



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 181 674**

51 Int. Cl.:  
**C12N 15/12** (2006.01)  
**C12N 15/63** (2006.01)  
**C12P 21/02** (2006.01)  
**C07K 14/47** (2006.01)  
**C07K 14/705** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA MODIFICADA

T5

86 Número de solicitud europea: **92911557 .4**  
86 Fecha de presentación : **01.05.1992**  
87 Número de publicación de la solicitud: **0576623**  
87 Fecha de publicación de la solicitud: **05.01.1994**

54 Título: **Dominios de unión en las proteínas *Notch*.**

30 Prioridad: **03.05.1991 US 695189**  
**14.11.1991 US 791923**

45 Fecha de publicación de la mención y de la traducción de patente europea: **01.03.2003**

45 Fecha de la publicación de la mención de la patente europea modificada BOPI: **01.03.2008**

45 Fecha de publicación de la traducción de patente europea modificada: **01.03.2008**

73 Titular/es: **Yale University**  
**216 Prospect Street**  
**New Haven, Connecticut 06511, US**  
**Advanced Research & Technology Institute**

72 Inventor/es: **Artavanis-Tsakonas, Spyridon;**  
**Muskavitch, Marc, Alan, Telander;**  
**Fehon, Richard, Grant;**  
**Rebay, Ilaria;**  
**Blaumueller, Christine, Marie y**  
**Shepard, Scott, Brockewell**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 181 674 T5

## DESCRIPCIÓN

Dominios de unión en las proteínas *Notch*.

5 Esta invención se hizo en parte con apoyo del gobierno con los números de concesión GM 19093 y NS 26084, otorgados por el Departamento de Sanidad y Servicios Humanos. El gobierno tiene ciertos derechos sobre la invención.

## 1. Introducción

10 La presente invención se refiere a los genes humanos *Notch* y a sus productos codificados. La invención también se refiere a secuencias (denominadas de aquí en adelante "secuencias adhesivas") dentro de las proteínas codificadas por genes humanos *Notch* que median en la unión heterotípica a secuencias dentro de proteínas codificadas por otros genes toporrítmicos. Tales genes incluyen, pero no están limitados a *Delta* y *Serrate*.

## 15 2. Antecedentes de la invención

Los análisis genéticos en *Drosophila* han sido extremadamente útiles para identificar la complejidad de las rutas de desarrollo y los loci interactuantes. Sin embargo, entender la naturaleza precisa de los procesos que sostienen las interacciones genéticas, requiere un conocimiento de las propiedades bioquímicas de los productos proteicos de los genes en cuestión.

25 Las mutaciones completas en uno cualquiera de los loci neurógenos cigóticos - *Notch* (*N*), *Delta* (*DI*), *mastermind* (*mam*), *Enhancer of Split* (*E(spl)*), *neuralized* (*neu*) y *big brain* (*bib*) - dan como resultado una hipertrofia del sistema nervioso a expensas de las estructuras epidérmicas ventrales y laterales. Este efecto es debido a la desviación de la ruta de las células precursoras epidérmicas a una ruta neuronal, e implica que la función de los genes neurógenos es necesaria para diversificar células dentro de la región neurógena desde un destino neuronal a un destino epitelial. Estudios que determinaban los efectos de la ablación con láser de neuroblastos embrionarios específicos, en saltamontes (Doe y Goodman 1985, *Dev. Biol.* 111, 206-219) han mostrado que interacciones celulares entre neuroblastos y células accesorias circundantes, sirven para inhibir que estas células accesorias adopten un destino de neuroblasto.

30 Estas observaciones genéticas y embriológicas han conducido ambas a la hipótesis de que los productos proteicos de los loci neurógenos funcionan como componentes de un mecanismo de interacción celular necesario para el correcto desarrollo epidérmico (Artavanis-Tsakonas, 1988, *Trends Genet.* 4, 95-100).

35 Los análisis de secuencias (Wharton y col., 1985, *Cell* 43, 567-581; Kidd y col., 1986, *Mol. Cell Biol.* 6, 3094-3108; Vassin y col., 1987, *EMBO J.* 6, 3431-3440; Kopczynski y col., 1988, *Genes Dev.* 2, 1723-1735) han mostrado que dos de los loci neurógenos, *Notch* y *Delta*, parecen codificar proteínas transmembranales que abarcan la membrana de una sola vez. El gen *Notch* codifica una proteína de ~300 kd (empleamos "Notch" para denominar esta proteína) con un dominio extracelular N-terminal extenso, que incluye 36 repeticiones en tándem similares al factor de crecimiento epidérmico (EGF), seguidas de otras tres repeticiones ricas en cisteína, denominadas repeticiones *Notch/lin-12* (Wharton y col., 1985, *Cell* 43, 567-581; Kidd y col., 1986, *Mol. Cell Biol.* 6, 3094-3108; Yochem y col., 1988, *Nature* 335, 547-550). *Delta* codifica una proteína de ~100 kd (empleamos "Delta" para denominar DLZM, el producto proteico de los transcritos cigóticos y maternos predominantes; Kopczynski y col., 1988, *Genes Dev.* 2, 1723-1735) que tiene nueve repeticiones similares a EGF en su dominio extracelular (Vassin y col., 1987, *EMBO J.* 6, 3431-3440; Kopczynski y col., 1988, *Genes Dev.* 2, 1723-1735). Aunque se conoce poco sobre la importancia funcional de estas repeticiones, el motivo similar a EGF se ha encontrado en una variedad de proteínas, incluyendo las implicadas en la cascada de la coagulación sanguínea (Furie y Furie, 1988, *Cell* 53, 505-518). En particular, este motivo se ha encontrado en proteínas extracelulares tales como los factores IX y X de la coagulación sanguínea (Rees y col., 1988, *EMBO J.* 7, 2053-2061; Furie y Furie, 1988, *Cell* 53, 505-518), en otros genes de *Drosophila* (Knust y col., 1987, *EMBO J.* 761-766; Rothberg y col., 1988, *Cell* 55, 1047-1059), y en algunas proteínas receptoras de la superficie celular, tales como el receptor de la trombomodulina (Suzuki y col., 1987, *EMBO J.* 6, 1891-1897) y el receptor de LDL (Sudhof y col., 1985, *Science* 228, 815-822). Se ha cartografiado un sitio de unión de proteína para el dominio de las repeticiones EGF en la trombomodulina y la uroquinasa (Kurosawa y col., 1988, *J. Biol. Chem.* 263, 5993-5996; Appella y col., 1987, *J. Biol. Chem.* 262, 4437-4440).

55 Se ha sido descrito un conjunto intrigante de interacciones entre mutaciones de *Notch* y *Delta* (Vassin y col., 1985, *J. Neurogenet.* 2, 292-308; Shepard y col., 1989, *Genetics* 122, 429-438; Xu y col., 1990, *Genes Dev.* 4, 464-475). Numerosos estudios genéticos (resumidos en Alton y col., 1989, *Dev. Genet.* 10, 261-272) han indicado que las dosis génicas de *Notch* y *Delta* en relación la una con la otra, son cruciales para el desarrollo normal. Una reducción del 50% de la dosis de *Delta* en un fondo *Notch* de tipo silvestre, provoca una extensión de las venas de las alas creando un "delta" en la base (Lindsley y Grell, 1968, Publication Number 627, Washington, D.C., Carnegie Institute of Washington). Un fenotipo similar es causado por un incremento del 50% de la dosis de *Notch* en un fondo *Delta* de tipo silvestre (un fenotipo de "confluencia"; Welshons, 1965, *Science*, 150, 1122-1129). Este fenotipo *Delta* se suprime parcialmente mediante una reducción de la dosis de *Notch*. Un trabajo reciente en nuestros laboratorios, ha mostrado que interacciones letales entre alelos que se correlacionan con alteraciones en las repeticiones similares a EGF en *Notch*, se pueden recuperar reduciendo las dosis de *Delta* (Xu y col., 1990, *Genes Dev.* 4, 464-475). Xu y col., (1990, *Genes Dev.* 4, 464-475) encontraron que mutaciones completas en *Delta* o *mam* suprimen las interacciones letales entre combinaciones heterocigotas de ciertos alelos *Notch*, conocidas como las mutaciones *Abruptex* (*Ax*). Los alelos

Ax están asociados con mutaciones sin sentido dentro de las repeticiones similares a EGF del dominio extracelular de Notch (Kelley y col., 1987, Cell 51, 539-548; Hartley y col., 1987, EMBO J. 6, 3407-3417).

5 Notch se expresa en los procesos axonales durante el crecimiento de neuronas embrionarias (Johansen y col., 1989, J. Cell Biol. 109, 2427-2440; Kidd y col., 1989, Genes Dev. 3, 1113-1129).

10 Un estudio ha mostrado que ciertos alelos Ax de *Notch* pueden alterar gravemente los datos de la ruta axonal durante el crecimiento neuronal sensitivo en los discos imaginales, aunque no se conoce hasta la fecha si la expresión de *Notch* aberrante es la responsable de este defecto en el axón mismo o en el epitelio a lo largo del cual crece (Palka y col., 1990, Development 109, 167-175).

### 3. Sumario de la invención

15 La presente invención se define en las reivindicaciones. En una realización específica de la invención, el fragmento adhesivo de Notch es el fragmento que comprende la secuencia de Notch más homóloga a las repeticiones similares a EGF 11 y 12 de Notch de *Drosophila*.

20 La presente memoria descriptiva describe las secuencias de nucleótidos de los genes *Notch* y *Delta* humanos, y secuencias de aminoácidos de sus proteínas codificadas, así como fragmentos de las mismas que contienen un determinante antigénico o que son funcionalmente activos. La memoria descriptiva también describe fragmentos (denominados en esta memoria "fragmentos adhesivos") y sus secuencias, de las proteínas ("proteínas toporrítmicas") codificadas por genes toporrítmicos que median en la unión homotípica o heterotípica a proteínas toporrítmicas. Los genes toporrítmicos, tal y como se emplean en esta memoria, se refieren a los genes *Notch*, *Delta* y *Serrate*, así como a otros miembros de la familia *Delta/Serrate* que se pueden identificar, por ejemplo, mediante métodos descritos en la Sección 5.3, más abajo. Se describen adicionalmente anticuerpos para Notch y para fragmentos adhesivos. En la siguiente descripción se hace referencia a una serie de genes toporrítmicos, incluyendo *Notch*, *Delta* y *Serrate*. Sin embargo, la presente invención se refiere solamente a proteínas relacionadas con Notch, fragmentos, proteínas químéricas, ácidos nucleicos, vectores, células hospedadoras, y métodos como se define en las reivindicaciones. Los otros genes toporrítmicos se describen para ayudar a entender la presente invención.

30 En los ejemplos específicos, el fragmento adhesivo de la unión heterotípica mediada por Delta, es el fragmento que comprende la secuencia más homóloga con los aminoácidos 1-230 de Delta de *Drosophila*; el fragmento adhesivo de la unión homotípica mediada por Delta, es el fragmento que comprende la secuencia más homóloga a los aminoácidos 32-230 de Delta de *Drosophila*; y el fragmento adhesivo de Serrate es el fragmento que comprende la secuencia más homóloga a los aminoácidos 85-283 o 79-282 de Serrate de *Drosophila*.

#### 3.1 Definiciones

40 Tal y como se emplea en esta memoria, las siguientes expresiones deben tener los significados indicados:

AA =	aminoácido
EGF =	factor de crecimiento epidérmico
45 ELR =	repetición similar a EGF (homóloga)
IC =	intracelular
PCR =	reacción en cadena de la polimerasa

50 Tal y como se emplea en esta memoria, el subrayado del nombre de un gen debe indicar al gen, en contraposición con su producto proteico codificado, que se indica con el nombre del gen sin ningún subrayado. Por ejemplo, "*Notch*" debe significar el gen *Notch*, mientras que "Notch" debe indicar el producto proteico del gen *Notch*.

### 55 4. Descripción de las figuras

Figura 1. Estructuras Artificiales de Expresión y Diseño Experimental para Examinar las Interacciones Notch-Delta. Se transfectaron de forma transitoria células S2 en la fase de crecimiento logarítmico con una de las tres estructuras artificiales mostradas. Notch codificada por el minigen MG11a (una estructura artificial quimérica de ADNc/genómico: 60 secuencias obtenidas a partir de ADNc se representan por un punteado, secuencias obtenidas genómicamente por un rayado diagonal (Ramos y col., 1989, Genetics 123, 337-348)) se expresó después de la inserción en el vector promotor de la metalotioneína pRmHa-3 (Bunch y col., 1988, Nucl. Acids Res. 16, 1043-1061). Delta codificada por el ADNc D11 (Kopczynski y col., 1988, Genes Dev. 2, 1723-1735) se expresó después de insertar en el mismo vector. La variante de Notch extracelular (ECN1) se obtuvo a partir de un cósmido genómico que contenía el locus completo de *Notch* 65 (Ramos y col., 1989, Genetics 123, 337-348) delecionando la secuencia codificadora de los aminoácidos 1790-2625 del dominio intracelular (indicado por  $\delta$ ; Wharton y col., 1985, Cell 43, 567-581), dejando 25 restos proximales a la membrana de la secuencia de tipo silvestre, fusionados con una cola de 59 aminoácidos nuevos (véase, Procedimientos Experimentales, Sección 6.1, más abajo). Esta estructura artificial se expresó bajo control de la región promotora de

## ES 2 181 674 T5

*Notch*. Para las estructuras artificiales que implicaban el vector de metalotioneína, la expresión se indujo con CuSO<sub>4</sub>, seguida de transfección. Las células se mezclaron a continuación, se incubaron bajo condiciones de agregación y se marcaron por su capacidad de agregarse, empleando antisueros específicos y microscopía de inmunofluorescencia para visualizar las células con expresión. MT, promotor de metalotioneína; ATG, sitio de inicio de la traducción; TM, dominio transmembranal; 3' N, señal de poliadenilación del gen *Notch*; 3' Adh, señal de poliadenilación del gen *Adh*, 5' N, región promotora del gen *Notch*.

Figura 2. Expresión de Notch y Delta en Células Cultivadas. (A) Material lisado de células no transfectadas (S2) y de células transfectadas con *Notch* (N), inducidas con CuSO<sub>4</sub> 0,7 mM, durante 12-16 h, se prepararon para electroforesis en gel de dodecil-sulfato sódico y poliacrilamida (SDS-PAGE), realizada en geles con gradiente de 3%-15%, y se realizó la transferencia a nitrocelulosa. Notch se visualizó empleando un anticuerpo monoclonal (MAb C17.9C6) contra el dominio intracelular de Notch. Bandas múltiples por debajo de la banda principal de 300 kd, pueden representar productos de degradación de Notch. (B) Material lisado de células no transfectadas (S2) y células transfectadas con *Delta* (DI) se visualizaron con un anticuerpo monoclonal (MAb 201) contra Delta. Se detecta una banda aislada de ~105 kd. En ambos casos, no hay Notch o Delta endógeno detectable en la línea celular S2, ni son especies con reactividad cruzada. En cada pista, se cargaron 10 µl de muestra (preparada según se describe en Procedimientos Experimentales).

Figura 3. Células S2 Que Expresan Notch y Delta Forman Agregados. En todos los paneles, Notch se muestra de color verde y Delta en rojo.

(A) Una célula aislada Notch<sup>+</sup>. Nótese la tinción intracelular visible, que incluye estructuras vesiculares así como un núcleo evidentemente sin teñir.

(A) Micrografía sobre fondo claro del mismo fondo, que muestra especificidad de la tinción del anticuerpo.

(B) Una célula aislada Delta<sup>+</sup>. La tinción está principalmente en la superficie celular.

(B) Micrografía de fondo claro del mismo fondo.

(C) Agregado de células Delta<sup>+</sup> procedentes de un experimento de agregación de 24 h. Nótese de nuevo que la tinción está principalmente en la superficie celular.

(D)-(F) Un agregado de células Notch<sup>+</sup> y Delta<sup>+</sup> formado a partir de una mezcla 1:1 de poblaciones celulares transfectadas una vez, que se dejaron agregar durante una noche a temperatura ambiente. (D) muestra células Notch<sup>+</sup> en este agregado; (E) muestra células Delta<sup>+</sup> y (F) es una exposición doble que muestra ambos tipos de células. Las bandas de Notch y Delta son prominentes en los puntos de contacto entre las células Notch<sup>+</sup> y Delta<sup>+</sup> (flechas). En (F), estas bandas aparecen de color amarillo debido a la coincidencia de verde y rojo en estos puntos. La célula aislada aparentemente con tinción doble (\*) es realmente dos células (una encima de la otra), una expresa Notch y la otra Delta.

(G) y (H) Micrografías confocales con pseudocolor de agregados de células Notch<sup>+</sup>-Delta<sup>+</sup>. Nótese que en (G) las extensiones (flechas) formadas por al menos dos células Delta<sup>+</sup>, circundan completamente la célula Notch<sup>+</sup> en el centro del agregado. (H) muestra un agregado formado a partir de un experimento de agregación durante 2 h, realizado a 4°C. Bandas intensas de Notch son evidentes dentro de regiones de contacto con células Delta<sup>+</sup>.

(I) Un agregado compuesto por células Delta<sup>+</sup> y células que expresan sólo el dominio extracelular de Notch (estructura artificial ECN1). Varilla graduada = 10 µm.

Figura 4. Notch y Delta están Asociadas en Células Cotransfectadas. La tinción de Notch se muestra en la columna izquierda (A, C y E) y la de Delta se muestra en la columna derecha (B, D y F).

(A) y (B) Célula S2 cotransfectada con ambas estructuras artificiales *Notch* y *Delta*. En general, había una buena correlación entre la localización de Notch y Delta en la superficie celular (flechas).

(C) y (D) Células cotransfectadas se expusieron a un antisuero policlonal anti-Notch (una dilución 1:250 de cada antisuero del dominio anti-extracelular) durante 1 h a temperatura ambiente antes de la fijación y la tinción con antisueros específicos. Nótese la tinción punteada de Notch y Delta y la correlación de su tinción respectiva (flechas).

(E) y (F) Células cotransfectadas con las estructuras artificiales de *Notch* (ECN1) y *Delta*, inducidas y después cultivadas en parches empleando antisueros policlonales anti-Notch. Había una estrecha correlación entre la tinción de ECN1 y Delta en la superficie, tal y como se observaba para Notch de longitud completa. Varilla graduada = 10 µm.

Figura 5. La Coimmunoprecipitación Muestra que Delta y Notch están Asociadas en Material Lisado procedente de Células S2 Transfectadas y Células Embrionarias de *Drosophila*. En todos los experimentos, Delta se precipitó a partir del material lisado con NP-40/desoxicolato empleando un antisuero policlonal anti-Delta de rata, precipitado con células fijadas con Staph A, y se visualizaron proteínas en la fracción precipitada con transferencias de tipo Western (para detalles, véanse los Procedimientos Experimentales). Pistas 1, 2, 3 y 5: Notch visible con MAb C17.9C6; pistas 4 y 6: Delta visible empleando MAb 201.

En (A), las pistas 1 y 2 eran testigos para estos experimentos. La pista 1 muestra una inmunoprecipitación con anti-Delta policlonal de células que expresan Notch, sólo hecha visible para Notch. No se podía detectar Notch en esta muestra, indicando que el anti-Delta policlonal no tenía reacción cruzada con Notch. La pista 2 muestra células cotransfectadas con *Notch/Delta*, inmunoprecipitadas con Staph A sin tratamiento inicial con antisuero anti-Delta y hechas visibles para Notch, demostrando que Notch no precipita inespecíficamente con Staph A o con anticuerpo secundario. La pista 3 muestra proteína precipitada con antisuero anti-Delta hecha visible para Delta (DI), y la pista 4 muestra la misma muestra hecha visible para Notch (N). La pista 4 muestra que Notch coprecipita con Delta inmunoprecipitada. Nótese que Notch aparece como un doblete, lo que es típico de Notch en inmunoprecipitados.

(B) muestra el mismo experimento empleando material lisado embrionario más que material lisado de células transfectadas. La pista 5 muestra proteína precipitada con antisuero anti-Delta hecha visible para Delta (DI), y la pista 6 muestra la misma muestra hecha visible para Notch (N). Estas pistas demuestran que Notch y Delta están asociadas de forma estable en material lisado embrionario. Las bandas (en todas las pistas) por debajo de la banda de Delta son de Staph A (SA) y las cadenas pesada (H) y ligera (L) del antisuero anti-Delta.

Figura 6. Estructuras Artificiales de Expresión de Notch y Cartografía de la Delección del Dominio de Unión de Delta/Serrate. Células S2 en la fase de crecimiento logarítmico se transfectaron transitoriamente con la serie de estructuras artificiales de expresión mostradas; los dibujos representan los productos proteicos pronosticados de los distintos mutantes creados por delección de *Notch*. Todas las estructuras artificiales de expresión se obtuvieron a partir de la estructura artificial nº 1 pMtNMg. Las células transfectadas de forma transitoria se mezclaron con células que expresaban Delta, procedentes de la línea transformada de forma estable L49-6-7 o con células que expresaban Serrate transfectadas de forma transitoria, se indujeron con CuSO<sub>4</sub>, se incubaron bajo condiciones de agregación y a continuación se marcaron por su capacidad de agregarse empleando antisueros específicos y microscopía de inmunofluorescencia. Los agregados se definieron como agrupaciones de cuatro o más células que contenían ambas células que expresaban Notch y Delta/Serrate. Los valores dados para el % de agregación se refieren al porcentaje de todas las células que expresan Notch encontradas en tales agrupaciones tanto con Delta (DI) (columna izquierda) como con Serrate (Ser) (columna derecha). Las distintas estructuras artificiales de la delección de Notch se representan en un diagrama con líneas de empalme que indican los puntos de unión de la ligación. Cada repetición EGF se representa por un rectángulo punteado y los números de las repeticiones EGF se marcan en cada lado de un punto de unión de ligación. En los puntos de unión de la ligación, las repeticiones EGF parciales producidas por distintas delecciones, se indican por rectángulos blancos y por corchetes negros (por ejemplo, véase nº 23  $\Delta$ Cl<sub>1</sub>+EGF(10-12)). Las estructuras artificiales nº 3-13 representan la serie de delecciones de Cl<sub>1</sub>. Tal y como se muestra en el diagrama, cuatro de los sitios Cl<sub>1</sub>, en las repeticiones 7, 9, 17 y 26, rompen la repetición en el centro, inmediatamente después de la tercera cisteína (indicada por repeticiones con un rectángulo blanco; véase Figura 7 para una mayor clarificación), mientras que la quinta y casi todo el sitio 3' se rompe limpiamente entre las repeticiones EGF 30 y 31 (indicadas por la repetición 31 con un rectángulo negro; véase de nuevo la Figura 7). En la estructura artificial nº 15 split, la repetición EGF 14 que es portadora de la mutación puntual *split*, se representa por un rectángulo rayado. En la estructura artificial nº 33  $\Delta$ Cl<sub>1</sub>+XEGF(10-13), las repeticiones EGF obtenidas a partir de Notch de *Xenopus*, se distinguen de las repeticiones de *Drosophila* por un patrón diferente de sombreado. SP, péptido señal; EGF, repetición del factor de crecimiento epidérmico; N, repetición *Notch/lin-12*; TM, dominio transmembranal; cdc10, repeticiones *cdc10/anquirina*; PA, secuencia de consenso de unión a nucleótido putativo; opa, extensión de poliglutamina denominada opa; DI, Delta; Ser, Serrate.

Figura 7. Estructura Detallada de las Estructuras Artificiales de Delección de Notch nº 19-24: Ambas Repeticiones EGF 11 y 12 son Necesarias para la Agregación de Notch-Delta. Las repeticiones EGF 10-13 se muestran en forma de diagrama en la parte superior, mostrando la separación regular de los seis restos de cisteína (C). Productos de la PCR, generados para estas estructuras artificiales (los nombres y los números se proporcionan en la Figura 6) se representan por las líneas negras oscuras y los puntos finales exactos se indican en relación con las distintas repeticiones EGF. La capacidad para agregarse con Delta se registra como (+) o (-) para cada estructura artificial. Los fragmentos de PCR, o bien rompen las repeticiones EGF en el centro, justo después de la tercera cisteína, en el mismo lugar que cuatro de los cinco sitios de Cl<sub>1</sub>, o bien exactamente entre dos repeticiones en el mismo lugar que la mayoría del sitio de Cl<sub>1</sub> más C-terminal.

Figura 8. Comparación de la Secuencia de Aminoácidos de las Repeticiones EGF 11 y 12 procedentes de Notch de *Drosophila* y de *Xenopus*. La secuencia de aminoácidos de las repeticiones EGF 11 y 12 de Notch de *Drosophila* (Wharton y col., 1985, Cell 43:567-581; Kidd y col., 1986, Mol. Cell Biol. 6:3094-3108) se alinean con las de las dos mismas repeticiones EGF procedentes de Notch de *Xenopus* (Coffman y col., 1990, Science 249:1438-1441). Los aminoácidos idénticos están incluidos en rectángulos. Los seis restos conservados de cisteína de cada repetición EGF y los restos de consenso de unión a Ca<sup>++</sup> (Rees y col., 1988, EMBO J. 7:2053-2061) se marcan con un asterisco (\*). El cambio de leucina a prolina encontrado en el clon de PCR de *Xenopus* que fallaba en la agregación, se indica abajo.

Figura 9. Estructuras Artificiales Empleadas en este Estudio. Se muestran diagramas esquemáticos de las variantes de Delta definidas en la Tabla IV. El extremo amino-proximal, extracelular está en cada caso a la izquierda. S, péptido señal; "EGF", motivos similares a EGF; M, hélice que atraviesa la membrana; H, secuencia de transferencia de detención; líneas continuas, otras secuencias de Delta; líneas sombreadas, secuencias de neuroglía. Las puntas de flecha indican sitios de inserción de engarces traducibles. Sca, ScaI; Nae, NaeI; Bam, BamHI; Bgl, BglII; ERL, repetición similar a EGF; Bst, BstEII; Dde, DdeI; Stu, StuI; NG1-NG5, quimeras de Delta de neuroglía.

## ES 2 181 674 T5

Figura 9A. Dependencia de la Agregación de las Cantidades Aportadas de ADN. A, Agregación heterotípica observada utilizando poblaciones de células S2 transfectadas transitoriamente, respectivamente, con cantidades variadas de ADN de pMTD11 (2, 4, 10 o 20  $\mu\text{g}/\text{placa}$ ) que se incubaron posteriormente bajo condiciones de agregación con poblaciones de células S2 transfectadas transitoriamente con una cantidad constante de ADN de pMtnMg (20  $\mu\text{g}/\text{placa}$ ). Los datos presentados son la fracción media (%) de células Delta en agregados de cuatro o más células  $\pm$  error típico para cada cantidad de ADN aportada (N = 3 duplicados, excepto aportaciones de 2  $\mu\text{g}$  y 10  $\mu\text{g}$  para las cuales N = 2). Se contaron un mínimo de 100 células que expresan Delta para cada duplicado. B, Agregación homotípica observada utilizando poblaciones de células S2 transfectadas transitoriamente, respectivamente, con cantidades variadas de ADN de pMTD11 (2, 4, 10 o 20  $\mu\text{g}/\text{placa}$ ) que se incubaron a continuación bajo condiciones de agregación. Los datos presentados son la fracción media (%) de células Delta en agregados de cuatro o más células  $\pm$  error típico para cada cantidad de ADN aportada (N = 3 duplicados). Se contó un mínimo de 500 células que expresan Delta para cada duplicado.

Figura 10. Alineación de la Secuencia Amino-Terminal de Delta-Serrate. Los restos se numeran en base a la traducción conceptual de las secuencias que codifican *Delta* (DI secuencia superior (SEQ ID NO:3); que comienza en el aminoácido 24, termina en el aminoácido 226) y *Serrate* (Ser, secuencia inferior (SEQ ID NO:4); que comienza en el aminoácido 85, termina en el aminoácido 283). Las líneas verticales entre las dos secuencias indican restos que son idénticos dentro de las secuencias de Delta y Serrate, tal y como están alineadas. Los puntos representan intersticios en la alineación. Los rectángulos incluyen restos de cisteína dentro de las regiones alineadas. N1, dominio 1 amino-proximal; N2, dominio 2 amino-proximal; N3, dominio 3 amino-proximal. Las inserciones traducibles asociadas con las estructuras artificiales STU B [sustitución del aminoácido 132 (A) de Delta por GKIFP] y NAE B [inserción de RKIF entre el aminoácido 197 de Delta y el aminoácido 198], respectivamente, se describen encima de la secuencia de Delta de tipo silvestre.

Figura 11. Geometrías Potenciales de las Interacciones Delta-Notch. A, Registro potencial de las moléculas Delta (izquierda) y Notch (derecha) que interaccionan entre membranas plasmáticas opuestas. B, Registro potencial de moléculas Delta (izquierda) y Notch (derecha) que interaccionan dentro de las mismas membranas plasmáticas. ELR, repetición similar a EGF; rectángulos blancos, repeticiones similares a EGF; rectángulos punteados, repeticiones LNR; rectángulos negros, hélices que atraviesan la membrana. El dominio amino-terminal de Delta y los dominios intracelulares de Delta y Notch están representados por óvalos.

Figura 12. Geometrías Potenciales de Interacciones Delta-Delta. A y B, Registro potencial de moléculas Delta que interaccionan entre membranas plasmáticas opuestas. B, Registro potencial de moléculas Delta que interaccionan dentro de las mismas membranas plasmáticas. Rectángulos blancos, repeticiones similares a EGF; rectángulos negros, hélices que atraviesan la membrana. Los dominios amino-terminales extracelulares e intracelulares de Delta están representados por óvalos.

Figura 13. Secuencia Primaria de Nucleótidos del ADNc de D11 de Delta (SEQ ID NO:5) y secuencia de aminoácidos de *Delta* (SEQ ID NO:6). Se muestra la secuencia de ADN de la hebra 5'-3' del ADNc de D11, que contiene una cantidad de correcciones en comparación con la presentada por Kopczynski y col. (1988, Genes Dev. 2, 1723-1735).

Figura 14. Secuencia Primaria de Nucleótidos de ADNc de Neuroglía 1B7A-250 (SEQ ID NO:7). Esta es una secuencia de ADN de una parte de la hebra 5'-3' del ADNc de 1B7A-250 (A.J. Bieber, pers. com.; Hortsch y col., 1990, Neuron 4, 697-709). El nucleótido 2890 se corresponde con el primer nucleótido de un codón de isoleucina que codifica el aminoácido 952 de la proteína de forma larga de neuroglía, traducida conceptualmente.

Figura 15. Homologías de la Secuencia de Ácidos Nucleicos entre *Serrate* y *Delta*. Se muestra una parte de la secuencia de nucleótidos de *Serrate* de *Drosophila* (SEQ ID NO:8), con la secuencia codificada de la proteína Serrate (SEQ ID NO:9) escrita a continuación, (Fleming y col., 1990, Genes & Dev. 4, 2188-2201 en 2193-94). Las cuatro regiones que muestran una alta homología de secuencia con la secuencia de Delta de *Drosophila*, se numeran por encima de la línea y se indican entre corchetes. La región total de homología abarca los nucleótidos números 627 hasta 1290 de la secuencia de nucleótidos de *Serrate* (numeración como en la Figura 4 de Fleming y col., 1990, Genes & Dev. 4, 2188-2201).

Figura 16. Cebadores empleados para PCR en la Clonación de *Notch* Humana. Se muestra la secuencia de los tres cebadores empleados en la PCR para multiplicar el ADN en una genoteca de ADNc de cerebro fetal humano. Los tres cebadores, cdc1 (SEQ ID NO:10), cdc2 (SEQ ID NO:11) y cdc3 (SEQ ID NO:12), se diseñaron para multiplicar tanto un fragmento de 200 pb como un fragmento de 400 pb en forma de parejas de cebadores cdc1/cdc2 o cdc1/cdc3, respectivamente. I: inosina.

Figura 17. Diagrama Esquemático de los Clones de *Notch* Humana. Se muestra un diagrama esquemático de *Notch* humana. Líneas muy negras por debajo del diagrama muestran la parte de la secuencia de *Notch* contenida en cada uno de los cuatro clones de ADNc. La localización de los cebadores empleados en la PCR, y su orientación, se indica con flechas.

Figura 18. Secuencias de *Notch* Humana Alineadas con la Secuencia de *Notch* de *Drosophila*. Las líneas verticales numeradas se corresponden con las coordenadas de *Notch* de *Drosophila*. Las líneas horizontales por debajo de cada mapa muestran donde se encuentran los clones en relación con extensiones de secuencia (líneas gruesas horizontales).

Figura 19. Secuencias de Nucleótidos de *Notch* Humana Contenidas en el Clon de ADNc Plasmídico hN2k. Figura 19A: La secuencia de ADN (SEQ ID NO:13) de una parte de la inserción de *Notch* humana se muestra, empezando por el sitio EcoRI en el extremo 3' y procediendo en la dirección 3' a 5'. Figura 19B: La secuencia de ADN (SEQ ID NO:14) de una parte de la inserción de *Notch* humana se muestra, comenzando por el sitio EcoRI en el extremo 5' y procediendo en la dirección 5' a 3'. Figura 19C: La secuencia de ADN (SEQ ID NO:15) de una parte de la inserción de *Notch* humana se muestra, comenzando en 3' de la secuencia mostrada en la Figura 19B y procediendo en la dirección 5' a 3'. Las secuencias mostradas son tentativas, están sujetas a confirmación por determinación de las secuencias de solapantes.

Figura 20. Secuencias de Nucleótidos de *Notch* Humana Contenidas en el Clon de ADNc de Plásmido hN3k. Figura 20A: La secuencia de ADN (SEQ ID NO:16) de una parte de la inserción de *Notch* humana se muestra, comenzando por el sitio EcoRI en el extremo 3' y procediendo en la dirección 3' a 5'. Figura 20B: La secuencia de ADN (SEQ ID NO:17) de una parte de la inserción de *Notch* humana se muestra, comenzando por el sitio EcoRI en el extremo 5' y procediendo en la dirección 5' a 3'. Figura 20C: La secuencia de ADN (SEQ ID NO:18) de una parte de la inserción de *Notch* humana se muestra, comenzando en 3' de la secuencia mostrada en la Figura 20B, y procediendo en la dirección 5' a 3'. Figura 20D: La secuencia de ADN (SEQ ID NO:19) de una parte de la inserción de *Notch* humana se muestra, comenzando 5' de la secuencia mostrada en la Figura 20A y procediendo en la dirección 3' a 5'. Las secuencias mostradas son tentativas, están sujetas a confirmación por determinación de las secuencias solapantes.

Figura 21. Secuencias de Nucleótidos de *Notch* Humana Contenidas en el Clon de ADNc Plasmídico hN4k. Figura 21A: La secuencia de ADN (SEQ ID NO:20) de una parte de la inserción de *Notch* humana se muestra, comenzando en el sitio EcoRI en el extremo 5' y procediendo en la dirección 5' a 3'. Figura 21B: La secuencia de ADN (SEQ ID NO:21) de una parte de la inserción de *Notch* humana se muestra, comenzando cerca del extremo 3', y procediendo en la dirección 3' a 5'. Las secuencias mostradas son tentativas y están sujetas a confirmación por determinación de las secuencias solapantes.

Figura 22. Secuencias de Nucleótidos de *Notch* Humana Contenidas en el Clon de ADNc de Plásmido hN5k. Figura 22A: La secuencia de ADN (SEQ ID NO:22) de una parte de la inserción de *Notch* humana se muestra, partiendo del sitio EcoRI en el extremo 5' y procediendo en la dirección 5' a 3'. Figura 22B: La secuencia de ADN (SEQ ID NO:23) de una parte de la inserción de *Notch* humana se muestra, partiendo cerca del extremo 3' y procediendo en la dirección 3' a 5'. Figura 22C: La secuencia de ADN (SEQ ID NO:24) de una parte de la inserción de *Notch* humana se muestra, partiendo en 3' de la secuencia mostrada en la Figura 22A, y procediendo en la dirección 5' a 3'. Figura 22D: La secuencia de ADN (SEQ ID NO:25) de una parte de la inserción de *Notch* humana se muestra, comenzando 5' de la secuencia mostrada en la Figura 22B y procediendo en la dirección 3' a 5'. Las secuencias mostradas son tentativas, están sujetas a confirmación por determinación de las secuencias solapantes.

Figura 23. Secuencias de ADN (SEQ ID NO:31) y Aminoácidos (SEQ ID NO:34) de *Notch* Humana Contenidas en el Clon de ADNc de Plásmido hN3k.

Figura 24. Secuencias de ADN (SEQ ID NO:33) y Aminoácidos (SEQ ID NO:34) de *Notch* Humana Contenidas en el Clon de ADNc de Plásmido hN5k.

Figura 25. Comparación de hN5k con otros Homólogos de Notch. Figura 25A. Representación esquemática de Notch de *Drosophila*. Se indican la secuencia señal (señal), las 36 repeticiones similares a EGF, las tres repeticiones *Notch/lin-12*, el dominio transmembranal (TM), las seis repeticiones *cdc10*, la repetición OPA y la región rica en PEST (prolina, ácido glutámico, serina, treonina). Figura 25B. Alineación de la secuencia de aminoácidos deducidos de hN5k con secuencias de otros homólogos de Notch. Los aminoácidos se numeran en el lado izquierdo. Las regiones *cdc10* y las regiones ricas en PEST están ambas enmarcadas por rectángulos y las repeticiones individuales *cdc10* están marcadas. Los aminoácidos que son iguales en tres o más secuencias, están iluminados. Los cebadores utilizados para clonar hN5k se indican debajo de las secuencias a partir de las cuales fueron diseñados. La secuencia de localización nuclear (NLS), los sitios de la quinasa de caseína II (CKII) y la quinasa *cdc2* (*cdc2*) del motivo CcN putativo de los homólogos de Notch de vertebrados están enmarcados por rectángulos. La posible secuencia bipartita de dirección al núcleo (BNTS) y los sitios de fosforilación proximales de Notch de *Drosophila*, también están enmarcados por rectángulos.

## 5. Descripción detallada de la invención

La presente invención proporciona secuencias de nucleótidos de los genes *Notch* humanos, y secuencias de aminoácidos de sus proteínas codificadas. La memoria descriptiva también describe secuencias de nucleótidos de genes *Delta* y secuencias de aminoácidos de sus proteínas codificadas. La memoria descriptiva describe adicionalmente fragmentos (denominados en esta memoria "fragmentos adhesivos") de las proteínas codificadas por genes toporrítmicos que median en la unión homotípica o heterotípica a proteínas toporrítmicas o a sus fragmentos adhesivos. Los genes toporrítmicos, tal y como se emplean en esta memoria, deben significar los genes *Notch*, *Delta* y *Serrate*, así como otros miembros de la familia *Delta/Serrate* que se pueden identificar, por ejemplo, por los métodos descritos en la Sección 5.3, más abajo.

Las secuencias de ácidos nucleicos y de aminoácidos y los anticuerpos para las mismas de la invención, se pueden utilizar para detectar y cuantificar el ARNm de moléculas de Notch humana, para estudiar su expresión, para producir secuencias de Notch humana, en el estudio y la manipulación de procesos de diferenciación.

## ES 2 181 674 T5

Para mayor claridad en la descripción, y no a modo de limitación, la descripción detallada siguiente se dividirá en las subsecciones siguientes:

- (i) La identificación de dominios de proteínas toporrítmicas y las secuencias de dominios de proteínas toporrítmicas que median en la unión a dominios de proteínas toporrítmicas;
- (ii) La clonación y la secuenciación de *Notch* y *Delta* humanas;
- (iii) La identificación de miembros adicionales de la familia *Delta/Serrate*;
- (iv) La expresión de genes toporrítmicos;
- (v) La identificación y la purificación del producto génico expresado; y
- (vi) La generación de anticuerpos para proteínas toporrítmicas y sus secuencias adhesivas.

### 5.1 Identificación de los dominios de proteínas toporrítmicas y las secuencias de dominios de proteínas toporrítmicas que median en la unión a dominios de proteínas toporrítmicas

Los fragmentos de proteínas toporrítmicas, que median en la unión homotípica o heterotípica (y que por tanto se denominan en esta memoria “adhesivos”), y las secuencias de ácidos nucleicos relacionadas con los anteriores, se describen a continuación.

En un ejemplo específico, el fragmento adhesivo de Notch que forma parte de la presente invención, es el que comprende la parte de Notch más homóloga con ELR 11 y 12, es decir, los aminoácidos número 447 hasta 527 (SEQ ID NO:1) de la secuencia de Notch de *Drosophila* (véase Figura 8). En otro ejemplo específico, el fragmento adhesivo de Delta que media en la unión homotípica es el que comprende la parte de Delta más homóloga con aproximadamente los aminoácidos número 32-230 de la secuencia de Delta de *Drosophila* (SEQ ID NO:6). En aún otro ejemplo específico, el fragmento adhesivo de Delta que media en la unión a Notch, es el que comprende la parte de Delta más homóloga con aproximadamente los aminoácidos número 1-230 de la secuencia de Delta de *Drosophila* (SEQ ID NO:6). En un ejemplo específico relacionado con un fragmento adhesivo de Serrate, dicho fragmento es el que comprende la parte de Serrate más homóloga con aproximadamente los aminoácidos número 85-283 o 79-282 de la secuencia de Serrate de *Drosophila* (véanse, Figura 10 (SEQ ID NO:4) y Figura 15 (SEQ ID NO:9)).

Las secuencias de ácidos nucleicos que codifican dominios adhesivos toporrítmicos se pueden aislar a partir de animales porcinos, bovinos, felinos, aviares, equinos o caninos, así como a partir de fuentes de primates y cualquier otra especie en la que se pueden identificar homólogos de genes toporrítmicos conocidos [incluyendo, pero no limitado a los siguientes genes (con la publicación de las secuencias entre paréntesis): *Notch* (Wharton y col., 1985, Cell 43, 567-581), *Delta* (Vassin y col., 1987, EMBO J. 6, 3431-3440; Kopczynski y col., 1988, Genes Dev. 2, 1723-1735; nótese las correcciones de la secuencia de Kopczynski y col. encontradas en la Figura 13 de esta memoria (SEQ ID NO:5 y SEQ ID NO:6)) y *Serrate* (Fleming y col., 1990, Genes & Dev. 4, 2188-2201)]. Tales secuencias se pueden alterar mediante sustituciones, adiciones o deleciones que proporcionan moléculas funcionalmente equivalentes (adhesivas). Debido a la degeneración de las secuencias que codifican nucleótidos, se pueden utilizar en la práctica de la presente invención, otras secuencias de ADN que codifican sustancialmente la misma secuencia de aminoácidos que las secuencias adhesivas. Estas incluyen, pero no están limitadas, a secuencias de nucleótidos que comprenden todas las partes de los genes *Notch* que están alteradas por la sustitución de diferentes codones que codifican un resto de aminoácidos funcionalmente equivalente dentro de la secuencia, produciendo de este modo un cambio silencioso. De forma similar, los fragmentos de proteína adhesiva o sus derivados incluyen, pero no están limitados, a los que contienen, como secuencia de aminoácidos primaria, toda o parte de la secuencia de aminoácidos de los dominios adhesivos que incluyen secuencias alteradas en las que restos de aminoácidos funcionalmente equivalentes están sustituidos por restos dentro de la secuencia, dando como resultado un cambio silencioso. Por ejemplo, uno o varios restos de aminoácidos dentro de la secuencia pueden estar sustituidos por otro aminoácido con una polaridad similar que actúa como un equivalente funcional, dando como resultado una alteración silenciosa. Los sustitutos de un aminoácido dentro de la secuencia se pueden seleccionar a partir de otros miembros de la clase a la que pertenece el aminoácido. Por ejemplo, los aminoácidos no polares (hidrófobos) incluyen alanina, leucina, isoleucina, valina, prolina, fenilalanina, triptófano y metionina. Los aminoácidos polares neutros incluyen glicina, serina, treonina, cisteína, tirosina, asparragina y glutamina. Los aminoácidos cargados positivamente (básicos) incluyen arginina, lisina e histidina. Los aminoácidos cargados negativamente (ácidos) incluyen ácido aspártico y ácido glutámico.

Los fragmentos adhesivos de proteínas toporrítmicas y péptidos relacionados con secuencias de proteínas toporrítmicas adhesivas, se pueden someter a ensayo para estudiar la actividad de unión deseada, por ejemplo, mediante ensayos de agregación *in vitro*, descritos en los ejemplos de esta memoria. Los fragmentos adhesivos de proteínas toporrítmicas incluyen, pero no están limitados a los péptidos que son sustancialmente homólogos a los fragmentos adhesivos, o cuyo ácido nucleico codificado es capaz de hibridarse con la secuencia de ácidos nucleicos que codifica los fragmentos adhesivos, y cuyos péptidos tienen una actividad de unión positiva, por ejemplo, cuando se someten a ensayo *in vitro* mediante un ensayo de agregación tal como el que se describe en las secciones de ejemplos, más abajo.

## ES 2 181 674 T5

Los péptidos relacionados con proteínas adhesivas, se pueden producir por distintos métodos conocidos en la técnica. Las manipulaciones que tienen lugar en su producción pueden ocurrir a nivel génico o proteico. Por ejemplo, la secuencia de genes que codifica la proteína adhesiva clonada se puede modificar mediante cualquiera de las numerosas estrategias conocidas en la técnica (Maniatis, T., 1990, *Molecular Cloning, A Laboratory Manual*, 2ª edición, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York). La secuencia se puede escindir en sitios apropiados con endonucleasa(s) de restricción, seguido de una modificación enzimática adicional, si se desea, aislamiento y ligación *in vitro*. En la producción del gen que codifica un péptido relacionado con un dominio adhesivo, se tiene que tener cuidado para asegurar que el gen modificado permanece dentro del mismo marco de lectura traduccional que la proteína adhesiva, sin interrupciones con señales de detención de la traducción, en la región génica en donde se codifica la actividad adhesiva deseada.

Adicionalmente, la secuencia de ácidos nucleicos que codifica el adhesivo, se puede mutar *in vitro* o *in vivo*, para crear y/o destruir secuencias de traducción, iniciación y/o terminación, o para crear variaciones en regiones codificadoras y/o formar nuevos sitios para endonucleasas de restricción o destruir sitios preexistentes, para facilitar una modificación adicional *in vitro*. Cualquier técnica de mutagénesis conocida en la técnica, se puede utilizar, incluyendo pero no limitada a mutagénesis dirigida al sitio *in vitro* (Hutchinson, C., y col., 1978, *J. Biol. Chem.* 253, 6551), uso de engarzadores TAB® (Pharmacia), etc.

También se pueden realizar manipulaciones de la secuencia adhesiva a nivel de proteína. Los fragmentos de proteínas toporrítmicas, se pueden modificar diferencialmente durante o después de la traducción, por ejemplo, mediante glicosilación, acetilación, fosforilación, escisión proteolítica, enlace a una molécula de anticuerpo u otro ligando celular, etc. Cualquiera de las numerosas modificaciones químicas se puede realizar mediante técnicas conocidas, que incluyen pero no están limitadas a escisión química específica mediante bromuro de cianógeno, tripsina, quimotripsina, papaína, proteasa V8, NaBH<sub>4</sub>; acetilación, formilación, oxidación, reducción; síntesis metabólica en presencia de tunicamicina; etc.

Además, los péptidos relacionados con fragmentos adhesivos se pueden sintetizar químicamente. Por ejemplo, un péptido correspondiente a una parte de una proteína toporrítmica, que media en la actividad de agregación deseada *in vitro*, se puede sintetizar empleando un sintetizador de péptidos.

Otro ejemplo específico descrito en esta memoria se refiere a fragmentos de una proteína Delta que tiene la capacidad de unirse a una segunda proteína Delta o a un fragmento o un derivado de la misma, pero que no se une a Notch. Una unión tal, o la falta de unión, se puede someter a ensayo *in vitro*, tal y como se describe en la Sección 8. A modo de ejemplo, pero no como limitación, un derivado tal de Delta es el que contiene una inserción del tetrapéptido Arg-Lys-Ile-Phe entre los restos 198 y 199 de Delta de la proteína de *Drosophila*.

### 5.2 La clonación y la secuenciación de Notch y Delta humanas

La invención se refiere adicionalmente a las secuencias de aminoácidos de Notch humana y a fragmentos de las mismas que comprenden un determinante antigénico (es decir, se pueden reconocer con un anticuerpo) o que son funcionalmente activas, así como a secuencias de ácidos nucleicos que codifican las anteriores. También se describe en esta memoria las secuencias de aminoácidos y de ácidos nucleicos de Delta y sus fragmentos y derivados correspondientes. Material "funcionalmente activo", tal y como se emplea en esta memoria, se refiere a aquel material que manifiesta una o más de las actividades funcionales conocidas, asociadas con el producto de la proteína de longitud completa (tipo silvestre), por ejemplo, en el caso de Notch, unión con Delta, unión con Serrate, antigenicidad (unión a un anticuerpo anti-Notch), etc.

En realizaciones específicas, la invención proporciona fragmentos de una proteína Notch humana que consta de al menos 40 aminoácidos, o de al menos 77 aminoácidos. En otras realizaciones, las proteínas de la invención comprenden o constan esencialmente del dominio intracelular, región transmembranal, dominio extracelular, región cdc10, repeticiones *Notch/lin-12* o las repeticiones homólogas a EGF, o cualquier combinación de los anteriores de una proteína Notch humana. También se proporcionan fragmentos, o proteínas que comprenden fragmentos, que carecen de alguna o de todas las repeticiones homólogas a EGF de Notch humana.

En otras realizaciones específicas, la invención se dirige adicionalmente a las secuencias y subsecuencias de nucleótidos de *Notch* humano que constan de al menos 25 nucleótidos, al menos 50 nucleótidos o al menos 121 nucleótidos. Los ácidos nucleicos que codifican las proteínas y los fragmentos proteicos descritos anteriormente, también se proporcionan, así como ácidos nucleicos complementarios y capaces de hibridarse con tales ácidos nucleicos. En una realización, una secuencia complementaria tal puede ser complementaria a una secuencia de ADNc de Notch humana de al menos 25 nucleótidos, o de al menos 121 nucleótidos. En un aspecto preferido, la invención se refiere a secuencias de ADNc que codifican Notch humana o una parte de la misma. En una realización específica, la invención se refiere a la secuencia de nucleótidos del gen *Notch* humano o ADNc, en particular, que comprende las secuencias descritas en las Figuras 19, 20, 21 y/o 22 (SEQ ID NO:13 hasta NO:25), o contenida en los plásmidos hN3k, hN4k o hN5k (véase Sección 9, más abajo), y las secuencias codificadas de la proteína de *Notch*. Tal y como es evidente, del modo en que se utiliza en esta memoria, un "ácido nucleico que codifica un fragmento o una parte de una proteína Notch", se debe interpretar como el que se refiere a un ácido nucleico que codifica sólo el fragmento mencionado o una parte de la proteína Notch y no otras partes de la proteína Notch.

## ES 2 181 674 T5

En un aspecto preferido de la invención, pero no limitante, una secuencia de ADN de *Notch* humano se puede clonar y secuenciar por el método descrito en la Sección 9, más abajo.

5 Un método preferido para clonar *Delta* humano, que se presenta como un ejemplo particular pero no a modo de limitación, es como sigue:

Se construye una genoteca humana de expresión por métodos conocidos en la técnica. Por ejemplo, se aísla ARNm humano, se prepara ADNc y se liga en un vector de expresión tal (por ejemplo, un derivado de bacteriófago) que es capaz de ser expresado por la célula hospedadora en la que se introduce éste a continuación. Después, se pueden emplear distintos ensayos de rastreo para seleccionar el producto expresado de *Delta* humana. En una realización, la selección se puede llevar a cabo en base a la unión positiva con el dominio adhesivo de *Notch* humana, (es decir, la parte de *Notch* humana más homóloga con ELR 11 y 12 de *Drosophila* (SEQ ID NO:1). En una realización alternativa, se pueden utilizar anticuerpos anti-*Delta* para la selección.

15 En otro aspecto preferido, se utiliza PCR para multiplicar la secuencia deseada en la genoteca, antes de la selección. Por ejemplo, cebadores de oligonucleótidos que representan parte de los dominios adhesivos codificados por un homólogo del gen deseado, se pueden utilizar como cebadores en la PCR.

20 Los métodos mencionados anteriormente no pretenden limitar la siguiente descripción general de los métodos mediante los cuales se pueden obtener clones de *Notch* y *Delta* humanos.

