1949    |
| Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro     |         |
| 640                                                     | 645     |
| GCA GGC CGC TTC TCG ACA CAG GAA GAA ATC CAG GCC AGG     | 1988    |
| Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile Gln Ala Arg     |         |
| 650                                                     | 655 660 |
| CTG TCT AGT GTA ATT GCT AAC CAA GAC CCT ATT GCT GTA TA  | 2029    |
| Leu Ser Ser Val Ile Ala Asn Gln Asp Pro Ile Ala Val     |         |
| 665                                                     | 670 675 |
| A AACCTAAATA AACACATAGA TTCACCTGTA AAACCTTTATT          | 2070    |
| TTATATAATA AAGTATTCCA CCTTAAATTA AACAAATTTAT TTTATTTTAG | 2120    |
| CAGTTCTGCA AATAGAAAAC AGGAAAAAAA CTTTTATAAA TTAAATATAT  | 2170    |
| GTATGTAAAA ATGAAAAAAA AAAAAAAA                          | 2199    |

**FIG.\_2E**

GTGGCTGCGG GGCAATTGAA AAAGAGCCGG CGAGGAGTTC CCGAAACTT 50  
 GTTGAACTC CGGGCTCGCG CGGAGGCCAG GAGCTGAGCG GCGGCGGCTG 100  
 CCGACGATG GGAGCGTGAG CAGGACGGTG ATAACCTCTC CCGATCGGG 150  
 TTGCGAGGGC GCCGGGCAGA GGCCAGGACG CGAGCCGCCA GCGGCGGGAC 200  
 CCATCGACGA CTTCCCGGGG CGACAGGAGC AGCCCCGAGA GCCAGGGCGA 250  
 GCGCCCGTTC CAGGTGGCCG GACCGCCCGC CGCGTCCGCG CCGCGCTCCC 300  
 TGCAGGCAAC GGGAGACGCC CCCGCGCAGC GCGAGCGCCT CAGCGCGGCC 350  
 GCTCGCTCTC CCCATCGAGG GACAACTTT TCCCAAACCC GATCCGAGCC 400  
 CTTGGACCAA ACTCGCCTGC GCCGAGAGCC GTCCGCGTAG AGCGCTCCGT 450  
 CTCCGGCGAG ATG TCC GAG CGC AAA GAA GGC AGA GGC AAA 490  
 Met Ser Glu Arg Lys Glu Gly Arg Gly Lys  
 1 5 10  
 GGG AAG GGC AAG AAG AAG GAG CGA GGC TCC GGC AAG AAG 529  
 Gly Lys Gly Lys Lys Lys Glu Arg Gly Ser Gly Lys Lys  
 15 20  
 CCG GAG TCC GCG GCG GGC AGC CAG AGC CCA GCC TTG CCT 568  
 Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro Ala Leu Pro  
 25 30 35  
 CCC CAA TTG AAA GAG ATG AAA AGC CAG GAA TCG GCT GCA 607  
 Pro Gln Leu Lys Glu Met Lys Ser Gln Glu Ser Ala Ala  
 40 45  
 GGT TCC AAA CTA GTC CTT CGG TGT GAA ACC AGT TCT GAA 646  
 Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser Ser Glu  
 50 55 60  
 TAC TCC TCT CTC AGA TTC AAG TGG TTC AAG AAT GGG AAT 685  
 Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn  
 65 70 75  
 GAA TTG AAT CGA AAA AAC AAA CCA CAA AAT ATC AAG ATA 724  
 Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile  
 80 85

**FIG. 3A**

# ES 2 329 467 T3

|                                                          |  |
|----------------------------------------------------------|--|
| CAA AAA AAG CCA GGG AAG TCA GAA CTT CGC ATT AAC AAA 763  |  |
| Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys      |  |
| 90 95 100                                                |  |
| GCA TCA CTG GCT GAT TCT GGA GAG TAT ATG TGC AAA GTG 802  |  |
| Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys Lys Val      |  |
| 105 110                                                  |  |
| ATC AGC AAA TTA GGA AAT GAC AGT GCC TCT GCC AAT ATC 841  |  |
| Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala Asn Ile      |  |
| 115 120 125                                              |  |
| ACC ATC GTG GAA TCA AAC GAG ATC ATC ACT GGT ATG CCA 880  |  |
| Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly Met Pro      |  |
| 130 135 140                                              |  |
| GCC TCA ACT GAA GGA GCA TAT GTG TCT TCA GAG TCT CCC 919  |  |
| Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu Ser Pro      |  |
| 145 150                                                  |  |
| ATT AGA ATA TCA GTA TCC ACA GAA GGA GCA AAT ACT TCT 958  |  |
| Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr Ser      |  |
| 155 160 165                                              |  |
| TCA TCT ACA TCT ACA TCC ACC ACT GGG ACA AGC CAT CTT 997  |  |
| Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu      |  |
| 170 175                                                  |  |
| GTA AAA TGT GCG GAG AAG GAG AAA ACT TTC TGT GTG AAT 1036 |  |
| Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn      |  |
| 180 185 190                                              |  |
| GGA GGG GAG TGC TTC ATG GTG AAA GAC CTT TCA AAC CCC 1075 |  |
| Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn Pro      |  |
| 195 200 205                                              |  |
| TCG AGA TAC TTG TGC AAG TGC CCA AAT GAG TTT ACT GGT 1114 |  |
| Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe Thr Gly      |  |
| 210 215                                                  |  |
| GAT CGC TGC CAA AAC TAC GTA ATG GCC AGC TTC TAC AAG 1153 |  |
| Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe Tyr Lys      |  |
| 220 225 230                                              |  |
| GCG GAG GAG CTG TAC CAG AAG AGA GTG CTG ACC ATA ACC 1192 |  |
| Ala Glu Glu Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr      |  |
| 235 240                                                  |  |
| GGC ATC TGC ATC GCC CTC CTT GTG GTC GGC ATC ATG TGT 1231 |  |
| Gly Ile Cys Ile Ala Leu Leu Val Val Gly Ile Met Cys      |  |
| 245 250 255                                              |  |
| GTG GTG GCC TAC TGC AAA ACC AAG AAA CAG CGG AAA AAG 1270 |  |
| Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys      |  |
| 260 265 270                                              |  |

**FIG.\_3B**

|                                                     |         |
|-----------------------------------------------------|---------|
| CTG CAT GAC CGT CTT CGG CAG AGC CTT CGG TCT GAA CGA | 1309    |
| Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg |         |
| 275                                                 | 280     |
| AAC AAT ATG ATG AAC ATT GCC AAT GGG CCT CAC CAT OCT | 1348    |
| Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro |         |
| 285                                                 | 290 295 |
| AAC CCA CCC CCC GAG AAT GTC CAG CTG GTG AAT CAA TAC | 1387    |
| Asn Pro Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr |         |
| 300                                                 | 305     |
| GTA TCT AAA AAC GTC ATC TCC AGT GAG CAT ATT GTT GAG | 1426    |
| Val Ser Lys Asn Val Ile Ser Ser Glu His Ile Val Glu |         |
| 310                                                 | 315 320 |
| AGA GAA GCA GAG ACA TCC TTT TCC ACC AGT CAC TAT ACT | 1465    |
| Arg Glu Ala Glu Thr Ser Phe Ser Thr Ser His Tyr Thr |         |
| 325                                                 | 330 335 |
| TCC ACA GCC CAT CAC TCC ACT ACT GTC ACC CAG ACT OCT | 1504    |
| Ser Thr Ala His His Ser Thr Thr Val Thr Gln Thr Pro |         |
| 340                                                 | 345     |
| AGC CAC AGC TGG AGC AAC GGA CAC ACT GAA AGC ATC CTT | 1543    |
| Ser His Ser Trp Ser Asn Gly His Thr Glu Ser Ile Leu |         |
| 350                                                 | 355 360 |
| TCC GAA AGC CAC TCT GTA ATC GTG ATG TCA TCC GTA GAA | 1582    |
| Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val Glu |         |
| 365                                                 | 370     |
| AAC AGT AGG CAC AGC AGC CCA ACT GGG GGC CCA AGA GGA | 1621    |
| Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly |         |
| 375                                                 | 380 385 |
| CGT CTT AAT GGC ACA GGA GGC CCT CGT GAA TGT AAC AGC | 1660    |
| Arg Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser |         |
| 390                                                 | 395 400 |
| TTC CTC AGG CAT GCC AGA GAA ACC CCT GAT TCC TAC CGA | 1699    |
| Phe Leu Arg His Ala Arg Glu Thr Pro Asp Ser Tyr Arg |         |
| 405                                                 | 410     |
| GAC TCT CCT CAT AGT GAA AGG TAT GTG TCA GCC ATG ACC | 1738    |
| Asp Ser Pro His Ser Glu Arg Tyr Val Ser Ala Met Thr |         |
| 415                                                 | 420 425 |
| ACC CCG GCT CGT ATG TCA OCT GTA GAT TTC CAC ACG CCA | 1777    |
| Thr Pro Ala Arg Met Ser Pro Val Asp Phe His Thr Pro |         |
| 430                                                 | 435     |
| AGC TCC CCC AAA TCG CCC CCT TCG GAA ATG TCT CCA CCC | 1816    |
| Ser Ser Pro Lys Ser Pro Pro Ser Glu Met Ser Pro Pro |         |
| 440                                                 | 445 450 |

**FIG. 3C**

GTG TCC AGC ATG ACG GTG TCC AAG CCT TCC ATG GCG GTC 1855  
 Val Ser Ser Met Thr Val Ser Lys Pro Ser Met Ala Val  
 455 460 465

AGC CCC TTC ATG GAA GAA GAG AGA OCT CTA CTT CTC GTG 1894  
 Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu Val  
 470 475

ACA CCA CCA AGG CTG CCG GAG AAG AAG TTT GAC CAT CAC 1933  
 Thr Pro Pro Arg Leu Arg Glu Lys Lys Phe Asp His His  
 480 485 490

CCT CAG CAG TTC AGC TCC TTC CAC CAC AAC CCC GCG CAT 1972  
 Pro Gln Gln Phe Ser Ser Phe His His Asn Pro Ala His  
 495 500

GAC AGT AAC AGC CTC CCT GCT AGC CCC TTG AGG ATA GTG 2011  
 Asp Ser Asn Ser Leu Pro Ala Ser Pro Leu Arg Ile Val  
 505 510 515

GAG GAT GAG GAG TAT GAA ACG ACC CAA GAG TAC GAG CCA 2050  
 Glu Asp Glu Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro  
 520 525 530

GCC CAA GAG CCT GTT AAG AAA CTC GCC AAT AGC CCG CGG 2089  
 Ala Gln Glu Pro Val Lys Lys Leu Ala Asn Ser Arg Arg  
 535 540

GCC AAA AGA ACC AAG CCC AAT GGC CAC ATT GCT AAC AGA 2128  
 Ala Lys Arg Thr Lys Pro Asn Gly His Ile Ala Asn Arg  
 545 550 555

TTG GAA GTG GAC AGC AAC ACA AGC TCC CAG AGC AGT AAC 2167  
 Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn  
 560 565

TCA GAG AGT GAA ACA GAA GAT GAA AGA GTA GGT GAA GAT 2206  
 Ser Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp  
 570 575 580

ACG CCT TTC CTG GGC ATA CAG AAC CCC CTG GCA GCC AGT 2245  
 Thr Pro Phe Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser  
 585 590 595

CTT GAG GCA ACA CCT GGC TTC CGC CTG GCT GAC AGC AGG 2284  
 Leu Glu Ala Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg  
 600 605

ACT AAC CCA GCA GGC CGC TTC TCG ACA CAG GAA GAA ATC 2323  
 Thr Asn Pro Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile  
 610 615 620

CAG GCC AGG CTG TCT AGT GTA ATT GCT AAC CAA GAC CCT 2362  
 Gln Ala Arg Leu Ser Ser Val Ile Ala Asn Gln Asp Pro  
 625 630

**FIG.\_3D**

ATT GCT GTA TAAACCTA AATAACACA TAGATTCACC TGTAAACTT 2410  
Ile Ala Val  
635 637

TATTTTATAT AATAAGTAT TCCACCTTAA ATTAAACAAT TTATTTTATT 2460

TTAGCAGTTC TGCAATATAA AAAAAAAAAA 2490

***FIG.\_3E***

GCGCCTGCCT CCAACCTGCG GGCGGGAGGT GGGTGGCTGC GGGGCAATTG 50  
 AAAAAGAGCC GGCGAGGAGT TCCCGAAAC TTGTTGGAAC TCCGGGCTCG 100  
 CGCGGAGGCC AGGAGCTGAG CGGCGGCGGC TGCCGGACGA TGGGAGCGTG 150  
 AGCAGGACGG TGATAACCTC TCCCGATCG GGTTCGAGG GCGCCGGGCA 200  
 GAGGCCAGGA CGCGAGCCGC CAGCGGCGGG ACCCATCGAC GACTTCCCGG 250  
 GGCGACAGGA GCAGCCCCGA GAGCCAGGGC GAGCGCCCGT TCCAGGTGGC 300  
 CGGACCGCCC GCCGCGTCCG CGCCGCGCTC CCTGCAGGCA ACGGGAGACG 350  
 CCCCCGCGCA GCGCGAGCGC CTCAGCGCGG CCGCTCGCTC TCCCCATCGA 400  
 GGGACAAACT TTTCCCAAAC CCGATCCGAG CCCTTGGACC AAACCTGCCT 450  
 GCGCCGAGAG CCGTCCGCGT AGAGCGCTCC GTCTCCGGCG AG ATG 495  
 Met  
 1  
 TCC GAG CGC AAA GAA GGC AGA GGC AAA GGG AAG GGC AAG 534  
 Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys  
 5 10  
 AAG AAG GAG CGA GGC TCC GGC AAG AAG CGG GAG TCC GCG 573  
 Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala  
 15 20 25  
 GCG GGC AGC CAG AGC CCA GGC TTG CCT CCC CAA TTG AAA 612  
 Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys  
 30 35 40  
 GAG ATG AAA AGC CAG GAA TCG GCT GCA GGT TCC AAA CTA 651  
 Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu  
 45 50  
 GTC CTT CGG TGT GAA ACC AGT TCT GAA TAC TCC TCT CTC 690  
 Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu  
 55 60 65  
 AGA TTC AAG TGG TTC AAG AAT GGG AAT GAA TTG AAT CGA 729  
 Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg  
 70 75

**FIG.\_4A**

AAA AAC AAA CCA CAA AAT ATC AAG ATA CAA AAA AAG CCA 768  
 Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro  
 80 85 90

GGG AAG TCA GAA CTT CGC ATT AAC AAA GCA TCA CTG GCT 807  
 Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala  
 95 100 105

GAT TCT GGA GAG TAT ATG TGC AAA GTG ATC AGC AAA TTA 846  
 Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu  
 110 115

GGA AAT GAC AGT GCC TCT GCC AAT ATC ACC ATC GTG GAA 885  
 Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu  
 120 125 130

TCA AAC GAG ATC ATC ACT GGT ATG CCA GCC TCA ACT GAA 924  
 Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu  
 135 140

GGA GCA TAT GTG TCT TCA GAG TCT CCC ATT AGA ATA TCA 963  
 Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser  
 145 150 155

GTA TCC ACA GAA GGA GCA AAT ACT TCT TCA TCT ACA TCT 1002  
 Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser  
 160 165 170

ACA TCC ACC ACT GGG ACA AGC CAT CTT GTA AAA TGT GCG 1041  
 Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala  
 175 180

GAG AAG GAG AAA ACT TTC TGT GTG AAT GGA GGG GAG TGC 1080  
 Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys  
 185 190 195

TTC ATG GTG AAA GAC CTT TCA AAC CCC TCG AGA TAC TTG 1119  
 Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu  
 200 205

TGC AAG TGC CCA AAT GAG TTT ACT GGT GAT CGC TGC CAA 1158  
 Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln  
 210 215 220

AAC TAC GTA ATG GCC AGC TTC TAC AGT ACG TCC ACT CCC 1197  
 Asn Tyr Val Met Ala Ser Phe Tyr Ser Thr Ser Thr Pro  
 225 230 235

TTT CTG TCT CTG CCT GAA TAGGA GCATGCTCAG TTGGTGCTGC 1240  
 Phe Leu Ser Leu Pro Glu  
 240 241

TTTCTTGTG CTGCACTCTCC CCTCAGATTC CACCTAGAGC TAGATGTGTC 1290

**FIG.\_4B**

TTACCAGATC TAATATTGAC TGCCTCTGCC TGTGCGCATGA GAACATTAAC 1340  
AAAAGCAATT GTATTACTTC CTCGTTCGC GACTAGTTGG CTCTGAGATA 1390  
CTAATAGGTG TGTGAGGCTC CGGATGTTTC TGGAATTGAT ATTGAATGAT 1440  
GTGATACAAA TTGATAGTCA ATATCAAGCA GTGAAATATG ATAATAAAGG 1490  
CATTTCAAAG TCTCACTTTT ATTGATAAAA TAAAAATCAT TCTACTGAAC 1540  
AGTCCATCTT CTTTATACAA TGACCACATC CTGAAAAGGG TGTGCTAAG 1590  
CTGTAACCGA TATGCACTTG AAATGATGGT AAGTTAATTT TGATTGAGAA 1640  
TGTGTTATTT GTCACAAATA AACATAATAA AAGGAGTTCA GATGTTTTTC 1690  
TTCATTAACC AAAAAAAAAA AAAAA 1715