Cualquier célula humana puede servir potencialmente como fuente de ácido nucleico para la clonación molecular de genes *Notch* y *Delta*. El ADN se puede obtener por procedimientos convencionales conocidos en la técnica a partir de ADN clonado (por ejemplo, una "genoteca" de ADN), mediante síntesis química, mediante clonación de ADNc o mediante la clonación de ADN genómico o de fragmentos del mismo, purificados de la célula humana deseada. (Véase, por ejemplo, Maniatis y col., 1982, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York; Glover, D.M. (compilador), 1985, *DNA Cloning: A Practical Approach*, MRL Press, Ltd., Oxford, G.B. vol. I, II). Los clones obtenidos a partir de ADN genómico, pueden contener regiones de ADN reguladoras e intrones además de las regiones codificadoras; los clones obtenidos a partir de ADNc contendrán sólo secuencias de exones. Cualquiera que sea la fuente, el gen debe ser clonado molecularmente en un vector adecuado para la multiplicación del gen.

35 En la clonación molecular del gen procedente de ADN genómico, se generan fragmentos de ADN, alguno de los cuales codificará el gen deseado. El ADN se puede escindir en sitios específicos empleando distintas enzimas de restricción. Alternativamente, la ADNasa se puede usar en presencia de manganeso para fragmentar el ADN, o el ADN se puede cortar físicamente, como por ejemplo, mediante exposición a ultrasonidos. Los fragmentos de ADN lineales se pueden separar entonces, según el tamaño, mediante técnicas normalizadas, que incluyen, pero no están limitadas a electroforesis en gel de agarosa y poliacrilamida y cromatografía en columna.

40 Una vez generados los fragmentos de ADN, se puede realizar de diversos modos la identificación del fragmento de ADN específico que contiene el gen deseado. Por ejemplo, si una cantidad de una parte de un gen *Notch* o *Delta* (de cualquier especie) o su ARN específico, o un fragmento del mismo, por ejemplo, el dominio adhesivo, está disponible y se puede purificar y marcar, los fragmentos de ADN generados se pueden escrutar mediante hibridación de los ácidos nucleicos con la sonda marcada (Benton, W. y Davis, R., 1977, *Science* 196, 180; Grunstein, M. y Hogness, D., 1975, *Proc. Natl. Acad. Sci. U.S.A.* 72, 3961). Aquellos fragmentos de ADN con una homología sustancial con la sonda se hibridarán. También es posible identificar el fragmento apropiado mediante digestión(es) con enzimas de restricción y comparar los tamaños de los fragmentos con los esperados, según un mapa de restricción conocido, si está disponible. Una selección adicional se puede realizar en base a las propiedades del gen. Alternativamente, la presencia del gen se puede detectar por ensayos basados en las propiedades físicas, químicas o inmunológicas de su producto expresado. Por ejemplo, clones de ADNc o clones de ADN que se hibridan o se seleccionan los ARNm adecuados, se pueden seleccionar los que producen una proteína que, por ejemplo, tiene una migración electroforética, comportamiento de enfoque isoeléctrico, mapas de digestión proteolítica, actividad de agregación *in vitro* ("adhesividad") o propiedades antigénicas, similares o idénticas, como las conocida para *Notch* o *Delta*. Si un anticuerpo para *Notch* o *Delta* está disponible, la proteína *Notch* o *Delta* se puede identificar uniendo el anticuerpo marcado a los clones putativos que sintetizan *Notch* o *Delta*, en un procedimiento de tipo ELISA (ensayo inmunoenzimático).

55 Los genes *Notch* o *Delta* también se pueden identificar por selección del ARNm mediante hibridación del ácido nucleico, seguido de traducción *in vitro*. En este procedimiento, los fragmentos se emplean para aislar ARNm complementarios por hibridación. Tales fragmentos de ADN pueden representar ADN purificado y disponible de *Notch* o de *Delta* de otra especie (por ejemplo, *Drosophila*). El análisis por inmunoprecipitación o los ensayos funcionales (por ejemplo, capacidad de agregación *in vitro*; véanse los ejemplos más abajo) de los productos de la traducción *in vitro* de los productos aislados de los ARNm aislados, identifica el ARNm y, por tanto, los fragmentos de ADN complementario que contienen las secuencias deseadas. Además, los ARNm específicos se pueden seleccionar por adsorción de polisomas aislados de células, para inmovilizar anticuerpos dirigidos específicamente contra la proteína *Notch* o *Delta*. Un ADNc radiomarcado de *Notch* o *Delta* se puede sintetizar empleando el ARNm seleccionado (procedente de los polisomas adsorbidos) como molde. El ARNm o ADNc marcado radiactivamente se puede utilizar a continuación como una sonda para identificar los fragmentos de ADN de *Notch* o *Delta* entre otros fragmentos de ADN genómico.

Las alternativas para aislar el ADN genómico de *Notch* o *Delta* incluyen, pero no están limitadas a la síntesis química de la secuencia génica misma, a partir de una secuencia conocida o preparando un ADNc del ARNm que codifica el gen *Notch* o *Delta*. Por ejemplo, ARN para la clonación de ADNc del gen *Notch* o *Delta*, se puede aislar a partir de células que expresan *Notch* o *Delta*. Otros métodos son posibles.

El gen identificado y aislado se puede insertar a continuación en un vector de clonación apropiado. Se pueden utilizar una gran cantidad de sistemas vector-hospedador conocidos en la técnica. Posibles vectores incluyen, pero no están limitados a plásmidos o virus modificados, pero el sistema del vector tiene que ser compatible con la célula hospedadora utilizada. Tales vectores incluyen, pero no están limitados a bacteriófagos tales como derivados de lambda, o plásmidos tales como pBR322 o derivados del plásmido pUC. La inserción en un vector de clonación se puede realizar, por ejemplo, ligando el fragmento de ADN en un vector de clonación que tiene extremos cohesivos complementarios. Sin embargo, si los sitios de restricción complementarios utilizados para fragmentar el ADN no están presentes en el vector de clonación, los extremos de las moléculas de ADN se pueden modificar enzimáticamente. Alternativamente, cualquier sitio deseado se puede producir ligando secuencias de nucleótidos (engarzadores) a los extremos del ADN; estos engarzadores ligados pueden comprender oligonucleótidos específicos sintetizados químicamente, que codifican secuencias de reconocimiento de las endonucleasas de restricción. En un método alternativo, el vector cortado y los genes *Notch* o *Delta* se pueden modificar adicionando una cola homopolímera. Las moléculas recombinantes se pueden introducir en las células hospedadoras mediante transformación, transfección, infección, electroporación, etc., de modo que se generan muchas copias de la secuencia génica.

En un método alternativo, el gen deseado se puede identificar y aislar después de la inserción en un vector de clonación adecuado, en un acercamiento de tipo “perdigonada” (del inglés, “shot gun”). El enriquecimiento del gen deseado, por ejemplo, mediante fraccionamiento por tamaño, se puede realizar antes de insertar en el vector de clonación.

En ejemplos específicos, la transformación de las células hospedadoras con moléculas de ADN recombinante que incorporan el gen *Notch* o *Delta* aislado, ADNc o una secuencia de ADN sintetizada, permite la generación de copias múltiples del gen. Por tanto, se puede obtener el gen en grandes cantidades haciendo crecer los transformantes, aislando las moléculas de ADN recombinante de los transformantes y, si es necesario, recuperar el gen insertado a partir del ADN recombinante aislado.

Las secuencias de *Notch* humano proporcionadas por la presente invención, incluyen aquellas secuencias de nucleótidos que codifican sustancialmente las mismas secuencias de aminoácidos que las encontradas en *Notch* humana, y aquellas secuencias de aminoácidos codificadas por aminoácidos funcionalmente equivalentes, todas ellas descritas más arriba, en la Sección 5.1 para partes adhesivas de proteínas toporrítmicas.

### 5.3 Identificación de miembros adicionales de la familia *Delta/Serrate*

Una búsqueda racional de miembros adicionales de la familia de genes *Delta/Serrate* se puede realizar empleando un método que toma ventaja de la existencia de segmentos conservados de fuerte homología entre *Serrate* y *Delta* (véase Figura 10, SEQ ID NO:3 y NO:4). Por ejemplo, miembros adicionales de esta familia génica se pueden identificar seleccionando, entre una diversidad de secuencias de ácidos nucleicos, las secuencias que son homólogas tanto con *Serrate* como con *Delta* (véase Figura 13 (SEQ ID NO:5) y Figura 15 (SEQ ID NO:8)), e identificando posteriormente, entre las secuencias seleccionadas, las que también contienen secuencias de ácidos nucleicos que no son homólogas con *Serrate* y con *Delta*. La expresión “no homóloga” se puede utilizar para indicar una región que contiene al menos aproximadamente 6 nucleótidos contiguos en los que al menos dos nucleótidos difieren de la secuencia de *Serrate* y *Delta*. Por ejemplo, un método preferido es como sigue: se pueden sintetizar grupos de sondas de oligonucleótidos degenerados de aproximadamente 10-20 nucleótidos, que se corresponden con dos segmentos conservados entre *Delta* y *Serrate*, los AA 63-73 de *Delta* y los AA 195-206 de *Delta* (véase Figura 13, SEQ ID NO:6), que representan todas las secuencias codificadoras posibles para los aminoácidos encontrados en *Delta* y *Serrate* para aproximadamente tres a siete codones contiguos. En otra realización, se pueden obtener oligonucleótidos que se corresponden con partes de las cuatro regiones altamente conservadas entre *Delta* y *Serrate* mostradas en la Figura 15 (SEQ ID NO:8 y NO:9), es decir, las representadas por los AA 124-134 de *Serrate*, 149-158, 214-219 y 250-259 de *Serrate*. Los oligonucleótidos sintéticos se pueden utilizar como cebadores para multiplicar secuencias con PCR a partir de una fuente (ARN o ADN) de potencial interés. (PCR se puede realizar, por ejemplo, empleando un variador térmico de Perkin-Elmer Cetus y polimerasa Taq (Gene Amp®)). Esta puede incluir ARNm o ADNc o ADN genómico procedente de especies eucariotas que podrían expresar un polipéptido estrechamente relacionado con *Serrate* y *Delta*. Al llevar a cabo las reacciones de la PCR, es posible detectar un gen o un producto génico que comparte los segmentos mencionados anteriormente de la secuencia conservada entre *Serrate* y *Delta*. Si se escoge sintetizar varios cebadores degenerados distintos, todavía puede ser posible realizar una búsqueda completa con un número razonablemente bajo de reacciones de PCR. También es posible variar la severidad de las condiciones de hibridación empleadas en el cebado de las reacciones de PCR, para permitir grados mayores o menores de similitud de la secuencia de nucleótidos entre el gen desconocido y *Serrate* o *Delta*. Si un segmento de un miembro previamente desconocido de la familia de genes *Serrate/Delta* se multiplica con éxito, este segmento se puede clonar molecularmente y secuenciar, y utilizar como sonda para aislar un ADNc completo o un clon genómico. Esto, a su vez, permitirá la determinación de la secuencia de nucleótidos completa del gen desconocido, el análisis de su expresión y la producción de su producto proteico para análisis funcionales. De este modo, se pueden identificar genes adicionales que codifican proteínas “adhesivas”.

Las homologías de la secuencia *Serrate/Delta* se pueden utilizar para el diseño de nuevas moléculas recombinantes que son miembros de la familia de genes *Serrate/Delta* pero que no están presentes en la naturaleza. Por ejemplo, y no a modo de limitación, una molécula recombinante se puede construir, comprendiendo partes de ambos genes *Serrate* y *Delta*. Una molécula tal podría mostrar propiedades asociadas con *Serrate* y *Delta* y retratar un nuevo perfil de actividades biológicas, incluyendo agonistas así como antagonistas. La secuencia primaria de *Serrate* y *Delta* también se puede utilizar para pronosticar la estructura terciaria de las moléculas utilizando simulación por ordenador (Hopp y Woods, 1981, Proc. Natl. Acad. Sci. U.S.A. 78, 3824-3828); genes recombinantes quiméricos de *Serrate/Delta* se podrían diseñar a la luz de correlaciones entre la estructura terciaria y la función biológica. De forma similar, se pueden construir genes quiméricos que comprenden partes de uno cualquiera o de varios miembros de la familia de genes toporrítmicos (por ejemplo, *Notch*).

#### 5.4 La expresión de genes toporrítmicos

La secuencia de nucleótidos que codifica un fragmento adhesivo de una proteína toporrítmica (preferentemente, *Notch*, *Serrate* o *Delta*), o *Notch* o *Delta* humana o un fragmento funcionalmente activo del mismo, se puede insertar en un vector de expresión apropiado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia que codifica la proteína insertada. Las señales necesarias para la transcripción y la traducción también pueden estar provistas por el gen toporrítmico natural y/o sus regiones flanqueantes. Se puede utilizar una variedad de sistemas hospedador-vector para expresar la secuencia que codifica la proteína. Estos incluyen, pero no están limitados a sistemas celulares de mamíferos infectados con virus (por ejemplo, virus vaccinia, adenovirus, etc.); sistemas celulares de insectos infectados con virus (por ejemplo, baculovirus); microorganismos tales como levadura que contiene vectores de levadura o bacterias transformadas con bacteriófago, ADN, ADN plasmídico o ADN de cósmido. Los elementos de expresión de vectores varían en sus intensidades y sus especificidades. Dependiendo del sistema de hospedador-vector utilizado, se puede emplear una cualquiera de una cantidad de elementos adecuados para la transcripción y la traducción. En un ejemplo específico, se expresa la parte adhesiva del gen *Notch*, por ejemplo, la que codifica las repeticiones similares a EGF 11 y 12. En otro ejemplo, se expresa la parte adhesiva del gen *Delta*, por ejemplo, la que codifica los aminoácidos 1-230. En otros ejemplos específicos, se expresa el gen *Notch* humano o *Delta* humano, o una secuencia que codifica una parte funcionalmente activa de *Notch* o *Delta* humana. En aún otro ejemplo, se expresa la parte adhesiva del gen *Serrate*.

Cualquiera de los métodos descritos previamente para la inserción de fragmentos de ADN en un vector, se puede utilizar para construir vectores de expresión que contienen un gen quimérico que consta de señales de control de la transcripción/traducción apropiadas y las secuencias que codifican la proteína. Estos métodos pueden incluir ADN recombinante *in vitro* y técnicas sintéticas y recombinantes *in vivo* (recombinación genética). La expresión de la secuencia de ácidos nucleicos que codifica una proteína toporrítmica o un fragmento peptídico, se puede regular con una segunda secuencia de ácidos nucleicos, de modo que la proteína o el péptido toporrítmico se expresa en un hospedador transformado con la molécula de ADN recombinante. Por ejemplo, la expresión de una proteína toporrítmica se puede controlar mediante cualquier elemento promotor/intensificador conocido en la técnica. Los promotores que se pueden utilizar para controlar la expresión de genes toporrítmicos incluyen, pero no están limitados a la región del promotor temprano de SV40 (Bernoist y Chambon, 1981, Nature 290, 304-310), el promotor contenido en la repetición 3' terminal larga del virus de sarcoma de Rous (Yamamoto y col., 1980, Cell 22, 787-797), el promotor de la quinasa de timidina de herpes (Wagner y col., 1981, Proc. Natl. Acad. Sci. U.S.A. 78, 1441-1445), las secuencias reguladoras del gen de la metalotioneína (Brinster y col., 1982, Nature 296, 39-42); los vectores de expresión procarióticos tales como el promotor de la  $\beta$ -lactamasa (Villa-Kamaroff y col., 1978, Proc. Natl. Acad. Sci. U.S.A. 75, 3727-3731), o el promotor *tac* (DeBoer y col., 1983, Proc. Natl. Acad. Sci. U.S.A. 80, 21-25); véase también "Useful proteins from recombinant bacteria" en Investigación y Ciencia, 1980, 242, 74-94; vectores de expresión vegetales que comprenden la región del promotor de la sintetasa de nopalina (Herrera-Estrella y col., Nature 303, 209-213) o el promotor del ARN 35S del virus del mosaico de la coliflor (Gardner y col., 1981, Nucl. Acids Res. 9, 2871) y el promotor de la enzima fotosintética ribulosa-bifosfato-carboxilasa (Herrera-Estrella y col., 1984, Nature 310, 115-120); elementos del promotor procedentes de levadura u otros hongos tales como el promotor Gal 4, el promotor ADC (deshidrogenasa de alcohol), el promotor PGK (quinasa de fosfoglicerol), el promotor de la fosfatasa alcalina y las siguientes regiones de control de la transcripción animal, que muestran especificidad de tejido y se han utilizado en animales transgénicos: región de control del gen de la elastasa I que es activa en células lobulares pancreáticas (Swift y col., 1984, Cell 38, 639-646; Ornitz y col., 1986, Cold Spring Harbor Symp. Quant. Biol. 50, 399-409; MacDonald, 1987, Hepatology 7, 425-515); región de control del gen de la insulina que es activa en células pancreáticas beta (Hanahan, 1985, Nature 315, 115-122), región de control del gen de la inmunoglobulina que es activa en células linfoides (Grosschedl y col., 1984, Cell 38, 647-658; Adames y col., 1985, Nature 318, 533-538; Alexander y col., 1987, Mol. Cell. Biol. 7, 1436-1444), región de control del virus de tumor mamario de ratón que es activa en células testiculares, mamarías, linfoides y mastocitos (Leder y col., 1986, Cell 45, 485-495), región de control del gen de la albúmina que es activa en el hígado (Pinkert y col., 1987, Genes and Devel. 1, 268-276), región de control del gen de la alfa-fetoproteína que es activa en el hígado (Krumlauf y col., 1985, Mol. Cell. Biol. 5, 1639-1648; Hammer y col., 1987, Science 235, 53-58); región de control del gen de la 1-antitripsina alfa que es activa en el hígado (Kelsey y col., 1987, Genes and Devel. 1, 161-171), región de control del gen de la beta-globina que es activa en células mieloides (Mogram y col., 1985, Nature 315, 338-340; Kollias y col., 1986, Cell 46, 89-94), región de control del gen de la proteína básica de mielina que es activa en oligodendrocitos en el cerebro (Readhead y col., 1987, Cell 48, 703-712), región de control del gen de la cadena-2 ligera que es activa en músculo esquelético (Sani, 1985, Nature 314, 283-286) y región de control del gen de la hormona liberadora de gonadotropina que es activa en el hipotálamo (Mason y col., 1986, Science 234, 1372-1378).

Los vectores de expresión que contienen inserciones de genes toporrítmicos se pueden identificar mediante tres métodos generales: (a) hibridación de ácidos nucleicos, (b) presencia o ausencia de funciones génicas “marcadoras” y (c) expresión de secuencias insertadas. En el primer método, la presencia de un gen extraño insertado en un vector de expresión se puede detectar mediante hibridación de ácidos nucleicos empleando sondas que comprenden secuencias que son homólogas a un gen toporrítmico insertado. En el segundo método, el sistema recombinante vector/hospedador se puede identificar y seleccionar en base a la presencia o ausencia de ciertas funciones génicas “marcadoras” (por ejemplo, actividad de quinasa de timidina, resistencia a antibióticos, fenotipo de transformación, formación de cuerpos de oclusión en baculovirus, etc.) debidas a la inserción de genes extraños en el vector. Por ejemplo, si el gen toporrítmico está insertado dentro de la secuencia génica marcadora del vector, se pueden identificar los recombinantes que contienen la inserción toporrítmica mediante la ausencia de la función génica marcadora. En el tercer método, los vectores de expresión recombinantes se pueden identificar sometiendo a ensayo el producto génico extraño expresado por el recombinante. Tales ensayos se pueden basar, por ejemplo, en las propiedades físicas o funcionales del producto génico toporrítmico en sistemas de ensayo *in vitro*, por ejemplo, capacidad de agregación (adhesiva) (véanse las Secciones 6-8, más abajo).

Una vez que se identifica y se aísla una molécula particular de ADN recombinante, se pueden emplear diversos métodos conocidos para su multiplicación. Una vez establecidos un sistema de hospedador adecuado y unas condiciones de crecimiento, los vectores de expresión recombinantes se pueden multiplicar y preparar en cantidad. Tal y como se ha explicado previamente, los vectores de expresión que se pueden utilizar incluyen, pero no están limitados a los siguientes vectores o sus derivados: virus humanos o animales tales como el virus vaccinia o el adenovirus; virus de insectos tales como baculovirus; vectores de levaduras; vectores de bacteriófagos (por ejemplo, lambda) y vectores de ADN de plásmidos y cósmidos, por mencionar unos pocos.

Además, se puede escoger una cepa celular hospedadora que module la expresión de las secuencias insertadas o que modifique y procese el producto génico en la forma específica deseada. La expresión de ciertos promotores se puede elevar en presencia de ciertos inductores; por tanto, la expresión de la proteína toporrítmica generada por ingeniería genética se puede controlar. Además, diferentes células hospedadoras tienen características y mecanismos específicos para el procesamiento traduccional y post-traduccional y para la modificación (por ejemplo, glicosilación, escisión) de proteínas. Líneas celulares adecuadas o sistemas hospedadores se pueden escoger para asegurar la modificación deseada y el procesamiento de la proteína extraña expresada. Por ejemplo, la expresión en un sistema bacteriano se puede utilizar para producir un producto proteico con el centro no glicosilado. La expresión en levaduras producirá un producto glicosilado. La expresión en células de mamífero se puede utilizar para asegurar una glicosilación “natural” de una proteína toporrítmica heteróloga de mamífero. Además, diferentes sistemas de expresión de vector/hospedador pueden efectuar reacciones de procesamiento tales como escisiones proteolíticas con diferentes grados.

En otros ejemplos específicos, la proteína o fragmento toporrítmico adhesivo se puede expresar como una fusión o como un producto proteico quimérico (que comprende la proteína o el fragmento unido a una secuencia de proteína heteróloga). Un producto quimérico tal se puede preparar ligando las secuencias de ácidos nucleicos adecuadas que codifican las secuencias de aminoácidos deseadas entre sí, por métodos conocidos en la técnica, en el marco de codificación adecuado y expresando el producto quimérico por métodos conocidos generalmente en la técnica. Alternativamente, dicho producto quimérico se puede preparar mediante técnicas de síntesis de proteínas, por ejemplo, empleando un sintetizador de péptidos.

Ambas secuencias de ADNc y genómica se pueden clonar y expresar.

En otros ejemplos, una secuencia de ADNc de Notch humana se puede integrar en el cromosoma y expresar. Se pueden emplear procedimientos de recombinación homóloga conocidos en la técnica.

#### 5.4.1 Identificación y purificación del producto génico expresado

Una vez identificado un recombinante que expresa la secuencia génica toporrítmica, se puede analizar el producto génico. Esto se puede conseguir con ensayos que se basan en las propiedades físicas o funcionales del producto, que incluyen la marcación radiactiva del producto seguida de análisis mediante electroforesis en gel.

Una vez identificada la proteína toporrítmica, se puede aislar y purificar por métodos convencionales que incluyen la cromatografía (por ejemplo, cromatografía de intercambio iónico, de afinidad y en columna por tamaños), la centrifugación, la solubilidad diferencial o mediante cualquier otra técnica convencional para la purificación de proteínas. Las propiedades funcionales se pueden evaluar empleando cualquier ensayo adecuado, que incluyen, pero no están limitados a, ensayos de agregación (véanse las Secciones 6-8).

#### 5.5 Generación de anticuerpos para proteínas toporrítmicas y sus secuencias adhesivas

Se pueden utilizar fragmentos de proteínas toporrítmicas que median en la unión homotípica o heterotípica, o proteínas Notch humanas o Delta humanas o sus fragmentos, como inmunógenos para generar anticuerpos anti-proteína toporrítmica. Tales anticuerpos pueden ser policlonales o monoclonales. En una realización específica, se pueden preparar anticuerpos específicos para las repeticiones 11 y 12 similares a EGF de Notch. En otras realizaciones se pueden generar anticuerpos reactivos con la “parte adhesiva” de Delta. Un ejemplo de tales anticuerpos puede evitar la agregación en un ensayo *in vitro*. En otra realización, se producen anticuerpos específicos para Notch humana.

## ES 2 181 674 T5

Se pueden emplear distintos procedimientos conocidos en la técnica para la producción de anticuerpos policlonales para una proteína o péptido toporrítmico. En una realización particular, se pueden obtener anticuerpos policlonales de conejo para un epítipo de la proteína Notch humana codificada por una secuencia descrita en las Figuras 19, 20, 21 o 22 (SEQ ID NO:13 hasta SEQ ID NO:25), o una subsecuencia de las mismas. Para la producción de anticuerpo, se pueden inmunizar varios animales hospedadores mediante inyección con la proteína toporrítmica natural, o una versión sintética, o un fragmento de la misma, que incluyen, pero no están limitados a conejos, ratones, ratas, etc. Se pueden utilizar distintos coadyuvantes para incrementar la respuesta inmunológica, dependiendo de las especies hospedadoras, y que incluyen, pero no están limitados a, Freund (completo e incompleto), geles minerales como hidróxido de aluminio, sustancias tensioactivas como liolecitina, polioles plurónicos, polianiones, péptidos, emulsiones de aceite, hemocianinas de lapa bocallave, dinitrofenol y coadyuvantes humanos potencialmente útiles como BCG (bacilo Calmette-Guerin) y *corynebacterium parvum*.

Para la preparación de anticuerpos monoclonales dirigidos a una secuencia de proteína toporrítmica, se puede utilizar cualquier técnica que proporciona la producción de moléculas de anticuerpo mediante líneas celulares continuas en cultivo. Por ejemplo, la técnica de hibridoma, desarrollada originalmente por Kohler y Milstein (1975, Nature 256, 495-497), así como la técnica de trioma, la técnica de hibridoma de células B humanas (Kozbor y col., 1983, Immunology Today 4, 72) y la técnica de hibridoma de EBV para producir anticuerpos monoclonales humanos (Cole y col., 1985, en Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., págs. 77-96).

Los fragmentos de anticuerpo que contienen el idiotipo de la molécula se pueden generar mediante técnicas conocidas. Por ejemplo, tales fragmentos incluyen pero no están limitados a: el fragmento  $F(ab')_2$  que se puede producir mediante digestión con pepsina de la molécula de anticuerpo; los fragmentos Fab' que se pueden generar reduciendo los puentes disulfuro del fragmento  $F(ab')_2$  y los fragmentos Fab que se pueden generar tratando la molécula de anticuerpo con papaína y un agente reductor.

En la producción de anticuerpos, se puede realizar un escrutinio en busca del anticuerpo deseado mediante técnicas conocidas, por ejemplo ELISA (ensayo inmunoenzimático). Por ejemplo, para seleccionar anticuerpos que reconocen el dominio adhesivo de una proteína toporrítmica, se pueden someter a ensayo hibridomas generados para un producto que se une a un fragmento de proteína que contiene dicho dominio. Para la selección de un anticuerpo específico de Notch humana, se puede seleccionar en base a la unión positiva con Notch humana y a una falta de unión con Notch de *Drosophila*.

Los anticuerpos anteriores se pueden utilizar en métodos conocidos en la técnica relacionados con la localización y la actividad de las secuencias proteicas de la invención. Por ejemplo, se pueden emplear diversos inmunoensayos conocidos en la técnica, que incluyen pero no están limitados a sistemas de ensayo competitivos y no competitivos que emplean técnicas tales como radioinmunoensayos, ELISA (ensayo inmunoenzimático), inmunoensayos de tipo "sándwich", reacciones de precipitina, reacciones de precipitina con difusión en gel, ensayos de inmunodifusión, ensayos de aglutinación, inmunoensayos fluorescentes, inmunoensayos con proteína A y ensayos de inmunoelectroforesis, por mencionar unos cuantos.

### 5.6 Entrega de agentes en células que expresan Notch

La invención también proporciona métodos para la entrega de agentes en células que expresan Notch. Como se expone en la Sección 8, más abajo, al unirse a una proteína Notch en la superficie de una célula que expresa Notch, la proteína Delta parece que se absorbe dentro de la célula que expresa Notch. Se puede conseguir la entrega de agentes en una célula que expresa Notch, mediante la conjugación de un agente con una proteína Delta o un fragmento adhesivo de la misma, capaz de unirse a Notch, y exponer una célula que expresa Notch al conjugado, de modo que el conjugado es absorbido por la célula. El agente conjugado puede ser, pero no está limitado a un marcador o un agente biológicamente activo. El agente biológicamente activo puede ser un agente terapéutico, una toxina, un agente quimioterapéutico, un factor de crecimiento, una enzima, una hormona, un fármaco, un ácido nucleico, (por ejemplo, ADN o ARN antisentido), etc. En una realización, el marcador puede ser un agente formador de imágenes, que incluye pero no está limitado a agentes de contraste de metales pesados para formación de imágenes por rayos X, agentes formadores de imagen por resonancia magnética y nucleidos radiactivos (es decir, isótopos) para gammagrafía. En un aspecto preferido, el agente se conjuga en un sitio en la mitad amino-terminal de la molécula Delta.

El conjugado Delta-agente se puede entregar a la célula que expresa Notch, exponiendo la célula que expresa Notch a células que expresan el conjugado Delta-agente o exponiendo la célula que expresa Notch al conjugado Delta-agente en una solución, una suspensión u otro vehículo. Cuando la entrega es *in vivo*, el conjugado Delta-agente se puede formular en un vehículo farmacéuticamente aceptable o un excipiente, para incluir una composición farmacéutica. El vehículo farmacéuticamente aceptable puede comprender solución salina, solución salina tamponada con fosfato, etc. El conjugado Delta-agente se puede formular como un líquido, comprimido, píldora, polvo, en una forma de liberación lenta, en un liposoma, etc., y se puede administrar por vía oral, intravenosa, intramuscular, subcutánea, intraperitoneal, por mencionar algunas vías, siendo la selección preferida fácil, basándose en los conocimientos de una persona experta en la técnica.

## 6. Interacciones moleculares entre los productos proteicos de los loci neurógenos *Notch* y *Delta*, dos genes homólogos a *EGF* en *Drosophila*

Para examinar la posibilidad de asociación intermolecular entre los productos de los genes *Notch* y *Delta*, estudiamos los efectos de su expresión sobre la agregación en células Schneider's 2 (S2) de *Drosophila* (Fehon y col., 1990, Cell 61, 523-534). En esta memoria ofrecemos una evidencia directa de interacciones intermoleculares entre *Notch* y *Delta* y describimos un sistema de ensayo que se puede utilizar en la disección de los componentes de esta interacción. Mostramos que células cultivadas S2 de *Drosophila* que normalmente no son adhesivas, que expresan *Notch*, se unen específicamente a células que expresan *Delta*, y que esta agregación es dependiente de calcio. Además, aun cuando las células que expresan *Notch* no se unen entre sí, las células que expresan *Delta* se unen entre sí, lo que sugiere que *Notch* y *Delta* pueden competir por la unión a *Delta* en la superficie celular. También presentamos evidencias que indican que *Notch* y *Delta* forman complejos solubles en detergente, ambos en células cultivadas y en células embrionarias, sugiriendo que *Notch* y *Delta* interactúan directamente a nivel molecular *in vitro* e *in vivo*. Nuestros análisis sugieren que las proteínas *Notch* y *Delta* interactúan en la superficie celular mediante sus dominios extracelulares.

### 6.1 Procedimientos experimentales

#### 6.1.1 Estructuras artificiales de expresión

Para la estructura artificial de expresión de *Notch*, se ligó con extremos romos el fragmento HpaI de 6 kb procedente del extremo 5' de la secuencia codificadora de *Notch* en MglIIa (Ramos y col., 1988, Genetics 123, 337-348) en el vector del promotor de la metalotioneína pRmHa-3 (Bunch y col., 1988, Nucl. Acids Res. 16, 1043-1061), después el vector se cortó con EcoRI y los extremos se rellenaron con el fragmento Klenow de la polimerasa I de ADN (Maniatis y col., 1982, Molecular Cloning; A Laboratory Manual (Cold Spring Harbor, Nueva York: Cold Spring Harbor Laboratory)). Se aisló un único transformante, orientado incorrectamente. El ADN de este transformante se digirió a continuación con SacI y se aisló un fragmento resultante de 3 kb que contenía el extremo 5' de la secuencia codificadora de *Notch* fusionada con el poliengarzador procedente de pRmHa-3. Este fragmento se ligó entonces en el sitio SacI de pRmHa-3 con la orientación correcta. El ADN procedente de esta estructura artificial se digirió con KpnI y XbaI para eliminar la mayoría de la secuencia de *Notch* y toda la secuencia de la señal de poliadenilación Adh en pRmHa-3 y se ligó con un fragmento KpnI-XbaI de 11 kb, procedente de MglIIa que contenía el resto de la secuencia codificadora de *Notch* y las secuencias 3' necesarias para la poliadenilación. En la estructura artificial resultante, denominada pMtNMg, el promotor de la metalotioneína en pRmHa-3 se fusiona con secuencias de *Notch*, que comienzan 20 nucleótidos aguas arriba del sitio de inicio de la traducción.

Para la estructura artificial de *Notch* extracelular (ECN1), el cósmido de *Notch* CosP479BE (Ramos y col., 1989, Genetics 123, 337-348), que contiene todas las secuencias genómicas de *Notch* necesarias para una función normal de *Notch* *in vivo*, se digirió parcialmente con AatII. Los extremos del fragmento se hicieron romos empleando la actividad exonucleasa de la polimerasa T4 de ADN (Maniatis y col., 1982, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, Nueva York: Cold Spring Harbor Laboratory)), y los fragmentos se digirieron de nuevo entonces completamente con StuI. Los fragmentos resultantes se separaron en un gel de agarosa con baja temperatura de fusión (SeaPlaque, FMC BioProducts), y el fragmento mayor se escindió. Este fragmento se ligó después con extremos romos consigo mismo. Esto dio como resultado una delección interna de las secuencias codificadoras de *Notch* desde el aminoácido 1790 hasta 2625 inclusive (Wharton y col., 1985, Cell 43, 567-581), y un cambio del marco de lectura pronosticado que produce un nuevo extremo carboxilo de 59 aminoácidos. (La unión ligada de esta estructura artificial no se ha comprobado por secuenciación).

Para la estructura artificial de expresión de *Delta*, el ADNc de D11 (Kopczynski y col., 1988, Genes Dev. 2, 1723-1735) que incluye la capacidad codificadora completa de *Delta*, se insertó en el sitio EcoRI de pRmHa-3. Esta estructura artificial se denominó pMTD11.

#### 6.1.2 Preparación de anticuerpos

La línea celular de hibridoma C17.9C6 se obtuvo a partir de un ratón inmunizado con una proteína de fusión basada en un fragmento Sall-HindIII de 2,1 kb que incluye las secuencias codificadoras de la mayoría del dominio intracelular de *Notch* (aminoácidos 1791-2504; Wharton y col., 1985, Cell 43, 567-581). El fragmento se subclonó en pUR289 (Ruther y Muller-Hill, 1983, EMBO J. 2, 1791-1794), y después se transfirió en el vector de expresión pATH 1 (Dieckmann y Tzagoloff, 1985, J. Biol. Chem. 260, 1513-1520) como un fragmento BglII-HindIII. La proteína de fusión soluble se expresó, se precipitó con  $(\text{NH}_4)_2\text{SO}_4$  al 25%, se resuspendió en urea 6 M y se purificó mediante enfoque isoeléctrico preparativo, empleando un Rotofor (BioRad) (para más detalles, véase Fehon, 1989, Rotofor Review N° 7, Boletín 1518, Richmond, California: Bio-Rad Laboratories).

Se produjeron antisueros policlonales de ratón contra el dominio extracelular de *Notch*, empleando cuatro fragmentos BstY1 de 0,8 kb (aminoácidos 237-501; Wharton y col., 1985, Cell 43, 567-581), 1,1 kb (aminoácidos 501-868), 0,99 kb (aminoácidos 868-1200) y 1,4 kb (aminoácidos 1465-1935) de longitud, que abarcaban desde la quinta repetición similar a EGF a través del dominio transmembranal, se insertaron individualmente en marco en el vector de expresión adecuado pGEX (Smith y Johnson, 1988, Gene 67, 31-40). Las proteínas de fusión se purificaron en lechos de glutatión-agarosa (SIGMA). Antisueros de ratón y rata se precipitaron con  $(\text{NH}_4)_2\text{SO}_4$  al 50% y se resuspendieron en PBS (NaCl 150 mM,  $\text{Na}_2\text{HPO}_4$  14 mM,  $\text{NaH}_2\text{PO}_4$  6 mM) con  $\text{NaN}_3$  al 0,02%.

## ES 2 181 674 T5

La línea celular de hibridoma 201 se obtuvo a partir de un ratón inmunizado con una proteína de fusión basada en un fragmento ClaI de 0,54 kb que incluye secuencias codificadoras procedentes del dominio extracelular de Delta (Kopczynski y col., 1988, Genes Dev. 2, 1723-1735) se subclonó en el sitio ClaI dentro del gen *lacZ* de pUR 288 (Rut-her y Muller-Hill, 1983, EMBO J. 2, 1791-1794). Este fragmento incluye secuencias que se extienden desde la cuarta hasta la novena repetición similar a EGF en Delta (aminoácidos 350-529). La proteína de fusión se preparó aislando cuerpos de inclusión (Gilmer y col., 1982, Proc. Natl. Acad. Sci. USA 79, 2152-2156); los cuerpos de inclusión se solubilizaron en urea (Carroll y Laughon, 1987, en DNA Cloning, Volumen III, D.M. Glover, compilador (Oxford: IRL Press), págs. 89-111) antes del uso en la inmunización.

Antisueros policlonales de rata se obtuvieron después de inmunizar con antígeno, obtenido a partir de la misma estructura artificial de proteína de fusión. En este caso, la proteína de fusión se preparó por lisis de células inducidas con IPTG en tampón SDS-Laemmli (Carroll y Laughon, 1987, en DNA Cloning, Volumen III, D.M. Glover, compilador (Oxford: IRL Press), págs. 89-111), separación de proteínas mediante SDS-PAGE, escisión de la banda apropiada del gel y electroelución del antígeno a partir del trozo cortado de gel para uso en inmunización (Harlow y Lane, 1988, Antibodies: A Laboratory Manual (Cold Spring Harbor, Nueva York: Cold Spring Harbor Laboratory)).

### 6.1.3. Cultivo celular y transfección

La línea celular S2 (Schneider, 1972, J. Embryol. Exp. Morph. 27, 353-365) se hizo crecer en medio M3 (preparado por Hazleton Co.) suplementado con 2,5 mg/ml de Bacto-Peptona (Difco), 1 mg/ml de TC Yeastolate (Difco), suero de ternera fetal inactivado con calor al 11% (FCS) (Hyclone) y 100 U/ml de penicilina-100 µg/ml de estreptomycin-0,25 µg/ml de fungizona (Hazleton). Las células que crecían en fase logarítmica hasta  $\sim 2 \times 10^6$  células/ml se transfectaron con 20 µg de coprecipitado de ADN-fosfato cálcico en 1 ml por 5 ml de cultivo tal y como se ha descrito previamente (Wigler y col., 1979, Proc. Natl. Acad. Sci. USA 78, 1373-1376) con la excepción de que se utilizó el tampón BES (SIGMA) en lugar de tampón HEPES (Chen y Okayama, 1987, Mol. Cell. Biol. 7, 2745-2572). Después de 16-18 h, las células se transfirieron a tubos de centrifuga cónicos, se sedimentaron en una centrifuga clínica a velocidad completa durante 30 segundos, se lavaron una vez con 1/4 del volumen de medio completo de nuevo aporte, se resuspendieron en su volumen original de medio completo y se llevaron de nuevo al frasco original. A las células transfectadas se les permitió recuperarse a continuación durante 24 h antes de la inducción.

### 6.1.4. Ensayos de agregación

La expresión de las estructuras artificiales de metalotioneína de *Notch* y *Delta* fue inducida por la adición de  $\text{CuSO}_4$  hasta 0,7 mM. Las células transfectadas con la estructura artificial ECN1, se trataron de forma similar. Se emplearon dos tipos de ensayos de agregación. En el primer ensayo, se colocó un total de 3 ml de células ( $5-10 \times 10^6$  células/ml) en un matraz Erlenmeyer de 25 ml y se hizo rotar a 40-50 rpm en un agitador con rotación durante 24-48 horas a temperatura ambiente. Para estos experimentos, las células se mezclaron 1-4 h después de que comenzara la inducción y la inducción continuó a lo largo del periodo de agregación. En el segundo ensayo, se colocaron  $\sim 0,6$  ml de células en un tubo Eppendorf de 0,6 ml (dejando una pequeña burbuja) después de inducir durante una noche (12-16 h) a temperatura ambiente y balancear suavemente durante 1-2 h a 4°C. Los experimentos de inhibición del anticuerpo y los experimentos de dependencia de  $\text{Ca}^{2+}$ , se realizaron empleando el último ensayo. Para los experimentos de dependencia de  $\text{Ca}^{2+}$  se recolectaron primero las células y se lavaron en solución salina equilibrada (BSS) con FCS al 11% (BSS-FCS; FCS se dializó contra NaCl al 0,9%, Tris 5 mM [pH 7,5]) o en BSS-FCS exento de  $\text{Ca}^{2+}$  que contenía EGTA 10 mM (Snow y col., 1989, Cell 59, 313-323) y después se resuspendieron en el mismo medio con el volumen original. Para los experimentos de inhibición del anticuerpo se recogieron células transfectadas con *Notch* y se lavaron en medio M3 y a continuación se trataron antes de la agregación en medio M3 durante 1 h a 4°C, con una dilución 1:250 de sueros inmunes o preinmunizados de cada uno de los cuatro ratones inmunizados con proteínas de fusión que contenían segmentos procedentes del dominio extracelular de Notch (véase, Preparación de Anticuerpos, arriba).

### 6.1.5. Inmunofluorescencia

Las células se recogieron por centrifugación (3000 rpm durante 20 segundos en una microcentrifuga de Eppendorf) y se fijaron en tubos Eppendorf de 0,6 ml con 0,5 ml de paraformaldehído al 2% preparado recientemente, en PBS durante 10 min a temperatura ambiente. Después de la fijación, se recogieron las células por centrifugación, se lavaron dos veces en PBS y se tiñeron durante 1 h con anticuerpo primario en PBS, con saponina al 0,1% (SIGMA) y suero de cabra normal al 1% (Pocono Rabbit Farm, Canadensis, PA). El material sobrenadante del anticuerpo monoclonal se diluyó 1:10 y los sueros de ratón o rata se diluyeron 1:1000 para esta etapa. Las células se lavaron a continuación una vez en PBS y se tiñeron durante 1 h con anticuerpos secundarios específicos (doble grado de marcación de cabra anti-ratón y cabra anti-rata, Jackson Immunoresearch) en suero de cabra normal y PBS-saponina. Después de esta incubación, las células se lavaron dos veces en PBS y se montaron en portaobjetos con 90% de glicerol, 10% de Tris 1 M (pH 8,0) y 0,5% de galato de n-propilo. Las células se visionaron bajo epifluorescencia en un microscopio Orthoplan 2 de Leitz.

Se tomaron micrografías confocales empleando el sistema de Bio-Rad MRC 500 conectado a un microscopio Axio-vert compuesto de Zeiss. Las imágenes se recogieron empleando grupos de filtros BHS y GHS, se alinearon empleando el programa ALIGN, y se fusionaron empleando MERGE. La extracción fluorescente desde el canal verde hasta el canal rojo, se redujo empleando el programa BLEED (todos los programas fueron proporcionados por Bio-Rad). Las fotografías se obtuvieron directamente de la pantalla del ordenador empleando una película Ectar 125 de Kodak.

6.1.6. *Materiales lisados celulares, inmunoprecipitaciones y transferencias de tipo western*

Material lisado no desnaturalizado con detergente de cultivo de tejido y de embriones de tipo silvestre Canton-S, se preparará en hielo con un volumen de tampón de lisis (NaCl 300 mM, Tris 50 mM [pH 8,0], 0,5% de NP-40, 0,5% de desoxicolato, CaCl<sub>2</sub> 1 mM, MgCl<sub>2</sub> 1 mM) de ~10 células con fluoruro de fenilmetilsulfonilo 1 mM (PMSF) y fluorofosfato de diisopropilo diluido 1:2500 como inhibidores de proteasas. El material lisado se trituró secuencialmente empleando agujas de 18G, 21G y 25G fijadas a jeringas de tuberculina de 1 cc y después se centrifugaron a velocidad máxima en una microfuga, 10 min a 4°C para separar el material insoluble. La inmunoprecipitación se realizó añadiendo ~1 µg de anticuerpo (1-2 µl de antisuero policlonal) a 250-500 µl de material lisado de células y se incubó durante 1 h a 4°C con agitación. A esta mezcla se añadieron 15 µg de anticuerpos de cabra anti-ratón (Jackson Immunoresearch; estos anticuerpos reconocen las IgG tanto de ratón como de rata) y se dejó incubar durante 1 h a 4°C con agitación. Esto fue seguido de la adición de 100 µl de bacterias *Staphylococcus aureus* (Staph A) fijadas (Zysorbin, Zymed; resuspendidas según las instrucciones del fabricante), que se recolectaron, se lavaron cinco veces en tampón de lisis y se incubaron durante una hora más. Los complejos de Staph A-anticuerpo se sedimentaron a continuación por centrifugación y se lavaron tres veces en tampón de lisis seguido de dos lavados de 15 min en tampón de lisis. Después de transferir a un tubo nuevo, el material precipitado se suspendió en 50 µl de tampón de muestra SDS-PAGE, se hirvió inmediatamente durante 10 min, se dejó correr en geles con gradiente 3%-15% y se transfirió a nitrocelulosa y se detectó empleando anticuerpos monoclonales y anticuerpos secundarios de cabra anti-ratón conjugados con HRP, tal y como se ha descrito previamente (Johansen y col., 1989, *J. Cell Biol.* 109, 2427-2440). Para las muestras de proteína celular total empleadas en las transferencias de tipo Western, (Figura 2), se recogieron células por centrifugación, se lisaron en un volumen de 10 células de tampón por muestra que contenía PMSF 1 mM y se hirvió inmediatamente.

6.2 *Resultados*6.2.1 *La expresión de notch y delta en células cultivadas*

Para detectar interacciones entre Notch y Delta, examinamos el comportamiento de células que expresaban estas proteínas en su superficie, empleando un ensayo de agregación. Escogimos la línea celular S2 (Schneider, 1972, *J. Embryol. Exp. Morph.* 27, 353-365) para estos estudios por distintas razones. Primera, estas células son relativamente no adhesivas, crecen en suspensión y se han utilizado previamente en un ensayo similar para estudiar la función de la fasciclina III (Snow y col., 1989, *Cell* 59, 313-323). Segunda, son fácilmente transfectables y está disponible un vector de promotor de metalotioneína inducible que se ha diseñado para la expresión de genes exógenos en células cultivadas de *Drosophila* (Bunch y col., 1988, *Nucl. Acids Res.* 16, 1043-1061). Tercera, las células S2 expresan un mensaje de *Notch* aberrante y *Notch* no detectable debido a una reorganización del extremo 5' de la secuencia codificadora de *Notch* (véase a continuación). Estas células también expresan Delta no detectable (véase a continuación).

Dibujos esquemáticos de las estructuras artificiales empleadas, se muestran en la Figura 1 (véase la Sección 6.1 de Procedimientos Experimentales, para más detalles). Para expresar Notch en células cultivadas, el minigen MG11a de *Notch*, descrito por Ramos y col. (1989, *Genetics* 123, 337-348) se insertó en el vector promotor de metalotioneína pRmHa-3 (Bunch y col., 1988, *Nucl. Acids Res.* 16, 1043-1061). La estructura artificial de expresión de *Delta* se preparó insertando ADNc de D11, que contiene la secuencia codificadora completa para Delta, procedente del transcrito de *Delta* embrionario principal (5.4Z; Kopczynski y col., 1988, *Genes Dev.* 2, 1723-1735) en el mismo vector. Una tercera estructura artificial, denominada ECN1 para "Notch 1 extracelular", contiene la región promotora 5' de *Notch* y la señal de poliadenilación 3' de *Notch*, junto con una capacidad codificadora de las regiones extracelulares y transmembranales del gen *Notch* procedente de secuencias genómicas, pero carece de secuencias codificadoras de 835 aminoácidos del dominio intracelular de ~1000 aminoácidos. Además, debido a un cambio del marco de lectura pronosticado, los 78 restos de aminoácidos carboxi-terminales restantes son reemplazados por una nueva cola carboxiterminal de 59 aminoácidos (véase, Procedimientos Experimentales).

Para todos los experimentos descritos en esta memoria, las estructuras artificiales de expresión se transfectaron en células S2 y se expresaron más transitoriamente que en transformantes estables. Las células con expresión estaban compuestas típicamente por 1%-5% del total de la población celular, como se deduce por la tinción inmunofluorescente (datos no mostrados). Una transferencia de tipo Western de proteínas expresadas después de la transfección, se muestra en la Figura 2. Las células no transfectadas no expresan niveles detectables de Notch o de Delta. Sin embargo, después de la transfección, proteínas con los pesos moleculares evidentes pronosticados, son fácilmente detectables empleando anticuerpos monoclonales específicos para cada una de estas proteínas, respectivamente. En el caso de Notch, múltiples bandas eran evidentes en células transfectadas por debajo del producto de longitud completa de ~300 kd. No sabemos hasta ahora si estas bandas representan una degradación de Notch durante la preparación de la muestra o quizás en la síntesis o en el procesamiento de los intermediarios de Notch, que están presentes dentro de células, pero las detectamos de forma estable en muestras procedentes de células transfectadas y de embriones. Además, realizamos una tinción inmunofluorescente de células transfectadas vivas con anticuerpos específicos de los dominios extracelulares de cada proteína, para someter a ensayo la expresión de la superficie celular de estas proteínas. En cada caso encontramos una tinción de superficie, tal y como se esperaba para un antígeno de superficie. Resumiendo, estos resultados muestran claramente que las estructuras artificiales de *Notch* y *Delta* mantienen la expresión de proteínas de los tamaños esperados y la localización subcelular.

## ES 2 181 674 T5

### 6.2.2. Células que expresan agregado de Notch y Delta

Para someter a ensayo la afirmación de que Notch y Delta interaccionan, diseñamos un ensayo de agregación sencillo para detectar estas interacciones entre proteínas expresadas en la superficie de células S2. Razonamos que si Notch y Delta eran capaces de formar complejos heterotípicos estables en la superficie celular, entonces células que expresan estas proteínas se podrían unir entre sí y formar agregados bajo condiciones adecuadas. Un sistema de ensayo similar se ha descrito recientemente para la proteína fasciclina III (Snow y col., 1989, Cell 59, 313-323).

Las células S2 en crecimiento en fase logarítmica se transfectaron separadamente con la estructura artificial del promotor de metalotioneína tanto de Notch como de Delta. Después de inducir con CuSO<sub>4</sub>, las células transfectadas se mezclaron en cantidades iguales y se dejaron agregar durante una noche a temperatura ambiente (para más detalles, véase la Sección 6.1 de Procedimientos Experimentales). Alternativamente, en algunos experimentos que pretendían reducir la actividad metabólica, las células se mezclaron cuidadosamente a 4°C durante 1-2 h. Para determinar si se han formado los agregados, las células se procesaron para microscopía inmunofluorescente empleando anticuerpos específicos para cada producto génico y anticuerpos secundarios marcados diferentemente con fluorescencia. Como se ha mencionado previamente, las células con expresión constituyen normalmente menos del 5% del total de la población celular debido a que empleamos transformantes transitorios más que transformantes estables. Las células restantes no expresaron una proteína dada o la expresaron a niveles demasiado bajos para la detección con microscopía inmunofluorescente. Como testigos, realizamos agregaciones con un solo tipo de célula transfectada.

La Figura 3 muestra fotomicrografías representativas de experimentos de agregación y la Tabla I presenta los resultados en forma numérica. Tal y como es evidente por la Figura 3C y la Tabla I, aunque las células que expresan Notch (Notch<sup>+</sup>) solas no forman agregados en nuestro ensayo, las células que expresan Delta (Delta<sup>+</sup>) si que los forman.

(Tabla pasa a página siguiente)

TABLA I  
PORCENTAJES DE CÉLULAS NOTCH' Y DELTA' EN AGREGADOS<sup>a</sup>

	Notch'			Delta'			Notch' - Delta'		
	Células tes- tigo <sup>b</sup>	Células agregadas <sup>c</sup>	Células testigo <sup>b</sup>	Células agregadas <sup>c</sup>	Células agregadas <sup>c</sup>	Global <sup>d</sup>	Células N°	Células DI <sup>e</sup>	
Experimento 1	0	0	19	37	32	26	42		
Experimento 2	-	1	-	40	54	47	79		
Experimento 3	0	-	12	-	43	42	44		
Experimento 4	5	5	20	-	47	41	59		
Experimento 5 <sup>g</sup>	-	2	-	48	71	66	82		
Experimento 6 <sup>h</sup>	0	0	13	61	63	60	73		

<sup>a</sup> Agregados definidos como agrupaciones de cuatro o más células con expresión. Para todos los valores, se marcaron al menos 100 unidades de células con expresión (células aisladas o agrupaciones celulares). Notch', que expresa Notch; Delta', que expresa Delta.

<sup>b</sup> Células testigo tomadas directamente de los matraces de transfección sin incubación en el ensayo de agregación.

<sup>c</sup> Células testigo después de incubar en el ensayo de agregación.

<sup>d</sup> Datos combinados de agregación para células Notch' y Delta' en agregados Notch'-Delta'.

<sup>e</sup> Datos de agregación para células Notch' en agregados Notch'-Delta'.

<sup>f</sup> Datos de agregación para células Delta' en agregados Notch'-Delta'.

<sup>g</sup> Células de este experimento precedentes de la misma transfección que el Experimento 4.

<sup>h</sup> Datos de experimentos de agregación de 48 h. Todos los demás datos son de experimentos de agregación de 24 h.

La tendencia de células Delta<sup>+</sup> a agregarse era evidente incluso en muestras de testigos no agregados (Tabla I), en donde agrupaciones celulares de 4-8 células que probablemente se originaron por la adherencia entre células mitóticas hermanas se presentaban comúnmente. Sin embargo, las agrupaciones eran más comunes después de incubación bajo condiciones de agregación (por ejemplo, 19% de células Delta<sup>+</sup> en agregados antes de la incubación, frente a 37% de células Delta<sup>+</sup> en agregados después de la incubación; Experimento 1 en la Tabla I), indicando que las células Delta<sup>+</sup> son capaces de formar contactos estables entre sí en este ensayo. Es importante tener en cuenta que, aunque las células sin teñir constituyen sobre 90% de las células en nuestras transfecciones transitorias, no las encontramos nunca dentro de los agregados. En raras ocasiones, se encontraron células sin teñir en el borde de un agregado. Debido a la presencia común de células que se tiñen débilmente en los bordes de agregados, es probable que estas células aparentemente sin expresión, estuvieran transfectadas pero expresaran niveles de Delta insuficientes para ser detectadas por inmunofluorescencia.

Como un contraste remarcable con experimentos testigos con células Notch<sup>+</sup> aisladas, la agregación de mezclas de células Notch<sup>+</sup> y Delta<sup>+</sup> daba como resultado la formación de agrupaciones de hasta 20 células o más (Figuras 3D-3H, Tabla I). Como muestra la Tabla I, la fracción de células con expresión encontradas en agrupaciones de cuatro o más células teñidas después de 24 h de agregación, fluctúa desde 32%-54% en mezclas de células Notch<sup>+</sup> y Delta<sup>+</sup>. Esta fluctuación era similar a la observada en células Delta<sup>+</sup> aisladas (37%-40%) pero muy diferente a la de células Notch<sup>+</sup> aisladas (sólo 0%-5%). Aunque se encontraron unas pocas agrupaciones que constaban sólo de células Delta<sup>+</sup>, las células Notch<sup>+</sup> no se encontraron nunca en agrupaciones mayores de cuatro a cinco células, a no ser que también estuvieran presentes células Delta<sup>+</sup>. De nuevo, ninguna de las células dentro de estas agrupaciones expresaban Notch ni Delta, aunque las células transfectadas estaban compuestas sólo por una pequeña fracción de la población celular total. A las 48 h (Tabla I, experimentos 5 y 6), el grado de agregación parecía mayor (63%-71%), sugiriendo que después de 24 h, la agregación no ha alcanzado todavía un máximo bajo esas condiciones. También, células cotransfectadas con estructuras artificiales de *Notch* y *Delta* (de modo que todas las células transfectadas expresan ambas proteínas) se agregaron de forma similar bajo las mismas condiciones experimentales.

Estos resultados indican que la agregación observada en estos experimentos requiere la expresión de Notch y Delta y que no es debida a la expresión fortuita de otra proteína interaccionante en células S2 no transfectadas. Sometimos posteriormente a ensayo la especificidad de esta interacción, diluyendo células Notch<sup>+</sup> y Delta<sup>+</sup> 10 veces con células S2 no transfectadas y permitiéndolas después agregarse durante 24 h a temperatura ambiente. En este experimento, 39% de las células con expresión se encontraron en agregados con otras células con expresión, aunque formaban menos de 0,1% del total de la población celular. Sin embargo no era sorprendente, que estos agregados fueran menores por término medio que los encontrados en experimentos de agregación convencionales. Además, para vigilar la posibilidad de que células Notch<sup>+</sup> fueran reclutadas de forma no específica en los agregados Delta<sup>+</sup>, debido a que sobreexpresan un tipo único de proteína en la superficie celular, mezclamos células Delta<sup>+</sup> con células que expresaban neuroglía, una proteína transmembranal de la superficie celular (Bieber y col., 1989, Cell 59, 447-460), bajo el control del promotor de metalotioneína (esta estructura artificial de metalotioneína-neuroglía la proporcionó amablemente A. Bieber y C. Goodman). No observamos ninguna tendencia de las células de neuroglía<sup>+</sup> a adherirse a agregados Delta<sup>+</sup>, indicando que la agregación Notch-Delta no es únicamente el resultado de altos niveles de expresión proteica sobre la superficie celular.

También sometimos a ensayo directamente la implicación de Notch en el proceso de agregación examinando, el efecto de una mezcla de antisueros policlonales dirigidos contra proteínas de fusión que abarcaban casi todo el dominio extracelular de Notch en agregación (véase la Sección 6.1 de Procedimientos Experimentales). Para minimizar los artefactos que pudieran surgir debido a la respuesta metabólica para arreglar los antígenos de superficie, se realizó un tratamiento con anticuerpos y el ensayo de agregación a 4°C en estos experimentos. Las células Notch<sup>+</sup> se incubaron con sueros de ratón preinmunes o inmunes durante 1 h, se añadieron las células Delta<sup>+</sup> y se realizó la agregación durante 1-2 h. Aunque las células Notch<sup>+</sup> tratadas previamente con sueros preinmunes se agregaban con células Delta<sup>+</sup> (en uno de los tres experimentos, 23% de las células Notch<sup>+</sup> estaban en agregados de células Notch<sup>+</sup>-Delta<sup>+</sup>), las tratadas con sueros inmunes no se agregaban (sólo 2% de las células Notch<sup>+</sup> estaban en agregados). Este resultado sugiere que el dominio extracelular de Notch es necesario para la agregación de células Notch<sup>+</sup>-Delta<sup>+</sup>, aunque no podemos excluir la posibilidad de que la reducida agregación se debiera a efectos estéricos inhibitorios o de estructura de membrana resultantes de la exposición de células Notch<sup>+</sup> al antisuero.

Otras tres observaciones dignas de ser mencionadas son evidentes en la Figura 3. Primera, aunque Delta era casi siempre evidente sólo en la superficie celular (Figuras 3B y 3C), la tinción de Notch era siempre evidente tanto en la superficie celular como intracelularmente, frecuentemente asociada con estructuras vesiculares (Figura 3A). Segunda, observamos consistentemente una diferencia morfológica entre células Delta<sup>+</sup> y Notch<sup>+</sup> en agregados mezclados que se incubaron durante una noche. Las células Delta<sup>+</sup> tenían frecuentemente amplias extensiones que rodeaban completamente las células Notch<sup>+</sup> adyacentes, mientras que células Notch<sup>+</sup> tenían casi siempre una apariencia redondeada sin extensiones citoplasmáticas observables (Figura 3G). Tercera, Notch y Delta aparecían frecuentemente agregadas dentro de regiones de contacto entre células Notch<sup>+</sup> y Delta<sup>+</sup>, produciendo una banda distinta de tinción inmunofluorescente (Figuras 3D-3F). Estas bandas eran fácilmente visibles en secciones ópticas visionadas con microscopio confocal (Figura 3H), indicando que no se debían únicamente a un artefacto del montaje íntegro. También observamos que estas bandas se formaban rápidamente (al cabo de 2 h de mezclar las células) y a 4°C, indicando que su formación no dependía probablemente del metabolismo celular. Estas observaciones se habrían esperado si, dentro de las regiones de contacto celular, Notch y Delta se unen entre sí y por tanto se quedan inmovilizadas. Este patrón de expresión

también es consistente con el observado para otras proteínas que median en la agregación celular (Takeichi, 1988, Development 102, 639-655; Snow y col., 1989, Cell 59, 313-323).

### 5 6.2.3. La agregación mediada por Notch-Delta es dependiente de calcio

Estudios previos han sugerido que repeticiones similares a EGF que contienen una secuencia de consenso particular pueden servir como dominios de unión a calcio ( $\text{Ca}^{2+}$ ) (Morita y col., 1984, J. Biol. Chem. 259, 5698-5704; Sugo y col., 1984, J. Biol. Chem. 259, 5705-5710; Rees y col., 1988, EMBO J. 7, 2053-2061; Handford y col., 1990, EMBO J. 9, 475-480). Para al menos dos de estas proteínas, C y C1, la unión a  $\text{Ca}^{2+}$  ha demostrado ser adicionalmente un componente necesario para sus interacciones con otras proteínas (Villiers y col., 1980, FEBS Lett. 117, 289-294; Esmon y col., 1983, J. Biol. Chem. 258, 5548-5553; Johnson y col., 1983, J. Biol. Chem. 258, 5554-5560). Muchas de las repeticiones homólogas a EGF dentro de Notch y la mayoría de estas dentro de Delta, contienen la secuencia de consenso necesaria para la unión a  $\text{Ca}^{2+}$  (Rees y col., 1988, EMBO J. 7, 2053-2061; Stenflo y col., 1987, Proc. Natl. Acad. Sci. USA 84, 368-372; Kopczynski y col., 1988, Genes Dev. 2, 1723-1735; Handford y col., 1990, EMBO J. 9, 475-480), aunque no se ha determinado todavía si estas proteínas se unen o no a calcio. Por tanto sometimos a ensayo la capacidad de células con expresión a agregarse en presencia o en ausencia de iones  $\text{Ca}^{2+}$ , para determinar si existe un requerimiento de ión  $\text{Ca}^{2+}$  para la agregación de Notch-Delta. Para minimizar posibles efectos inespecíficos debidos a respuestas metabólicas por la eliminación de  $\text{Ca}^{2+}$ , realizamos estos experimentos a 4°C. Las mezclas testigos de células Notch<sup>+</sup> y células Delta<sup>+</sup> incubadas con condiciones de agregación en un medio que contenía  $\text{Ca}^{2+}$  a 4°C, formaban rápidamente agregados (un promedio de 34% ± 13%, media ± SD, n = 3; Tabla II). Por el contrario, células mezcladas en medio que carecía de iones  $\text{Ca}^{2+}$  y que contenía EGTA formaban pocos agregados (5% ± 5%). Estos resultados demuestran claramente una dependencia de  $\text{Ca}^{2+}$  en la agregación mediada por Notch-Delta, y contrastan marcadamente con los publicados recientemente para las proteínas fasciclina III y fasciclina I de Drosophila en células S2 (Snow y col., 1989, Cell 59, 313-323; Elkins y col., 1990 J. Cell Biol. 110, 1825-1832), en los que no se detectaba ningún efecto de la eliminación de iones  $\text{Ca}^{2+}$  sobre la agregación mediada por ambas proteínas.

30 TABLA II

Efecto de  $\text{Ca}^{2+}$  exógeno sobre la agregación<sup>a</sup> de Notch<sup>+</sup>-Delta<sup>+</sup>

	Sin iones $\text{Ca}^{2+}$			Con iones $\text{Ca}^{2+}$		
	Global <sup>b</sup>	Células N <sup>c</sup>	Células D1 <sup>d</sup>	Global <sup>b</sup>	Células N <sup>c</sup>	Células D1 <sup>d</sup>
Experimento 1	4	2	5	28	28	27
Experimento 2	12	0	13	53	63	50
Experimento 3	0	0	0	22	28	17

45 <sup>a</sup> Datos presentados como porcentajes de células con expresión encontradas en agregados (como en la Tabla I).

<sup>b</sup> Datos de agregación combinada para células Notch<sup>+</sup> y Delta<sup>+</sup>.

50 <sup>c</sup> Datos de agregación para células Notch<sup>+</sup> en agregados Notch<sup>+</sup>-Delta<sup>+</sup>.

55 <sup>d</sup> Datos de agregación para células Delta<sup>+</sup> en agregados Notch<sup>+</sup>-Delta<sup>+</sup>.

### 60 6.2.4. Notch y Delta interaccionan dentro de una sola célula

Nos preguntamos si Notch y Delta están asociadas dentro de la membrana de una célula que expresa ambas proteínas, examinando las distribuciones de Notch y Delta en células cotransfectadas. Como se muestra en las Figuras 4A y 4B, estas dos proteínas muestran frecuentemente distribuciones muy similares en la superficie de células cotransfectadas. Para someter a ensayo si la co-localización observada era una coincidencia o representaba una interacción estable entre Notch y Delta, tratamos células vivas con un exceso de antisuero policlonal anti-Notch. Este tratamiento daba como resultado la "formación de placas" de Notch en la superficie de células con expresión, en placas discretas detectadas por inmunofluorescencia. Había una correlación distinta entre las distribuciones de Notch y Delta en las superficies de estas células después de este tratamiento (Figuras 4C y 4D), indicando que estas proteínas están asocia-

das dentro de la membrana. Es importante tener en cuenta que estos experimentos no cuestionan si esta asociación es directa o está mediada por otros componentes, tales como el citoesqueleto. Para controlar la posibilidad de que Delta forme placas de forma no específica en este experimento, cotransfectamos células con Notch y con la estructura artificial de neuroglía mencionada anteriormente (A. Bieber y C. Goodman, datos no publicados) y se formaron parches con antisueros anti-Notch. En este caso, no había una correlación evidente entre Notch y neuroglía.

### 6.2.5. Las interacciones con Delta no requieren el dominio intracelular de Notch

Además de un dominio extracelular grande que contiene repeticiones similares a EGF, Notch tiene un dominio intracelular (IC) medible de ~940 aminoácidos. El dominio IC incluye un sitio de fosforilación (Kidd y col., 1989, Genes Dev. 3, 1113-1129), un dominio de unión putativa a nucleótidos, una extensión de poliglutamina (Wharton y col., 1985, Cell 43, 567-581; Kidd y col., 1986, Mol. Cell. Biol. 6, 3094-3108) y secuencias homólogas al gen *cdc10* de levadura, que está implicado en el control del ciclo celular en levaduras (Breden y Nasmyth, 1987, Nature 329, 651-654). Una vez dado el tamaño y la complejidad estructural de este dominio, nos preguntamos si era necesario para las interacciones de Notch-Delta. Para ello usamos una variante de la estructura artificial de Notch, de la cual se han eliminado las secuencias codificadoras de ~835 aminoácidos del dominio IC, incluyendo todas las características estructurales mencionadas antes, (dejando 25 aminoácidos proximales a la membrana y un nuevo extremo carboxi-terminal de 59 aminoácidos; véase Procedimientos Experimentales y la Figura 1 para más detalles). Esta estructura artificial, denominada ECN1, se expresaba constitutivamente bajo el control del promotor normal de *Notch* en células transfectadas a un nivel inferior que el observado para las estructuras artificiales del promotor de metalotioneína, pero todavía era detectable fácilmente con inmunofluorescencia.

En los ensayos de agregación, las células que expresaban la estructura artificial ECN1, formaban consistentemente agregados con células Delta<sup>+</sup> (31% de células que expresaban ECN1 estaban en agregados en uno de tres experimentos; véase también la Figura 3I), pero no con ellas mismas (sólo 4% en agregados), lo mismo que cuando observamos células que expresaban Notch intacta. También observamos bandas intensas de tinción de ECN1 dentro de regiones de contacto con células Delta<sup>+</sup>, indicando de nuevo una localización de ECN1 dentro de regiones de contacto entre células. Para someter a ensayo las interacciones dentro de la membrana, repetimos los experimentos de formación conjunta de placas con antígenos de superficie, empleando células cotransfectadas con las estructuras artificiales ECN1 y *Delta*. Tal y como se observó para Notch intacta, encontramos que cuando en ECN1 se formaban placas empleando antisueros policlonales contra el dominio extracelular de Notch, ECN1 y Delta se localizaban conjuntamente en la superficie celular (Figuras 4E y 4F). Estos resultados demostraban que las interacciones observadas entre Notch y Delta dentro de la membrana no requieren la parte deletada del dominio IC de Notch y están mediadas probablemente por el dominio extracelular. Sin embargo, es posible que las secuencias de la transmembrana o del dominio IC restantes en ECN1 sean suficientes para intervenir en las interacciones dentro de una sola célula.

### 6.2.6. Notch Y Delta forman complejos intermoleculares solubles en detergente

Simultáneamente, tomamos los resultados anteriores para indicar interacciones moleculares entre Notch y Delta presentes dentro de la misma membrana y entre estas proteínas expresadas en células diferentes. Como ensayo adicional para tales interacciones, nos preguntamos si estas proteínas coprecipitarían desde extractos de células que expresan Notch y Delta, sin desnaturalizar con detergente. Si Notch y Delta forman un complejo intermolecular estable entre o dentro de las células, debería ser posible precipitar ambas proteínas desde extractos celulares, empleando antisueros específicos dirigidos contra una de estas proteínas. Realizamos este análisis inmunoprecipitando Delta con antisueros policlonales procedentes de material lisado con NP-40/deso-xicolato (véase Procedimientos Experimentales) de células cotransfectadas con las estructuras artificiales de Notch y Delta, a las que se había permitido agregarse durante una noche o de embriones de tipo silvestre de 0-24 h. No fuimos capaces de realizar los inmunoprecipitados inversos porque no era posible discernir sin ambigüedad una banda débil de Delta entre la señal de fondo de las bandas de Staph A. Es importante tener en cuenta que sometimos a ensayo este antisuero policlonal anti-Delta para estudiar la reactividad cruzada contra Notch en material lisado de células (Figura 5A, pista 1) y por inmunofluorescencia (por ejemplo, comparar las Figuras 3D y 3E) y no encontramos ninguna. Después de repetidos lavados para eliminar proteínas que se adherían inespecíficamente, sometimos a ensayo la coprecipitación de Notch empleando un anticuerpo monoclonal (MAb C17.9C6) contra Notch en transferencias tipo Western.

Como se muestra en la Figura 5, no detectamos coprecipitación de Notch en inmunoprecipitados con Delta procedentes de células cotransfectadas y embriones. Sin embargo, Notch coprecipitada parecía estar presente en cantidades mucho menores que Delta y por eso era difícil de detectar. Esta disparidad se debe lo más probable a la rotura de complejos Notch-Delta durante las etapas de lisis y de lavado del procedimiento. Sin embargo, también es posible que esta disparidad refleje una interacción no equimolar entre Notch y Delta o afinidades muy diferentes de los antisueros empleados para detectar estas proteínas. El hecho de que la inmunoprecipitación de Delta dé como resultado la coprecipitación de Notch, constituye una evidencia directa de que estas dos proteínas forman complejos intermoleculares estables en células S2 transfectadas y en células embrionarias.

### 6.3. Discusión

Hemos estudiado las interacciones entre los productos proteicos de dos de los loci neurógenos, *Notch* y *Delta*, para entender mejor sus funciones celulares. Empleando un ensayo de agregación *in vitro* que utiliza normalmente células S2 no adhesivas, mostramos que células que expresan Notch y Delta se adhieren específicamente entre sí. La

especificidad de esta interacción es evidente por la observación de que agregados de células Notch<sup>+</sup>-Delta<sup>+</sup> raramente contienen células sin expresión, aún cuando las células sin expresión componen la amplia mayoría de la población total celular en estos experimentos. Proponemos que esta agregación está mediada por la unión heterotípica entre los dominios extracelulares de Notch y Delta presentes en las superficies de células con expresión. De acuerdo con esta propuesta, encontramos que antisueros dirigidos contra el dominio extracelular de Notch, inhiben la agregación mediada por Notch-Delta, y que la variante de Notch ECN1, que carece de casi todo el dominio intracelular de Notch, puede mediar en la agregación con células que expresan Delta. También encontramos que células que expresan sólo Delta se agregan entre sí, aunque las que expresan sólo Notch no lo hacen. Estos resultados sugieren que Delta puede participar en una interacción homotípica cuando está presente en superficies celulares yuxtapuestas, pero que Notch no puede bajo nuestras condiciones de ensayo.

La propuesta de que Notch y Delta interaccionan en la superficie celular se apoya adicionalmente por tres líneas de evidencia. Primera, encontramos una intensa localización de ambas proteínas dentro de regiones de contacto entre células Notch<sup>+</sup> y Delta<sup>+</sup>, lo que implica que Notch y Delta interaccionan directamente, incluso cuando se expresan en células diferentes. Segunda, Notch y Delta se localizan conjuntamente en la superficie de células que expresan ambas proteínas, lo que sugiere que estas proteínas pueden interaccionar dentro de la membrana celular. Tercera, Notch y Delta se pueden coprecipitar desde extractos no desnaturizados con detergente no desnaturizante de células cultivadas que expresan ambas proteínas, así como desde extractos de células embrionarias. Simultáneamente, estos resultados apoyan fuertemente la hipótesis de que Notch y Delta pueden interaccionar heterotípicamente cuando se expresan en las superficies de las mismas células o de células diferentes.

La base subyacente de las interacciones genéticas observadas entre *Notch* y *Delta* y entre *Notch* y *mam* (Xu y col., 1990, Genes Dev. 4, 464-475) puede ser una interacción sensible a la dosis entre las proteínas codificadas por estos genes.

Dos líneas de evidencia sugieren que las proteínas Notch y Delta funcionan similarmente *in vitro* e *in vivo*. Primera, los análisis genéticos han indicado que la estequiometría de Notch y Delta es crucial para su función en el desarrollo. Nuestras observaciones de que ambas asociaciones Notch-Delta y Delta-Delta pueden tener lugar *in vitro*, implican que Notch y Delta pueden competir por la unión a Delta. Por tanto, interacciones genéticas sensibles a la dosis entre Notch y Delta pueden ser el resultado de interacciones por unión competitiva entre sus productos proteicos. Segunda, hemos sido capaces de detectar asociación Notch-Delta en material lisado de células cultivadas y en material lisado de embriones de *Drosophila*, empleando inmunoprecipitación. Resumiendo, estos análisis genéticos y bioquímicos sugieren que Notch y Delta se asocian *in vivo* de una forma similar a la que proponemos basándonos en nuestros ensayos de agregación.

Los análisis genéticos y moleculares de *Notch* también han mostrado la posibilidad de que puede haber interacciones entre proteínas Notch individuales (Portin, 1975, Genetics 81, 121-133; Kelley y col., 1987, Cell 51, 539-548; Artavanis-Tsakonas, 1988, Trends Genet. 4, 95-100). Realmente, Kidd y col., (1989, Genes Dev. 3, 1113-1129) han propuesto que esta proteína forma dímeros con puentes disulfuro, aunque este punto no se ha probado todavía rigurosamente. Con o sin la formación de enlaces covalentes, tales interacciones podrían tener lugar presumiblemente dentro de una sola célula o entre células. Sin embargo, nuestro descubrimiento de que células Notch<sup>+</sup> no se agregan homotípicamente, sugiere que las asociaciones Notch-Notch ocurren probablemente dentro de una sola célula y no entre células. Alternativamente, es posible que las interacciones homotípicas de Notch requieran productos génicos que no se expresan en células S2.

Las interacciones Notch-Delta indicadas por nuestros análisis están mediadas probablemente por los dominios extracelulares de estas proteínas. Los experimentos de agregación que emplean la estructura artificial ECN1, de la cual se ha eliminado casi todo el dominio intracelular de Notch o se ha alterado por mutagénesis *in vitro*, confirman esta conclusión. Experimentos adicionales que demuestran las asociaciones ECN1-Delta dentro de la membrana sobre la base de su capacidad de formar placas conjuntamente, indicaban que estas interacciones estaban mediadas también probablemente por los dominios extracelulares de Notch y Delta, aunque en este caso no podemos excluir una posible implicación del dominio transmembranal o de la parte restante del dominio intracelular de Notch. Estos resultados son especialmente interesantes de cara al hecho de que Notch y Delta tienen repeticiones similares a EGF dentro de sus dominios extracelulares (Wharton y col., 1985, Cell 43, 567-581; Kidd y col., 1986, Mol. Cell Biol. 6, 3094-3108; Vassin y col., 1987, EMBO J. 6, 3431-3440; Kopczynski y col., 1988, Genes Dev. 2, 1723-1735).

Una segunda conclusión de interés en relación con los dominios de EGF es la propuesta de que pueden servir como dominios de unión a Ca<sup>2+</sup> cuando contienen una secuencia de consenso que consta de los restos Asp, Asp/Asn, Asp/Asn y Tyr/Phe en posiciones conservadas dentro de las repeticiones similares a EGF (Rees y col., 1988, EMBO J. 7, 2053-2061; Handford y col., 1990, EMBO J. 9, 475-480). Comparaciones con una secuencia de consenso propuesta para la unión a Ca<sup>2+</sup>, han revelado que secuencias similares se encuentran dentro de la mayoría de las repeticiones similares a EGF de Notch (Rees y col., 1988, EMBO J. 7, 2053-2061) y dentro de las repeticiones similares a EGF de Delta (Kopczynski y col., 1988, Genes Dev. 2, 1723-1735). Además, análisis de secuencias de mutaciones de *Notch* han mostrado que ciertos alelos de Ax están asociados con cambios en aminoácidos dentro de este dominio putativo de unión a Ca<sup>2+</sup> (Kelley y col., 1987, Cell 51, 539-548; Hartley y col., 1987, EMBO J. 6, 3407-3417; Rees y col., 1988, EMBO J. 7, 2053-2061). Por ejemplo, la mutación Ax<sup>E2</sup>, que se correlaciona con un cambio de His a Tyr en la 29ª repetición similar a EGF, parece mover esta repetición hacia la secuencia de consenso para la unión a Ca<sup>2+</sup>. Recíprocamente, la mutación Ax<sup>9B2</sup> parece mover la 24ª repetición similar a EGF lejos de esta secuencia de consenso

como resultado de un cambio de Asp a Val. Por tanto, las interacciones genéticas entre los alelos *Ax* y las mutaciones de *Delta* (Xu y col., 1990, Genes Dev. 4, 464-475) muestran la posibilidad de que los iones  $Ca^{2+}$  tengan un papel en las interacciones Notch-Delta. Nuestro descubrimiento de que  $Ca^{2+}$  exógeno es necesario para la agregación mediada por Notch-Delta de células S2 transfectadas, apoya esta afirmación.

Tal y como hemos argumentado (Johansen y col., 1989, J. Cell Biol. 109, 2427-2440; Alton y col., 1989, Dev. Genet. 10, 261-272), sobre la base de análisis moleculares y genéticos previos, no se podría pronosticar con ninguna certeza la función celular de Notch o Delta más allá de su implicación en las interacciones célula-célula. Sin embargo, dados los resultados presentados en esta memoria, ahora parece razonable sugerir que Notch y Delta pueden funcionar *in vivo*, para mediar en interacciones adhesivas entre células. Al mismo tiempo, es totalmente posible que las interacciones Notch-Delta observadas no puedan reflejar sólo una función adhesiva y puedan reflejar además interacciones de la unión de receptor-ligando que tienen lugar *in vivo*. Realmente, la presencia de un complejo estructural de 1000 aminoácidos del dominio intracelular dentro de Notch, puede ser más consistente con un papel en la transducción de señales que con interacciones puramente adhesivas. Dado que Notch puede tener una función adhesiva en concierto con Delta, la expresión axonal de Notch puede tener algún papel en la guía de los axones.

### 7. Las repeticiones EGF 11 y 12 de Notch son necesarias y suficientes para la agregación mediada por Notch-Delta

En este estudio, utilizamos el mismo ensayo de agregación que el descrito en la Sección 6, junto con mutantes por delección de Notch para identificar regiones dentro del dominio extracelular de Notch necesarias para las interacciones con Delta. Presentamos la evidencia de que las repeticiones EGF de Notch están directamente implicadas en esta interacción y que sólo dos de las 36 repeticiones EGF parecen ser necesarias. Demostramos que estas dos repeticiones EGF son suficientes para la unión con Delta y que la dependencia de calcio de la agregación mediada por Notch-Delta, también está asociada con estas dos repeticiones. Finalmente, las dos repeticiones EGF correspondientes con las homólogas de *Xenopus*, también median en la agregación con Delta, implicando que no sólo la estructura de Notch se ha conservado evolutivamente, pero también su función. Estos resultados sugieren que el dominio extracelular de Notch es sorprendentemente modular y podría unirse potencialmente a una variedad de proteínas además de a Delta.

#### 7.1. Procedimientos experimentales

##### 7.1.1 Estructuras artificiales de expresión

Las estructuras artificiales descritas se obtienen todas a partir de la estructura artificial de expresión nº 1 de Notch de longitud completa, pMtNMg (véase Sección 6, más arriba). Todas las ligaciones se realizaron usando fragmentos de ADN cortados de geles de agarosa de bajo punto de fusión (Sea Plaque, FMC BioProducts). El fragmento EcoRI-XhoI de 6 kb procedente de pMtNMg que contenía el dominio extracelular completo de Notch, se ligó en los sitios EcoRI-XhoI del vector Bluescript (Stratagene), y se denominó RI/XBS. Todas las delecciones y las inserciones posteriores de las repeticiones EGF se realizaron en este subclon. La secuencia de *Notch* que contenía el fragmento EcoRI/XhoI de estos derivados de RI/XBS se mezcló a continuación con el fragmento XhoI-XbaI de 5,5 kb procedente de pMtNMg, que contenía el dominio intracelular y las secuencias 3' necesarias para la poliadenilación y después se insertó en el sitio EcoRI-XbaI de pRMHa-3 (Bunch y col., 1988, Nucl. Acids Res. 16, 1043-1061) en una ligación de tres trozos. Todos los números siguientes se refieren a las coordenadas de nucleótidos de la secuencia de *Notch* según Wharton y col. (1985, Cell 43, 567-581).

Para la estructura artificial nº 2 DSph, se digirió RI/XBS completamente con SphI y después se recircularizó, dando como resultado una delección en marco de 3,5 kb desde SphI(996) hasta SphI(4545).

Para la estructura artificial nº 3 ΔCla, RI/XBS se digirió completamente con ClaI y después se ligó de nuevo, produciendo una delección en marco de 2,7 kb desde ClaI(1668) hasta ClaI(4407). La unión por ligación se comprobó mediante secuenciación de la doble cadena (como describen Xu y col., 1990, Genes Dev. 4, 464-475) empleando el equipo de reactivos de Sequenase (U.S. Biochemical Corp., Cleveland). Encontramos que aunque existe el sitio ClaI en la posición 4566 de acuerdo con la secuencia, no era reconocido bajo nuestras condiciones por la enzima de restricción ClaI.

Para las estructuras artificiales nº 4-12, RI/XBS se digirió parcialmente con ClaI y después se ligó de nuevo para producir todas las combinaciones posibles de delecciones en marco: en la estructura artificial nº 4 ΔEFG7-17, se eliminaba la secuencia entre ClaI(1668) y ClaI(2820); en la estructura artificial nº 5 ΔEGF9-26 se eliminaba la secuencia entre ClaI(1905) y ClaI(3855); en la estructura artificial nº 6 ΔEGF17-31 se eliminaba la secuencia entre ClaI(2820) y ClaI(4407); en la estructura artificial nº 7 ΔEGF7-9 se eliminaba la secuencia entre ClaI(1668) y ClaI(1905); en la estructura artificial nº 8 ΔEGF9-17 se eliminaba la secuencia entre ClaI(1905) y ClaI(2820); en la estructura artificial nº 9 ΔEGF17-26 se eliminaba la secuencia entre ClaI(2820) y ClaI(3855); en la estructura artificial nº 10 ΔEGF 26-30 se eliminaba la secuencia entre ClaI(3855) y ClaI(4407); en la estructura artificial nº 11 ΔEGF9-30 se eliminaba la secuencia entre ClaI(1905) y ClaI(4407); en la estructura artificial nº 12 ΔEGF 7-26 se eliminaba la secuencia entre ClaI(1668) y ClaI(3855).

Para las estructuras artificiales nº 13 ΔCla+EGF9-17 y nº 14 ΔCla+EGF17-26, el fragmento de ~0,9 kb entre ClaI(1905) y ClaI(2820), y el fragmento de ~1,0 kb entre ClaI(2820) y ClaI(3855), respectivamente, se insertaron en el único sitio ClaI de la estructura artificial nº 3 ΔCla.

## ES 2 181 674 T5

Para la estructura artificial nº 16 split, el fragmento KpnI/XbaI de 11 kb de pMtNMg se reemplazó por el fragmento correspondiente KpnI/XbaI procedente de una estructura artificial de un minigen de Notch, que contenía la mutación split en la repetición EGF 14.

5 Para las estructuras artificiales nº 17-25, se diseñaron cebadores sintéticos para la reacción en cadena de la polimerasa (PCR) para multiplicar extensiones de repeticiones EGF mientras que se rompían las repeticiones EGF en los extremos del trozo multiplicado en el mismo lugar que los sitios comunes de ClaI, justo después de la tercera cisteína de la repetición (véase Figura 7). Los productos de la PCR se purificaron en gel de forma convencional y se ligaron en el sitio ClaI de la estructura artificial nº 3 ΔClaI, a la que se había puesto extremos romos, rellenando con el fragmento Klenow de la polimerasa I de ADN (Maniatis y col., 1990, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York). La orientación correcta de las inserciones se determinó con PCR, usando un cebador en la hebra transcrita dentro de la inserción junto con un cebador en la hebra complementaria en la repetición EGF 35. Todos los cebadores eran 20-meros, y se denominaron con el número del nucleótido de su extremo 5', según las coordenadas de nucleótidos de la secuencia de *Notch* en Wharton y col. (1985, Cell 43, 567-581), y S se refiere a un cebador de la hebra transcrita mientras que A se refiere a un cebador de la hebra complementaria. La estructura artificial nº 16 ΔCla+EGF(9-13) empleaba los cebadores S1917 y A2367. La estructura artificial nº 17 ΔCla+EGF(11-15) empleaba los cebadores S2141 y A2591. La estructura artificial nº 18 ΔCla+EGF(13-17) empleaba los cebadores S2375 y A2819. La estructura artificial nº 19 ΔCla+EGF(10-13) empleaba los cebadores S2018 y A2367. La estructura artificial nº 20 ΔCla+EGF(11-13) empleaba los cebadores S2141 y A2367. La estructura artificial nº 21 ΔCla+EGF(10-12) empleaba los cebadores S2018 y A2015. La estructura artificial nº 22 ΔCla+EGF(10-11) empleaba los cebadores S2018 y A2322. La estructura artificial nº 23 ΔCla+EGF(10-12) empleaba los cebadores S2018 y A2322. La estructura artificial nº 24 ΔCla+EGF(11-12) empleaba los cebadores S2018 y A2322.

Para la estructura artificial nº 25 ΔEGF, la estructura artificial R1/XBS se digirió completamente con SphI(996) y se digirió parcialmente con BamHI(5135). Los extremos resultantes incompatibles se unieron utilizando un engarzador sintético diseñado para crear un único sitio ClaI. Esto producía una deleción en marco que eliminaba las 36 repeticiones EGF con excepción de la primera mitad de la repetición 1. Para las estructuras artificiales nº 26-29, los fragmentos EGF se insertaron en este sitio ClaI tal y como se ha descrito previamente para las estructuras artificiales correspondientes nº 13, 16, 19 y 23.

30 Para la estructura artificial nº 30 ΔECN, la estructura artificial R1/XBS se digirió completamente con BglI, EcoRI y XhoI. El fragmento EcoRI-BglI de ~0,2 kb (722-948) y los fragmentos BglI-XhoI de ~0,7 kb (5873-6627) se ligaron con el vector Bluescript cortado con EcoRI-XhoI y un engarzador sintético diseñado para crear un único sitio ClaI, dando como resultado una deleción en marco desde BglI(941) hasta BglI(5873) que eliminaba las 36 repeticiones EGF, excepto el primer tercio de la repetición 1 así como las 3 repeticiones *Notch/lin-12*. Para las estructuras artificiales nº 31 y 32, los fragmentos EGF se insertaron en el único sitio ClaI, tal y como se ha descrito previamente para las estructuras artificiales nº 19 y 23.

40 Para las estructuras artificiales nº 33 y 34, cebadores de PCR S1508 y A1859, basados en la secuencia de Notch de *Xenopus* (Coffman y col., 1990, Science 249, 1438-1441; los números se refieren a las coordenadas de nucleótidos empleadas en este artículo), se usaron para multiplicar las repeticiones EGF 11 y 12, fuera de una genoteca de ADNc de *Xenopus* en el estado 17 (la genoteca fue preparada por D. Melton y amablemente suministrada por M. Danilchek). El fragmento se ligó en la estructura artificial nº 3 ΔCla y se secuenció.

### 45 7.1.2. Cultivo celular y transfección

La línea celular S2 de *Drosophila* se hizo crecer y se transfectó tal y como se ha descrito en la Sección 6, más arriba. La línea celular S2 transformada de forma estable y que expresa Delta, L-49-6-7 (establecida amablemente por L. Cherbas) se hizo crecer en medio M3 (preparado por Hazleton Co.) suplementado con 11% de suero de ternera fetal (FCS) inactivado con calor (Hyclone), 100 U/ml de penicilina-100 μg/ml de estreptomina-0,25 μg/ml de fungizona (Hazleton), metotrexato  $2 \times 10^{-7}$  M, hipoxantina 0,1 mM y timidina 0,016 mM.

### 55 7.1.3. Ensayos de agregación e inmunofluorescencia

Los ensayos de agregación y los experimentos de dependencia de Ca<sup>++</sup> fueron como se ha descrito más arriba, Sección 6. las células se tiñeron con el anticuerpo monoclonal anti-Notch 9C6.C17 y con antisueros policlonales de rata anti-Delta (detalles descritos en la Sección 6, más arriba). La expresión en superficie de las estructuras artificiales de Notch en células impermeabilizadas se sometió a ensayo empleando antisueros policlonales de rata, producidos contra el fragmento BstYI de 0,8 kb (aminoácidos 237-501; Wharton y col., 1985, Cell 43, 567-581) procedente del dominio extracelular de Notch. Las células se visionaron bajo epifluorescencia en un microscopio Orthoplan 2 de Leitz.

## 65 7.2. Resultados

### 7.2.1. Las repeticiones EGF 11 y 12 de Notch son necesarias para la agregación mediada por Notch-Delta

Realizamos un análisis extenso de la deleción del dominio extracelular de la proteína Notch, que hemos mostrado (más arriba, Sección 6) que está implicada en las interacciones Notch-Delta, para identificar el dominio preciso de Notch que media en estas interacciones. Sometimos a ensayo la capacidad de las células transfectadas con distintas es-

estructuras artificiales de delección, de interactuar con Delta, empleando el ensayo de agregación descrito en la Sección 6. Brevemente, las estructuras artificiales de delección de *Notch* se transfectaron de forma transitoria en células S2 de *Drosophila*, se indujeron con CuSO<sub>4</sub>, y luego se agregaron durante una noche a temperatura ambiente con una pequeña cantidad de células de la línea celular que expresa Delta, transformadas de forma estable, L49-6-7 (Cherbas), consiguiendo una población compuesta típicamente por ~1% de células que expresan Notch y ~5% de células que expresan Delta, y el resto de las células no expresan ninguna proteína. Para someter a ensayo el grado de agregación, se tiñeron las células con antisueros específicos para cada producto génico y se examinaron con microscopía inmunofluorescente (véase Procedimientos Experimentales para más detalles). Los agregados se definieron como agrupaciones de cuatro o más células que contenían células que expresan Notch y Delta, y los valores mostrados en la Figura 6 representan el porcentaje de todas las células que expresan Notch encontradas en dichas agrupaciones. Todos los números reflejan el resultado promedio de al menos dos experimentos de transfección distintos, en los que al menos se marcaron 100 unidades celulares que expresan Notch (tanto células aisladas como agrupaciones).

Dibujos esquemáticos de las estructuras artificiales sometidas a ensayo y los resultados de los experimentos de agregación se muestran en la Figura 6 (véase Procedimientos Experimentales para más detalles). Todas las estructuras artificiales de expresión eran derivadas de la estructura artificial de expresión de Notch de longitud completa, n° 1 pMtNMg (descrita en la Sección 6, más arriba).

Las estructuras artificiales iniciales (n° 2 DSph y n° 3 ΔCla) delecionaban partes grandes de las repeticiones EGF. Su incapacidad para favorecer la agregación de Notch-Delta sugería que las repeticiones EGF de Notch estaban implicadas en la interacción con Delta. Nos aprovechamos de una serie de seis sitios de restricción de ClaI en marco, para diseccionar adicionalmente la región entre las repeticiones EGF 7 y 30. Debido a la homología de secuencias entre las repeticiones, cinco de los sitios ClaI estaban en el mismo lugar relativo dentro de la repetición EGF, justo después de la tercera cisteína, mientras que el sexto sitio estaba justo antes de la tercera cisteína de la repetición EGF 31 (Figura 7). Por tanto, al realizar una digestión parcial con ClaI y después ligar de nuevo, obtuvimos delecciones que no sólo conservaban el marco de lectura abierto de la proteína Notch, pero que además mantenían frecuentemente la integridad estructural y conservaban los espaciadores, al menos teóricamente, de tres enlaces disulfuro en las repeticiones EGF quiméricas producidas por la religación (Figura 6, estructuras artificiales n° 4-14). Desgraciadamente, el sitio ClaI más 3' era resistente a la digestión, mientras que el siguiente sitio ClaI más 3' se rompía entre las repeticiones EGF 30 y 31. Por tanto, cuando se insertaban de nuevo varios fragmentos de digestión con ClaI en el esqueleto de la digestión completa con ClaI (estructura artificial n° 3 ΔCla), la estructura global de las repeticiones EGF se interrumpía aparentemente en la unión 3'.

Varios puntos sobre esta serie de estructuras artificiales son dignos de mención. Primero, la eliminación del fragmento de restricción ClaI que rompe las repeticiones EGF 9 y 17 (estructura artificial n° 8 ΔEGF9-17) suprime la agregación con Delta, mientras que la reinserción de este trozo en la estructura artificial n° 3 ΔCla, que carece de las repeticiones EGF 7-30, restaura la agregación hasta aproximadamente los niveles del tipo silvestre (estructura artificial n° 13 ΔCla+EGF9-17), sugiriendo que las repeticiones EGF 9 a 17 contienen secuencias importantes para la unión a Delta. Segundo, todas las estructuras artificiales en esta serie (n° 4-14) eran compatibles con el sitio de unión cartografiado en las repeticiones EGF 9 a 17. Las estructuras artificiales de expresión que contienen estas repeticiones (n° 6, 7, 9, 10, 13) favorecen las interacciones Notch-Delta, mientras que las estructuras artificiales que carecen de estas repeticiones (n° 4, 5, 8, 11, 12, 14) no las favorecen. Para confirmar que la incapacidad de agregación con células Delta no se debía simplemente a un fallo de la proteína Notch mutagenizada para alcanzar la superficie celular, pero que realmente reflejaba la delección del sitio de unión necesario, sometimos a ensayo la expresión en la superficie celular de todas las estructuras artificiales mediante tinción inmunofluorescente de células vivas transfectadas con anticuerpos específicos del dominio extracelular de Notch. Todas las estructuras artificiales que fallaban en la mediación de las interacciones Notch-Delta, producían una proteína que parecía que se expresaba normalmente en la superficie celular. Tercero, aunque el ensayo de agregación no es cuantitativo, dos estructuras artificiales que contenían las repeticiones EGF 9-17, n° 9 ΔEGF17-26 o la más visible n° 10 ΔEGF26-30, se agregaban a un nivel aparentemente inferior. Las células transfectadas con las estructuras artificiales n° 9 ΔEGF17-26 y 10 ΔEGF26-30, mostraban una tinción de superficie considerablemente menor que la normal, aunque las células fijadas y permeabilizadas reaccionaban con el mismo anticuerpo teñido normalmente, indicando que no se habían delecionado simplemente los epítopos reconocidos por los antisueros. Comparando el porcentaje de células transfectadas en las poblaciones de células permeabilizadas o vivas, encontramos que aproximadamente 50% de las células transfectadas con la estructura artificial n° 9 ΔEGF17-26 y 10% con la estructura artificial n° 10 ΔEGF26-30 producían una proteína detectable en la superficie celular. Por tanto, estas dos estructuras artificiales producían proteínas que fallaban frecuentemente en alcanzar la superficie celular, quizás debido a un plegamiento defectuoso, reduciendo de este modo, pero no suprimiendo, la capacidad de agregación de las células transfectadas con células que expresan Delta.

Habiendo cartografiado el sitio de unión para las repeticiones EGF 9 a 17, comprobamos si se había cartografiado en esta región alguna mutación de Notch cuya lesión molecular estuviera determinada. La única de tales mutaciones era *split*, un alelo semidominante de *Notch* que se correlaciona con una mutación puntual en la repetición EGF 14 (Hartley y col., 1987, EMBO J. 6, 3407-3417; Kelley y col., 1987, Mol. Cell. Biol. 6, 3094-3108). De hecho, un rastreo genético en busca de modificadores secundarios del sitio de *split*, reveló varios alelos de *Delta*, sugiriendo una especial relación entre el alelo *split* de *Notch*, y *Delta* (Brand y Campus-Ortega, 1990, Roux's Arch. Dev. Biol. 198 (5), 275-285). Para estudiar los posibles efectos de la mutación *split* sobre la agregación mediada por Notch-Delta, un fragmento de 11 kb que contenía la mutación sin sentido asociada con *split* se clonó en la estructura artificial de expresión de *Notch* (n° 15 *split*). Sin embargo, la agregación con células que expresan Delta no estaba afectada en

esta estructura artificial, lo que sugiere, como se confirmó con las siguientes estructuras artificiales, que la repetición EGF 14 de Notch no estaba implicada en las interacciones con Delta ejemplificadas con nuestro ensayo de cultivo de tejidos.

5 Por tanto, para cartografiar adicionalmente el dominio de unión a Delta dentro de las repeticiones EGF 9-17, empleamos cebadores específicos de oligonucleótidos y la técnica de PCR para generar varios subfragmentos de esta región. Para ser compatible con las estructuras artificiales n° 4-14 que producían proteínas que eran capaces de interactuar con Delta, diseñamos los cebadores para cortar y empalmar las repeticiones EGF justo después de la tercera cisteína, en el mismo lugar que el sitio común de ClaI (Figura 7). Los productos resultantes de la PCR se  
10 ligaron en el sitio ClaI de la estructura artificial n° 3 ΔCla. Se produjeron tres estructuras artificiales solapantes, n° 16, 17 y 18, de las cuales, sólo una la n° 16 ΔCla+EGF9-13, permitía la agregación con células Delta cuando se transfectaba en las células S2. La estructura artificial n° 19 ΔCla+EGF(10-13), que carece de la repetición EGF 9, definía adicionalmente las repeticiones EGF 10-13 como la región necesaria para las interacciones Notch-Delta.

15 Las estructuras artificiales n° 20-24 representaban intentos de romper este dominio, empleando adicionalmente incluso la misma estrategia de PCR (véase Figura 7). Nos preguntamos primero si ambas repeticiones EGF 11 y 12 eran necesarias, y segundo, si las secuencias flanqueantes de las repeticiones EGF 10 y 13 estaban implicadas directamente en la unión a Delta. Las estructuras artificiales n° 20 ΔCla+EGF(11-13), en la que la repetición EGF 12 es la única repetición completa añadida, y la n° 21 ΔCla+EGF(10-12), en la que la repetición EGF 11 es la única repetición  
20 completa añadida, fallaban en la mediación de la agregación, sugiriendo que la presencia de la repetición EGF 11 o 12 aislada no era suficiente para las interacciones Notch-Delta. Sin embargo, puesto que la unión 3' de la ligación de estas estructuras artificiales interrumpía la estructura global de las repeticiones EGF, era posible que una pequeña zona “tampón” fuera necesaria para permitir que la repetición crucial funcionara normalmente. Así, por ejemplo, en la estructura artificial n° 19 ΔCla+EGF(10-13), la repetición EGF 12 podría no estar implicada directamente en  
25 la unión a Delta, pero en su lugar, podría contribuir con la cantidad mínima necesaria de secuencia tampón para proteger la estructura de la repetición EGF 11, permitiendo de este modo las interacciones con Delta. Las estructuras artificiales n° 22-24 se dedicaban a este fin. Diseñamos cebadores de PCR que rompían en el extremo de la repetición EGF y por tanto eran menos idóneos para romper la formación disulfuro de EGF en la unión de ligación 3'. Las estructuras artificiales n° 22 ΔCla+EGF(10-11), que no mediaba en la agregación, y la n° 23 ΔCla+EGF(10-12), que  
30 si mediaba, sugerían de nuevo que ambas repeticiones 11 y 12 eran necesarias, mientras que la secuencia flanqueante de la repetición 13 claramente no lo era. Finalmente, la estructura artificial n° 24 ΔCla+EGF(11-12), aunque ahora rompía potencial y estructuralmente en la unión 5', demostraba convincentemente que las secuencias procedentes de la repetición EGF 10 no son cruciales. Por tanto, basándose en los datos totalmente consecuentes de las 24 estructuras artificiales, proponemos que las repeticiones EGF 11 y 12 de Notch juntas, definen la unidad funcional más pequeña  
35 obtenible a partir de este análisis que contiene los sitios necesarios para unirse a Delta en células S2 transfectadas.

### 7.2.2. Las repeticiones EGF 11 y 12 de Notch son suficientes para la agregación mediada por Notch-Delta

40 La delección extensa de ClaI en la que se insertaron fragmentos de PCR (n° 3 ΔCla), retiene aproximadamente 1/3 de las 36 repeticiones EGF originales, así como las tres repeticiones *Notch/lin-12*. Aunque estas no son claramente suficientes para favorecer la agregación, es posible que formen un esqueleto necesario dentro del cual repeticiones EGF específicas puedan interactuar con Delta. Para someter a ensayo si sólo unas pocas repeticiones EGF eran de hecho suficientes para favorecer la agregación, diseñamos dos estructuras artificiales, la n° 25 ΔEGF que delecionaba  
45 las 36 repeticiones EGF excepto los dos tercios primeros de la repetición EGF 1, y la n° 30 ΔECN que delecionaba la parte extracelular completa de Notch, excepto el primer tercio de la repetición 1 y ~35 aminoácidos justo antes del dominio transmembranal. Los fragmentos que habían mediado en la agregación Notch-Delta en el fondo de la estructura artificial n° 3 ΔCla, cuando se insertaba en la estructura artificial n° 25 ΔEGF, eran capaces de nuevo de favorecer las interacciones con Delta (estructuras artificiales n° 26-30). Estructuras artificiales análogas (n° 31, 32)  
50 en las que las repeticiones *Notch/lin-12* también estaban ausentes, también mediaban con éxito en la agregación de Notch-Delta. Por tanto, las repeticiones EGF 11 y 12 parecen funcionar como unidades modulares independientes que son suficientes para mediar en las interacciones entre Notch-Delta en células S2, incluso en ausencia de la mayoría del dominio extracelular de Notch.

55 *7.2.3. Las repeticiones EGF 11 y 12 de Notch mantienen la dependencia de calcio de la agregación mediada por Notch-Delta*

Tal y como se describe en la Sección 6, más arriba, (Fehon y col., 1990, Cell 61, 523-534), mostramos que la  
60 agregación de células S2, mediada por Notch-Delta, es dependiente de calcio. Por tanto examinamos la capacidad de las células que expresan ciertas estructuras artificiales con delecciones, de agregarse con células que expresan Delta, en presencia o en ausencia de iones Ca<sup>++</sup>. Sometimos a ensayo las estructuras artificiales n° 1 pMtNMg como testigo, y las n° 13, 16, 19, 23, 24, 26, 27 y 28, y encontramos que células mezcladas en medio que contenía Ca<sup>++</sup>, a 4°C, formaban fácilmente agregados, mientras que células mezcladas en medio exento de Ca<sup>++</sup> que contenía EGTA, fallaban en la  
65 agregación (Tabla III).