***FIG.\_4C***

GAGGGCCCTG CCTCCAACT GGGGGCGGA GGTGGGTGGC TCCGGGGCAA 50

TTGAAAAAGA GCGGGCGAGG AGTTCCCGA AACTTGTGG AACTCCGGGC 100

TCGCGCGGAG GCCAGGAGCT GAGCGCGGC GGCTGCCGA CGATGGGAGC 150

GTGAGCAGGA CCGTGATAAC CTCTCCCGA TCGGGTTGG AGGGCGCCGG 200

GCAGAGGCCA GGACGCGAGC CGCCAGCGGC GGGACCCATC GACGACTTCC 250

CGGGGCGACA GGAGCAGCCC CGAGAGCCAG GGGAGCGCC CGTTCCAGGT 300

GGCCGGACCG CCCGCCCGT CCGCGCCGG CTCCCTGCAG GCAACGGGAG 350

ACGCCCCCGC GCAGCGGAG CGCCTCAGCG CGGCCGCTCG CTCTCCCAT 400

CGAGGGACAA ACTTTTCCA AACCCGATCC GAGCCCTTGG ACCAACTCG 450

CCTGCGCCGA GAGCCGTCCG CGTAGAGCGC TCCGTCTCCG GCGAG AT 497  
Met  
1

G TCC GAG CGC AAA GAA GGC AGA GGC AAA GGG AAG GGC AAG 537  
Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys  
5 10

AAG AAG GAG CGA GGC TCC GGC AAG AAG CCG GAG TCC GCG 576  
Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala  
15 20 25

GCG GGC AGC CAG AGC CCA GCC TTG CCT CCC CAA TTG AAA 615  
Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys  
30 35 40

GAG ATG AAA AGC CAG GAA TCG GCT GCA GGT TCC AAA CTA 654  
Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu  
45 50

GTC CTT CGG TGT GAA ACC AGT TCT GAA TAC TCC TCT CTC 693  
Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu  
55 60 65

AGA TTC AAG TGG TTC AAG AAT GGG AAT GAA TTG AAT CGA 732  
Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg  
70 75

**FIG. 5A**

|                                                          |  |
|----------------------------------------------------------|--|
| AAA AAC AAA CCA CAA AAT ATC AAG ATA CAA AAA AAG CCA 771  |  |
| Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro      |  |
| 80 85 90                                                 |  |
| GGG AAG TCA GAA CTT CGC ATT AAC AAA GCA TCA CTG GCT 810  |  |
| Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala      |  |
| 95 100 105                                               |  |
| GAT TCT GGA GAG TAT ATG TGC AAA GTG ATC AGC AAA TTA 849  |  |
| Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu      |  |
| 110 115                                                  |  |
| GGA AAT GAC AGT GGC TCT GCC AAT ATC ACC ATC GTG GAA 888  |  |
| Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu      |  |
| 120 125 130                                              |  |
| TCA AAC GAG ATC ATC ACT GGT ATG CCA GCC TCA ACT GAA 927  |  |
| Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu      |  |
| 135 140                                                  |  |
| GGA GCA TAT GTG TCT TCA GAG TCT CCC ATT AGA ATA TCA 966  |  |
| Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser      |  |
| 145 150 155                                              |  |
| GTA TCC ACA GAA GGA GCA AAT ACT TCT TCA TCT ACA TCT 1005 |  |
| Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser      |  |
| 160 165 170                                              |  |
| ACA TCC ACC ACT GGG ACA AGC CAT CTT GTA AAA TGT GCG 1044 |  |
| Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala      |  |
| 175 180                                                  |  |
| GAG AAG GAG AAA ACT TTC TGT GTG AAT GGA GGG GAG TGC 1083 |  |
| Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys      |  |
| 185 190 195                                              |  |
| TTC ATG GTG AAA GAC CTT TCA AAC CCC TCG AGA TAC TTG 1122 |  |
| Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu      |  |
| 200 205                                                  |  |
| TGC AAG TGC CCA AAT GAG TTT ACT GGT GAT CGC TGC CAA 1161 |  |
| Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln      |  |
| 210 215 220                                              |  |
| AAC TAC GTA ATG GGC AGC TTC TAC AAG GCG GAG GAG CTG 1200 |  |
| Asn Tyr Val Met Ala Ser Phe Tyr Lys Ala Glu Glu Leu      |  |
| 225 230 235                                              |  |
| TAC CAG AAG AGA GTG CTG ACC ATA ACC GGC ATC TGC ATC 1239 |  |
| Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys Ile      |  |
| 240 245                                                  |  |
| GCC CTC CTT GTG GTC GGC ATC ATG TGT GTG GTG GCC TAC 1278 |  |
| Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala Tyr      |  |
| 250 255 260                                              |  |

FIG. 5B

|     |     |     |       |            |            |            |      |     |     |     |     |     |      |
|-----|-----|-----|-------|------------|------------|------------|------|-----|-----|-----|-----|-----|------|
| TGC | AAA | ACC | AAG   | AAA        | CAG        | CGG        | AAA  | AAG | CTG | CAT | GAC | CGT | 1317 |
| Cys | Lys | Thr | Lys   | Lys        | Gln        | Arg        | Lys  | Lys | Leu | His | Asp | Arg |      |
|     |     |     | 265   |            |            |            |      | 270 |     |     |     |     |      |
| CTT | CGG | CAG | AGC   | CTT        | CGG        | TCT        | GAA  | CGA | AAC | AAT | ATG | ATG | 1356 |
| Leu | Arg | Gln | Ser   | Leu        | Arg        | Ser        | Glu  | Arg | Asn | Asn | Met | Met |      |
| 275 |     |     |       |            | 280        |            |      |     |     | 285 |     |     |      |
| AAC | ATT | GCC | AAT   | GGG        | CCT        | CAC        | CAT  | CCT | AAC | CCA | CCC | CCC | 1395 |
| Asn | Ile | Ala | Asn   | Gly        | Pro        | His        | His  | Pro | Asn | Pro | Pro | Pro |      |
|     |     | 290 |       |            |            |            | 295  |     |     |     |     | 300 |      |
| GAG | AAT | GTC | CAG   | CTG        | GTG        | AAT        | CAA  | TAC | GTA | TCT | AAA | AAC | 1434 |
| Glu | Asn | Val | Gln   |            | Val        | Asn        | Gln  | Tyr | Val | Ser | Lys | Asn |      |
|     |     |     | 305   |            |            |            |      | 310 |     |     |     |     |      |
| GTC | ATC | TCC | AGT   | GAG        | CAT        | ATT        | GTT  | GAG | AGA | GAA | GCA | GAG | 1473 |
| Val | Ile | Ser | Ser   | Glu        | His        | Ile        | Val  | Glu | Arg | Glu | Ala | Glu |      |
|     | 315 |     |       |            |            | 320        |      |     |     |     | 325 |     |      |
| ACA | TCC | TTT | TCC   | ACC        | AGT        | CAC        | TAT  | ACT | TCC | ACA | GCC | CAT | 1512 |
| Thr | Ser | Phe | Ser   | Thr        | Ser        | His        | Tyr  | Thr | Ser | Thr | Ala | His |      |
|     |     |     | 330   |            |            |            | 335  |     |     |     |     |     |      |
| CAC | TCC | ACT | ACT   | GTC        | ACC        | CAG        | ACT  | CCT | AGC | CAC | AGC | TGG | 1551 |
| His | Ser | Thr | Thr   | Val        | Thr        | Gln        | Thr  | Pro | Ser | His | Ser | Trp |      |
| 340 |     |     |       |            | 345        |            |      |     |     | 350 |     |     |      |
| AGC | AAC | GGA | CAC   | ACT        | GAA        | AGC        | ATC  | CTT | TCC | GAA | AGC | CAC | 1590 |
| Ser | Asn | Gly | His   | Thr        | Glu        | Ser        | Ile  | Leu | Ser | Glu | Ser | His |      |
|     |     | 355 |       |            |            |            | 360  |     |     |     |     | 365 |      |
| TCT | GTA | ATC | GTG   | ATG        | TCA        | TCC        | GTA  | GAA | AAC | AGT | AGG | CAC | 1629 |
| Ser | Val | Ile | Val   | Met        | Ser        | Ser        | Val  | Glu | Asn | Ser | Arg | His |      |
|     |     |     | 370   |            |            |            |      | 375 |     |     |     |     |      |
| AGC | AGC | CCA | ACT   | GGG        | GGC        | CCA        | AGA  | GGA | CGT | CTT | AAT | GGC | 1668 |
| Ser | Ser | Pro | Thr   | Gly        | Gly        | Pro        | Arg  | Gly | Arg | Leu | Asn | Gly |      |
|     |     | 380 |       |            |            | 385        |      |     |     |     | 390 |     |      |
| ACA | GGA | GGC | CCT   | CGT        | GAA        | TGT        | AAC  | AGC | TTC | CTC | AGG | CAT | 1707 |
| Thr | Gly | Gly | Pro   | Arg        | Glu        | Cys        | Asn  | Ser | Phe | Leu | Arg | His |      |
|     |     |     | 395   |            |            |            |      | 400 |     |     |     |     |      |
| GCC | AGA | GAA | ACC   | CCT        | GAT        | TCC        | TAC  | CGA | GAC | TCT | CCT | CAT | 1746 |
| Ala | Arg | Glu | Thr   | Pro        | Asp        | Ser        | Tyr  | Arg | Asp | Ser | Pro | His |      |
| 405 |     |     |       |            | 410        |            |      |     |     | 415 |     |     |      |
| AGT | GAA | AGG | TAAAA | CCGAAGGCAA | AGCTACTGCA | GAGGAGAAAC | 1790 |     |     |     |     |     |      |
| Ser | Glu | Arg |       |            |            |            |      |     |     |     |     |     |      |
|     |     |     | 420   |            |            |            |      |     |     |     |     |     |      |

**FIG.\_5C**

TCAGTCAGAG AATCCCTGTG AGCACCTGCG GTCTCACCTC AGGAAATCTA 1840  
 CTCTAATCAG AATAAGGGGC GGCAGTTACC TGTTC TAGGA GTGCTCCTAG 1890  
 TTGATGAAGT CATCTCTTTG TTTGACGGAA CTTATTTCTT CTGAGCTTCT 1940  
 CTCGTGCTCC CAGTGACTGA CAGGCAACAG ACTCTTAAAG AGCTGGGATG 1990  
 CTTTGATGCG GAAGGTGCAG CACATGGAGT TTCCAGCTCT GGCCATGGGC 2040  
 TCAGACCCAC TCGGGGTCTC AGTGTCTCA GTTGTAACAT TAGAGAGATG 2090  
 GCATCAATGC TTGATAAGGA CCCTTCTATA ATTCCAATTG CCAGTTATCC 2140  
 AAACCTCTGAT TCGGTGGTCG AGCTGGCCTC GTGTTCTTAT CTGCTAACCC 2190  
 TGTCTTACCT TCCAGCCTCA GTTAAGTCAA ATCAAGGGCT ATGTCATTGC 2240  
 TGAATGTCAT GGGGGGCAAC TGCTTGCCCT CCACCCTATA GTATCTATTT 2290  
 TATGAAATTC CAAGAAGGGA TGAATAATA AATCTCTTGG ATGCTGCGTC 2340  
 TGGCAGTCTT CACGGGTGGT TTTCAAAGCA GAAAAAAAAA AAAAAAAAAA 2390  
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A 2431

**FIG.\_5D**

|    |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|----|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16 | 1   | M | S | E | R | K | E | Q | R | Q | K | Q | K | K | E | R | G | S | G | K | P | E | S | A | A | G | S | Q | S | P | A | L | P | P | R | L | K | E | M | K | S | Q | E | S | A | A | G |   |   |   |   |   |
| 11 | 1   | M | S | E | R | K | E | Q | R | Q | K | Q | K | K | E | R | G | S | G | K | P | E | S | A | A | G | S | Q | S | P | A | L | P | P | Q | L | K | E | M | K | S | Q | E | S | A | A | G |   |   |   |   |   |
| 76 | 1   | M | S | E | R | K | E | Q | R | Q | K | Q | K | K | E | R | G | S | G | K | P | E | S | A | A | G | S | Q | S | P | A | L | P | P | Q | L | K | E | M | K | S | Q | E | S | A | A | G |   |   |   |   |   |
| 84 | 1   | M | S | E | R | K | E | Q | R | Q | K | Q | K | K | E | R | G | S | G | K | P | E | S | A | A | G | S | Q | S | P | A | L | P | P | Q | L | K | E | M | K | S | Q | E | S | A | A | G |   |   |   |   |   |
| 78 | 1   | M | S | E | R | K | E | Q | R | Q | K | Q | K | K | E | R | G | S | G | K | P | E | S | A | A | G | S | Q | S | P | A | L | P | P | Q | L | K | E | M | K | S | Q | E | S | A | A | G |   |   |   |   |   |
| 16 | 61  | S | K | L | V | L | R | C | E | T | S | S | E | Y | S | S | L | R | F | K | W | F | K | N | G | N | E | L | N | R | K | N | K | P | Q | N | I | K | I | Q | K | K | P | Q | K | S | E | L | R | I | N |   |
| 11 | 61  | S | K | L | V | L | R | C | E | T | S | S | E | Y | S | S | L | R | F | K | W | F | K | N | G | N | E | L | N | R | K | N | K | P | Q | N | I | K | I | Q | K | K | P | Q | K | S | E | L | R | I | N |   |
| 76 | 61  | S | K | L | V | L | R | C | E | T | S | S | E | Y | S | S | L | R | F | K | W | F | K | N | G | N | E | L | N | R | K | N | K | P | Q | N | I | K | I | Q | K | K | P | Q | K | S | E | L | R | I | N |   |
| 84 | 61  | S | K | L | V | L | R | C | E | T | S | S | E | Y | S | S | L | R | F | K | W | F | K | N | G | N | E | L | N | R | K | N | K | P | Q | N | I | K | I | Q | K | K | P | Q | K | S | E | L | R | I | N |   |
| 78 | 61  | S | K | L | V | L | R | C | E | T | S | S | E | Y | S | S | L | R | F | K | W | F | K | N | G | N | E | L | N | R | K | N | K | P | Q | N | I | K | I | Q | K | K | P | Q | K | S | E | L | R | I | N |   |
| 16 | 101 | K | A | S | L | A | D | S | Q | E | Y | M | C | K | V | I | S | K | L | Q | N | D | S | A | S | A | N | I | T | I | V | E | S | N | E | I | I | T | Q | M | P | A | S | T | E | G | A | Y | V | S |   |   |
| 11 | 101 | K | A | S | L | A | D | S | Q | E | Y | M | C | K | V | I | S | K | L | Q | N | D | S | A | S | A | N | I | T | I | V | E | S | N | E | I | I | T | Q | M | P | A | S | T | E | G | A | Y | V | S |   |   |
| 76 | 101 | K | A | S | L | A | D | S | Q | E | Y | M | C | K | V | I | S | K | L | Q | N | D | S | A | S | A | N | I | T | I | V | E | S | N | E | I | I | T | Q | M | P | A | S | T | E | G | A | Y | V | S |   |   |
| 84 | 101 | K | A | S | L | A | D | S | Q | E | Y | M | C | K | V | I | S | K | L | Q | N | D | S | A | S | A | N | I | T | I | V | E | S | N | E | I | I | T | Q | M | P | A | S | T | E | G | A | Y | V | S |   |   |
| 78 | 101 | K | A | S | L | A | D | S | Q | E | Y | M | C | K | V | I | S | K | L | Q | N | D | S | A | S | A | N | I | T | I | V | E | S | N | E | I | I | T | Q | M | P | A | S | T | E | G | A | Y | V | S |   |   |
| 16 | 151 | E | S | P | I | R | I | S | V | S | T | E | G | A | N | T | S | S | S | T | S | T | S | T | S | T | T | Q | T | S | H | L | V | K | C | A | E | K | E | K | T | F | C | V | N | Q | G | E | C | F | M | V |
| 11 | 151 | E | S | P | I | R | I | S | V | S | T | E | G | A | N | T | S | S | S | T | S | T | S | T | S | T | T | Q | T | S | H | L | V | K | C | A | E | K | E | K | T | F | C | V | N | Q | G | E | C | F | M | V |
| 76 | 151 | E | S | P | I | R | I | S | V | S | T | E | G | A | N | T | S | S | S | T | S | T | S | T | S | T | T | Q | T | S | H | L | V | K | C | A | E | K | E | K | T | F | C | V | N | Q | G | E | C | F | M | V |
| 84 | 151 | E | S | P | I | R | I | S | V | S | T | E | G | A | N | T | S | S | S | T | S | T | S | T | S | T | T | Q | T | S | H | L | V | K | C | A | E | K | E | K | T | F | C | V | N | Q | G | E | C | F | M | V |
| 78 | 151 | E | S | P | I | R | I | S | V | S | T | E | G | A | N | T | S | S | S | T | S | T | S | T | S | T | T | Q | T | S | H | L | V | K | C | A | E | K | E | K | T | F | C | V | N | Q | G | E | C | F | M | V |