TABLA III

Efecto de  $Ca^{++}$  exógeno sobre la agregación<sup>a</sup> de Notch-Delta

	<u>Sin iones <math>Ca^{++}</math></u>	<u>Con iones <math>Ca^{++}</math></u>
5		
10		
15		
20		

<sup>a</sup> Datos presentados como porcentaje de células que expresan Notch encontradas en agregados (como en la Figura 6).

Claramente, la dependencia de calcio de la interacción se ha conservado incluso en la estructura artificial más pequeña, de acuerdo con la idea de que las estructuras artificiales mínimas que contienen las repeticiones EGF 11 y 12, se unen a Delta de una forma similar a la de Notch de longitud completa. Este resultado también es interesante desde el punto de vista de estudios recientes que sugieren que repeticiones similares a EGF con una secuencia de consenso particular, pueden actuar como dominios de unión a  $Ca^{++}$  (Morita y col., 1984, J. Biol. Chem. 259, 5698-5704; Sugo y col., 1984, J. Biol. Chem. 259, 5705-5710; Rees y col., 1988, EMBO J. 7, 2053-2061; Handford y col., 1990, EMBO J. 9, 475-480). Sobre la mitad de las repeticiones EGF en Notch, incluyendo las repeticiones 11 y 12, se ajustan a este consenso, confirmando adicionalmente el argumento de que las repeticiones EGF 11 y 12 son responsables de favorecer las interacciones entre Notch-Delta.

#### 7.2.4. La función de unión a Delta de las repeticiones EGF 11 y 12 de Notch está conservada en el homólogo de Notch de *Xenopus*

Habiendo cartografiado el sitio de unión a Delta de las repeticiones EGF 11 y 12 de Notch, estábamos interesados en saber si esta función estaba conservada en el homólogo de Notch que se había identificado en *Xenopus* (Coffman y col., 1990, Science 249, 1438-1441). Esta proteína muestra una notable similitud con Notch de *Drosophila* en toda la estructura y la organización. Por ejemplo, dentro de la región de las repeticiones EGF, tanto el número como la organización lineal de las repeticiones están conservados, sugiriendo una posible conservación también funcional. Para someter esto a ensayo, preparamos cebadores con PCR en base a la secuencia de Notch de *Xenopus* (Coffman y col., 1990, Science 249, 1438-1441) y empleamos éstos para obtener un fragmento de ~350 pb a partir de una genoteca de ADNc de *Xenopus* en el estado 17, que incluye las repeticiones EGF 11 y 12 flanqueadas por la mitad de las repeticiones 10 y 13 en ambos lados. Este fragmento se clonó en la estructura artificial n° 3 ΔCla y se sometieron a ensayo tres clones independientes para estudiar la capacidad de interactuar con Delta en el ensayo de agregación de cultivo celular. Dos de los clones, n° 33a & b ΔCla+XEGF(10-13), cuando se transfectaron en células S2, eran capaces de mediar en las interacciones Notch-Delta a un nivel aproximadamente equivalente al de la estructura artificial de Notch análoga de *Drosophila*, n° 19 ΔCla+EGF(10-13), y de nuevo en forma dependiente de calcio (Tabla III). Sin embargo, el tercer clon n° 33c ΔCla+XEGF(10-13), fallaba en la mediación de las interacciones Notch-Delta aunque la proteína se expresaba normalmente en la superficie celular, como se estimaba por la tinción de células vivas impermeabilizadas. La comparación de secuencias del producto de la PCR de *Xenopus* en las estructuras artificiales n° 33a y 33c, revelaba una mutación sin sentido originada en un cambio de leucina a prolina (aminoácido n° 453, Coffman y col., 1990, Science 249, 1438-1441) en la repetición EGF 11 de la estructura artificial n° 33c. Aunque este resto no está conservado entre la Notch de *Drosophila* y de *Xenopus* (Figura 8), la introducción de un resto de prolina podría romper fácilmente la estructura de la repetición EGF, y por tanto evitar interacciones adecuadas con Delta.

La comparación de la secuencia de aminoácidos de las repeticiones EGF 11 y 12 de Notch de *Drosophila* y de *Xenopus*, revela un alto grado de identidad de aminoácidos, incluyendo la secuencia de consenso de unión a calcio (Figura 8, SEQ ID NO:1 y NO:2). Sin embargo, el nivel de homología no es notablemente diferente del compartido entre la mayoría de las otras repeticiones EGF, que muestran globalmente aproximadamente 50% de identidad a nivel de aminoácidos. La correspondencia de uno a uno entre las repeticiones EGF individuales sugiere que quizá pueden comprender también unidades funcionales conservadas. Las interacciones con Delta, de nuevo en una forma dependiente del ión calcio.

7.3. *Discusión*

Hemos continuado nuestro estudio de las interacciones entre los productos proteicos de los genes *Notch* y *Delta*, empleando el ensayo de agregación de células S2 *in vitro*, descrito en la Sección 6, más arriba. Sobre la base del análisis de una extensa delección del dominio extracelular de Notch, mostramos que las regiones de Notch que contienen repeticiones homólogas a EGF 11 y 12 son ambas necesarias y suficientes para la agregación mediada por Notch-Delta, y que esta capacidad de unión a Delta se ha conservado en las dos mismas repeticiones EGF de Notch de *Xenopus*. Nuestro hallazgo de que la agregación se cartografiaba en las repeticiones EGF 11 y 12, demuestra que las repeticiones EGF de Notch también funcionan como dominios de unión a proteínas específicos.

Estudios recientes han demostrado que los dominios de EGF que contienen una secuencia de consenso específica se pueden unir a iones  $Ca^{++}$  (Morita y col., 1984, J. Biol. Chem. 259, 5698-5704; Sugo y col., 1984, J. Biol. Chem. 259, 5705-5710; Rees y col., 1988, EMBO J. 7, 2053-2061; Handford y col., 1990, EMBO J. 9, 475-480). De hecho, aproximadamente la mitad de las repeticiones EGF en Notch, que incluyen las repeticiones 11 y 12, están de acuerdo con este consenso. Hemos mostrado que es necesario  $Ca^{++}$  exógeno para la agregación mediada por Notch-Delta de células S2 transfectadas (véase Sección 6; Fehon y col., 1990, Cell 61, 523-534). Sometimosa ensayo una subpoblación de nuestras estructuras artificiales de delección y encontramos que las repeticiones EGF 11 y 12 aisladas (n° 32 $\Delta$ ECN+EGF(11-12)) eran suficientes para mantener la dependencia de  $Ca^{++}$  de las interacciones Notch-Delta.

Una cantidad de estudios han sugerido que las interacciones genéticas entre *Notch* y *Delta* pueden reflejar una interacción sensible a la dosis entre sus productos proteicos. Estudios genéticos han indicado que las dosificaciones génicas relativas de *Notch* y *Delta* son cruciales para el desarrollo normal. Por ejemplo, Xu y col., (1990, Genes Dev. 4, 464-475) encontraron que mutaciones completas en *Delta*, podían suprimir interacciones letales entre combinaciones heterocigotas de alelos *Abruptex* (*Ax*), una clase de mutaciones de *Notch* que se correlacionan con mutaciones sin sentido dentro de las repeticiones EGF (Hartley y col., 1987, EMBO J. 6, 3407-3417; Kelley y col., 1987, Mol. Cell Biol. 6, 3094-3108). Las interacciones *in vitro* que hemos descrito en las que observamos asociaciones Notch-Delta y Delta-Delta (véase Sección 6) implican que una interacción competitiva entre Notch y Delta para unirse a Delta, puede reflejar la base que sostiene las interacciones genéticas observadas. Además, fuimos capaces de coimmunoprecipitar Notch y Delta a partir de cultivo de tejidos y de extractos celulares embrionarios (véase Sección 6), indicando una posible asociación *in vivo* de las dos proteínas. Además, análisis *in situ* del ARNm de los patrones de expresión de *Notch* y *Delta* en embriones, sugieren que la expresión de los dos es solapante pero no idéntica (Kopczynski y Muskavitch, 1989, Development 107, 623-636; Hartley y col., 1987, EMBO J. 6, 3407-3417). Análisis detallados de anticuerpos de la expresión de la proteína Notch durante el desarrollo, han revelado recientemente que la expresión de Notch está más restringida a niveles de tejido y subcelulares que lo que habían indicado estudios previos (Johansen y col., 1989, J. Cell Biol. 109, 2427-2440; Kidd y col., 1989, Genes Dev. 3, 1113-1129).

Nuestro resultado de que las dos mismas repeticiones EGF procedentes de Notch homóloga de *Xenopus* eran también capaces de mediar en interacciones con Delta en células en cultivo de tejidos indica firmemente que una función similar se debe haber conservado *in vivo*. Aunque estas dos repeticiones EGF son suficientes *in vitro*, por supuesto que es posible que *in vivo* sea necesario algo más que la molécula de Notch, para facilitar las interacciones Notch-Delta. De hecho, estábamos algo sorprendidos por dos razones al encontrar que el sitio de unión a Delta no se cartografiaba en las repeticiones EGF en donde se había mostrado que se encontraban varias de las mutaciones *Ax*, primero, debido al rastreo genético (Xu y col., 1990, Genes Dev. 4, 464-475) que demostraba interacciones entre alelos *Ax* y mutaciones *Delta*, y segundo, porque análisis de secuencias han mostrado que ciertos alelos *Ax* están asociados con cambios aislados de aminoácidos dentro del consenso putativo de unión a  $Ca^{++}$  de las repeticiones EGF. Por ejemplo, la mutación AX<sup>E2</sup> cambia la repetición EGF 29 hacia la secuencia de consenso de unión a  $Ca^{++}$ , mientras que la mutación AX<sup>9B2</sup> mueve la repetición EGF 24 fuera del consenso. Es posible que estas regiones *in vivo* de la proteína *Notch* puedan estar implicadas en interacciones, con Delta y/o con otras proteínas, pero que no se pueden ejemplificar correctamente con nuestro ensayo de cultivo de células.

Nuestro cartografiado *in vitro* del dominio de unión a Delta para las repeticiones EGF 11 y 12 de Notch representa la primera asignación de función a un dominio estructural de Notch. De hecho, las distintas estructuras artificiales de delección sugieren que estas dos repeticiones EGF funcionan como una unidad modular, independientemente del contexto inmediato en el que están situadas. Por tanto, ni las 34 repeticiones EGF restantes ni las tres repeticiones *Notch/lin-12* parecen necesarias para establecer un esqueleto estructural necesario para que funcionen las repeticiones EGF 11 y 12. De forma interesante, se observó casi el efecto opuesto: aunque nuestro ensayo de agregación no mide la fuerza de la interacción, cuando fuimos limitando el sitio de unión a fragmentos cada vez más pequeños, observamos un incremento de la capacidad de las células transfectadas para agregarse con células que expresan Delta, sugiriendo que las secuencias EGF normales flanqueantes realmente impiden la asociación entre las proteínas. En dos series separadas de estructuras artificiales, tanto en el fondo de la estructura artificial n° 3  $\Delta$ Cla (compárese con n° 9, 16, 19, 23) como en el fondo de la estructura artificial n° 25  $\Delta$ EGF (compárese con n° 26, 27, 28), observamos un incremento de la capacidad de agregación tal, que las estructuras artificiales más pequeñas (n° 19, 23, 28, 29) se agregaban consistentemente por encima de los niveles del tipo silvestre (n° 1 pMtNMg). Estos resultados implican que las repeticiones EGF del contorno pueden servir para limitar la capacidad de las repeticiones EGF 11 y 12 para acceder a Delta, modulando quizá de este modo las interacciones Notch-Delta *in vivo*.

Notch codifica una proteína transmembranal, estructuralmente compleja que se ha propuesto que tiene un papel pleotrópico en todo el desarrollo de *Drosophila*. El hecho de que las repeticiones EGF 11 y 12 parezcan funcionar como una unidad modular independiente, que es suficiente, al menos en cultivos celulares, para las interacciones con Delta, presenta inmediatamente la cuestión del papel de la hipótesis de si éstas también pueden formar dominios de unión modulares para otras proteínas que interaccionan con Notch en distintas etapas durante el desarrollo.

Además de *Notch* de *Xenopus*, *lin-12* y *glp-1*, dos genes que se pensaba que funcionaban en las interacciones célula-célula implicadas en la especificación de ciertos destinos celulares durante el desarrollo de *C. elegans*, codifican proteínas transmembranales homólogas a EGF que son estructuralmente totalmente similares a Notch de *Drosophila* y de *Xenopus*. Las cuatro proteínas contienen repeticiones homólogas a EGF seguidas de otras tres repeticiones ricas en cisteína (repeticiones Notch/*lin-12*) en el dominio extracelular, un dominio transmembranal sencillo y seis repeticiones *cdc10/anquirina* en la región intracelular. A diferencia de *Notch* de *Xenopus*, que en base a la comparación de ambas secuencias como en los resultados de nuestro ensayo de unión a Delta, parece probable que codifica la parte que corresponde a *Notch* de *Drosophila*, *lin-12* y *glp-1* codifican probablemente miembros distintos de la misma familia génica. La comparación de los productos proteicos pronosticados de *lin-12* y *glp-1* con Notch, revela diferencias específicas a pesar de una organización global similar de los motivos estructurales. La diferencia más obvia es que las proteínas de *lin-12* y *glp-1* contienen sólo 13 y 10 repeticiones EGF, respectivamente, comparando con las 36 de Notch de *Xenopus* y de *Drosophila*. Además, en los genes del nematodo el conjunto de repeticiones EGF está interrumpido después de la primera repetición EGF por una extensión diferente de la secuencia, ausente en Notch. Adicionalmente, en relación con el dominio de unión a Delta que hemos definido como las repeticiones EGF 11 y 12 de Notch, no hay dos repeticiones EGF contiguas en las proteínas *lin-12* o *glp-1* que muestren la secuencia de consenso de unión a Ca<sup>++</sup>, ni cualquiera de las dos repeticiones contiguas que muestran una similitud notable con las repeticiones EGF 11 y 12, sugiriendo de nuevo que los productos génicos de *lin-12* y *glp-1* son probablemente funcionalmente distintos de Notch.

Nuestro resultado de que las repeticiones EGF 11 y 12 de Notch forman una unidad de unión a Delta discreta representa la primera evidencia concreta que apoya la idea de que cada repetición EGF o una subpoblación pequeña de repeticiones, puede tener un papel único durante el desarrollo, posiblemente a través de interacciones directas con otras proteínas. Las homologías vistas entre el dominio adhesivo de Delta y Serrate (véase la Sección 8.3.4, más abajo) sugieren que la parte homóloga de Serrate es “adhesiva” porque media en la unión de otras proteínas toporrítmicas. Además, el gen *scabrous*, que codifica una proteína secretada con similitud al fibrinógeno, puede interaccionar con Notch.

Además de la repetición EGF, múltiples copias de otros motivos estructurales están presentes habitualmente en una variedad de proteínas. Un ejemplo relevante es el motivo *cdc10/anquirina*, seis copias del mismo se encuentran en el dominio intracelular de Notch. La anquirina contiene 22 de estas repeticiones. Quizá conjuntos ordenados repetidos de motivos estructurales pueden representar en general un conjunto lineal de una serie de unidades modulares de unión a proteínas. Resumiendo estos resultados junto con la complejidad estructural, genética y del desarrollo conocida de *Notch*, Notch puede interaccionar con una cantidad de ligandos diferentes en un patrón regulado temporal y espacialmente durante todo el desarrollo. Tales interacciones específicas del contexto con proteínas extracelulares podrían estar mediadas por las repeticiones EGF y *Notch/lin-12*, mientras que las interacciones con las proteínas del citoesqueleto y citoplásmicas podrían estar mediadas por los motivos *cdc10/anquirina* intracelulares.

## 8. El extremo amino de Delta es un dominio de unión a EGF que interacciona con Notch y Delta

La agregación de células cultivadas programadas para expresar proteínas Delta de tipo silvestre y variantes, se ha empleado para delinear secuencias de Delta necesarias para la interacción heterotípica con Notch y la interacción homotípica con Delta. Hemos encontrado que el extremo amino del dominio extracelular de Delta es necesario y suficiente para la participación de Delta en las interacciones heterotípicas (Delta-Notch) y las homotípicas (Delta-Delta). Deducimos que el extremo amino de Delta es un dominio de unión a un motivo de Delta (EBD), dado que las secuencias similares a EGF de Notch son suficientes para mediar en la interacción heterotípica con Delta. El EBD de Delta posee aparentemente dos actividades: la capacidad de unirse a secuencias relacionadas con EGF y la capacidad de autoasociarse. También encontramos que Delta es absorbida por las células cultivadas que expresan Notch, lo que puede ser un reflejo de un mecanismo por el que estas proteínas interaccionan *in vivo*.

### 8.1. Materiales y métodos

#### 8.1.1 Líneas celulares

La línea celular S2 de *Drosophila* (Schneider y col., 1972, J. Embryol. Exp. Morph. 27, 353-365) empleada en estos experimentos, se hizo crecer tal y como se ha descrito en la Sección 6.

## 8.1.2. Muestras inmunológicas

La inmunohistoquímica se realizó tal y como se ha descrito en la Sección 6, más arriba, o a veces con ligeras modificaciones de este procedimiento. Los antisueros y los anticuerpos empleados incluyen sueros policlonales anti-Delta de ratón, producidos contra un segmento del conjunto ELR de Delta que se extiende desde la cuarta hasta la novena ELR (véase la Sección 6); sueros policlonales de rata anti-Delta producidos contra el mismo segmento de Delta (véase la Sección 6); sueros policlonales anti-Notch de rata producidos contra un segmento del conjunto de ELR de Notch que se extiende desde la quinta hasta la treceava ELR; anticuerpo monoclonal de ratón C17.9C6 (véase la Sección 6), que reconoce el dominio intracelular de Notch; y anticuerpo monoclonal de ratón EP-104 (Hortsch y col., 1990, Neuron 4, 697-709), que reconoce la forma larga de neuroglía de *Drosophila*.

## 8.1.3. Estructuras artificiales de vector de expresión

Las estructuras artificiales empleadas para programar la expresión de Delta de tipo silvestre (pMTD11) y Notch de tipo silvestre (pMTNMg) se describen en la Sección 6, más arriba. Las estructuras artificiales que dirigen la expresión de proteínas de Delta variantes se generaron empleando pMTD11, el ADNc D11 clonado en Bluescript+ (pBSD11; Kopczynski y col., 1988, Genes Dev. 2, 1723-1735) y pRmHa3-104 (A.J. Bieber, pers. comm.), que consta de la inserción del ADNc 1B7A-250 en el vector promotor de la metalotioneína pRmHa-3 (Bunch y col., 1988, Nucl. Acids Res. 16, 1043-1061) y favorece la expresión inducible de la forma larga de neuroglía de *Drosophila* (Hortsch y col., 1990, Neuron 4, 697-709).

Brevemente, las estructuras artificiales se prepararon del modo siguiente:

Del(Sca-Nae) - Se corta pBSD11 con Sall (digestión completa) y con ScaI (parcial), se aísla el fragmento que contiene el vector. Se corta pBSD11 con NaeI (parcial) y con Sall (completa), se aísla el fragmento que codifica el carboxilo-terminal de Delta. Se ligan los fragmentos, se transforman y se aíslan los clones. Se transfiere la inserción EcoRI a pRmHa-3.

Del(Bam-Bgl) - Se corta pBSD11 con BglII (completa) y con BamHI (parcial), se rellenan los extremos con polimerasa Klenow de ADN, se ligan, se transforman y se aíslan los clones. Se transfiere la inserción EcoRI a pRmHa-3.

Del(ELR1-ELR3) - Se multiplican con PCR los pares de bases 236-830 del ADNc D11 empleando 5-ACTTCAG-CAACGATCACGGG-3' (SEQ ID NO:26) y 5'-TTGGGTATGTGACAGTAATCG-3' (SEQ ID NO:27), se tratan con polimerasa T4 de ADN, se ligan en pBSD11 cortado con ScaI (parcial) y BglII (completa) y se rellenan los extremos con polimerasa Klenow de ADN, se transforma y se aíslan los clones. Se transfiere el fragmento BamHI-Sall que codifica el carboxilo terminal de Delta a pRmHa-3.

Del(ELR4-ELR5) - pBSD11 se digirió completamente con BglII y parcialmente con PstI. Se aisló el fragmento que contenía el vector de 5,6 kb, se circularizó empleando la ligasa T4 de ADN en presencia de un exceso molar de 100x del oligonucleótido 5'-GATCTGCA-3' y se transformó y los clones se aislaron. La inserción EcoRI resultante se transfirió después a pRmHa-3.

Ter(Dde) - Se cortó pBSD11 con DdeI (parcialmente), y se rellenan los extremos con la polimerasa Klenow de ADN, se ligó con exceso molar de 100x de 5'-TTAAGTTAACTTAA-3' (SEQ ID NO:28), se transformó y se aislaron los clones. Se transfirió la inserción EcoRI a pRmHa-3.

Ins(Nae)A - pMTD11 se cortó con NaeI (parcialmente), se aisló el fragmento que contenía el vector, se ligó con un exceso molar de 100x de 5'-GGAAGATCTTCC-3' (SEQ ID NO:29), se transformó y se aislaron los clones.

NAE B - pMTD11 se digirió parcialmente con NaeI y se aisló la población de círculos tentativamente linealizados de aproximadamente 5,8 kb de longitud. Los fragmentos se circularizaron de nuevo empleando la ligasa T4 de ADN en presencia de un exceso molar de 100x del oligonucleótido 5'-GGAAGATCTTCC-3' (SEQ ID NO:29) y se transformaron, y se aisló un clon (NAE A) que contenía inserciones múltiples del engarzador. NAE A se digirió completamente con BglII y se aislaron los fragmentos resultantes de 0,4 kb y 5,4 kb, se ligaron y se transformaron, y los clones se aislaron.

Ins(Stu) - Se cortó pMTD11 con StuI (completamente), se aisló el fragmento que contenía el vector, se ligó con un exceso molar de 100x de 5'-GGAAGATCTTCC-3' (SEQ ID NO:29), se transformó y se aislaron los clones.

STU B - pMTD11 se digirió completamente con StuI, y se aisló el fragmento resultante de 5,8 kb. El fragmento se circularizó de nuevo empleando la ligasa T4 de ADN en presencia de un exceso molar de 100x del oligonucleótido 5'-GGAAGATCTTCC-3' (SEQ ID NO:29) y se transformó, y se aisló un clon (STU B) que contenía inserciones múltiples del engarzador. STU B se digirió completamente con BglII y se aislaron los fragmentos resultantes de 0,6 kb y 5,2 kb, se ligaron y se transformaron, y se aislaron los clones.

## ES 2 181 674 T5

NG1 - pRmHa3-104 se cortó con BglIII (completamente) y EcoRI (completamente), se aisló el fragmento que contenía el vector. Se cortó Ins(Nae)A con EcoRI (completamente) y con BglIII (completamente), se aisló el fragmento que codificaba el amino-terminal de Delta. Se ligaron los fragmentos, se transformaron y se aislaron los clones.

5 NG2 - Se cortó pRmHa3-104 con BglIII (completamente) y con EcoRI (completamente), se aisló el fragmento que contenía el vector. Se cortó Del(ELR1-ELR3) con EcoRI (completamente) y con BglIII (completamente), se aisló el fragmento que codificaba el amino-terminal de Delta. Se ligaron los fragmentos, se transformaron y se aislaron los clones.

10 NG3 - pRmHa3-104 se cortó con BglIII (completamente) y con EcoRI (completamente), se aisló el fragmento que contiene el vector. Se cortó pMTD11 con EcoRI (completamente) y con BglIII (completamente), se aisló el fragmento que codificaba el amino-terminal de Delta. Se ligaron los fragmentos, se transformaron y se aislaron los clones.

15 NG4 - pRmHa3-104 se cortó con BglIII (completamente) y con EcoRI (completamente), se aisló el fragmento que contenía el vector. Se cortó Del(Sca-Nae) con EcoRI (completamente) y con BglIII (completamente), se aisló el fragmento que codificaba el amino-terminal de Delta. Se ligaron los fragmentos, se transformaron y se aislaron los clones.

20 NG5 - Se generó Del(Sca-Stu) del modo siguiente: se cortó pMTD11 con ScaI (completamente) y con StuI (completamente), se aisló el fragmento ScaI-ScaI que codificaba el amino-terminal y el fragmento StuI-ScaI que codificaba el carboxilo-terminal, se ligaron, se transformaron y se aislaron los clones. Se cortó Del(Sca-Stu) con EcoRI (completamente) y con BglIII (completamente), se aisló el fragmento que codificaba el amino-terminal de Delta. Se cortó pRmHa3-104 con BglIII (completamente) y con EcoRI (completamente), se aisló el fragmento que contenía el vector. Se ligaron los fragmentos, se transformaron y se aislaron los clones.

25 Los contenidos de la secuencia de las distintas variantes de Delta se muestran en la Tabla IV. Los diagramas esquemáticos de las variantes de Delta definidas en la Tabla IV, se muestran en la Figura 9.

30

(Tabla pasa a página siguiente)

35

40

45

50

55

60

65

## ES 2 181 674 T5

TABLA IV

*Contenidos de las secuencias de las variantes de delta empleadas en este estudio*

5		<u>Nucleótidos</u>	<u>Aminoácidos</u>
	Tipo silvestre	1-2892 <sup>A</sup>	1-833
	Del (Sca-Nae)	1-235/734-2892	1-31/W/199-833
10	Del (Bam-Bgl)	1-713/1134-2892	1-191/332-833
	Del (ELR1-ELR3)	1-830/1134-2892	1-230/332-833
	Del (ELR4-ELR5)	1-1137/1405-2892	1-332/422-833
15	Ter (Dde)	1-2021/TTAAGTTAACTTAA <sup>E</sup> / 2227-2892	1-626/H
	Ins (Nae) A	1-733/ (GGAAGATCTTCC) <sub>n</sub> <sup>F</sup> / 734-2892 <sup>B</sup>	1-197/ (RKIF) <sub>n</sub> 198-833
20	NAE B	1-733/GGAAGATCTTCC <sup>F</sup> / 734-2892	1-197/RKIF 198-833
25	Ins (Stu)	1-535/ (GGAAGATCTTCC) <sub>n</sub> <sup>F</sup> / 536-2892 <sup>B</sup>	1-131/ G (KFIR) <sub>n-1</sub> KIFP/133-833
30	STU B	1-535/GGAAGATCTTCC <sup>F</sup> / 536-2892	1-131/GKIFP 133-833
35	NG1	1-733/GGAA/ <b>2889-3955 (NG)</b> <sup>C</sup>	1-198/K/ <b>952-1302</b> <sup>D</sup>
	NG2	1-830/ <b>2889-3955 (NG)</b>	1-230/ <b>952-1302</b>
40	NG3	1-1133/ <b>2889-3955 (NG)</b>	1-331/ <b>952-1302</b>
45	NG4	1-235/734-1133/ 2889-3955 (NG)	1-31/199-331/ <b>952-1302</b>
50	NG5	1-235/536-1133/ <b>2889-3955 (NG)</b>	1-31/S/133- <b>952-1302</b>

A Coordinadas de las secuencias de Delta que se corresponden con la secuencia del ADNc DII (Figura 12).

55 B El número exacto de engarzadores insertados no se ha determinado para esta estructura artificial.

C Coordinadas para las secuencias de nucleótidos de neuroglía (Bieber y col., 1989, Cell 59, 447-460; Hortsch y col., 1990, Neuron 4, 697-709) presentes en quimeras de Delta-neuroglía que se corresponden con la secuencia del ADNc 1B7A-250 (Figura 13, SEQ ID NO:5) y que se indican con negrita.

60 D Las secuencias de aminoácidos de neuroglía se obtienen a partir de la traducción conceptual de la secuencia de nucleótidos del ADNc 1B7A-250 (Figura 13, SEQ ID NO:5) y se indican con negrita.

E SEQ ID NO:28

65 F SEQ ID NO:29

8.1.4. *Protocolos de agregación*

La transfección y la agregación celulares se realizaron tal y como se ha descrito en la Sección 6, más arriba, o con ligeras modificaciones de las mismas.

5

8.2. *Resultados*

8.2.1. *Las secuencias amino-terminales dentro del dominio extracelular de delta son necesarias y suficientes para la interacción heterotípica con Notch*

10

Ya que se anticipó que algunas variantes de Delta no se podían localizar eficazmente en la superficie celular, investigamos la relación entre el nivel de expresión de Delta de tipo silvestre y el grado de agregación con células que expresan Notch, haciendo variar la cantidad aportada de estructura artificial de expresión de Delta en diferentes transfecciones. Encontramos que la interacción heterotípica Delta-Notch mostraba sólo una ligera dependencia del nivel de aportación de Delta sobre 10 veces un margen en este ensayo (Figura 9A). Dada la solidez de la interacción heterotípica sobre el margen sometido a ensayo y nuestras observaciones de que cada una de las variantes de Delta que empleamos, mostraba una acumulación sustancial en la superficie en células transfectadas, deducimos que la incapacidad de una variante de Delta dada para favorecer la agregación heterotípica, refleja lo más probable una deficiencia funcional mostrada por esa variante, como contrapunto con el impacto de niveles reducidos de expresión en la superficie en la agregación heterotípica.

15

20

Los resultados de los experimentos de agregación heterotípica mediada por variantes de Delta y de Notch de tipo silvestre, se muestran en la Tabla V.

25

(Tabla pasa a página siguiente)

30

35

40

45

50

55

60

65

TABLA V

## AGREGACIÓN HETEROTÍPICA MEDIADA POR VARIANTES DE DELTA Y POR NOTCH DE TIPO SILVESTRE

Estructura artificial	Total agregado <sup>A</sup>	Total no agregado	N° de Exp.	Agregación		Sin agregación	
				Notch <sup>+</sup>	Delta <sup>+B</sup>	Notch <sup>+</sup>	Delta <sup>+</sup>
Tipo silvestre	33 (H) <sup>C</sup>	179	1	15	18	67	112
	58 (H)	247	2	37	21	218	29
	38 (H)	209	3	21	17	148	61
	29 (H)	174	4	18	11	95	79
	175 (B)	68	5	84	91	37	31
Del (Sca-Nae)	0 (H)	207	1	0	0	125	82
	0 (H)	226	2	0	0	215	11
	0 (H)	287	3	0	0	215	72
	0 (H)	200	4	ND <sup>D</sup>	ND	ND	ND
Del (Bam-Bgl)	4 (H)	245	1	3	1	171	74
	0 (H)	200	2	0	0	110	90
	0 (H)	200	3	ND	ND	ND	ND
Del (ELR1-ELR3)	28 (B)	296	1	11	17	139	157
	20 (B)	90	2	9	11	53	37
	22 (B)	227	3	19	13	114	113
	127 (B)	97	4	19	78	66	61
Del (ELR4-ELR5)	38 (H)	188	1	26	12	141	47
	36 (H)	204	2	20	16	90	114
Ter (Dde)	53 (H)	236	1	24	29	144	92
	51 (H)	214	2	30	21	126	88

	52 (H)	190	3	30	22	110	80
Ins (Nae)A	0 (B)	205	1	0	0	111	94
	0 (B)	254	2	0	0	161	93
	0 (B)	201	3	0	0	121	80
NG1	0 (B)	208	1	0	0	140	68
	0 (B)	114	2	0	0	38	76
	0 (B)	218	3	0	0	76	142
NG2	14 (B)	106	1	7	7	54	52
	50 (B)	216	2	35	15	94	122
	36 (B)	168	3	12	24	29	139
NG3	71 (B)	175	1	43	28	84	91
NG4	0 (B)	254	1	0	0	150	104
	0 (B)	215	2	0	0	35	180
	0 (B)	200	3	0	0	93	107

- A Número total de células con expresión en agregados que contienen cuatro o más células.
- B Células que expresan estructuras artificiales basadas en neuroglia (NGn) fueron detectadas empleando un anticuerpo monoclonal que reconoce el dominio intracelular de neuroglia (véase Materiales y Métodos).
- C (H) indica que se agregaron células en un matraz Erlenmeyer de 25 ml, (B) indica que se agregaron células en una placa de microtitulación de 12 pocillos (véase Materiales y Métodos).
- D Datos para tipos de células individuales (es decir, Delta<sup>+</sup> y Notch<sup>+</sup>) con agregación y sin agregación no fueron registradas.

## ES 2 181 674 T5

Los aminoácidos Delta (AA) 1-230 son el intervalo de secuencia mínima actual, definido como suficiente para la interacción con Notch. Esto se basa en el éxito de la agregación de NG2-Notch. Dentro de este intervalo, los AA198-230 de Delta son decisivos porque su delección en la estructura artificial NG1 inactivaba la actividad de unión a Notch observada para la estructura artificial NG2. También dentro de este intervalo, los AA32-198 de Delta son decisivos porque su delección en la estructura artificial NG4, también inactivaba la actividad de unión a Notch, observada para la estructura artificial NG3. La importancia de los AA192-230 de Delta también se apoya en la observación de que la variante Del(ELR1-ELR3), que contiene todos los aminoácidos de Delta, excepto los AA231-331, posea actividad de unión a Notch, mientras que la variante Del(Bam-Bgl), que contiene todos los aminoácidos de Delta excepto los AA192-331, tenía inactivada aparentemente la actividad de unión a Notch.

La configuración y/o la secuencia primaria en la proximidad de los AA197-198 de Delta es aparentemente decisiva, debido a que una inserción multimérica del tetrapéptido - Arg-Lys-Ile-Phe [en código de una sola letra (véase, por ejemplo, Lehninger y col., 1975, Biochemistry, 2ª ed., pág. 72), RKIF] (SEQ ID NO:30) - entre estos dos restos, como en la estructura artificial Ins(Nae)A, inactivaba la actividad de unión a Notch, observada en Delta de tipo silvestre.

Además, la observación de que la estructura artificial Del(ELR1-ELR3) favorecía la agregación, implica que ELR1-ELR3 no es necesaria para la interacción Delta-Notch; la observación de que la estructura artificial Del(ELR4-ELR5) favorecía la agregación, implica que ELR4 y ELR5 no son necesaria para la interacción Delta-Notch, y la observación de que la estructura artificial Ter(Dde) favorece la agregación, implica que el dominio intracelular de Delta no es necesario para la interacción Delta-Notch.

*8.2.2. Las secuencias amino-terminales dentro del dominio extracelular de Delta son necesarias y suficientes para la interacción homotípica*

Los resultados de los experimentos de agregación homotípica mediada por variantes de Delta, se muestran en la Tabla VI

TABLA VI  
*Agregación homotípica mediada por variantes de Delta*

<u>Estructura arti- ficial</u>	<u>Agregación</u>	<u>Sin agregación</u>	<u>Exp. N°</u>
Tipo silvestre	38 (H) <sup>A</sup>	175	1
	48 (H)	171	2
	13 (H)	95	3
	33 (H)	173	4
	134 (B)	72	5
Del (Sca-Nae)	0 (H)	200	1
	0 (H)	200	2
	0 (H)	200	3
Del (Bam-Bgl)	0 (H)	200	1
	0 (H)	200	2
	0 (H)	200	3
Del (ELR1-ELR3)	160 (B)	62	1
	55 (B)	80	2
	0 (B)	200	3
	4 (B)	203	4
	41 (B)	234	5
	4 (B)	366	6 <sup>B</sup>
	23 (B)	325 (1:20)	
	0 (B)	400	7 <sup>B</sup>
	5 (B)	347 (1:5)	
	10 (B)	228 (1:20)	
0 (B)	400	8 <sup>B</sup>	

## ES 2 181 674 T5

	16 (B)	346 (1:5)	
	4 (B)	268 (1:20)	
5	4 (B)	500	9 <sup>c</sup>
	18 (B)	500 (1:5)	
10	12 (B)	271 (1:20)	
	7 (B)	128 (1:50)	
	0 (B)	500	10 <sup>c</sup>
15	0 (B)	500 (1:5)	
	0 (B)	500 (1:20)	
20	21 (B)	246 (1:50)	
	0 (B)	500	11 <sup>c</sup>
	5 (B)	500 (1:5)	
25	8 (B)	177 (1:20)	
	4 (B)	69 (1:50)	
30	Del (ELR4-ELR5)	21 (H)	175
		29 (H)	243
		35 (H)	179
35	Ter (Dde)	53 (H)	164
		33 (H)	178
		36 (H)	203
40	Ins (Nae) A	0 (B)	200
		0 (B)	200
45		0 (B)	200

- A (H) indica que las células se agregaron en un matraz Erlenmeyer de 25 ml; (B) indica que las células se agregaron en una placa de microtitulación de 12 pocillos (véase Materiales y Métodos).
- B Las células transfectadas se incubaron bajo condiciones de agregación durante una noche, a continuación se diluyeron en el volumen apropiado de células S2 en fase logarítmica, en presencia de inductor y se incubaron bajo condiciones de agregación durante cuatro a seis horas adicionales.
- C Las células transfectadas a las que se había añadido inductor, se diluyeron en el volumen adecuado de células S2 en fase logarítmica a las que se había añadido inductor, y la mezcla de células se incubó bajo condiciones de agregación durante una noche.

60 La delección de los AA32-198 de Delta [Del(Sca-Nae)] o los AA192-331 de Delta [Del(Bam-Bgl)] de la proteína Delta de longitud completa, eliminaba la interacción Delta-Delta. La delección de los AA231-331 de Delta [Del(ELR1-ELR3)] no eliminaba la interacción Delta-Delta. Por tanto, las secuencias dentro de los AA32-230 de Delta son necesarias para la interacción Delta-Delta.

65 La configuración y/o la secuencia primaria en la proximidad de los AA197/198 de Delta son evidentemente decisivas para la interacción Delta-Delta, porque una inserción múltimera del tetrapéptido -Arg-Lys-Ile-Phe- (SEQ ID NO:30) entre estos dos restos, como en la estructura artificial Ins(Nae)A, inactivaba la interacción Delta-Delta.

## ES 2 181 674 T5

Además, la observación de que la estructura artificial Del(ELR1-ELR3) podía favorecer la agregación, implica que ELR1-ELR3 no son necesarias para la interacción Delta-Delta; la observación de que la estructura artificial Del (ELR4-ELR5) favorecía la agregación, implica que ELR4 y ELR5 no son necesarias para la interacción Delta-Delta, y la observación de que la estructura artificial Ter(Dde) favorece la agregación, implica que el dominio intracelular de Delta no es necesario para la interacción Delta-Delta.

Un resumen de los resultados de los ensayos para la agregación heterotípica y homotípica con distintas estructuras artificiales, se muestra en la Tabla VI A.

TABLA VI A

AGREGACIÓN MEDIADA POR PROTEÍNAS DELTA DE TIPO SILVESTRE Y VARIANTES DE DELTA			
ESTRUCTURA ARTIFICIAL	AGREGACIÓN HETEROTÍPICA <sup>a</sup>		AGREGACIÓN HOMOTÍPICA <sup>b</sup>
	DELTA	NOTCH	DELTA
Tipo silvestre	33 ± 12 <sup>c</sup>	26 ± 11 <sup>c</sup>	27 ± 10 <sup>c</sup>
Del (Sca-Nae)	0	0	0
Del (Bam-Bgl)	0,4 ± 0,4	0,6 ± 0,6	0
Del (ELR1-ELR3)	25 ± 11 <sup>d</sup>	15 ± 3 <sup>d</sup>	32 ± 15 <sup>d</sup>
Del (ELR4-ELR5)	17 ± 2	18 ± 2	13 ± 2
Ter (Dde)	22 ± 1	18 ± 2	18 ± 3
NAE B	25 ± 5	0	27 ± 7
STU B	0	0	0
NG1	0	0	0
NG2	13 ± 1	23 ± 6	4 ± 1 <sup>d</sup>
NG3	16 ± 1	13 ± 1	27 ± 17
NG4	0	0	0,5 ± 0,3

a: Fracción media (%) de células Delta o Notch en agregados de cuatro o más células (± error típico). N = 3 duplicados, si no se indica de otro modo.

b: Fracción media (%) de células Delta en agregados de cuatro o más células (± error típico), N = 3 duplicados, si no se indica de otro modo.

c: N = 5 duplicados.

d: N = 4 duplicados.

### 8.2.3 Las secuencias de Delta implicadas en las interacciones heterotípicas y homotípicas son cualitativamente distintas

Las características respectivas de las secuencias de Delta restañadas para la interacción heterotípica y homotípica se definieron adicionalmente empleando variantes de Delta, en las que se introdujeron inserciones cortas, en marco, con engarzador traducible, dentro del extremo amino de Delta (es decir, NAE B y STU B; Figura 9, Tabla VI A). La sustitución del resto de Delta 132 (A) por el pentapéptido GKIFP (variante STU B) conduce a la inactivación

de las actividades de la interacción heterotípica y homotípica del extremo amino de Delta. Esto sugiere que algunas secuencias de Delta necesarias para estas dos interacciones distintas, coinciden y residen en la proximidad del resto 132. Por otro lado, la inserción del tetrapéptido RKIF entre los restos de Delta 198 y 199 (variante NAE B) elimina la capacidad del extremo amino de Delta de mediar en la interacción heterotípica con Notch, pero no tiene un efecto evidente sobre la capacidad del extremo amino alterado para mediar en la interacción homotípica. El descubrimiento de que la inserción NAE B afecta sólo a una de las dos actividades del extremo amino de Delta, implica que las secuencias de Delta median en las interacciones heterotípicas y homotípicas, aunque coinciden, son cualitativamente distintas.

#### 8.2.4. *Delta es absorbida por las células que expresan Notch*

En el transcurso de muchos experimentos de agregación heterotípica, hemos observado que la proteína Delta se puede encontrar a veces dentro de células que han sido programadas para expresar Notch, pero no Delta. Realizamos ensayos de agregación heterotípica mezclando poblaciones inicialmente separadas de células S2 que se habían transfectado independientemente con estructuras artificiales de expresión que programan la expresión de Delta o de Notch. Entonces, detectamos frecuentemente una tinción punteada de Delta dentro de células que expresan Notch en agregados heterotípicos, empleando antisueros específicos de Delta. Nuestras observaciones son compatibles con la unión de Delta directamente a Notch en la superficie celular y una eliminación posterior de este complejo Delta-Notch de la superficie celular, mediante endocitosis.

### 8.3. *Discusión*

#### 8.3.1. *Las secuencias amino-terminales no relacionadas con EGF están implicadas en la interacción entre Delta y Notch*

Hemos empleado ensayos de agregación celular para definir una región dentro de la región amino-proximal del dominio extracelular de Delta que es necesaria y suficiente para mediar en la interacción Delta-Notch. Análisis funcionales de una combinación de deleciones y estructuras artificiales suficientes revelaron que esta región se extiende, como máximo, desde el AA1 hasta el AA230. Es sorprendente que esta región no incluya ninguna de las secuencias similares a EGF que residen dentro del dominio extracelular de Delta. Es probable que las secuencias particulares de Delta dentro del intervalo suficiente necesario para la interacción con Notch, incluyan los AA198-230, porque la deleción de estos restos elimina la actividad de unión a Notch. El hecho de que la deleción de los AA32-198 también inactive la actividad de unión a Notch, sugiere que secuencias amino-proximales del AA198, también son necesarias, aunque el impacto nocivo-deletéreo de esta deleción podría dar como resultado la eliminación de aminoácidos adicionales en estrecha proximidad al AA198.

Las secuencias dentro de Delta suficientes para la interacción con Notch se pueden agrupar en tres subdominios - N1, N2 y N3 - que difieren en sus contenidos respectivos de restos de cisteína (Figura 10, SEQ ID NO:3). Los dominios N1 y N3 contienen cada uno seis restos de cisteína, mientras que el dominio N2 contiene uno. El número par de cisteínas presentes en N1 y N3, respectivamente, permite la posibilidad de que las estructuras respectivas de estos subdominios estén dictadas, en parte, por la formación de enlaces disulfuro particulares. El amplio patrón de organización de los extremos amino-terminales de Delta es generalmente también análogo al del dominio extracelular del receptor de EGF de vertebrados (Lax y col., 1988, Mol. Cell. Biol. 8, 1970-1978), en el que secuencias que se piensa que interaccionan con EGF, están unidas por dos subdominios ricos en cisteína.

#### 8.3.2. *Las secuencias de delta necesarias para interacciones homotípicas y para interacciones homotípicas heterotípicas parecen coincidir*

Nuestros resultados indican que secuencias esenciales para la interacción homotípica de Delta, residen dentro del intervalo AA32-230. La deleción de secuencias o la inserción de aminoácidos adicionales dentro de este dominio amino-proximal, eliminan la capacidad de tales variantes de Delta de favorecer aisladamente la agregación celular. Por tanto, las secuencias necesarias para la interacción Delta-Delta se cartografían dentro del mismo dominio de la proteína que las necesarias para la interacción Delta-Notch.

#### 8.3.3. *El extremo amino de Delta constituye un motivo de unión a EGF*

El trabajo descrito en los ejemplos más arriba, ha revelado que las secuencias de Notch necesarias para la interacción Delta-Notch en el ensayo de agregación celular, se cartografían dentro del conjunto de repeticiones similares a EGF del dominio extracelular de Notch. Este resultado implica que Delta y Notch interaccionan en virtud de la unión del extremo amino de Delta a las secuencias similares a EGF dentro de Notch, y de este modo, el extremo amino del dominio extracelular de Delta constituye un dominio de unión a EGF (Figura 11).

Estos resultados aumentan la posibilidad de que la interacción homotípica de Delta implique la unión del extremo amino de Delta a secuencias similares a EGF dentro del dominio extracelular de Delta (Figura 12). Sin embargo, ninguna de las repeticiones similares a EGF dentro del dominio extracelular de Delta es idéntica a cualquiera de las repeticiones similares a EGF dentro del dominio extracelular de Notch (Figuras 13, SEQ ID NO:6; Wharton y col., 1985, Cell 43, 567-581). En base a esto, si las interacciones homotípicas de Delta están mediadas realmente por la interacción entre el extremo amino de Delta y las repeticiones similares a EGF de Delta, entonces el dominio de unión a EGF de Delta tiene la capacidad de interaccionar con al menos dos secuencias distintas similares a EGF.

#### 8.3.4. *Las secuencias de delta implicadas en la interacción Delta-Notch están conservadas en la proteína Serrate*

La alineación de las secuencias de aminoácidos desde los extremos amino-terminales de Delta (Figura 13, SEQ ID NO:6, y Figura 15, SEQ ID NO:9) y Serrate (Fleming y col., 1990, Genes Dev. 4, 2188-2201; Thomas y col., 1991, Devel. 111, 749-761) revela una conservación sorprendente del carácter estructural y de la composición de las secuencias. La estructura general de los subdominios N1-N2-N3 del extremo amino-terminal de Delta, también se observa dentro del extremo amino-terminal de Serrate, como es el caso específico de seis restos de cisteína dentro de los dominios homólogos a N1 de Delta y homólogo a N3 de Delta, de la proteína Serrate. Dos bloques de conservación notables se corresponden con los AA63-73 de Delta (restos 8/11 idénticos) y los AA195-206 de Delta (restos 10/11 idénticos). El último bloque tiene un interés particular debido a que la inserción de aminoácidos adicionales en este intervalo, puede eliminar la capacidad de Delta de unirse a Notch o a Delta.

#### 8.3.5. *Las interacciones cis y trans entre Delta y Notch pueden implicar secuencias diferentes dentro de Notch*

La inspección de todas las estructuras de Delta y Notch, sugiere que la interacción Delta-Notch podría implicar contactos entre el dominio de unión a EGF de Delta con las dos regiones dentro de Notch, dependiendo de si la interacción era entre moléculas que residen en membranas opuestas o dentro de la misma membrana (Figura 11). Los ensayos de agregación celular, que detectan presumiblemente la interacción de moléculas en membranas opuestas, implican que el dominio de unión a EGF de Delta interacciona con las repeticiones similares a EGF 11 y 12 (véanse los ejemplos más arriba). Si las agrupaciones en tándem de motivos similares a EGF forman estructuras similares a varillas (Engel, 1989, FEBS Lett. 251, 1-7) dentro de las proteínas Delta y Notch, entonces el desplazamiento estimado del dominio de unión a EGF de Delta desde la superficie celular sería presumiblemente suficiente para acomodar el conjunto rígido de las repeticiones similares a EGF de Notch 1-10. También es curioso indicar que el desplazamiento del dominio de unión a EGF de Delta desde la superficie celular, podría colocar este dominio en la proximidad de las repeticiones similares a EGF (25-29) de Notch que están afectadas por mutaciones Abruptex (Hartley y col., 1987, EMBO J. 6, 3407-3417; Kelley y col., 1987, Mol. Cell. Biol. 6, 3094-3108) y podría permitir la interacción de las proteínas de Delta y Notch presentes en la misma membrana.

#### 8.3.6. *Interacciones análogas a la interacción Delta-Notch en vertebrados*

Dada la interacción entre Delta y Notch en *Drosophila*, es bastante probable que un homólogo de Delta (*¿Helta?*) exista en vertebrados y que los aspectos cualitativos y moleculares de las interacciones Delta-Notch y Delta-Delta que hemos definido en *Drosophila*, estén altamente conservados en vertebrados, incluyendo los seres humanos. Tales homólogos se pueden clonar y secuenciar tal y como se ha descrito más arriba, Sección 5.2.

### 9. *Secuencias que median en las interacciones Notch-Serrate*

Hemos descrito una nueva interacción molecular entre Notch y Serrate, y hemos mostrado que las dos repeticiones EGF de Notch que median en las interacciones con Delta, a saber, las repeticiones EGF 11 y 12, también constituyen un dominio de unión a Serrate.

Para someter a ensayo si Notch y Serrate interaccionan directamente, se transfectaron células S2 con una estructura artificial de expresión de Serrate y se mezclaron con células que expresan Notch en nuestro ensayo de agregación. Para la estructura artificial de expresión de Serrate, se empleó un cebador sintético que contenía un sitio BamHI artificial, inmediatamente 5' del iniciador AUG en la posición 442 (todos los números de secuencias están de acuerdo con Fleming y col., 1990, Genes & Dev. 4:2188-2201) y era homólogo hasta la posición 464, junto con un segundo cebador desde la posición 681-698 para generar un fragmento de ADN de ~260 pares de bases. Este fragmento se cortó con BamHI y KpnI (posición 571) y se ligó en Bluescript KS+ (Stratagene). Esta estructura artificial, BTSer5'PCR, se examinó por secuenciación y luego se cortó con KpnI. El fragmento KpnI de Serrate (571-2981) se insertó y se seleccionó la orientación adecuada, para generar BTSer5'PCR-Kpn. El fragmento 5' SacII de BTSer5'PCR-Kpn (sitios SacII en el poliengarzador Bluescript y en Serrate (1199)) se aisló y se empleó para reemplazar el fragmento SacII 5' del ADNc C1 (Fleming y col., 1990, Genes & Dev. 4:2188-2201), regenerando de este modo el ADNc de Serrate de longitud completa, menos las regiones 5' no traducidas. Esta inserción se aisló mediante digestión con SalI y digestión parcial con BamHI y se transportó a los sitios BamHI y SalI de pRmHa-3, para generar la estructura artificial de expresión final, Ser-mtn.

Encontramos que las células que expresan Serrate se adhieren a células que expresan Notch en un modo dependiente de calcio (Figura 6 y Tabla VII). Sin embargo, al contrario que Delta, bajo las condiciones experimentales sometidas a ensayo, Serrate no parece que interaccione homotípicamente. Además, no detectamos interacciones entre Serrate y Delta.

TABLA VII

*Efecto de Ca<sup>++</sup> Exógeno sobre la Agregación<sup>a</sup> de Notch-Serrate*

	<u>Notch-Serrate</u>	
	<u>Sin Ca<sup>++</sup></u>	<u>Con Ca<sup>++</sup></u>
5 1. pMtNMg	0	15
10 32. ΔECN+EGF(10-12)	0	13
15 33. ΔC1a+XEGF(10-13)	0	15

15  
20  
25  
<sup>a</sup> Datos presentados como porcentaje de células que expresan Notch encontradas en agregados (como en la Figura 6). Todos los números son de experimentos de transfección aislados (más que un promedio de valores de distintos experimentos separados, como en la Figura 6).

30 Hemos sometido a ensayo una subpoblación de nuestras estructuras artificiales de delección de Notch para cartografiar el dominio de unión a Serrate y hemos encontrado que las repeticiones EGF 11 y 12, además de unirse a Delta, también median en las interacciones con Serrate (Figura 6; estructuras artificiales n° 1, 7-10, 13, 16, 17, 19, 28 y 32). Además, la función de unión a Serrate de estas repeticiones parece también que se ha conservado en las dos repeticiones EGF correspondientes a Notch de *Xenopus* (n° 33ΔC1a+XEGF(10-13)). Estos resultados muestran sin ambigüedad que Notch interacciona con Delta y con Serrate, y que las dos mismas repeticiones EGF de Notch median en ambas interacciones. También fuimos capaces de definir la región de Serrate que es esencial para la agregación de Notch/Serrate. La delección de los nucleótidos 676-1287 (es decir, los aminoácidos 79-282) (véase Figura 15), elimina la capacidad de la proteína Serrate para agregarse con Notch.

40 Notch y Serrate parece que se agregan menos eficazmente que Notch y Delta, quizá debido a que la interacción Notch-Serrate es más débil. Por ejemplo, cuando marcamos los agregados Notch-Delta, detectamos ~40% de todas las células que expresan Notch en agrupaciones con células que expresan Delta (Figura 6, n° 1 pMtNMg) y ~40% de todas las células que expresan Delta en contacto con células que expresan Notch. Para Notch-Serrate, sólo encontramos ~20% de todas las células que expresan Notch (Figura 6; pMtNMg) y ~15% de todas las células que expresan Serrate en agregados. Para las distintas estructuras artificiales de delección de Notch que sometimos a ensayo, detectamos consistentemente una reducción en la cantidad de agregación entre Notch y Serrate al compararlas con los niveles correspondientes de Notch-Delta (Figura 6), con la posible excepción de las estructuras artificiales n° 9 y 10 que muestran niveles de agregación muy reducidos, incluso con Delta. Una explicación trivial para esta cantidad reducida de agregación podría ser que nuestra estructura artificial de Serrate simplemente no exprese tanta proteína en la superficie celular como la estructura artificial de Delta, disminuyendo de este modo la fuerza de la interacción. Alternativamente, la diferencia en la fuerza de la interacción puede indicar una diferencia funcional fundamental entre las interacciones Notch-Delta y Notch-Serrate que puede ser significativa *in vivo*.

## 10. La clonación, secuenciación y expresión de Notch de ser humano

### 10.1. Aislamiento y secuenciación de Notch humana

55 Los clones para la secuencia de *Notch* humana se obtuvieron originalmente empleando la reacción en cadena de la polimerasa (PCR), para multiplicar el ADN de una genoteca de ADNc de cerebro de feto humano de 17-18 semanas, en el vector Lambda Zap II (Stratagene). Los cebadores degenerados que se iban a emplear en esta reacción, se diseñaron comparando las secuencias de aminoácidos del homólogo de *Notch* de *Xenopus* con *Notch* de *Drosophila*. Se diseñaron tres cebadores (cdc1 (SEQ ID NO:10), cdc2 (SEQ ID NO:11) y cdc3 (SEQ ID NO:12); Figura 16) para multiplicar un fragmento de 200 pb o un fragmento de 400 pb, como parejas de cebadores cdc1/cdc2, o cdc1/cdc3, respectivamente.

65 El fragmento de 400 pb obtenido de este modo, se empleó a continuación como una sonda con la que rastrear la misma genoteca en busca de clones *Notch* humanos. El rastreo original produjo tres únicos clones, hN3k, hN2k y hN5k, todos ellos mostraron, por un análisis posterior de la secuencia, que se encontraban en el extremo 3' de *Notch* humana (Figura 17). Un segundo rastreo que empleaba el extremo 5' de hN3k como sonda, se realizó para buscar

clones que rodeaban el extremo 5' de *Notch* humana. Se obtuvo un único clon, hN4k, a partir de este rastreo, y datos de una secuenciación preliminar indican que contiene la mayoría del extremo 5' del gen (Figura 17). Resumiendo, los clones hN4k, hN3k y hN5k abarcan aproximadamente 10 kb del homólogo de *Notch* humana, comenzando al principio en las repeticiones EGF y extendiéndose en la región 3' no traducida del gen. Los tres clones son inserciones de ADNc en el sitio EcoRI de pBluescript SK (Stratagene). La cepa hospedadora de *E. coli* es XL1-Blue (véase Maniatis, T., 1990, Molecular Cloning, A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, pág. A12).

La secuencia de las distintas partes de *Notch* contenidas en los clones de ADNc, se determinó (empleando Sequenase®, U.S. Biochemical Corp.) y se muestra en las Figuras 19-22 (SEQ ID NO:13 hasta NO:25).

Las secuencias de nucleótidos completas del ADNc de *Notch* humana contenidas en hN3k y hN5k, se determinaron mediante el método de terminación de cadena didesoxi empleando el equipo de reactivos de Sequenase® (U.S. Biochemical Corp.). Las secuencias de nucleótidos que codifican *Notch* humana, en el marco de lectura adecuado, se identificaron fácilmente puesto que la traducción en sólo uno de los tres marcos de lectura posibles, produce una secuencia que, después de comparar con la secuencia publicada de aminoácidos, deducidos de *Notch* de *Drosophila*, produce una secuencia con un grado sustancial de homología con la secuencia de *Notch* de *Drosophila*. Puesto que no hay intrones, la traducción de los tres marcos de lectura posibles y la comparación con *Notch* de *Drosophila* se realizó fácilmente, conduciendo a la identificación terminada de la región codificadora. El ADN y las secuencias de proteínas deducidas del ADNc de *Notch* humana en hN3k y hN5k, se presentan en las Figuras 23 y 24, respectivamente. El clon hN3k codifica una parte de un polipéptido de *Notch* que comienza aproximadamente en la tercera repetición *Notch/lin-12* hasta varios aminoácidos, cerca del aminoácido carboxi-terminal. El clon hN5k codifica una parte de un polipéptido de *Notch* que comienza aproximadamente antes de la región cdc10 hasta el final del polipéptido, y también contiene una región 3' sin traducir.

Comparando el ADN y las secuencias de proteínas presentadas en la Figura 23 (SEQ ID NO:31 y NO:32) con las de la Figura 23 (SEQ ID NO:33 y NO:34) se revelan diferencias importantes entre las secuencias, sugiriendo que hN3k y hN5k representan parte de dos genes distintos homólogos a *Notch*. Nuestros datos sugieren por tanto que el genoma humano contiene más de un gen homólogo a *Notch*. Esto es improbable en *Drosophila*, en donde *Notch* parece ser un gen con una única copia.

La comparación del ADN y de las secuencias de aminoácidos de los homólogos de *Notch* humano contenidos en hN3k y hN5k con las secuencias correspondientes de *Notch* de *Drosophila* (tal y como han publicado Wharton y col., 1985, Cell 43:567-581) y con las secuencias correspondientes de *Notch* de *Xenopus* (tal y como han publicado Coffman y col., 1990, Science 249:1438-1441 o disponibles a través de Genbank® (número de orden M33874)), también revelan diferencias.

La secuencia de aminoácidos mostrada en la Figura 23 (hN3k) se comparó con la secuencia pronosticada del polipéptido TAN-1 mostrado en la Figura 2 de Ellisen y col., Agosto 1991, Cell 66:649-661. Se encontraron algunas diferencias entre las secuencias de aminoácidos deducidas; sin embargo, toda la secuencia del polipéptido de *Notch* de hN3k es idéntica en 99% a la región TAN-1 correspondiente (aminoácidos TAN-1 1455 hasta 2506). Se observaron cuatro diferencias; en la región entre la tercera repetición *Notch/lin-12* y el primer motivo cdc10, hay una arginina (hN3k) en lugar de un X (TAN-1, aminoácido 1763); (2) hay una prolina (hN3k) en lugar de un X (TAN-1, aminoácido 1787); (3) hay un cambio conservador de un resto de ácido aspártico (hN3k) en lugar de un resto de ácido glutámico (TAN-1, aminoácido 2495); y (4) la región carboxilo-terminal difiere sustancialmente entre los aminoácidos TAN-1 2507 y 2535.

La secuencia de aminoácidos mostrada en la Figura 24 (hN5k) se comparó con la secuencia pronosticada del polipéptido TAN-1 mostrado en la Figura 2 de Ellisen y col., Agosto 1991, Cell 66:649-661. Se encontraron diferencias entre las secuencias de aminoácidos deducidas. El polipéptido de *Notch* deducido de hN5k es idéntico en 79% al polipéptido de TAN-1 (idéntico en 64% con *Notch* de *Drosophila*) en la región cdc10 que abarca el motivo cdc10 (aminoácidos TAN-1 1860 hasta 2217) y las regiones flanqueantes bien conservadas (Fig. 25). La región cdc10 cubre los aminoácidos 1860 hasta 2217 de la secuencia TAN-1. Además, el polipéptido codificado de hN5k es idéntico en 65% al polipéptido de TAN-1 (idéntico en 44% a *Notch* de *Drosophila*) en el extremo carboxi-terminal de la molécula que contiene una región (aminoácidos TAN-1 2482 hasta 2551) rica en PEST (prolina, ácido glutámico, serina, treonina) (Fig. 25B). La extensión de 215 aminoácidos que se encuentra entre las regiones anteriormente mencionadas no está bien conservada en ninguno de los clones homólogos de *Notch*, representados por hN3k, hN5k y TAN-1. Ni el polipéptido hN5k ni *Notch* de *Drosophila* muestran niveles significativos de identidad de aminoácidos con las otras proteínas en esta región (por ejemplo, hN5k/TAN-1 = 24% de identidad; hN5k/*Notch* de *Drosophila* = 11% de identidad; TAN-1/*Notch* de *Drosophila* = 17% de identidad). Por el contrario, *Notch* de *Xenopus* (Xotch) (SEQ ID NO:35), *Notch* de rata (SEQ ID NO:36) y TAN-1 (SEQ ID NO:37) continúan compartiendo niveles significativos de identidad de secuencia con otra (por ejemplo, TAN-1/*Notch* de rata = 75% de identidad, TAN-1/*Notch* de *Xenopus* = 45% de identidad, *Notch* de rata/*Notch* de *Xenopus* = 50% de identidad).

Finalmente, el examen de la secuencia de los dominios intracelulares de los homólogos de *Notch* en vertebrados mostrados en la Figura 25B, revelaba un descubrimiento inesperado: todas esas proteínas, incluyendo hN5k, contienen un motivo CcN putativo, asociado con una función de dirección a la diana nuclear, en la región conservada después de la última de las seis repeticiones cdc10 (Fig. 25B). Aunque *Notch* de *Drosophila* carece de un motivo tal definido,

## ES 2 181 674 T5

un estudio más a fondo de su secuencia reveló la presencia de una posible secuencia de localización nuclear bipartita (Robbins y col., 1991, Cell 64:615-623), así como de posibles sitios de fosforilación de CK II y cdc2, todos en proximidad relativa entre sí, definiendo posiblemente de este modo un tipo alternativo de motivo CcN (Fig. 25B).

### 5 10.2. Expresión de Notch humana

Las estructuras artificiales de expresión se prepararon empleando clones de ADNc de *Notch* humana, descritos en la Sección 10.1, anterior. En los casos de hN3k y hN2k, el clon completo se escindió de su vector como un fragmento de restricción EcoRI y se subclonó en el sitio de restricción EcoRI de cada uno de los tres vectores pGEX (vectores de expresión de la S-transferasa de glutatión; Smith y Johnson, 1988, Gene 7, 31-40). Esto permite la expresión del producto de la proteína Notch desde el subclon en el marco de lectura correcto. En el caso de hN5k, el clon contiene dos sitios de restricción internos de EcoRI, que producen fragmentos de 2,6, 1,5 y 0,6 kb. Los dos fragmentos de 2,6 kb y 1,5 kb también se han subclonado en cada uno de los vectores pGEX.

El sistema del vector pGEX se empleó para obtener la expresión de proteínas de fusión (quiméricas) de Notch humana desde las estructuras artificiales descritas anteriormente. El ADN de *Notch* clonado en cada caso se insertó, en fase, en el vector pGEX apropiado. Cada estructura artificial se electroporó a continuación en bacterias (*E. coli*), y se expresó como una proteína de fusión que contenía las secuencias de la proteína Notch fusionadas con el extremo carboxilo de la proteína S-transferasa de glutatión. La expresión de las proteínas de fusión se confirmó por análisis de extractos proteicos bacterianos mediante electroforesis en gel de poliacrilamida, comparando los extractos proteicos obtenidos de las bacterias que contenían los plásmidos pGEX con y sin el ADN de *Notch* insertado. Las proteínas de fusión eran solubles en solución acuosa, y se purificaron a partir de material lisado bacteriano mediante cromatografía de afinidad, empleando agarosa revestida con glutatión (puesto que el extremo carboxilo de la S-transferasa de glutatión se une a glutatiónina). Las proteínas de fusión expresadas se unieron a un anticuerpo de Notch de *Drosophila*, tal y como se sometió a ensayo por transferencia de tipo Western.

Las estructuras artificiales empleadas para preparar proteínas de fusión de Notch/S-transferasa de glutatión, eran del modo siguiente:

30 *hNFPn<sup>o</sup>2* - Se empleó PCR para obtener un fragmento que empezaba justo antes de las repeticiones cdc10 en el nucleótido 192 de la inserción hN5k, hasta justo antes de la región rica en PEST, en el nucleótido 1694. El ADN se digirió a continuación con BamHI y SmaI y el fragmento resultante se ligó en pGEX-3. Después de la expresión, la proteína de fusión se purificó por unión a agarosa de glutatión. El polipéptido purificado se cuantificó en un gel de poliacrilamida con un gradiente de 4-15%. La proteína de fusión resultante tenía un peso molecular aproximado de 83 kD.

*hN3FPn<sup>o</sup>1* - La inserción de ADN de hN3k completa (nucleótidos 1 hasta 3235) se escindió del vector Bluescript (SK) digiriendo con EcoRI. El ADN se ligó en pGEX-3.

40 *hN3FPn<sup>o</sup>2* - Se cortó del vector Bluescript (SK) un segmento 3' de ADN de hN3k (nucleótidos 1847 hasta 3235) más alguno de los poliengarzadores, digiriendo con XmaI. El fragmento se ligó en pGEX-1.

Después de la purificación, estas proteínas de fusión se emplearon para preparar anticuerpos policlonales y/o monoclonales de Notch humana.

45

### 11. Depósito de microorganismos

Las siguientes bacterias recombinantes, siendo portadora cada una de un plásmido que codifica una parte de Notch humana, se depositaron el 2 de Mayo de 1991 en la American Type Culture Collection, 1201 Parklawn Drive, Rockville, Maryland 20852, bajo las condiciones del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes.

55	<u>Bacterias</u>	<u>Plásmido incluido</u>	<u>Nº de orden ATCC</u>
	<i>E. coli</i> XL1-Blue	hN4k	68610
	<i>E. coli</i> XL1-Blue	hN3k	68609
60	<i>E. coli</i> XL1-Blue	hN5k	68611

La presente invención no debe estar limitada en el alcance por los microorganismos depositados o las realizaciones específicas descritas en esta memoria. Claro está que diversas modificaciones de la invención además de las descritas en esta memoria serán evidentes para los expertos en la técnica, a partir de la descripción anterior y las figuras acompañantes. Tales modificaciones pretenden entrar dentro del alcance de las reivindicaciones anexas.

65

En esta memoria se citan diversas publicaciones, cuyas descripciones se incorporan como referencia en su totalidad.

## ES 2 181 674 T5

### REIVINDICACIONES

- 5 1. Una proteína Notch humana purificada, codificada por un primer ácido nucleico, que es obtenible a partir de ADN humano o ADNc y que es hibridable con un segundo ácido nucleico que consiste en la secuencia de Notch humana, o su complemento, contenido en (a) el plásmido hN3k, tal y como se ha depositado en la ATCC con el número de orden asignado 68609, o (b) el plásmido hN5k, tal y como se ha depositado en la ATCC con el número de orden asignado 68611; proteína la cual es capaz de unirse a una proteína Delta o es capaz de unirse a un anticuerpo específico de la proteína Notch humana.
- 10 2. Una proteína purificada que comprende la secuencia de aminoácidos descrita en la Figura 23, o la secuencia de aminoácidos descrita en la Figura 24.
- 15 3. Una proteína purificada que comprende la secuencia de aminoácidos de Notch
- codificada por la secuencia de ADN de Notch humana contenida en el plásmido hN3k, tal y como se ha depositado en la ATCC, con el número de orden asignado 68609; o
  - codificada por la secuencia de ADN de Notch humana contenida en el plásmido hN5k, tal y como se ha depositado en la ATCC, con el número de orden asignado 68611.
- 20 4. Una proteína purificada que es un fragmento de una proteína Notch humana de la reivindicación 1, fragmento el cual, es capaz de unirse a una proteína Delta, o es capaz de unirse a un anticuerpo específico de la proteína Notch humana.
- 25 5. Un fragmento de proteína purificada de acuerdo con la reivindicación 4, en la que dicho fragmento es una parte de la secuencia de la proteína Notch codificada por la secuencia de ADNc humana contenida en el plásmido hN3k, tal y como se ha depositado en la ATCC con el número de orden asignado 68609, o en el plásmido hN5k, tal y como se ha depositado en la ATCC con el número de orden asignado 68611.
- 30 6. Un fragmento de proteína purificada de acuerdo con la reivindicación 4, que comprende:
- un fragmento de una proteína Notch humana que consiste en el dominio extracelular de la proteína;
  - 35 - un fragmento de una proteína Notch humana que consiste en el dominio intracelular de la proteína;
  - un fragmento de una proteína Notch humana que consiste en los dominios extracelular y transmembranal de la proteína;
  - 40 - un fragmento de una proteína Notch humana que consiste en la región que contiene las repeticiones cdc10 de la proteína;
  - una región de una proteína Notch humana que consiste en las repeticiones homólogas al factor de crecimiento epidérmico (EGF) de la proteína; o
  - 45 - una región de una proteína Notch humana que consiste en las repeticiones Notch/lin-12 de la proteína.
- 50 7. Un fragmento de proteína purificada de acuerdo con la reivindicación 6, en la que la secuencia del dominio intracelular de la proteína y la secuencia de la región que contiene las repeticiones cdc10, están codificadas por una parte del plásmido hN3k, tal y como se ha depositado en la ATCC con el número de orden asignado 68609, o por una parte del plásmido hN5k, tal y como se ha depositado en la ATCC con el número de orden asignado 68611.
- 55 8. Un fragmento de proteína purificada de acuerdo con la reivindicación 4, que comprende la parte de una proteína Notch humana de la reivindicación 1, con la mayor homología con las repeticiones similares al EGF 11 y 12 de la secuencia de Notch de *Drosophila*, tal y como se muestra en la Figura 8 (SEQ ID NO:1).
- 60 9. Un fragmento de proteína purificada de la reivindicación 4:
- que consiste en el dominio intracelular de la proteína;
  - que consiste en el dominio extracelular de la proteína; o
  - que consiste en los dominios extracelular y transmembranal de la proteína.
- 65 10. Una proteína purificada de la reivindicación 2 ó 3, o un fragmento de proteína purificada de la reivindicación 4, que se **caracteriza** por la capacidad *in vitro*, cuando se expresa en la superficie de una primera célula, de unirse a una proteína Delta expresada en la superficie de una segunda célula.

## ES 2 181 674 T5

11. Una proteína purificada de la reivindicación 2 ó 3, o un fragmento de proteína purificada de la reivindicación 4, fragmento el cual, es capaz de unirse a un anticuerpo específico de Notch humana.

5 12. Una proteína quimérica que comprende una proteína de acuerdo con una cualquiera de las reivindicaciones 1 a 3, o un fragmento de proteína de acuerdo con una cualquiera de las reivindicaciones 4 a 11, asociada con una secuencia de proteína heteróloga.

13. Un ácido nucleico purificado que codifica

10 - una proteína Notch humana de acuerdo con una cualquiera de las reivindicaciones 1 a 3;

- un fragmento de proteína de acuerdo con una cualquiera de las reivindicaciones 4 a 11; o

15 - una proteína quimérica de acuerdo con la reivindicación 12.

14. Un ácido nucleico purificado de acuerdo con la reivindicación 13, que comprende una secuencia que codifica la secuencia de aminoácidos descrita en la Figura 23 o en la Figura 24.

20 15. Un ácido nucleico purificado que comprende el ADNc de Notch humana contenido en el plásmido hN3k, tal y como se ha depositado en la ATCC con el número de orden asignado 68609, o en el plásmido hN5k, tal y como se ha depositado en la ATCC con el número de orden asignado 68611, o una secuencia de nucleótidos complementaria al mismo.

25 16. Un ácido nucleico purificado de acuerdo con la reivindicación 13, que comprende una secuencia de ADNc que codifica el fragmento de proteína de la reivindicación 6.

17. Un ácido nucleico purificado de acuerdo con la reivindicación 13, que codifica el fragmento de proteína de la reivindicación 8.

30 18. Un vector de ácido nucleico que comprende un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 13 a 17.

19. Una célula recombinante que contiene el vector de ácido nucleico de la reivindicación 18.

35 20. Un método para producir:

- una proteína Notch humana de acuerdo con una cualquiera de las reivindicaciones 1 a 3;

40 - un fragmento de proteína de acuerdo con una cualquiera de las reivindicaciones 4 a 11; o

- una proteína quimérica de acuerdo con la reivindicación 12,

45 que comprende hacer crecer una célula recombinante de acuerdo con la reivindicación 19, de modo que la proteína, el fragmento de proteína o la proteína quimérica se expresa por la célula; y aislar la proteína, el fragmento de proteína o la proteína quimérica expresada por la célula.

21. Un anticuerpo específico de Notch humana y que se une a una proteína de Notch humana de acuerdo con la reivindicación 1, o un fragmento de proteína purificada de acuerdo con la reivindicación 4.

50 22. Un método para producir un anticuerpo de acuerdo con la reivindicación 21, que comprende inmunizar un animal hospedador con una proteína de acuerdo con una cualquiera de las reivindicaciones 1 a 3, un fragmento de proteína de acuerdo con una cualquiera de las reivindicaciones 4 a 11, o una proteína quimérica de acuerdo con la reivindicación 12, y seleccionar un anticuerpo que se une a dicha proteína, a dicho fragmento de proteína o a dicha proteína quimérica.

55 23. Un método de acuerdo con la reivindicación 22, en el que dicho anticuerpo es un anticuerpo policlonal.

24. Un método de acuerdo con la reivindicación 22, en el que dicho anticuerpo es un anticuerpo monoclonal.

60 25. Un fragmento o un derivado del anticuerpo de acuerdo con la reivindicación 21, que contiene el idiotipo del anticuerpo.

26. Un anticuerpo específico de Notch humana, que se une a la secuencia de aminoácidos de Notch de la proteína de acuerdo con la reivindicación 3.

65

## ES 2 181 674 T5

27. Un método para aislar un ácido nucleico de Notch humana, que comprende:

5 a. poner en contacto una muestra que contiene ácidos nucleicos que comprenden secuencias de nucleótidos humanos con una sonda de ácido nucleico, comprendiendo dicha sonda el ácido nucleico de acuerdo con la reivindicación 13 o una parte del mismo, bajo condiciones tales que la hibridación puede tener lugar; y

b. aislar un ácido nucleico en la muestra que se hibrida con la sonda.

10 28. El método de acuerdo con la reivindicación 27, en el que la muestra contiene ADNc de Notch humana, en el que el ADNc comprende la secuencia de ADN mostrada en la Figura 23 o en la Figura 24.

29. Un método para identificar un análogo de una proteína toporrítmica adhesiva, que comprende:

15 a. someter a ensayo la unión de una molécula con una proteína Notch humana de acuerdo con la reivindicación 1, o una parte de la misma que comprende las repeticiones similares a EGF 11 y 12; y

b. identificar una molécula que se une en la etapa (a); en la que la unión con una proteína Notch humana indica que la molécula es un análogo de una proteína toporrítmica adhesiva.

20

25

30

35

40

45

50

55

60

65

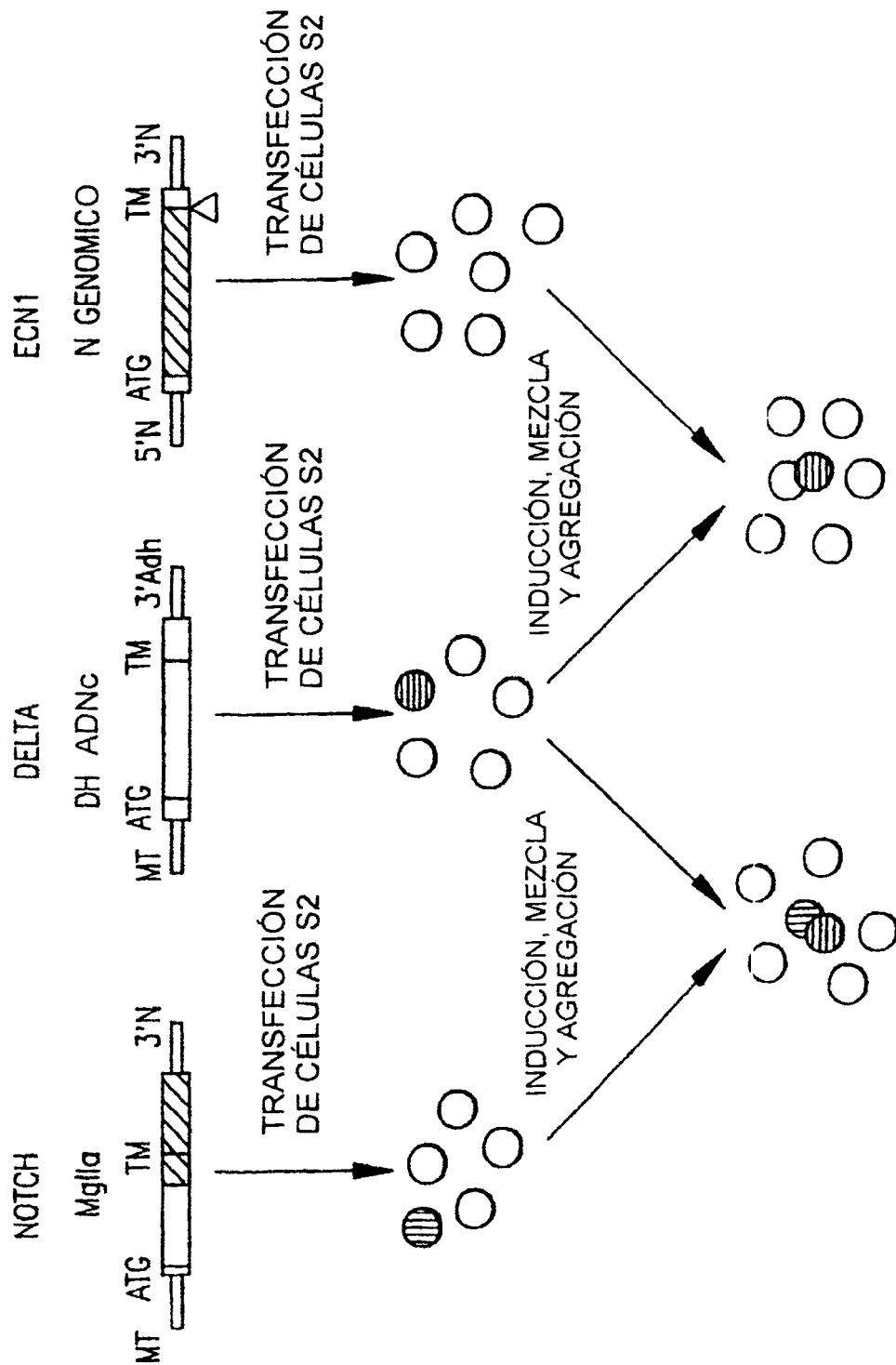
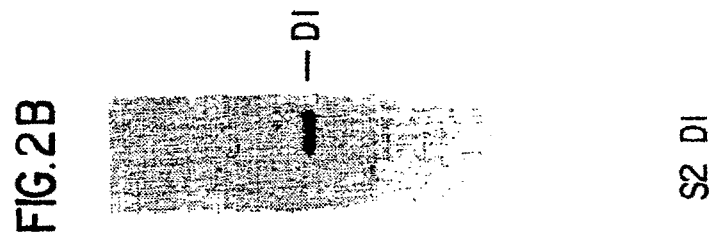


FIG.1



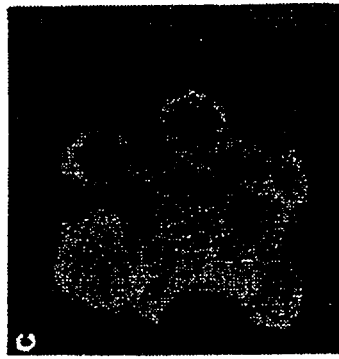


FIG. 3C

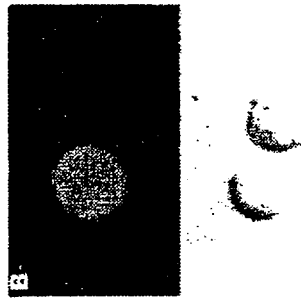


FIG. 3B

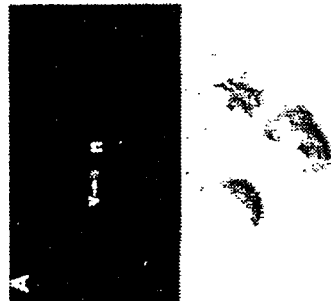


FIG. 3A



FIG.3F

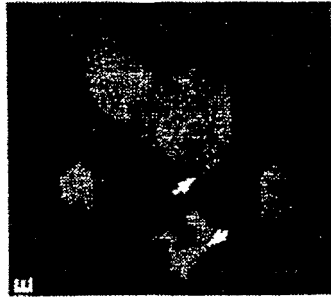


FIG.3E

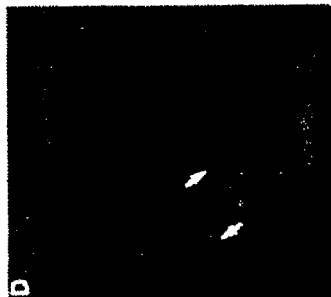


FIG.3D

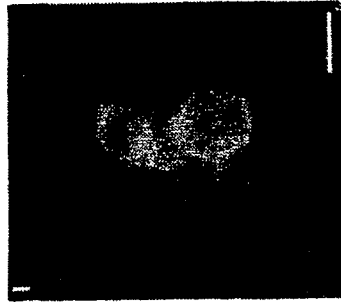


FIG.3I

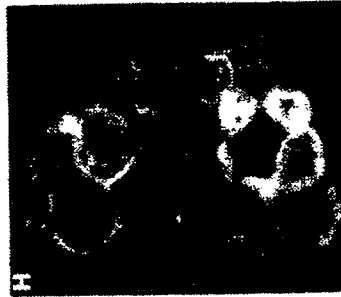


FIG.3H

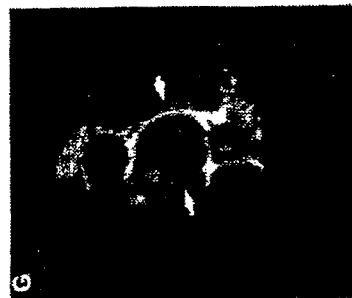


FIG.3G

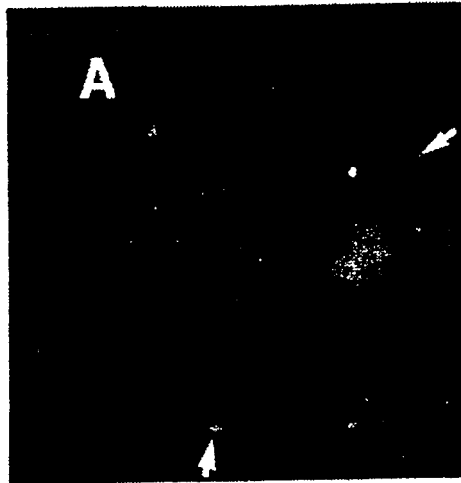


FIG. 4A

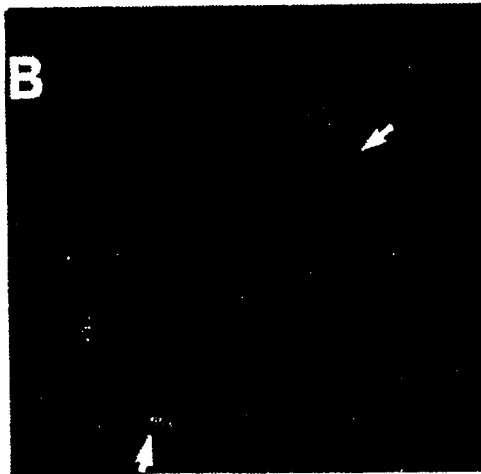


FIG. 4B

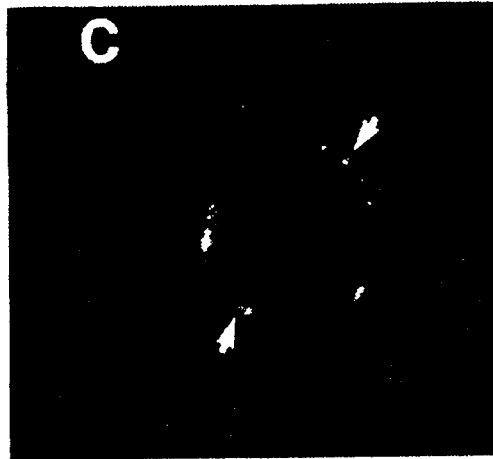


FIG.4C

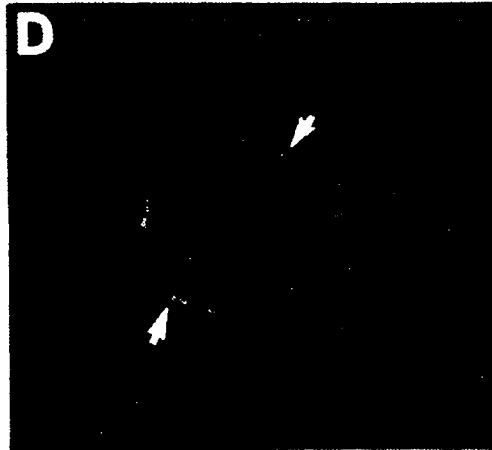


FIG. 4D

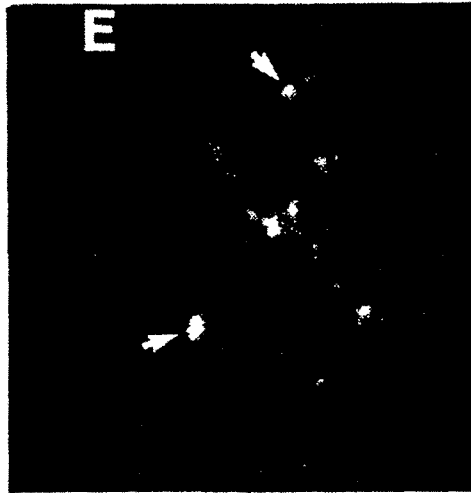


FIG. 4E

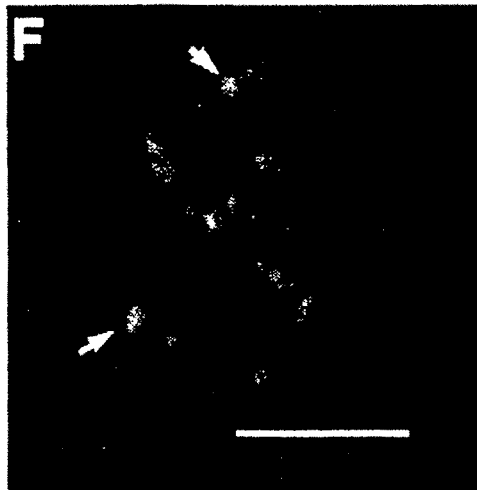


FIG. 4F

FIG. 5B

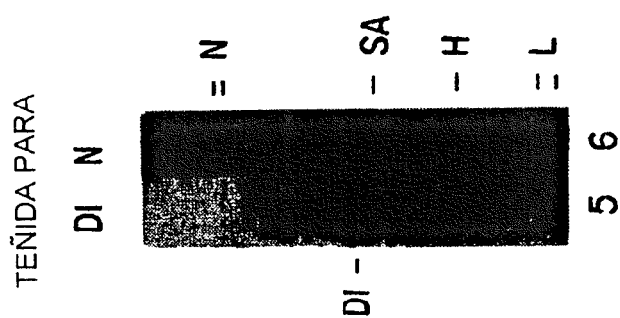
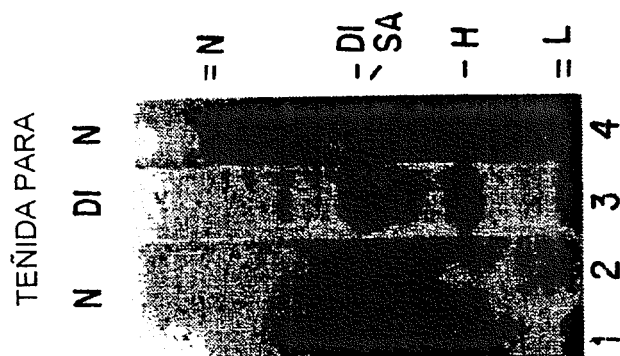


FIG. 5A



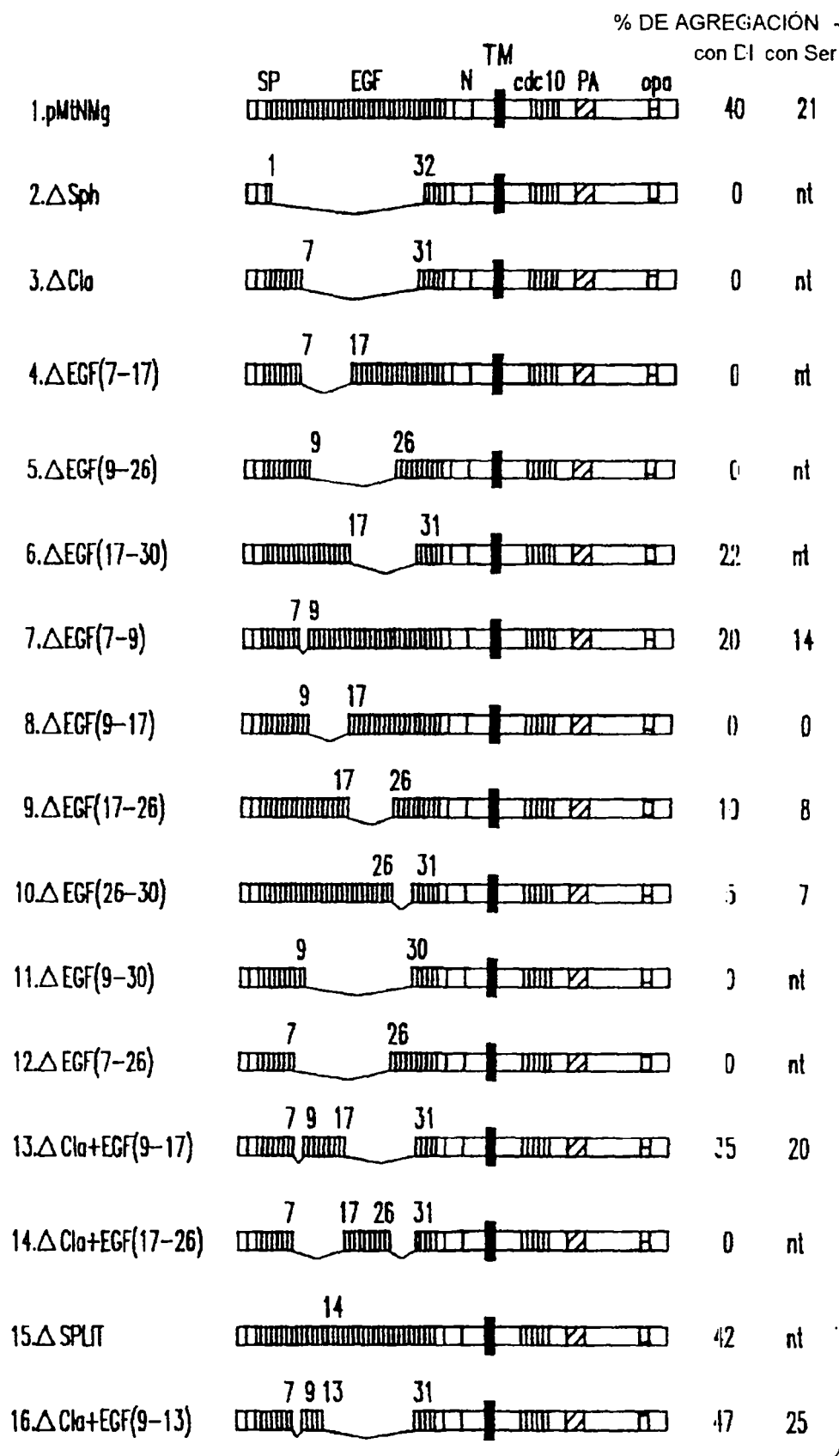


FIG.6A

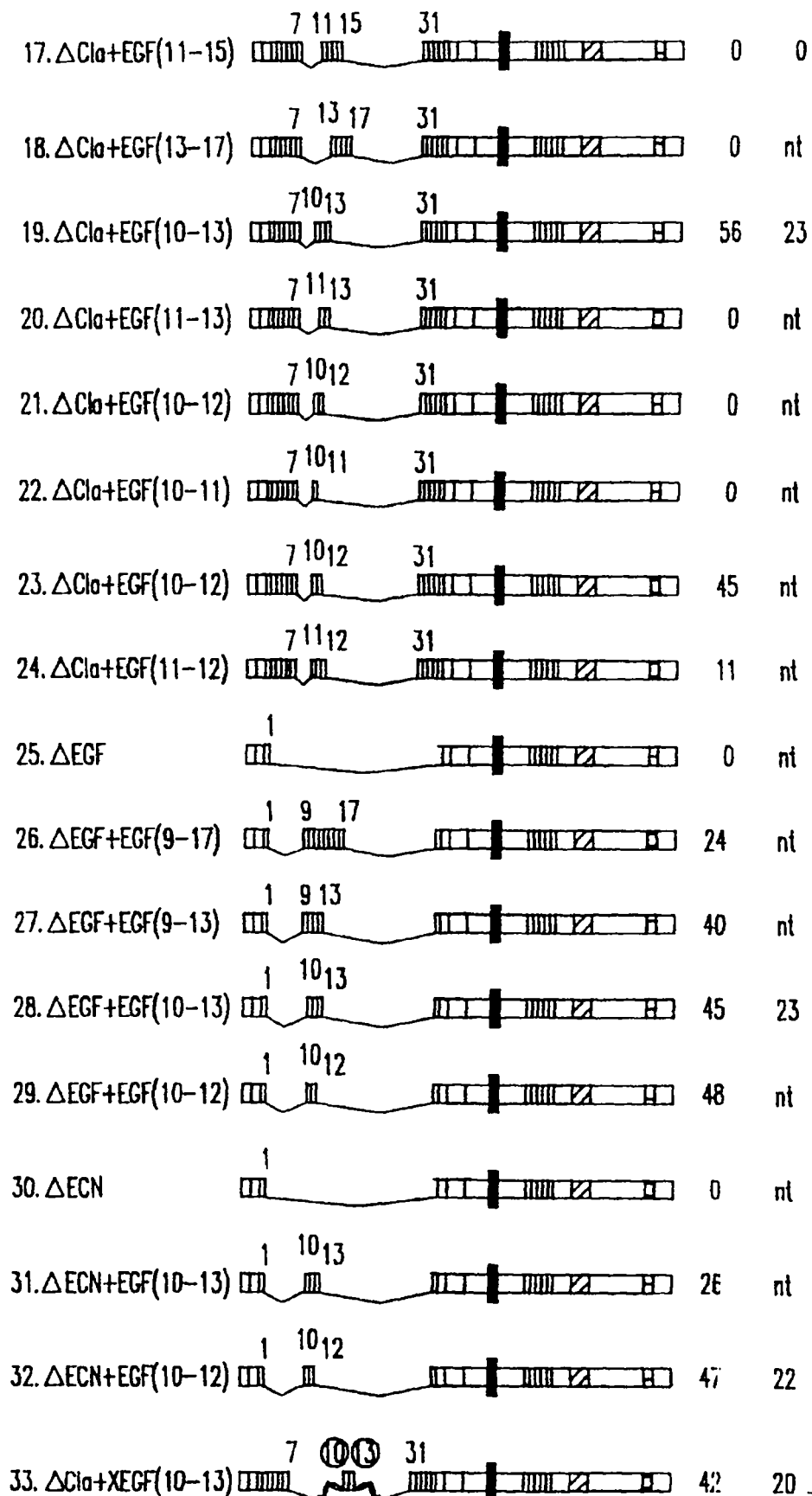


FIG.6B

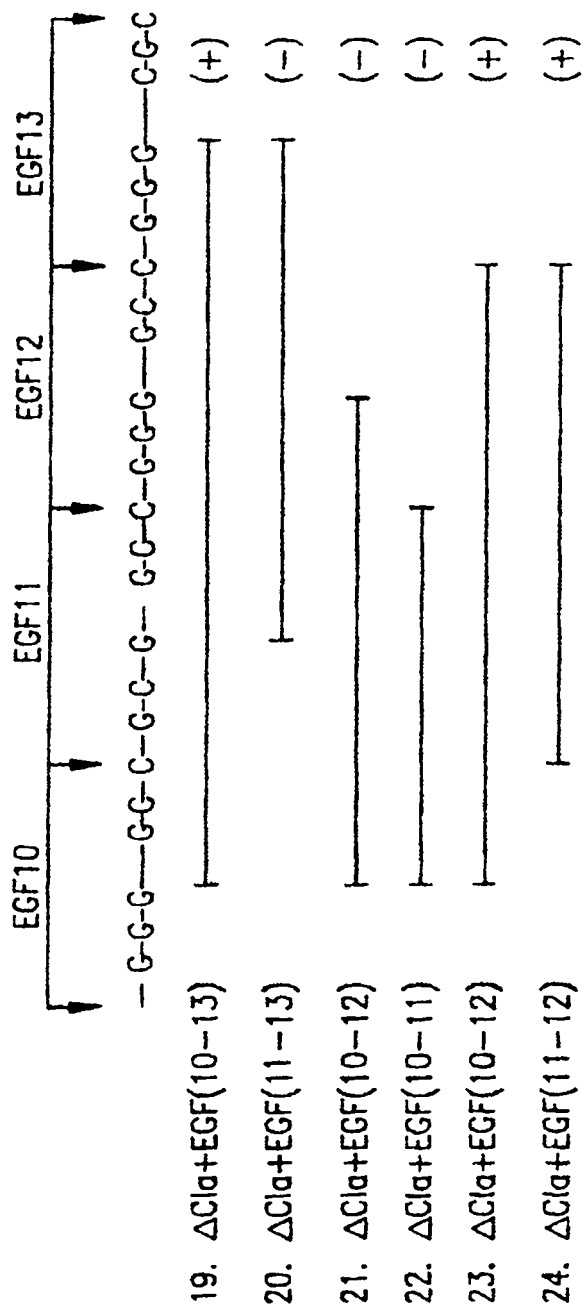


FIG.7

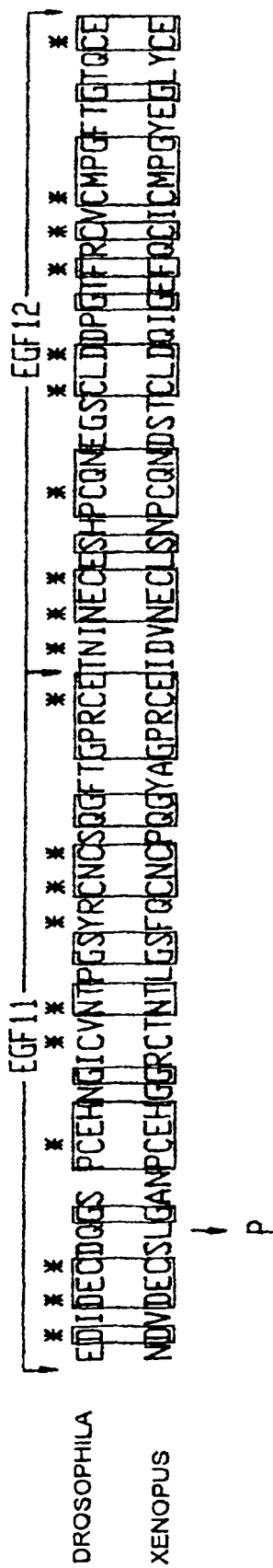


FIG. 8

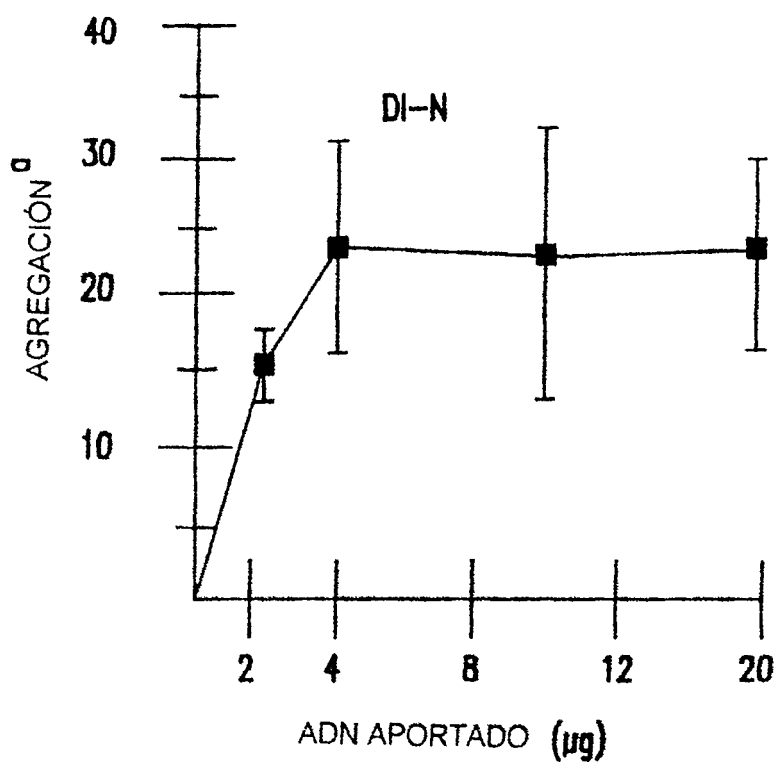


FIG.9A

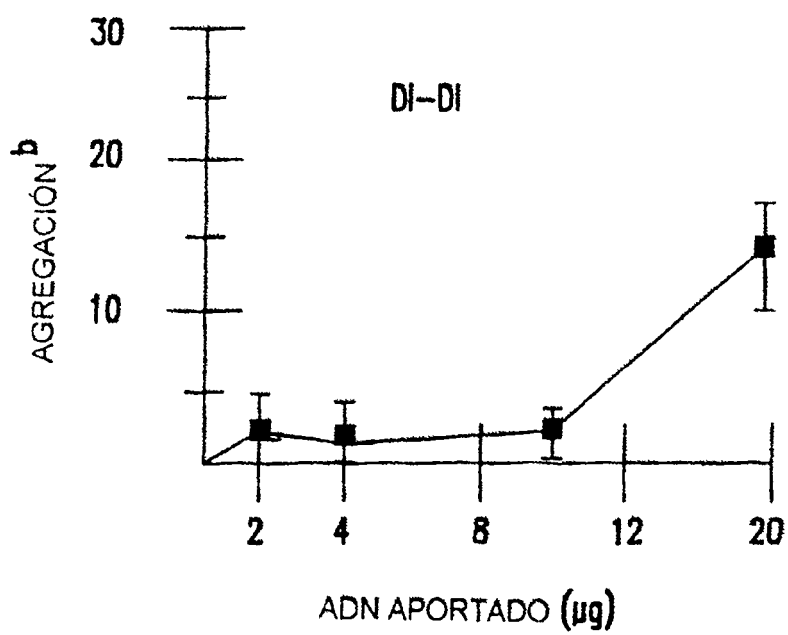


FIG.9B



DI 24

GSFELRLKYFSNDHGRDNEGRDCCSGESDGGATGKGLG.SCKTRFRVCLKHYQATIDTTS...QC  
 N1  
 GNFELEILEISNTNSHLLNGYDGGMPAELRATKIIGSPQTTAFRLCLKEYQTTEGGASISTGC  
 SER85