FIG.\_6A

**FIG. 6B**

|    |     |                                                        |
|----|-----|--------------------------------------------------------|
| 16 | 388 | GGPRECNSFLRHARETPDSYRDSPHSERYVSAMTTPARMSPVDFHTPSSP     |
| 11 | 401 | GGPRECNSFLRHARETPDSYRDSPHSERYVSAMTTPARMSPVDFHTPSSP     |
| 76 | 393 | GGPRECNSFLRHARETPDSYRDSPHSERYVSAMTTPARMSPVDFHTPSSP     |
| 84 | 393 | GGPRECNSFLRHARETPDSYRDSPHSER-----                      |
| 16 | 448 | KSPPSSEMSPPVSSMTVSMPSMAVSPPFMEEERPLLLVTPPRLREKKFDHHP   |
| 11 | 451 | KSPPSSEMSPPVSSMTVSMPSMAVSPPFMEEERPLLLVTPPRLREKKFDHHP   |
| 76 | 443 | KSPPSSEMSPPVSSMTVS[K]PSMAVSPPFMEEERPLLLVTPPRLREKKFDHHP |
| 16 | 488 | QQFSSFHNNPAHDSNSLPASPLRIVEDEEYETTQEYEPAQEPVKKLANSR     |
| 11 | 501 | QQFSSFHNNPAHDSNSLPASPLRIVEDEEYETTQEYEPAQEPVKKLANSR     |
| 76 | 483 | QQFSSFHNNPAHDSNSLPASPLRIVEDEEYETTQEYEPAQEPVKKLANSR     |
| 16 | 548 | RAKRTKPNQGHIANRLEVDNSTSSQSSNSESETERVQEDTFFLQIQNPL      |
| 11 | 551 | RAKRTKPNQGHIANRLEVDNSTSSQSSNSESETERVQEDTFFLQIQNPL      |
| 76 | 543 | RAKRTKPNQGHIANRLEVDNSTSSQSSNSESETERVQEDTFFLQIQNPL      |
| 16 | 588 | AASLEATPAFRLADSRTPNPAGRFSTQEEIQ-----                   |
| 11 | 601 | AASLEATPAFRLADSRTPNPAGRFSTQEEIQARLSSSVIANQDPIAV        |
| 76 | 593 | AASLEATPAFRLADSRTPNPAGRFSTQEEIQARLSSSVIANQDPIAV        |

FIG..6C

1 GGGTACCATGGGTGCGTGAGCGCGTTTCCCGCCTGAGCGCAACTAGCGGC  
 51 GGGTCGTGGGCACCTCCAGAAAAGATCCCGCACCATCCTCCAGGATCCAA  
 101 TGGCCTTGGAGAGAGGGCTGCAGGGCCCACGGACATTGCTGACTCTTCAQ  
 151 AACGTGCTGACATGGAGCCAGGTAGACTGAAATTATCATGTGTCCAAATT  
 201 AAAATTGCATACTTCAAGGATTATTTGAAGGACTATTCTTAGACCCTTTT  
 251 AAGAAGATTTAAAGAAAAACCACTCGGCCCTGAGTGCGGCGAGGACCCTG  
  
 301 TTTGTGGATGTGGAGGAGCGCGGCCCGGAGGCCATGGACGTGAAGGAGAG  
 1 M D V K E R  
  
 351 GAAGCCTTACCGCTCGCTGACCCGGCGCCGCGACGCCGAGCGCCGCTACA  
 7 K P Y R S L T R R R D A E R R Y T  
  
 401 CCAGCTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCCGAGAAATCGTAC  
 24 S S S A D S E E G K A P Q K S Y  
  
 451 AGCTCCAGCGAGACCCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCTA  
 40 S S S E T L K A Y D Q D A R L A Y  
  
 501 TGGCAGCCGCGTCAAGGACATTGTGCCGAGGAGGCCGAGGAATTCTGCC  
 57 G S R V K D I V P Q E A E E F C R  
  
 551 GCACAGGTGCCA<sup>•</sup>ACTTCCACCCTGCGGGAGCTGGGGCTGGAAGAAGTAACG  
 74 T G A N F T L R E L G L E E V T  
  
 601 CCCCCTCACGGGACCCTGTACCGGACAGACATTGGCCTCCCCCACTGCGG  
 90 P P H G T L Y R T D I G L P H C G  
  
 651 CTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGC  
 107 Y S M G A G S D A D M E A D T V L  
  
 701 TGTCCCCTGAGCACCCCGTGCCTCTGTGGGGCCCGAGCACACGGTCAGGG  
 124 S P E H P V R L W G R S T R S G  
  
 751 CGCAGCTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGA  
 140 R S S C L S S R A N S N L T L T D  
  
 801 CACCGAGCATGAAAACACTGAGACTGATCATCCGGGCGGCCTGCAGAACC  
 157 T E H E N T E T D H P G G L Q N H  
  
 851 ACGCGCGGCTCCGGACGCCGCGCCGCGCCTCTCGCACGCCACACCCCC  
 174 A R L R T P P P P L S H A H T P  
  
 901 AACCAGCACCACGCGGCCTCCATTA<sup>•</sup>ACTCCCTGAACCGGGGCAACTTCAC  
 190 N Q H H A A S I N S L N R G N F T  
  
 951 GCCGAGGAGCAACCCAGCCCGGCCCCCACGGACCACTCGCTCTCCGGAG  
 207 P R S N P S P A P T D H S L S G E  
  
 1001 AGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCAGGAGA<sup>•</sup>ACTGG  
 224 P P A G G A Q E P A H A Q E N W  
  
 1051 CTGCTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCC  
 240 L L N S N I P L E T R N L G K Q P

FIG. 7A

1101 ATTCCTAGGGACATTGCAGGACAACCTCATTGAGATGGACATTCTCGGCG  
 257 F L G T L Q D N L I E M D I L G A  
 1151 CCTCCCGCCATGATGGGGCTTACAGTGACGGGCACTTCCTCTTCAAGCCT  
 274 S R H D G A Y S D G H F L F K P  
 1201 GGAGGCACCTCCCCGCTCTTCTGCACCACATCACCAGGGTACCCACTGAC  
 290 G G T S P L F C T T S P G Y P L T  
 1251 GTCCAGCACAGTGTACTCTCCTCCGCCCCGACCCCTGCCCCGCAGCACCT  
 307 S S T V Y S P P P R P L P R S T F  
 1301 TCGCCCGGCGGCGCTTTAACCTCAAGAAGCCCTCCAAGTACTGTAAGTGG  
 324 A R P A F N L K K P S K Y C N W  
 1351 AAGTGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCCT  
 340 K C A A L S A I V I S A T L V I L  
 1401 GCTGGCATACTTTGTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGC  
 357 L A Y F V A M H L F G L N W H L Q  
 1451 AGCCGATGGAGGGGCAGATGTATGAGATCACGGAGGACACAGCCAGCAGT  
 374 P M E G Q M Y E I T E D T A S S  
 1501 TGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGGCACTGGCTT  
 390 W P V P T D V S L Y P S G G T G L  
 1551 AGAGACCCCTGACAGGAAAGGCAAAGGAACACAGAAGGAAAGCCAGTA  
 407 E T F D R K G K G T T E G K P S S  
 1601 GTTCTTTTCCAGAGGACAGTTTCATAGATTCTGGAGAAATTGATGTGGGA  
 424 F F P E D S F I D S G E I D V G  
 1651 AGGCGAGCTTCCCAGAAGATTCCCTCCTGGCACTTTCTGGAGATCTCAAGT  
 440 R R A S Q K I P P G T F W R S Q V  
 1701 GTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCTCTGGGAAAGG  
 457 F I D H P V H L K F N V S L G K A  
 1751 CAGCCCTGGTTGGCATTATATGGCAGAAAAGGCCTCCCTCCTTCACATACA  
 474 A L V G I Y G R K G L P P S H T  
 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGA  
 490 Q F D F V E L L D G R R L L T Q E  
 1851 GGCGCGGAGCCTAGAGGGGACCCCGCGCCAGTCTCGGGGAAGTGTGCCCC  
 507 A R S L E G T P R Q S R G T V P P  
 1901 CCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTCAGGAATCTGG  
 524 S S H E T G F I Q Y L D S G I W  
 1951 CACTTGGCTTTTTACAATGACGGAAAGGAGTCAGAAGTGGTTTCCTTTCT  
 540 E L A F Y N D G K E S E V V S F L

FIG.\_7B

2001 CACCACTGCCATTGCCTTGCCTCCCCGATTGAAAGAGATGAAAAGCCAGG  
 557 T T A I A L P P R L K E M K S Q E  
 2051 AATCGGCTGCAGGTTCCAACTAGTCCTTCGGTGTGAAACCAGTTCTGAA  
 574 S A A G S K L V L R C E T S S E  
 2101 TACTCCTCTCTCAGATTCAAGTGGTTCAAGAATGGGAATGAATTGAATCG  
 590 Y S S L R F K W F K N G N E L N R  
 2151 AAAAAACAAACCACAAAATATCAAGATACAAAAAAGCCAGGGAAGTCAG  
 607 K N K P Q N I K I Q K K P G K S E  
 2201 AACTTCGCATTAACAAAGCATCACTGGCTGATTCTGGAGAGTATATGTGC  
 624 L R I N K A S L A D S G E Y M C  
 2251 AAAGTGATCAGCAAATTAGGAAATGACAGTGCCTCTGCCAATATCACCAT  
 640 K V I S K L G N D S A S A N I T I  
 2301 CGTGGAAATCAAACGAGATCATCACTGGTATGCCAGCCTCAACTGAAGGAG  
 657 V E S N E I I T G M P A S T E G A  
 2351 CATATGTGTCTTCAGAGTCTCCCATAGAAATATCAGTATCCACAGAAGGA  
 674 Y V S S E S P I R I S V S T E G  
 2401 GCAAATACTTCTTCATCTACATCTACATCCACCCTGGGACAAGCCATCT  
 690 A N T S S S T S T S T T G T S H L  
 2451 TCTAAATGTGCGGAGAAGGAGAAACTTTCTGTGTGAATGGAGGGGAGT  
 707 V K © A E K E K T F © V W G G E ©  
 2501 GCTTCATGGTGAAAGACCTTTCAAACCCCTCGAGATACTTGTGCAAGTGC  
 724 F M V K D L S N P S R Y L © K ©  
 2551 CCAAATGAGTTTACTGGTGATCGCTGCCAAAACCTACGTAATGGCCAGCTT  
 740 P N E P T G D R © Q N Y V M A S F  
 2601 CTACAGTACGTCCACTCCCTTTCTGTCTCTGCCTGAATAGGAGCATGCTC  
 757 Y S T S T P F L S L P E  
 2651 AGTTGGTGCTGCTTTCTTGTGCTGCATCTCCCCTCAGATTCCACCTAGA  
 2701 GCTAGATGTGTCTTACCAGATCTAATATTGACTGCCTCTGCCTGTGCGCAT  
 2751 GAGAACATTAAACAAAAGCAATTGTATTACTTCTCTGTTCCGCGACTAGTT  
 2801 GGCTCTGAGATACTAATAGGTGTGTGAGGCTCCGGATGTTTCTGGAAATG  
 2851 ATATTGAATGATGTGATACAAATTGATAGTCAATATCAAGCAGTGAATA  
 2901 TGATAATAAAGGCATTTCAAAGTCTCACTTTTATTGATAAAATAAAATC  
 2951 ATTCTACTGAACAGTCCATCTTCTTTATACAAATGACCACATCCTGAAAA  
 3001 GGTGTTGCTAAGCTGTAACCGATATGCACTTGAAATGATGGTAAGTTAAT  
 3051 TTTGATTCAAGATGTGTTATTTGTCAAAATAAACATAATAAAAGGAAAA  
 3101 AAAAAAAAAA

FIG.\_7C

**FIG.-8A**

976 TGCAAGGTGACAJGGCTGTGTCTCTCTTTGAAACCATCAGCGGCACCGACACCGAAGAAATCGTATTTTGCCTTTT  
 158 Q G D K A V V S F E P S A A P T P K N R I F A F S  
  
 1051 CTTTCTTGCCGTCACACTGCGCCATCCTTCCCTTACCCACCCGGAACCCCTGAGGTGAGNACGCCCAAGTCAGCAA  
 183 F L P S T A P S F P S P T R N P E V R T P K S A T  
  
 1126 CTCAGCCACAAACACAGAACTAATCTCCAACTGCTCTCTAACTTTCTACATCTACATCCACCACCTGGGACAA  
 208 Q P Q T T E T N L Q T A P K L S T S T S T T G T S  
  
 1201 GCCATCTTGTAAATGTGCGGAGAGGAGAAACTTTCTGTGTAATGGAGGGAGTGCTTCATGTGTAAGACC  
 233 H L V K C A E K E K T F C V N G G E C F H V K D L  
  
 1276 TTTCAAACCCCTCGAGATACTTGTGCAAGTGCCCAATGAGTTTACTGTGATCGCTGCCAAAACTACGTAATGG  
 258 S N P S R Y L C K C P N E F T G D R C Q N Y V H A  
  
 1351 CCAGCTTCTACAGTAGCTCCACTCCCTTTCTGTGCTCTGCTGAATAGGAGCATGCTCAGTTGGTGCTGCTTTCTT  
 283 S F Y S T S T P F L S L P E O  
  
 1426 GTTCTGCATCTCCCTCAGATTCCACCTAGAGCTAGATGTGTCTTACCAGATCTAATATTGACTGCCCTCTGCCT  
  
 1501 GTCCGATGAGAACATTTAACAAAGCAATTGTATTACTTCTCTCTCTGTTCCGGACTAGTTGGCTCTGAGATACTAATA  
  
 1576 CGTGTGTGAGGCTCCGGATGTTTCTGGAAATTGATATTGAAATGATGTGATACAAATTGATAGTCAATATCAAGCAG  
  
 1651 TGAATATGATAATAAGGCCATTTTCAAGTCTCCTCTTTTATTGATAAATATAAATCATTTCTACTGAACAGTCCA  
  
 1726 TCTTCTTTATACAAAGACCACATCCTGAAAGGGGTGTTGCTAAGCTGTAAACCGATATGCACCTTGAAATGATGTA  
  
 1801 AGTTAATTTTGATTTCAGAAATGTGTTATTGTGCACAAATTAACATATAAAGGAAATAAACCAGAAATTC

EGF  
-like

FIG.\_8B

# ES 2 329 467 T3

## LISTA DE SECUENCIAS

<110> Genentech, Inc.

<120> Use of Heregulin as a Growth Factor

<130> P1145PCT

<141> 1999-02-03

<150> US 09/020,598

<151> 1998-02-04

<160> 14

<210> 1

<211> 669

<212> PRT

<213> *Homo sapiens*

<400> 1

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

# ES 2 329 467 T3

|    | 170                                                         | 175 | 180 |
|----|-------------------------------------------------------------|-----|-----|
|    | Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu |     |     |
| 5  | 185                                                         | 190 | 195 |
|    | Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr Ser |     |     |
|    | 200                                                         | 205 | 210 |
| 10 | Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys |     |     |
|    | 215                                                         | 220 | 225 |
|    | Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys |     |     |
|    | 230                                                         | 235 | 240 |
| 15 | Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys Lys |     |     |
|    | 245                                                         | 250 | 255 |
|    | Cys Gln Pro Gly Phe Thr Gly Ala Arg Cys Thr Glu Asn Val Pro |     |     |
| 20 | 260                                                         | 265 | 270 |
|    | Met Lys Val Gln Asn Gln Glu Lys Ala Glu Glu Leu Tyr Gln Lys |     |     |
|    | 275                                                         | 280 | 285 |
| 25 | Arg Val Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu Val Val |     |     |
|    | 290                                                         | 295 | 300 |
|    | Gly Ile Met Cys Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg |     |     |
|    | 305                                                         | 310 | 315 |
| 30 | Lys Lys Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg |     |     |
|    | 320                                                         | 325 | 330 |
|    | Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro |     |     |
| 35 | 335                                                         | 340 | 345 |
|    | Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn |     |     |
|    | 350                                                         | 355 | 360 |
| 40 | Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser |     |     |
|    | 365                                                         | 370 | 375 |
|    | Phe Ser Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr Thr |     |     |
|    | 380                                                         | 385 | 390 |
| 45 | Val Thr Gln Thr Pro Ser His Ser Trp Ser Asn Gly His Thr Glu |     |     |
|    | 395                                                         | 400 | 405 |
|    | Ser Ile Leu Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val |     |     |
|    | 410                                                         | 415 | 420 |
| 50 | Glu Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg |     |     |
|    | 425                                                         | 430 | 435 |
| 55 | Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg |     |     |
|    | 440                                                         | 445 | 450 |

# ES 2 329 467 T3

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
|    | His | Ala | Arg | Glu | Thr | Pro | Asp | Ser | Tyr | Arg | Asp | Ser | Pro | His | Ser |  |
|    |     |     |     |     |     |     | 455 |     |     |     | 460 |     |     |     | 465 |  |
| 5  | Glu | Arg | Tyr | Val | Ser | Ala | Met | Thr | Thr | Pro | Ala | Arg | Met | Ser | Pro |  |
|    |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |  |
| 10 | Val | Asp | Phe | His | Thr | Pro | Ser | Ser | Pro | Lys | Ser | Pro | Pro | Ser | Glu |  |
|    |     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |  |
| 15 | Met | Ser | Pro | Pro | Val | Ser | Ser | Met | Thr | Val | Ser | Met | Pro | Ser | Met |  |
|    |     |     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |  |
|    | Ala | Val | Ser | Pro | Phe | Met | Glu | Glu | Glu | Arg | Pro | Leu | Leu | Leu | Val |  |
|    |     |     |     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |  |
| 20 | Thr | Pro | Pro | Arg | Leu | Arg | Glu | Lys | Lys | Phe | Asp | His | His | Pro | Gln |  |
|    |     |     |     |     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |  |
| 25 | Gln | Phe | Ser | Ser | Phe | His | His | Asn | Pro | Ala | His | Asp | Ser | Asn | Ser |  |
|    |     |     |     |     | 545 |     |     |     |     | 550 |     |     |     |     | 555 |  |
| 30 | Leu | Pro | Ala | Ser | Pro | Leu | Arg | Ile | Val | Glu | Asp | Glu | Glu | Tyr | Glu |  |
|    |     |     |     |     | 560 |     |     |     |     | 565 |     |     |     |     | 570 |  |
|    | Thr | Thr | Gln | Glu | Tyr | Glu | Pro | Ala | Gln | Glu | Pro | Val | Lys | Lys | Leu |  |
|    |     |     |     |     | 575 |     |     |     |     | 580 |     |     |     |     | 585 |  |
| 35 | Ala | Asn | Ser | Arg | Arg | Ala | Lys | Arg | Thr | Lys | Pro | Asn | Gly | His | Ile |  |
|    |     |     |     |     | 590 |     |     |     |     | 595 |     |     |     |     | 600 |  |
| 40 | Ala | Asn | Arg | Leu | Glu | Val | Asp | Ser | Asn | Thr | Ser | Ser | Gln | Ser | Ser |  |
|    |     |     |     |     | 605 |     |     |     |     | 610 |     |     |     |     | 615 |  |
| 45 | Asn | Ser | Glu | Ser | Glu | Thr | Glu | Asp | Glu | Arg | Val | Gly | Glu | Asp | Thr |  |
|    |     |     |     |     | 620 |     |     |     |     | 625 |     |     |     |     | 630 |  |
|    | Pro | Phe | Leu | Gly | Ile | Gln | Asn | Pro | Leu | Ala | Ala | Ser | Leu | Glu | Ala |  |
|    |     |     |     |     | 635 |     |     |     |     | 640 |     |     |     |     | 645 |  |
| 50 | Thr | Pro | Ala | Phe | Arg | Leu | Ala | Asp | Ser | Arg | Thr | Asn | Pro | Ala | Gly |  |
|    |     |     |     |     | 650 |     |     |     |     | 655 |     |     |     |     | 660 |  |
| 55 | Arg | Phe | Ser | Thr | Gln | Glu | Glu | Ile | Gln |     |     |     |     |     |     |  |
|    |     |     |     |     | 665 |     |     |     | 669 |     |     |     |     |     |     |  |

<210> 2

<211> 2226

60 <212> DNA

<213> *Homo sapiens*

65

# ES 2 329 467 T3

<400> 2

```

5      gg  gcg cga gcg cct cag cgc ggc cgc tcg ctc tcc ccc 38
        Ala Arg Ala Pro Gln Arg Gly Arg Ser Leu Ser Pro
          1           5           10

10     tcg agg gac aaa ctt ttc cca aac ccg atc cga gcc ctt 77

        Ser Arg Asp Lys Leu Phe Pro Asn Pro Ile Arg Ala Leu
          15           20           25

15     gga cca aac tcg cct gcg ccg aga gcc gtc cgc gta gag 116
        Gly Pro Asn Ser Pro Ala Pro Arg Ala Val Arg Val Glu
          30           35

20     cgc tcc gtc tcc ggc gag atg tcc gag cgc aaa gaa ggc 155
        Arg Ser Val Ser Gly Glu Met Ser Glu Arg Lys Glu Gly
          40           45           50

25     aga ggc aaa ggg aag ggc aag aag gag cga gcc tcc 194
        Arg Gly Lys Gly Lys Gly Lys Lys Glu Arg Gly Ser
          55           60

30     ggc aag aag ccg gag tcc gcg gcg ggc agc cag agc cca 233
        Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro
          65           70           75

35     gcc ttg cct ccc cga ttg aaa gag atg aaa agc cag gaa 272
        Ala Leu Pro Pro Arg Leu Lys Glu Met Lys Ser Gln Glu
          80           85           90

40     tcg gct gca ggt tcc aaa cta gtc ctt cgg tgt gaa acc 311
        Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr
          95           100

45     agt tct gaa tac tcc tct ctc aga ttc aag tgg ttc aag 350
        Ser Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys
          105           110           115

50     aat ggg aat gaa ttg aat cga aaa aac aaa cca caa aat 389
        Asn Gly Asn Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn
          120           125

55     atc aag ata caa aaa aag cca ggg aag tca gaa ctt cgc 428
        Ile Lys Ile Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg
          130           135           140

60     att aac aaa gca tca ctg gct gat tct gga gag tat atg 467
        Ile Asn Lys Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met
          145           150           155

65     tgc aaa gtg atc agc aaa tta gga aat gac agt gcc tct 506
        Cys Lys Val Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser
          160           165

70     gcc aat atc acc atc gtg gaa tca aac gag atc atc act 545
        Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile Thr
          170           175           180

75     ggt atg cca gcc tca act gaa gga gca tat gtg tct tca 584
        Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser
          185           190

80     gag tct ccc att aga ata tca gta tcc aca gaa gga gca 623

```

# ES 2 329 467 T3

Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala  
 195 200 205  
 aat act tct tca tct aca tct aca tcc acc act ggg aca 662  
 5 Asn Thr Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr  
 210 215 220  
 agc cat ctt gta aaa tgt gcg gag aag gag aaa act ttc 701  
 10 Ser His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe  
 225 230  
 tgt gtg aat gga ggg gag tgc ttc atg gtg aaa gac ctt 740  
 Cys Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu  
 15 235 240 245  
 tca aac ccc tgc aga tac ttg tgc aag tgc caa cct gga 779  
 Ser Asn Pro Ser Arg Tyr Leu Cys Lys Cys Gln Pro Gly  
 250 255  
 ttc act gga gca aga tgt act gag aat gtg ccc atg aaa 818  
 20 Phe Thr Gly Ala Arg Cys Thr Glu Asn Val Pro Met Lys  
 260 265 270  
 gtc caa aac caa gaa aag gcg gag gag ctg tac cag aag 857  
 25 Val Gln Asn Gln Glu Lys Ala Glu Glu Leu Tyr Gln Lys  
 275 280 285  
 aga gtg ctg acc ata acc ggc atc tgc atc gcc ctc ctt 896  
 30 Arg Val Leu Thr Ile Thr Gly Ile Cys Ile Ala Leu Leu  
 290 295  
 gtg gtc ggc atc atg tgt gtg gtg gcc tac tgc aaa acc 935  
 Val Val Gly Ile Met Cys Val Val Ala Tyr Cys Lys Thr  
 35 300 305 310  
 aag aaa cag cgg aaa aag ctg cat gac cgt ctt cgg cag 974  
 Lys Lys Gln Arg Lys Lys Leu His Asp Arg Leu Arg Gln  
 315 320  
 agc ctt cgg tct gaa cga aac aat atg atg aac att gcc 1013  
 40 Ser Leu Arg Ser Glu Arg Asn Asn Met Met Asn Ile Ala  
 325 330 335  
 aat ggg cct cac cat cct aac cca ccc ccc gag aat gtc 1052  
 45 Asn Gly Pro His His Pro Asn Pro Pro Pro Glu Asn Val  
 340 345 350  
 cag ctg gtg aat caa tac gta tct aaa aac gtc atc tcc 1091  
 50 Gln Leu Val Asn Gln Tyr Val Ser Lys Asn Val Ile Ser  
 355 360  
 agt gag cat att gtt gag aga gaa gca gag aca tcc ttt 1130  
 Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser Phe  
 365 370 375  
 55 tcc acc agt cac tat act tcc aca gcc cat cac tcc act 1169

60

65

# ES 2 329 467 T3

Ser Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr  
380 385

5 act gtc acc cag act cct agc cac agc tgg agc aac gga 1208  
Thr Val Thr Gln Thr Pro Ser His Ser Trp Ser Asn Gly  
390 395 400

10 cac act gaa agc atc ctt tcc gaa agc cac tct gta atc 1247  
His Thr Glu Ser Ile Leu Ser Glu Ser His Ser Val Ile  
405 410 415

15 gtg atg tca tcc gta gaa aac agt agg cac agc agc cca 1286  
Val Met Ser Ser Val Glu Asn Ser Arg His Ser Ser Pro  
420 425

20 act ggg ggc cca aga gga cgt ctt aat ggc aca gga ggc 1325  
Thr Gly Gly Pro Arg Gly Arg Leu Asn Gly Thr Gly Gly  
430 435 440

25 cct cgt gaa tgt aac agc ttc ctc agg cat gcc aga gaa 1364  
Pro Arg Glu Cys Asn Ser Phe Leu Arg His Ala Arg Glu  
445 450

30 acc cct gat tcc tac cga gac tct cct cat agt gaa agg 1403  
Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser Glu Arg  
455 460 465

35 tat gtg tca gcc atg acc acc ccg gct cgt atg tca cct 1442  
Tyr Val Ser Ala Met Thr Thr Pro Ala Arg Met Ser Pro  
470 475 480

40 gta gat ttc cac acg cca agc tcc ccc aaa tcg ccc cct 1481  
Val Asp Phe His Thr Pro Ser Ser Pro Lys Ser Pro Pro  
485 490

45 tcg gaa atg tct cca ccc gtg tcc agc atg acg gtg tcc 1520  
Ser Glu Met Ser Pro Pro Val Ser Ser Met Thr Val Ser  
495 500 505

50 atg cct tcc atg gcg gtc agc ccc ttc atg gaa gaa gag 1559  
Met Pro Ser Met Ala Val Ser Pro Phe Met Glu Glu Glu  
510 515

55 aga cct cta ctt ctc gtg aca cca cca agg ctg cgg gag 1598  
Arg Pro Leu Leu Leu Val Thr Pro Pro Arg Leu Arg Glu  
520 525 530

60 aag aag ttt gac cat cac cct cag cag ttc agc tcc ttc 1637  
Lys Lys Phe Asp His His Pro Gln Gln Phe Ser Ser Phe  
535 540 545

65 cac cac aac ccc gcg cat gac agt aac agc ctc cct gct 1676  
His His Asn Pro Ala His Asp Ser Asn Ser Leu Pro Ala  
550 555

agc ccc ttg agg ata gtg gag gat gag gag tat gaa acg 1715

# ES 2 329 467 T3

Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr Glu Thr  
560 565 570

5 acc caa gag tac gag cca gcc caa gag cct gtt aag aaa 1754  
Thr Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys  
575 580

10 ctc gcc aat agc cgg cgg gcc aaa aga acc aag ccc aat 1793  
Leu Ala Asn Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn  
585 590 595

15 ggc cac att gct aac aga ttg gaa gtg gac agc aac aca 1832  
Gly His Ile Ala Asn Arg Leu Glu Val Asp Ser Asn Thr  
600 605 610

20 agc tcc cag agc agt aac tca gag agt gaa aca gaa gat 1871  
Ser Ser Gln Ser Ser Asn Ser Glu Ser Glu Thr Glu Asp  
615 620

25 gaa aga gta ggt gaa gat acg cct ttc ctg ggc ata cag 1910  
Glu Arg Val Gly Glu Asp Thr Pro Phe Leu Gly Ile Gln  
625 630 635

30 aac ccc ctg gca gcc agt ctt gag gca aca cct gcc ttc 1949  
Asn Pro Leu Ala Ala Ser Leu Glu Ala Thr Pro Ala Phe  
640 645

35 cgc ctg gct gac agc agg act aac cca gca ggc cgc ttc 1988  
Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala Gly Arg Phe  
650 655 660

40 tcg aca cag gaa gaa atc cag g ccaggctgtc tagtgtaatt 2030  
Ser Thr Gln Glu Glu Ile Gln  
665 669

45 gctaaccaag accctattgc tgtataaaac cttaaataaac acatagattc 2080

50 acctgtaaaaa ctttatttta tataataaag tattccacct taaattaaac 2130

aatttatattt attttagcag ttctgcaaat agaaaacagg aaaaaaactt 2180

55 ttataaatta aatatatgta tgtaaaaatg aaaaaaaaaa aaaaaa 2226

<210> 3

<211> 675

60 <212> PRT

<213> *Homo sapiens*

65

# ES 2 329 467 T3

<400> 3

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

# ES 2 329 467 T3

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
|    | Ser | Leu | Arg | Ser | Glu | Arg | Asn | Asn | Met | Met | Asn | Ile | Ala | Asn | Gly |  |
|    |     |     |     |     | 320 |     |     |     |     | 325 |     |     |     |     | 330 |  |
| 5  | Pro | His | His | Pro | Asn | Pro | Pro | Pro | Glu | Asn | Val | Gln | Leu | Val | Asn |  |
|    |     |     |     |     | 335 |     |     |     |     | 340 |     |     |     |     | 345 |  |
|    | Gln | Tyr | Val | Ser | Lys | Asn | Val | Ile | Ser | Ser | Glu | His | Ile | Val | Glu |  |
|    |     |     |     |     | 350 |     |     |     |     | 355 |     |     |     |     | 360 |  |
| 10 | Arg | Glu | Ala | Glu | Thr | Ser | Phe | Ser | Thr | Ser | His | Tyr | Thr | Ser | Thr |  |
|    |     |     |     |     | 365 |     |     |     |     | 370 |     |     |     |     | 375 |  |
|    | Ala | His | His | Ser | Thr | Thr | Val | Thr | Gln | Thr | Pro | Ser | His | Ser | Trp |  |
| 15 |     |     |     |     | 380 |     |     |     |     | 385 |     |     |     |     | 390 |  |
|    | Ser | Asn | Gly | His | Thr | Glu | Ser | Ile | Leu | Ser | Glu | Ser | His | Ser | Val |  |
|    |     |     |     |     | 395 |     |     |     |     | 400 |     |     |     |     | 405 |  |
| 20 | Ile | Val | Met | Ser | Ser | Val | Glu | Asn | Ser | Arg | His | Ser | Ser | Pro | Thr |  |
|    |     |     |     |     | 410 |     |     |     |     | 415 |     |     |     |     | 420 |  |
|    | Gly | Gly | Pro | Arg | Gly | Arg | Leu | Asn | Gly | Thr | Gly | Gly | Pro | Arg | Glu |  |
|    |     |     |     |     | 425 |     |     |     |     | 430 |     |     |     |     | 435 |  |
| 25 | Cys | Asn | Ser | Phe | Leu | Arg | His | Ala | Arg | Glu | Thr | Pro | Asp | Ser | Tyr |  |
|    |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |     | 450 |  |
|    | Arg | Asp | Ser | Pro | His | Ser | Glu | Arg | Tyr | Val | Ser | Ala | Met | Thr | Thr |  |
| 30 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     | 465 |  |
|    | Pro | Ala | Arg | Met | Ser | Pro | Val | Asp | Phe | His | Thr | Pro | Ser | Ser | Pro |  |
|    |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |  |
| 35 | Lys | Ser | Pro | Pro | Ser | Glu | Met | Ser | Pro | Pro | Val | Ser | Ser | Met | Thr |  |
|    |     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |  |
|    | Val | Ser | Met | Pro | Ser | Met | Ala | Val | Ser | Pro | Phe | Met | Glu | Glu | Glu |  |
|    |     |     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |  |
| 40 | Arg | Pro | Leu | Leu | Leu | Val | Thr | Pro | Pro | Arg | Leu | Arg | Glu | Lys | Lys |  |
|    |     |     |     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |  |
|    | Phe | Asp | His | His | Pro | Gln | Gln | Phe | Ser | Ser | Phe | His | His | Asn | Pro |  |
| 45 |     |     |     |     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |  |
|    | Ala | His | Asp | Ser | Asn | Ser | Leu | Pro | Ala | Ser | Pro | Leu | Arg | Ile | Val |  |
|    |     |     |     |     | 545 |     |     |     |     | 550 |     |     |     |     | 555 |  |
| 50 | Glu | Asp | Glu | Glu | Tyr | Glu | Thr | Thr | Gln | Glu | Tyr | Glu | Pro | Ala | Gln |  |
|    |     |     |     |     | 560 |     |     |     |     | 565 |     |     |     |     | 570 |  |
|    | Glu | Pro | Val | Lys | Lys | Leu | Ala | Asn | Ser | Arg | Arg | Ala | Lys | Arg | Thr |  |
|    |     |     |     |     | 575 |     |     |     |     | 580 |     |     |     |     | 585 |  |
| 55 | Lys | Pro | Asn | Gly | His | Ile | Ala | Asn | Arg | Leu | Glu | Val | Asp | Ser | Asn |  |
|    |     |     |     |     | 590 |     |     |     |     | 595 |     |     |     |     | 600 |  |