TYGDVITPILGENSVNLTDQRFQNKGFNPIQFPFSFSWPGTFSLIVEAW  
 SFGNATTKILGGSSFVLSDPG.....YGAIVLPFTFRVTKSFTLILQAL  
 STU B  
 EGKIFPW

N2

HDTNNSGNARTNKLLIQRLLYQQVLEVSSEWKTNKSEQYTSLEYDFRVT  
 DMYNTS..YPDAERLIEETSYSGVILPSPWKTLDHIGNARITYRVRVQ

NAE B  
 CRKIFR  
 226  
 DDLYYGGGCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYIC  
 CAVTYNNTTGTTCRPRDDQFGHYACGSEGGKLCNGWQGVNQC  
 283

N3

FIG.10

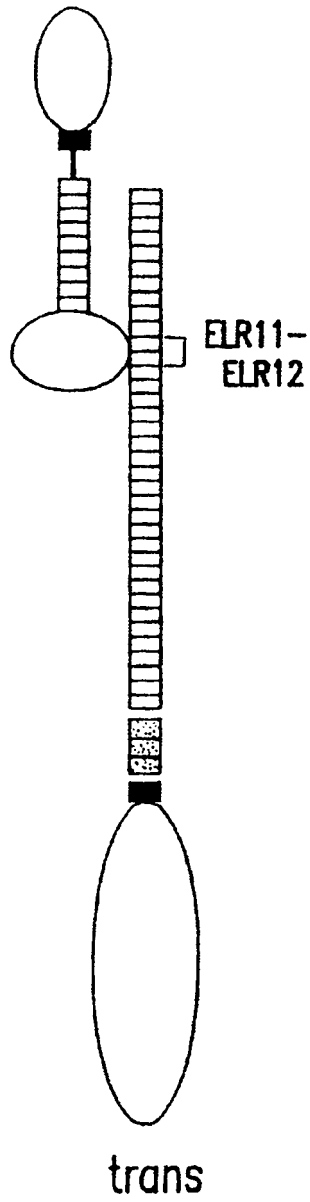


FIG.11A

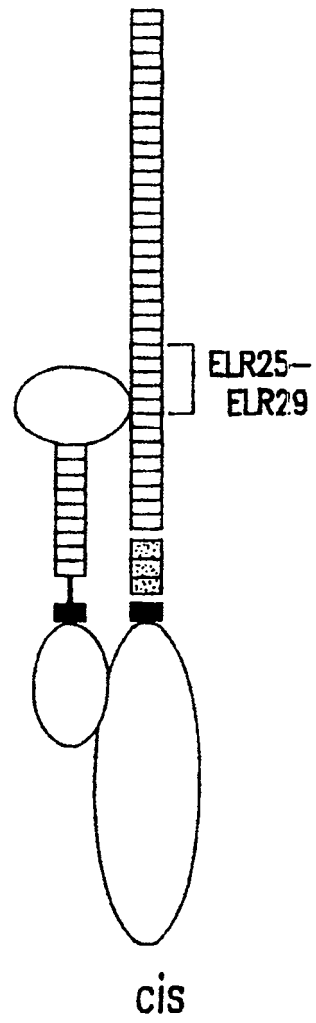
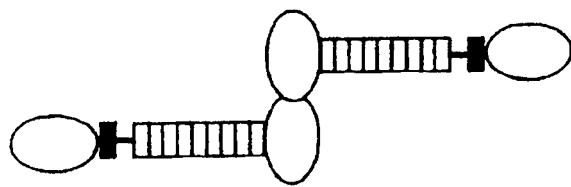
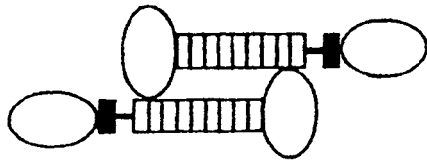


FIG.11B



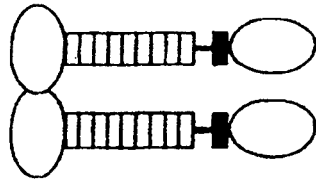
trans

FIG.12A



trans

FIG.12B



cis

FIG.12C

ES 2 181 674 T5

GAATTCGGAG GAATTATTCA AACATAAAC ACAATAAACA ATTTGAGTAG TTGCCGCACA	60
CACACACACA CACAGCCCGT GGATTATTAC ACTAAAAGCG ACACTCAATC CAAAAAATCA	120
GCAACAAAAA CATCAATAAA C ATG CAT TGG ATT AAA TGT TTA TTA ACA GCA	171
Met His Trp Ile Lys Cys Leu Leu Thr Ala	
1 5 10	
TTC ATT TGC TTC ACA GTC ATC GTG CAG GTT CAC AGT TCC GGC AGC TTT	219
Phe Ile Cys Phe Thr Val Ile Val Gln Val His Ser Ser Gly Ser Phe	
15 20 25	
GAG TTG CGC CTG AAG TAC TTC AGC AAC GAT CAC GGG CGG GAC AAC GAG	267
Glu Leu Arg Leu Lys Tyr Phe Ser Asn Asp His Gly Arg Asp Asn Glu	
30 35 40	
GGT CGC TGC TGC AGC GGG GAG TCG GAC GGA GCG ACG GGC AAG TGC CTG	315
Gly Arg Cys Cys Ser Gly Glu Ser Asp Gly Ala Thr Gly Lys Cys Leu	
45 50 55	
GGC AGC TGC AAG ACG CGG TTT CGC GTC TGC CTA AAG CAC TAC CAG GCC	363
Gly Ser Cys Lys Thr Arg Phe Arg Val Cys Leu Lys His Tyr Gln Ala	
60 65 70	
ACC ATC GAC ACC ACC TCC CAG TGC ACC TAC GGG GAC GTG ATC ACG CCC	411
Thr Ile Asp Thr Thr Ser Gln Cys Thr Tyr Gly Asp Val Ile Thr Pro	
75 80 85 90	
ATT CTC GGC GAG AAC TCG GTC AAT CTG ACC GAC GCC CAG CGC TTC CAG	459
Ile Leu Gly Glu Asn Ser Val Asn Leu Thr Asp Ala Gln Arg Phe Gln	
95 100 105	
AAC AAG GGC TTC ACG AAT CCC ATC CAG TTC CCC TTC TCG TTC TCA TGG	507
Asn Lys Gly Phe Thr Asn Pro Ile Gln Phe Pro Phe Ser Phe Ser Trp	
110 115 120	

FIG. 13A

ES 2 181 674 T5

CCG GGT ACC TTC TCG CTG ATC GTC GAG GCC TGG CAT GAT ACG AAC AAT	555
Pro Gly Thr Phe Ser Leu Ile Val Glu Ala Trp His Asp Thr Asn Asn	
125 130 135	
AGC GGC AAT GCG CGA ACC AAC AAG CTC CTC ATC CAG CGA CTC TTG GTG	603
Ser Gly Asn Ala Arg Thr Asn Lys Leu Leu Ile Gln Arg Leu Leu Val	
140 145 150	
CAG CAG GTA CTG GAG GTG TCC TCC GAA TGG AAG ACG AAC AAG TCG GAA	651
Gln Gln Val Leu Glu Val Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu	
155 160 165 170	
TCG CAG TAC ACG TCG CTG GAG TAC GAT TTC CGT GTC ACC TGC GAT CTC	699
Ser Gln Tyr Thr Ser Leu Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu	
175 180 185	
AAC TAC TAC GGA TCC GGC TGT GCC AAG TTC TGC CCG CCC CGC GAC GAT	747
Asn Tyr Tyr Gly Ser Gly Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp	
190 195 200	
TCA TTT GGA CAC TCG ACT TGC TCG GAG ACG GGC GAA ATT ATC TGT TTG	795
Ser Phe Gly His Ser Thr Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu	
205 210 215	
ACC GGA TGG CAG GGC GAT TAC TGT CAC ATA CCC AAA TGC GCC AAA GGC	843
Thr Gly Trp Gln Gly Asp Tyr Cys His Ile Pro Lys Cys Ala Lys Gly	
220 225 230	
TGT GAA CAT GGA CAT TGC GAC AAA CCC AAT CAA TGC GTT TGC CAA CTG	891
Cys Glu His Gly His Cys Asp Lys Pro Asn Gln Cys Val Cys Gln Leu	
235 240 245 250	
GGC TGG AAG GGA GCC TTG TGC AAC GAG TGC GTT CTG GAA CCG AAC TGC	939
Gly Trp Lys Gly Ala Leu Cys Asn Glu Cys Val Leu Glu Pro Asn Cys	
255 260 265	

FIG.13B

ES 2 181 674 T5

ATC CAT GGC ACC TGC AAC AAA CCC TGG ACT TGC ATC TGC AAC GAG GGT	987
Ile His Gly Thr Cys Asn Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly	
270 275 280	
TGG GGA GGC TTG TAC TGC AAC CAG GAT CTG AAC TAC TGC ACC AAC CAC	1035
Trp Gly Gly Leu Tyr Cys Asn Gln Asp Leu Asn Tyr Cys Thr Asn His	
285 290 295	
AGA CCC TGC AAG AAT GGC GGA ACC TGC TTC AAC ACC GGC GAG GGA TTG	1083
Arg Pro Cys Lys Asn Gly Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu	
300 305 310	
TAC ACA TGC AAA TGC GCT CCA GGA TAC AGT GGT GAT GAT TGC GAA AAT	1131
Tyr Thr Cys Lys Cys Ala Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn	
315 320 325 330	
GAG ATC TAC TCC TGC GAT GCC GAT GTC AAT CCC TGC CAG AAT GGT GGT	1179
Glu Ile Tyr Ser Cys Asp Ala Asp Val Asn Pro Cys Gln Asn Gly Gly	
335 340 345	
ACC TGC ATC GAT GAG CCG CAC ACA AAA ACC GGC TAC AAG TGT CAT TGC	1227
Thr Cys Ile Asp Glu Pro His Thr Lys Thr Gly Tyr Lys Cys His Cys	
350 355 360	
GCC AAC GGC TGG AGC GGA AAG ATG TGC GAG GAG AAA GTG CTC ACG TGT	1275
Ala Asn Gly Trp Ser Gly Lys Met Cys Glu Glu Lys Val Leu Thr Cys	
365 370 375	
TCG GAC AAA CCC TGT CAT CAG GGA ATC TGC CGC AAC GTT CGT CCT GGC	1323
Ser Asp Lys Pro Cys His Gln Gly Ile Cys Arg Asn Val Arg Pro Gly	
380 385 390	
TTG GGA AGC AAG GGT CAG GGC TAC CAG TGC GAA TGT CCC ATT GGC TAC	1371
Leu Gly Ser Lys Gly Gln Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr	
395 400 405 410	

FIG. 13C

ES 2 181 674 T5

AGC GGA CCC AAC TGC GAT CTC CAG CTG GAC AAC TGC AGT CCG AAT CCA Ser Gly Pro Asn Cys Asp Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro 415 420 425	1419
TGC ATA AAC GGT GGA AGC TGT CAG CCG AGC GGA AAG TGT ATT TGC CCA Cys Ile Asn Gly Gly Ser Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro 430 435 440	1467
GCG GGA TTT TCG GGA ACG AGA TGC GAG ACC AAC ATT GAC GAT TGT CTT Ala Gly Phe Ser Gly Thr Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu 445 450 455	1515
GGC CAC CAG TGC GAG AAC GGA GGC ACC TGC ATA GAT ATG GTC AAC CAA Gly His Gln Cys Glu Asn Gly Gly Thr Cys Ile Asp Met Val Asn Gln 460 465 470	1563
TAT CCG TGC CAA TGC GTT CCC GGT TTC CAT GGC ACC CAC TGT AGT AGC Tyr Arg Cys Gln Cys Val Pro Gly Phe His Gly Thr His Cys Ser Ser 475 480 485 490	1611
AAA GTT GAC TTG TGC CTC ATC AGA CCG TGT GCC AAT GGA GGA ACC TGC Lys Val Asp Leu Cys Leu Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys 495 500 505	1659
TTG AAT CTC AAC AAC GAT TAC CAG TGC ACC TGT CGT GCG GGA TTT ACT Leu Asn Leu Asn Asn Asp Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr 510 515 520	1707
GGC AAG GAT TGC TCT GTG GAC ATC GAT GAG TGC AGC AGT GGA CCC TGT Gly Lys Asp Cys Ser Val Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys 525 530 535	1755
CAT AAC GGC GGC ACT TGC ATG AAC CGC GTC AAT TCG TTC GAA TGC GTG His Asn Gly Gly Thr Cys Met Asn Arg Val Asn Ser Phe Glu Cys Val 540 545 550	1803

FIG. 13D



ES 2 181 674 T5

GCA GCG GCG GCG GCG GCA GCA GCG GCG GAC GAG TGT CTC ATG TAC GCC Ala Ala Ala Ala Ala Ala Ala Ala Ala Asp Glu Cys Leu Met Tyr Gly 700 705 710	2283
GGA TAT GTG GCC TCG GTG GCG GAT AAC AAC AAT GCC AAC TCA GAC TTT Gly Tyr Val Ala Ser Val Ala Asp Asn Asn Asn Ala Asn Ser Asp Phe 715 720 725 730	2331
TGT GTG GCT CCG CTA CAA AGA GCC AAG TCG CAA AAG CAA CTC AAC ACC Cys Val Ala Pro Leu Gln Arg Ala Lys Ser Gln Lys Gln Leu Asn Thr 735 740 745	2379
GAT CCC ACG CTC ATG CAC CGC GGT TCG CCG GCA GGC AGC TCA GCC AAG Asp Pro Thr Leu Met His Arg Gly Ser Pro Ala Gly Ser Ser Ala Lys 750 755 760	2427
GGA GCG TCT GGC GGA GGA CCG GGA GCG GCG GAG GGC AAG AGG ATC TCT Gly Ala Ser Gly Gly Gly Pro Gly Ala Ala Glu Gly Lys Arg Ile Ser 765 770 775	2475
GTT TTA GGC GAG GGT TCC TAC TGT AGC CAG CGT TGG CCC TCG TTG GCG Val Leu Gly Glu Gly Ser Tyr Cys Ser Gln Arg Trp Pro Ser Leu Ala 780 785 790	2523
GCG GCG GGA GTG GCC GGA GCC TGT TCA TCC CAG CTA ATG GCT GCA GCT Ala Ala Gly Val Ala Gly Ala Cys Ser Ser Gln Leu Met Ala Ala Ala 795 800 805 810	2571
TCG GCA GCG GGC AGC GGA GCG GGG ACG GCG CAA CAG CAG CGA TCC GTG Ser Ala Ala Gly Ser Gly Ala Gly Thr Ala Gln Gln Gln Arg Ser Val 815 820 825	2619
GTC TGC GGC ACT CCG CAT ATG TAACTCCAAA AATCCGGGAG GGCTCCTGGT Val Cys Gly Thr Pro His Met 830	2670
AAATCCGGAG AAATCCGCAT GGAGGAGCTG ACAGCACATA CACAAAGAAA AGACTGGGTT	2730
GGGTTCAAAA TGTGAGAGAG ACGCCAAAAT GTTGTGTGTG ATTGAAGCAG TTTAGTCGTC	2790
ACGAAAAATG AAAATCTGT AACAGGCATA ACTCGTAAAC TCCCTAAAAA ATTTGTATAG	2850
TAATTAGCAA AGCTGTGACC CAGCCGTTTC GATCCCGAAT TC	2892

FIG. 13F

2889 GATC:TACTAC

GAGGAGGTTAAGGAGAGCTATGTGGGCGAGCGACGCGAATACGATCCCCACATCACCGATCCCAGGGTC  
 ACACGCATGAAGATGGCCGGCCTGAAGCCCAACTCCAAATACCGCATCTCCATCACTGCCACCACGAAA  
 ATGGGCGAGGGATCTGAACACTATATCGAAAAGACCACGCTCAAGGATGCCGTCAATGTGGCCCTGCC  
 ACGCCATCTTTCTCCTGGGAGCAACTGCCATCCGACAATGGACTAGCCAAGTTCCGCATCAACTGGCTG  
 CCAAGTACCGAGGGTCATCCAGGCACTCACTTCTTTACGATGCACAGGATCAAGGGCGAAACCCAATGG  
 ATACGCGAGAATGAGGAAAAGAACTCCGATTACCAGGAGGTGGGTGGCTTAGATCCGGAGACCGCTAC  
 GAGTTCGGGTGGTGTCCGTGGATGGCCACTTTAACACGGAGAGTGGCACGCAGGAGATCGACACGAAC  
 ACCGTTGAGGGACCAATAATGGTGGCCAACGAGACGCTGGCCAATGCCGGATGGTTCATTGSCATGATG  
 CTGGCCCTGGCCTTCATCATCATCCTCTTCATCATCATCTGCATTATCCGACGCAATCGGGGCGGAAAG  
 TACGATGTCCAGATCGGGAGCTGGCCAACGGCCGGCGGATTATCCCGAAGAGGGCGGATTCACAGAG  
 TACTCGCAACCGTTGGATAACAAGAGCGCTGGTCGCCAATCCGTGAGTTCAGCGAACAACCGGGCGTG  
 GAAAGCGATACTGATTGATGGCCGAATACGGTGATGGCGATACAGGACAATTTACCGAGGATGGCTCC  
 TTCATTGGCCAATATGTTCCCTGGAAAGCTCCAACCGCCGTTAGCCACAGCCACTGAAC/AATTCGGCT  
 GCGGGCATCAGGGGGCCCAACTGCCGGAGGATCGGGAGCAGCCGGATCGGCAGCAGCAICCGGAGCA  
 TCGGGTGGAGCATCGTCCGGGAGGAGCAGCTGCCAGCAATGGAGGAGCTGCAGCCGGAGCCGTGGCC  
 ACCTACGTCTAAGCTTGGTACC 3955

FIG. 14

1 CCGAGTCGAGCGCGGTGCTCGAGGGGTGATGAGCCCTTTTCTGTCAACGCTAAGATC  
 121 AAGCACATAAGGTCCATATAAATAATAAATAAATGTTGTTGATACCAACATTAT  
 241 GGCCGTTATTCAGCTATCCAGCAAGTGTAGTGGCAAAATAGAACAAACAAGGCA  
 361 CAATCCAGAGTGAATCCGAACAACAACTCCATCTAGATGCCAACCCAGCATCAGCTCGCA  
 481 TCGTCGTTGGAGTCAACAATAGAAATCAGCAGACAGCCCTGGGAATGTCCAAGAACGGCG  
 SerSerLeuGluSerThrIleGluSerAlaAspSerLeuGlyMetSerLysLysThrAla  
 601 CCGGATTGTCGATCATTAAAGTCGCTGCAACTTAATTGCTTTAATTTAATACTGTTA  
 ArgAspCysArgSerLeuLysSerAlaCysAsnLeuIleAlaLeuIleLeuIleLeuLeu  
 721 AACAGCCATCTACTCAACGGCTATTGCTGGGGCATGCCAGCGGAACCTAGGGCCACCAAG  
 AsnSerHisLeuLeuAsnGlyTyrCysCysGlyMetProAlaGluLeuArgAlaThrLys  
 841 ACCGAGCAGGGTGCCAGCATATCCACGGGCTGTTGCTTGGCAAGGCCACCACCAAGATA  
 ThrGluGlnGlyAlaSerIleSerThrGlyCysSerPheGlyAsnAlaThrThrLysIle  
 961 ACGTTTCGTTGGACGAAGTCGTTTACGCTGATACTGCAGGCGTTGGATATGTACAACACA  
 ThrPheArgTrpThrLysSerPheThrLeuIleLeuGlnAlaLeuAspMetTyrAsnThr  
 1081 TCGCCGGAGTGGAGACGCTGGACCACATCGGGGGAAACGGCGGATCACCTACCGTGTG  
 SerProGluTrpLysThrLeuAspHisIleGlyArgAsnAlaArgIleThrTyrArgVal  
 1201 GACGATCAGTTCGGTCACTACGCCCTGGGGCTCCGAGGGTCAGAGCTCTGCCTGAATGGC  
 AspAspGlnPheGlyHisThrAlaCysGlySerGluGlyGlnLysLeuCysLeuAsnGly

FIG. 15A

TACAAACATCAGCGCCTATCAAGTGGAAAGTCAAGTGTGAACAAAACAAAACGAGAG  
 CCAAAACAAAACAAAACGAAAGGCAAGTGGAGAAAATGATACAGCATCCAGAGTAC  
 CCAAAATCTGCATACATGGCTAATAAGGCTGCCAGCGAATTTACATTTGGTGGTGC  
 AACGCCCCAGAAATGTACAAAATGTTAGGAAACATTTGCGCGAAAACCCAGCTACGTGG  
 MetPheArgLysHisPheArgArgLysProAlaThrSer 13  
 ACAAAGGCGAGGTCGGAGGCATCGGTACCCAAAATCGGACCCCTGCCATCGACGGATC  
 ThrLysArgGlnArgProArgHisArgValProLysIleAlaThrLeuProSerThrIle 53  
 GTCCATAAGATATCCGCAGCTGGTAACCTCGAGCTGGAATATTAGAAATCTCAAATACC  
 ValHisLysIleSerAlaAlaGlyAsnPheGluLeuGluIleLeuGluIleSerAsnThr 93  
 -----  
 n° 1  
 ACGATAGGCTGCTCGCCATGCACGAGGCAATCCGGCTGTGCCGTGAGGAGTACCAGACC  
 ThrIleGlyCysSerProCysThrThralaPheArgLeuCysLeuLysGluTyrGlnThr 133  
 n° 2  
 CTGGGTGGCTCCAGGTTTGTGCTCAGCGATCCGGGTGTGGAGCCATTGTGCTGCCCTTT  
 LeuGlyGlySerSerPheValLeuSerAspProGlyValGlyAlaIleValLeuProPhe 173  
 TCCTATCCAGATGGGAGAGGTTAATTGAGGAAACATCATCTGGGGGTGATACTGCCG  
 SerTyrProAspAlaGluArgLeuIleGluGluThrSerTyrSerGlyValIleLeuPro 213  
 n° 4  
 CGGGTGCAATGGCCGTTACCTACTACAAACACGACCTGCACGACCTTTGCCGTCCGGCG  
 ArgValGlnCysAlaValThrTyrTyrAsnThrThrCysThrThrPheCysArgProArg 253  
 TGGCAGGGCGTCAACTGGGAGGGCCATATGCAAGCGGGGCTGGACCCCGTCCACGGC  
 TrpGlnGlyValAsnCysGluGluAlaIleCysLysAlaGlyCysAspProValHisGly 293

FIG. 15B



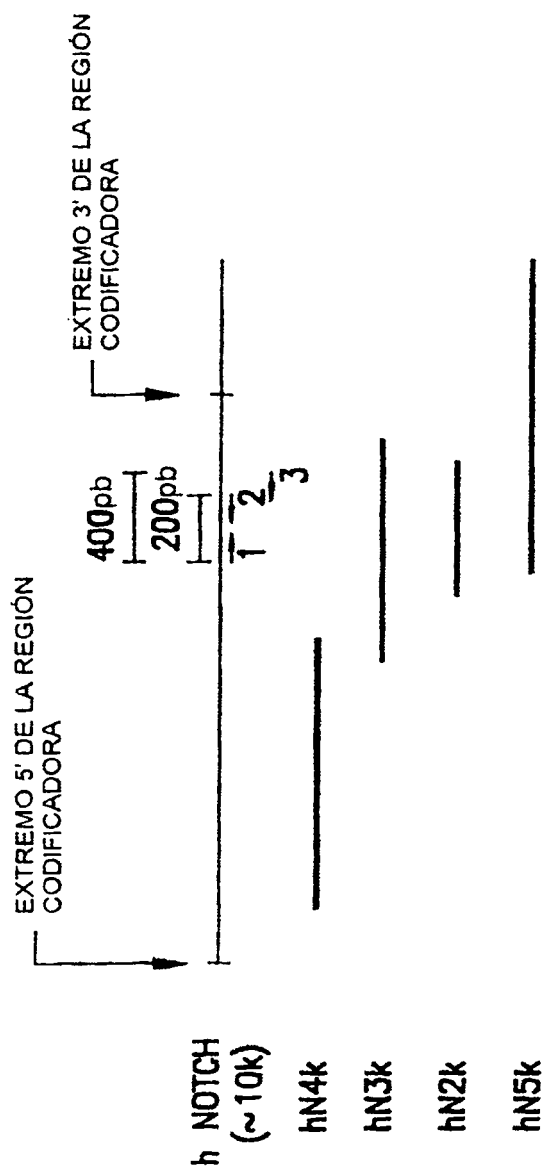
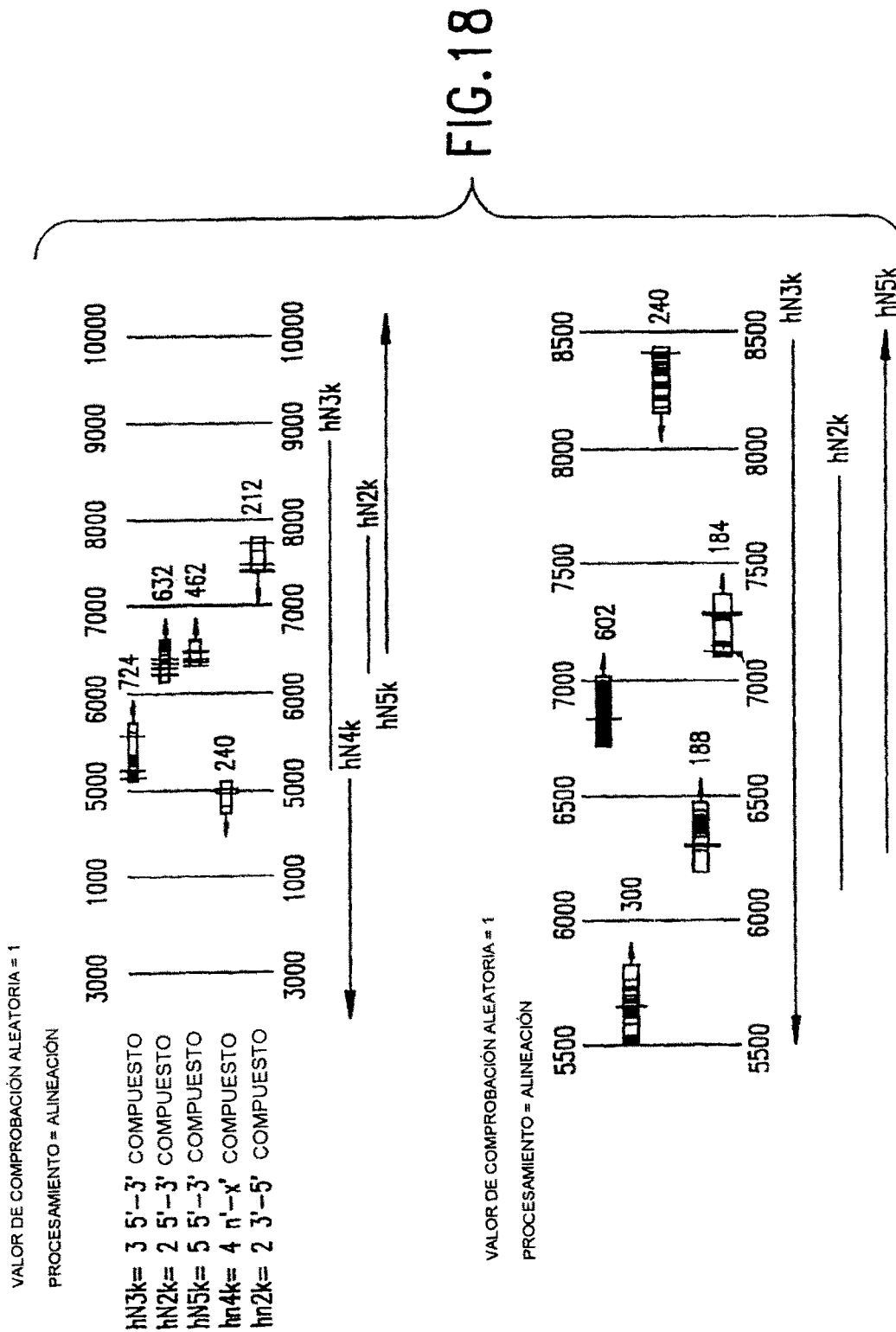


FIG.17



1 GAATTCGGCT GGGAGAATGG TCTGAGCTAC CTGCCCGTCC TGCTGGGGCA TCAATGGCAA  
 61 GTGGGGAAAG CCACACTGGG CAAACGGGCC AGGCCATTTC TGGAAATGTGG TACATGGTGG  
 121 GCAGGGGGCC CGCAACAGCT GGAGGGCAGG TGGACTGAGG CTGGGGATCC CCGCTGGTT  
 181 GGGCAATACT GCCTTTACCC ATGAGCTGGA AAGTCACAAT GGGGGGCAAG GGCTCCCGAG  
 241 GGTGGTTATG TGCTTCCTTC AGGTGGC

## FIG.19A

1 GAATTCCTTC CATTATACGT GACTTTTCTG AAACGTAGC CACCCTAGTG TCTCTAACTC  
 61 CCTCTGGAGT TTGTCAGCTT TGGTCTTTTC AAAGAGCAGG CTCTCTTCAA GTCCTTAAAT  
 121 GCGGGCATGC TCCAGTTTGG TCTGCGTCTC AAGATCACCT TTGGTAATTG ATTCTTCTTC  
 181 AACCCGGAAC TGAAGGCTGG CTCTCACCCCT CTAGGCAGAG CAGGAATTCC GAGGTGGATG  
 241 TGTTAGATGT GAATGTCCGT GGCCAGATG GCTGCACCCC ATTGATGTTG GCTTCTCTCC  
 301 GAGGAGGCAG CTCAGATTTG AGTGATGAAG ATGAAGATGC AGAGGACTGT TCTGCTAACA  
 361 TCATCACAGA CTTGGTCTAC CAGGGTGCCA GCCTCCAGNC CAGACAGACC GGACTGGTGA  
 421 GATGGCCCTG CACCTTGCAg CCCGCTACTC ACGGGCTGAT GCTGCCAAGC GTCTCCTGGA  
 481 TGCAGGTGCA GATGCCAATG CCCAGGACAA CATGGGCCCG TGTCCTACTCC ATGCTGCAGT  
 541 GGCACGTGAT GCCAAGGTGT ATTCAGATCT GTTA

## FIG.19B

1 TCCAGATTCT GATTGGCAAC CGAGTAACTG ATCTAGATGC CAGGATGAAT GATGGTACTA  
 61 CACCCCTGAT CCTGGCTGCC CGCCTGGCTG TGGAGGGAAT GGTGGCAGAA CTGATCAACT  
 121 GCCAAGCGGA TGTGAATGCA GTGGATGACC ATGGAAAATC TGCTCTTAC TGGGCAGCTG  
 181 CTGTCAATAA TGTGGAGGCA ACTCTTTTGT TGTGAAAAA TGGGGCCAAC CGAGACATGC  
 241 AGGACAACAA GGAAGAGACA CCTCTGTTTC TTGCTGCCCG GGAGGAGCTA TAAGC

## FIG.19C

1 GAATTCCTGC CAGGAGGACG CCGGCAACAA GGTCTGCAGC CTGCAGTGCA ACAACCCAGC  
 61 GTGCGGCTGG GACGGCGGTG ACTGCTCCCT CAACTTCACA ATGACCCCTG GAGAACTGC  
 121 ACGCAGTCTC TGCAGTGCTG GAAGTACTTC AGTGACGGCC ACTGTGACAC CCACTGCAAC  
 181 TCAGCCGGCT GCCTCTCGA CCGCTTTGAC TGCCAGCGGC GGAAGGCCAG TTGCAACCCC  
 241 CTGTACGACC AGTACTGCAA GGACCACTTG AGCGACGGGC ACTGCGACCA GGGCTGCAAC  
 301 AGCGCGGAGT NCAGNTGGGA CCGGCTGGAC TGTGCGGCAG TGTACCCGAG AACTGGCGGC  
 361 GCACGCTGGT GGTGGTGGTG CTGATGCCGC CCGAGCAGCT GCGCAACAGC TCCTTCCACT  
 421 TCCTGCGGGA CGTCAGCCGC GTGCTGCACA CCAACGTGTC TTCAGCGTG ACGCACACGG  
 481 CCAGCAGATG ATGTTCCCTT ACTACGGCCG CAGAGGAGAG CTGCGCAAGC CCCATCAAGC  
 541 GTGCCGCCGA GGGCTGGGCC GCACCTGACG CCTGCTGGGC CA

FIG.20A

1 GAATTCCTGC CAGGAGGACG CCGGCAACAA GGTCTGCAGC CTGCAGTGCA ACAACCCAGC  
 61 GTGCGGCTGG GACGGCGGTG ACTGCTCCCT CAACTTCACA ATGACCCCTG GAGAACTGC  
 121 ACGCAGTCTC TGCAGTGCTG GAAGTACTTC AGTGACGGCC ACTGTGACAC CCACTGCAAC  
 181 TCAGCCGGCT GCCTCTCGA CCGCTTTGAC TGCCAGCGGC GGAAGGCCAG TTGCAACCCC  
 241 CTGTACGACC AGTACTGCAA GGACCACTTG AGCGACGGGC ACTGCGACCA GGGCTGCAAC  
 301 AGCGCGGAGT NCAGNTGGGA CCGGCTGGAC TGTGCGGCAG TGTACCCGAG AACTGGCGGC  
 361 GCACGCTGGT GGTGGTGGTG CTGATGCCGC CCGAGCAGCT GCGCAACAGC TCCTTCCACT  
 421 TCCTGCGGGA CGTCAGCCGC GTGCTGCACA CCAACGTGTC TTCAGCGTG ACGCACACGG  
 481 CCAGCAGATG ATGTTCCCTT ACTACGGCCG CAGAGGAGAG CTGCGCAAGC CCCATCAAGC  
 541 GTGCCGCCGA GGGCTGGGCC GCACCTGACG CCTGCTGGGC CA

FIG.20B

1 TCAGCCGAGT GCTGCACACC AACGTGTCTT CAAGCGTGAC GCACACGGCC AGCAGATGAT  
 61 GTTCCCTTAC TACGCGCCGC AGGAGGAGCT GCGCAAGCCC CATCAAGCGT GCGCGCGAGG  
 121 GCTGGGCCGC ACCTGACGCC TGCTGGGCCA

FIG.20C

1 TTACCATTAC AGGAACAGGC ACCTGTAGCT GGTGGCTGGG GGTGTTGTCC ACAGGCGAGG  
 61 AGTAGCTGTG CTGCGAGGGG GCGTCAAGG ACTGGGCTGC GGTACGGGT GGGACCAGCG  
 121 AGGATGGCAG CGACGTGGGC AGGGCGGGGC TCTCCTGGGG CAGAATAGTG TGCACCGCCA  
 181 GCTGCTGGGG CCCAGTGTG CAGCTCTGCC TGGCTCGGCT CTCCACTCAG GAAGCTCCGG  
 241 CCCAGGT

FIG.20D

1 GAATTCATT CAGGAGGAAA GGGTGGGGAG AGAAGCAGGC ACCCACTTTC CCGTGGCTGG  
61 ACTCGTTCCC AGGTGGCTCC ACCGGCAGCT GTGACCGCCG CAGGTGGGGG CCGAGTGCCA  
121 TTCAGAAAAT TCCAGAAAAG CCTACCCCA ACTCGGACGG CAACGTCACA CCCGTGGGTA  
181 GCAACTGGCA CACAAACAGC CAGCGTGTCT GGGGCACGGG GGGATGGCAC CCCCTGCAGG  
241 CAGAGCTG

FIG.21A

1 CTAAGGGAA CAAAAGCNGG AGCTCCACCG CGGGCGGCNC NGCTCTAGAA CTAGTGGANN  
61 NCCCGGGCTG CAGGAATTCC GCGGACTGG GCTCGGGCTC AGAGCGGCGC TGTGGAAGAG  
121 ATTCTAGACC GGGAGAACA GCGAATGGCT GACAGCTGGC CTCCAAAGTC ACCAGGCTCA  
181 AATCGCTCGC CCTGGACATC GAGGGATGCA GAGGATCAGA ACCGGTACCT GGATGGCATG  
241 ACTCGGATTT ACAAGCATGA CCAGCCTGCT TACAGGGAGC GTGANNTTTT CACATGCAGT  
301 CGACAGACAC GAGCTCTATG CAT

FIG.21B

1 GAATTCGAG GTGGATGTGT TAGATGTGAA TGTCCGTGGC CCAGATGGCT GCACCCCAT  
61 GATGTTGGCT TCTCTCCGAG GAGGCAGCTC AGATTTGAGT GATGAAGATG AAGATGCAGA  
121 GGACTCTTCT GCTAACATCA TCACAGACTT GGTCTTACCA GGGTGCCAGC CTTCCAGGCC  
181 CAAGAACAGA CCGGACTTGG TGAGATGGCC CTGCACCTTG CAGCCCGCTA CTACGGGCTG  
241 ATGCTGCCAA GGTTCCTGGAT GCAGGTGCAG ATGCCAATGC CCAGGACAAC ATGGGCCGCT  
301 GTCCACTCCA TGCTGCAGTG GCACTGATGC

## FIG.22A

1 CAGAGGATGG TGAGGGTCCA TGCAGATAGG TTCTCCCAT CCTGTGAATA ATAAATGGGT  
61 GCAAGGGCAG AGAGTACCA TTTAGAATGA TAAAATGTTT GCACACTATG AAGAGGCTG  
121 ACAGAATGTT GCCACATGGA GAGATAAAGC AGAGAATGAA CAAACTT

## FIG.22B

1 AGGATGAATG ATGGTACTAC ACCCCTGATC CTGGCTGCC GCCTGGCTGT GGAGGGAAATG  
61 GTGGCAGAAC TGATCAACTG CCAAGCGGAT GTGAATGCAG TGGATGACCA TGGAAAATCT  
121 GCTCTTCACT GGGCAGCTGC TGCAATAAT GTGGAGGCAA CTCTTTTGT GTTGAAAAAT  
181 GGGCCAACC GAGACATGCA GGACAACAAG GAAGAGACAC CTCTG

## FIG.22C

1 AATAATAAAT GGGTGCAAGG GCAGAGATC ACCATTTAGA ATGATAAAAT GTTGCACAC  
61 TATGAAAGAG GCTGACAGAA TGTTGCCACA TGGAGAGATA AAGCAGAGAA TGAACAACT  
121 T

## FIG.22D

```

*      *      *      *      *      *      *      *      *      *
10  *      *      *      *      *      *      *      *      *      *
   *      *      *      *      *      *      *      *      *      *
TGC CAG GAG GAC GCG GGC AAC AAG AAG GTC TGC AGC CTG CAG TGC AAC AAC
C  Q  E  D  A  G  N  K  V  C  C  S  L  Q  C  C  N  N  >

50  *      *      *      *      *      *      *      *      *      *
   *      *      *      *      *      *      *      *      *      *
CAG GCG TGC GGC TGG GAC GGC GGT GAC TGC TCC CTC AAC TTC AAT GAC
H  A  C  G  W  D  G  G  D  C  C  S  L  N  F  N  D  >

100 *      *      *      *      *      *      *      *      *      *
   *      *      *      *      *      *      *      *      *      *
CCC TGG AAG AAC TGC ACG CAG TCT CTG CAG TGC TGG AAG TAC TTC AGT
P  W  K  N  C  T  Q  S  L  Q  C  W  K  Y  F  S  >

150 *      *      *      *      *      *      *      *      *      *
   *      *      *      *      *      *      *      *      *      *
GAC GGC CAC TGT GAC AGC CAG TGC AAC TCA GCC GGC TGC CTC TTC GAC
D  G  H  C  D  S  Q  C  N  S  A  G  C  L  F  D  >

```

FIG.23A

```

200 * * * * * 210 * * * * * 220 * * * * * 230 * * * * * 240 * *
    * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
    GGC TTT GAC TGC CAG CGT GCG GAA GGC CAG TGC AAC CCC CTG TAC GAC
    G F D C Q R A E G Q C N P L Y D> -

250 * * * * * 260 * * * * * 270 * * * * * 280 * * * * *
    * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
    CAG TAC TGC AAG GAC CAC TTC AGC GAC GGG CAC TGC GAC CAG GGC TGC
    Q Y C K D H F S D G H C D Q G C>

290 * * * * * 300 * * * * * 310 * * * * * 320 * * * * * 330 * * * * *
    * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
    AAC AGC GCG GAG TGC GAG TGG GAC GGC CTG GAC TGT GCG GAG CAT GTA
    N S A E C E W D G L D C A E H V>

340 * * * * * 350 * * * * * 360 * * * * * 370 * * * * * 380 * * * * *
    * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
    CCC GAG AGG CTG GCG GCC GGC ACG CTG GTG GTG GTG CTG ATG CCG
    P E R L A A G T L V V V V L M P>

```

FIG.23B

```

390 * * * * * 400 * * * * * 410 * * * * * 420 * * * * * 430 *
    CCG GAG CAG CTG CGC AAC AGC ICC TTC CAC TTC CCG GAG CTC AGC
    P E Q L R N S S F H F L R E L S>

    440 * * * * * 450 * * * * * 460 * * * * * 470 * * * * * 480 *
    CGC GTG CTG CAC ACC AAC GTG GTC TTC AAG CGT GAC GCA CAC GGC CAG
    R V L H T N V V F K R D A H G Q>

    490 * * * * * 500 * * * * * 510 * * * * * 520 * * * * *
    CAG ATG ATC TTC CCC TAC TAC GGC CGC GAG GAG GAG CTG CGC AAG CAC
    Q M I F P Y Y G R E E L R K HD>

    530 * * * * * 540 * * * * * 550 * * * * * 560 * * * * * 570 * * * * *
    CCC ATC AAG CGT GCC GAG GGC TGG GCC GCA CCT GAC GCC CTG CTG
    P I K R A A E G W A A P D A L L>

```

FIG.23C

```

580 * * * * * 590 * * * * * 600 * * * * * 610 * * * * * 620 * *
    GGC CAG GTG AAG GCC TCG CTG CTC CCT GGT GGC AGC GAG GGT GGG CGG
    G Q V K A S L L L P G G S E G G R>

    630 * * * * * 640 * * * * * 650 * * * * * 660 * * * * * 670 * *
    CCG CGG AGG GAG CTG GAC CCC ATG GAC GTC CGC GGC TCC ATC GTC TAC
    R R E L D P M D V R G S I V Y>

    680 * * * * * 690 * * * * * 700 * * * * * 710 * * * * * 720 * *
    CTG GAG ATT GAC AAC CCG CAG TGT GTG CAG GCC TCC TCG CAG TGC TTC
    L E I D N R Q C V Q A S S Q C F>

    730 * * * * * 740 * * * * * 750 * * * * * 760 * * * * *
    CAG AGT GCC ACC GAC GTG GCC GCA TTC CTG GGA GGC CTC GCC TCG CTG
    Q S A T D V A A A F L L G A L A S L>

```

FIG.23D



```

970      *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
AAA GTG TCT GAG GCC AGC AAG AAG AAG AAG CGG CGG GAG CCC CTC GGC GAG
K  V  S  E  A  S  K  K  K  K  K  R  R  E  P  L  G  E>

1010     1020     1030     1040     1050
      *      *      *      *      *
GAC TCC GTG GGC CTC AAG CCC CTG AAG AAC GCT TCA GAC GGT GCC CTC
D  S  V  G  L  K  P  L  K  N  A  S  D  G  A  L>

1060     1070     1080     1090     1100
      *      *      *      *      *
ATG GAC GAC AAC CAG AAT GAG TGG GGG GAC GAG GAC CTG GAG ACC AAG
M  D  D  N  Q  N  E  W  G  D  E  D  L  E  T  K>

1110     1120     1130     1140     1150
      *      *      *      *      *
AAG TTC CCG TTC GAG GAG CCC GTG GTT CTG CCT GAC CTG GAC GAC CAG
K  F  R  F  E  E  P  V  V  L  P  D  L  D  D  Q>

```

FIG.23F

```

1160  * 1170  * 1180  * 1190  * 1200  *
ACA GAC CAC CGG CAG TGG ACT CAG CAC CTG GAT GCC GCT GAC CTG
T D H R Q W T Q Q H L D A A D L>

1210  * 1220  * 1230  * 1240  *
CGC ATG TCT GCC ATG GCC CCC ACA CCG CCC CAG GGT GAG GTT GAC GCC
R M S A M A P T P P Q G E V D A>

1250  * 1260  * 1270  * 1280  * 1290  *
GAC TGC ATG GAC GTC AAT GTC CGC GGG CCT GAT GGC TTC ACC CCG CTC
D C M D V N V R G P D G F T P L>

1300  * 1310  * 1320  * 1330  * 1340  *
ATG ATC GCC TCC TGC AGC GGC GGC GGC CTG GAG ACG GGC AAC AGC GAG
M I A S C S G G G L E T G N S E>

```

FIG. 23G

```

1350 * * * * * 1360 * * * * * 1370 * * * * * 1380 * * * * * 1390 * *
GAA GAG GAG GAC GCG CCG GCC GTC ATC TCC GAC TTC ATC TAC CAG GGC
E E E D A P A V I S D F I Y Q G>

1400 * * * * * 1410 * * * * * 1420 * * * * * 1430 * * * * * 1440 * *
GCC AGC CTG CAC AAC CAG ACA GAC CGC ACG GGC GAG ACC GCC TTG CAC
A S L H N Q T D R T G E T A L H>

1450 * * * * * 1460 * * * * * 1470 * * * * * 1480 * * * * *
CTG GCC GCC CGC TAC TCA CGC TCT GAT GCC GCC AAG CGC CTG CTG GAG
L A A R Y S R S D A A K R L L E>

1490 * * * * * 1500 * * * * * 1510 * * * * * 1520 * * * * * 1530 * *
GCC AGC GCA GAT GCC AAC ATC CAG GAG AAC ATG GGC CGC ACC CCG CTG
A S A D A N I Q D N M G R T P L>

```

FIG. 23H

```

1540 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAT GCG GCT GTG TCT GCC GAC GCA CAA GGT GTC TTC CAG ATC CTG ATC
H  A  A  V  S  A  D  A  Q  G  V  F  Q  I  L  I >

      1590      1600      1610      1620      1630
*      *      *      *      *
CGG AAC CGA GCC ACA GAC CTG GAT GCC CGC ATG CAT GAT GGC ACG ACG
R  N  R  A  T  D  L  D  A  R  M  H  D  G  T  I >

      1640      1650      1660      1670      1680
*      *      *      *      *
CCA CTG ATC CTG GCT GCC CGC CTG GCC GTG GAG GGC ATG CTG GAG GAC
P  L  I  L  A  A  R  L  A  V  E  G  M  L  E  D >

      1690      1700      1710      1720
*      *      *      *
CTC ATC AAC TCA CAC GCC GAC GTC AAC GCC GTA GAT GAC CTG GGC AAG
L  I  N  S  H  A  D  V  N  A  V  D  D  L  G  K >

      1730      1740      1750      1760      1770
*      *      *      *      *
TCC GCC CTG CAC TGG GCC GCC GCG AAC AAT GTG GAT GCC GCA GTT
S  A  L  H  W  A  A  A  V  N  N  V  D  A  A  V >

```

FIG. 231

```

1780      *      1790      *      1800      *      1810      *      1820      *
GIG CTC CTG AAG AAC GGG GCT AAC AAA GAT ATG CAG AAC AAC AGG GAG
V  L  L  K  N  G  A  N  K  D  M  Q  N  N  R  E>

1830      *      1840      *      1850      *      1860      *      1870      *
GAG ACA CCC CTG TTT CTG GCC GCC CGG GAG GGC AGC TAC GAG ACC GCC
E  T  P  L  F  L  A  A  R  E  G  S  Y  E  T  A>

1880      *      1890      *      1900      *      1910      *      1920      *
AAG GTG CTG CTG GAC CAC TTT GCC AAC CGG GAC ATC ACG GAT CAT ATG
K  V  L  L  D  H  F  A  N  R  D  I  T  D  H  M>

1930      *      1940      *      1950      *      1960      *
GAC CGC CTG CCG GAC ATC GCA CAG GAG CGC ATG CAT CAC GAC ATC
D  R  L  P  R  D  I  A  Q  E  R  M  H  H  D  I>

```

FIG.23J

```

1970 1980 1990 2000 2010 2060
* * * * *
GTG AGG CTG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC
V R L L D E Y N L V R S P Q L HD>

2020 2030 2040 2050 2060
* * * * *
GGA GCC CCG CTG GGG GGC AGG CCC ACC CTG TCG CCC CCG CTC TGC TCG
G A P L G G T P T L S P P L C S>

2070 2080 2090 2100 2110
* * * * *
CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG
P N G Y L G S L K P G V Q G K K>

2120 2130 2140 2150 2160
* * * * *
GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC
V R K P S S K G L A C G S K E A>

```

FIG.23K



2360	2370	2380	2390	2400
* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
CCC GAC ACC CAC CTG GGC ATC GGG CAC CTG AAC GTG GCG GCC AAG CCC				
P D T H L G I G H L L N V A A K P>				
2410	2420	2430	2440	
* * * * *	* * * * *	* * * * *	* * * * *	
GAG ATG GCG GCG CTG GGT GGG GGC GGC CCG CIG GCC TTT GAG ACT GGC				
E M A A L G G G R L A F E T G>				
2450	2460	2470	2480	2490
* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
CCA CCT CGT CTC TCC CAC CTG CCT GTG GCC TCT GGC ACC AGC ACC GTC				
P P R L S H L P V A S G T S T V>				
2500	2510	2520	2530	2540
* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
CTG GGC TCC AGC AGC GGA GGG GCC CTG AAT TTC ACT GTG GGC GGG TCC				
L G S S S G G A L N F T V G G S>				

FIG.23 M

```

2550 * 2560 2570 2580 2590
* * * * *
ACC AGT TTG AAT GGT CAA TGC GAG TGG CTG TCC CGG CTG CAG AGC GGC
T S L N G Q C E W L S R L Q S G>

2600 2610 2620 2630 2640
* * * * *
ATG GTG CCG AAC CAA TAC AAC CCT CTG CCG GGG AGT GTG GCA CCA GGC
M V P N Q Y N P L R G S V A P G>

2650 2660 2670 2680
* * * * *
CCC CTG AGC ACA CAG GCC CCC TCC CTG CAG CAT GGC ATG GTA GGC CCG
P L S T Q A P S L Q H G M V G P>

2690 2700 2710 2720 2730
* * * * *
CTG CAC AGT AGC CTT GCT GCC AGC GCC CTG TCC CAG ATG ATG AGC TAC
L H S S L A A S A L S Q M M S Y>

```

FIG.23N

```

2740 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG
Q  G  L  P  S  T  R  L  A  T  Q  P  H  L  V  Q>
      *      *      *      *      *      *      *      *      *      *
2750 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG
Q  G  L  P  S  T  R  L  A  T  Q  P  H  L  V  Q>
      *      *      *      *      *      *      *      *      *      *
2760 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG
Q  G  L  P  S  T  R  L  A  T  Q  P  H  L  V  Q>
      *      *      *      *      *      *      *      *      *      *
2770 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG
Q  G  L  P  S  T  R  L  A  T  Q  P  H  L  V  Q>
      *      *      *      *      *      *      *      *      *      *
2780 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG
Q  G  L  P  S  T  R  L  A  T  Q  P  H  L  V  Q>
      *      *      *      *      *      *      *      *      *      *
2790 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG
T  Q  Q  V  Q  P  Q  N  L  Q  M  Q  Q  Q  N  L>
      *      *      *      *      *      *      *      *      *      *
2800 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG
T  Q  Q  V  Q  P  Q  N  L  Q  M  Q  Q  Q  N  L>
      *      *      *      *      *      *      *      *      *      *
2810 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG
T  Q  Q  V  Q  P  Q  N  L  Q  M  Q  Q  Q  N  L>
      *      *      *      *      *      *      *      *      *      *
2820 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG
T  Q  Q  V  Q  P  Q  N  L  Q  M  Q  Q  Q  N  L>
      *      *      *      *      *      *      *      *      *      *
2830 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG
T  Q  Q  V  Q  P  Q  N  L  Q  M  Q  Q  Q  N  L>
      *      *      *      *      *      *      *      *      *      *
2840 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA
Q  P  A  N  I  Q  Q  Q  Q  S  L  Q  P  P  P  P>
      *      *      *      *      *      *      *      *      *      *
2850 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA
Q  P  A  N  I  Q  Q  Q  Q  S  L  Q  P  P  P  P>
      *      *      *      *      *      *      *      *      *      *
2860 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA
Q  P  A  N  I  Q  Q  Q  Q  S  L  Q  P  P  P  P>
      *      *      *      *      *      *      *      *      *      *
2870 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA
Q  P  A  N  I  Q  Q  Q  Q  S  L  Q  P  P  P  P>
      *      *      *      *      *      *      *      *      *      *
2880 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA
Q  P  A  N  I  Q  Q  Q  Q  S  L  Q  P  P  P  P>
      *      *      *      *      *      *      *      *      *      *
2890 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG
P  P  Q  P  H  L  G  V  S  S  A  A  S  G  H  L>
      *      *      *      *      *      *      *      *      *      *
2900 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG
P  P  Q  P  H  L  G  V  S  S  A  A  S  G  H  L>
      *      *      *      *      *      *      *      *      *      *
2910 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG
P  P  Q  P  H  L  G  V  S  S  A  A  S  G  H  L>
      *      *      *      *      *      *      *      *      *      *
2920 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG
P  P  Q  P  H  L  G  V  S  S  A  A  S  G  H  L>
      *      *      *      *      *      *      *      *      *      *
2930 *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
GCC CGG AGC TTC CTG AGT GGA GAG CCG AGC CAG GCA GAC GIG CAG CCA
      *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
G  R  S  F  L  S  G  E  P  S  Q  A  D  V  Q  P>