# ES 2 329 467 T3

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|    | Thr | Ser | Ser | Gln | Ser | Ser | Asn | Ser | Glu | Ser | Glu | Thr | Glu | Asp | Glu |
|    |     |     |     |     | 605 |     |     |     |     | 610 |     |     |     |     | 615 |
| 5  | Arg | Val | Gly | Glu | Asp | Thr | Pro | Phe | Leu | Gly | Ile | Gln | Asn | Pro | Leu |
|    |     |     |     |     | 620 |     |     |     |     | 625 |     |     |     |     | 630 |
|    | Ala | Ala | Ser | Leu | Glu | Ala | Thr | Pro | Ala | Phe | Arg | Leu | Ala | Asp | Ser |
| 10 |     |     |     |     | 635 |     |     |     |     | 640 |     |     |     |     | 645 |
|    | Arg | Thr | Asn | Pro | Ala | Gly | Arg | Phe | Ser | Thr | Gln | Glu | Glu | Ile | Gln |
|    |     |     |     |     | 650 |     |     |     |     | 655 |     |     |     |     | 660 |
| 15 | Ala | Arg | Leu | Ser | Ser | Val | Ile | Ala | Asn | Gln | Asp | Pro | Ile | Ala | Val |
|    |     |     |     |     | 665 |     |     |     |     | 670 |     |     |     |     | 675 |

<210> 4  
 20 <211> 2199  
 <212> DNA  
 <213> *Homo sapiens*  
 25 <400> 4

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|    | gg  | gac | aaa | ctt | ttc | cca | aac | ccg | atc | cga | gcc | ctt | gga | 38  |
|    | Asp | Lys | Leu | Phe | Pro | Asn | Pro | Ile | Arg | Ala | Leu | Gly |     |     |
| 30 |     | 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |
|    | cca | aac | tcg | cct | gcg | ccg | aga | gcc | gtc | cgc | gta | gag | cgc | 77  |
|    | Pro | Asn | Ser | Pro | Ala | Pro | Arg | Ala | Val | Arg | Val | Glu | Arg |     |
| 35 |     |     | 15  |     |     |     | 20  |     |     |     |     | 25  |     |     |
|    | tcc | gtc | tcc | ggc | gag | atg | tcc | gag | cgc | aaa | gaa | ggc | aga | 116 |
|    | Ser | Val | Ser | Gly | Glu | Met | Ser | Glu | Arg | Lys | Glu | Gly | Arg |     |
| 40 |     |     |     |     | 30  |     |     |     |     | 35  |     |     |     |     |
|    | ggc | aaa | ggg | aag | ggc | aag | aag | aag | gag | cga | ggc | tcc | ggc | 155 |
|    | Gly | Lys | Gly | Lys | Gly | Lys | Lys | Lys | Glu | Arg | Gly | Ser | Gly |     |
| 45 |     | 40  |     |     |     | 45  |     |     |     |     | 50  |     |     |     |
|    | aag | aag | ccg | gag | tcc | gcg | gcg | ggc | agc | cag | agc | cca | gcc | 194 |
|    | Lys | Lys | Pro | Glu | Ser | Ala | Ala | Gly | Ser | Gln | Ser | Pro | Ala |     |
| 50 |     |     |     |     | 55  |     |     |     | 60  |     |     |     |     |     |
|    | ttg | cct | ccc | caa | ttg | aaa | gag | atg | aaa | agc | cag | gaa | tcg | 233 |
|    | Leu | Pro | Pro | Gln | Leu | Lys | Glu | Met | Lys | Ser | Gln | Glu | Ser |     |
| 55 |     | 65  |     |     |     | 70  |     |     |     |     | 75  |     |     |     |
|    | gct | gca | ggt | tcc | aaa | cta | gtc | ctt | cgg | tgt | gaa | acc | agt | 272 |
|    | Ala | Ala | Gly | Ser | Lys | Leu | Val | Leu | Arg | Cys | Glu | Thr | Ser |     |
| 60 |     |     | 80  |     |     |     |     | 85  |     |     |     | 90  |     |     |
|    | tct | gaa | tac | tcc | tct | ctc | aga | ttc | aag | tgg | ttc | aag | aat | 311 |
|    | Ser | Glu | Tyr | Ser | Ser | Leu | Arg | Phe | Lys | Trp | Phe | Lys | Asn |     |
| 65 |     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     |
|    | ggg | aat | gaa | ttg | aat | cga | aaa | aac | aaa | cca | caa | aat | atc | 350 |
|    | Gly | Asn | Glu | Leu | Asn | Arg | Lys | Lys | Pro | Gln | Asn | Ile |     |     |
|    |     | 105 |     |     |     |     | 110 |     |     |     |     | 115 |     |     |

# ES 2 329 467 T3

aag ata caa aaa aag cca ggg aag tca gaa ctt cgc att 389  
 Lys Ile Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile  
 120 125

5

aac aaa gca tca ctg gct gat tct gga gag tat atg tgc 428  
 Asn Lys Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys  
 130 135 140

10

aaa gtg atc agc aaa tta gga aat gac agt gcc tct gcc 467  
 Lys Val Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala  
 145 150 155

15

aat atc acc atc gtg gaa tca aac gag atc atc act ggt 506  
 Asn Ile Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly  
 160 165

20

atg cca gcc tca act gaa gga gca tat gtg tct tca gag 545  
 Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu  
 170 175 180

25

tct ccc att aga ata tca gta tcc aca gaa gga gca aat 584  
 Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn  
 185 190

30

act tct tca tct aca tct aca tcc acc act ggg aca agc 623  
 Thr Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser  
 195 200 205

35

cat ctt gta aaa tgt gcg gag aag gag aaa act ttg tgt 662  
 His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys  
 210 215 220

40

gtg aat gga ggg gag tgc ttc atg gtg aaa gac ctt tca 701  
 Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser  
 225 230

45

aac ccc tcg aga tac ttg tgc aag tgc cca aat gag ttt 740  
 Asn Pro Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe  
 235 240 245

50

act ggt gat cgc tgc caa aac tac gta atg gcc agc ttc 779  
 Thr Gly Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe  
 250 255

55

tac aag cat ctt ggg att gaa ttt atg gag gcg gag gag 818  
 Tyr Lys His Leu Gly Ile Glu Phe Met Glu Ala Glu Glu  
 260 265 270

60

ctg tac cag aag aga gtg ctg acc ata acc ggc atc tgc 857  
 Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys  
 275 280 285

65

atc gcc ctc ctt gtg gtc ggc atc atg tgt gtg gtg gcc 896  
 Ile Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala  
 290 295

# ES 2 329 467 T3

tac tgc aaa acc aag aaa cag cgg aaa aag ctg cat gac 935  
 Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp  
 300 305 310

5

cgt ctt cgg cag agc ctt cgg tct gaa cga aac aat atg 974  
 Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn Met  
 315 320

10

atg aac att gcc aat ggg cct cac cat cct aac cca ccc 1013  
 Met Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro  
 325 330 335

15

ccc gag aat gtc cag ctg gtg aat caa tac gta tct aaa 1052  
 Pro Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys  
 340 345 350

20

aac gtc atc tcc agt gag cat att gtt gag aga gaa gca 1091  
 Asn Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala  
 355 360

25

gag aca tcc ttt tcc acc agt cac tat act tcc aca gcc 1130  
 Glu Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr Ala  
 365 370 375

30

cat cac tcc act act gtc acc cag act cct agc cac agc 1169  
 His His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser  
 380 385

35

tgg agc aac gga cac act gaa agc atc ctt tcc gaa agc 1208  
 Trp Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser  
 390 395 400

40

cac tct gta atc gtg atg tca tcc gta gaa aac agt agg 1247  
 His Ser Val Ile Val Met Ser Ser Val Glu Asn Ser Arg  
 405 410 415

45

cac agc agc cca act ggg ggc cca aga gga cgt ctt aat 1286  
 His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn  
 420 425

50

ggc aca gga ggc cct cgt gaa tgt aac agc ttc ctc agg 1325  
 Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg  
 430 435 440

55

cat gcc aga gaa acc cct gat tcc tac cga gac tct cct 1364  
 His Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro  
 445 450

60

cat agt gaa agg tat gtg tca gcc atg acc acc ccg gct 1403  
 His Ser Glu Arg Tyr Val Ser Ala Met Thr Thr Pro Ala  
 455 460 465

65

cgt atg tca cct gta gat ttc cac acg cca agc tcc ccc 1442  
 Arg Met Ser Pro Val Asp Phe His Thr Pro Ser Ser Pro  
 470 475 480

# ES 2 329 467 T3

aaa tcg ccc cct tcg gaa atg tct cca ccc gtg tcc agc 1481  
 Lys Ser Pro Pro Ser Glu Met Ser Pro Pro Val Ser Ser  
 485 490

5

atg acg gtg tcc atg cct tcc atg gcg gtc agc ccc ttc 1520  
 Met Thr Val Ser Met Pro Ser Met Ala Val Ser Pro Phe  
 495 500 505

10

atg gaa gaa gag aga cct cta ctt ctc gtg aca cca cca 1559  
 Met Glu Glu Glu Arg Pro Leu Leu Leu Val Thr Pro Pro  
 510 515

15

agg ctg cgg gag aag aag ttt gac cat cac cct cag cag 1598  
 Arg Leu Arg Glu Lys Lys Phe Asp His His Pro Gln Gln  
 520 525 530

20

ttc agc tcc ttc cac cac aac ccc gcg cat gac agt aac 1637  
 Phe Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn  
 535 540 545

25

agc ctc cct gct agc ccc ttg agg ata gtg gag gat gag 1676  
 Ser Leu Pro Ala Ser Pro Leu Arg Ile Val Glu Asp Glu  
 550 555

30

gag tat gaa acg acc caa gag tac gag cca gcc caa gag 1715  
 Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro Ala Gln Glu  
 560 565 570

35

cct gtt aag aaa ctc gcc aat agc cgg cgg gcc aaa aga 1754  
 Pro Val Lys Lys Leu Ala Asn Ser Arg Arg Ala Lys Arg  
 575 580

40

acc aag ccc aat ggc cac att gct aac aga ttg gaa gtg 1793  
 Thr Lys Pro Asn Gly His Ile Ala Asn Arg Leu Glu Val  
 585 590 595

45

gac agc aac aca agc tcc cag agc agt aac tca gag agt 1832  
 Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn Ser Glu Ser  
 600 605 610

50

gaa aca gaa gat gaa aga gta ggt gaa gat acg cct ttc 1871  
 Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr Pro Phe  
 615 620

55

ctg ggc ata cag aac ccc ctg gca gcc agt ctt gag gca 1910  
 Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala  
 625 630 635

60

aca cct gcc ttc cgc ctg gct gac agc agg act aac cca 1949  
 Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro  
 640 645

65

gca ggc cgc ttc tcg aca cag gaa gaa atc cag gcc agg 1988  
 Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile Gln Ala Arg  
 650 655 660

# ES 2 329 467 T3

ctg tct agt gta att gct aac caa gac cct att gct gta ta 2029  
 Leu Ser Ser Val Ile Ala Asn Gln Asp Pro Ile Ala Val  
           665                                670                                675

5

a aacctaaata aacacataga ttcacctgta aaactttatt 2070

10

ttatataata aagtattcca ccttaaatta aacaatttat tttatttttag 2120

cagttctgca aatagaaaac aggaaaaaaaa cttttataaa ttaaatatat 2170

15

gtatgtaaaa atgaaaaaaaa aaaaaaaaaa 2199

<210> 5

20 <211> 637

<212> PRT

<213> *Homo sapiens*

25

30

35

40

45

50

55

60

65

# ES 2 329 467 T3

<400> 5

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

# ES 2 329 467 T3

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

# ES 2 329 467 T3

|    |                                                             |         |
|----|-------------------------------------------------------------|---------|
|    | Pro Arg Leu Arg Glu Lys Lys Phe Asp His His Pro Gln Gln Phe |         |
|    | 485                                                         | 490 495 |
| 5  | Ser Ser Phe His His Asn Pro Ala His Asp Ser Asn Ser Leu Pro |         |
|    | 500                                                         | 505 510 |
|    | Ala Ser Pro Leu Arg Ile Val Glu Asp Glu Glu Tyr Glu Thr Thr |         |
|    | 515                                                         | 520 525 |
| 10 | Gln Glu Tyr Glu Pro Ala Gln Glu Pro Val Lys Lys Leu Ala Asn |         |
|    | 530                                                         | 535 540 |
|    | Ser Arg Arg Ala Lys Arg Thr Lys Pro Asn Gly His Ile Ala Asn |         |
|    | 545                                                         | 550 555 |
| 15 | Arg Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn Ser |         |
|    | 560                                                         | 565 570 |
|    | Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp Thr Pro Phe |         |
| 20 | 575                                                         | 580 585 |
|    | Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser Leu Glu Ala Thr Pro |         |
|    | 590                                                         | 595 600 |
|    | Ala Phe Arg Leu Ala Asp Ser Arg Thr Asn Pro Ala Gly Arg Phe |         |
| 25 | 605                                                         | 610 615 |
|    | Ser Thr Gln Glu Glu Ile Gln Ala Arg Leu Ser Ser Val Ile Ala |         |
|    | 620                                                         | 625 630 |
| 30 | Asn Gln Asp Pro Ile Ala Val                                 |         |
|    | 635 637                                                     |         |

<210> 6

<211> 2490

35 <212> DNA

<213> *Homo sapiens*

<400> 6

|    |                                                            |
|----|------------------------------------------------------------|
| 40 | gtggctgcgg ggcaattgaa aaagagccgg cgaggagttc cccgaaactt 50  |
|    | gttggaactc cgggctcgcg cggaggccag gagctgagcg gcggcggtcg 100 |
| 45 | ccggacgatg ggagcgtgag caggacggtg ataacctctc cccgatcggg 150 |
|    | ttgcgagggc gccgggcaga ggccaggacg cgagccgcca gcggcgggac 200 |
| 50 | ccatcgacga cttcccgggg cgacaggagc agccccgaga gccagggcga 250 |
|    | gcgcccgttc caggtggccg gaccgcccgc cgcgtccgcg ccgcgctccc 300 |
| 55 | tgcaggcaac gggagacgcc cccgcgcagc gcgagcgctt cagcgcggcc 350 |
|    | gctcgtcttc cccatcgagg gacaaacttt tcccaaacc gatccgagcc 400  |
| 60 | cttggaacaa actcgcctgc gccgagagcc gtccgcgtag agcgtccgt 450  |

65

# ES 2 329 467 T3

ctccggcgag atg tcc gag cgc aaa gaa ggc aga ggc aaa 490  
Met Ser Glu Arg Lys Glu Gly Arg Gly Lys  
1 5 10

5 ggg aag ggc aag aag aag gag cga ggc tcc ggc aag aag 529  
Gly Lys Gly Lys Lys Lys Glu Arg Gly Ser Gly Lys Lys  
15 20

10 ccg gag tcc gcg gcg ggc agc cag agc cca gcc ttg cct 568  
Pro Glu Ser Ala Ala Gly Ser Gln Ser Pro Ala Leu Pro  
25 30 35

15 ccc caa ttg aaa gag atg aaa agc cag gaa tcg gct gca 607  
Pro Gln Leu Lys Glu Met Lys Ser Gln Glu Ser Ala Ala  
40 45

20 ggt tcc aaa cta gtc ctt cgg tgt gaa acc agt tct gaa 646  
Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser Ser Glu  
50 55 60

25 tac tcc tct ctc aga ttc aag tgg ttc aag aat ggg aat 685  
Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn  
65 70 75

30 gaa ttg aat cga aaa aac aaa cca caa aat atc aag ata 724  
Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile  
80 85

35 caa aaa aag cca ggg aag tca gaa ctt cgc att aac aaa 763  
Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys  
90 95 100

40 gca tca ctg gct gat tct gga gag tat atg tgc aaa gtg 802  
Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys Lys Val  
105 110

45 atc agc aaa tta gga aat gac agt gcc tct gcc aat atc 841  
Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala Asn Ile  
115 120 125

50 acc atc gtg gaa tca aac gag atc atc act ggt atg cca 880  
Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly Met Pro  
130 135 140

55 gcc tca act gaa gga gca tat gtg tct tca gag tct ccc 919  
Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu Ser Pro  
145 150

60 att aga ata tca gta tcc aca gaa gga gca aat act tct 958  
Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr Ser  
155 160 165

65 tca tct aca tct aca tcc acc act ggg aca agc cat ctt 997  
Ser Ser Thr Ser Thr Thr Thr Gly Thr Ser His Leu  
170 175

# ES 2 329 467 T3

gta aaa tgt gcg gag aag gag aaa act ttc tgt gtg aat 1036  
 Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn  
 180 185 190

5

gga ggg gag tgc ttc atg gtg aaa gac ctt tca aac ccc 1075  
 Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn Pro  
 195 200 205

10

tcg aga tac ttg tgc aag tgc cca aat gag ttt act ggt 1114  
 Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe Thr Gly  
 210 215

15

gat cgc tgc caa aac tac gta atg gcc agc ttc tac aag 1153  
 Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe Tyr Lys  
 220 225 230

20

gcg gag gag ctg tac cag aag aga gtg ctg acc ata acc 1192  
 Ala Glu Glu Leu Tyr Gln Lys Arg Val Leu Thr Ile Thr  
 235 240

25

ggc atc tgc atc gcc ctc ctt gtg gtc ggc atc atg tgt 1231  
 Gly Ile Cys Ile Ala Leu Leu Val Val Gly Ile Met Cys  
 245 250 255

30

gtg gtg gcc tac tgc aaa acc aag aaa cag cgg aaa aag 1270  
 Val Val Ala Tyr Cys Lys Thr Lys Lys Gln Arg Lys Lys  
 260 265 270