```

FIG.23 0

```

2980 * * * * * 2990 * * * * * 3000 * * * * * 3010 * * * * * 3020 *
CTG GGC CCC AGC AGC CTG GCG GTG CAC ACT ATT CTG CCC CAG GAG AGC
L G P S S L A V H T I L P Q E S>

3030 * * * * * 3040 * * * * * 3050 * * * * * 3060 * * * * * 3070 *
* * * * * * * * * * * * * * * * * * * * * * *
CCC GCC CTG CCC ACG TCG CTG CCA TCC TCG CTG GTC CCA CCC GTG ACC
P A L P T S L P S S L V P P V T>

3080 * * * * * 3090 * * * * * 3100 * * * * * 3110 * * * * * 3120 *
* * * * * * * * * * * * * * * * * * * * * * *
GCA GCC CAG TTC CTG ACG CCC CCC TCG CAG CAC AGC TAC TCC TCG CCT
A A Q F L T P P S Q H S Y S S P>

```

FIG.23P



ES 2 181 674 T5

G GAG GTG GAT GTG TTA GAT GTG AAT GTC CGT GGC CCA GAT GGC TGC	46
Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys	
1 5 10 15	
ACC CCA TTG ATG TTG GCT TCT CTC CGA GGA GGC AGC TCA GAT TTG AGT	94
Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser	
20 25 30	
GAT GAA GAT GAA GAT GCA GAG GAC TCT TCT GCT AAC ATC ATC ACA GAC	142
Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp	
35 40 45	
TTG GTC TAC CAG GGT GCC AGC CTC CAG GCC CAG ACA GAC CGG ACT GGT	190
Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly	
50 55 60	
GAG ATG GCC CTG CAC CTT GCA GCC CGC TAC TCA CGG GCT GAT GCT GCC	238
Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala	
65 70 75	
AAG CGT CTC CTG GAT GCA GGT GCA GAT GCC AAT GCC CAG GAC AAC ATG	286
Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met	
80 85 90 95	
GGC CGC TGT CCA CTC CAT GCT GCA GTG GCA GCT GAT GCC CAA GGT GTC	334
Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val	
100 105 110	
TTC CAG ATT CTG ATT CGC AAC CGA GTA ACT GAT CTA GAT GCC AGG ATG	382
Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met	
115 120 125	
AAT GAT GGT ACT ACA CCC CTG ATC CTG GCT GCC CGC CTG GCT GTG GAG	430
Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu	
130 135 140	

FIG.24A

ES 2 181 674 T5

GGA ATG GTG GCA GAA CTG ATC AAC TGC CAA GCG GAT GTG AAT GCA GTG	478
Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val	
145 150 155	
GAT GAC CAT GGA AAA TCT GCT CTT CAC TGG GCA GCT GCT GTC AAT AAT	526
Asp Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn	
160 165 170 175	
GTG GAG GCA ACT CTT TTG TTG TTG AAA AAT GGG GCC AAC CGA GAC ATG	574
Val Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met	
180 185 190	
CAG GAC AAC AAG GAA GAG ACA CCT CTG TTT CTT GCT GCC CGG GAG GGG	622
Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly	
195 200 205	
AGC TAT GAA GCA GCC AAG ATC CTG TTA GAC CAT TTT GCC AAT CGA GAC	670
Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp	
210 215 220	
ATC ACA GAC CAT ATC GAT CGT CTT CCC CGG GAT GTG GCT CGG GAT CGC	718
Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg	
225 230 235	
ATG CAC CAT GAC ATT GTG CGC CTT CTG GAT GAA TAC AAT GTG ACC CCA	766
Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro	
240 245 250 255	
AGC CCT CCA GGC ACC GTG TTG ACT TCT GCT CTC TCA CCT GTC ATC TGT	814
Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys	
260 265 270	
GGG CCC AAC AGA TCT TTC CTC AGC CTG AAG CAC ACC CCA ATG GGC AAG	862
Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys	
275 280 285	

FIG.24B

ES 2 181 674 T5

AAG TCT AGA CGG CCC AGT GCC AAG AGT ACC ATG CCT ACT AGC CTC CCT	910
Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro	
290 295 300	
AAC CTT GCC AAG GAG GCA AAG GAT GCC AAG GGT AGT AGG AGG AAG AAG	958
Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys	
305 310 315	
TCT CTG AGT GAG AAG GTC CAA CTG TCT GAG AGT TCA GTA ACT TTA TCC	1006
Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser	
320 325 330 335	
CCT GTT GAT TCC CTA GAA TCT CCT CAC ACG TAT GTT TCC GAC ACC ACA	1054
Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr	
340 345 350	
TCC TCT CCA ATG ATT ACA TCC CCT GGG ATC TTA CAG GCC TCA CCC AAC	1102
Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn	
355 360 365	
CCT ATG TTG GCC ACT GCC GCC CCT CCT GCC CCA GTC CAT GCC CAG CAT	1150
Pro Met Leu Ala Thr Ala Ala Pro Pro Ala Pro Val His Ala Gln His	
370 375 380	
GCA CTA TCT TTT TCT AAC CTT CAT GAA ATG CAG CCT TTG GCA CAT GGG	1198
Ala Leu Ser Phe Ser Asn Leu His Glu Met Gln Pro Leu Ala His Gly	
385 390 395	
GCC AGC ACT GTG CTT CCC TCA GTG AGC CAG TTG CTA TCC CAC CAC CAC	1246
Ala Ser Thr Val Leu Pro Ser Val Ser Gln Leu Leu Ser His His His	
400 405 410 415	
ATT GTG TCT CCA GGC AGT GGC AGT GCT GGA AGC TTG AGT AGG CTC CAT	1294
Ile Val Ser Pro Gly Ser Gly Ser Ala Gly Ser Leu Ser Arg Leu His	
420 425 430	
CCA GTC CCA GTC CCA GCA GAT TGG ATG AAC CGC ATG GAG GTG AAT GAG	1342
Pro Val Pro Val Pro Ala Asp Trp Met Asn Arg Met Glu Val Asn Glu	
435 440 445	

FIG.24C

ES 2 181 674 T5

ACC CAG TAC AAT GAG ATG TTT GGT ATG GTC CTG GCT CCA GCT GAG GGC	1390
Thr Gln Tyr Asn Glu Met Phe Gly Met Val Leu Ala Pro Ala Glu Gly	
450 455 460	
ACC CAT CCT GGC ATA GCT CCC CAG AGC AGG CCA CCT GAA GGG AAG CAC	1438
Thr His Pro Gly Ile Ala Pro Gln Ser Arg Pro Pro Glu Gly Lys His	
465 470 475	
ATA ACC ACC CCT CGG GAG CCC TTG CCC CCC ATT GTG ACT TTC CAG CTC	1486
Ile Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile Val Thr Phe Gln Leu	
480 485 490 495	
ATC CCT AAA GGC AGT ATT GCC CAA CCA GCG GGG GCT CCC CAG CCT CAG	1534
Ile Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly Ala Pro Gln Pro Gln	
500 505 510	
TCC ACC TGC CCT CCA GCT GTT GCG GGC CCC CTG CCC ACC ATG TAC CAG	1582
Ser Thr Cys Pro Pro Ala Val Ala Gly Pro Leu Pro Thr Met Tyr Gln	
515 520 525	
ATT CCA GAA ATG GCC CGT TTG CCC AGT GTG GCT TTC CCC ACT GCC ATG	1630
Ile Pro Glu Met Ala Arg Leu Pro Ser Val Ala Phe Pro Thr Ala Met	
530 535 540	
ATG CCC CAG CAG GAC GGG CAG GTA GCT CAG ACC ATT CTC CCA GCC TAT	1678
Met Pro Gln Gln Asp Gly Gln Val Ala Gln Thr Ile Leu Pro Ala Tyr	
545 550 555	
CAT CCT TTC CCA GCC TCT GTG GGC AAG TAC CCC ACA CCC CCT TCA CAG	1726
His Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro Thr Pro Pro Ser Gln	
560 565 570 575	
CAC AGT TAT GCT TCC TCA AAT GCT GCT GAG CGA ACA CCC AGT CAC AGT	1774
His Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg Thr Pro Ser His Ser	
580 585 590	
GGT CAC CTC CAG GGT GAG CAT CCC TAC CTG ACA CCA TCC CCA GAG TCT	1822
Gly His Leu Gln Gly Glu His Pro Tyr Leu Thr Pro Ser Pro Glu Ser	
595 600 605	

FIG.24D

ES 2 181 674 T5

CCT GAC CAG TGG TCA AGT TCA TCA CCC CAC TCT GCT TCT GAC TGG TCA	1870
Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser Ala Ser Asp Trp Ser	
610 615 620	
GAT GTG ACC ACC AGC CCT ACC CCT GGG GGT GCT GGA GGA GGT CAG CGG	1918
Asp Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gly Gln Arg	
625 630 635	
GGA CCT GGG ACA CAC ATG TCT GAG CCA CCA CAC AAC AAC ATG CAG GTT	1966
Gly Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val	
640 645 650 655	
TAT GCG TGAGAGAGTC CACCTCCAGT GTAGAGACAT AACTGACTTT TGAAATGCT	2022
Tyr Ala	
GCTGAGGAAC AAATGAAGGT CATCCGGGAG AGAAATGAAG AAATCTCTGG AGCCAGCTTC	2082
TAGAGGTAGG AAAGAGAAGA TGTTCTTATT CAGATAATGC AAGAGAAGCA ATTCGTCAST	2142
TTCACTGGGT ATCTGCAAGG CTTATTGATT ATTCTAATCT AATAAGACAA GTTTGTGGAA	2202
ATGCAAGATG AATACAAGCC TTGGGTCCAT GTTACTCTC TTCTATTTGG AGAATAAGAT	2262
GGATGCTTAT TGAAGCCCAG ACATTCTTGC AGCTTGGACT GCATTTTAAG CCCTGCAGGC	2322
TTCTGCCATA TCCATGAGAA GATTCTACAC TAGCGTCTG TTGGGAATTA TGCCCTGGAA	2382
TTCTGCCTGA ATTGACCTAC GCATCTCCTC CTCCTGGAC ATTCTTTTGT CTTCATTTGG	2442
TGCTTTTGGT TTTGCACCTC TCCGTGATTG TAGCCCTACC AGCATGTTAT AGGGCAAGAC	2502
CTTTGTGCTT TTGATCATTG TGGCCCATGA AAGCAACTTT GGTCTCCTTT CCCCTCCTGT	2562
CTTCCCGTA TCCCTTGGAG TCTCACAAGG TTTACTTTGG TATGGTTCTC AGCACAACC	2622
TTCAAGTAT GTTGTTCCTT TGGAAAATGG ACATACTGTA TTGTGTTCTC CTGCATATAT	2682
CATTCTGGA GAGAGAAGGG GAGAAGAATA CTTTCTTCA ACAAATTTG GGGCAGGAG	2742
ATCCCTCAA GAGGCTGCAC CTTAATTTTT CTGTCTGTG TGCAGGTCTT CATATAAAT	2802

FIG.24E

TTACCAGGAA GAAGGGTGTG AGTTTGTGT TTTTCTGTGT ATGGGCCTGG TCAGTGTA	2862
GTTTTATCCT TGATAGTCTA GTTACTATGA CCCTCCCCAC TTTTTAAAA CCAGAAAAAG	2922
GTTTGGAAATG TTGGAATGAC CAAGAGACAA GTTAACTCGT GCAAGAGCCA GTTACCCACC	2982
CACAGGTCCC CCTACTTCCT GCCAAGCATT CCATTGACTG CCTGTATGGA ACACATTTGT	3042
CCCAGATCTG AGCATTCTAG GCCTGTTTCA CTCACTCACC CAGCATATGA AACTAGTCTT	3102
AACTGTTGAG CCTTTCCTTT CATATCCACA GAAGACACTG TCTCAAATGT TGTACCCTTG	3162
CCATTTAGGA CTGAACTTTC CTTAGCCCAA GGGACCCAGT GACAGTTGTC TTCCGTTTGT	3222
CAGATGATCA GTCTCTACTG ATTATCTTGC TGCTTAAAGG CCTGCTCACC AATCTTCTT	3282
TCACACCGTG TGGTCCGTGT TACTGGTATA CCCAGTATGT TCTCACTGAA GACATGGACT	3342
TTATATGTTT AAGTGCAGGA ATTGGAAAGT TGGACTTGTT TTCTATGATC CAAAACAGCC	3402
CTATAAGGAG GTTGGAAAAG GAGGAACTAT ATAGCAGCCT TTGCTATTTT CTGCTACCAT	3462
TTCTTTTCCT CTGAAGCGGC CATGACATTC CCTTTGGCAA CTAACGTAGA AACTCACAG	3522

FIG.24F

AACATTTTCC TTTCCTAGAG TCACCTTTTA GATGATAATG GACAACATA GACTTGCTCA	3582
TTGTTGAGAC TGATTGCCCC TCACCTGAAT CCACTCTCTG TATTCATGCT CTTGGCAATT	3642
TCTTTGACTT TCTTTAAGG GCAGAAGCAT TTTAGTTAAT TGAGATAAA GAATAGTTT	3702
CTTCCTCTTC TCCTTGGGCC AGTTAATAAT TGGTCCATGG CTACACTGCA ACTTCCGTCC	3762
AGTGCTGTGA TGCCCATGAC ACCTGCAAAA TAAGTTCTGC CTGGGCATTT TGAGATATT	3822
AACAGGTGAA TTCCCGACTC TTTTGGTTG AATGACAGTT CTCATTCCTT CTATGGCTGC	3882
AAGTATGCAT CAGTGCTTCC CACTTACCTG ATTTGTCTGT CGGTGGCCCC ATATGGAAAC	3942
CCTGCGTGTG TGTTGGCATA ATAGTTTACA AATGGTTTTT TCAGTCCTAT CCAAATTTAT	4002
TGAACCAACA AAAATAATTA CTCTGCCCT GAGATAAGCA GATTAAGTTT GTTCATTCTC	4062
TGCTTTATTC TCTCCATGTG GCAACATTCT GTCAGCCTCT TTCATAGTGT GCAAACATTT	4122
TATCATTCTA AATGGTGACT CTCTGCCCTT GGACCCATTT ATTATTCACA GATGGGGAGA	4182
ACCTATCTGC ATGGACCCTC ACCATCCTCT GTGCAGCACA CACAGTGCAG GGAGCCAGTG	4242
GCGATGGCGA TGACTTTCTT CCCCTG	4268

FIG.24G

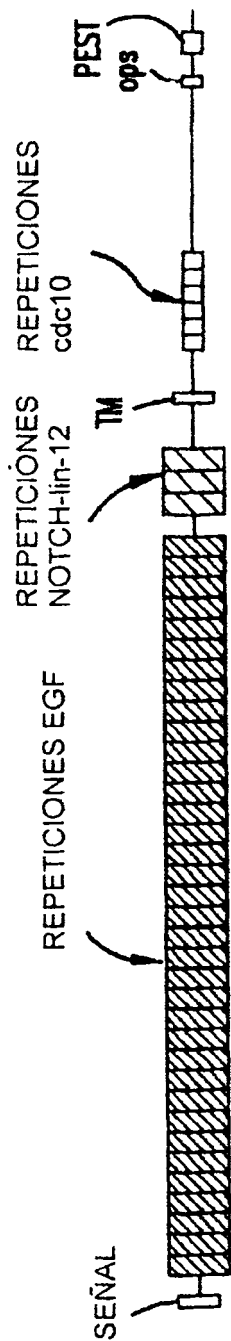
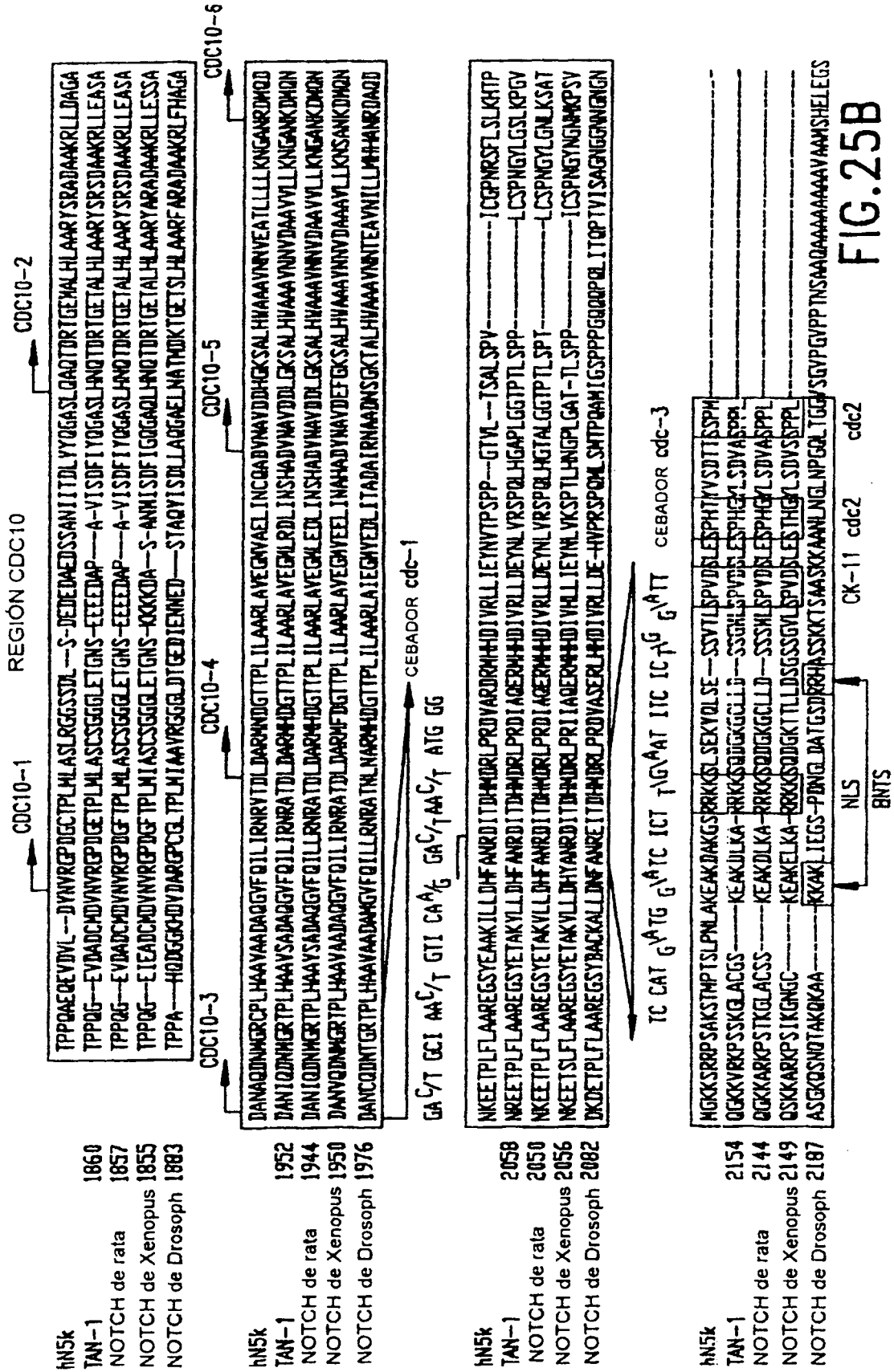


FIG.25A





**LISTA DE SECUENCIAS**

(1) INFORMACIÓN GENERAL:

- 5 (i) SOLICITANTE: Artavanis-Tsakonas, Spyridon *et al.*
- (ii) TÍTULO DE LA INVENCIÓN: Notch y Delta Humanas, Dominios de Unión en Proteínas Toporrítmicas y Métodos Basados en las Mismas
- 10 (iii) NÚMERO DE SECUENCIAS: 37
- (iv) DIRECCIÓN DE LA CORRESPONDENCIA:
- (A) DESTINATARIO: Pennie & Edmonds
- 15 (B) CALLE: 1155 Avenue of the Americas
- (C) CIUDAD: Nueva York
- (D) ESTADO: Nueva York
- 20 (E) PAÍS: EE.UU.
- (F) C.P.: 10036
- (v) FORMA LEGIBLE EN ORDENADOR:
- (A) TIPO DE MEDIO: Disquete
- 25 (B) ORDENADOR: Compatible con PC IBM
- (C) SISTEMA OPERADOR: PC-DOS/MS-DOS
- (D) PROGRAMA: PatentIn Release nº 1.0, Version nº 1.25
- 30 (vi) DATOS DE LA SOLICITUD ACTUAL:
- (A) NÚMERO DE SOLICITUD:
- (B) FECHA DE PRESENTACIÓN:
- 35 (C) CLASIFICACIÓN:
- (viii) INFORMACIÓN SOBRE EL APODERADO/AGENTE:
- (A) NOMBRE: Misrock, S. Leslie
- 40 (B) NÚMERO DE INSCRIPCIÓN: 18,872
- (C) REFERENCIA/NÚMERO DE ORDEN: 7326-009
- (ix) INFORMACIÓN SOBRE TELECOMUNICACIONES:
- 45 (A) TELÉFONO: 212 790-9090
- (B) TELEFAX: 212 8698864/9741
- (C) TELEX: 66141 PENNIE

50 (2) INFORMACIÓN DE LA SEQ ID NO:1:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 77 aminoácidos
- 55 (B) TIPO: aminoácido
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida
- 60 (ii) TIPO DE MOLÉCULA: péptido
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:1:

65



ES 2 181 674 T5

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

(2) INFORMACIÓN DE LA SEQ ID NO:4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 199 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: péptido

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

ES 2 181 674 T5

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

40 (2) INFORMACIÓN DE LA SEQ ID NO:5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 2892 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: bicatenaria
- (D) TOPOLOGÍA: desconocida

50 (ii) TIPO DE MOLÉCULA: ADNc

(ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 142..2640

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

60

65



ES 2 181 674 T5

	Ser	Phe	Gly	His	Ser	Thr	Cys	Ser	Glu	Thr	Gly	Glu	Ile	Ile	Cys	Leu	
			205					210					215				
5	ACC	GGA	TGG	CAG	GGC	GAT	TAC	TGT	CAC	ATA	CCC	AAA	TGC	GCC	AAA	GGC	843
	Thr	Gly	Trp	Gln	Gly	Asp	Tyr	Cys	His	Ile	Pro	Lys	Cys	Ala	Lys	Gly	
			220					225				230					
10	TGT	GAA	CAT	GGA	CAT	TGC	GAC	AAA	CCC	AAT	CAA	TGC	GTT	TGC	CAA	CTG	891
	Cys	Glu	His	Gly	His	Cys	Asp	Lys	Pro	Asn	Gln	Cys	Val	Cys	Gln	Leu	
						240					245					250	
15	GGC	TGG	AAG	GGA	GCC	TTG	TGC	AAC	GAG	TGC	GTT	CTG	GAA	CCG	AAC	TGC	939
	Gly	Trp	Lys	Gly	Ala	Leu	Cys	Asn	Glu	Cys	Val	Leu	Glu	Pro	Asn	Cys	
					255					260					265		
20	ATC	CAT	GGC	ACC	TGC	AAC	AAA	CCC	TGG	ACT	TGC	ATC	TGC	AAC	GAG	GGT	987
	Ile	His	Gly	Thr	Cys	Asn	Lys	Pro	Trp	Thr	Cys	Ile	Cys	Asn	Glu	Gly	
				270					275					280			
25	TGG	GGA	GGC	TTG	TAC	TGC	AAC	CAG	GAT	CTG	AAC	TAC	TGC	ACC	AAC	CAC	1035
	Trp	Gly	Gly	Leu	Tyr	Cys	Asn	Gln	Asp	Leu	Asn	Tyr	Cys	Thr	Asn	His	
				285				290					295				
30	AGA	CCC	TGC	AAG	AAT	GGC	GGA	ACC	TGC	TTC	AAC	ACC	GGC	GAG	GGA	TTG	1083
	Arg	Pro	Cys	Lys	Asn	Gly	Gly	Thr	Cys	Phe	Asn	Thr	Gly	Glu	Gly	Leu	
							305					310					
35	TAC	ACA	TGC	AAA	TGC	GCT	CCA	GGA	TAC	AGT	GGT	GAT	GAT	TGC	GAA	AAT	1131
	Tyr	Thr	Cys	Lys	Cys	Ala	Pro	Gly	Tyr	Ser	Gly	Asp	Asp	Cys	Glu	Asn	
						320					325					330	
40	GAG	ATC	TAC	TCC	TGC	GAT	GCC	GAT	GTC	AAT	CCC	TGC	CAG	AAT	GGT	GGT	1179
	Glu	Ile	Tyr	Ser	Cys	Asp	Ala	Asp	Val	Asn	Pro	Cys	Gln	Asn	Gly	Gly	
					335					340					345		
45	ACC	TGC	ATC	GAT	GAG	CCG	CAC	ACA	AAA	ACC	GGC	TAC	AAG	TGT	CAT	TGC	1227
	Thr	Cys	Ile	Asp	Glu	Pro	His	Thr	Lys	Thr	Gly	Tyr	Lys	Cys	His	Cys	
				350					355					360			
50	GCC	AAC	GGC	TGG	AGC	GGA	AAG	ATG	TGC	GAG	GAG	AAA	GTG	CTC	ACG	TGT	1275
	Ala	Asn	Gly	Trp	Ser	Gly	Lys	Met	Cys	Glu	Glu	Lys	Val	Leu	Thr	Cys	
				365				370					375				
55	TCG	GAC	AAA	CCC	TGT	CAT	CAG	GGA	ATC	TGC	CGC	AAC	GTT	CGT	CCT	GGC	1323
	Ser	Asp	Lys	Pro	Cys	His	Gln	Gly	Ile	Cys	Arg	Asn	Val	Arg	Pro	Gly	
				380			385					390					
60	TTG	GGA	AGC	AAG	GGT	CAG	GGC	TAC	CAG	TGC	GAA	TGT	CCC	ATT	GGC	TAC	1371
	Leu	Gly	Ser	Lys	Gly	Gln	Gly	Tyr	Gln	Cys	Glu	Cys	Pro	Ile	Gly	Tyr	
						400				405					410		
65	AGC	GGA	CCC	AAC	TGC	GAT	CTC	CAG	CTG	GAC	AAC	TGC	AGT	CCG	AAT	CCA	1419
	Ser	Gly	Pro	Asn	Cys	Asp	Leu	Gln	Leu	Asp	Asn	Cys	Ser	Pro	Asn	Pro	
				415					420						425		
70	TGC	ATA	AAC	GGT	GGA	AGC	TGT	CAG	CCG	AGC	GGA	AAG	TGT	ATT	TGC	CCA	1467
	Cys	Ile	Asn	Gly	Gly	Ser	Cys	Gln	Pro	Ser	Gly	Lys	Cys	Ile	Cys	Pro	
				430					435					440			
75	GCG	GGA	TTT	TCG	GGA	ACG	AGA	TGC	GAG	ACC	AAC	ATT	GAC	GAT	TGT	CTT	1515
	Ala	Gly	Phe	Ser	Gly	Thr	Arg	Cys	Glu	Thr	Asn	Ile	Asp	Asp	Cys	Leu	
				445			450						455				
80	GGC	CAC	CAG	TGC	GAG	AAC	GGA	GGC	ACC	TGC	ATA	GAT	ATG	GTC	AAC	CAA	1563
	Gly	His	Gln	Cys	Glu	Asn	Gly	Gly	Thr	Cys	Ile	Asp	Met	Val	Asn	Gln	
				460			465					470					
85	TAT	CGC	TGC	CAA	TGC	GTT	CCC	GGT	TTC	CAT	GGC	ACC	CAC	TGT	AGT	AGC	1611
	Tyr	Arg	Cys	Gln	Cys	Val	Pro	Gly	Phe	His	Gly	Thr	His	Cys	Ser	Ser	
						480					485					490	

ES 2 181 674 T5

AAA GTT GAC TTG TGC CTC ATC AGA CCG TGT GCC AAT GGA GGA ACC TGC 1659  
 Lys Val Asp Leu Cys Leu Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys  
 495 500 505

5 TTG AAT CTC AAC AAC GAT TAC CAG TGC ACC TGT CGT GCG GGA TTT ACT 1707  
 Leu Asn Leu Asn Asn Asp Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr  
 510 515 520

10 GGC AAG GAT TGC TCT GTG GAC ATC GAT GAG TGC AGC AGT GGA CCC TGT 1755  
 Gly Lys Asp Cys Ser Val Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys  
 525 530 535

CAT AAC GGC GGC ACT TGC ATG AAC CGC GTC AAT TCG TTC GAA TGC GTG 1803  
 His Asn Gly Gly Thr Cys Met Asn Arg Val Asn Ser Phe Glu Cys Val  
 540 545 550

15 TGT GCC AAT GGT TTC AGG GGC AAG CAG TGC GAT GAG GAG TCC TAC GAT 1851  
 Cys Ala Asn Gly Phe Arg Gly Lys Gln Cys Asp Glu Glu Ser Tyr Asp  
 555 560 565 570

20 TCG GTG ACC TTC GAT GCC CAC CAA TAT GGA GCG ACC ACA CAA GCG AGA 1899  
 Ser Val Thr Phe Asp Ala His Gln Tyr Gly Ala Thr Thr Gln Ala Phe  
 575 580 585

GCC GAT GGT TTG ACC AAT GCC CAG GTA GTC CTA ATT GCT GTT TTC TCC 1947  
 Ala Asp Gly Leu Thr Asn Ala Gln Val Val Leu Ile Ala Val Phe Ser  
 590 595 600

25 GTT GCG ATG CCT TTG GTG GCG GTT ATT GCG GCG TGC GTG GTC TTC TGC 1995  
 Val Ala Met Pro Leu Val Ala Val Ile Ala Ala Cys Val Val Phe Cys  
 605 610 615

30 ATG AAG CGC AAG CGT AAG CGT GCT CAG GAA AAG GAC GAC GCG GAG GCC 2043  
 Met Lys Arg Lys Arg Lys Arg Ala Gln Glu Lys Asp Asp Ala Glu Ala  
 620 625 630

35 AGG AAG CAG AAC GAA CAG AAT GCG GTG GCC ACA ATG CAT CAC AAT GGC 2091  
 Arg Lys Gln Asn Glu Gln Asn Ala Val Ala Thr Met His His Asn Gly  
 635 640 645 650

AGT GGG GTG GGT GTA GCT TTG GCT TCA GCC TCT CTG GGC GGC AAA ACT 2139  
 Ser Gly Val Gly Val Ala Leu Ala Ser Ala Ser Leu Gly Gly Lys Thr  
 655 660 665

40 GGC AGC AAC AGC GGT CTC ACC TTC GAT GGC GGC AAC CCG AAT ATC ATC 2187  
 Gly Ser Asn Ser Gly Leu Thr Phe Asp Gly Gly Asn Pro Asn Ile Ile  
 670 675 680

45 AAA AAC ACC TGG GAC AAG TCG GTC AAC AAC ATT TGT GCC TCA GCA GCA 2235  
 Lys Asn Thr Trp Asp Lys Ser Val Asn Asn Ile Cys Ala Ser Ala Ala  
 685 690 695

GCA GCG GCG GCG GCG GCA GCA GCG GCG GAC GAG TGT CTC ATG TAC GGC 2283  
 Ala Ala Ala Ala Ala Ala Ala Ala Ala Asp Glu Cys Leu Met Tyr Gly  
 700 705 710

50 GGA TAT GTG GCC TCG GTG GCG GAT AAC AAC AAT GCC AAC TCA GAC TTT 2331  
 Gly Tyr Val Ala Ser Val Ala Asp Asn Asn Asn Ala Asn Ser Asp Phe  
 715 720 725 730

55 TGT GTG GCT CCG CTA CAA AGA GCC AAG TCG CAA AAG CAA CTC AAC ACC 2379  
 Cys Val Ala Pro Leu Gln Arg Ala Lys Ser Gln Lys Gln Leu Asn Thr  
 735 740 745

GAT CCC ACG CTC ATG CAC CGC GGT TCG CCG GCA GGC AGC TCA GCC AAG 2427  
 Asp Pro Thr Leu Met His Arg Gly Ser Pro Ala Gly Ser Ser Ala Lys  
 750 755 760

60 GGA GCG TCT GGC GGA GGA CCG GGA GCG GCG GAG GGC AAG AGG ATC TCT 2475  
 Gly Ala Ser Gly Gly Gly Pro Gly Ala Ala Glu Gly Lys Arg Ile Ser

ES 2 181 674 T5

	765	770	775			
5	GTT TTA GGC GAG GGT TCC TAC TGT AGC CAG CGT TGG CCC TCG TTG GCG Val Leu Gly Glu Gly Ser Tyr Cys Ser Gln Arg Trp Pro Ser Leu Ala	780	785	790	2523	
	GCG GCG GGA GTG GCC GGA GCC TGT TCA TCC CAG CTA ATG GCT GCA GCT Ala Ala Gly Val Ala Gly Ala Cys Ser Ser Gln Leu Met Ala Ala Ala	795	800	805	810	2571
10	TCG GCA GCG GGC AGC GGA GCG GGG ACC GCG CAA CAG CAG CGA TCC GTG Ser Ala Ala Gly Ser Gly Ala Gly Thr Ala Gln Gln Gln Arg Ser Val	815	820	825	2619	
15	GTC TGC GGC ACT CCG CAT ATG TAACTCCAAA AATCCGGAAG GGCTCCTGG Val Cys Gly Thr Pro His Met	830			2670	
	AAATCCGGAG AAATCCGCAT GGAGGAGCTG ACAGCACATA CACAAAGAAA AGACTGGGTT				2730	
	GGGTTCAAAA TGTGAGAGAG ACGCCAAAAT GTTGTGTGTTG ATTGAAGCAG TTTAGTCGTC				2790	
20	ACGAAAAATG AAAAATCTGT AACAGGCATA ACTCGTAAAC TCCCTAAAAA ATTTGTATAG				2850	
	TAATTAGCAA AGCTGTGACC CAGCCGTTTC GATCCCGAAT TC				2892	

25 (2) INFORMACIÓN DE LA SEQ ID NO:6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 833 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: proteína

35 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:6:

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

ES 2 181 674 T5

Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu Asn Tyr Tyr Gly Ser Gly  
 180 185 190  
 Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp Ser Phe Gly His Ser Thr  
 195 200 205  
 Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu Thr Gly Trp Gln Gly Asp  
 210 215 220  
 Tyr Cys His Ile Pro Lys Cys Ala Lys Gly Cys Glu His Gly His Cys  
 225 230 235 240  
 Asp Lys Pro Asn Gln Cys Val Cys Gln Leu Gly Trp Lys Gly Ala Leu  
 245 250 255  
 Cys Asn Glu Cys Val Leu Glu Pro Asn Cys Ile His Gly Thr Cys Asn  
 260 265 270  
 Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly Trp Gly Gly Leu Tyr Cys  
 275 280 285  
 Asn Gln Asp Leu Asn Tyr Cys Thr Asn His Arg Pro Cys Lys Asn Gly  
 290 295 300  
 Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu Tyr Thr Cys Lys Cys Ala  
 305 310 315 320  
 Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn Glu Ile Tyr Ser Cys Asp  
 325 330 335  
 Ala Asp Val Asn Pro Cys Gln Asn Gly Gly Thr Cys Ile Asp Glu Pro  
 340 345 350  
 His Thr Lys Thr Gly Tyr Lys Cys His Cys Ala Asn Gly Trp Ser Gly  
 355 360 365  
 Lys Met Cys Glu Glu Lys Val Leu Thr Cys Ser Asp Lys Pro Cys His  
 370 375 380  
 Gln Gly Ile Cys Arg Asn Val Arg Pro Gly Leu Gly Ser Lys Gly Gln  
 385 390 395 400  
 Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr Ser Gly Pro Asn Cys Asp  
 405 410 415  
 Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro Cys Ile Asn Gly Gly Ser  
 420 425 430  
 Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro Ala Gly Phe Ser Gly Thr  
 435 440 445  
 Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu Gly His Gln Cys Glu Asn  
 450 455 460  
 Gly Gly Thr Cys Ile Asp Met Val Asn Gln Tyr Arg Cys Gln Cys Val  
 465 470 475 480  
 Pro Gly Phe His Gly Thr His Cys Ser Ser Lys Val Asp Leu Cys Leu  
 485 490 495  
 Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys Leu Asn Leu Asn Asn Asp  
 500 505 510  
 Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr Gly Lys Asp Cys Ser Val  
 515 520 525  
 Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys His Asn Gly Gly Thr Cys  
 530 535 540  
 Met Asn Arg Val Asn Ser Phe Glu Cys Val Cys Ala Asn Gly Phe Arg



ES 2 181 674 T5

	<b>GATCTACTAC GAGGAGGTTA AGGAGAGCTA TGTGGGCGAG CGACGCGAAT ACGATCCCCA</b>	<b>60</b>
	<b>CATCACCGAT CCCAGGGTCA CACGCATGAA GATGGCCGGC CTGAAGCCCA ACTCCAATA</b>	<b>120</b>
5	<b>CCGCATCTCC ATCACTGCCA CCACGAAAAT GGGCGAGGGA TCTGAACACT ATATCGAAAA</b>	<b>180</b>
	<b>GACCACGCTC AAGGATGCCG TCAATGTGGC CCCTGCCACG CCATCTTTCT CCTGGGAGCA</b>	<b>240</b>
	<b>ACTGCCATCC GACAATGGAC TAGCCAAGTT CCGCATCAAC TGGCTGCCAA GTAACGAGGG</b>	<b>300</b>
10	<b>TCATCCAGGC ACTCACTTCT TTACGATGCA CAGGATCAAG GGCGAAACCC AATGGATACG</b>	<b>360</b>
	<b>CGAGAATGAG GAAAAGAACT CCGATTACCA GGAGGTCGGT GGCTTAGATC CGGAGACCGC</b>	<b>420</b>
	<b>CTACGAGTTC CGCGTGGTGT CCGTGGATGG CCACTTTAAC ACGGAGAGTG CCAACGAGGA</b>	<b>480</b>
15	<b>GATCGACACG AACACCGTTG AGGGACCAAT AATGGTGGCC AACGAGACGG TGGCCAATGC</b>	<b>540</b>
	<b>CGGATGGTTC ATTGGCATGA TGCTGGCCCT GGCCTTCATC ATCATCCTCT TCAATCATCAT</b>	<b>600</b>
	<b>CTGCATTATC CGACGCAATC GGGGCGGAAA GTACGATGTC CACGATCGGG AGCTGGCCAA</b>	<b>660</b>
20	<b>CGGCCGGCGG GATTATCCCG AAGAGGGCGG ATTCCACGAG TACTCGCAAC CGTIGGATAA</b>	<b>720</b>
	<b>CAAGAGCGCT GGTCCGCAAT CCGTGAGTTC AGCGAACAA CCGGGCGTGG AAAGCGATAC</b>	<b>780</b>
	<b>TGATTCGATG GCCGAATACG GTGATGGCGA TACAGGACAA TTTACCGAGG ATGGCTCCTT</b>	<b>840</b>
25	<b>CATTGGCCAA TATGTTCTTG GAAAGCTCCA ACCGCCGTT AGCCCACAGC CACTGAACAA</b>	<b>900</b>
	<b>TTCCGCTGCG GCGCATCAGG CGGCGCCAAC TGCCGGAGGA TCGGGAGCAG CCGGATCGGC</b>	<b>960</b>
	<b>AGCAGCAGCC GGAGCATCGG GTGGAGCATC GTCCGCCGGA GGAGCAGCTG CCAGCAATGG</b>	<b>1020</b>
30	<b>AGGAGCTGCA GCCGGAGCCG TGGCCACCTA CGTCTAAGCT TGGTACC</b>	<b>1067</b>

(2) INFORMACIÓN DE LA SEQ ID NO:8:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1320 pares de bases

(B) TIPO: ácido nucleico

40 (C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

45 (ix) CARACTERÍSTICAS:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 442..1320

50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:8:

	<b>CCGAGTCGAG CGCCGTGCTT CGAGCGGTGA TGAGCCCCTT TTCTGTCAAC GCTAAAGATC</b>	<b>60</b>
	<b>TACAAAACAT CAGCGCCTAT CAAGTGGGAAG TGTCAAGTGT GAACAAAACA AAACGAGAG</b>	<b>120</b>
55	<b>AAGCACATAC TAAGGTCCAT ATAAATAATA AATAATAATT GTGTGTGATA ACMACATTAT</b>	<b>180</b>
	<b>CCAAACAAAA CCAAACAAA CGAAGGCAAA GTGGAGAAAA TGATACAGCA TCCAGAGTAC</b>	<b>240</b>
	<b>GGCCGTTATT CAGCTATCCA GAGCAAGTGT AGTGTGGCAA AATAGAAACA AACAAAGGCA</b>	<b>300</b>
60	<b>CCAAAATCTG CATAATGGG CTAATTAAGG CTGCCACCG AATTACATT TGTTGGGTGC</b>	<b>360</b>
	<b>CAATCCAGAG TGAATCCGAA ACAAACCTCCA TCTAGATCGC CAACCAGCAT CAGCTCGCA</b>	<b>420</b>
65	<b>AACGCCCCCA GAATGTACAA A ATG TTT AGG AAA CAT TTT CGG CGA AAA CCA</b>	<b>471</b>
	<b>Met Phe Arg Lys His Phe Arg Arg Lys Pro</b>	

ES 2 181 674 T5

	1	5	10	
5	GCT ACG TCG TCG TCG TTG GAG TCA ACA ATA GAA TCA GCA GAC AGC CTG Ala Thr Ser Ser Ser Leu Glu Ser Thr Ile Glu Ser Ala Asp Ser Leu 15 20 25	519		
	GGA ATG TCC AAG AAG ACG GCG ACA AAA AGG CAG CGT CCG AGG CAT CCG Gly Met Ser Lys Lys Thr Ala Thr Lys Arg Gln Arg Pro Arg His Arg 30 35 40	567		
10	GTA CCC AAA ATC GCG ACC CTG CCA TCG ACG ATC CGC GAT TGT CGA TCA Val Pro Lys Ile Ala Thr Leu Pro Ser Thr Ile Arg Asp Cys Arg Ser 45 50 55	615		
15	TTA AAG TCT GCC TGC AAC TTA ATT GCT TTA ATT TTA ATA CTG TTA GTC Leu Lys Ser Ala Cys Asn Leu Ile Ala Leu Ile Leu Ile Leu Leu Val 60 65 70	663		
	CAT AAG ATA TCC GCA GCT GGT AAC TTC GAG CTG GAA ATA TTA GAA ATC His Lys Ile Ser Ala Ala Gly Asn Phe Glu Leu Glu Ile Leu Glu Ile 75 80 85 90	711		
20	TCA AAT ACC AAC AGC CAT CTA CTC AAC GGC TAT TGC TGC GGC ATG CCA Ser Asn Thr Asn Ser His Leu Leu Asn Gly Tyr Cys Cys Gly Met Pro 95 100 105	759		
25	GCG GAA CTT AGG GCC ACC AAG ACG ATA GGC TGC TCG CCA TGC ACG ACG Ala Glu Leu Arg Ala Thr Lys Thr Ile Gly Cys Ser Pro Cys Thr Thr 110 115 120	807		
	GCA TTC CCG CTG TGC CTG AAG GAG TAC CAG ACC ACG GAG CAG GGT GCC Ala Phe Arg Leu Cys Leu Lys Glu Tyr Gln Thr Thr Thr Gln Gly Ala 125 130 135	855		
30	AGC ATA TCC ACG GGC TGT TCG TTT GGC AAC GCC ACC ACC AAG ATA CTG Ser Ile Ser Thr Gly Cys Ser Phe Gly Asn Ala Thr Thr Lys Ile Leu 140 145 150	903		
35	GGT GGC TCC AGC TTT GTG CTC AGC GAT CCG GGT GTG GGA GCC ATT GTG Gly Gly Ser Ser Phe Val Leu Ser Asp Pro Gly Val Gly Ala Ile Val 155 160 165 170	951		
	CTG CCC TTT ACG TTT CGT TGG ACG AAG TCG TTT ACG CTG ATA CTG CAG Leu Pro Phe Thr Phe Arg Trp Thr Lys Ser Phe Thr Leu Ile Leu Gln 175 180 185	999		
40	GCG TTG GAT ATG TAC AAC ACA TCC TAT CCA GAT GCG GAG AGG TTA ATT Ala Leu Asp Met Tyr Asn Thr Ser Tyr Pro Asp Ala Glu Arg Leu Ile 190 195 200	1047		
45	GAG GAA ACA TCA TAC TCG GGC GTG ATA CTG CCG TCG CCG GAG TGG AAG Glu Glu Thr Ser Tyr Ser Gly Val Ile Leu Pro Ser Pro Glu Trp Lys 205 210 215	1095		
	ACG CTG GAC CAC ATC GGG CCG AAC GCG CCG ATC ACC TAC CGT GTC CCG Thr Leu Asp His Ile Gly Arg Asn Ala Arg Ile Thr Tyr Arg Val Arg 220 225 230	1143		
50	GTG CAA TGC GCC GTT ACC TAC TAC AAC ACG ACC TGC ACG ACC TTC TGC Val Gln Cys Ala Val Thr Tyr Tyr Asn Thr Cys Thr Thr Phe Cys 235 240 245 250	1191		
55	CGT CCG CCG GAC GAT CAG TTC GGT CAC TAC GCC TGC GGC TCC GAG GGT Arg Pro Arg Asp Asp Gln Phe Gly His Tyr Ala Cys Gly Ser Glu Gly 255 260 265	1239		
	CAG AAG CTC TGC CTG AAT GGC TGG CAG GGC GTC AAC TGC GAG GAG GCC Gln Lys Leu Cys Leu Asn Gly Trp Gln Gly Val Asn Cys Glu Glu Ala 270 275 280	1287		
60	ATA TGC AAG GCC GGC TGC GAC CCC GTC CAC GGC Ile Cys Lys Ala Gly Cys Asp Pro Val His Gly 285 290	1320		

## ES 2 181 674 T5

(2) INFORMACIÓN DE LA SEQ ID NO:9:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 293 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: proteína

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

```

Met Phe Arg Lys His Phe Arg Arg Lys Pro Ala Thr Ser Ser Ser Leu
 1           5           10           15

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

Ala Thr Lys Arg Gln Arg Pro Arg His Arg Val Pro Lys Ile Ala Thr
 35           40           45

Leu Pro Ser Thr Ile Arg Asp Cys Arg Ser Leu Lys Ser Ala Cys Asn
 50           55           60

Leu Ile Ala Leu Ile Leu Ile Leu Leu Val His Lys Ile Ser Ala Ala
 65           70           75           80

Gly Asn Phe Glu Leu Glu Ile Leu Glu Ile Ser Asn Thr Asn Ser His
 85           90           95

Leu Leu Asn Gly Tyr Cys Cys Gly Met Pro Ala Glu Leu Arg Ala Thr
100          105          110

Lys Thr Ile Gly Cys Ser Pro Cys Thr Thr Ala Phe Arg Leu Cys Leu
115          120          125

Lys Glu Tyr Gln Thr Thr Glu Gln Gly Ala Ser Ile Ser Thr Gly Cys
130          135          140

Ser Phe Gly Asn Ala Thr Thr Lys Ile Leu Gly Gly Ser Ser Phe Val
145          150          155          160

Leu Ser Asp Pro Gly Val Gly Ala Ile Val Leu Pro Phe Thr Phe Arg
165          170          175

Trp Thr Lys Ser Phe Thr Leu Ile Leu Gln Ala Leu Asp Met Tyr Asn
180          185          190

Thr Ser Tyr Pro Asp Ala Glu Arg Leu Ile Glu Glu Thr Ser Tyr Ser
195          200          205

Gly Val Ile Leu Pro Ser Pro Glu Trp Lys Thr Leu Asp His Ile Gly
210          215          220

Arg Asn Ala Arg Ile Thr Tyr Arg Val Arg Val Gln Cys Ala Val Thr
225          230          235          240

Tyr Tyr Asn Thr Thr Cys Thr Thr Phe Cys Arg Pro Arg Asp Asp Gln
245          250          255

Phe Gly His Tyr Ala Cys Gly Ser Glu Gly Gln Lys Leu Cys Leu Asn
260          265          270

Gly Trp Gln Gly Val Asn Cys Glu Glu Ala Ile Cys Lys Ala Gly Cys
275          280          285

Asp Pro Val His Gly
290
    
```

## ES 2 181 674 T5

### (2) INFORMACIÓN DE LA SEQ ID NO:10:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 26 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 6
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 12
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:10:

GAYGCNAAYG TNCARGAYAA YATGGC

26

### (2) INFORMACIÓN DE LA SEQ ID NO:11:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 25 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 3
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 12
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 18
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:11:

ATNARRTCYT CNACCATNCC YTCDA

25

## ES 2 181 674 T5

### (2) INFORMACIÓN DE LA SEQ ID NO:12:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 26 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 12
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 18
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: modified\_base
- (B) LOCALIZACIÓN: 21
- (D) OTRA INFORMACIÓN: /mod\_base= i /label= N

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:12:

TCCATRTGRT CNGTDATNTC NCKRTT

26

### (2) INFORMACIÓN DE LA SEQ ID NO:13:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 267 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: bicatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:13:

```
GAATTCGGCT GGGAGAATGG TCTGAGCTAC CTGCCCGTCC TGCTGGGGCA TCAATGGCAA 60
GTGGGGAAAG CCACACTGGG CAAACGGGCC AGGCCATTTC TGGAAATGTGG TACATGGTGG 120
GCAGGGGGCC CGCAACAGCT GGAGGGCAGG TGGACTGAGG CTGGGGATCC CCGCTGGTT 180
GGGCAATACT GCCTTTACCC ATGAGCTGGA AAGTCACAAT GGGGGGCAAG GGTCTCCGAG 240
GGTGGTTATG TGCTTCCTTC AGGTGGC 267
```

### (2) INFORMACIÓN DE LA SEQ ID NO:14:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 574 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: bicatenaria

## ES 2 181 674 T5

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

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

	<b>GAATTCCTTC CATTATACGT GACTTTTCTG AACTGTAGC CACCCTAGTG TCTCTAACTC</b>	<b>60</b>
	<b>CCTCTGGAGT TTGTCAGCTT TGGTCTTTTC AAGAGCAGG CTCTCTTCAI GTCCTTAAT</b>	<b>120</b>
10	<b>GCGGGCATGC TCCAGTTTGG TCTGCGTCTC AAGATCACCT TTGGTAATTG ATTCTTCTTC</b>	<b>180</b>
	<b>AACCCGGAAC TGAAGGCTGG CTCTCACCCCT CTAGGCAGAG CAGGAATTCC GAGGTGGATG</b>	<b>240</b>
15	<b>TGTTAGATGT GAATGTCCGT GGCCCAGATG GCTGCACCCC ATTGATGTTG GCTTCTCTCC</b>	<b>300</b>
	<b>GAGGAGGCAG CTCAGATTG AGTGATGAAG ATGAAGATGC AGAGGACTGT TCTGCTAACA</b>	<b>360</b>
	<b>TCATCACAGA CTTGGTCTAC CAGGGTGCCA GCCTCCAGNC CAGACAGAC GACTGGTGA</b>	<b>420</b>
20	<b>GATGGCCCTG CACCTTGACG CCCGCTACTC ACGGGCTGAT GCTGCCAAG GTCTCCTGGA</b>	<b>480</b>
	<b>TGCAGGTGCA GATGCCAATG CCCAGGACAA CATGGCCCGC TGTCCACTCC ATGCTGCAGT</b>	<b>540</b>
25	<b>GGCACGTGAT GCCAAGGTGT ATTCAGATCT GTTA</b>	<b>574</b>