35

ctg cat gac cgt ctt cgg cag agc ctt cgg tct gaa cga 1309  
 Leu His Asp Arg Leu Arg Gln Ser Leu Arg Ser Glu Arg  
 275 280

40

aac aat atg atg aac att gcc aat ggg cct cac cat cct 1348  
 Asn Asn Met Met Asn Ile Ala Asn Gly Pro His His Pro  
 285 290 295

45

aac cca ccc ccc gag aat gtc cag ctg gtg aat caa tac 1387  
 Asn Pro Pro Pro Glu Asn Val Gln Leu Val Asn Gln Tyr  
 300 305

50

gta tct aaa aac gtc atc tcc agt gag cat att gtt gag 1426  
 Val Ser Lys Asn Val Ile Ser Ser Glu His Ile Val Glu  
 310 315 320

55

aga gaa gca gag aca tcc ttt tcc acc agt cac tat act 1465  
 Arg Glu Ala Glu Thr Ser Phe Ser Thr Ser His Tyr Thr  
 325 330 335

60

tcc aca gcc cat cac tcc act act gtc acc cag act cct 1504  
 Ser Thr Ala His His Ser Thr Thr Val Thr Gln Thr Pro  
 340 345

65

agc cac agc tgg agc aac gga cac act gaa agc atc ctt 1543  
 Ser His Ser Trp Ser Asn Gly His Thr Glu Ser Ile Leu  
 350 355 360

# ES 2 329 467 T3

tcc gaa agc cac tct gta atc gtg atg tca tcc gta gaa 1582  
 Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val Glu  
 365 370

5 aac agt agg cac agc agc cca act ggg ggc cca aga gga 1621  
 Asn Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly  
 375 380 385

10 cgt ctt aat ggc aca gga ggc cct cgt gaa tgt aac agc 1660  
 Arg Leu Asn Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser  
 390 395 400

15 ttc ctc agg cat gcc aga gaa acc cct gat tcc tac cga 1699  
 Phe Leu Arg His Ala Arg Glu Thr Pro Asp Ser Tyr Arg  
 405 410

20 gac tct cct cat agt gaa agg tat gtg tca gcc atg acc 1738  
 Asp Ser Pro His Ser Glu Arg Tyr Val Ser Ala Met Thr  
 415 420 425

acc ccg gct cgt atg tca cct gta gat ttc cac acg cca 1777  
 Thr Pro Ala Arg Met Ser Pro Val Asp Phe His Thr Pro  
 430 435

25 agc tcc ccc aaa tcg ccc cct tcg gaa atg tct cca ccc 1816  
 Ser Ser Pro Lys Ser Pro Pro Ser Glu Met Ser Pro Pro  
 440 445 450

30 gtg tcc agc atg acg gtg tcc aag cct tcc atg gcg gtc 1855  
 Val Ser Ser Met Thr Val Ser Lys Pro Ser Met Ala Val  
 455 460 465

35 agc ccc ttc atg gaa gaa gag aga cct cta ctt ctc gtg 1894  
 Ser Pro Phe Met Glu Glu Glu Arg Pro Leu Leu Leu Val  
 470 475

40 aca cca cca agg ctg cgg gag aag aag ttt gac cat cac 1933  
 Thr Pro Pro Arg Leu Arg Glu Lys Lys Phe Asp His His  
 480 485 490

cct cag cag ttc agc tcc ttc cac cac aac ccc gcg cat 1972  
 Pro Gln Gln Phe Ser Ser Phe His His Asn Pro Ala His  
 495 500

45 gac agt aac agc ctc cct gct agc ccc ttg agg ata gtg 2011  
 Asp Ser Asn Ser Leu Pro Ala Ser Pro Leu Arg Ile Val  
 505 510 515

50 gag gat gag gag tat gaa acg acc caa gag tac gag cca 2050  
 Glu Asp Glu Glu Tyr Glu Thr Thr Gln Glu Tyr Glu Pro  
 520 525 530

55 gcc caa gag cct gtt aag aaa ctc gcc aat agc cgg cgg 2089  
 Ala Gln Glu Pro Val Lys Lys Leu Ala Asn Ser Arg Arg  
 535 540

# ES 2 329 467 T3

gcc aaa aga acc aag ccc aat ggc cac att gct aac aga 2128  
 Ala Lys Arg Thr Lys Pro Asn Gly His Ile Ala Asn Arg  
 545 550 555  
 5  
 ttg gaa gtg gac agc aac aca agc tcc cag agc agt aac 2167  
 Leu Glu Val Asp Ser Asn Thr Ser Ser Gln Ser Ser Asn  
 560 565  
 10  
 tca gag agt gaa aca gaa gat gaa aga gta ggt gaa gat 2206  
 Ser Glu Ser Glu Thr Glu Asp Glu Arg Val Gly Glu Asp  
 570 575 580  
 15  
 acg cct ttc ctg ggc ata cag aac ccc ctg gca gcc agt 2245  
 Thr Pro Phe Leu Gly Ile Gln Asn Pro Leu Ala Ala Ser  
 585 590 595  
 20  
 ctt gag gca aca cct gcc ttc cgc ctg gct gac agc agg 2284  
 Leu Glu Ala Thr Pro Ala Phe Arg Leu Ala Asp Ser Arg  
 600 605  
 25  
 act aac cca gca ggc cgc ttc tcg aca cag gaa gaa atc 2323  
 Thr Asn Pro Ala Gly Arg Phe Ser Thr Gln Glu Glu Ile  
 610 615 620  
 30  
 cag gcc agg ctg tct agt gta att gct aac caa gac cct 2362  
 Gln Ala Arg Leu Ser Ser Val Ile Ala Asn Gln Asp Pro  
 625 630  
 35  
 att gct gta taaaaccta aataaacaca tagattcacc tgtaaaactt 2410  
 Ile Ala Val  
 635 637  
 40  
 tattttatat aataaagtat tccaccttaa attaaacaat ttattttatt 2460  
 45  
 ttagcagttc tgcaaataaa aaaaaaaaaa 2490

50 <210> 7  
 <211> 241  
 <212> PRT  
 55 <213> *Homo sapiens*

60

65

# ES 2 329 467 T3

<400> 7

|    |                                                             |  |
|----|-------------------------------------------------------------|--|
|    | Met Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys Lys |  |
|    | 1 5 10 15                                                   |  |
| 5  | Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala Ala Gly Ser |  |
|    | 20 25 30                                                    |  |
| 10 | Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys Glu Met Lys Ser Gln |  |
|    | 35 40 45                                                    |  |
|    | Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser |  |
|    | 50 55 60                                                    |  |
| 15 | Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly Asn |  |
|    | 65 70 75                                                    |  |
| 20 | Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys |  |
|    | 80 85 90                                                    |  |
|    | Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala |  |
|    | 95 100 105                                                  |  |
| 25 | Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu Gly Asn |  |
|    | 110 115 120                                                 |  |
|    | Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu Ser Asn Glu Ile |  |
|    | 125 130 135                                                 |  |
| 30 | Ile Thr Gly Met Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser |  |
|    | 140 145 150                                                 |  |
|    | Glu Ser Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr |  |
|    | 155 160 165                                                 |  |
| 35 | Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His Leu Val |  |
|    | 170 175 180                                                 |  |
| 40 | Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu |  |
|    | 185 190 195                                                 |  |
|    | Cys Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys |  |
|    | 200 205 210                                                 |  |
| 45 | Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln Asn Tyr Val |  |
|    | 215 220 225                                                 |  |
|    | Met Ala Ser Phe Tyr Ser Thr Ser Thr Pro Phe Leu Ser Leu Pro |  |
|    | 230 235 240                                                 |  |
| 50 | Glu                                                         |  |
|    | 241                                                         |  |

<210> 8

<211> 1715

<212> DNA

<213> *Homo sapiens*

# ES 2 329 467 T3

<400> 8

```

ggcgcctgcct ccaacctgcg ggcgggaggt ggggtgctgc ggggcaattg 50
aaaaagagcc ggcgaggagt tccccgaaac ttgttggaac tccgggctcg 100
5  cgcgagggcc aggagctgag cggcggcggc tgccggacga tgggagcgtg 150
agcaggacgg tgataacctc tccccgatcg ggttgcgagg gcgccgggca 200
10  gaggccagga cgcgagccgc cagcggcggg acccatcgac gacttcccgg 250
ggcgacagga gcagccccga gagccagggc gagcgcccg tccaggtggc 300
cggaccgccc gccgcgtccg cgcgcgctc cctgcaggca acgggagacg 350
15  cccccgcgca gcgcgagcgc ctccgcggc cgcgtcgctc tccccatga 400
gggacaaact tttcccaaac ccgatccgag cccttggaac aaactcgct 450
gcgcgagag cgcgtccgct agagcgctcc gtctccggcg ag atg 495
20                                     Met
                                     1

tcc gag cgc aaa gaa ggc aga ggc aaa ggg aag ggc aag 534
Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys
5                                     10

aag aag gag cga ggc tcc ggc aag aag ccg gag tcc gcg 573
Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala
15                                     20                                     25

gcg ggc agc cag agc cca gcc ttg cct ccc caa ttg aaa 612
Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys
30                                     35                                     40

gag atg aaa agc cag gaa tcg gct gca ggt tcc aaa cta 651
Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu
35                                     45                                     50

gtc ctt cgg tgt gaa acc agt tct gaa tac tcc tct ctc 690
Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu
40                                     55                                     60                                     65

aga ttc aag tgg ttc aag aat ggg aat gaa ttg aat cga 729
Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg
70                                     75

aaa aac aaa cca caa aat atc aag ata caa aaa aag cca 768
Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro
80                                     85                                     90

ggg aag tca gaa ctt cgc att aac aaa gca tca ctg gct 807
Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala
95                                     100                                     105

gat tct gga gag tat atg tgc aaa gtg atc agc aaa tta 846
Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu
110                                     115

gga aat gac agt gcc tct gcc aat atc acc atc gtg gaa 885
Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu
120                                     125                                     130

tca aac gag atc atc act ggt atg cca gcc tca act gaa 924
Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu
135                                     140

gga gca tat gtg tct tca gag tct ccc att aga ata tca 963
Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser
145                                     150                                     155

```

# ES 2 329 467 T3

gta tcc aca gaa gga gca aat act tct tca tct aca tct 1002  
Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser  
160 165 170

5 aca tcc acc act ggg aca agc cat ctt gta aaa tgt gcg 1041  
Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala  
175 180

gag aag gag aaa act ttc tgt gtg aat gga ggg gag tgc 1080  
Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys  
10 185 190 195

ttc atg gtg aaa gac ctt tca aac ccc tcg aga tac ttg 1119  
Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu  
200 205

15 tgc aag tgc cca aat gag ttt act ggt gat cgc tgc caa 1158  
Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln  
210 215 220

aac tac gta atg gcc agc ttc tac agt acg tcc act ccc 1197  
Asn Tyr Val Met Ala Ser Phe Tyr Ser Thr Ser Thr Pro  
225 230 235

20 ttt ctg tct ctg cct gaa tagga gcatgctcag ttggtgctgc 1240  
Phe Leu Ser Leu Pro Glu  
240 241

tttcttgttg ctgcatctcc cctcagattc cacctagagc tagatgtgtc 1290

25 ttaccagatc taatattgac tgccctctgcc tgtcgcatga gaacattaac 1340

aaaagcaatt gtattacttc ctctgttcgc gactagtgg ctctgagata 1390

ctaataagggtg tgtgaggctc cggatgttcc tggaattgat attgaatgat 1440

30 gtgatacaaa ttgatagtca atatcaagca gtgaaatatg ataataaagg 1490

catttcaaag tctcactttt attgataaaa taaaaatcat tctactgaac 1540

agtcacatctt ctttatacaa tgaccacatc ctgaaaaggg tgttgctaag 1590

35 ctgtaaccga tatgcacttg aaatgatggt aagttaattt tgattcagaa 1640

tgtgttattt gtcacaaata aacataataa aaggagtcca gatgttttcc 1690

ttcattaacc aaaaaaaaaa aaaaa 1715

40 <210> 9  
<211> 420  
<212> PRT  
45 <213> *Homo sapiens*

50

55

60

65

# ES 2 329 467 T3

<400> 9

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

65

# ES 2 329 467 T3

Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn Val Ile  
 305 310 315  
 5 Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu Thr Ser Phe Ser  
 320 325 330  
 Thr Ser His Tyr Thr Ser Thr Ala His His Ser Thr Thr Val Thr  
 335 340 345  
 10 Gln Thr Pro Ser His Ser Trp Ser Asn Gly His Thr Glu Ser Ile  
 350 355 360  
 15 Leu Ser Glu Ser His Ser Val Ile Val Met Ser Ser Val Glu Asn  
 365 370 375  
 Ser Arg His Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn  
 380 385 390  
 20 Gly Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg His Ala  
 395 400 405  
 25 Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His Ser Glu Arg  
 410 415 420  
 <210> 10  
 <211> 2431  
 30 <212> DNA  
 <213> *Homo sapiens*  
 <400> 10  
 35 gaggcgcctg cctccaacct gcgggcggga ggtgggtggc tgcggggcaa 50  
 ttgaaaaaga gccggcgagg agttccccga aacttggttg aactccgggc 100  
 40 tcgcgcgggag gccaggagct gagcggcggc ggctgccgga cgatgggagc 150  
 gtgagcagga cgggtataac ctctccccga tcgggttgcg agggcgccgg 200  
 45 gcagaggcca ggacgcgagc cgccagcggc gggacccatc gacgacttcc 250  
 cggggcgaca ggagcagccc cgagagccag ggcgagcgcc cgttccaggt 300  
 50 ggccggaccg cccgcgcgt ccgcgccgcg ctccctgcag gcaacgggag 350  
 acgccccgc gcagcgcgag cgctcagcg cggccgctcg ctctcccat 400  
 cgagggacaa acttttccca aaccgatcc gagcccttg accaaactcg 450  
 55 cctgcgccga gagccgtccg cgtagagcgc tccgtctccg gcgag atg 498  
 Met  
 1  
 60 tcc gag cgc aaa gaa ggc aga ggc aaa ggg aag ggc aag 537  
 Ser Glu Arg Lys Glu Gly Arg Gly Lys Gly Lys Gly Lys  
 5 10

# ES 2 329 467 T3

aag aag gag cga ggc tcc ggc aag aag ccg gag tcc gcg 576  
 Lys Lys Glu Arg Gly Ser Gly Lys Lys Pro Glu Ser Ala  
 15 20 25

5

gcg ggc agc cag agc cca gcc ttg cct ccc caa ttg aaa 615  
 Ala Gly Ser Gln Ser Pro Ala Leu Pro Pro Gln Leu Lys  
 30 35 40

10

gag atg aaa agc cag gaa tcg gct gca ggt tcc aaa cta 654  
 Glu Met Lys Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu  
 45 50

15

gtc ctt cgg tgt gaa acc agt tct gaa tac tcc tct ctc 693  
 Val Leu Arg Cys Glu Thr Ser Ser Glu Tyr Ser Ser Leu  
 55 60 65

20

aga ttc aag tgg ttc aag aat ggg aat gaa ttg aat cga 732  
 Arg Phe Lys Trp Phe Lys Asn Gly Asn Glu Leu Asn Arg  
 70 75

25

aaa aac aaa cca caa aat atc aag ata caa aaa aag cca 771  
 Lys Asn Lys Pro Gln Asn Ile Lys Ile Gln Lys Lys Pro  
 80 85 90

30

ggg aag tca gaa ctt cgc att aac aaa gca tca ctg gct 810  
 Gly Lys Ser Glu Leu Arg Ile Asn Lys Ala Ser Leu Ala  
 95 100 105

35

gat tct gga gag tat atg tgc aaa gtg atc agc aaa tta 849  
 Asp Ser Gly Glu Tyr Met Cys Lys Val Ile Ser Lys Leu  
 110 115

40

gga aat gac agt gcc tct gcc aat atc acc atc gtg gaa 888  
 Gly Asn Asp Ser Ala Ser Ala Asn Ile Thr Ile Val Glu  
 120 125 130

45

tca aac gag atc atc act ggt atg cca gcc tca act gaa 927  
 Ser Asn Glu Ile Ile Thr Gly Met Pro Ala Ser Thr Glu  
 135 140

50

gga gca tat gtg tct tca gag tct ccc att aga ata tca 966  
 Gly Ala Tyr Val Ser Ser Glu Ser Pro Ile Arg Ile Ser  
 145 150 155

55

gta tcc aca gaa gga gca aat act tct tca tct aca tct 1005  
 Val Ser Thr Glu Gly Ala Asn Thr Ser Ser Ser Thr Ser  
 160 165 170

60

aca tcc acc act ggg aca agc cat ctt gta aaa tgt gcg 1044  
 Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala  
 175 180

65

gag aag gag aaa act ttc tgt gtg aat gga ggg gag tgc 1083  
 Glu Lys Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys  
 185 190 195

# ES 2 329 467 T3

ttc atg gtg aaa gac ctt tca aac ccc tcg aga tac ttg 1122  
 Phe Met Val Lys Asp Leu Ser Asn Pro Ser Arg Tyr Leu  
 200 205

5

tgc aag tgt cca aat gag ttt act ggt gat cgc tgc caa 1161  
 Cys Lys Cys Pro Asn Glu Phe Thr Gly Asp Arg Cys Gln  
 210 215 220