(2) INFORMACIÓN DE LA SEQ ID NO:15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 295 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:15:

	<b>TCCAGATTCT GATTGCGAAC CGAGTAACTG ATCTAGATGC CAGGATGAAT GATGGTACTA</b>	<b>60</b>
	<b>CACCCCTGAT CCTGGCTGCC CGCCTGGCTG TGGAGGGAAT GGTGGCAGAA CTGATCAACT</b>	<b>120</b>
45	<b>GCCAAGCGGA TGTGAATGCA GTGGATGACC ATGGAATAATC TGCTCTTCAC TGGGCAGCTG</b>	<b>180</b>
	<b>CTGTCAATAA TGTGGAGGCA ACTCTTTTGT TGTGAAAAA TGGGGCCAAC CGAGACATGC</b>	<b>240</b>
50	<b>AGGACAACAA GGAAGAGACA CCTCTGTTTC TTGCTGCCCC GGAGGAGCTA TAAGC</b>	<b>295</b>

(2) INFORMACIÓN DE LA SEQ ID NO:16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:16:

ES 2 181 674 T5

GAATTC~~CCAT~~ GAGTCGGGAG CTTCGATCAA AATTGATGAG CCTTTAGAAG GATCCGAAGA 60  
 TCGGATCATT ACCATTACAG GAACAGGCAC CTGTAGCTGG TGGCTGGGGG TGTTGTCCAC 120  
 5 AGGCGAGGAG TAGCTGTGCT GCGAGGGGGG CGTCAGGAAC TGGGCTGCGG TCACGGGTGG 180  
 GACCAGCGAG GATGGCAGCG ACGTGGGCAG GCGGGGGCTC TCCTGGGGCA GAATAGTGTG 240  
 CACCGCCAGG CTGCTGGGGC CCAGTACTGC ACGTCTGCCT GGCTCGGCTC TCCACTCAGG 300  
 10 AAGCTCCGGC CCAGGTGGCC GCTGGCTGCT GAG 333

(2) INFORMACIÓN DE LA SEQ ID NO:17:

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 582 pares de bases

(B) TIPO: ácido nucleico

20 (C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:17:

GAATTCCTGC CAGGAGGACG CGGGCAACAA GGTCTGCAGC CTGCAGTGCA ACAACCACGC 60  
 GTGCGGCTGG GACGGCGGTG ACTGCTCCCT CAACTTCACA ATGACCCCTG GAAGAACTGC 120  
 30 ACGCAGTCTC TGCAGTGCTG GAAGTACTTC AGTGACGGCC ACTGTGACAG CCAGTGCAAC 180  
 TCAGCCGGCT GCCTCTTCGA CGGCTTTGAC TGCCAGCGGC GGAAGGCCAG TTGCAACCCC 240  
 35 CTGTACGACC AGTACTGCAA GGACCACTTC AGCGACGGGC ACTGCGACCA GGGCTGCAAC 300  
 AGCGCGGAGT NCAGNTGGGA CGGCTGGAC TGTGCGGCAG TGTACCCGAG AGCTGGCGGC 360  
 GCACGCTGGT GGTGGTGGTG CTGATGCCGC CGGAGCAGCT GCGCAACAGC TCCTTCCACT 420  
 40 TCCTGCGGGA CGTCAGCCGC GTGCTGCACA CCAACGTGTC TTCAAGCGTG ACGCACACGG 480  
 CCAGCAGATG ATGTTCCCTT ACTACGGCCG CGAGGAGGAG CTGCGCAAGC CCCATCAAGC 540  
 45 GTGCCGCCGA GGGCTGGGCC GCACCTGACG CCTGCTGGGC CA 582

(2) INFORMACIÓN DE LA SEQ ID NO:18:

50 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 150 pares de bases

(B) TIPO: ácido nucleico

55 (C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

60 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:18:

TCAGCCGAGT GCTGCACACC AACGTGTCTT CAAGCGTGAC GCACACGGCC AGCAGATGAT 60  
 GTTCCCCTAC TACGGCCGCG AGGAGGAGCT GCGCAAGCCC CATCAAGCGT GCCGCCGAGG 120  
 65 GCTGGGCCGC ACCTGACGCC TGCTGGGCCA 150

## ES 2 181 674 T5

### (2) INFORMACIÓN DE LA SEQ ID NO:19:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 247 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:19:

```
15 TTACCATTAC AGGAACAGGC ACCTGTAGCT GGTGGCTGGG GGTGTTGTCC ACAGGCGAGG 60  
AGTAGCTGTG CTGCGAGGGG GCGTCAGGA ACTGGGCTGC GGTACCGGT GGGACCAGCG 120  
AGGATGGCAG CGACGTGGGC AGGGCGGGGC TCTCTGGGG CAGAATAGTG TGACCCGCCA 180  
20 GCTGCTGGGG CCCAGTGCTG CACGTCTGCC TGGCTCGGCT CTCCACTCAG GAGCTCCGG 240  
CCCAGGT 247
```

### (2) INFORMACIÓN DE LA SEQ ID NO:20:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 248 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:20:

```
40 GAATTCATT CAGGAGGAAA GGTGGGGAG AGAAGCAGGC ACCCACTTC CCGTGGCTGG 60  
ACTCGTTCC AGGTGGCTCC ACCGGCAGCT GTGACCGCCG CAGGTGGGGG CCGAGTGCCA 120  
TTCAGAAAAT TCCAGAAAAG CCCTACCCCA ACTCGGACGG CAACGTCACA CCGTGGGTA 180  
45 GCAACTGGCA CACAAACAGC CAGCGTGTCT GGGGCACGGG GGGATGGCAC CCCCTGCAGG 240  
CAGAGCTG 248
```

### (2) INFORMACIÓN DE LA SEQ ID NO:21:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 323 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:21:

## ES 2 181 674 T5

	<b>CTAAAGGGAA CAAAAGCNGG AGCTCCACCG CGGGCGGCNC NGCTCTAGAA CTAGTGGANN</b>	<b>60</b>
	<b>NCCCGGGCTG CAGGAATTCC GCGGACTGG GCTCGGGCTC AGAGCGGCGC TGTGGAAGAG</b>	<b>120</b>
5	<b>ATTCTAGACC GGGAGAACAA GCGAATGGCT GACAGCTGGC CTCCAAAGTC ACCAGGCTCA</b>	<b>180</b>
	<b>AATCGCTCGC CCTGGACATC GAGGGATGCA GAGGATCAGA ACCGGTACCT GGATGGCATG</b>	<b>240</b>
	<b>ACTCGGATTT ACAAGCATGA CCAGCCTGCT TACAGGGAGC GTGANNTTTT CACATGCAGT</b>	<b>300</b>
10	<b>CGACAGACAC GAGCTCTATG CAT</b>	<b>323</b>

(2) INFORMACIÓN DE LA SEQ ID NO:22:

- 15 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 330 pares de bases
  - (B) TIPO: ácido nucleico
  - 20 (C) TIPO DE CADENA: bicatenaria
  - (D) TOPOLOGÍA: desconocida

25 (ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:22:

	<b>GAATCCGAG GTGGATGTGT TAGATGTGAA TGTCCGTGGC CCAGATGGCT GCACCCCAT</b>	<b>60</b>
	<b>GATGTTGGCT TCTCTCCGAG GAGGCAGCTC AGATTTGAGT GATGAAGATG AAGATGCAGA</b>	<b>120</b>
30	<b>GGACTCTTCT GCTAACATCA TCACAGACTT GGTCTTACCA GGGTGCCAGC CTTCCAGGCC</b>	<b>180</b>
	<b>CAAGAACAGA COGGACTTGG TGAGATGGCC CTGCACCTG CAGCCCGCTA CTACGGGCTG</b>	<b>240</b>
35	<b>ATGCTGCCAA GGTCTGGAT GCAGGTGCAG ATGCCAATGC CCAGGACAAC ATGGGCCGCT</b>	<b>300</b>
	<b>GTCCACTCCA TGCTGCAGTG GCACTGATGC</b>	<b>330</b>

40 (2) INFORMACIÓN DE LA SEQ ID NO:23:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 167 pares de bases
  - 45 (B) TIPO: ácido nucleico
  - (C) TIPO DE CADENA: bicatenaria
  - (D) TOPOLOGÍA: desconocida

50 (ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:23:

	<b>CAGAGGATGG TGAGGGTCCA TGCAGATAGG TTCTCCCAT CCTGTGAATA ATAATGGGT</b>	<b>60</b>
	<b>GCAAGGGCAG AGAGTCACCA TTTAGAATGA TAAATGTTT GCACACTATG AAAGAGGCTG</b>	<b>120</b>
55	<b>ACAGAATGTT GCCCATGGA GAGATAAAGC AGAGAATGAA CAACTT</b>	<b>167</b>

60 (2) INFORMACIÓN DE LA SEQ ID NO:24:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- 65 (A) LONGITUD: 225 pares de bases
  - (B) TIPO: ácido nucleico
  - (C) TIPO DE CADENA: bicatenaria

## ES 2 181 674 T5

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:24:

5  
10  
15  
20  
25  
30  
35

<b>AGGATGAATG ATGGTACTAC ACCCCTGATC CTGGCTGCC</b>	<b>GCCTGGCTGT GGAGGGAATG</b>	<b>60</b>
<b>GTGGCAGAAC TGATCAACTG CCAAGCGGAT GTGAATGCAG</b>	<b>TGGATGACCA TGGAAAATCT</b>	<b>120</b>
<b>GCTCTTCACT GGGCAGCTGC TGTCAATAAT GTGGAGGCAA</b>	<b>CTCTTTTCTT GTTGAAAAT</b>	<b>180</b>
<b>GGGGCCAACC GAGACATGCA GGACAACAAG GAAGAGACAC</b>	<b>CTCTG</b>	<b>225</b>

(2) INFORMACIÓN DE LA SEQ ID NO:25:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 121 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: bicatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:25:

30  
35

<b>AATAATAAT GGGTGCAAGG GCAGAGAGTC ACCATTTAGA</b>	<b>ATGATAAAAT GTTGCACAC</b>	<b>60</b>
<b>TATGAAGAG GCTGACAGAA TGTGCCACA TGGAGAGATA</b>	<b>AAGCAGAGAA TGAACAACT</b>	<b>120</b>
<b>T</b>		<b>121</b>

(2) INFORMACIÓN DE LA SEQ ID NO:26:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: monocatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:26:

50  
20

<b>ACTTCAGCAA CGATCACGGG</b>	<b>20</b>
------------------------------	-----------

(2) INFORMACIÓN DE LA SEQ ID NO:27:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 21 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: monocatenaria

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:27:

65  
21

<b>TTGGGTATGT GACAGTAATC G</b>	<b>21</b>
--------------------------------	-----------

## ES 2 181 674 T5

### (2) INFORMACIÓN DE LA SEQ ID NO:28:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 14 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:28:

TTAAGTTAAC TTAA

14

### (2) INFORMACIÓN DE LA SEQ ID NO:29:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 12 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:29:

GGAAGATCTT CC

12

### (2) INFORMACIÓN DE LA SEQ ID NO:30:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 4 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: péptido

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:30:

Arg Lys Ile Phe  
1

### (2) INFORMACIÓN DE LA SEQ ID NO:31:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 3234 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: bicatenaria
- (D) TOPOLOGÍA: desconocida

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) CARACTERÍSTICAS:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..3234

ES 2 181 674 T5

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:31:

5	TGC CAG GAG GAC GCG GGC AAC AAG GTC TGC AGC CTG CAG TGC AAC AAC	48
	Cys Gln Glu Asp Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn	
	1 5 10 15	
10	CAC GCG TGC GGC TGG GAC GGC GGT GAC TGC TCC CTC AAC TTC AAT GAC	96
	His Ala Cys Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp	
	20 25 30	
15	CCC TGG AAG AAC TGC ACG CAG TCT CTG CAG TGC TGG AAG TAC TTC AGT	144
	Pro Trp Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser	
	35 40 45	
20	GAC GGC CAC TGT GAC AGC CAG TGC AAC TCA GCC GGC TGC CTC TTC GAC	192
	Asp Gly His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp	
	50 55 60	
25	GGC TTT GAC TGC CAG CGT GCG GAA GGC CAG TGC AAC CCC CTC TAC GAC	240
	Gly Phe Asp Cys Gln Arg Ala Glu Gly Gln Cys Asn Pro Leu Tyr Asp	
	65 70 75 80	
30	CAG TAC TGC AAG GAC CAC TTC AGC GAC GGG CAC TGC GAC CAG GGC TGC	288
	Gln Tyr Cys Lys Asp His Phe Ser Asp Gly His Cys Asp Gln Gly Cys	
	85 90 95	
35	AAC AGC GCG GAG TGC GAG TGG GAC GGG CTG GAC TGT GCG GAG CAT GTA	336
	Asn Ser Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala Glu His Val	
	100 105 110	
40	CCC GAG AGG CTG GCG GCC GGC ACG CTG GTG GTG GTG GTG CTG ATG CCG	384
	Pro Glu Arg Leu Ala Ala Gly Thr Leu Val Val Val Val Leu Met Pro	
	115 120 125	
45	CCG GAG CAG CTG CGC AAC AGC TCC TTC CAC TTC CTG CGG GAG CTC AGC	432
	Pro Glu Gln Leu Arg Asn Ser Ser Phe His Phe Leu Arg Glu Leu Ser	
	130 135 140	
50		
55		
60		
65		

ES 2 181 674 T5

	CGC GTG CTG CAC ACC AAC GTG GTC TTC AAG CGT GAC GCA CAC GGC CAG	480
	Arg Val Leu His Thr Asn Val Val Phe Lys Arg Asp Ala His Gly Gln	
	145 150 155 160	
5	CAG ATG ATC TTC CCC TAC TAC GGC CGC GAG GAG GAG CTG CGC AAG CAC	528
	Gln Met Ile Phe Pro Tyr Tyr Gly Arg Glu Glu Glu Leu Arg Lys His	
	165 170 175	
10	CCC ATC AAG CGT GCC GCC GAG GGC TGG GCC GCA CCT GAC GCC CTG CTG	576
	Pro Ile Lys Arg Ala Ala Glu Gly Trp Ala Ala Pro Asp Ala Leu Leu	
	180 185 190	
15	GGC CAG GTG AAG GCC TCG CTG CTC CCT GGT GGC AGC GAG GGT GGG CGG	624
	Gly Gln Val Lys Ala Ser Leu Leu Pro Gly Gly Ser Glu Gly Gly Arg	
	195 200 205	
20	CGG CGG AGG GAG CTG GAC CCC ATG GAC GTC CGC GGC TCC ATC GTC TAC	672
	Arg Arg Arg Glu Leu Asp Pro Met Asp Val Arg Gly Ser Ile Val Tyr	
	210 215 220	
25	CTG GAG ATT GAC AAC CCG CAG TGT GTG CAG GCC TCC TCG CAG TGC TTC	720
	Leu Glu Ile Asp Asn Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe	
	225 230 235 240	
30	CAG AGT GCC ACC GAC GTG GCC GCA TTC CTG GGA GCG CTC GCC TCG CTG	768
	Gln Ser Ala Thr Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu	
	245 250 255	
35	GGC AGC CTC AAC ATC CCC TAC AAG ATC GAG GCC GTG CAG AGT GAG ACC	816
	Gly Ser Leu Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr	
	260 265 270	
40	GTG GAG CCG CCC CCG CCG GCG CAG CTG CAC TTC ATG TAC GTG GCG GCG	864
	Val Glu Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala	
	275 280 285	
45	GCC GCC TTT GTG CTT CTG TTC TTC GTG GGC TGC GGG GTG CTG CTG TCC	912
	Ala Ala Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser	
	290 295 300	
50	CGC AAG CGC CCG CCG CAG CAT GGC CAG CTC TGG TTC CCT GAG GGC TTC	960
	Arg Lys Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu Gly Phe	
	305 310 315 320	
55	AAA GTG TCT GAG GCC AGC AAG AAG AAG CCG CCG GAG CCC CTC GGC GAG	1008
	Lys Val Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Pro Leu Gly Glu	
	325 330 335	
60	GAC TCC GTG GGC CTC AAG CCC CTG AAG AAC GCT TCA GAC GGT GCC CTC	1056
	Asp Ser Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly Ala Leu	
	340 345 350	
65	ATG GAC GAC AAC CAG AAT GAG TGG GGG GAC GAG GAC CTG GAG ACC AAG	1104
	Met Asp Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu Thr Lys	
	355 360 365	
70	AAG TTC CCG TTC GAG GAG CCC GTG GTT CTG CCT GAC CTG GAC GAC CAG	1152
	Lys Phe Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp Asp Gln	
	370 375 380	
75	ACA GAC CAC CCG CAG TGG ACT CAG CAG CAC CTG GAT GCC GCT GAC CTG	1200
	Thr Asp His Arg Gln Trp Thr Gln Gln His Leu Asp Ala Ala Asp Leu	
	385 390 395 400	
80	CGC ATG TCT GCC ATG GCC CCC ACA CCG CCC CAG GGT GAG GTT GAC GCC	1248
	Arg Met Ser Ala Met Ala Pro Thr Pro Pro Gln Gly Glu Val Asp Ala	
	405 410 415	
85	GAC TGC ATG GAC GTC AAT GTC CGC GGG CCT GAT GGC TTC ACC CCG CTC	1296
	Asp Cys Met Asp Val Asn Val Arg Gly Pro Asp Gly Phe Thr Pro Leu	

ES 2 181 674 T5

	420	425	430	
5	ATG ATC GCC TCC TGC AGC GGG GGC GGC CTG GAG ACG GGC AAC AGC GAG Met Ile Ala Ser Cys Ser Gly Gly Gly Leu Glu Thr Gly Asn Ser Glu 435 440 445			1344
	GAA GAG GAG GAC GCG CCG GCC GTC ATC TCC GAC TTC ATC TAC CAG GGC Glu Glu Glu Asp Ala Pro Ala Val Ile Ser Asp Phe Ile Tyr Gln Gly 450 455 460			1392
10	GCC AGC CTG CAC AAC CAG ACA GAC CGC ACG GGC GAG ACC GCC TTG CAC Ala Ser Leu His Asn Gln Thr Asp Arg Thr Gly Glu Thr Ala Leu His 465 470 475 480			1440
15	CTG GCC GCC CGC TAC TCA CGC TCT GAT GCC GCC AAG CGC CTG CTG GAG Leu Ala Ala Arg Tyr Ser Arg Ser Asp Ala Ala Lys Arg Leu Leu Glu 485 490 495			1488
20	GCC AGC GCA GAT GCC AAC ATC CAG GAC AAC ATG GGC CGC ACC CCG CTG Ala Ser Ala Asp Ala Asn Ile Gln Asp Asn Met Gly Arg Thr Pro Leu 500 505 510			1536
25	CAT GCG GCT GTG TCT GCC GAC GCA CAA GGT GTC TTC CAG ATC CTG ATC His Ala Ala Val Ser Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile 515 520 525			1584
30	CGG AAC CGA GCC ACA GAC CTG GAT GCC CGC ATG CAT GAT GGC ACG ACG Arg Asn Arg Ala Thr Asp Leu Asp Ala Arg Met His Asp Gly Thr Thr 530 535 540			1632
35	CCA CTG ATC CTG GCT GCC CGC CTG GCC GTG GAG GGC ATG CTG GAG GAC Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu Gly Met Leu Glu Asp 545 550 555 560			1680
40	CTC ATC AAC TCA CAC GCC GAC GTC AAC GCC GTA GAT GAC CTG GGC AAG Leu Ile Asn Ser His Ala Asp Val Asn Ala Val Asp Asp Leu Gly Lys 565 570 575			1728
45	TCC GCC CTG CAC TGG GCC GCC GCC GTG AAC AAT GTG GAT GCC GCA GTT Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn Val Asp Ala Ala Val 580 585 590			1776
50	GTG CTC CTG AAG AAC GGG GCT AAC AAA GAT ATG CAG AAC AAC AGG GAG Val Leu Leu Lys Asn Gly Ala Asn Lys Asp Met Gln Asn Asn Arg Glu 595 600 605			1824
55	GAG ACA CCC CTG TTT CTG GCC GCC CGG GAG GGC AGC TAC GAG ACC GCC Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly Ser Tyr Glu Thr Ala 610 615 620			1872
60	AAG GTG CTG CTG GAC CAC TTT GCC AAC CGG GAC ATC ACG GAT CAT ATG Lys Val Leu Leu Asp His Phe Ala Asn Arg Asp Ile Thr Asp His Met 625 630 635 640			1920
65	GAC CGC CTG CCG CGC GAC ATC GCA CAG GAG CGC ATG CAT CAC GAC ATC Asp Arg Leu Pro Arg Asp Ile Ala Gln Glu Arg Met His His Asp Ile 645 650 655			1968
	GTG AGG CTG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC Val Arg Leu Leu Asp Glu Tyr Asn Leu Val Arg Ser Pro Gln Leu His 660 665 670			2016
	GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG Gly Ala Pro Leu Gly Gly Thr Pro Thr Leu Ser Pro Pro Leu Cys Ser 675 680 685			2064
	CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG Pro Asn Gly Tyr Leu Gly Ser Leu Lys Pro Gly Val Gln Gly Lys Lys 690 695 700			2112

ES 2 181 674 T5

	GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC	2160
	Val Arg Lys Pro Ser Ser Lys Gly Leu Ala Cys Gly Ser Lys Glu Ala	
	705 710 715 720	
5	AAG GAC CTC AAG GCA CGG AGG AAG AAG TCC CAG GAT GGC AAG GGC TGC	2208
	Lys Asp Leu Lys Ala Arg Arg Lys Lys Ser Gln Asp Gly Lys Gly Cys	
	725 730 735	
10	CTG CTG GAC AGC TCC GGC ATG CTC TCG CCC GTG GAC TCC CTG GAG TCA	2256
	Leu Leu Asp Ser Ser Gly Met Leu Ser Pro Val Asp Ser Leu Glu Ser	
	740 745 750	
15	CCC CAT GGC TAC CTG TCA GAC GTG GCC TCG CCG CCA CTG CTG CCC TCC	2304
	Pro His Gly Tyr Leu Ser Asp Val Ala Ser Pro Pro Leu Leu Pro Ser	
	755 760 765	
20	CCG TTC CAG CAG TCT CCG TCC GTG CCC CTC AAC CAC CTG CCT GGG ATG	2352
	Pro Phe Gln Gln Ser Pro Ser Val Pro Leu Asn His Leu Pro Gly Met	
	770 775 780	
25	CCC GAC ACC CAC CTG GGC ATC GGG CAC CTG AAC GTG GCG GCC AAG CCC	2400
	Pro Asp Thr His Leu Gly Ile Gly His Leu Asn Val Ala Ala Lys Pro	
	785 790 795 800	
30	GAG ATG GCG GCG CTG GGT GGG GGC GGC CGG CTG GCC TTT GAG ACT GGC	2448
	Glu Met Ala Ala Leu Gly Gly Gly Gly Arg Leu Ala Phe Glu Thr Gly	
	805 810 815	
35	CCA CCT CGT CTC TCC CAC CTG CCT GTG GCC TCT GGC ACC AGC ACC GTC	2496
	Pro Pro Arg Leu Ser His Leu Pro Val Ala Ser Gly Thr Ser Thr Val	
	820 825 830	
40	CTG GGC TCC AGC AGC GGA GGG GCC CTG AAT TTC ACT GTG GGC GGG TCC	2544
	Leu Gly Ser Ser Ser Gly Gly Ala Leu Asn Phe Thr Val Gly Gly Ser	
	835 840 845	
45	ACC AGT TTG AAT GGT CAA TGC GAG TGG CTG TCC CCG CTG CAG AGC GGC	2592
	Thr Ser Leu Asn Gly Gln Cys Glu Trp Leu Ser Arg Leu Gln Ser Gly	
	850 855 860	
50	ATG GTG CCG AAC CAA TAC AAC CCT CTG CGG GGG AGT GTG GCA CCA GGC	2640
	Met Val Pro Asn Gln Tyr Asn Pro Leu Arg Gly Ser Val Ala Pro Gly	
	865 870 875 880	
55	CCC CTG AGC ACA CAG GCC CCC TCC CTG CAG CAT GGC ATG GTA GGC CCG	2688
	Pro Leu Ser Thr Gln Ala Pro Ser Leu Gln His Gly Met Val Gly Pro	
	885 890 895	
60	CTG CAC AGT AGC CTT GCT GCC AGC GCC CTG TCC CAG ATG ATG AGC TAC	2736
	Leu His Ser Ser Leu Ala Ala Ser Ala Leu Ser Gln Met Met Ser Tyr	
	900 905 910	
65	CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG	2784
	Gln Gly Leu Pro Ser Thr Arg Leu Ala Thr Gln Pro His Leu Val Gln	
	915 920 925	
70	ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG	2832
	Thr Gln Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln Asn Leu	
	930 935 940	
75	CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA	2880
	Gln Pro Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro Pro Pro	
	945 950 955 960	
80	CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG	2928
	Pro Pro Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu	
	965 970 975	
85	GGC CGG AGC TTC CTG AGT GGA GAG CCG AGC CAG GCA GAC GTG CAG CCA	2976
	Gly Arg Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro	

ES 2 181 674 T5

	980	985	990	
5	CTG GGC CCC AGC AGC CTG GCG GTG CAC ACT ATT CTG CCC CAG GAG AGC Leu Gly Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser 995 1000 1005			3024
	CCC GCC CTG CCC ACG TCG CTG CCA TCC TCG CTG GTC CCA CCC GTG ACC Pro Ala Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr 1010 1015 1020			3072
10	GCA GCC CAG TTC CTG ACG CCC CCC TCG CAG CAC AGC TAC TCC TCG CCT Ala Ala Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro 1025 1030 1035 1040			3120
15	GTG GAC AAC ACC CCC AGC CAC CAG CTA CAG GTG CCT GTT CCT GTA ATG Val Asp Asn Thr Pro Ser His Gln Leu Gln Val Pro Val Pro Val Met 1045 1050 1055			3168
20	GTA ATG ATC CGA TCT TCG GAT CCT TCT AAA GGC TCA TCA ATT TTG ATC Val Met Ile Arg Ser Ser Asp Pro Ser Lys Gly Ser Ser Ile Leu Ile 1060 1065 1070			3216
25	GAA GCT CCC GAC TCA TGG Glu Ala Pro Asp Ser Trp 1075			3234

(2) INFORMACIÓN DE LA SEQ ID NO:32:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1078 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:32:

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

ES 2 181 674 T5

Pro Ile Lys Arg Ala Ala Glu Gly Trp Ala Ala Pro Asp Ala Leu Leu  
 180 185 190  
 5 Gly Gln Val Lys Ala Ser Leu Leu Pro Gly Gly Ser Glu Gly Gly Arg  
 195 200 205  
 Arg Arg Arg Glu Leu Asp Pro Met Asp Val Arg Gly Ser Ile Val Tyr  
 210 215 220  
 10 Leu Glu Ile Asp Asn Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe  
 225 230 235 240  
 Gln Ser Ala Thr Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu  
 245 250 255  
 15 Gly Ser Leu Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr  
 260 265 270  
 Val Glu Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala  
 275 280 285  
 20 Ala Ala Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser  
 290 295 300  
 Arg Lys Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu Gly Phe  
 305 310 315 320  
 Lys Val Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Pro Leu Gly Glu  
 325 330 335  
 30 Asp Ser Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly Ala Leu  
 340 345 350  
 Met Asp Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu Thr Lys  
 355 360 365  
 35 Lys Phe Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp Asp Gln  
 370 375 380  
 Thr Asp His Arg Gln Trp Thr Gln Gln His Leu Asp Ala Ala Asp Leu  
 385 390 395 400  
 Arg Met Ser Ala Met Ala Pro Thr Pro Pro Gln Gly Glu Val Asp Ala  
 405 410 415  
 45 Asp Cys Met Asp Val Asn Val Arg Gly Pro Asp Gly Phe Thr Pro Leu  
 420 425 430  
 Met Ile Ala Ser Cys Ser Gly Gly Gly Leu Glu Thr Gly Asn Ser Glu  
 435 440 445  
 50 Glu Glu Glu Asp Ala Pro Ala Val Ile Ser Asp Phe Ile Tyr Gln Gly  
 450 455 460  
 Ala Ser Leu His Asn Gln Thr Asp Arg Thr Gly Glu Thr Ala Leu His  
 465 470 475 480  
 55 Leu Ala Ala Arg Tyr Ser Arg Ser Asp Ala Ala Lys Arg Leu Leu Glu  
 485 490 495  
 Ala Ser Ala Asp Ala Asn Ile Gln Asp Asn Met Gly Arg Thr Pro Leu  
 500 505 510  
 60 His Ala Ala Val Ser Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile  
 515 520 525  
 Arg Asn Arg Ala Thr Asp Leu Asp Ala Arg Met His Asp Gly Thr Thr  
 530 535 540  
 65 Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu Gly Met Leu Glu Asp



ES 2 181 674 T5

5  
10  
15  
20  
25

Thr Gln Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln Asn Leu  
930 935 940

Gln Pro Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro Pro Pro  
945 950 955 960

Pro Pro Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu  
965 970 975

Gly Arg Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro  
980 985 990

Leu Gly Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser  
995 1000 1005

Pro Ala Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr  
1010 1015 1020

Ala Ala Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro  
1025 1030 1035 1040

Val Asp Asn Thr Pro Ser His Gln Leu Gln Val Pro Val Pro Val Met  
1045 1050 1055

Val Met Ile Arg Ser Ser Asp Pro Ser Lys Gly Ser Ser Ile Leu Ile  
1060 1065 1070

Glu Ala Pro Asp Ser Trp  
1075

(2) INFORMACIÓN DE LA SEQ ID NO:33:

- 30 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 4268 pares de bases
- (B) TIPO: ácido nucleico
- 35 (C) TIPO DE CADENA: bicatenaria
- (D) TOPOLOGÍA: desconocida
- (ii) TIPO DE MOLÉCULA: ADNc
- 40 (ix) CARACTERÍSTICAS:
- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 2..1972
- 45 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:33:

50  
55  
60  
65

G GAG GTG GAT GTG TTA GAT GTG AAT GTC CGT GGC CCA GAT GGC TGC 46  
Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys  
1 5 10 15

ACC CCA TTG ATG TTG GCT TCT CTC CGA GGA GGC AGC TCA GAT TTG AGT 94  
Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser  
20 25 30

GAT GAA GAT GAA GAT GCA GAG GAC TCT TCT GCT AAC ATC ATC ACA GAC 142  
Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp  
35 40 45

TTG GTC TAC CAG GGT GCC AGC CTC CAG GCC CAG ACA GAC CGG ACT GGT 190  
Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly  
50 55 60

GAG ATG GCC CTG CAC CTT GCA GCC CGC TAC TCA CGG GCT GAT GCT GCC 238  
Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala  
65 70 75

AAG CGT CTC CTG GAT GCA GGT GCA GAT GCC AAT GCC CAG GAC AAC ATG 286  
Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met

ES 2 181 674 T5

	80		85		90		95	
5	GGC CGC TGT CCA CTC CAT GCT GCA GTG GCA GCT GAT GCC CAA GGT GTC							334
	Gly Arg Cys Pro		Leu His Ala Ala		Val Ala Ala Asp Ala		Gln Gly Val	
			100		105		110	
10	TTC CAG ATT CTG ATT CGC AAC CGA GTA ACT GAT CTA GAT GCC AGG ATG							382
	Phe Gln Ile Leu Ile Arg Asn Arg		Val Thr Asp Leu Asp Ala Arg		Met			
			115		120		125	
15	AAT GAT GGT ACT ACA CCC CTG ATC CTG GCT GCC CGC CTG GCT GTG GAG							430
	Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu							
			130		135		140	
20	GGA ATG GTG GCA GAA CTG ATC AAC TGC CAA GCG GAT GTG AAT GCA GTG							478
	Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val							
			145		150		155	
25	GAT GAC CAT GGA AAA TCT GCT CTT CAC TGG GCA GCT GCT GTC AAT AAT							526
	Asp Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn							
			160		165		170	175
30	GTG GAG GCA ACT CTT TTG TTG TTG AAA AAT GGG GCC AAC CGA GAC ATG							574
	Val Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met							
			180		185		190	
35	CAG GAC AAC AAG GAA GAG ACA CCT CTG TTT CTT GCT GCC CGG GAG GGG							622
	Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly							
			195		200		205	
40	AGC TAT GAA GCA GCC AAG ATC CTG TTA GAC CAT TTT GCC AAT CGA GAC							670
	Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp							
			210		215		220	
45	ATC ACA GAC CAT ATG GAT CGT CTT CCC CGG GAT GTG GCT CGG GAT CGC							718
	Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg							
			225		230		235	
50	ATG CAC CAT GAC ATT GTG CGC CTT CTG GAT GAA TAC AAT GTG ACC CCA							766
	Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro							
			240		245		250	255
55	AGC CCT CCA GGC ACC GTG TTG ACT TCT GCT CTC TCA CCT GTC ATC TGT							814
	Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys							
			260		265		270	
60	GGG CCC AAC AGA TCT TTC CTC AGC CTG AAG CAC ACC CCA ATG GGC AAG							862
	Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys							
			275		280		285	
65	AAG TCT AGA CGG CCC AGT GCC AAG AGT ACC ATG CCT ACT AGC CTC CCT							910
	Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro							
			290		295		300	
70	AAC CTT GCC AAG GAG GCA AAG GAT GCC AAG GGT AGT AGG AGG AAG AAG							958
	Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys							
			305		310		315	
75	TCT CTG AGT GAG AAG GTC CAA CTG TCT GAG AGT TCA GTA ACT TTA TCC							1006
	Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser							
			320		325		330	335
80	CCT GTT GAT TCC CTA GAA TCT CCT CAC ACG TAT GTT TCC GAC ACC ACA							1054
	Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr							
			340		345		350	
85	TCC TCT CCA ATG ATT ACA TCC CCT GGG ATC TTA CAG GCC TCA CCC AAC							1102
	Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn							
			355		360		365	

ES 2 181 674 T5

	CCT	ATG	TTG	GCC	ACT	GCC	GCC	CCT	CCT	GCC	CCA	GTC	CAT	GCC	CAG	CAT	1150
	Pro	Met	Leu	Ala	Thr	Ala	Ala	Pro	Pro	Ala	Pro	Val	His	Ala	Gln	His	
			370					375					380				
5	GCA	CTA	TCT	TTT	TCT	AAC	CTT	CAT	GAA	ATG	CAG	CCT	TTG	GCA	CAT	GGG	1198
	Ala	Leu	Ser	Phe	Ser	Asn	Leu	His	Glu	Met	Gln	Pro	Leu	Ala	His	Gly	
			385				390					395					
10	GCC	AGC	ACT	GTG	CTT	CCC	TCA	GTG	AGC	CAG	TTG	CTA	TCC	CAC	CAC	CAC	1246
	Ala	Ser	Thr	Val	Leu	Pro	Ser	Val	Ser	Gln	Leu	Leu	Ser	His	His	His	
						405					410					415	
15	ATT	GTG	TCT	CCA	GGC	AGT	GGC	AGT	GCT	GGA	AGC	TTG	AGT	AGG	CTC	CAT	1294
	Ile	Val	Ser	Pro	Gly	Ser	Gly	Ser	Ala	Gly	Ser	Leu	Ser	Arg	Leu	His	
					420					425					430		
20	CCA	GTC	CCA	GTC	CCA	GCA	GAT	TGG	ATG	AAC	CGC	ATG	GAG	GTG	AAT	GAG	1342
	Pro	Val	Pro	Val	Pro	Ala	Asp	Trp	Met	Asn	Arg	Met	Glu	Val	Asn	Glu	
						435			440					445			
25	ACC	CAG	TAC	AAT	GAG	ATG	TTT	GGT	ATG	GTC	CTG	GCT	CCA	GCT	GAG	GGC	1390
	Thr	Gln	Tyr	Asn	Glu	Met	Phe	Gly	Met	Val	Leu	Ala	Pro	Ala	Glu	Gly	
			450					455					460				
30	ACC	CAT	CCT	GGC	ATA	GCT	CCC	CAG	AGC	AGG	CCA	CCT	GAA	GGG	AAG	CAC	1438
	Thr	His	Pro	Gly	Ile	Ala	Pro	Gln	Ser	Arg	Pro	Pro	Glu	Gly	Lys	His	
			465				470					475					
35	ATA	ACC	ACC	CCT	CGG	GAG	CCC	TTG	CCC	CCC	ATT	GTG	ACT	TTC	CAG	CTC	1486
	Ile	Thr	Thr	Pro	Arg	Glu	Pro	Leu	Pro	Pro	Ile	Val	Thr	Phe	Gln	Leu	
						485					490					495	
40	ATC	CCT	AAA	GGC	AGT	ATT	GCC	CAA	CCA	GCG	GGG	GCT	CCC	CAG	CCT	CAG	1534
	Ile	Pro	Lys	Gly	Ser	Ile	Ala	Gln	Pro	Ala	Gly	Ala	Pro	Gln	Pro	Gln	
					500					505					510		
45	TCC	ACC	TGC	CCT	CCA	GCT	GTT	GCG	GGC	CCC	CTG	CCC	ACC	ATG	TAC	CAG	1582
	Ser	Thr	Cys	Pro	Pro	Ala	Val	Ala	Gly	Pro	Leu	Pro	Thr	Met	Tyr	Gln	
				515					520					525			
50	ATT	CCA	GAA	ATG	GCC	CGT	TTG	CCC	AGT	GTG	GCT	TTC	CCC	ACT	GCC	ATG	1630
	Ile	Pro	Glu	Met	Ala	Arg	Leu	Pro	Ser	Val	Ala	Phe	Pro	Thr	Ala	Met	
				530				535					540				
55	ATG	CCC	CAG	CAG	GAC	GGG	CAG	GTA	GCT	CAG	ACC	ATT	CTC	CCA	GCC	TAT	1678
	Met	Pro	Gln	Gln	Asp	Gly	Gln	Val	Ala	Gln	Thr	Ile	Leu	Pro	Ala	Tyr	
						545							555				
60	CAT	CCT	TTC	CCA	GCC	TCT	GTG	GGC	AAG	TAC	CCC	ACA	CCC	CCT	TCA	CAG	1726
	His	Pro	Phe	Pro	Ala	Ser	Val	Gly	Lys	Tyr	Pro	Thr	Pro	Pro	Ser	Gln	
						565					570					575	
65	CAC	AGT	TAT	GCT	TCC	TCA	AAT	GCT	GCT	GAG	CGA	ACA	CCC	AGT	CAC	AGT	1774
	His	Ser	Tyr	Ala	Ser	Ser	Asn	Ala	Ala	Glu	Arg	Thr	Pro	Ser	His	Ser	
						580				585					590		
70	GGT	CAC	CTC	CAG	GGT	GAG	CAT	CCC	TAC	CTG	ACA	CCA	TCC	CCA	GAG	TCT	1822
	Gly	His	Leu	Gln	Gly	Glu	His	Pro	Tyr	Leu	Thr	Pro	Ser	Pro	Glu	Ser	
					595				600					605			
75	CCT	GAC	CAG	TGG	TCA	AGT	TCA	TCA	CCC	CAC	TCT	GCT	TCT	GAC	TGG	TCA	1870
	Pro	Asp	Gln	Trp	Ser	Ser	Ser	Ser	Pro	His	Ser	Ala	Ser	Asp	Trp	Ser	
				610				615					620				
80	GAT	GTG	ACC	ACC	AGC	CCT	ACC	CCT	GGG	GGT	GCT	GGA	GGA	GGT	CAG	CGG	1918
	Asp	Val	Thr	Thr	Ser	Pro	Thr	Pro	Gly	Gly	Ala	Gly	Gly	Gly	Gln	Arg	
						625		630				635					
85	GGA	CCT	GGG	ACA	CAC	ATG	TCT	GAG	CCA	CCA	CAC	AAC	AAC	ATG	CAG	GTT	1966
	Gly	Pro	Gly	Thr	His	Met	Ser	Glu	Pro	Pro	His	Asn	Asn	Met	Gln	Val	

ES 2 181 674 T5

	640	645	650	655	
	TAT GCG TGAGAGAGTC CACCTCCAGT GTAGAGACAT AACTGACTTT TGTAATGCT				2022
5	Tyr Ala				
	GCTGAGGAAC AAATGAAGGT CATCCGGGAG AGAAATGAAG AAATCTCTGG AGCCAGCTTC				2082
	TAGAGGTAGG AAAGAGAAGA TGTTCCTTATT CAGATAATGC AAGAGAAGCA ATTCCGTCAGT				2142
10	TTCACTGGGT ATCTGCAAGG CTTATTGATT ATTCTAATCT AATAAGACAA GTTTGTGGAA				2202
	ATGCAAGATG AATACAAGCC TTGGGTCCAT GTTACTCTC TTCTATTTGG AGAATAAGAT				2262
15	GGATGCTTAT TGAAGCCCAG ACATTCTTGC AGCTTGGACT GCATTTTAAG CCCTGCAGGC				2322
	TTCTGCCATA TCCATGAGAA GATTCTACAC TAGCGTCCCTG TTGGGAATTA TGCCCTGGAA				2382
	TTCTGCCTGA ATTGACCTAC GCATCTCCTC CTCCTTGGAC ATTCTTTTGT CTTCAATTTGG				2442
20	TGCTTTTGGT TTTGCACCTC TCCGTGATTG TAGCCCTACC AGCATGTTAT AGGGCAAGAC				2502
	CTTGTGCTT TTGATCATTG TGGCCCATGA AAGCAACTTT GGTCTCCTTT CCCCTCCTGT				2562
	CTTCCCGGTA TCCCTTGGAG TCTCACAAGG TTTACTTTGG TATGGTTCTC AGCACAAACC				2622
25	TTTCAAGTAT GTTGTTCCTT TGGAAAATGG ACATACTGTA TTGTGTTCTC CTGCATATAT				2682
	CATTCCTGGA GAGAGAAGGG GAGAAGAATA CTTTTCTTCA ACAAATTTTG GGGGCAGGAG				2742
	ATCCCTTCAA GAGGCTGCAC CTTAATTTTT CTGTCTGTG TGCAGGTCTT CATATAAACT				2802
30	TTACCAGGAA GAAGGGTGTG AGTTTGTGT TTTTCTGTGT ATGGGCCTGG TCAGTGTAAA				2862
	GTTTTATCCT TGATAGTCTA GTTACTATGA CCCTCCCCAC TTTTTTAAA CCAGAAAAG				2922
35	GTTTGGAAATG TTGGAATGAC CAAGAGACAA GTTAACCTGT GCAAGAGCCA GTTACCACC				2982
	CACAGGTCCC OCTACTTCCT GCCAAGCATT CCATTGACTG CCTGTATGGA ACACATTTGT				3042
	CCCAGATCTG AGCATTCTAG GCCTGTTTCA CTCACTCACC CAGCATATGA AACTAGTCTT				3102
40	AACTGTTGAG CCTTTCCTTT CATACCACA GAAGACTG TCTCAAATGT TGTACCCTTG				3162
	CCATTTAGGA CTGAACTTTC CTAGCCCAA GGGACCCAGT GACAGTTGTC TTCCGTTTGT				3222
	CAGATGATCA GTCTCTACTG ATTATCTTGC TGCTTAAAGG CCTGCTCACC AATCTTCTT				3282
45	TCACACCGTG TGGTCCGTGT TACTGGTATA CCCAGTATGT TCTCACTGAA GACATGGACT				3342
	TTATATGTTT AAGTGCAGGA ATTGGAAAGT TGGACTTGT TTCTATGATC CAAAACAGCC				3402
50	CTATAAGAAG GTTGGAAAAG GAGGAACTAT ATAGCAGCCT TTGCTATTTT CTGCTACCAT				3462
	TTCTTTTCTT CTGAAGCGGC CATGACATTC CCTTTGGCAA CTAACGTAGA AACTCAACAG				3522
	AACATTTTCC TTTCCTAGAG TCACCTTTTA GATGATAATG GACAACATA GACTTGCTCA				3582
55	TTGTTGAGAC TGATTGCCCC TCACCTGAAT CCACTCTCTG TATTGATGCT CTGGCAATT				3642
	TCTTTGACTT TCTTTTAAAG GCAGAAGCAT TTTAGTTAAT TGTAGATAAA GAATAGTTTT				3702
	CTTCTCTTTC TCCTTGGGCC AGTTAATAAT TGGTCCATGG CTACACTGCA ACTTCCGTCC				3762
60	AGTGCTGTGA TGCCCATGAC ACCTGCAAAA TAAGTTCTGC CTGGGCATTT TGTAGATATT				3822
	AACAGGTGAA TTCCCGACTC TTTTGGTTTG AATGACAGTT CTCATTCCTT CTATGGCTGC				3882
65	AAGTATGCAT CAGTGCCTCC CACTTACCTG ATTTGTCTGT CGGTGGCCCC ATATGGAAAC				3942





ES 2 181 674 T5

Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gly Gln Arg Gly  
625 630 635 640

Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val Tyr  
645 650 655

Ala

(2) INFORMACIÓN DE LA SEQ ID NO:35:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 654 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADENA: monocatenaria
- (D) TOPOLOGÍA: desconocida

(ii) TIPO DE MOLÉCULA: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:35:

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

ES 2 181 674 T5

Ile Ala Gln Glu Arg Met His His Asp Ile Val His Leu Leu Asp Glu  
 245 250 255  
 Tyr Asn Leu Val Lys Ser Pro Thr Leu His Asn Gly Pro Leu Gly Ala  
 5 260 265 270  
 Thr Thr Leu Ser Pro Pro Ile Cys Ser Pro Asn Gly Tyr Met Gly Asn  
 275 280 285  
 Met Lys Pro Ser Val Gln Ser Lys Lys Ala Arg Lys Pro Ser Ile Lys  
 10 290 295 300  
 Gly Asn Gly Cys Lys Glu Ala Lys Glu Leu Lys Ala Arg Arg Lys Lys  
 15 305 310 315 320  
 Ser Gln Asp Gly Lys Thr Thr Leu Leu Asp Ser Gly Ser Ser Gly Val  
 325 330 335  
 Leu Ser Pro Val Asp Ser Leu Glu Ser Thr His Gly Tyr Leu Ser Asp  
 20 340 345 350  
 Val Ser Ser Pro Pro Leu Met Thr Ser Pro Phe Gln Gln Ser Pro Ser  
 355 360 365  
 Met Pro Leu Asn His Leu Thr Ser Met Pro Glu Ser Gln Leu Gly Met  
 25 370 375 380  
 Asn His Ile Asn Met Ala Thr Lys Gln Glu Met Ala Ala Gly Ser Asn  
 385 390 395 400  
 Arg Met Ala Phe Asp Ala Met Val Pro Arg Leu Thr His Leu Asn Ala  
 30 405 410 415  
 Ser Ser Pro Asn Thr Ile Met Ser Asn Gly Ser Met His Phe Thr Val  
 420 425 430  
 Gly Gly Ala Pro Thr Met Asn Ser Gln Cys Asp Trp Leu Ala Arg Leu  
 35 435 440 445  
 Gln Asn Gly Met Val Gln Asn Gln Tyr Asp Pro Ile Arg Asn Gly Ile  
 40 450 455 460  
 Gln Gln Gly Asn Ala Gln Gln Ala Gln Ala Leu Gln His Gly Leu Met  
 465 470 475 480  
 Thr Ser Leu His Asn Gly Leu Pro Ala Thr Thr Leu Ser Gln Met Met  
 45 485 490 495  
 Thr Tyr Gln Ala Met Pro Asn Thr Arg Leu Ala Asn Gln Pro His Leu  
 500 505 510  
 Met Gln Ala Gln Gln Met Gln Gln Gln Gln Asn Leu Gln Leu His Gln  
 50 515 520 525  
 Ser Met Gln Gln Gln His His Asn Ser Ser Thr Thr Ser Thr His Ile  
 530 535 540  
 Asn Ser Pro Phe Cys Ser Ser Asp Ile Ser Gln Thr Asp Leu Gln Gln  
 55 545 550 555 560  
 Met Ser Ser Asn Asn Ile His Ser Val Met Pro Gln Asp Thr Gln Ile  
 565 570 575  
 Phe Ala Ala Ser Leu Pro Ser Asn Leu Thr Gln Ser Met Thr Thr Ala  
 580 585 590  
 Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro Met Asp  
 595 600 605  
 Asn Thr Pro Ser His Gln Leu Gln Val Pro Asp His Pro Phe Leu Thr



ES 2 181 674 T5

Asn Leu Val Arg Ser Pro Gln Leu His Gly Thr Ala Leu Gly Gly Thr  
 260 265 270  
 5 Pro Thr Leu Ser Pro Thr Leu Cys Ser Pro Asn Gly Tyr Leu Gly Asn  
 275 280 285  
 Leu Lys Ser Ala Thr Gln Gly Lys Lys Ala Arg Lys Pro Ser Thr Lys  
 290 295 300  
 10 Gly Leu Ala Cys Ser Ser Lys Glu Ala Lys Asp Leu Lys Ala Arg Arg  
 305 310 315 320  
 Lys Lys Ser Gln Asp Gly Lys Gly Cys Leu Leu Asp Ser Ser Ser Met  
 325 330 335  
 15 Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His Gly Tyr Leu Ser Asp  
 340 345 350  
 Val Ala Ser Pro Pro Leu Pro Ser Pro Phe Gln Gln Ser Pro Ser Met  
 355 360 365  
 20 Pro Leu Ser His Leu Pro Gly Met Pro Asp Thr His Leu Gly Ile Ser  
 370 375 380  
 His Leu Asn Val Ala Ala Lys Pro Glu Met Ala Ala Leu Ala Gly Gly  
 385 390 395 400  
 25 Ser Arg Leu Ala Phe Glu Pro Pro Pro Arg Leu Ser His Leu Pro  
 405 410 415  
 Val Ala Ser Ser Ala Ser Thr Val Leu Ser Thr Asn Gly Thr Gly Ala  
 420 425 430  
 30 Met Asn Phe Thr Val Gly Ala Pro Ala Ser Leu Asn Gly Gln Cys Glu  
 435 440 445  
 Trp Leu Pro Arg Leu Gln Asn Gly Met Val Pro Ser Gln Tyr Asn Pro  
 450 455 460  
 35 Leu Arg Pro Gly Val Thr Pro Gly Thr Leu Ser Thr Gln Ala Ala Gly  
 465 470 475 480  
 40 Leu Gln His Gly Met Met Ser Pro Ile His Ser Ser Leu Ser Thr Asn  
 485 490 495  
 Thr Leu Ser Pro Ile Ile Tyr Gln Gly Leu Pro Asn Thr Arg Leu Ala  
 500 505 510  
 45 Thr Gln Pro His Leu Val Gln Thr Gln Gln Val Gln Pro Gln Asn Leu  
 515 520 525  
 Gln Ile Gln Pro Gln Asn Leu Gln Pro Pro Ser Gln Pro His Leu Ser  
 530 535 540  
 50 Val Ser Ser Ala Ala Asn Gly His Leu Gly Arg Ser Phe Leu Ser Gly  
 545 550 555 560  
 55 Glu Pro Ser Gln Ala Asp Val Gln Pro Leu Gly Pro Ser Ser Leu Pro  
 565 570 575  
 Val His Thr Ile Leu Pro Gln Glu Ser Gln Ala Leu Pro Thr Ser Leu  
 580 585 590  
 60 Pro Ser Ser Met Val Pro Pro Met Thr Thr Thr Gln Phe Leu Thr Pro  
 595 600 605  
 Pro Ser Gln His Ser Tyr Ser Ser Ser Pro Val Asp Asn Thr Pro Ser  
 610 615 620  
 65 His Gln Leu Gln Val Pro Glu His Pro Phe Leu Thr Pro Ser Pro Glu





ES 2 181 674 T5

	<b>625</b>					<b>630</b>					<b>635</b>				<b>640</b>	
	Gln	Leu	Gln	Val	Pro	Glu	His	Pro	Phe	Leu	Thr	Pro	Ser	Pro	Glu	Ser
5					645					650					655	
	Pro	Asp	Gln	Trp	Ser	Ser	Ser	Ser	Pro	His	Ser	Asn	Val	Ser	Asp	Trp
				660					665					670		
10	Ser	Glu	Gly	Val	Ser	Ser	Pro	Pro	Thr							
			675					680								
15																
20																
25																
30																
35																
40																
45																
50																
55																
60																
65																