10

aac tac gta atg gcc agc ttc tac aag gcg gag gag ctg 1200  
 Asn Tyr Val Met Ala Ser Phe Tyr Lys Ala Glu Glu Leu  
 225 230 235

15

tac cag aag aga gtg ctg acc ata acc ggc atc tgc atc 1239  
 Tyr Gln Lys Arg Val Leu Thr Ile Thr Gly Ile Cys Ile  
 240 245

20

gcc ctc ctt gtg gtc ggc atc atg tgt gtg gtg gcc tac 1278  
 Ala Leu Leu Val Val Gly Ile Met Cys Val Val Ala Tyr  
 250 255 260

25

tgc aaa acc aag aaa cag cgg aaa aag ctg cat gac cgt 1317  
 Cys Lys Thr Lys Lys Gln Arg Lys Lys Leu His Asp Arg  
 265 270

30

ctt cgg cag agc ctt cgg tct gaa cga aac aat atg atg 1356  
 Leu Arg Gln Ser Leu Arg Ser Glu Arg Asn Asn Met Met  
 275 280 285

35

aac att gcc aat ggg cct cac cat cct aac cca ccc ccc 1395  
 Asn Ile Ala Asn Gly Pro His His Pro Asn Pro Pro Pro  
 290 295 300

40

gag aat gtc cag ctg gtg aat caa tac gta tct aaa aac 1434  
 Glu Asn Val Gln Leu Val Asn Gln Tyr Val Ser Lys Asn  
 305 310

45

gtc atc tcc agt gag cat att gtt gag aga gaa gca gag 1473  
 Val Ile Ser Ser Glu His Ile Val Glu Arg Glu Ala Glu  
 315 320 325

50

aca tcc ttt tcc acc agt cac tat act tcc aca gcc cat 1512  
 Thr Ser Phe Ser Thr Ser His Tyr Thr Ser Thr Ala His  
 330 335

55

cac tcc act act gtc acc cag act cct agc cac agc tgg 1551  
 His Ser Thr Thr Val Thr Gln Thr Pro Ser His Ser Trp  
 340 345 350

60

agc aac gga cac act gaa agc atc ctt tcc gaa agc cac 1590  
 Ser Asn Gly His Thr Glu Ser Ile Leu Ser Glu Ser His  
 355 360 365

65

tct gta atc gtg atg tca tcc gta gaa aac agt agg cac 1629  
 Ser Val Ile Val Met Ser Ser Val Glu Asn Ser Arg His  
 370 375

# ES 2 329 467 T3

agc agc cca act ggg ggc cca aga gga cgt ctt aat ggc 1668  
Ser Ser Pro Thr Gly Gly Pro Arg Gly Arg Leu Asn Gly  
380 385 390

aca gga ggc cct cgt gaa tgt aac agc ttc ctc agg cat 1707  
Thr Gly Gly Pro Arg Glu Cys Asn Ser Phe Leu Arg His  
395 400

gcc aga gaa acc cct gat tcc tac cga gac tct cct cat 1746  
Ala Arg Glu Thr Pro Asp Ser Tyr Arg Asp Ser Pro His  
405 410 415

agt gaa agg taaaa ccgaaggcaa agctactgca gaggagaaac 1790  
Ser Glu Arg  
420

tcagtcagag aatccctgtg agcacctgcg gtctcacctc aggaaatcta 1840

ctctaatacag aataaggggc ggcagttacc tgttctagga gtgctcctag 1890

ttgatgaagt catctctttg tttagcggaa cttatttctt ctgagcttct 1940

ctcgctgctcc cagtgactga caggcaacag actcttaaag agctgggatg 1990

ctttgatgcg gaaggtgcag cacatggagt ttccagctct ggccatgggc 2040

tcagaccac tcgggggtctc agtgcctca gttgtaacat tagagagatg 2090

gcatcaatgc ttgataagga ccttctata attccaattg ccagttatcc 2140

aaactctgat tcgggtggtcg agctggcctc gtgttcttat ctgctaacc 2190

tgtcttacct tccagctca gtttaagtcaa atcaagggt atgtcattgc 2240

tgaatgtcat ggggggcaac tgcttgccct ccacctata gtatctattt 2290

tatgaaattc caagaaggga tgaataaata aatctcttgg atgctgcgtc 2340

tggcagtctt cacgggtggt tttaaagca gaaaaaaaaa aaaaaaaaaa 2390

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 2431

<210> 11

<211> 768

<212> PRT

<213> *Homo sapiens*

<400> 11

Met Asp Val Lys Glu Arg Lys Pro Tyr Arg Ser Leu Thr Arg Arg  
1 5 10 15

Arg Asp Ala Glu Arg Arg Tyr Thr Ser Ser Ser Ala Asp Ser Glu  
20 25 30

Glu Gly Lys Ala Pro Gln Lys Ser Tyr Ser Ser Ser Glu Thr Leu  
35 40 45

# ES 2 329 467 T3

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

# ES 2 329 467 T3

|    |                                                             |     |     |     |
|----|-------------------------------------------------------------|-----|-----|-----|
|    | Lys Lys Pro Ser Lys Tyr Cys Asn Trp Lys Cys Ala Ala Leu Ser | 335 | 340 | 345 |
| 5  | Ala Ile Val Ile Ser Ala Thr Leu Val Ile Leu Leu Ala Tyr Phe | 350 | 355 | 360 |
|    | Val Ala Met His Leu Phe Gly Leu Asn Trp His Leu Gln Pro Met | 365 | 370 | 375 |
| 10 | Glu Gly Gln Met Tyr Glu Ile Thr Glu Asp Thr Ala Ser Ser Trp | 380 | 385 | 390 |
| 15 | Pro Val Pro Thr Asp Val Ser Leu Tyr Pro Ser Gly Gly Thr Gly | 395 | 400 | 405 |
|    | Leu Glu Thr Pro Asp Arg Lys Gly Lys Gly Thr Thr Glu Gly Lys | 410 | 415 | 420 |
| 20 | Pro Ser Ser Phe Phe Pro Glu Asp Ser Phe Ile Asp Ser Gly Glu | 425 | 430 | 435 |
|    | Ile Asp Val Gly Arg Arg Ala Ser Gln Lys Ile Pro Pro Gly Thr | 440 | 445 | 450 |
| 25 | Phe Trp Arg Ser Gln Val Phe Ile Asp His Pro Val His Leu Lys | 455 | 460 | 465 |
|    | Phe Asn Val Ser Leu Gly Lys Ala Ala Leu Val Gly Ile Tyr Gly | 470 | 475 | 480 |
| 30 | Arg Lys Gly Leu Pro Pro Ser His Thr Gln Phe Asp Phe Val Glu | 485 | 490 | 495 |
| 35 | Leu Leu Asp Gly Arg Arg Leu Leu Thr Gln Glu Ala Arg Ser Leu | 500 | 505 | 510 |
|    | Glu Gly Thr Pro Arg Gln Ser Arg Gly Thr Val Pro Pro Ser Ser | 515 | 520 | 525 |
| 40 | His Glu Thr Gly Phe Ile Gln Tyr Leu Asp Ser Gly Ile Trp His | 530 | 535 | 540 |
|    | Leu Ala Phe Tyr Asn Asp Gly Lys Glu Ser Glu Val Val Ser Phe | 545 | 550 | 555 |
| 45 | Leu Thr Thr Ala Ile Ala Leu Pro Pro Arg Leu Lys Glu Met Lys | 560 | 565 | 570 |
| 50 | Ser Gln Glu Ser Ala Ala Gly Ser Lys Leu Val Leu Arg Cys Glu | 575 | 580 | 585 |
|    | Thr Ser Ser Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn | 590 | 595 | 600 |
| 55 | Gly Asn Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys Ile | 605 | 610 | 615 |

# ES 2 329 467 T3

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
|    | Gln | Lys | Lys | Pro | Gly | Lys | Ser | Glu | Leu | Arg | Ile | Asn | Lys | Ala | Ser |  |
|    |     |     |     |     | 620 |     |     |     |     | 625 |     |     |     |     | 630 |  |
| 5  | Leu | Ala | Asp | Ser | Gly | Glu | Tyr | Met | Cys | Lys | Val | Ile | Ser | Lys | Leu |  |
|    |     |     |     |     | 635 |     |     |     |     | 640 |     |     |     |     | 645 |  |
|    | Gly | Asn | Asp | Ser | Ala | Ser | Ala | Asn | Ile | Thr | Ile | Val | Glu | Ser | Asn |  |
| 10 |     |     |     |     | 650 |     |     |     |     | 655 |     |     |     |     | 660 |  |
|    | Glu | Ile | Ile | Thr | Gly | Met | Pro | Ala | Ser | Thr | Glu | Gly | Ala | Tyr | Val |  |
|    |     |     |     |     | 665 |     |     |     |     | 670 |     |     |     |     | 675 |  |
| 15 | Ser | Ser | Glu | Ser | Pro | Ile | Arg | Ile | Ser | Val | Ser | Thr | Glu | Gly | Ala |  |
|    |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |     | 690 |  |
|    | Asn | Thr | Ser | Ser | Ser | Thr | Ser | Thr | Ser | Thr | Thr | Gly | Thr | Ser | His |  |
| 20 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     | 705 |  |
|    | Leu | Val | Lys | Cys | Ala | Glu | Lys | Glu | Lys | Thr | Phe | Cys | Val | Asn | Gly |  |
|    |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |  |
| 25 | Gly | Glu | Cys | Phe | Met | Val | Lys | Asp | Leu | Ser | Asn | Pro | Ser | Arg | Tyr |  |
|    |     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |  |
|    | Leu | Cys | Lys | Cys | Pro | Asn | Glu | Phe | Thr | Gly | Asp | Arg | Cys | Gln | Asn |  |
| 30 |     |     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |  |
|    | Tyr | Val | Met | Ala | Ser | Phe | Tyr | Ser | Thr | Ser | Thr | Pro | Phe | Leu | Ser |  |
|    |     |     |     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |  |
| 35 | Leu | Pro | Glu |     |     |     |     |     |     |     |     |     |     |     |     |  |
|    |     |     |     |     | 768 |     |     |     |     |     |     |     |     |     |     |  |

<210> 12

<211> 3111

40 <212> DNA

<213> *Homo sapiens*

<400> 12

|    |                 |             |             |             |             |     |
|----|-----------------|-------------|-------------|-------------|-------------|-----|
| 45 | gggtaccatg      | ggtcggtgag  | cgcggtttccc | gcctgagcgc  | aactagcggc  | 50  |
|    | gggtcgtggg      | cacctccaga  | aaagatcccg  | caccatcctc  | caggatccaa  | 100 |
| 50 | tggccttgga      | gagagggctg  | cagggccac   | ggacattgct  | gactcttcag  | 150 |
|    | aacgtgctga      | catggagcca  | ggtagactga  | aattatcatg  | tgtccaaatt  | 200 |
| 55 | aaaattgcat      | acttcaagga  | ttatttgaag  | gactattctt  | agaccctttt  | 250 |
|    | aagaagattt      | aaagaaaaac  | cactcggccc  | tgagtgcggc  | gaggaccctg  | 300 |
| 60 | tttgtggatg      | tggaggagcg  | cgggcccggag | gcc         | atg gac gtg | 342 |
|    |                 |             |             |             | Met Asp Val |     |
|    |                 |             |             |             | 1           |     |
| 65 | aag gag agg aag | cct tac cgc | tcg ctg acc | cgg cgc cgc |             | 381 |

# ES 2 329 467 T3

Lys Glu Arg Lys Pro Tyr Arg Ser Leu Thr Arg Arg Arg  
           5                          10                          15

5      gag gcc gag cgc cgc tac acc agc tcg tcc gcg gac agc 420  
       Asp Ala Glu Arg Arg Tyr Thr Ser Ser Ser Ala Asp Ser  
                   20                          25

10     gag gag ggc aaa gcc ccg cag aaa tcg tac agc tcc agc 459  
       Glu Glu Gly Lys Ala Pro Gln Lys Ser Tyr Ser Ser Ser  
           30                          35                          40

15     gag acc ctg aag gcc tac gac cag gac gcc cgc cta gcc 498  
       Glu Thr Leu Lys Ala Tyr Asp Gln Asp Ala Arg Leu Ala  
           45                          50                          55

20     tat ggc agc cgc gtc aag gac att gtg ccg cag gag gcc 537  
       Tyr Gly Ser Arg Val Lys Asp Ile Val Pro Gln Glu Ala  
                   60                          65

25     gag gaa ttc tgc cgc aca ggt gcc aac ttc acc ctg cgg 576  
       Glu Glu Phe Cys Arg Thr Gly Ala Asn Phe Thr Leu Arg  
           70                          75                          80

30     gag ctg ggg ctg gaa gaa gta acg ccc cct cac ggg acc 615  
       Glu Leu Gly Leu Glu Glu Val Thr Pro Pro His Gly Thr  
           85                          90

35     ctg tac cgg aca gac att ggc ctc ccc cac tgc ggc tac 654  
       Leu Tyr Arg Thr Asp Ile Gly Leu Pro His Cys Gly Tyr  
           95                          100                          105

40     tcc atg ggg gct ggc tct gat gcc gac atg gag gct gac 693  
       Ser Met Gly Ala Gly Ser Asp Ala Asp Met Glu Ala Asp  
           110                          115                          120

45     acg gtg ctg tcc cct gag cac ccc gtg cgt ctg tgg ggc 732  
       Thr Val Leu Ser Pro Glu His Pro Val Arg Leu Trp Gly  
                   125                          130

50     cgg agc aca cgg tca ggg cgc agc tcc tgc ctg tcc agc 771  
       Arg Ser Thr Arg Ser Gly Arg Ser Ser Cys Leu Ser Ser  
           135                          140                          145

55     cgg gcc aat tcc aat ctc aca ctc acc gac acc gag cat 810  
       Arg Ala Asn Ser Asn Leu Thr Leu Thr Asp Thr Glu His  
           150                          155

60     gaa aac act gag act gat cat ccg ggc ggc ctg cag aac 849  
       Glu Asn Thr Glu Thr Asp His Pro Gly Gly Leu Gln Asn  
           160                          165                          170

65     cac gcg cgg ctc cgg acg ccg ccg ccg ctc tcg cac 888  
       His Ala Arg Leu Arg Thr Pro Pro Pro Pro Leu Ser His  
           175                          180                          185

927    gcc cac acc ccc aac cag cac cac gcg gcc tcc att aac 927

# ES 2 329 467 T3

Ala His Thr Pro Asn Gln His His Ala Ala Ser Ile Asn  
190 195

5 tcc ctg aac cgg ggc aac ttc acg ccg agg agc aac ccc 966  
Ser Leu Asn Arg Gly Asn Phe Thr Pro Arg Ser Asn Pro  
200 205 210

10 agc ccg gcc ccc acg gac cac tcg ctc tcc gga gag ccc 1005  
Ser Pro Ala Pro Thr Asp His Ser Leu Ser Gly Glu Pro  
215 220

15 cct gcc ggc ggc gcc cag gag cct gcc cac gcc cag gag 1044  
Pro Ala Gly Gly Ala Gln Glu Pro Ala His Ala Gln Glu  
225 230 235

20 aac tgg ctg ctc aac agc aac atc ccc ctg gag acc aga 1083  
Asn Trp Leu Leu Asn Ser Asn Ile Pro Leu Glu Thr Arg  
240 245 250

aac cta ggc aag cag cca ttc cta ggg aca ttg cag gac 1122  
Asn Leu Gly Lys Gln Pro Phe Leu Gly Thr Leu Gln Asp  
255 260

25 aac ctc att gag atg gac att ctc ggc gcc tcc cgc cat 1161  
Asn Leu Ile Glu Met Asp Ile Leu Gly Ala Ser Arg His  
265 270 275

30 gat ggg gct tac agt gac ggg cac ttc ctc ttc aag cct 1200  
Asp Gly Ala Tyr Ser Asp Gly His Phe Leu Phe Lys Pro  
280 285

35 gga ggc acc tcc ccg ctc ttc tgc acc aca tca cca ggg 1239  
Gly Gly Thr Ser Pro Leu Phe Cys Thr Thr Ser Pro Gly  
290 295 300

tac cca ctg acg tcc agc aca gtg tac tct cct ccg ccc 1278  
Tyr Pro Leu Thr Ser Ser Thr Val Tyr Ser Pro Pro Pro  
305 310 315

40 cga ccc ctg ccc cgc agc acc ttc gcc ccg ccg gcc ttt 1317  
Arg Pro Leu Pro Arg Ser Thr Phe Ala Arg Pro Ala Phe  
320 325

45 aac ctc aag aag ccc tcc aag tac tgt aac tgg aag tgc 1356  
Asn Leu Lys Lys Pro Ser Lys Tyr Cys Asn Trp Lys Cys  
330 335 340

50 gca gcc ctg agc gcc atc gtc atc tca gcc act ctg gtc 1395  
Ala Ala Leu Ser Ala Ile Val Ile Ser Ala Thr Leu Val  
345 350

55 atc ctg ctg gca tac ttt gtg gcc atg cac ctg ttt ggc 1434  
Ile Leu Leu Ala Tyr Phe Val Ala Met His Leu Phe Gly  
355 360 365

cta aac tgg cac ctg cag ccg atg gag ggg cag atg tat 1473

# ES 2 329 467 T3

Ala His Thr Pro Asn Gln His His Ala Ala Ser Ile Asn  
190 195

5 tcc ctg aac cgg ggc aac ttc acg ccg agg agc aac ccc 966  
Ser Leu Asn Arg Gly Asn Phe Thr Pro Arg Ser Asn Pro  
200 205 210

10 agc ccg gcc ccc acg gac cac tcg ctc tcc gga gag ccc 1005  
Ser Pro Ala Pro Thr Asp His Ser Leu Ser Gly Glu Pro  
215 220

15 cct gcc ggc ggc gcc cag gag cct gcc cac gcc cag gag 1044  
Pro Ala Gly Gly Ala Gln Glu Pro Ala His Ala Gln Glu  
225 230 235

20 aac tgg ctg ctc aac agc aac atc ccc ctg gag acc aga 1083  
Asn Trp Leu Leu Asn Ser Asn Ile Pro Leu Glu Thr Arg  
240 245 250

25 aac cta ggc aag cag cca ttc cta ggg aca ttg cag gac 1122  
Asn Leu Gly Lys Gln Pro Phe Leu Gly Thr Leu Gln Asp  
255 260

30 gat ggg gct tac agt gac ggg cac ttc ctc ttc aag cct 1200  
Asp Gly Ala Tyr Ser Asp Gly His Phe Leu Phe Lys Pro  
280 285

35 gga ggc acc tcc ccg ctc ttc tgc acc aca tca cca ggg 1239  
Gly Gly Thr Ser Pro Leu Phe Cys Thr Thr Ser Pro Gly  
290 295 300

40 tac cca ctg acg tcc agc aca gtg tac tct cct ccg ccc 1278  
Tyr Pro Leu Thr Ser Ser Thr Val Tyr Ser Pro Pro Pro  
305 310 315

45 cga ccc ctg ccc cgc agc acc ttc gcc cgg ccg gcc ttt 1317  
Arg Pro Leu Pro Arg Ser Thr Phe Ala Arg Pro Ala Phe  
320 325

50 aac ctc aag aag ccc tcc aag tac tgt aac tgg aag tgc 1356  
Asn Leu Lys Lys Pro Ser Lys Tyr Cys Asn Trp Lys Cys  
330 335 340

55 gca gcc ctg agc gcc atc gtc atc tca gcc act ctg gtc 1395  
Ala Ala Leu Ser Ala Ile Val Ile Ser Ala Thr Leu Val  
345 350

60 atc ctg ctg gca tac ttt gtg gcc atg cac ctg ttt ggc 1434  
Ile Leu Leu Ala Tyr Phe Val Ala Met His Leu Phe Gly  
355 360 365

65 cta aac tgg cac ctg cag ccg atg gag ggg cag atg tat 1473

# ES 2 329 467 T3

Ser Glu Val Val Ser Phe Leu Thr Thr Ala Ile Ala Leu  
 550 555 560  
  
 5 cct ccc cga ttg aaa gag atg aaa agc cag gaa tcg gct 2058  
 Pro Pro Arg Leu Lys Glu Met Lys Ser Gln Glu Ser Ala  
 565 570 575  
  
 10 gca ggt tcc aaa cta gtc ctt cgg tgt gaa acc agt tct 2097  
 Ala Gly Ser Lys Leu Val Leu Arg Cys Glu Thr Ser Ser  
 580 585  
  
 15 gaa tac tcc tct ctc aga ttc aag tgg ttc aag aat ggg 2136  
 Glu Tyr Ser Ser Leu Arg Phe Lys Trp Phe Lys Asn Gly  
 590 595 600  
  
 aat gaa ttg aat cga aaa aac aaa cca caa aat atc aag 2175  
 Asn Glu Leu Asn Arg Lys Asn Lys Pro Gln Asn Ile Lys  
 605 610  
  
 20 ata caa aaa aag cca ggg aag tca gaa ctt cgc att aac 2214  
 Ile Gln Lys Lys Pro Gly Lys Ser Glu Leu Arg Ile Asn  
 615 620 625  
  
 25 aaa gca tca ctg gct gat tct gga gag tat atg tgc aaa 2253  
 Lys Ala Ser Leu Ala Asp Ser Gly Glu Tyr Met Cys Lys  
 630 635 640  
  
 30 gtg atc agc aaa tta gga aat gac agt gcc tct gcc aat 2292  
 Val Ile Ser Lys Leu Gly Asn Asp Ser Ala Ser Ala Asn  
 645 650  
  
 atc acc atc gtg gaa tca aac gag atc atc act ggt atg 2331  
 Ile Thr Ile Val Glu Ser Asn Glu Ile Ile Thr Gly Met  
 655 660 665  
  
 35 cca gcc tca act gaa gga gca tat gtg tct tca gag tct 2370  
 Pro Ala Ser Thr Glu Gly Ala Tyr Val Ser Ser Glu Ser  
 670 675  
  
 40 ccc att aga ata tca gta tcc aca gaa gga gca aat act 2409  
 Pro Ile Arg Ile Ser Val Ser Thr Glu Gly Ala Asn Thr  
 680 685 690  
  
 45 tct tca tct aca tct aca tcc acc act ggg aca agc cat 2448  
 Ser Ser Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser His  
 695 700 705  
  
 50 ctt gta aaa tgt gcg gag aag gag aaa act ttc tgt gtg 2487  
 Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys Val  
 710 715  
  
 aat gga ggg gag tgc ttc atg gtg aaa gac ctt tca aac 2526  
 Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser Asn  
 720 725 730  
  
 55 ccc tcg aga tac ttg tgc aag tgc cca aat gag ttt act 2565

# ES 2 329 467 T3

Pro Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe Thr  
735 740

5 ggt gat cgc tgc caa aac tac gta atg gcc agc ttc tac 2604  
Gly Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe Tyr  
745 750 755

10 agt acg tcc act ccc ttt ctg tct ctg cct gaa tag 2640  
Ser Thr Ser Thr Pro Phe Leu Ser Leu Pro Glu  
760 765 768

15 gagcatgctc agttggtgct gctttcttgt tgctgcatct cccctcagat 2690

20 tccacctaga gctagatgtg tcttaccaga tctaatttg actgcctctg 2740

25 cctgtcgcac gagaacatta acaaaagcaa ttgtattact tcctctgttc 2790

30 gcgactagtt ggctctgaga tactaatagg tgtgtgaggc tccggatggt 2840

35 tctggaattg atattgaatg atgtgataca aattgatagt caatatcaag 2890

40 cagtgaaata tgataataaa ggcatttcaa agtctcactt ttattgataa 2940

aataaaaatc attctactga acagtccatc ttctttatac aatgaccaca 2990

tcctgaaaag ggtggttgcta agctgtaacc gatatgcact tgaaatgatg 3040

gtaagttaat ttgattcag aatgtgttat ttgtcacaaa taaacataat 3090

aaaaggaaaa aaaaaaaaaa a 3111

<210> 13

<211> 1872

<212> DNA

<213> *Homo sapiens*

## ES 2 329 467 T3

<400> 13

5        gaattcggga cagcctctcc tgccgceget gctgctgccg ccgcegccac 50  
      cgccggctgg tcctccttct gcttttactt ctctgcatg acagttgttt 100  
10       tcttcactctg agcagacacc agcttcagat gctcgaggtg agaaacatgc 150  
      ctttcagttt gggctactgg ttctacttaat taatcagccg gcagctccgt 200  
15       cgatctatctt tcgtccctgt cctcttgacg agcccgggat ggtttggagt 250  
      agcatttaaa agaactagaa aagtggccca gaaacagcag cttaaagaat 300  
20       tattacgata tactttgatt ttgtagttgc taggagcttt tcttcccccc 350  
      ttgcatcttt ctgaactctt cttgatttta ataatggcct tggacttgga 400  
25       cgatttatcg atttccccct gtaagatgct gtatcatttg gttggggggg 450  
      cctctgcgtg gtaatggacc gtgagagcgg ccaggccttc ttctggaggt 500  
30  
  
35  
  
40  
  
45  
  
50  
  
55  
  
60  
  
65

# ES 2 329 467 T3

gagccg atg gag att tat tcc cca gac atg tct gag gtc 539  
 Met Glu Ile Tyr Ser Pro Asp Met Ser Glu Val  
 1 5 10

gcc gcc gag agg tcc tcc agc ccc tcc act cag ctg agt 578  
 Ala Ala Glu Arg Ser Ser Ser Pro Ser Thr Gln Leu Ser  
 15 20

gca gac cca tct ctt gat ggg ctt ccg gca gca gaa gac 617  
 Ala Asp Pro Ser Leu Asp Gly Leu Pro Ala Ala Glu Asp  
 25 30 35

atg cca gag ccc cag act gaa gat ggg aga acc cct gga 656  
 Met Pro Glu Pro Gln Thr Glu Asp Gly Arg Thr Pro Gly  
 40 45 50

ctc gtg ggc ctg gcc gtg ccc tgc tgt gcg tgc cta gaa 695  
 Leu Val Gly Leu Ala Val Pro Cys Cys Ala Cys Leu Glu  
 55 60

gct gag cgc ctg aga ggt tgc ctc aac tca gag aaa atc 734  
 Ala Glu Arg Leu Arg Gly Cys Leu Asn Ser Glu Lys Ile  
 65 70 75

tgc att gtc ccc atc ctg gct tgc ctg gtc agc ctc tgc 773  
 Cys Ile Val Pro Ile Leu Ala Cys Leu Val Ser Leu Cys  
 80 85

ctc tgc atc gcc ggc ctc aag tgg gta ttt gtg gac aag 812  
 Leu Cys Ile Ala Gly Leu Lys Trp Val Phe Val Asp Lys  
 90 95 100

atc ttt gaa tat gac tct cct act cac ctt gac cct ggg 851  
 Ile Phe Glu Tyr Asp Ser Pro Thr His Leu Asp Pro Gly  
 105 110 115

ggg tta ggc cag gac cct att att tct ctg gac gca act 890  
 Gly Leu Gly Gln Asp Pro Ile Ile Ser Leu Asp Ala Thr  
 120 125

gct gcc tca gct gtg tgg gtg tgg tct gag gca tac act 929  
 Ala Ala Ser Ala Val Trp Val Ser Ser Glu Ala Tyr Thr  
 130 135 140

tca cct gtc tct agg gct caa tct gaa agt gag gtt caa 968  
 Ser Pro Val Ser Arg Ala Gln Ser Glu Ser Glu Val Gln  
 145 150

gtt aca gtg caa ggt gac aag gct gtt gtc tcc ttt gaa 1007  
 Val Thr Val Gln Gly Asp Lys Ala Val Val Ser Phe Glu  
 155 160 165

cca tca gcg gca ccg aca ccg aag aat cgt att ttt gcc 1046  
 Pro Ser Ala Ala Pro Thr Pro Lys Asn Arg Ile Phe Ala  
 170 175 180

# ES 2 329 467 T3

ttt tct ttc ttg ccg tcc act gcg cca tcc ttc cct tca 1085  
 Phe Ser Phe Leu Pro Ser Thr Ala Pro Ser Phe Pro Ser  
 185 190  
 5  
 ccc acc cgg aac cct gag gtg aga acg ccc aag tca gca 1124  
 Pro Thr Arg Asn Pro Glu Val Arg Thr Pro Lys Ser Ala  
 195 200 205  
 10  
 act cag cca caa aca aca gaa act aat ctc caa act gct 1163  
 Thr Gln Pro Gln Thr Thr Glu Thr Asn Leu Gln Thr Ala  
 210 215  
 15  
 cct aaa ctt tct aca tct aca tcc acc act ggg aca agc 1202  
 Pro Lys Leu Ser Thr Ser Thr Ser Thr Thr Gly Thr Ser  
 220 225 230  
 20  
 cat ctt gta aaa tgt gcg gag aag gag aaa act ttc tgt 1241  
 His Leu Val Lys Cys Ala Glu Lys Glu Lys Thr Phe Cys  
 235 240 245  
 25  
 gtg aat gga ggg gag tgc ttc atg gtg aaa gac ctt tca 1280  
 Val Asn Gly Gly Glu Cys Phe Met Val Lys Asp Leu Ser  
 250 255  
 30  
 aac ccc tcg aga tac ttg tgc aag tgc cca aat gag ttt 1319  
 Asn Pro Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu Phe  
 260 265 270  
 35  
 act ggt gat cgc tgc caa aac tac gta atg gcc agc ttc 1358  
 Thr Gly Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe  
 275 280  
 40  
 tac agt acg tcc act ccc ttt ctg tct ctg cct gaa taggag 1400  
 Tyr Ser Thr Ser Thr Pro Phe Leu Ser Leu Pro Glu  
 285 290 295 296  
 45  
 catgctcagt tgggtgctgct ttcttggtgc tgcattctccc ctcagattcc 1450  
 50  
 acctagagct agatgtgtct taccagatct aatattgact gcctctgcct 1500  
 55  
 gtcgcatgag aacattaaca aaagcaattg tattacttcc tctgttcgag 1550  
 60  
 actagttggc tctgagatac taataggtgt gtgaggctcc ggatgtttct 1600  
 65  
 ggaattgata ttgaatgatg tgatacaaat tgatagtcaa tatcaagcag 1650  
 70  
 tgaaatatga taataaaggc atttcaaagt ctcactttta ttgataaaat 1700  
 75  
 aaaaatcatt ctactgaaca gtccatcttc ttatataaat gaccacatcc 1750  
 80  
 tgaaaagggt gttgctaagc tgtaaccgat atgcacttga aatgatggta 1800  
 85  
 agttaatttt gattcagaat gtgttatttg tcacaaataa acataataaa 1850  
 90  
 aggaaaaaaa aaaccggaat tc 1872

<210> 14

<211> 296

60 <212> PRT

<213> *Homo sapiens*

65

# ES 2 329 467 T3

<400> 14

|    |                                                             |  |
|----|-------------------------------------------------------------|--|
|    | Met Glu Ile Tyr Ser Pro Asp Met Ser Glu Val Ala Ala Glu Arg |  |
|    | 1 5 10 15                                                   |  |
| 5  | Ser Ser Ser Pro Ser Thr Gln Leu Ser Ala Asp Pro Ser Leu Asp |  |
|    | 20 25 30                                                    |  |
|    | Gly Leu Pro Ala Ala Glu Asp Met Pro Glu Pro Gln Thr Glu Asp |  |
| 10 | 35 40 45                                                    |  |
|    | Gly Arg Thr Pro Gly Leu Val Gly Leu Ala Val Pro Cys Cys Ala |  |
|    | 50 55 60                                                    |  |
| 15 | Cys Leu Glu Ala Glu Arg Leu Arg Gly Cys Leu Asn Ser Glu Lys |  |
|    | 65 70 75                                                    |  |
|    | Ile Cys Ile Val Pro Ile Leu Ala Cys Leu Val Ser Leu Cys Leu |  |
|    | 80 85 90                                                    |  |
| 20 | Cys Ile Ala Gly Leu Lys Trp Val Phe Val Asp Lys Ile Phe Glu |  |
|    | 95 100 105                                                  |  |
|    | Tyr Asp Ser Pro Thr His Leu Asp Pro Gly Gly Leu Gly Gln Asp |  |
| 25 | 110 115 120                                                 |  |
|    | Pro Ile Ile Ser Leu Asp Ala Thr Ala Ala Ser Ala Val Trp Val |  |
|    | 125 130 135                                                 |  |
| 30 | Ser Ser Glu Ala Tyr Thr Ser Pro Val Ser Arg Ala Gln Ser Glu |  |
|    | 140 145 150                                                 |  |
|    | Ser Glu Val Gln Val Thr Val Gln Gly Asp Lys Ala Val Val Ser |  |
|    | 155 160 165                                                 |  |
| 35 | Phe Glu Pro Ser Ala Ala Pro Thr Pro Lys Asn Arg Ile Phe Ala |  |
|    | 170 175 180                                                 |  |
|    | Phe Ser Phe Leu Pro Ser Thr Ala Pro Ser Phe Pro Ser Pro Thr |  |
| 40 | 185 190 195                                                 |  |
|    | Arg Asn Pro Glu Val Arg Thr Pro Lys Ser Ala Thr Gln Pro Gln |  |
|    | 200 205 210                                                 |  |
| 45 | Thr Thr Glu Thr Asn Leu Gln Thr Ala Pro Lys Leu Ser Thr Ser |  |
|    | 215 220 225                                                 |  |
|    | Thr Ser Thr Thr Gly Thr Ser His Leu Val Lys Cys Ala Glu Lys |  |
|    | 230 235 240                                                 |  |
| 50 | Glu Lys Thr Phe Cys Val Asn Gly Gly Glu Cys Phe Met Val Lys |  |
|    | 245 250 255                                                 |  |
|    | Asp Leu Ser Asn Pro Ser Arg Tyr Leu Cys Lys Cys Pro Asn Glu |  |
| 55 | 260 265 270                                                 |  |
|    | Phe Thr Gly Asp Arg Cys Gln Asn Tyr Val Met Ala Ser Phe Tyr |  |
|    | 275 280 285                                                 |  |
| 60 | Ser Thr Ser Thr Pro Phe Leu Ser Leu Pro Glu                 |  |
|    | 290 295 296                                                 |  |

65